

PROCEEDINGS

NCPS VOLUME 4 (2024)

TABLE OF CONTENT

1. Development of Information Technology (IT) Security System for Telemarketing: I.E. Eteng, Abuokwen, V. A.	1-16
2. Internet of Things (IoT)-Based Food Crop Security Model for Grazing Monitoring and Reporting using Convolutional Neural Network: M. A. Agana and O. A. Akima	17-30
3. A Stack Heterogenous Ensemble (SHE) model for predicting road traffic violations in South-Eastern Nigeria: E. A. Emmanuel, Ibobo N. S., Odo K.	31-44
4. Development of a remote emergency call mobile tracking system: Uke, Ernest M. and F. U. Ogban	45-53
5. A survey of Markov Chain on Metric Space: U. E. Edeke and G. I. Attah	54-57
6. Asymmetric Power Autoregressive Conditional Heteroscedasticity Modelling of Interest Rate Return in Nigeria: Andrew Ekpung Ntu and Eric Monfung Egomu	58-66
7. Operators defined by Convolution on the Euclidean Motion Group: U. E. Edeke, U.N. Bassey	67-71
8. Partial survey of Fourier transform on the Euclidean Motion Group: U.E.Edeke¹, U.N.Bassey	72-76
9. A review of Sobolev space on Locally Compact group: U. E. Edeke and M. Joy	77-80
10. Improving the performance of hybridized random forest and long short term memory algorithms using smote, for intrusion detection: M. A. Agana, E. Edim and A. I. Bassey	81-93
11. Virtual Communication using Centralized Geolocation Network Data: N. N. Eku, O. A. Akima and O.A. Ofem	94-106
12. A Binary Integer Programming Model for Detecting Arbitrage Opportunities in Currency Arbitrage Problems: E. O. Effanga, P. A. Ayuk	107-116
13. Advanced Techniques for Solving Specific Classes of Third-Order Nonlinear Partial Differential Equations Using the Balance Method: Theoretical Developments and Practical Applications: Chikwe, C.F. Ezeorah, J.N and Essang, S.O	117-129
14. Radiation Shielding Properties of Clam Shell Mortar for use in Diagnostic Radiology Facility: Abeh J. U, and S. O. Inyang.	130-138
15. Production and characterization of microalgae bio-oil: an energy transition approach and alternative for fossil fuel: Ebri, Robert Utum	139-146
16. An intelligent network intrusion detection system using multilayered technique: B. I. Ele, O. E. Ofem & D. O. Egete	147-158
17. The effects of plant extracts on the proximate composition of Bonga Shad (<i>Ethmalosa fimbriata</i>) in Marian Market, Calabar, Nigeria: Inyang- Etoh, A.P. and Asinde, J.	159-164
18. An Improved Estimator for Estimating the Current Population Mean in a Two Occasion Successive Sampling Using One Auxiliary Variable: Agbebia, C. S., Eric M. E, Etorti, I. J.	165-175
19. Theoretical investigation of metals (Ni, Zn, Fe) doped zeolites for adsorption of volatile gases (Toluene, Xylene and Formaldehyde): B. E. Inah, and S. J. Oduma	176-185
20. Physicochemical properties of Nigerian-grown <i>Pachira aquatic</i> seed oil and its potentials for national socio-economic recovery: B. Sam-Uket Okori and M. A. Akpe	186-192
21. Hydrogeochemistry and Water Assessment of Shallow Groundwater in Parts of Benue State, North Central Nigeria: C.I. Adamu, I. Akadi, E. Kudamnya, M. Asen, F. Adamu	193-199
22. Geospatial assessment of landslide vulnerability mapping using remotely sensed data and fuzzy-AHP in urban and rural regions of Orumba North, southeastern Nigeria: G. I. Joshua, E. A. Kudamnya, A. S. Ekwere, B. O. Onwanyi, A. E. Edet	200-204
23. Measurement of seasonal variation of radon concentration levels associate with building types in some parts of Cross River State, Nigeria: W.E. Azogor, Inyang, I. O. Akpan	205-210
24. Conglomerate Facies of the Awi Formation, Calabar Flank, Southeast Nigeria: Implications for Paleoenvironments and Economic Geology: Okon E. E., Kudamnya, E. A., Essien, N. U., Inyang, D. O, Nton, M. E. and Ugbaja, A. N.	211-217
25. Yttrium-encapsulated Tin-decorated carbon-based fullerene (Sn-Y@C ₈₀) nanostructured as a sensor for AsH ₃ , NH ₃ , PH ₃ and SbH ₃ : Insight from DFT and QTAIM: C. G. Apebende, I. O. Amodu, Ubua U. P., S. James, M. A. Igbilishi, M. A. Sunday and S. A. Adalikwu	218-224

26. Geochemistry of the Awi sandstone, Calabar Flank, Southeastern (SE) Nigeria: Constraints on the metal enrichment, provenance, and tectonic setting: **B.O. Omang, A.T. Okah, L.T. Ojikutu, G.E. Abang and G.T. Kave** 225-231
27. Compositional Signatures of Gold from two Pan-African settings: Igarra Schist Belt, Southwest Nigeria and Lom Series Schist Belt, East Cameroon: **G.T. Kave and B.O. Omang** 232-235
28. Assessment of Groundwater Suitability for Irrigation in Uburu and its Environs, SE Nigeria: **Nnodim C.J., Kudamnya E.A., Ochelebe I., Eyong G.A., Akambende E.A.** 236-240
29. Assessment of Groundwater Potentials using Dar-Zarrouk Parameters in parts of Akamkpa, Southeast Nigeria: **Akambende E. A., Asinya E. A., Kudamnya E. A., Okon, E. E., Ochelebe I. Nnodim C. J.** 241-244
30. An integrated study of biogeochemical controls on heavy metals transfer from soil to food chain using cassava tubers around major rock quarrying sites in western Oban Massif, south-eastern Nigeria: **A. S. Ekwere, B.B Edet and A.N. Ugbaja** 245-251
31. Conglomerate facies of the Awi formation, Calabar flank, southeast Nigeria: implications for paleoenvironments and economic geology: **Okon E. E., Kudamnya, E. A., Essien, N. U., Inyang, D. O, Nton, M. E. and Ugbaja, A. N.** 252-257
32. Suitability assessment of water resources around Akamkpa and environs for irrigation purposes: an integrated approach: **B.B. Edet and A. S. Ekwere** 258-262
33. Statistical and Decadal Prediction of Temperature Trend for 121 Years (1901-2022) Using ARIMA (2,2,1) Model in Calabar Nigeria: **M.A. Okono, I.O. Ewona, E.A. Awak, J.E. Osang** 263-277
34. Climatic variability over the vegetation zones in Nigeria for 31 years (1991 – 2022): **M.A. Okono, I.O. Ewona, E.A. Awak, J.E. Osang** 278-297
35. Major and trace element geochemistry of pegmatites in Umai, Oban Massif, SE-Nigeria; constraints on its characterization and mineralization potential: **B.O. Omang, L.T. Ojikutu, A.T. Okah, G.E. Abang and G.T. Kave** 298-303
36. Trend and latitudinal variation of wind speed distribution in three selected locations in Nigeria for 41 years (1981-2022). An insight from atmospheric circulation: **M.A. Okono, I.O. Ewona, E.A. Awak, J.E. Osang** 304-318
37. Assessing the Correlation between Seasonal Atmospheric Heating and Ionospheric TEC Dynamics: Insights from Spring Equinox and Summer Solstice and Fall: **Awak, E. A., George, A. M, Okono, M. A., and Odong, P. O.** 319-330
38. Deep Learning Framework for Predicting Alternative Routes for Information Management in Intra-city Road Traffic Network: **Ifreke J. Udoeka, Moses A. Agana, Ofem A. Ofem and Ifiok J. Udo** 331-339
39. AI-Powered Geophysical Assessment of Dumpsites for Economic Recovery: A Case Study of Umusedeli and Azunze, Kwale, South-South Nigeria: **Okorafor, C. N., Salami, A. S., Ugbe, C. F., Okorafor G. N.** 340-350
40. Mathematical Analysis of Poor Gas Mileage in Automobiles Using Cause-effect Diagram: **Bassey Echeng Bassey** 351-361
41. Exploring patterns of breast cancer tumor characteristics using factor analysis: **E. I. Enang and E. A. Essien** 362-372

Development of Information Technology (IT) Security System for Telemarketing

Idongesit E. Eteng. Abuokwen, Victoria Awhokiema
Department of Computer Science, University of Calabar, Calabar, Nigeria
abuokwenv31@gmail.co

ABSTRACT

The gradual loss of customer confidence in telemarketing strategies necessitated the development of this IT security system for telemarketing. This research employed the Dynamic System Development Methodology to create a system designed to control telemarketing activities and document incidents. It has been identified that telemarketing issues can only be effectively addressed by implementing a security system at the service provider's end, rather than relying solely on mobile phone applications. A comprehensive data text was compiled for both back-end and front-end users, including a database of phone numbers representing telemarketers and "Do Not Call" (DNC) numbers. Users were categorized into telemarketers, normal callers, and DNC users. A security application system was developed to act as an interceptor, capable of terminating calls initiated to DNC numbers and identifying telemarketers from a customer's phone with the interceptor app installed. Kotlin and Java were used as the primary programming languages, with Android Studio as the Integrated Development Environment (IDE). The user interface was designed using XML for layout files, with Android Studio's visual layout editor facilitating drag-and-drop UI design and easy previewing. The results demonstrated that telemarketing calls could be detected, curbing fraud and ensuring that only those who consent to telemarketing calls would be contacted. While telemarketing remains a functional marketing tool, data integrity and security measures are crucial for secure business transactions and effective information dissemination across various domains and business platforms.

I Introduction

Organizations and enterprises produce goods and services geared towards consumer satisfaction. Satisfying consumer needs in this regard must be considered as it helps organizations proffer better solutions to teeming consumer needs. To do this, there ought to be appropriate communication between the producers and the consumers, who are the final recipients. Therefore, appropriate marketing strategies must be implemented to enhance consumer-organization communication.

Marketing remains an approach followed by firms to link the consumers and inform them about several types of products they produce and their work. These innovative techniques include Personalized Marketing, Direct Mail Marketing, Partner Marketing, Public Relations (PR) Marketing, Word of Mouth Marketing, Stealth Marketing, Cross Media Marketing, Database Marketing, Digital Marketing, Evangelism Marketing, Free Sample Marketing, Freebie Marketing, Guerrilla Marketing, Telemarketing or Direct Marketing. (Laoyan, 2024).

Telemarketing is the marketing technique used by companies to contact potential customers and talk to them about their product and services over the telephone or internet (Nerea, 2023). It is a vital marketing tool for selling goods and services to customers of various business persuasions. The widespread adoption of telemarketing and its acceptance of business transactions came with the emergence of the Internet as the de facto standard for converged business transactions. Most importantly, telemarketing is the means customers can be reached at great distances without recourse to rigorous transportation procedures.

All types of advertising activities are currently used by marketers for their brands to be differentiated in the market. At the same time, competitiveness among marketers is booming. As a result of population and economic growth, marketers have expanded their marketing activities to reach consumers (Kenneth, 2017). Peattie and Peattie (2019) stated that "Promotional activities usually specific to a period, place or customer group, which encourage a direct response from consumers or marketing intermediaries, through the offer of additional benefits" with this urge to reach prospective consumers on time gave rise to Telemarketing.

III Methodology

A survey methodology, which uses Dynamic System Development Technique (DSDT), was used for this research, because the method is capable of providing information on the attitudes, feelings, beliefs, behaviour, personal characteristics and other descriptive terms of respondents (Chisnall, 2018). The population comprised of the major distributors of consumer products. (DSDT) has five phases which are feasibility, business study, functional model iteration, design and build iteration and implementation phase. It has the following life cycle as shown in figure 1.

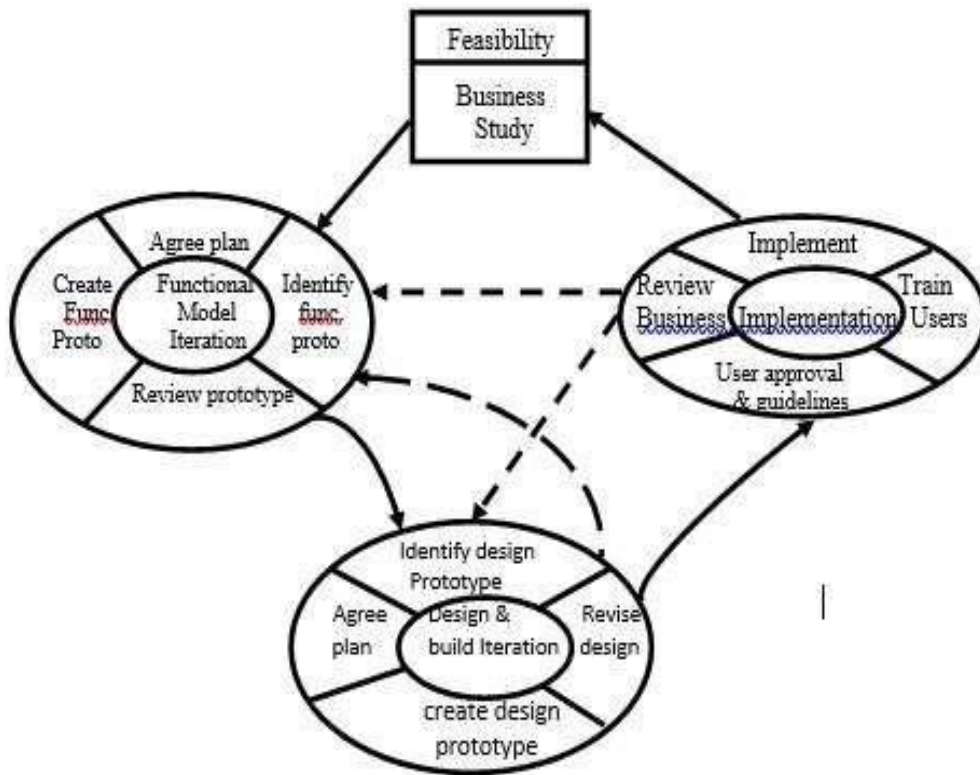


FIG. 1: Five-phase Life Cycle

Proposed High Level Model for Solution to Problems Identified

The security concerns of the existing system have necessitated the development of a new system that incorporates a security module for logging and reporting incidents that may arise during the cause of transacting telemarketing businesses. The model for the proposed system is depicted in figure 4.

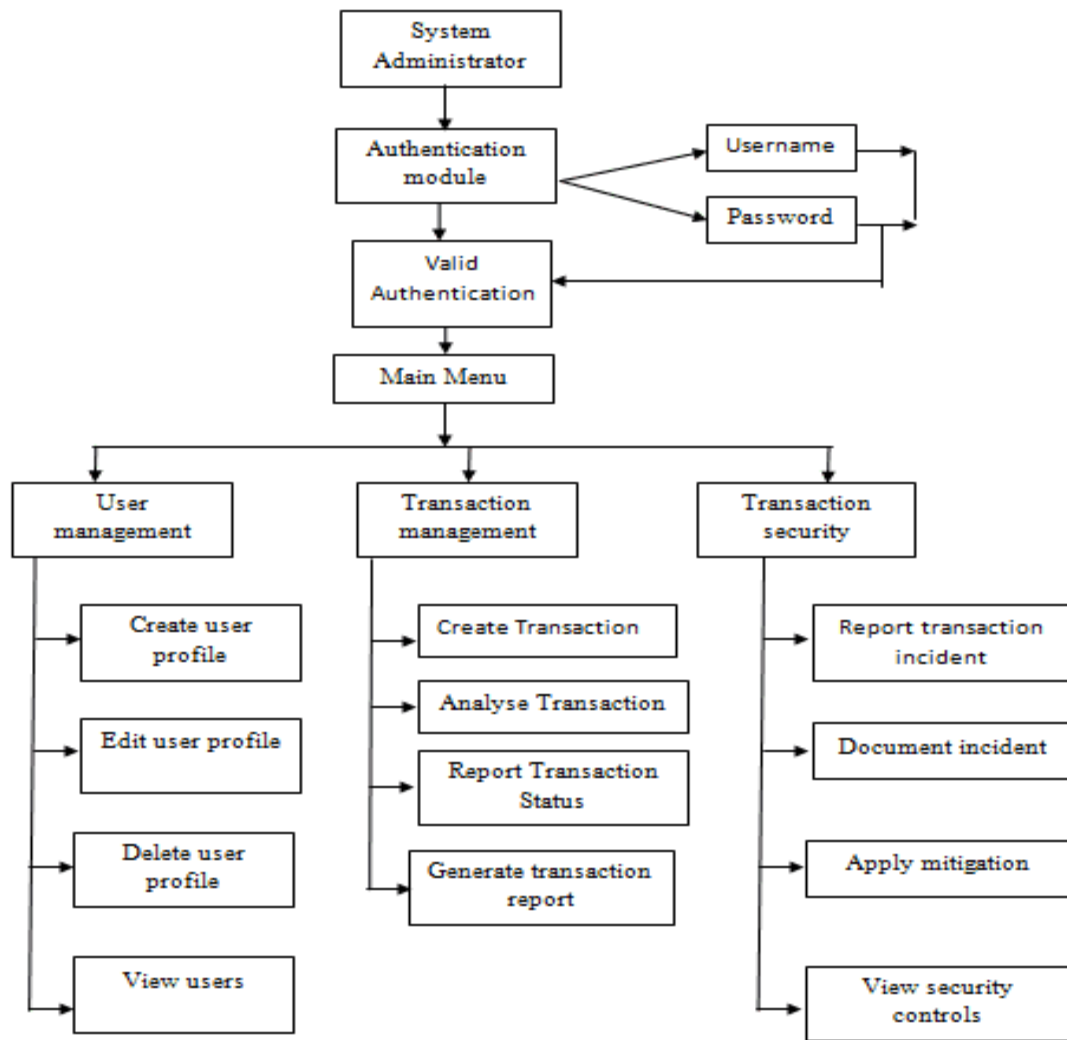


FIG. 3: High level model of the proposed solution

Figure 3 shows the high-level model of the proposed system. The system is designed to have an administrator who initiates the process of creating profiles for telemarketers and tracking the contents they push out to consumers in order to control fraudulent messages.

System Design

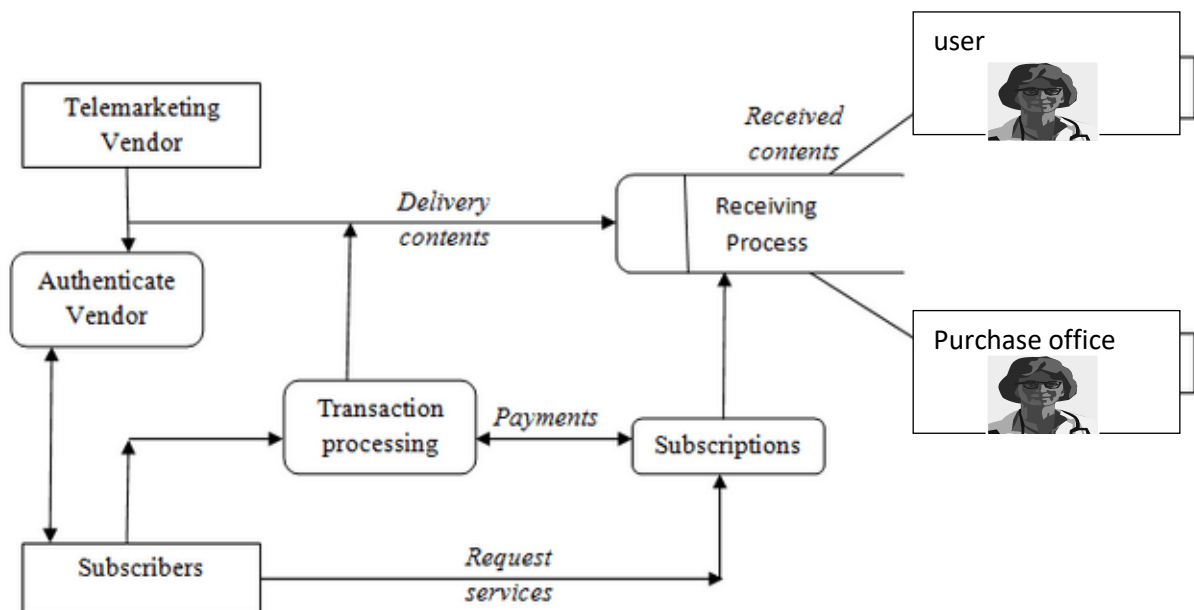


FIG. 4: System Design

Analysis of the new system

- A database is created for all SIM registered as Telemarketers in the Nigeria communication commission.
- Equally a database is also created for “Do not call SIM” these are members of the public who will have come to the “Do not call registry” in the Nigeria communication commission (NCC) to register their SIM that they do not want calls from Telemarketers. The numbers will be blocked from all telemarketers calls.
- Telemarketing lines cannot be used for normal calls (Do not call lines) i.e. calls that are not related to telemarketing. Noting that logging into internet even without SIM can grant one access to telemarketing sites where there is a lot of advert and ADS to lure customers. But the researcher is concerned with telephone calls since a customer who could log into a site has the capability to logout if he/she so wishes.
- Any telemarketing calls made with normal SIMS will be deemed criminal and the individual will be tracked and arrested.
This strategy provides affective control against unsolicited calls and messages
- A system program is coded using Java programming language. This system software detects and identifies Telemarketing calls, do a check if the call channeled to a do not call number, it ends the call or allow if the call is to another telemarketing line, the call is allowed.
- All normal calls made to each other are allowed. Then the following activities shall be done by the system.
- All numbers not registered under Do not call registry are free to telemarketers’ calls
- Call to customers will be done through the company’s customer care lines via the security system (software).

The activities below shall also be considered

- (1) Time of communication.
- (2) Who made the call
- (3) Customer calls
- (4) Content of communication

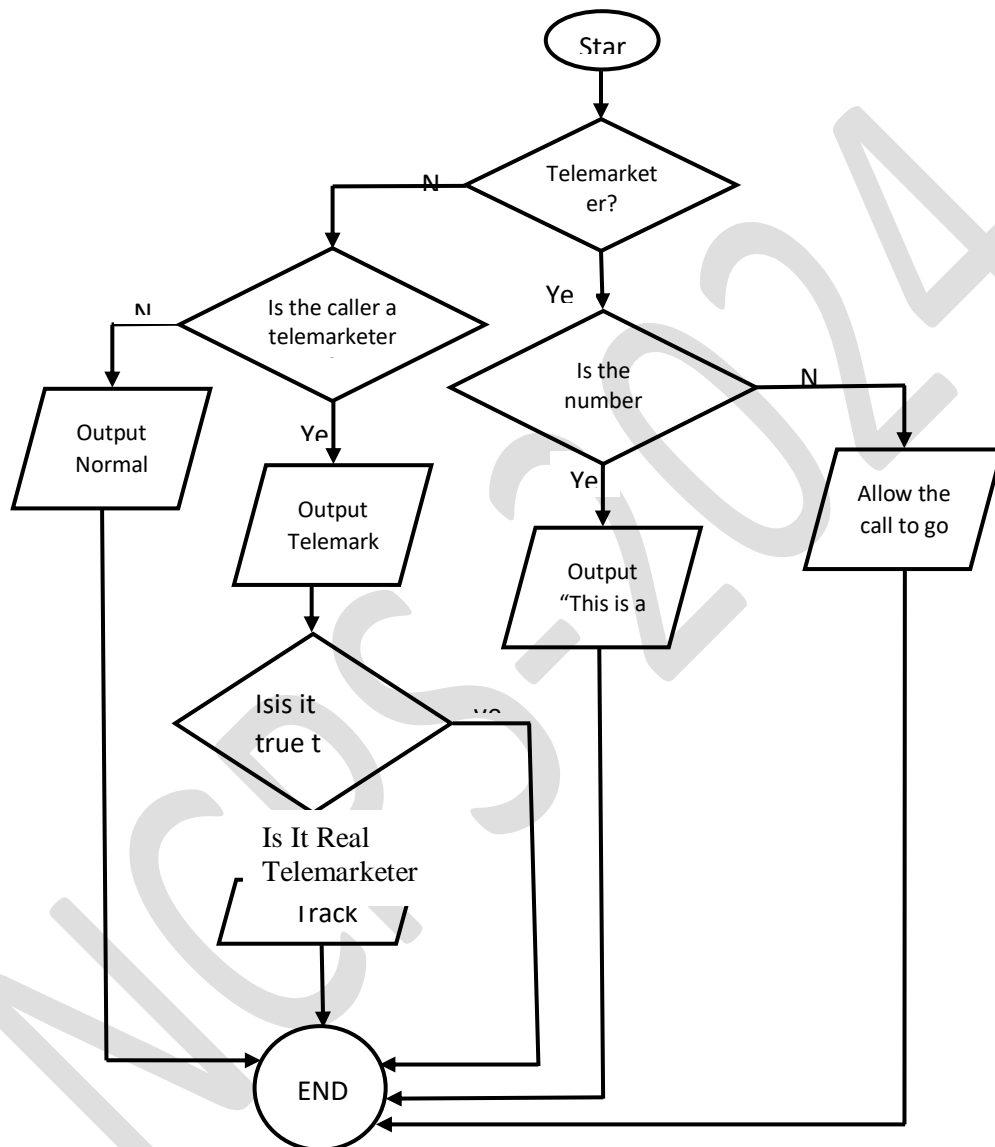


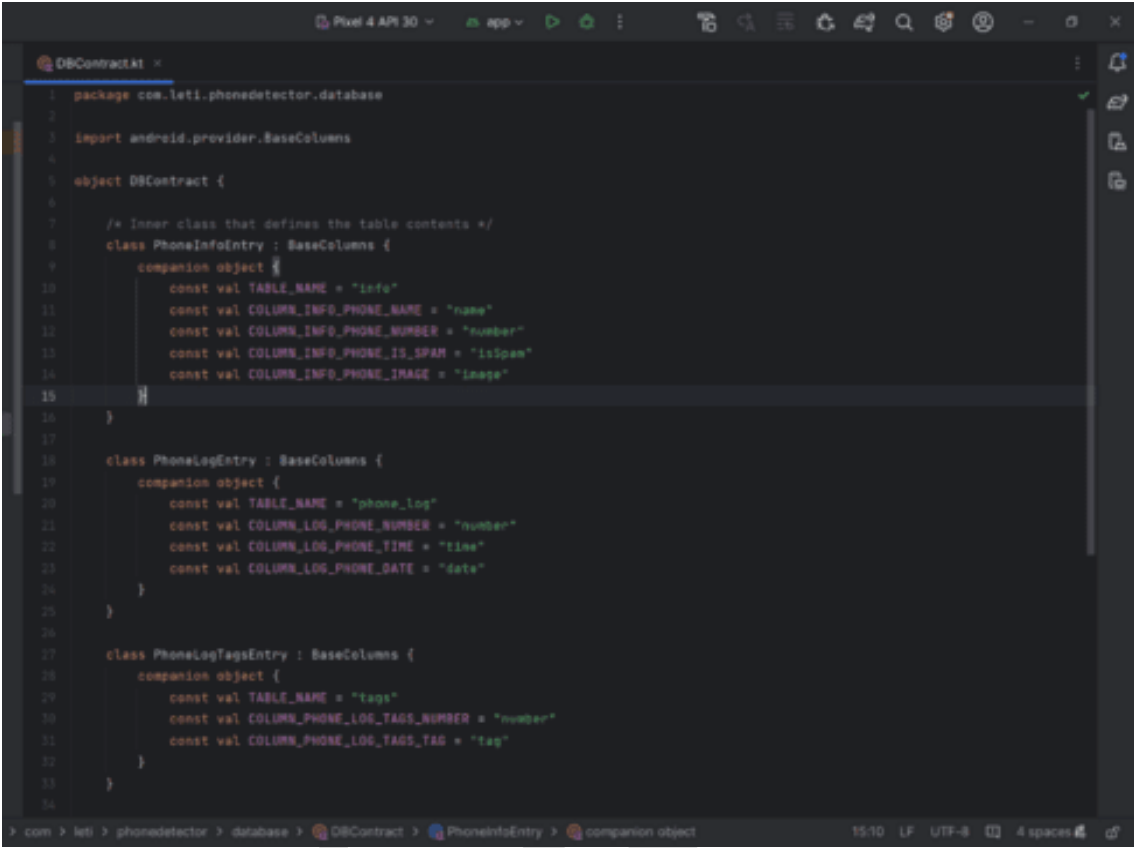
FIG. 7: System Flowchart

Results

Datatest phoneID detector launched for IT Telemarketing security system in the programming environment is shown. Fig 3 below.



Datatext for IT Telemarketing *security system* in the programming environment is shown below



```
1 package com.leti.phonedetector.database
2
3 import android.provider.BaseColumns
4
5 object DBContract {
6
7     /* Inner class that defines the table contents */
8     class PhoneInfoEntry : BaseColumns {
9         companion object {
10             const val TABLE_NAME = "info"
11             const val COLUMN_INFO_PHONE_NAME = "name"
12             const val COLUMN_INFO_PHONE_NUMBER = "number"
13             const val COLUMN_INFO_PHONE_IS_SPAM = "isSpam"
14             const val COLUMN_INFO_PHONE_IMAGE = "image"
15         }
16     }
17
18     class PhoneLogEntry : BaseColumns {
19         companion object {
20             const val TABLE_NAME = "phone_log"
21             const val COLUMN_LOG_PHONE_NUMBER = "number"
22             const val COLUMN_LOG_PHONE_TIME = "time"
23             const val COLUMN_LOG_PHONE_DATE = "date"
24         }
25     }
26
27     class PhoneLogTagsEntry : BaseColumns {
28         companion object {
29             const val TABLE_NAME = "tags"
30             const val COLUMN_PHONE_LOG_TAGS_NUMBER = "number"
31             const val COLUMN_PHONE_LOG_TAGS_TAG = "tag"
32         }
33     }
34 }
```

S

FIG. 9: Datatext for IT Telemarketing security system

This is the Datatext for IT Telemarketing security code it serves as a gateway security to either permit the telemarketer to call or deny the caller to initiate the call. It enables telemarketer to know the status he he/she want to call by indicating this is a Do Not call number or allow the call to be made.

Table 1: Below Result showing lists of Do Not Call and Telemarketer numbers in the Database

Table 1. List of Do Not Call and Registered Telemarketers in Database

Telemarketer	Do not Call
08086279901	8058422573
07084851517	7055912435
07084851517	8065965027
08086279901	7086103461
09095318499	8051866179
09095318499	8023574579
08094723473	8055504324
08094723473	8038722277
09077697628	8038480737
09077697628	7069449787
09080337271	9055317248
09080337271	8033852870
08110248613	8105777255
08110248613	7030215958
07035911154	8078598531
07035911154	8038106202
08038722277	9137337051
08038722277	8062469224
07065276157	8035818685
07065276157	
09011460634	
09011460634	
08133723714	
08133723714	
07033384879	
07033384879	
07036471882	
07036471882	
07036673324	
08160872320	
08037134722	
08052366321	
07038437153	
08076830805	
07033097446	
08035418759	
08032436857	
08164804533	
08131140169	
07089445787	
08062658416	
07048098282	
09049001492	
08183481353	

This a table showing the list of registered Telemarketer who are officially allowed to operate as Telemarketer and a list of Do Not call Number that are officially registered against calls from Telemarketers

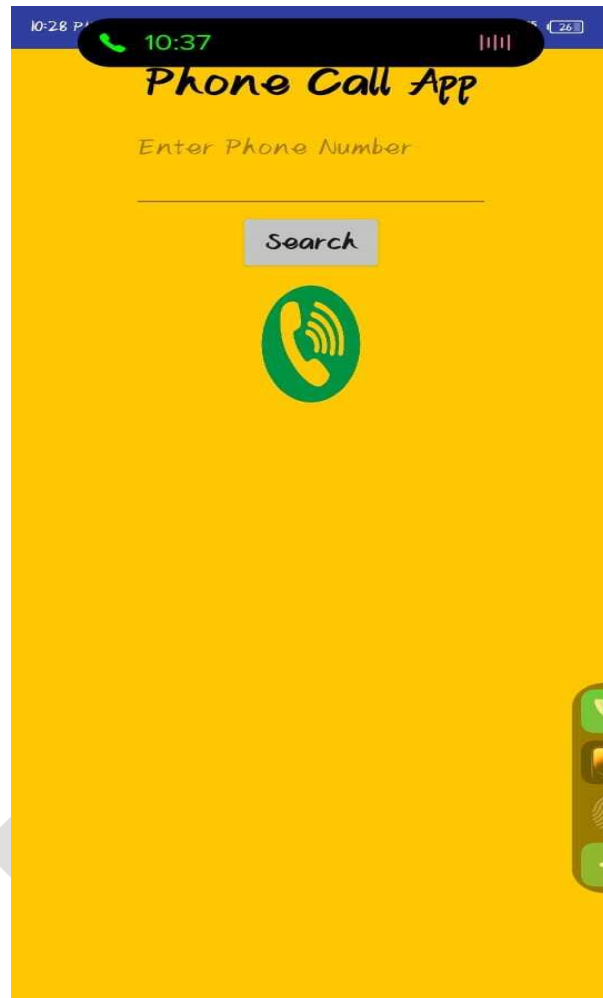


FIG. 10: Default caller Screen IT Security System

Fig 10: shows the screen shot of the result of the default screen where the registered Telemarketer will search for a number he wants to call to authenticate the status of the customer.

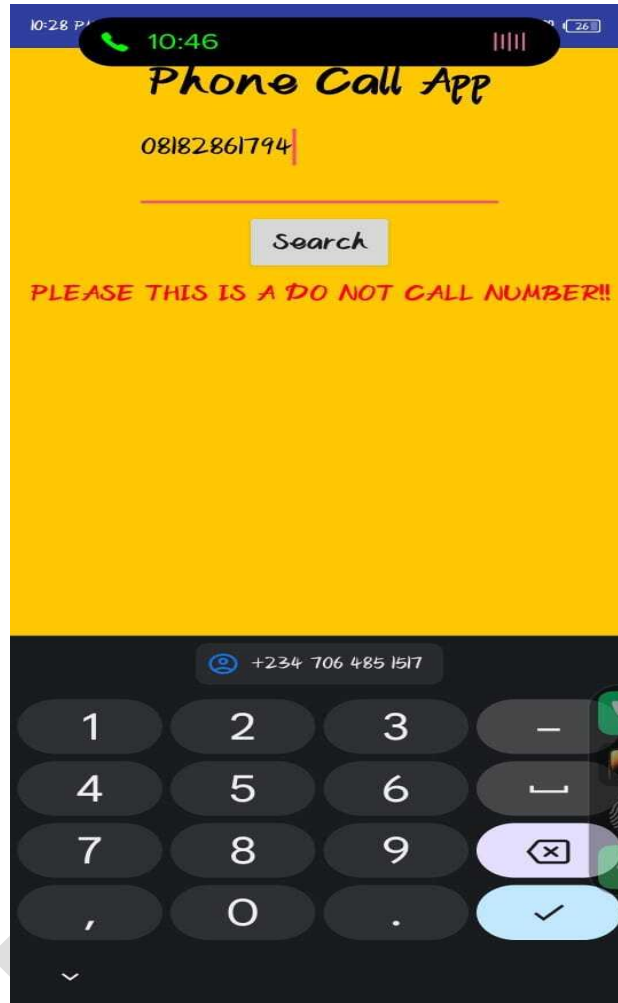


FIG.11: Attempting to Call a DNC Number from IT security App

Fig 11: Shows the screen shot of the Telemarketer attempting to call a Do Not Call (DNC) number and the reply on his phone as a security measure to curb telemarketers calling customers who have registered against calls from them.

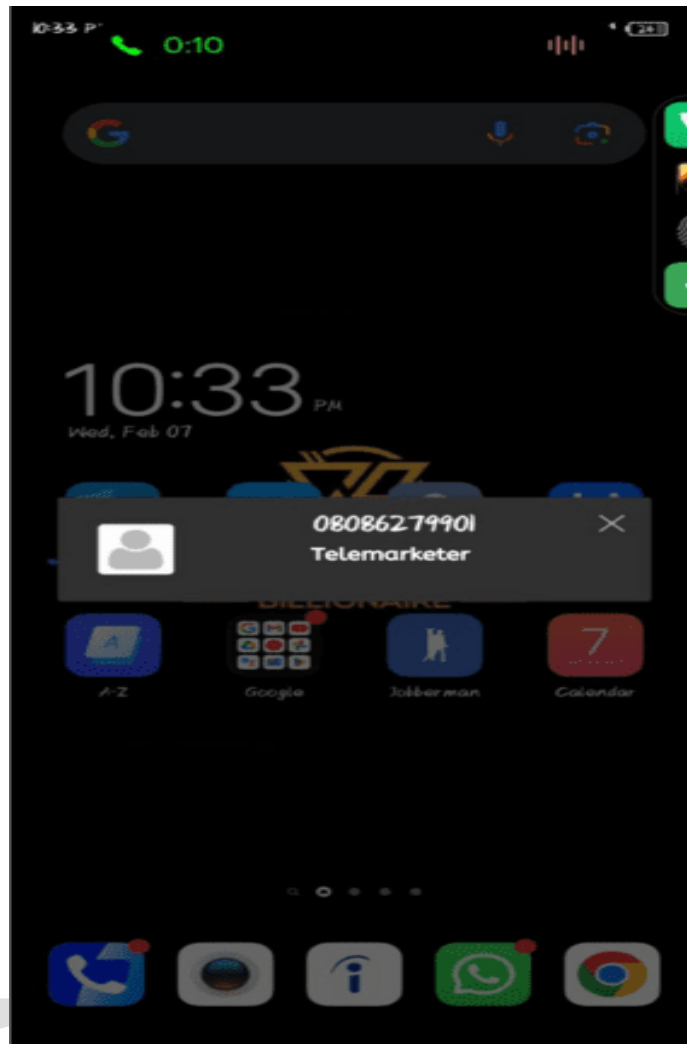


FIG 12: *Incoming Call from a Telemarketer*

Figure 12. Shows the screen of customers who have the IT. Telemarketing security app installed in their phone. This will always alert the customer when a telemarketer call is coming to their phone as a security measure.

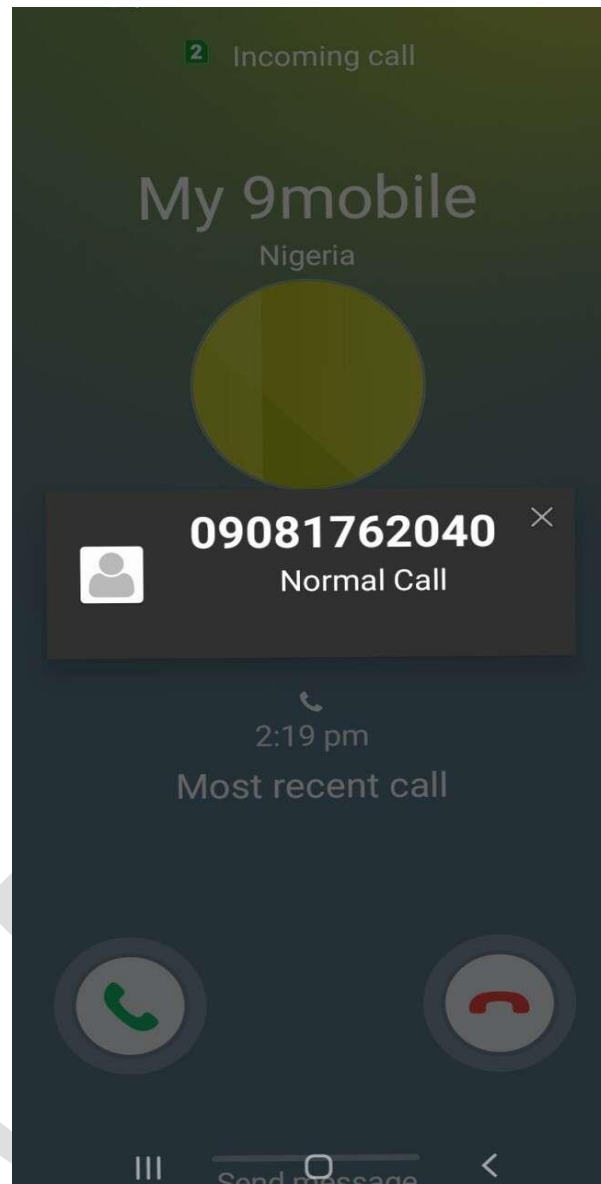


FIG. 13: Incoming normal call

Figure 13 shows the screen shot of an incoming call from a customer who is not registered as a telemarketer to a customer who has the IT security for Telemarketers App installed in the phone

Discussions and findings

In this research work the researcher has highlighted on the problems that existed in the present telemarketing business system. The researcher indicated that the Nigerian internet is fill with fraught, scams and malicious activities.

The researcher has developed a good information technology (IT) Security system for telemarketing. The IT security System has given different approaches to telemarketing and how this information can be managed securely to prevent security breaches on sensitive data and information resources. The researcher said the functionality of data integrity is dependent on the security measure that have been put in place to allow for secured business transactions and how proper information dissemination across domain and across business plate formed will yield better result.

References

- Adams, R. L. (2017). *10 Marketing Strategies to Fuel Your Business Growth*. Entrepreneur. <https://www.entrepreneur.com/growing-a-business/10-marketing-strategies-to-fuel-your-business-growth/299335>
- Adela, M., & Mian, M. B. (2011). *Consumers' Behavior Towards Telemarketing: A Case Study of Developed and Developing Countries*. In *liu.diva-portal.org*. <https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-70946>
- Adeyemi, A. (2023). *Telemarketers, loan app firms face prosecution over data breaches*. Retrieved from <https://guardian.ng/telemarketers-loan-app-firms-face-prosecution-over-data-breaches/>
- Alex, F. (2019). *Undeniable value of phone calls*. Infinity. Retrieved from <https://www.infinity.co/us/resources/news-and-views/mary-jane-copps-the-undeniable-value-of-a-phone-call-interview>
- Andrea, M., & Tara S. (2021) *Telemarketing history*. Retrieved from <https://www.en.wikipedia.org/wiki/Telemarketing>
- Artaya, I Putu. (2019). *Telemarketing Media Communication In Attracting Customers*. https://www.researchgate.net/publication/331429032_Telemarketing_Media_Communication_In_Attracting_Customers.
- Baker, M. (2017). *Marketing : an introductory text : Baker, Michael J : Free Download, Borrow, and Streaming : Internet Archive*. Internet Archive. <https://archive.org/details/marketingintrodu0000bake>
- Better Business Bureau. (2021). *Online Purchase Scams 2021 | BBB Institute for Marketplace Trust*. International Association of Better Business Bureaus. <https://www.bbb.org/all/bbbi/online-purchase-scams-2021>
- Bischoff B. (2019) a paper on *Common Telemarketing scam and how to spot them, privacy advocate and VPN expert @ Pabischoff 2019*.
- Bischoff, P. (2019). *Common telemarketing scams and how to spot them | FightIdentityTheft*. Comparitech. <https://www.comparitech.com/identity-theft-protection/common-telemarketing-scams/>
- Bourgeois, D. T. (2019). *Information Systems for Business and Beyond*. Biola University.
- Chisnall, P. (2018). *Marketing research; analysis and measurement: Chisnall, Peter M: Free Download, Borrow, and Streaming: Internet Archive*. Internet Archive. <https://archive.org/details/marketingresearch00chis>
- Cliford, C. (2020). *The Ultimate List of Types of Marketing [40 and Counting]*. Blog.hubspot.com. <https://blog.hubspot.com/marketing/marketing-types>

- Dharmani, I. G. A. A. N. (2010). *Telemarketing skills*. Retrieved from <http://ayurai-management.blogspot.com>
- Dynamic System Development Method (DSDM). (2024). DSDM. SlideShare; Slideshare. <https://www.slideshare.net/slideshow/dsdm-12836937/12836937>
- Enyia, & Abang, T. A. (2018). Telemarketing and Consumer protection in Nigeria: A Case for the Domestication of the U.S. Telephone Consumer Protection Act, (TCPA, 1991). *European Journal of Business and Management*, 10(9), 112–126.

NCPS-2024

Internet of Things (IoT)-Based Food Crop Security Model for Grazing Monitoring and Reporting using Convolutional Neural Network

Moses Adah Agana and Ogar Akima Akima

Department of Computer Science, University of Calabar, Calabar, Nigeria

[Corresponding Email: akimaogar@gmail.com](mailto:akimaogar@gmail.com)

Abstract

The persistent conflicts between crop farmers and herders in Nigeria, rooted in disputes over arable land resources, have led to significant security challenges, surpassing the toll of the Boko Haram insurgency. This conflict, marked by the frequent killing of crop farmers and the destruction of crops during open grazing by Fulani herders with their cattle, prompted the development of an innovative solution. In response, an Internet of Things (IoT)-based food crop security system for grazing monitoring and reporting was developed, using a network of interconnected computing devices, mechanical and virtual machines, objects, animals, and living beings that can be remotely controlled via the Internet, utilizing a Mobile phone as a user interface and reporting system for crop farmers. Additionally, a Flask server serves as an Internet gateway for IoT devices. Applying the Agile methodology and a deep learning convolutional neural network (CNN), specifically the You Only Look Once (YOLO) v5 algorithm, a YOLOv5 model was adopted. This model targets the detection of grazing animals entering farmland, thereby preventing damage to farmers' food crops. The YOLOv5 algorithm model underwent comprehensive training, validation, and testing, comparing various weights for speed, accuracy, and efficiency in real-time animal detection, identification, and reporting of Fulani herders' cattle on farmlands. Results from the pre-trained YOLOv5 algorithm model demonstrate the Pi Camera capturing and reporting cow images with confidence scores of 0.82 and 0.72. This unequivocally confirms the model's effective identification and localization of objects within the images. Implementing this automated system is expected to alleviate conflicts between farmers and herders while enhancing food security. The proactive nature of the system, combined with its real-time monitoring capabilities, signifies a pivotal step toward addressing the intricate challenges posed by crop farmers-herders conflicts in Nigeria.

Keywords: Agriculture, Farmer, Herder, Internet of Things (IoT), Machine Learning, Computer Vision, Convolutional Neural Network (CNN), You Look Only Once Algorithm

1. Introduction

Agriculture, farming, and animal husbandry stand as a fundamental occupation throughout human history. The term "agriculture" encompasses all facets within the linear progression of the food chain essential for human sustenance (Mohit & Rahul, 2017). It represents a pivotal human endeavor primarily focused on producing food, fiber, fuel, and various other commodities through the controlled utilization of terrestrial plants and animals. Food remains the fundamental energy source for all living organisms, for millennia, human civilization has thrived within and exploited rangelands, spanning over 40% of the earth's land surface (FAO, 2016). Agriculture is the amalgamation of science, art, and practice involving soil cultivation, crop production, and livestock rearing. The discipline bifurcates into two principal branches: crop production and livestock production. Crop production constitutes the sector dedicated to cultivating crops for food

and fiber. It is a widely applied science among global farmers, enabling them to manage crops effectively. By enhancing crop production and management techniques, this field empowers farmers to augment food output (Collegedunia Team, 2022).

Livestock production from farming stands as one of the world's key sectors impacting the economy, food security, and livelihoods (Falvey, 2017). It is essentially the management and breeding of domesticated animals for their meat and other derivatives (Vigneswarria et al., 2021). In Nigeria, livestock farming assumes critical importance within the agricultural landscape, expected to contribute to national development significantly. Grazing, a traditional method in animal husbandry, involves allowing domestic livestock to freely roam and consume natural vegetation, transforming indigestible cellulose within grass into usable farm produce. The pastoralist practices in Nigeria predominantly involve Fulani herders who guide their cattle on foot across extended periods, traversing from the northern regions to the verdant, rain-fed southern areas of the country, enabling their herds to graze en route (Chukwuemeka et al., 2018).

In Nigeria, conflicts arise due to pastoralist practices wherein livestock graze on food crops, leading to significant clashes between farmers and herders (Oli, et al., 2018). The conflict between farmers and herdsman stands as a persistently prevalent and violent issue in Nigeria, notably involving disputes over grazing fields and land-related matters, constituting a significant proportion of ongoing national insecurity. In modern times, the Internet of Things (IoT) plays a pivotal role in various sectors, especially in environmental and agro-industrial domains (Talavera et al., 2017). It facilitates enhanced insights and the ability to alter our surroundings, notably revolutionizing smart agricultural practices (Andzio&Xiong, 2019). Within the realm of technological advancements, computer vision and machine learning emerge as essential tools (Kolesnikova, 2023). Computer vision seeks to decipher visual scenes and objects through signals from video cameras, while machine learning enables tasks without explicit programming (Redmon & Farhadi, 2017).

Deep learning, a subset of machine learning, leverages algorithms and statistical methods to transform input data into meaningful output. Convolutional neural networks (CNNs) within deep learning have demonstrated high accuracy and effectiveness, particularly in image recognition and object detection (Ren & Wang, 2022). Among the networks within computer vision, the YOLO (You Only Look Once) algorithm stands out for its real-time object detection capabilities using convolutional neural networks, acclaimed for its speed and accuracy (Karimi, 2021). Present research endeavors aim to employ an IoT-based embedded system to safeguard food crops from destruction by cattle entering farms, particularly attributed to Fulani herders

2. Methodology

The Agile methodology was adopted for this study due to its iterative and incremental approach to software development. Agile methodology prioritizes the rapid and frequent delivery of functional products, accommodating evolving requirements and fostering continuous feedback from end users. This approach facilitates flexible responses to change, emphasizing continual planning, learning, improvement, team collaboration,

evolutionary development, and early product delivery (Hamilton, 2023). In the prevalent situation, various approaches are employed to detect animal intrusion into farmlands:

- i. Human observation serves as a primary method for detecting animal intrusion.
- ii. Another method involves employing repellents, such as chemicals, smoke, organic substances, scarecrows, and sonic electronic repellants, to mitigate the destruction of food crops by animals.
- iii. The utilization of microcontroller-based systems like Arduino constitutes another approach.

The proposed system is meant to address animal intrusion on farmland and involves an Internet of Things (IoT)-based food security system designed to detect and respond to such incidents effectively. The system comprises of several components and functionalities such as:

- i. Utilization of an ultrasonic sensor capable of detecting animals from distances exceeding two meters.
- ii. Implementation of the YOLOv5 algorithm to detect the presence of cattle and individuals carrying weapons like sticks, guns, and knives.
- iii. Image capture by a camera followed by image processing through convolutional neural network layers to compare with pre-trained datasets of the YOLOv5 model.
- iv. Activation of live recording upon correspondence with the YOLOv5 model datasets, accompanied by the activation of a buzzer to deter the animal.
- v. Integration of a Raspberry Pi 3b+ with Wi-Fi capabilities to facilitate communication with the crop farmer's Mobile phone using Flask, enabling the crop farmer to receive push notifications and access a live stream of their farm.
- vi. Communication between the crop farmer and security personnel or villagers to report incidents occurring on the farmland.

2.1 System Schematic Diagram

Figure 1 delineates the functionality and operational workflow of the system: The image recognition algorithm is designed to undergo training for the identification of specific entities, including cattle, herdsman, and various weapons such as rifles, shotguns, and knives. Upon detection of these predefined objects, the system initiates a notification to a mobile phone using Flask. Subsequently, the Farmer can directly access a live video stream through this mobile phone interface and notify security guards to respond promptly to prevent illegal grazing on food crops.

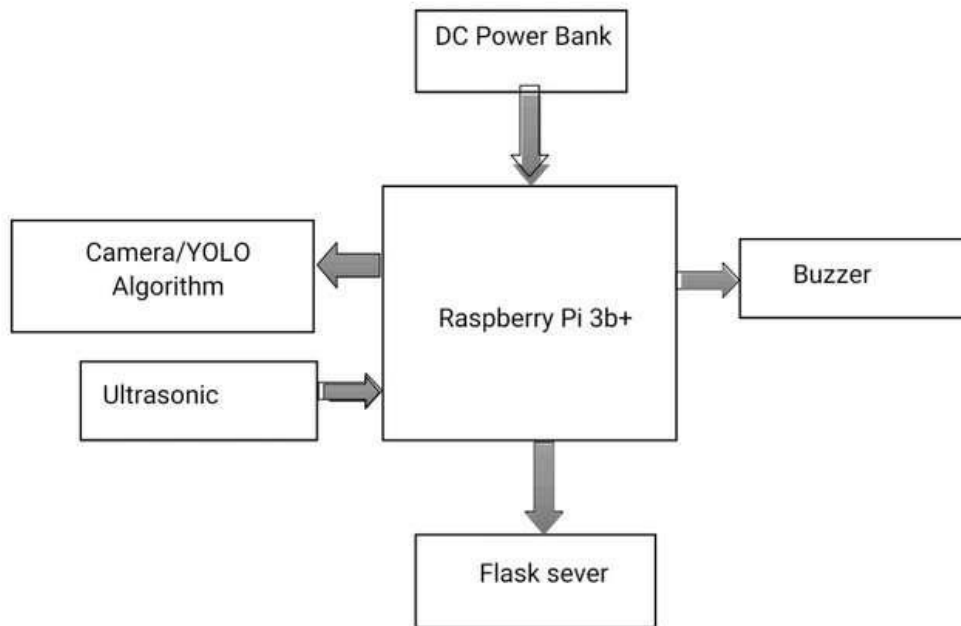


Figure 1: Operational workflow of the IoT-Based Food Crop Security Model for Grazing Monitoring and Reporting

The Use Case diagram in Figure 2 visually illustrates the system's functionalities from the user's perspective. These components collectively facilitate the following operations: detecting animals from a distance using the ultrasonic sensor, capturing animal images with the camera, activating the buzzer as a deterrent measure, and sending notifications to both the crop farmer and security guard when an animal is detected inside the farm.

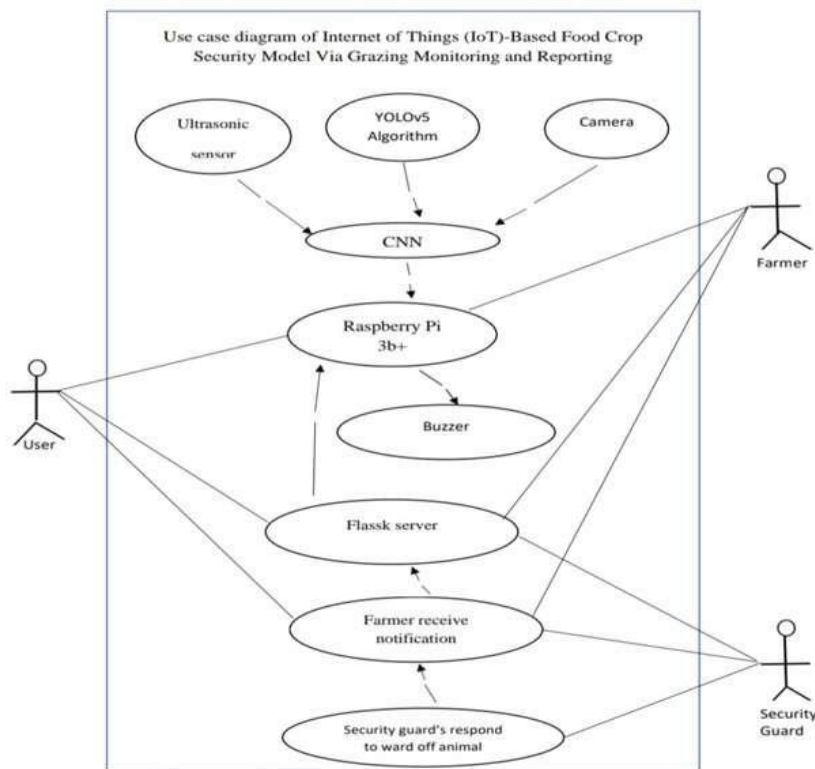


Figure 2: Use Case Diagram of Internet of Things (IoT)-Based Food Crop Security Model for Grazing Monitoring and Reporting

The Class diagram in Figure 3 illustrates the structure, interrelationships, attributes, and methods inherent in this system. It delineates the classes, interfaces, and associations between components such as the camera, YOLOv5 algorithm, ultrasonic sensors, Raspberry Pi 3b+, buzzer, and mobile phone, highlighting their interconnections and functionalities within the system.

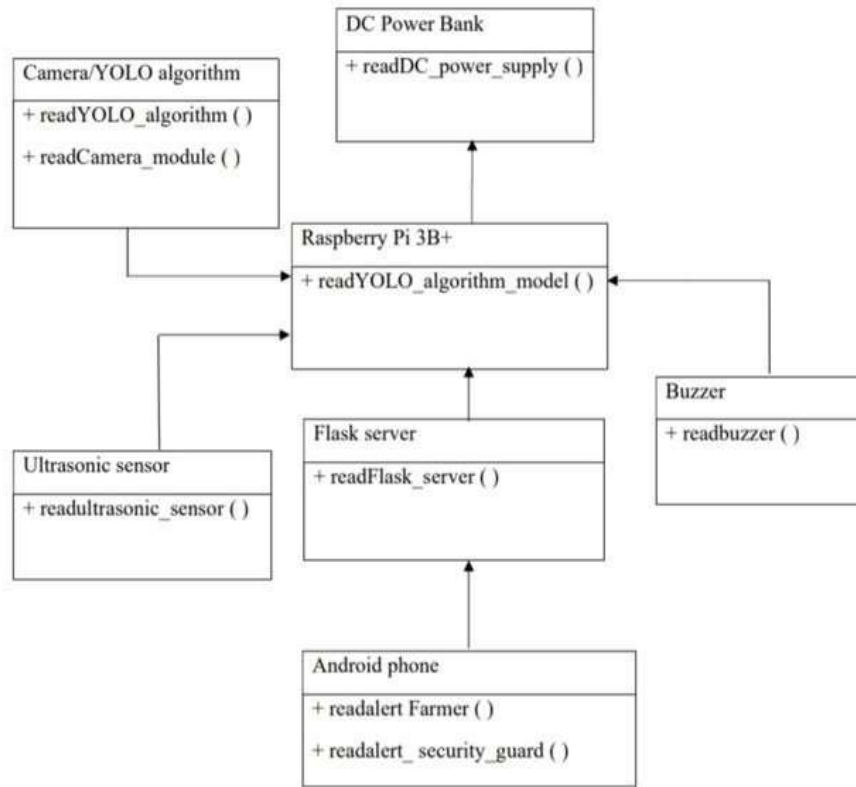


Figure 3: Class Diagram of Internet of Things (IoT)-Based Food Crop Security Model for Grazing Monitoring and Reporting

The image detection flowchart involves a series of crucial steps essential for the system's functionality. It consists of various phases, namely: image collection of data, image annotation, model training, and weight inference. This comprehensive flowchart encompasses essential steps for image detection, facilitating the development and training of a robust object detection model. The diagram depicted in the Figure 4 illustrates an image detection flowchart.

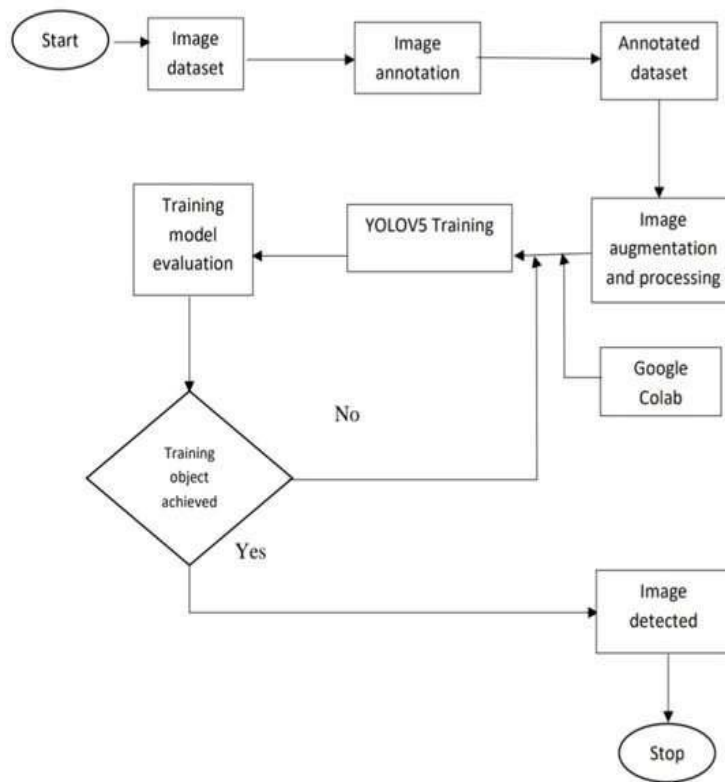


Figure 4: Image Detection Flowchart

The system flowchart delineates the sequential stages integral to the developed system aimed at detecting and preventing animal intrusion on farmland. Figure 5 illustrates the flowchart depicting the process designed to detect and avert animal intrusion on farmland.

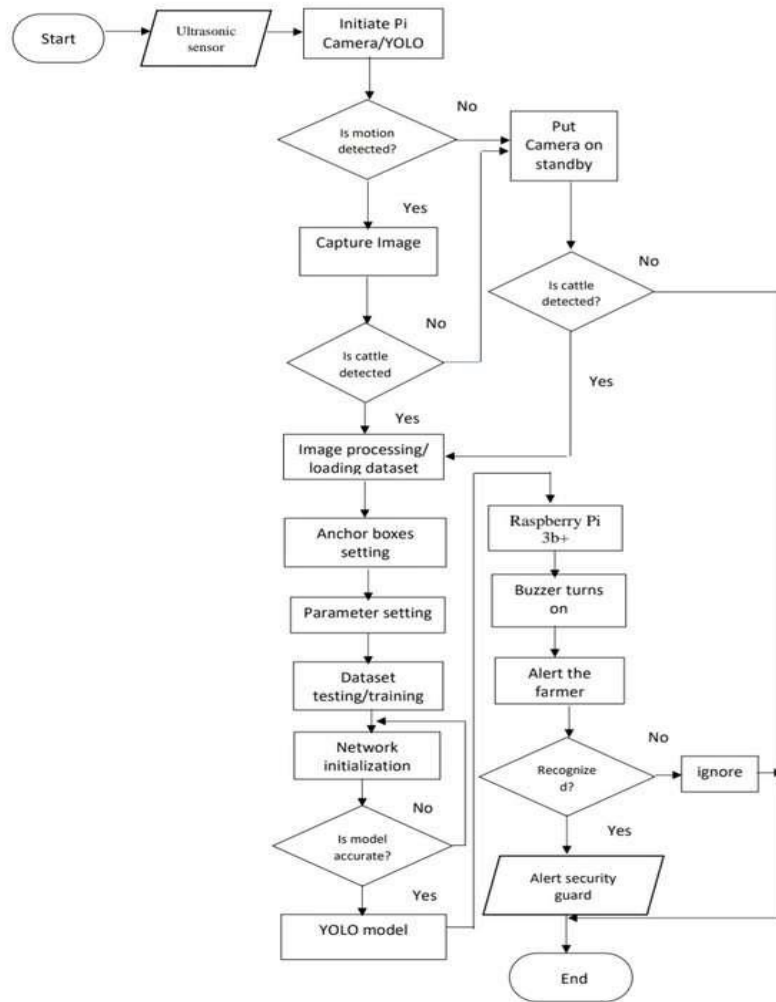


Figure 5: System flow chart

Figure 6 depicts a visual representation of the security flowchart within the Flask application. The system is equipped with user login credentials to gain access to the Mobile phone, necessitating the input of an email address and password. In case of three consecutive failed login attempts, the system prompts an alert message indicating an incorrect email address and password, instructing the user to retry.

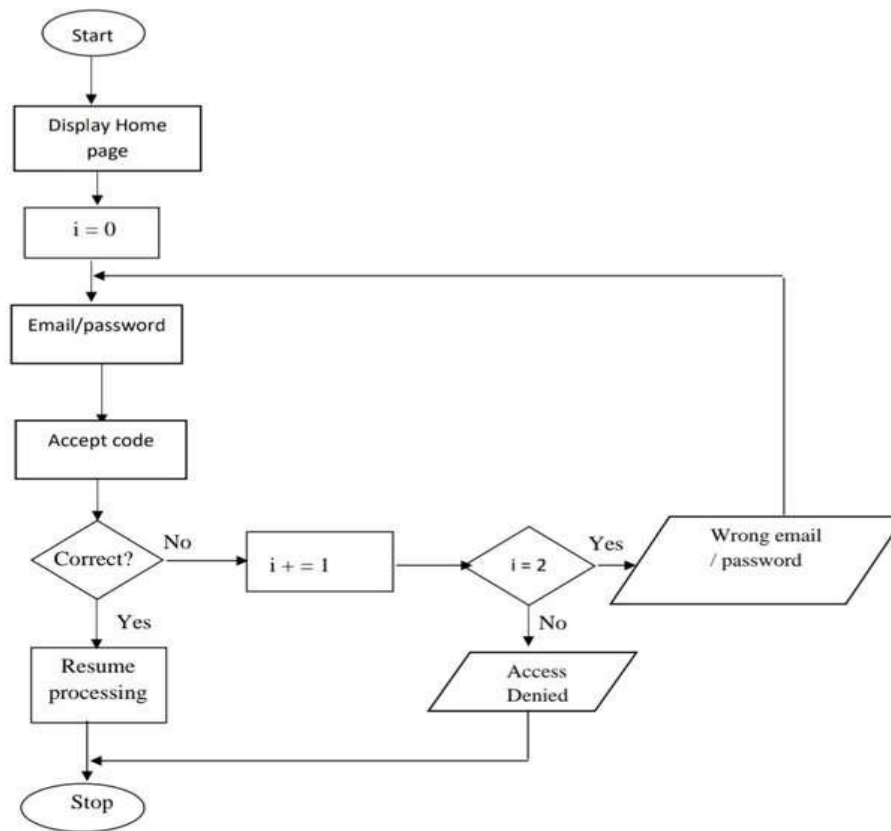


Figure 6: Security flow chart

3. Results

The system, crafted to fortify food crop security through meticulous grazing monitoring, exemplifies a professional application of cutting-edge technology. With a seamless integration of ultrasonic sensors, a buzzer, and a Raspberry Pi 3b+ paired with a Pi Camera using Python and MySQL, the implementation of Flask ensures the delivery of real-time Telegram notifications to crop farmers upon animal detection. The system's robust performance is underpinned by the utilization of YOLOv5, with each weight trained through 200 epochs on Google Colab, optimizing both speed and accuracy for enhanced efficiency in animal detection. Results of table 1 show, the YOLOv5 nano, small, medium, large, and extra-large (n, s, m, l, x) model and Screen captures that showcase the system's user-friendly design, highlighting a commitment to innovative and efficient food crop security.

The general implication in comparison with existing works shows that, this system introduces significant advancements in terms of speed, accuracy, and usability, primarily due to the choice of the YOLOv5 object detection model, ultrasonic sensors for precise environmental awareness, and an integrated mobile interface for remote management. These improvements enhance both the real-time monitoring of farmlands and the responsiveness to potential threats, offering a more reliable solution for food crop security.

Table 1: Result of YOLOv5 n, s, m, l, x Model

YOLOv5 model	YOLOv5n	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv5x
Precision	0.98	0.969	0.884	0.973	0.912
Recall	0.912	0.98	0.98	0.96	0.97
F1-score	0.91	0.93	0.91	0.92	0.93
mAP_0.5	0.932	0.930	0.929	0.928	0.935

The assessment of the YOLOv5x model is presented below, providing a comprehensive overview of object detection metrics. This includes the graphical confusion matrix, precision, recall, and F1 score, offering a thorough evaluation of the model's performance.

Raspberry Pi Home Page

Figure 7 of the Raspberry Pi home page serves as the main interface for the Raspbian operating system, providing the platform for configuring IoT settings.



Figure 7: Raspberry Pi Home Page

Login Page

Figure 8 of the login page serves as a secure and controlled entry point to the farm monitoring system platform, enabling a user-friendly interface for secure authentication.

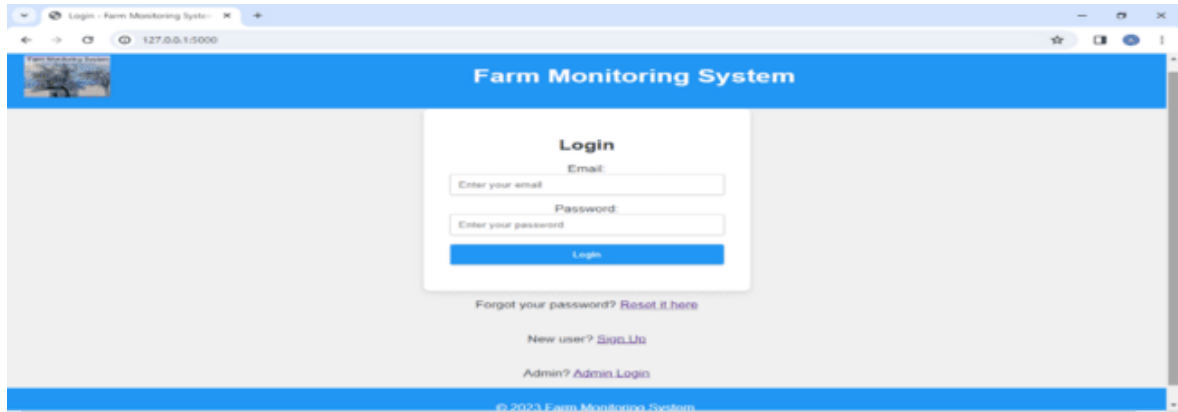


Figure 8: Login Page

Home Page

Figure 9 of the home page of the farm monitoring system is designed to offer users an informative overview of the platform.

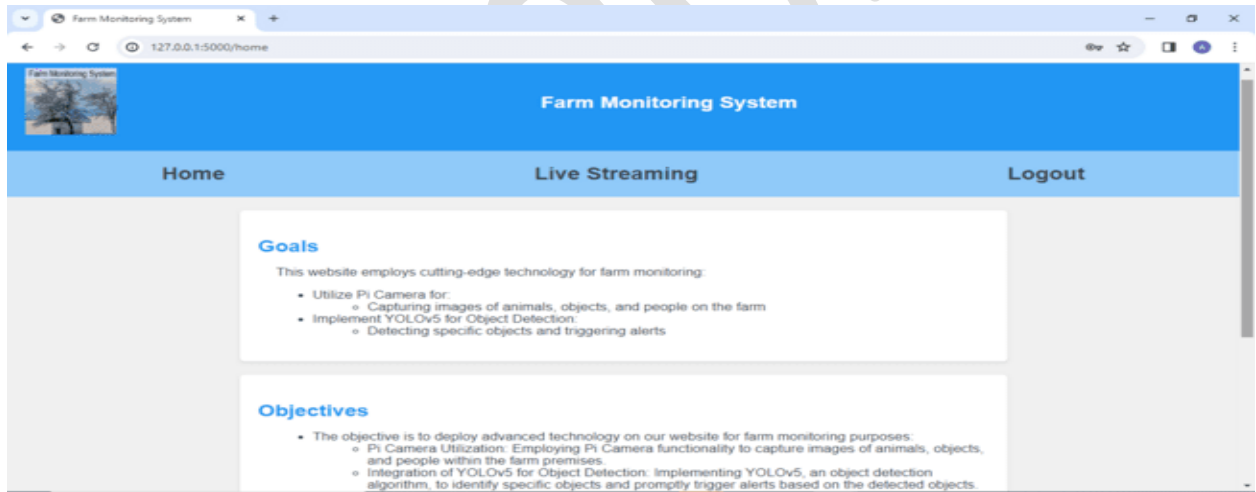


Figure 9: Home Page

Figure 10 of the live streaming page of the farm monitoring system provides users with real-time video content captured by the Pi Camera, along with additional features for saving, capturing, and sharing media. Animal detection from the images displayed on the live streaming page of the farm monitoring system is illustrated in Figure above. The YOLOv5 algorithm model demonstrates the Pi Camera capturing and reporting cow images with confidence scores of 0.82 and 0.72. This signifies a high level of confidence in the model's accurate identification and localization of an object within the image.



Figure 10: Live Streaming Page

Telegram Notification

Figure 11 shows Telegram notification received on the mobile phones offarmers and security personnel, when animals are detected on their farmland.

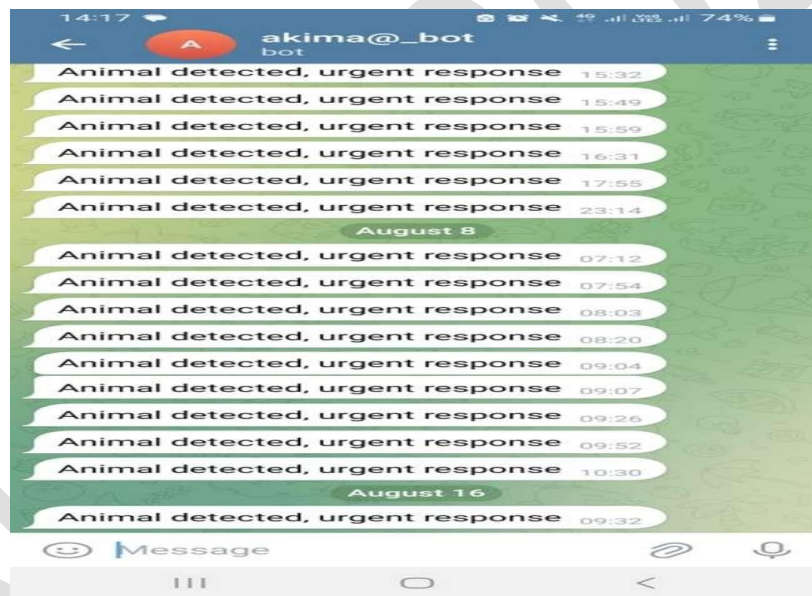


Figure 11: Telegram Notification

Conclusion

In conclusion, this study conducted a thorough and exhaustive examination of an Internet of Things (IoT)-based food crop security model designed for grazing monitoring and reporting, employing a convolutional neural network (CNN). The primary focus was on establishing preventive measures against animal attacks on farmers' food crops, thereby mitigating potential conflicts between farmers and herders. The implementation of this IoT-based food crop security system empowers crop farmers to vigilantly monitor

their crops, protecting against the intrusion of Fulani herders and their cattle. The system integrates a diverse array of technologies and components to bolster farm security, enhance situational awareness, and facilitate rapid responses to Fulani herders and their cattle encroaching on farmland. Key highlights include:

- i. **Robust Security Measures:** The system incorporates robust security measures, encompassing real-time monitoring, detection of Fulani herders and animals, and immediate alerts. This capability empowers crop farmers to respond promptly to animal invasions on their farmland.
- ii. **Efficiency Enhancement with YOLOv5:** YOLOv5 object detection technology significantly improves the system's efficiency by accurately identifying objects within the camera feed. This results in the reduction of false alarms and an enhancement in the accuracy of detecting herders and animals.
- iii. **Mobile Phone Interface:** A mobile phone acts as a user-friendly interface, enabling crop farmers to seamlessly monitor, control, and respond to animal invasions on their farmland.

REFERENCES

- Anand, R., Anand, V., Mridhul, M., Roshan, D., Jisha, J. & Robin, J., (2022). Animal Trespassing Detection System. *Journal of Engineering Research Technology (IJERT)*, 10 (4), Volume 10, Issue 4, 690-694
- Andzio, E. & Xiong, S., (2019). Agriculture monitoring system using smart and innovative farming: A real-time study. *International Journal of Scientific & Technology Research* Volume 8, Issue 12, 1214-1219
- Chukwuemeka, E., Aloysius, A. & Eneh, M., (2018). The Logic of Open Grazing in Nigeria: Interrogating the Effect on Sustainable Development. *International Journal of Family Business and Management Studies*, 2(1), 1-17.
- Collegedunia Team, (2022). Crop Production: Definition, Factors, Harvesting, and Storage. Retrieved October 11, from <https://collegedunia.com/exams/crop-production-definitionfactors-harvesting-storage-science-articleid-1160>
- Falvey, L., (2017). Food Security: The Contribution of Livestock. Retrieved April 05, from <https://www.researchgate.net/publication/276308353>
- FAO., (2016). State of the World's Forests 2016. Forests and agriculture: land-use challenges and opportunities. Rome.
- Ibam, E., Boyinbode, O. & Aladesiun, H., (2023). IOT-Based Farmland Intrusion Detection System. *Buana Information Technology and Computer Sciences (BIT and CS)*. Vol.4 No.2
- Karimi, G., (2021). Introduction to YOLO Algorithm for Object Detection. Retrieved April 2021, from <https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/>
- Kolesnikova, I., (2023). Revolutionizing Industries: Computer Vision Applications and Real-world Case Studies. Retrieve August 15, from <https://mindtitan.com/resources/blog/computer-vision-applications/#main-content>

- Mohit, k. & Rahul, G., (2017). A Review Paper on Internet of Things-based Application Smart Agricultural System. *International Journal of Latest Engineering and Management Research (IJLEMR)*, 2 (4), 69-71.
- Neale, B., Parlewar, P., Sharma, R., Zararia, D., Mishra, J., Itankar N., Wazalwar, R. &Harkare, A., (2021). Farm Protection and Monitoring with the Smart Animal Intrusion System. *International Journal of Mechanical Engineering*, 6 (3), 2851-285
- Oli, N., Ibekwe, C. & Nwankwo, I., (2018). Prevalence of Herdsmen and Farmers Conflict in Nigeria. *International Journal of Innovative Studies in Sociology and Humanities (IJISSH)*, Volume 3, Issue 1.
- Paramasivam, K., Krishnaveni, S., Sowndarya, S. & Kavipriya, E., (2020). Convolution Neural Network Based Animal Detection Algorithm for Partial Image. *Egaeum Journal*, Volume 8, Issue 6, 1461-1469
- Redmon, J. & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 6517–6525. <https://doi.org/10.1109/CVPR.2017.690>
- Ren, J. & Wang, Y. (2022) Overview of Object Detection Algorithms Using Convolutional Neural Networks. *Journal of Computer and Communications*, 10, 115-132. doi: 10.4236/jcc.2022.101006.
- Sabeenian, R., Deivanai, N. & Mythili, B., (2020). Wild animals' intrusion detection using deep learning techniques. *Int. J. Pharm. Res.*, vol. 12, no. 4, 1053-1058
- Santhoshi, K. & Bhavan, S., (2018). Intruder recognition in a farm through wireless sensor network. *International Journal of Advance Research, Ideas and Innovations in Technology*. Volume 4, Issue 3, 667-669
- Selvamuthukumaran, N., Evangeline, J., Thriambika, B. & Vishali, B., (2021). Intelligent Animal Detection System Using IoT and Deep Learning. *International Research Journal of Modernization in Engineering Technology and Science*, 3 (4), 2433-2438
- Suresh, B., Kamrujama, A. & Dhanush, V., (2022). Animal Tracking and Alert System. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*. Volume 10, Issue 7, 675-681
- Talavera J., Tobón L., Gómez J., Culman M., Aranda J., Parra DL., Quiroz L., Hoyos A. & Garreta L., (2017). Review of IoT applications in agro-industrial and environmental fields. *Computers and Electronics in Agriculture*. 142, 283–297
- Vigneswaria, T., Kalaiselvib, N., Mathumithac, K., Nivedithac, A. &Sowmianc, A., (2021). Smart IoT Cloud Based Livestock Monitoring System: A Survey. *Turkish Journal of Computer and Mathematics Education*, Vol.12 No.10, 3308-3315.

A STACK HETEROGENOUS ENSEMBLE (SHE) MODEL FOR PREDICTING ROAD TRAFFIC VIOLATIONS IN SOUTH-EASTERN NIGERIA

E. A. Emmanuel, Ibobo N. Simon, Odo Kelvin

Department of Computer Science, Faculty of Physical Science, University of Calabar

Corresponding Emails: edimazom@unical.edu.ng;

ibobosimon@gmail.com;ikechukwukelvin972@gmail.com

Abstract

Road traffic violations is among the leading cause of mortality rate and destruction of public infrastructure in Nigeria. Managing traffic offenses by traffic agencies is essential for the smooth movement of goods and services around the globe. Relying on traffic personnel to avert road crashes because of traffic violations, introduces so many uncertainties. To address these challenges, this research develops a SHE method to predict road traffic violations and draw insights from indigenous dataset gathered. Moreover, preprocessing routines, such as data imputation, duplicates and outliers' removal, and normalization, were employed. We evaluated four machine learning models- Random Forest, Extreme Gradient Boost, Multi-Layer Perceptron, and Stack Heterogenous Ensemble, using FRSC traffic dataset, an unexplored resource for road traffic crashes prediction. Following a comprehensive machine learning developmental life cycle, the researchers assessed the models. Findings from the implementation shows the successive outcomes: Multi-Layer Perceptron – 50% accuracy, 48% precision, 50% sensitivity; Random Forest – 59% accuracy, 56% - precision, 59% - sensitivity; Extreme Gradient Boost – 65% - accuracy, 63% - precision, 65% sensitivity; Stack Heterogenous Ensemble – 87% accuracy, 85% precision, 87% sensitivity. However, Stack Heterogenous Ensemble Classifier perform better in terms of accuracy, precision, and sensitivity than the other three algorithms used. Additionally, further analysis using feature importance gives insights how certain attributes interact and influences the base model's prediction outcomes. Finally, utilizing a local dataset reinforces the relevance of our findings. Implementing the Stack Heterogenous Ensemble algorithm in road traffic violations settings may enhance road users and outcomes. Given sufficient datasets from similar sources from each of the geo-political regions, SHE could be adapted to develop a life-saving predictive system.

Keywords: Accuracy Score, CNN, Deep Learning, FRSC, LSTM, ML, MLP, Random Forest, Road Traffic Offences, SHE, Violations, XGBoost.

1 Introduction

In Nigeria and the globe in general, road traffic violations occur when road users don't follow the rules of the road. These rules, just like everyday routines that keeps us safe, are there to protect everyone. Adhering to road principles helps prevent accidents, which is why there are consequences for not doing so. Consequently, road transport industry is an essential component of any economy in the world, contributing to the welfare and development of people and serving as a necessary condition for growth of any nation's economy. According to the World Health Organization, road transportation assists nations and individuals by facilitating the movement of

goods and people, enabling deeper access to jobs, economic markets, education, recreation, and healthcare, all of which have direct and indirect effects on the health of a population (Siyan *et al.*, 2019). However, despite these numerous benefits, road traffic violations seriously harm any country's social, political, and economic development.

All over the world, there are traffic laws created for tackling traffic offenses as well as penalties at different levels of governance to discourage traffic violation (FRSC, 2007). However, the genesis of road traffic accidents or offenses in Nigeria is intricately linked with road transport development in the country (Isa & Siyan, 2016). Nigeria recorded her first traffic accident in Lagos in 1906 (Oluduro, 1999). In Nigeria, road safety laws violators violate mostly the use of seatbelts, exceeding expected speed limits, reckless driving, use of vehicles with unauthorized plate number, jettisoning the use of fire extinguisher, overtaking at ill points, overloading, phoning while on the wheel and failure to comply with traffic lights and signs (Emenike and Akpu, 2017). Over the last decades, the federal road safety corps (FRSC) saddled with the responsibility of safeguarding major roads in Nigeria, recorded over a hundred thousand road offences and accident cases, with over sixty thousand death cases and more than eighty thousand injury cases ranging from minor, severe and fatal injuries. Advances in the use of artificial intelligence (AI) and machine learning (ML), has been greatly deployed in solving most of the real-World problems, which has presented new opportunities for developing sustainable and effective system. As a result of increase in road violations and crashes, traditional statistics techniques have been widely used to address the global problem of RTO over the last century.

2 Materials and methods

2.1 Research design

This is basically the steps taken in research and how the intendent objectives are implemented to achieve the overall goal of a particular study. Burn and Gove (2017) defined research design as the blueprint for a study, guiding researchers to make sound and actionable decisions and avoid issues that could invalidate results. However, this research work employs a quantitative research design, utilizing machine learning techniques to predict road traffic violations based on the data gathered from Federal Road Safety Corps of Nigeria. Also, the primary models used in this study are Extreme Gradient Boost (XGBoost), Multi-Layer Perceptron (MLP), and Random Forest (RF) as well as leveraging data analytics process to gain insights from the data gathered. The adoption of ensemble learning model approach leverage on the strengths of these models in improving the predictive accuracy of the combine algorithm.

2.2 Requirement elicitation

To gain useful insight, this study was conducted to understand and identify the trends patterning road traffic violation between Onitsha and Enugu route. In this study, qualitative method was used to conduct basic statistical and exploratory analysis. Other requirements were gathered mostly through descriptive study in which secondary data was widely exploited. Consequently, keyword searches were conducted utilizing Internet search engines, scientific and open access journals and social media to gather information. However, the essence of this process is to identify the data needed as well as determining the necessary tools and techniques analysis in defining the specific machine learning models to be used.

2.2.1 Ethical consideration and data collection

The research prioritizes ethical data collection practices aligned with the Nigeria Data Protection Regulation (NDPR) of 2019. To ensure compliance, the researcher collaborated with the zonal

headquarter of Federal Road Safety Corps (FRSC) to anonymize all data and implementing robust security measures for drivers' privacy which aligns with the growing emphasis on data protection in Nigeria in strengthening public trust in research initiatives. Following ethical approval, the researcher obtained clearance from the sector commander to access datasets reflecting road traffic violations. Utilizing a stratified sampling technique, a representative sample records was extracted from both manual archives and electronic traffic crash database. This approach, increasingly adopted in Nigeria research ensures the dataset reflects the diversity of road users, whereby enhancing the study's generalizability and relevance to national traffic safety concerns.

2.2.2 Datasets description

This research used historical dataset gathered from Federal Road Safety Corps (FRSC) from the period January 2020 to December 2023 of road traffic event that happened between Onitsha and Enugu expressway. The dataset comprises of one thousand six hundred and forty-eight (1648) records, each representing individual observations, and fifteen (15) columns, each describing different features of the dataset. Below is table 1 depicting the columns header of the data extracted.

TABLE 1
Description of dataset features

Column Header	Column Description
S/N	Unique identification for each record
Year	The year incident occurs
Month	The exact month of the year incident occurs
Age	Age of drivers
Gender	Status of drivers if they are either male or female
Vehicle Type	Type of vehicle if it is either car, jeep, pick-up, bus or truck
Vehicle Category	The categories of vehicles if they are private or commercial
Seatbelt	If the driver is putting on his or her seatbelt or not
Road Condition	Condition section of the road if it is good, poor or bad
Aggressive Driving	If the driver was driving aggressively or not
Weather Condition	The weather condition as of the time of incident
Speed	If the driver of a particular vehicle is on high speed or not
Phone Use	Is the driver on a call or not
Time of Day	What time of day traffic incident happened
Violations	Traffic offence committed by drivers

3.3 Model design

The research work leverages a cutting-edge ensemble architecture known as "Stacked Heterogenous Ensemble" (SHE). This approach strategically feeds individual classification algorithms (Random Forest and Gradient Boosting Machines), with carefully selected features from the dataset. The outputs from each model are then concatenated and fed into a state-of-the-art stacking ensemble model (Multi-layer Perceptron). This stacked ensemble acts as a meta-learner, combining the strengths of the individual models and potentially leading to more robust

predictions of road traffic violations in the chosen region. Figure 1 below depict the Stack Heterogenous Ensemble architecture.

Level 0 (base learner)

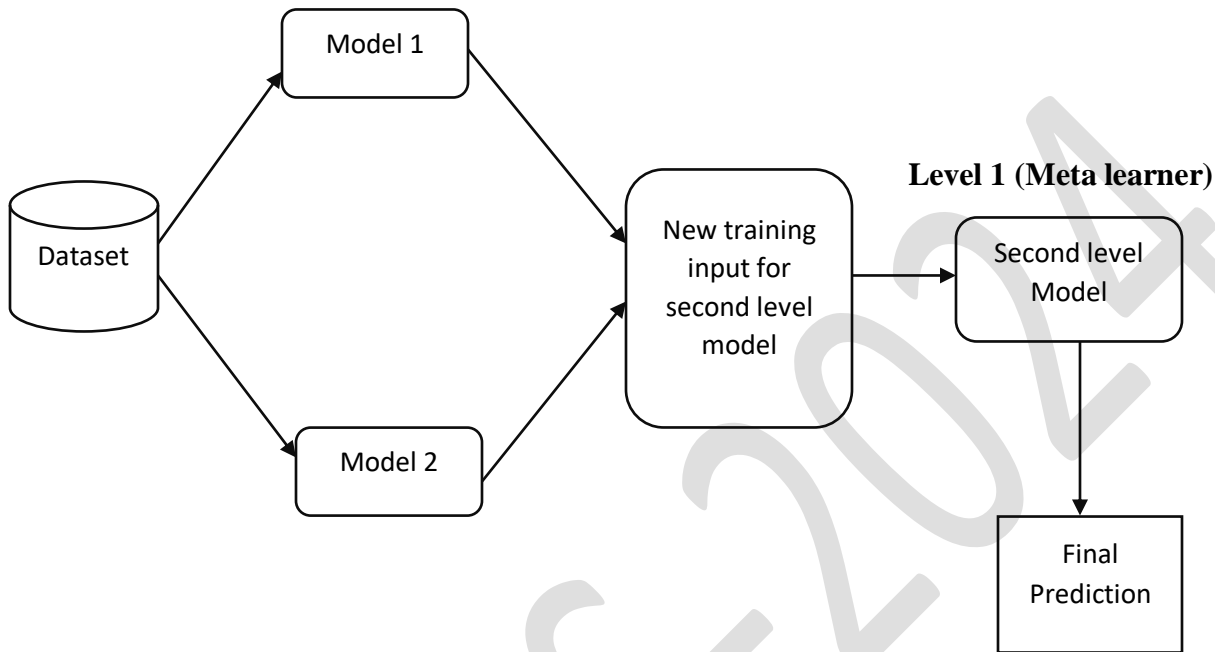


Fig. 1 Stacked Heterogenous Ensemble Model Architecture

2.4 SHE mathematical abstraction

Below is the stacked mathematical model adopted for this research, combining Random Forest, Extreme Gradient Boosting (XGBoost) as base learners, and a multi-layer perceptron (MLP) as the meta-learner. The given abstraction captures the key steps of the stacked ensemble model. The base learners (h_1 and h_2) provide initial predictions, which are then combined with the original features to create a new input for the meta-learner (MLP). The MLP, with its weights (W_i), integrates the information from both base learners and the original features to produce the final prediction ($f(x)$):

Sets and Notation:

X : Training data – a set of n data points, each with d features, represented as:

$$X = \{x_1, x_2, \dots, x_n\} \text{ where } x_i \in \mathbb{R}^d \text{ is the } i\text{-th data point.}$$

Y : Target labels – a vector of length n containing the true labels for each data point in X . represented as:

$$Y \in \mathbb{R}^n \dots\dots\dots (3.1)$$

M : Number of base learners

$h_m(x)$: Prediction function of the m -th base learner (RF or XGBoost).

Model Staging:

Base learner Training Phase:

Training data (X, y) i.e train RF and XGBoost

Meta-learner training Phase:

for each data point x_i in X:

obtain prediction from both base learners:

$z_i^1 = h_1(x_i)$ and $z_i^2 = h_2(x_i)$

Combine predictions with the original features (x_i) to form a new feature vector:

$v_i = [x_i, z_i^1, z_i^2]$

Prediction Phase (given new data point):

$Z^1 = h_1(x)$ and $Z^2 = h_2(x)$

Combine predictions with the original features (x) to form a new feature vector:

$v = [x, Z^1, Z^2]$

output:

$f(x) = f_{\text{meta}}(v; w_1, w_2, \dots, W_L) \dots \dots \dots (3.2)$

where:

w_i denotes the weight matrix of the i -th layer in the MLP

and:

$v_i = [x_i, h_1(x_i), h_2(x_i)] \dots \dots \dots (3.3)$

3 Results and discussion

3.1 Data presentation and summarization

The dataset used for this research work was collected from FRSC as earlier mentioned in chapter 3 and was pre-processed using Microsoft Excel spreadsheet. It comprises 1649 rows (records), each representing a road traffic incident that occur between Anambra-Enugu highway. Key attributes that make up our dataset includes vehicle type, category, driver gender, age, seatbelt usage, road conditions, aggressive driving, weather conditions, speeding, phone use, time of day, and the corresponding violation. The data spans from 2020 to 2023 (four years), capturing male-dominated population of drivers primarily plying and operating commercial vehicles. Finally, Table 2 below depicts summary of the basic statistic conducted on our dataset.

Table 2 Summary of basic statistic

	Count	Mean	Standard Deviation
Vehicle Type	1648	2.18	1.36
Age	1648	38.65	8.08
Seatbelt	1648	0.45	0.50
Road Condition	1648	1.83	0.95
Aggressive Drive	1648	0.25	0.43
Weather Cond	1648	1.07	0.25
Speeding	1648	0.12	0.33
Phone Use	1648	0.04	0.21
Time of Day	1648	1.59	0.49

3.2 Data and features visualization

To draw useful insights from the dataset collected from FRSC, the researcher visualizes the relationship between attributes such as “Age”, “Vehicle Type”, “Gender”, “Category”, and “Violations” using bar chart, Pie chart, Correlation matrix and stacked bar chart. Below are the given visualization figures from the data analysis conducted.

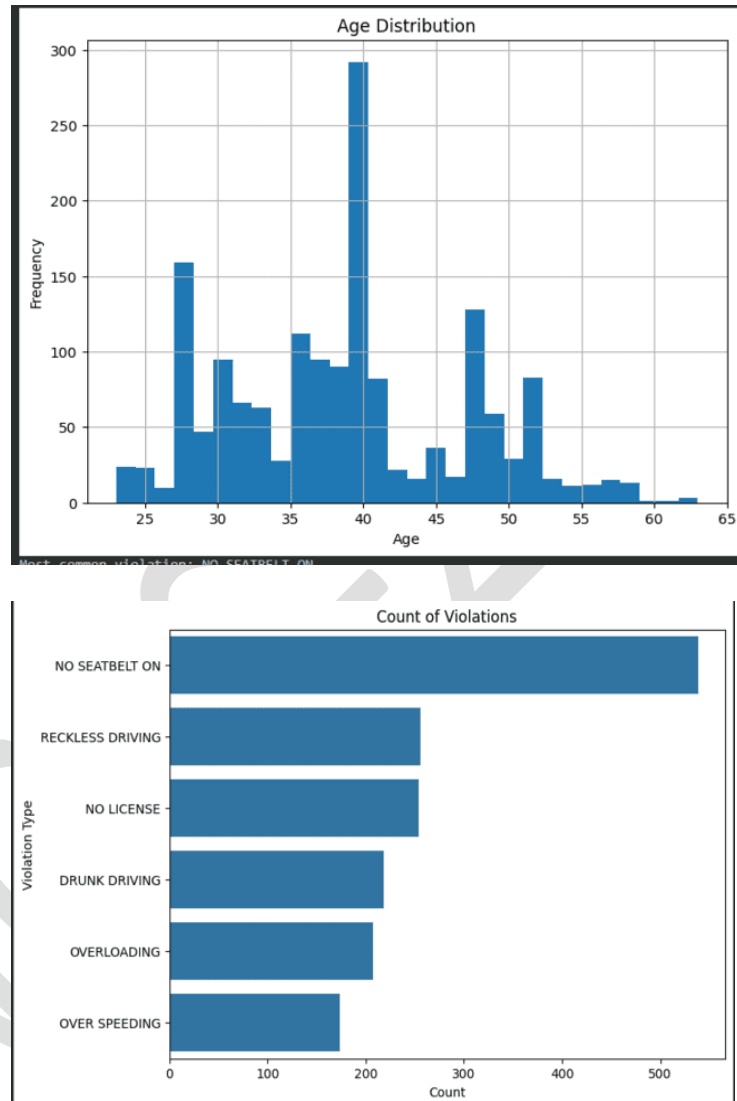


Fig. 2 Frequency and bar chart distribution of age and count of violations

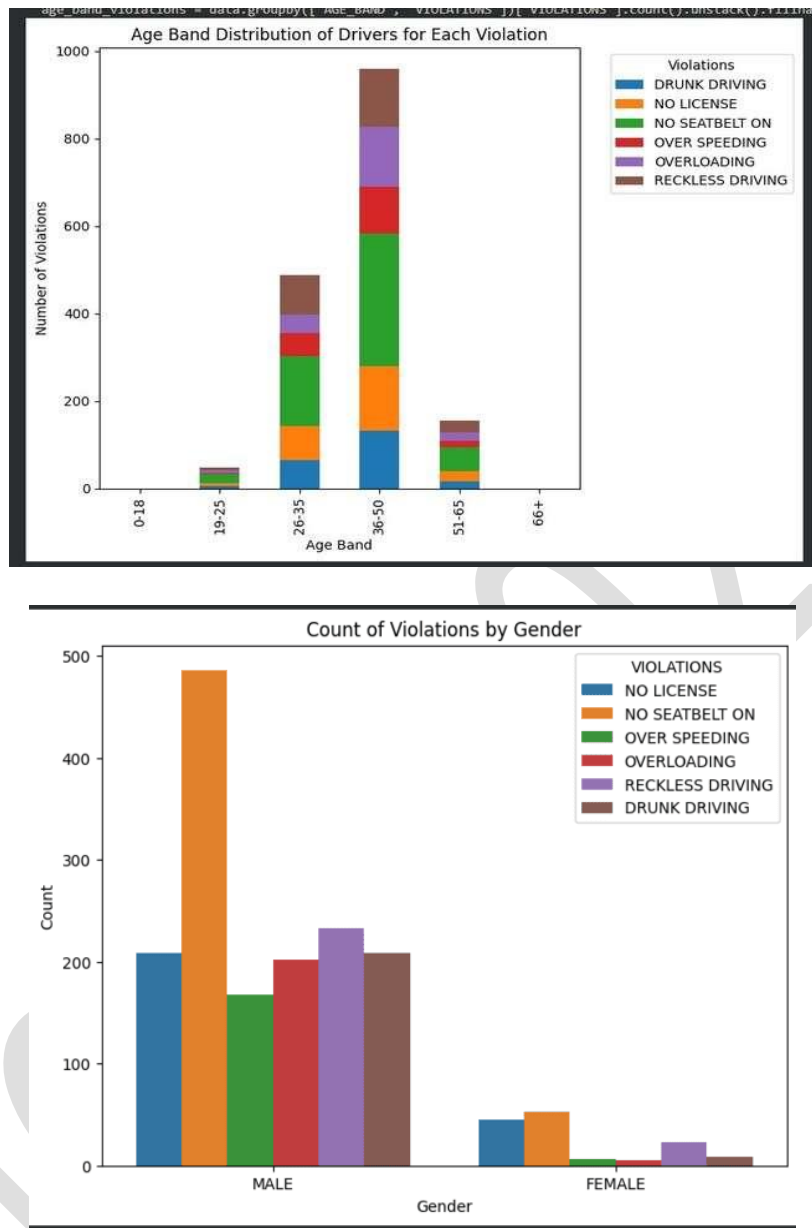


Fig. 3 Bar chart distribution depicting violations by gender and age

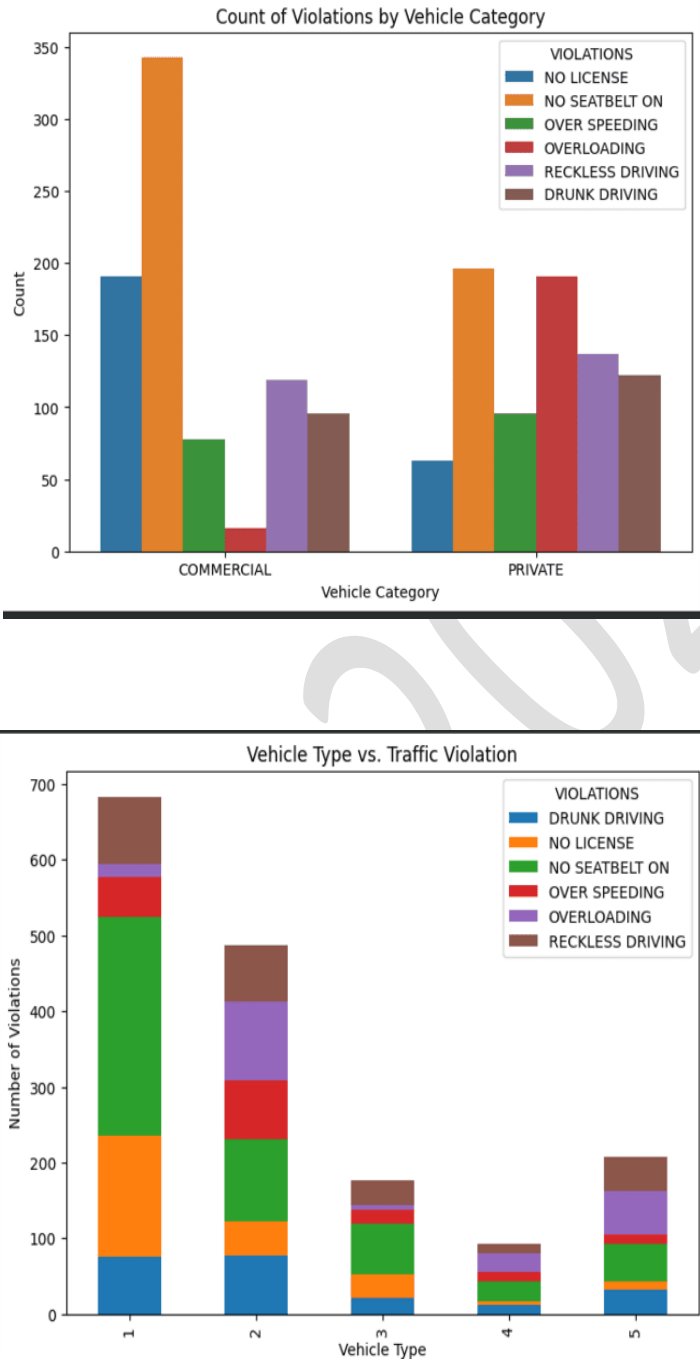


Fig. 4 Bar chart distribution depicting violations by vehicle category and type

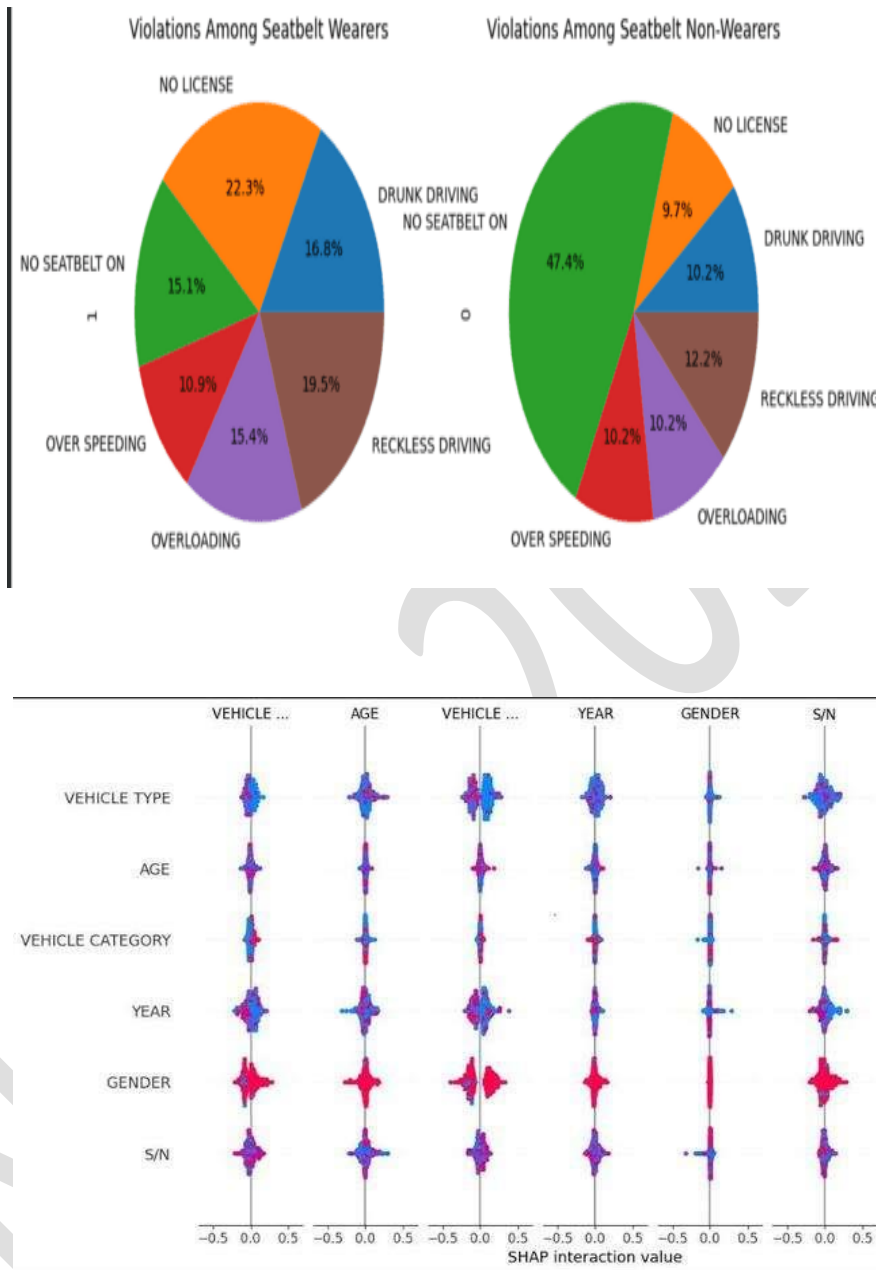


Fig. 5 Pia chart and bubble distribution depicting violations and attribute features

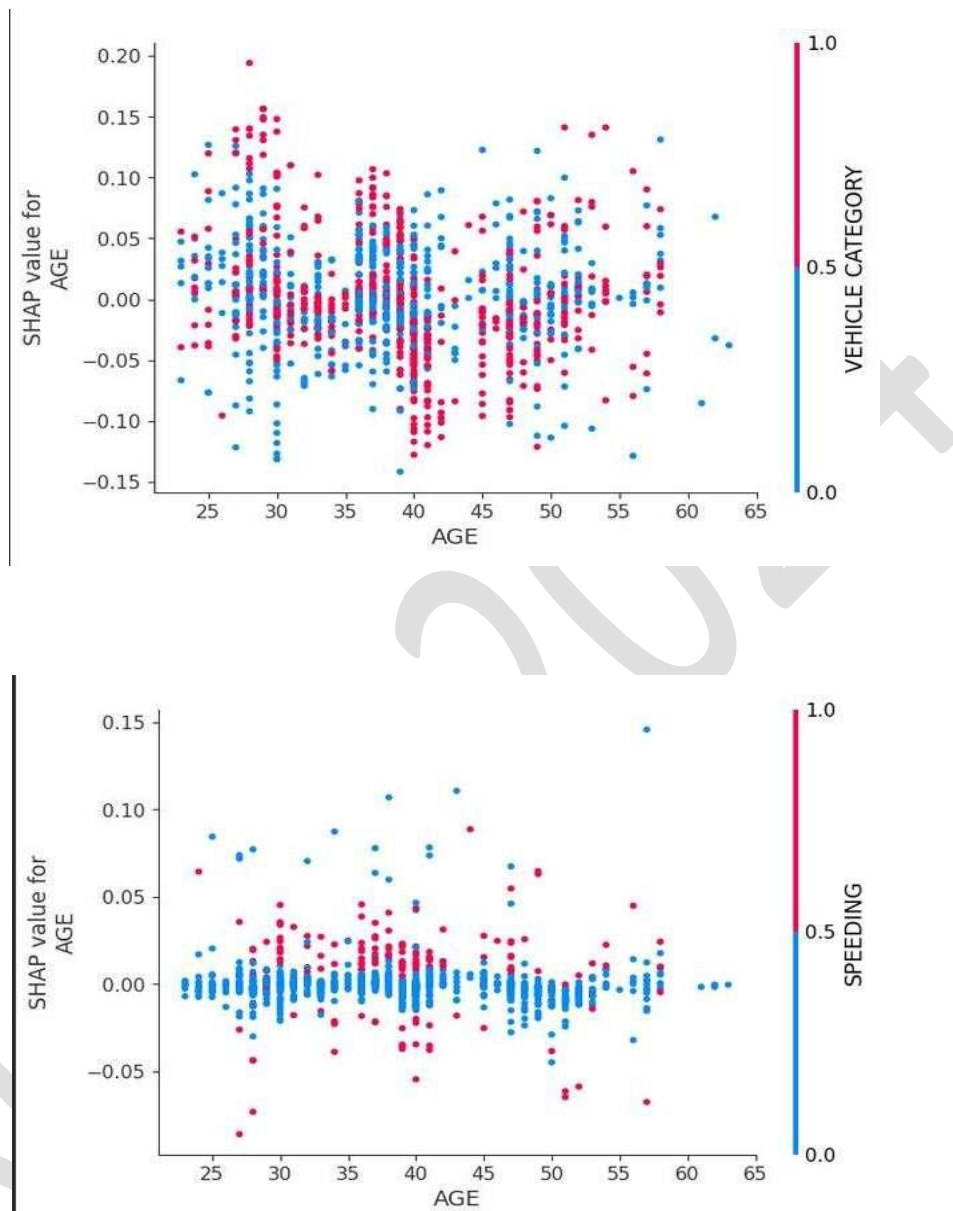


Fig. 6 Scatter-plot chart distribution depicting violations by age, category and speeding

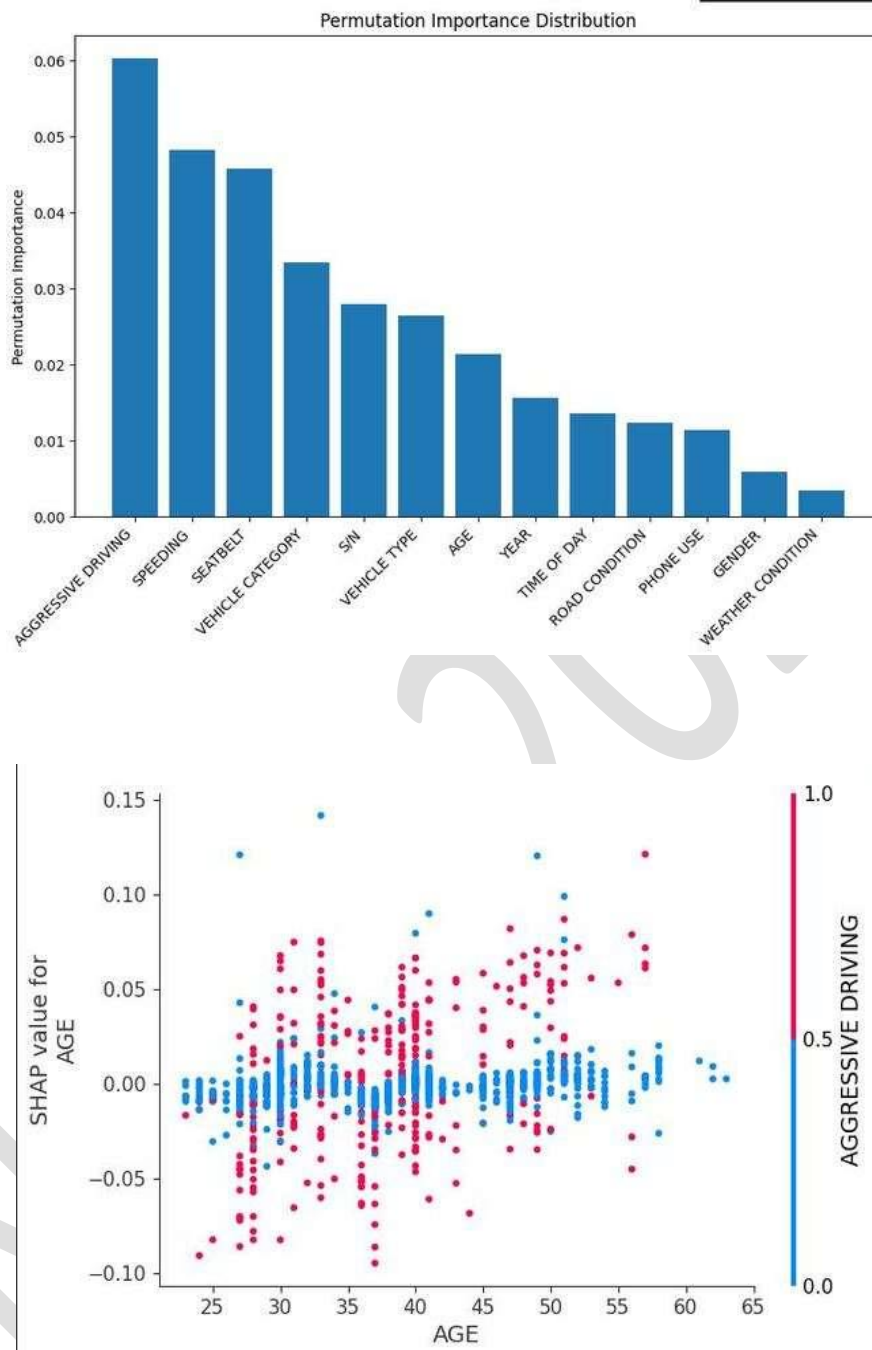


Fig. 7 Bar chart and scatter plot distribution depicting features importance

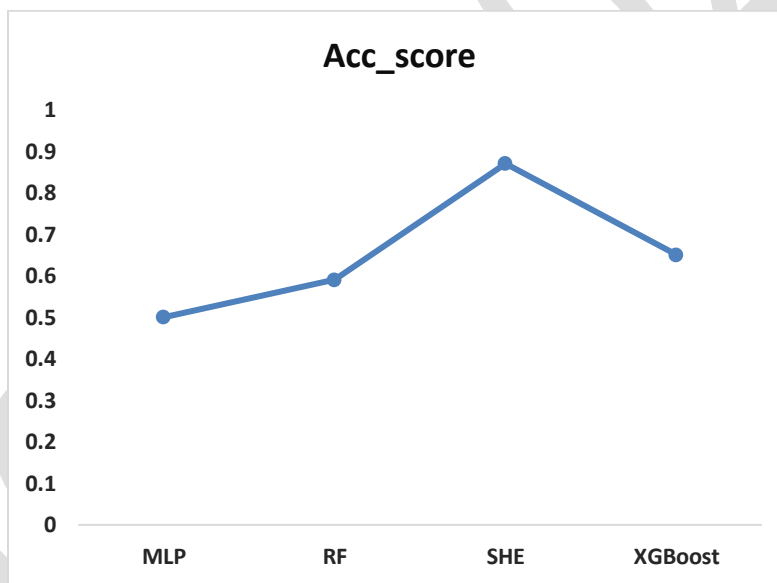
4 Discussion

After applying machine learning algorithms on the FRSC traffic datasets. The researcher utilizes accuracy score, precision, sensitivity and area under curve (AUC) as performance metrics to compare and evaluate the models in determining the best performing algorithms for predicting traffic violations. Also, model interpretability techniques were employed to dissect the behavior on how different features influences the prediction outcomes of Stacked Heterogenous Ensemble

model. Hence, accuracy score in this context defines the number of correct predictions made as a ratio of all predictions made. Consequently, precision defines the number of correct instances returned, and while sensitivity (Recall) defines the number of positive outcomes returned by SHE. Table 3 and figure 23 and 24 below shows the performance metrics value for FRSC datasets. From the results of training and testing set, all classifiers (models) have varying metrics value in terms of accuracy score, precision, and sensitivity. However, SHE method has higher accuracy testing set score of 87% than the other classifiers. Also, figure 8 depicts the receiver operating characteristic (ROC) curves for the multi-class instances.

Table 3 Model comparison

Models	Accuracy Score	Precision	Sensitivity
SHE	0.87	0.85	0.87
XGBoost	0.65	0.63	0.65
MLP	0.50	0.48	0.50
Random Forest	0.59	0.56	0.59



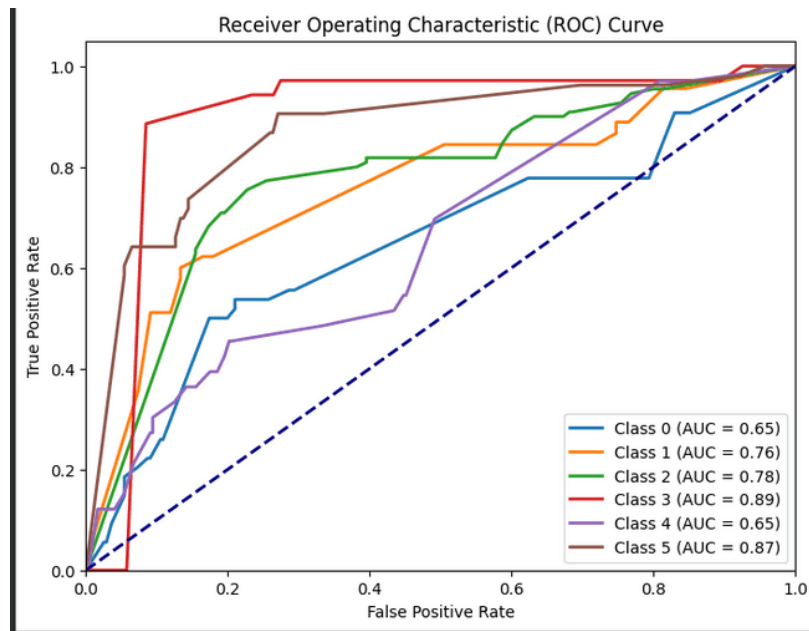


Fig. 8 Line and ROC chart distribution depicting metrics score

REFERENCES

- Abdulalah, H. A, Osama, J. R, (2022). Using data analytics and machine learning for road traffic accident prediction and casual factor analysis. *Journal of King Abdulaziz University, (JKAU: Comp. IT. Sci)*. 11(2): 53-66.
- Abdulmalik, M. M, (2023). Analysis of road traffic violations on passengers and freights transport in Minna metropolis. *Lapai International Journal of Management and Social Science*. 15(1): 2006-6473.
- Azad, A, (2017). Road crash prediction models: Different statistical modeling approaches. *Journal of Transportation Technologies (JTT)*. 7: 190-205.
- Bara, W. A., Ahmad, H. A., Rana, I. & Mohammad, M. (2022). Using machine learning models to forecast severity of traffic crashes by R studio and ArcGis”, *Frontiers in Built Environment*.
- Dabiri, S., & Heaslip, K. (2019). Developing a Twitter-based traffic event detection model using deep learning architectures. *Expert Systems with Applications (ESA)*. 118: 425-439. <https://doi.org/10.1016/j.eswa.2018.10.017>.
- Ebrahim, S. Milon-Islam, M.D. & Quazi, S.H. (2021). A review on neural network techniques for the prediction of road traffic accident severity, *Asian Transport Studies*. www.sciencedirect.com/journal/asian-transport-studies.
- Emenike, G. C. & Akpu, D. N., (2017). Assessment of road traffic violations in Port Harcourt metropolis, Nigeria. *International Journal of New Technology and Research (IJNR)*. 3(4): 55-62.
- Eric, L. M, Ibai, L., & Javier D. S, (2021). Deep learning model for road traffic forecasting: Does it make a difference. *IEEE Transaction on Intelligent Transportation Systems*.

- Evwiekpaefe, A. E, & Umar, S. N, (2021). Predicting road traffic crash severity in Kudana metropolis using selected machine learning techniques. *Nigerian Journal of Technology (NIJOTECH)*. 40(5): 888-900.
- Federal Road Safety Corps (2012). Nigerian Road Safety Strategy (NRSS) 2012-2016.
- Francisca, N. O, Toochukwu, C. O, Sanjay, M. & Fernandez-Sanz, L. (2014). An artificial neural network model for road accident prediction: A case study of a developing country”, *Acta Polytechnica Hungarica*. 11(5).
- Iyabode, O. F, (2016). Statistical analysis of pattern on monthly reported road accidents in Nigeria. *Scientific Journal of Applied Mathematics and Statistic*. 4(4): 119.
- Joshua, O. Oluwashina, O. Lawrence, O. Adeolu, O. (2020). Adoption of deep learning for computer aided driving and road safety in Nigeria. *International Journal of Engineering and Artificial Intelligence (IJEAI)*. 1(3): 25-31.
- Oluwafunbi, A. M, Alexander, H., Killian, A., Amalia, V., Ritu, S., Michael, S., Molloy, MB. & Gregory, R. C. (2021). Trend analysis on road traffic collision occurrence in Nigeria”, www.cambridge.org/dmp.
- Omolola, F. A., Sanjay, M. & Akshat, A. (2023). Improving real time intelligent transportation systems in predicting road accident.
- Radzi, N. H. M, Gwari, L. S. B, Mustaffa, N. H. & Sallehuddin, R. (2019). Support vector machine with principal component analysis for road traffic crashes severity classification. *Joint Conference on Green Technology and Applied Computing*. 551(2019).
- Salahadin, S. Y, & Pooja. (2020). Road accident prediction and model interpretation using hybrid K-means and random forest algorithm approach”, *Springer Nature of Applied Science*.
- Siyan, P. Boboye, O., & Olabode, A. (2019). Road accident analysis and prevention in Nigeria: Experimental and numerical approaches. *International Journal of Advanced Scientific Research and Development (IJASRD)*. 6(1): 01-12.
- Yunxuan, L, Meng, L, Jinghui, Y, Jian, Lu, & Mohamed, A. (2022). Analysis and prediction of intersection traffic violations using automated enforcement system data. *Department of Civil Engineering, Tsinghua University, Beijing 100084, PR. China*.
- Yannis, G. Theofilatos, A. Pispiringos, G. (2017). Investigation of road accident severity per vehicle type. *World Conference on Transport Research- WCTR 2016 Shanghai*. 25(2017): 2076-2083.

Development of a remote emergency call mobiletracking system

¹Uke, Ernest M. and ²F. U. Ogban

^{1,2}Department of Computer Science, University of Calabar, Calabar, Cross River State.

Abstract

Emergency situations are situations that require quick intervention so as to save lives, assets or properties from loss or destruction. Emergency situations may also be referred to as emergencies. Ideally, in today's technology-driven society, emergency service should be accessible remotely. This study involved the development of a remote emergency mobile call tracking system. Object Oriented Analysis and Design (OOAD) was used for analysis and design of the system (software), while Entity Relationship Modelling (ERM) was used for the database component of the system. The system consists of both front-end and back-end. The front-end was built with Cordova framework, while the back-end was built with PHP. Cordova framework is used to develop User Interface (UI) for mobile apps. Cordova framework for mobile development integrates with Android SDK for this purpose. Regarding the back-end of the system, PHP is used as the technology for the implementation of this study. The App has a User Interface (UI). The UI is built as screens (that is, activities) which are accessible via the App when installed on the User's Android phone. From the findings of the study, the App is able to identify automatically every emergency reporter, and pick the exact location of the report. This App has the capability of making emergency services provision more effective, efficient and affordable. Based on the findings of this study and the system developed, it is recommended among others that the government of Nigeria being a major stakeholder in the emergency services sector, (by virtue of its ownership of emergency services such as the police, fire services, hospitals, and so on) should support the deployment and use of the technology (software) in this sector.

Keywords: Remote, Emergency, Tracking, Integrate.

1.0. INTRODUCTION

Emergency situations are situations that require quick intervention so as to save lives, assets or properties from loss or destruction. Emergency situations may also be referred to as emergencies. Emergencies are common occurrence in most modern societies. Wherever humans occur, emergency situations always occur. It is as though emergencies are a part of human societies and life. According to Qureshi (2014), human societies are vulnerable to multiple natural and human-induced disasters and they have been recurrently hit by natural and man-made disasters which has led to loss of lives, livelihoods and properties. This view is corroborated by UNDP (2001). In modern societies, there are emergency response services whose duties involve receiving emergency reports from the citizenry, document them, and make the appropriate response.

In this study, an Emergency Response system is developed. It is a mobile App that is installable on mobile (android) devices (Phones, Tabs). On installation of the App in the mobile phone, the owner of the phone can register to become a registered user of the system, and can therefore send report from any remote location to the emergency response system. On accessing the system, the system is able to decipher and detect the location the report is coming from.

Emergency reporting system as a sample of technology, is capable of enhancing and adding value to humans in the modern era. Figuracion et. al., (2016) explained that an emergency report system can be used to report on emergencies ranging from fires to typhoon-related incidents, vehicle accidents with casualties, health-related concerns (e.g. unidentified person suffering from heart attack), community related incidents and concerns, and any other occurrences which requires immediate and concrete response from the concerned agencies.

2. METHODOLOGY

In the existing system of emergency response in Nigeria and other Sub-Saharan nations, emergency responders could be said to include the Police, Fire Service, Ambulance Services, Federal Road Safety Corp (FRSC), etc. These emergency responders only receive reports via phone calls, and have no means of automatically knowing the identity of the caller and the location of the call. This puts lots of strain on individuals calling to report emergencies and needing immediate response.

In the proposed system of this study, the manual documentation is eliminated. Documentation regarding the description of the emergency, the location, the identity of the caller, etc., is immediate and accurate because it is automated. This turns out to be a significant improvement upon the existing system.

3. Design Considerations

Object Oriented Analysis and Design (OOAD) is used for analysis and design of the system (software) while Entity Relationship Modelling (ERM) is used for the database component of the system. In OOAD, everything is based on the concept of classes and objects. A class is defined as a template for an object. An object is referred to as the instance of a class. A Class contains the data fields (attributes) and methods which the object will use. OOAD is the current convention (and highly recommended) in the software development industry globally.

Object-Oriented Analysis and Design (OOAD) is the procedure of identifying software engineering requirements and developing software specifications in terms of a software system's object model, which comprises of interacting objects. The main difference between object-oriented analysis and other forms of analysis is that in object-oriented approach, requirements are organized around Classes (and their objects), which integrate both data (attributes) and functions (methods).

Grady Booch (1994) has defined OOAD as, "Object-oriented analysis is a method of analysis that examines requirements from the perspective of the classes and objects found in the vocabulary of the problem domain". The primary tasks in object-oriented analysis and design (OOAD) are:

- Identifying Classes.
- Implementation of methods, i.e., internal data structures and algorithms.
- Organizing the classes by creating object model diagram.
- Defining the internals of the classes, or object attributes.
- Defining the behaviours of the classes, i.e., object actions.
- Describing how the classes interact.

The common models used in OOAD are use cases and object models.

3.1 Agile Model

Agile—denoting “the quality of being agile; readiness for motion; nimbleness, activity, dexterity in motion”—software development methods are attempting to offer once again an answer to the eager business community asking for lighter weight along with faster and nimbler software development processes.

Agile development methods are understood to be reaction to plan-based or traditional methods, which emphasize “a rationalized, engineering-based approach” incorporating extensive planning, codified processes, and rigorous reuse (Dybå, 2000; Boehm 2002). By contrast, agile methods address the challenge of an unpredictable world; by recognizing the value competent people and their relationships bring to software development (Nerur and Balijepally, 2007). In that light, Ericksson et al. (2005) underline the importance of lightweight processes in agile development, defining agility as to “strip away as much of the heaviness, commonly associated with the traditional software-development methodologies, as possible to promote quick response to changing environments, changes in user requirements, accelerated project dead-lines and the like”.

According to Abrahamsson, Salo, Ronkainen and Warsta (2002), the central values that the agile community adheres to are: First, the agile movement emphasizes the relationship and communality of software developers and the human role reflected in the contracts, as opposed to institutionalized processes and development tools. In the existing agile practices, this manifests itself in close team relationships, close working environment arrangements, and other procedures boosting team spirit.

Second, the vital objective of the software team is to continuously turn out tested working software. New releases are produced at frequent intervals, in some approaches even hourly or daily, but more usually bi-monthly or monthly. The developers are urged to keep the code simple, straightforward, and technically as advanced as possible, thus lessening the documentation burden to an appropriate level. Third, the relationship and cooperation between the developers and the clients is given the preference over strict contracts, although the importance of well-drafted contracts does grow at the same pace as the size of the software project. The negotiation process itself should be seen as a means of achieving and maintaining a viable relationship. From a business point of view, agile development is focused on delivering business value immediately as the project starts, thus reducing the risks of non-fulfillment regarding the contract.

Fourth, the development group, comprising both software developers and customers representatives, should be well-informed, competent and authorized to consider possible adjustment needs emerging during the development process life-cycle. This means that the participants are prepared to make changes and that also the existing contracts are formed with tools that support and allow these enhancements to be made.

3.2 Implementation technology

The proposed system consists of both front-end and back-end. The front-end is built with Cordova framework, while the back-end is built with PHP. Cordova framework is used to develop User Interface (UI) for mobile apps. Cordova framework for mobile development integrates with Android SDK for this purpose. With Cordova, UI can be coded in HTML, CSS, JavaScript/jQuery. Cordova framework includes a Browser platform for testing while developing. On successful testing with the Browser platform, the developer can build installer (apk) file with the build tools of Cordova framework. The apk can be used (as installer) to install the App in any Android mobile phone.

Regarding the back-end of the proposed system, PHP is used as the technology for the implementation proposed in this study. PHP is a full-fledged object-oriented technology implementing the concepts of inheritance, polymorphism, encapsulation, etc. In other words, the system is fully class - and object-based. PHP is used for the system because globally it is accepted in the software development industry as web technology. PHP web technology is widely deployable in several web servers of renown. PHP is generally known to be a very secure web technology.

3.3 Information Flow Diagrams

A flowchart is a type of diagram that represents an algorithm, workflow or process, showing the steps as boxes of various kinds, and their order by connecting them with arrows. This diagrammatic representation illustrates a solution model to a given problem. In the proposed system of this study, several process (actions) are done by various actors (roles) on the system. These processes (actions) are hereby represented by the following flow charts:

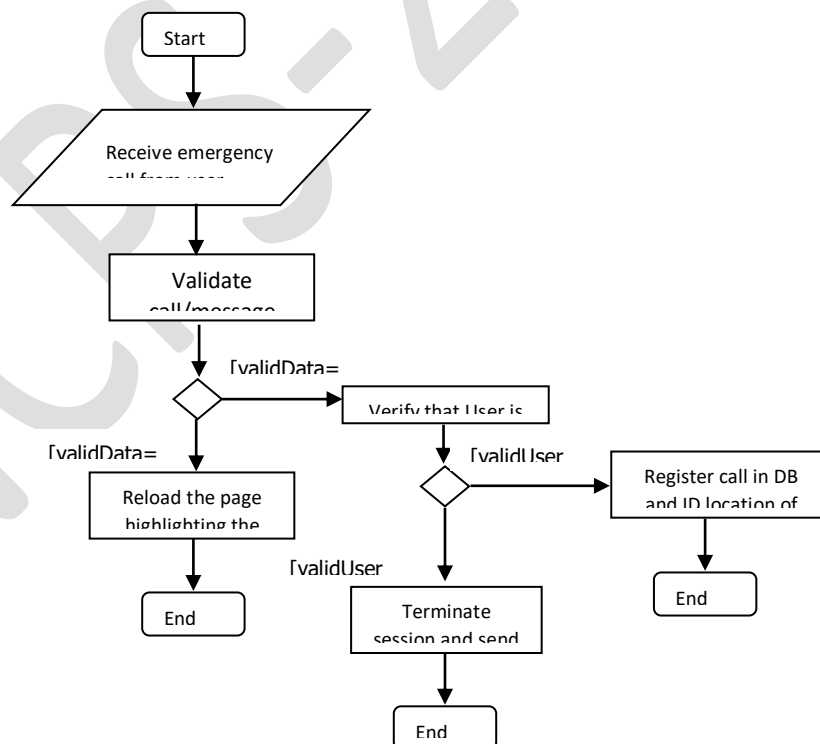


Figure 3.2. Flow chart for receive Emergency Call

Process Model Diagram

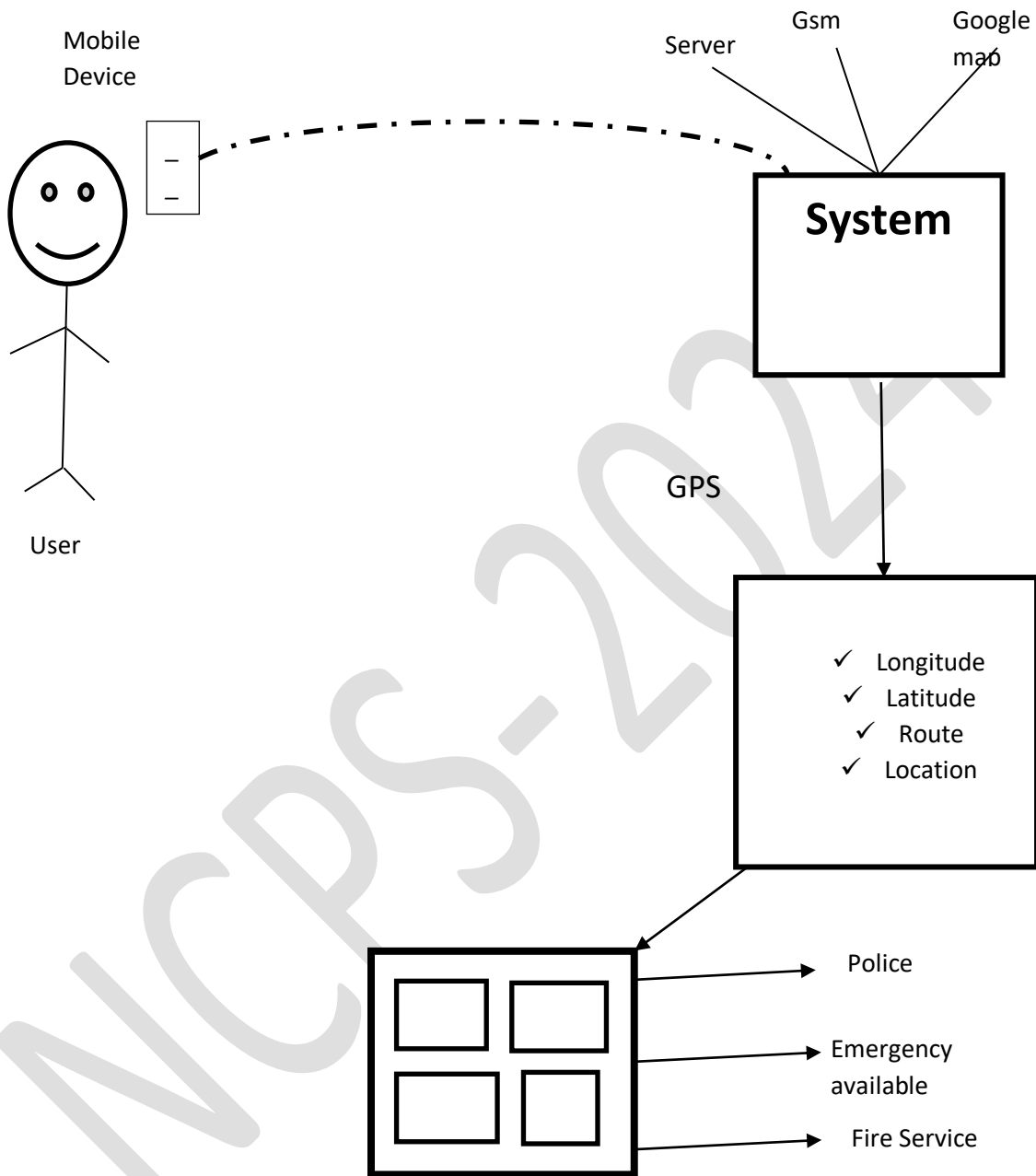


Figure 3.3. Process Model Diagram

3.4Voronoi Diagram Algorithm and Triangulations

A *triangulation* of an emergency report made from points **A** in a plane is a partition of the convex hull (the smallest convex set that contains **A**) into triangles. So, for any report **A** of discrete points in a plane, a *Delaunay triangulation* **DT(A)** is a triangulation such that no point in **A** is inside the circumcircle of any triangle in **DT(A)**. Interestingly enough, it turns out that **DT(A)** corresponds to the *dual graph* of the Voronoi diagram for **A**.

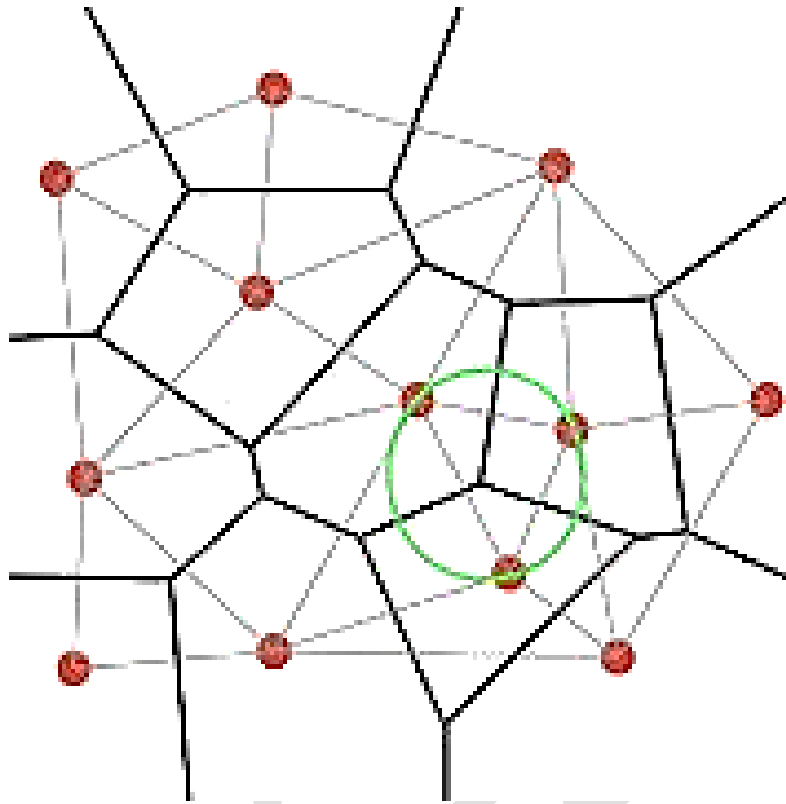


Figure 3.4 Triangulation for emergency report A

Based on the hierarchical network model and the Delaunay triangulation, and considering that the cluster heads and the lower nodes are moving, a topology control based on Voronoi diagram (TC-VD) algorithm is adopted, which creates a topology with high connectivity and strong robustness for emergency reports.

Shown below is a Delaunay triangulation with all the circumcircles. Connecting the circumcenters produces the Voronoi diagram. In particular, if two triangles share an edge in the Delaunay triangulation, their circumcenters are connected with an edge in the Voronoi diagram.

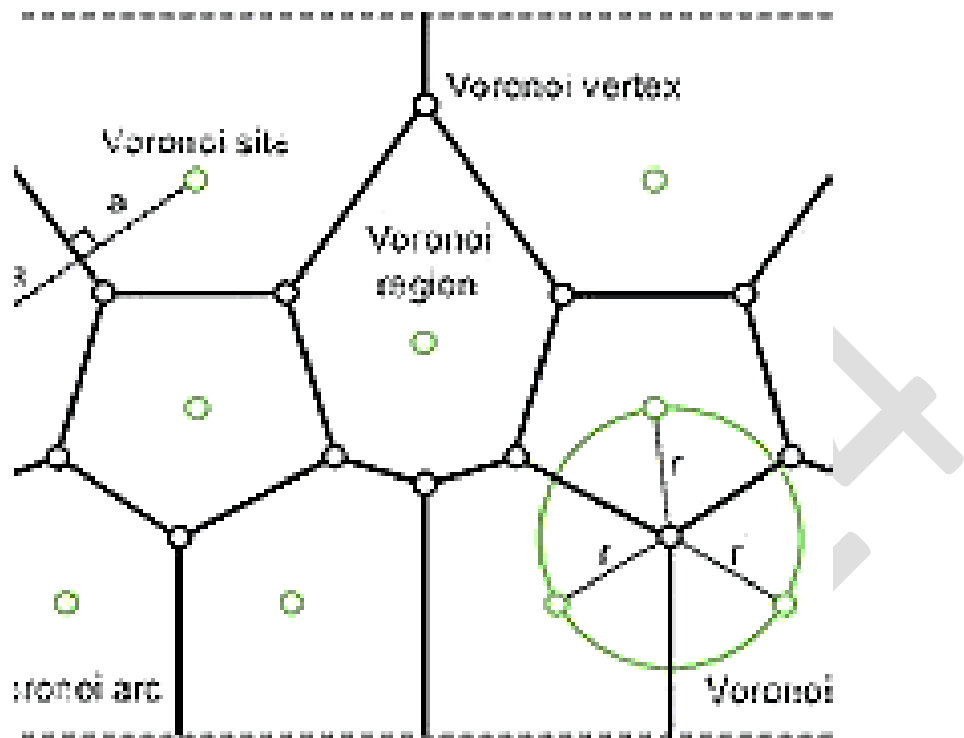


Figure 3.5 Voronoi diagram for emergency reports

4. SYSTEM IMPLEMENTATION

This is the phase where the programmers transform details of the design specification into real software (a real Application). This chapter entails the implementation of the proposed system (Emergency Report System). The system is developed implementing the findings of previous phases such as Requirements gathering, and Analysis and design. The implementation was done as a Mobile (Android) Application. The App was implemented using Cordova framework. Cordova framework is a framework for mobile development. Cordova integrates with Android SDK at the background to develop Apps for the Android platform. Android is Java technology (programming language).

The App involved a User Interface (UI). The UI is built as screens (i.e. Activities) which are accessible via the App when installed on the User's Android phone. System development is explained in detail. In this section, the researcher describes the basic tools employed to carry out the development. Also, testing and debugging of the application is done. This is done to ensure that the application meets the requirement arrived at during the Analysis & Design process. This chapter also describes the client system requirements. Finally, maintenance issues are discussed. The researcher enumerates how the software would be sustained, managed or scaled (up or down) if necessary.

4.1. Features and Choice of Implementation Language

Implementation refers to translating algorithm gotten from analysis and design phase of the system development life cycle, into codes in a particular programming language (technology) adopted for the system. As already mentioned, Cordova framework integrated with Android SDK is used in this study. Android SDK is Java technology (programming language). Therefore, for the purpose of this study, Java technology is used. In other words, Packages, Classes and Functions—in an ideal Object-Oriented Analysis and Design (OOAD) paradigm—are

implemented with the Java language, codes are written in the language to implement the logic (algorithm, flow chart) designed, the codes are interpreted and executed by the Java Interpreter. There are several advantages to developers and users in implementing algorithms (i.e. developing software) with the Java language. The reasons for the use of Java language for the implementation of the system are listed below:

- Java is strict-typed
- Java is object-oriented
- Java is compiled
- Security
- Reliability
- Java is superb for Network-based project such as this
- Java is Robust
- Java is Dynamic
- Java is widely circulated (deployed) in the software industry globally.

4.2. System Testing Strategies

Coding of the proposed software is done with Java technology with the help of industry-standard Integrated Development Environment (IDE), Apache Netbeans 12. Coding in Java for Android platform with the help of an IDE such as Apache Netbeans 12 can be done when the Java SDK and Android SDK has been installed with all necessary utilities activated in the computer of the software (system) developer. With this, the developer can take advantage of the very rich Java native and Android SDK libraries for the purpose of the project. The developer may also make use of external (third-party) Java libraries for the purpose of the project.

In addition to making coding in Java language possible, the Java and Android SDKs also make it possible for the developer to debug the codes written for the purpose of the project with the use of debugger for Java applications. Simply put, Debugging refers to investigating unintended behaviours within a software. For the purpose of this study, debugging was done during the implementation and testing phase of the software development life cycle. By using a Java-friendly IDE such as Apache Netbeans 12, the developer was able to investigate the values each variable/object within the code is holding, and make changes or corrections were necessary.

4.3. Target Computer System Requirements

This section examines the issues of ensuring software compatibility. For this system developed, there are minimum requirements to be satisfied if the application (software) is to run on a particular client computer. The basic requirements for a computer that can run the proposed software can be grouped into hardware and software and requirements. Each category is described below:

4.3.1. Hardware Requirements

Hardware requirement for the purpose of this system includes:

- 35GB Internal Memory or higher,
- Dual Core Processor or higher,
- 512MB Random Access Memory (RAM) or higher and,
- Recommended memory 2GB,
- Minimum Disk Space 500MB,
- Recommended Disk Space 1GB

4.3.2. Software Requirements

Software requirement for the purpose of this system includes Windows XP Operating System or later, Browser (Mozilla/Chrome/Opera). With these, any computer or mobile device can install and run (use without hitches) the Emergency Report System developed in this study.

5. Results

In the Implementation phase of the software development life cycle, programmers translated details in the analysis and design to real software. The basic tools employed in system development are the development framework and associated development tools. The software development framework used in this is Cordova. With Cordova, mobile App User Interface (UI) can be written in HTML, CSS, JavaScript/jQuery. Cordova framework is configured to use Android SDK at the background thus converting HTML pages into real Android activities (screens) in the mobile App. Android is Java technology (programming language).

This served as a transition from the design specification to the actual system implementations (the App or Software). System testing involved testing and debugging of the program and its operations so as to meet the requirements of the overall system objectives. During this phase, it may be necessary to make changes to the design specifications. If so, the changes are made. It is best that changes are made before the system actually goes in to operation (i.e. deployed for use by users/clients). Changes at this stage are less expensive and disruptive than if they have to be later on. As a general rule, the earlier the change, the less expensive it will be.

Map Screen: GoogleMap showing location of report made/received.

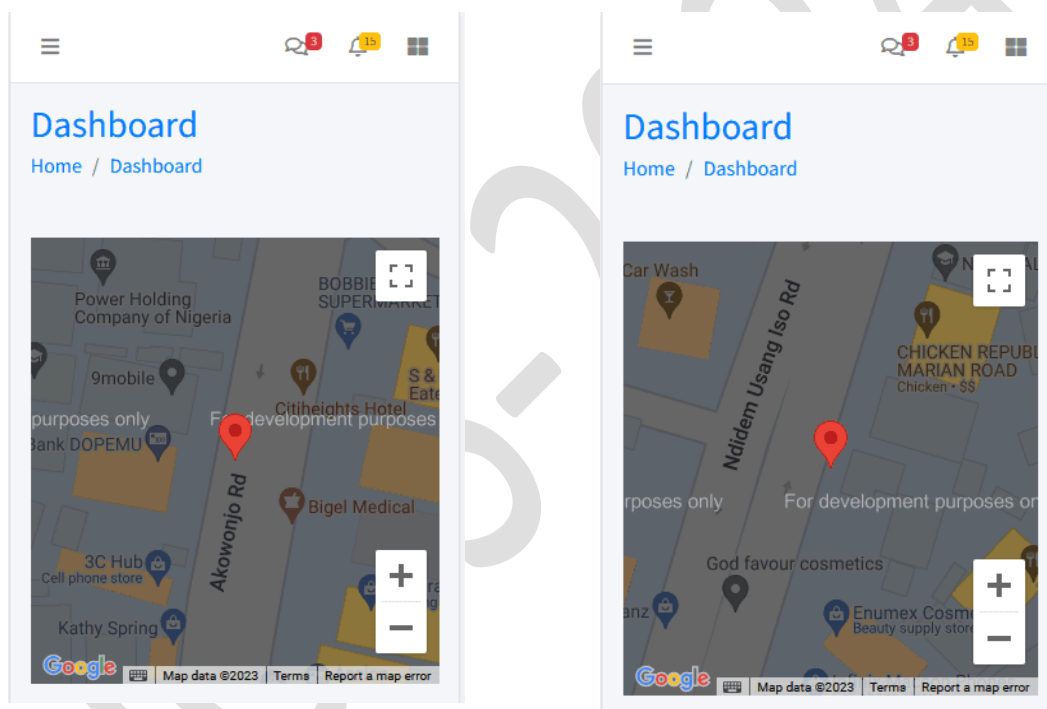


Figure 3.7 Google Map of Emergency report made from Akowonjo Road area of Lagos State.

Figure 3.8 Google Map of Emergency report made from Chicken Republic area of Marian Road, Calabar

5.2. Conclusion

The goal of this research is focused more on the design and development of Emergency Report System for use by every member of society and emergency service providers. The researcher shows that an Emergency Report System is Win-Win situation for every member of society. This includes individuals, households, communities, emergency service providers, government agencies, etc. The experience of the emergency reporter and service provider is improved as the later (and their organisations) are able to interact with the Application and get real-time accurate and immediate data regarding identity of emergency reporters and exact location of report.

Easy access to records, data and information removes frustration among emergency service providers. It is therefore not an exaggeration to say that software can be used to improve service delivery in the emergency response service sector.

REFERENCES

- Abrahamsson, P., Salo, O., Ronkainen, J. & Warsta, J. (2002) Agile software development methods: Review and analysis, VTT publication 478, Espoo, Finland, 107p.C
- Adusei, K., Kyamakya, K., & Erbas, F (2004). Location Based Services: Advances and challenges. May, CCECE2004-CCGEI2004, pp. 1-7. Niagara Falls: IEEE. [online] Available at: <http://IEEEExplore.ieee.org>
- Ani, K. & Onyebukwa, C. F. (2016). Nigerian Security Challenges And Recommendations For Sustainable National Development. Urbanization, Security and Development Issues in Nigeria.
- Barbeau, S. (2007). Wireless Emergency-Reporting System. Retrieved from www.research.usf.edu/absolute-news/templates/?a=2513&z=1
- Behzad, M., Sana, A., Khan, M. A., Walayat, Z., Qasim, U., Khan, Z. A. & Javaid, N. (2014). Design and Development of a Low Cost Ubiquitous Tracking System. *Procedia Computer Science* 34: 220 – 227
- Boehm, B. (2002) Get ready for agile methods, with care, *IEEE Computer*, vol. 35.
- Booch, G. (1994). Object-oriented Analysis and Design. California: Addison Wesley Longman, Inc
- Cellular News (2014). Available on the Internet: <http://www.cellular-news.com/story/29824.php>
- Conboy, Kieran (2009) Agility from First Principles: Reconstructing the Concept of Agility in Information Systems Development, *Information Systems Research*, 20, 329-354.
- Conti, J.P (2008). The Androids are coming. *IEEE Engineering and Technology*, Vol.3 (9): 72-7.
- Conti, J.P. (2008). The Androids are coming. *IEEE Engineering and Technology*, Vol.3(9).
- Dai, J., Teng, J., Bai, X., Shen, Z & Xuan, D (2010). Mobile phone based drunk driving detection. In 4th Int. Conf. On Pervasive Computing Technologies for Healthcare 2010, pp.1-8, 22-25.
- Dybå, Tore (2000) Improvisation in Small Software Organizations, *IEEE Software*, 17, 82-87
- Erickson, J., Lyytinen, K. & Siau, K. (2005) Agile modeling, agile software development, and extreme programming: The state of research, *Journal of Database Management*, vol. 16, pp 22-24.
- Figuracion, E., Palaoag, T., Ignacio, D & Doblon, M. J. (2016). Albay Emergency Response and Report Tool (ALERTT). NETCOM, NCS, WiMoNe, GRAPH-HOC, SPM, CSEIT – 2016, pp. 161– 177
- Gonzalez, J. (2008). First Google's android phone launched. *IEEE Vehicular Technology Magazine*, Vol.3.
- Hou, Q (2010). Research and implementation of remote heart rate monitoring system based on GSM and MCU. In 2nd Int. Conf. On Information Science and Engineering, Hangzhou, 2010
- Karimul, M. M. & Alam, M. J. (2017). Mobile Tracking System using Web Application and Android Apps. *International Journal of Engineering Research & Technology (IJERT)*, Vol. 6 (2).
- Lagazio, M. (2012). The Evolution of the Concept of Security. *The Thinker*, Vol. 43: 36-41.
- Lin, J.Y. Yang, B.K. & Do, T.A. (2013). The Accuracy Enhancement of GPS Track in Google Map. Eighth International Conference on Broadband, Wireless Computing, Communication and Applications, 2013.
- Nerur, Sridhar & Balijepally, Venu Gopal (2007) Theoretical Reflections on Agile Development Methodologies, *Communications of the ACM*, 50, 79-83.
- Qureshi, A. B. (2014). The Afghanistan Profile of Natural and Technological Disasters. *Emergency and Disaster Reports*, Vol. 1 (1)
- Samuel, J. (2015). Implementation of GPS Based Object Location and Route Tracking on Android Device. *International Journal of Information System and Engineering*, Vol. 3 (No.2)
- Sana, Khan, Walayat, Qasim, D., Khan & Javaid (2014). Design and Development of a Low-Cost Ubiquitous Tracking System. *Procedia Computer Science* 34, 220 – 227
- Sharma, A., Sharma, J., Monga, D. & Aggarwal, R. (2014). Mobile Tracing Software for Android Phone. *International Journal of Engineering Research and General Science*, Vol. 2 (5).
- Tamboli, A. S. (2013). Incident Reporting System Using GIS. Retrieved from www.estij.org/papers/vol3no12013/10vol3no1.pdf.
- Whipple, J., Arensman, W. & Boler, M.S (2009). A Public Safety Application of GPS-Enabled smart phones and the Android Operating System. *IEEE Int. Conf. on System, Man and Cybernetics*, San Antonio, 2009, pp. 2059-2061.
- Yang, S., Lee, D. & Chen, K. (2011). A new ubiquitous information agent system for cloud computing - Example on GPS and Bluetooth techniques in Google Android platform. in *Int. Conf. On Electric Information and Control Engineering*, 2011, pp. 1929-1932.

A survey of Markov Chain on Metric Space.

U. E. Edeke and G. I. Attah

Mathematics Department, University of Calabar, Calabar.

Abstract

This work provides a comprehensive review of Markov chain on the Euclidean Metric Space. Properties and convergence theorem on metric space with applications in stochastic modelling are considered. Kolmogorov equations and other diffusion processes are majorly considered, among others, as case studies.

1. Introduction

Markov chains are stochastic processes commonly used to model systems evolving over discrete time steps. Traditionally, these chains are defined on discrete state spaces, where transitions between states occur with certain probabilities. However, extending Markov chains to metric spaces allows for the modeling of systems with continuous state spaces, providing a more comprehensive framework for understanding dynamic processes (Tijms, 2011). The motivation for studying Markov chains on metric spaces stems from the need to accurately model systems characterized by continuous variables. In many real-world scenarios, state transitions occur gradually and smoothly, necessitating a representation that accommodates continuous state spaces. By extending Markov chains to metric spaces, researchers can capture the nuanced behavior of such systems, offer insights into their dynamics and behavior. In physics, for example, Markov chains on metric spaces can be used to model the diffusion of particles in continuous space (Enting, 2013). In biology, they can represent the evolution of populations with continuously varying traits (Waxman, 2009) Markov models of evolution. In economics: (Ozer and Wei, 2017). Markov chain modeling for economic system, they can model the behavior of financial markets with continuous price dynamics. In computer science, they can be applied to optimization problems involving continuous variables.

2. Preliminaries.

The major parameter used for this study is a sequence of random variables X_0, X_1, \dots (taking values in I) whose joint distribution is determined by simple rules. Let I be a countable set, $\{i, j, k, \dots\}$, each $i \in I$ is called a state and I is called the state space. We work in a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$. Here Ω is a set of outcomes, \mathcal{F} is a set of subsets of Ω , and for $A \in \mathcal{F}$, $\mathbb{P}(A)$ is the probability of A . We have the following definitions.

2.1 Markov Chains

Stochastic processes that model the evolution of systems over discrete time steps, characterized by the Markov property, where the future state depends only on the current state and not on the history of the system is known as the Markov Chain. (Kijimi, 2013)

2.2 Metric Spaces

Mathematical structures consisting of a set of points along with a distance function (metric) that defines the distance between any two points in the space, satisfying certain properties such as non-negativity, symmetry, and the triangle inequality. Mathematically, it is defined as: The pair $\{M, \rho\}$ is a metric space if M is a non-empty set and $\rho: M \times M \rightarrow \mathbb{R}^+$ is a map, such that

- 1) $\rho(x, y) \geq 0, \rho(x, y) = 0 \leftrightarrow x = y$ for all $x, y \in M$
- 2) $\rho(x, y) = \rho(y, x)$ for all $x, y \in M$ (symmetric property)
- 3) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for all $x, y, z \in M$ (Edeke and Ousseinni, 2023).

For example, let X be a nonempty set; a map $d: x \times x \rightarrow R$ is defined as

$$\begin{aligned} d(x, y) &= 0 \text{ if } x = y, \\ d(x, y) &= 1 \text{ if } x \neq y = 1 \end{aligned}$$

Here d is called discrete metric on X .

2.3 Transition Probabilities

Transition probability is a probabilities associated with transitioning from one state to another in a Markov chain, indicating the likelihood of moving from a current state to a future state within a single time step. A transition matrix, also known as a stochastic matrix is a square matrix used to describe the transitions of a Markov chain. It is a matrix where each element represents the probability of transitioning from one state to another (Gahiiagher and Finn, 2018).

2.4 Stochastic process

Let T be an ordered set, $(\Omega, \mathcal{F}, \rho)$ a probability measure and (E, G) a measurable space. A stochastic process is a collection of random variables $X = \{X_t, t \in T\}$ where for each fixed $t \in T$ (Pavliotis, 2015).

X_t is a random variable from (Ω, \mathcal{F}, P) to (E, G) . Where Ω is known as the sample space, E is the state space of the stochastic process X_t .

A Markov process is a stochastic process for which given the present, the past and future are statistically independent.

2.5 Stationary Distribution

A probability distribution is known as a stationary distribution if it remains unchanged by the transition dynamics of a Markov chain, representing the long-term behavior of the system. A property of a Markov chain where every state in the state space is reachable from every other state with positive probability, ensuring that the chain does not become trapped in a subset of states is known as its irreducibility property. (Meyn and Tweedie, 2012; Hairer and Mattingly, 2011).

2.6 Probability Density Function

This is a function that describes the likelihood of a continuous random variable taking on a particular value within a given range, often used to define transition probabilities in continuous state spaces (Resnick, 2013). A mathematical function that specifies the probability of transitioning from one state to another in a Markov chain on a metric space, extending the notion of transition probabilities to continuous state spaces is known as its transition kernel. (Meyn and Tweedie, 2012).

3. Review of some results concerning Markov Chain on Metric Space.

Let M be a separable metric space with metric d that is equipped with its Borel σ -field $\beta(M)$. Let $\beta(M)$ denote the set of real-valued bounded measurable functions of M equipped with the norm.

$$\|f\|_{\infty} := \sup_{x \in M} |f(x)| \quad (3.1)$$

If μ is a (non-negative) measure on M and $f \in L^1(\mu)$ (or $f \geq 0$ measurable), we let

$$\mu f := \int_M f(x) \mu(dx) \quad (3.2)$$

denote the integral of f with respect of μ . We have the following definitions.

3.1 Markov kernel

A Markov kernel on M is a family of measures

$$P = \{P(x, \cdot)\}_{x \in M} \quad (3.3)$$

such that

- (i) For all $x \in M$, $P(x, \cdot): \beta(M) \rightarrow [0, 1]$ is a probability measure;
- (ii) For all $G \in \beta(M)$, the mapping $x \in M \mapsto P(x, G) \in \mathbb{R}$ is measurable.

The Markov kernel P acts on functions $g \in B(M)$ and measures (respectively probability measures (according to the formular:

$$Pg(x) := \int_M P(x, dy) g(y) \quad (3.4)$$

$$\mu P(G) := \int_M \mu(dx) P(x, G) \quad (3.5)$$

$$P_{n+1}(x, \beta) = \int P_1(x, dy) P_n(y, \beta), n = 1, 2, \dots \quad (3.6)$$

3.2 Markov chain

Let (Ω, F, P) be a probability space. A filtration $\mathbb{F} = (F_n)_{n \geq 0}$ is an increasing sequence of σ -fields: $F_n \subset F_{n+1} \subset F$ for all $n \in \mathbb{N}$. $(\Omega, F, \mathbb{F}, P)$ is called a filtered probability space. An M -valued adapted stochastic process on $(\Omega, F, \mathbb{F}, P)$ is a family $(X_n)_{n \geq 0}$ of random variables defined on (Ω, F, P) , taking values in M and such that X_n is F_n measurable for all $n \in \mathbb{N}$. Given a filtered probability space $(\Omega, F, \mathbb{F}, P)$ and a Markov kernel P on M , a Markov chain with kernel P on $(\Omega, F, \mathbb{F}, P)$ is an M -valued adapted stochastic process (X_n) on $(\Omega, F, \mathbb{F}, P)$ such that

$$P(X_{n+1} \in G | F_n) = P(X_n, G) \quad (3.7)$$

for all $n \in \mathbb{N}$ and for all $g \in \beta(M)$. Equivalently,

$$E(g(X_{n+1}) | F_n) = Pg(X_n) \quad (3.8)$$

for all $n \in \mathbb{N}$ and for all $g \in \beta(M)$ (or all functions $g: M \rightarrow \mathbb{R}$ that are measurable and nonnegative). $E(\cdot | F_n)$ denotes conditional expectation with respect to F_n , and $P(X_{n+1} \in G | F_n) := E(1_{X_{n+1} \in G} | F_n)$.

The Kolmogorov Equation, which is also known as the forward equation that describes the time-evolution of the probability distribution of the Markov process is presented as a major reviewed result in the following theorem.

3.3 Theorem (Kolmogorov). Let $f(k) \in C_b(\mathbb{R})$ and let $u(x, s) := E(f(X_t) | X_s = x) = \int f(y) p(dy, t/x, s)$, with t fixed. Assume furthermore that the functions $b(x, s), \Sigma(x, s)$ are smooth in both x and s . Then $u(x, s)$ solve the final value problem for $S \in [0, t]$.

$$\begin{aligned} \frac{\partial u}{\partial s} &= b(x, s) \frac{\partial u}{\partial x} + \frac{1}{2} \Sigma(x, s) \frac{\partial^2 u}{\partial x^2}, \\ u(t, x) &= f(x). \end{aligned}$$

The Kolmogorov Equation above, that describes the time-evolution of the probability distribution of the Markov process can be given in a compact form as

$$\frac{\partial}{\partial t} P(t, x, A) = Q P(t, x, A)$$

Where Q = generator of the Markov process,

$P(t, x, A)$ = Transition probability from X to A at time t .

REFERENCES

- Enting, I. G. (2013). *Markov chain Monte Carlo and climate modeling*. In Bayesian Theory and application, 509 – 526. Oxford University Press.
- Evans, L. C. (2013). *An introduction to stochastic differential equations*. American Mathematical Society, Providence, RI.
- Gardiner, C. W. (1985). *Handbook of stochastic methods*. Springer-Verlag, Berlin, second edition. For physics, chemistry and the natural sciences.
- Hairer, M. & Mattingly, J. C. (2006). Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math*, (2)164(3), 993-1032.
- Hairer, M. & Mattingly, J. C. (2011). *Yet another look at Harris Ergodic theorem for Markov chains*. In seminar on Stochastic Processes, 2011. Springer.
- Hairer, M. (2006). Ergodic properties of Markov processes, Lecture given at the University of Warwick.
- Meyn, S. P. & Tweedie, R. L. (2009). *Markov Chains and Stochastic Stability*. Cambridge University Press.
- Meyn, S. P. (2008). *Control Techniques for Complex Networks*. Cambridge University Press.
- Ozer, O. & Wei, W. (2017). Markov Chain Models for Complex Systems: Theory and Applications. *IEEE Transactions on Automatic Control*, 62(6), 2831-2846.
- Pavliotis, G. A. (2015). Stochastic processes and applications. Department of Mathematics, Imperial College London, November 11, 2015.
- Robert, C. P. (2012). *Monte Carlo Statistical methods*. Springer.
- Robert, P. (2003). Stochastic networks and queues, french ed., Applications of Mathematics (New York), 52, Springer-Verlag, Berlin, 2003, *Stochastic Modelling and Applied Probability*.
- Tijms, H. C. (2011). *Stochastic process*. An Introduction. Wiley.
- Tweedie, R. L. (2001). Markov Chains: Statistical Structure and Applications. *Journal of Applied Probability*, 38(A), 1-15.
- Wax, N. (1954). *Selected Papers on Noise and Stochastic Processes*. Dover, New York.

Asymmetric Power Autoregressive Conditional Heteroscedasticity Modelling of Interest Rate Return in Nigeria

Andrew Ekpung Ntu and Eric Monfung Egomo

Department of Statistics, University of Cross River State

Corresponding Author: andrewntul@unicross.edu.ng

ABSTRACT

This study models Nigerian commercial bank interest rates from January 1997 to December 2023 using Generalised Autoregressive Conditional Heteroscedasticity (GARCH). The Central Bank of Nigeria (CBN) statistical bulletin provided the data. The original series' time plot demonstrated the exchange rate return series' trend and metamorphosis, which caused it to become stationary. The variables were stationary at lag one, according to the results of the series' Augmented Dickey-Fuller test. To simulate the interest rate return series, the study used both symmetric and asymmetric (GARCH) models, which capture characteristics of financial time-series data such as volatility clustering and leverage impact. Asymmetric power autoregressive conditional heteroscedasticity (APARCH (1,1)), however, was chosen as the best model for the return series after six models were calculated for the conditional variance. Ultimately, the asymmetric APARCH model is the most suitable model to estimate the interest rate volatility. The model's post-estimation revealed that its distribution was normal.

Keywords: Interest Rate, APARCH, GARCH, Volatility and Leverage Effect.

1. Introduction

One clever way to increase your savings is to make timely investments in deposit money banks. Deposit money banks, as opposed to traditional savings accounts, safeguard investors' funds and enable them to generate better returns; these returns are known as interest. Sunday (2012) states that this interest is a payment that a borrower makes to a lender in exchange for using money that has been put in the bank for a predetermined amount of time. He went on to say that interest is the payment made to those who contribute the money needed to buy capital goods. Conversely, the rate can also be interpreted as a gauge or regularity of a certain fee paid or assessed for the total amount of money placed in the bank for a specific time frame.

It was discovered that the set price paid or charged fluctuates over time, regardless of the amount or frequency of the payment made against the total amount of money deposited in the bank for a specific amount of time. It is observed to be very low at times and high at others. Sunday (2012) disclosed that investors benefit from low interest rates on money deposited in banks, as high interest rates put depositors in danger of losing money.

Volatility in finances is one of the greatly important concept, which measures variances of variable asset return. Volatility is the uncertainty surrounding potential price movement and it is many a time used as a basic measure of the total risk of financial assets. A number of models have been developed to investigate volatility across different countries. The most common model to estimate exchange rate volatility is the GARCH model developed independently by Bollerslev (1986), Taylor (1986) and Engle (1986).

Volodymyr (2002) recorded that, interest rates behave differently in various nations under various conditions. For instance, certain nations impose interest levies on bank deposit facilities and other related economic obligations. Shocks can result from various factors, such as fast changes in interest rates, unanticipated swings, and all expenses. According to Dmytro (2017), interest rates have been sharply declining globally since 2008. This was the initial reaction to the 2008–2009 global financial crisis shock brought on by the US subprime mortgage crisis. According to Dmytro (2017), central banks all over the world have lowered discount rates and implemented open market operations (OMO) to expand the money supply. This has allowed commercial banks to borrow money at lower rates to satisfy reserve requirements and expand lending. However, since the first response to the shock of the global financial experience, it has been often believed that interest rate reacts more dramatically to changes in volatility. This reality calls for a more exact volatility modelling. Besides, an accurate volatility model and forecasting can allow a more appropriate estimation of the value at risk. Therefore, it becomes necessary to develop an appropriate model for modelling shocks in interest rate levies in a developing country like Nigeria. According to Volodymyr (2002), the theoretical approach to modelling interest rates differs according to countries. Some interest rate models developed and tested in one country may not be appropriate when they are applied in a transitory or inflationary economy. Nonetheless, it has long been held since the initial reaction to the global financial crisis that interest rates respond more sharply to fluctuations in volatility. This fact necessitates more precise volatility modelling. Furthermore, a more precise evaluation of the value at risk can be made possible by reliable volatility modelling and forecasting. As a result, it becomes essential to create a suitable model for simulating interest rate shocks in developing nations like Nigeria. Volodymyr (2002) asserts that the theoretical framework for interest rate modelling varies by nation.

2.0 METHODOLOGY

2.1 Conditional Mean Model

2.1.1 ARMA (p,q)

Both an AR and an MA procedure are involved in the movement of the Nigerian Naira (NN) to the US Dollar (USD). Consequently, ARMA (p,q), where p denotes the autoregressive term order and q denotes the moving-average term order, may be expressed as

$$y_t = \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t + \sum_{j=1}^q \beta_j \varepsilon_{t-j} \quad (1)$$

An autoregressive moving average model of orders p and q, denoted ARMA (p, q), is stated to be followed by a series $\{y_t\}$. The model is stationary and invertible due to the constants β_j and α_i , and $\{\varepsilon_t\}$ is a white noise process.

2.2. Volatility Models

There are two primary categories of volatility modelling techniques: symmetric models and asymmetric models. While the asymmetric model's negative and positive size shocks have differing effects on future volatility, the symmetric model's conditional variance solely depends on the asset's magnitude and not its sign. Brook, Chris (2008).

2.2.1 Symmetric Models

2.2.2 ARCH (1)

The conditional variance of a series is modelled using the autoregressive conditional heteroscedastic (ARCH) model. This approach is frequently applied to explain variations that rise and decrease. Assume that the variance of a series y_t is to be modelled. The ARCH (1) model states that the variance of y_t depends on y_{t-1} at time t .

The ARCH model is represented mathematically as follows.

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 \quad (2)$$

We impose the constraints that $\alpha_0 \geq 0$ and $\alpha_1 \geq 0$ to avoid negative variance

2.2.3. GARCH (1, 1) Model

The variance at time t is modelled by the generalised autoregressive conditional heteroscedastic model, which permits the conditional variance to depend on prior lags. The model uses the value of the past squared observation and past variance. The replicas calculate the amount that a volatility shock from today influences the volatility in the following period. Determines how quickly this effect changes over time. Below is an illustration of a GARCH (1, 1) model.

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (3)$$

This is a GARCH (1, 1) model. σ_t^2 is known as the conditional variance since it is a one-period ahead estimate for the variance.

2.2.4 GARCH- M (1,1)

In the financial markets, a high level of risk is thought to yield a high return. One could think about the GARCH IN MEAN model in these kinds of situations. The model permits a sequence's condition mean to rely on its conditional variance.

The following is the model.

$$r_t = \mu + \lambda \sigma_t^2 + y_t \quad (4)$$

$$y_t = \sigma_t \varepsilon_t$$

$$\varepsilon_t \sim (0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (5)$$

Where λ and μ are constant, if λ is positive the return is also positive related to volatility

2.3 Asymmetric Models

It is necessary to discuss the ASYMMETRIC GARCH model because negative shocks, or bad news, typically have a greater impact on volatility than positive shocks. To that aim, we limited our research to the most well-known asymmetric GARCH models, including EGARCH, TS-GARCH, and APARCH.

2.4 EGARCH Model:

The model is known as exponential GARCH (EGARCH). Compared to the pure GARCH specification, the model has several advantages. First, even if the parameters are negative,

(σ_i^2) will still be positive because the $\log(\sigma_i^2)$ is simulated. Therefore, the model parameters do not require the imposition of non-negativity requirements in an unnatural manner. Second, asymmetries are permitted by the EGARCH formulation since the model has the following representation if the link between volatility and returns is negative, γ will be negative:

$$\text{Log}h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{|\varepsilon_{t-1}| + \gamma_i \varepsilon_{t-1}}{h_{t-1}^{\frac{1}{2}}} + \sum_{j=1}^q \beta_j \log h_{t-j} \quad (6)$$

Where γ is leverage effect co-efficient. (If $\gamma > 0$ it indicates the presence of leverage effect). Note that when ε is positive there is good news, when ε is negative there is bad news

2.5 Ts GARCH Model:

Another GARCH method that is capable of modelling leverage effects is the Threshold GARCH (TGARCH) model, which has the following form:

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^p \gamma_i \varepsilon_{t-1}^2 s_{t-1} + \sum_{j=1}^q \beta_j h_{t-1} \quad (7)$$

Where $s_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{if } \varepsilon_{t-1} \geq 0 \end{cases}$

γ is the leverage effects coefficient. (if $\gamma > 0$ it indicates the presence of leverage effect). That is depending on whether ε is above or below the threshold value of zero, ε_t^2 has different effects on conditional variance h_t when ε_{t-1} is positive.

2.6 APARCH (1,1)

The Asymmetric Power ARCH (APARCH) model can handle both power transformation of the variance and asymmetric effects. The following is its specification for the conditional variance.

$$\sigma_t^2 = \alpha_0 z_t + \sum_{i=1}^q \alpha_i (|u_t| - \gamma_i u_t) + \sum_{j=1}^p \beta_j \sigma_{t-1}^2 \quad (8)$$

Where $\sigma_t = \sqrt{h_t}$, the parameter γ (assumed positive and ranging between 1 and 2)

3.0 RESULTS

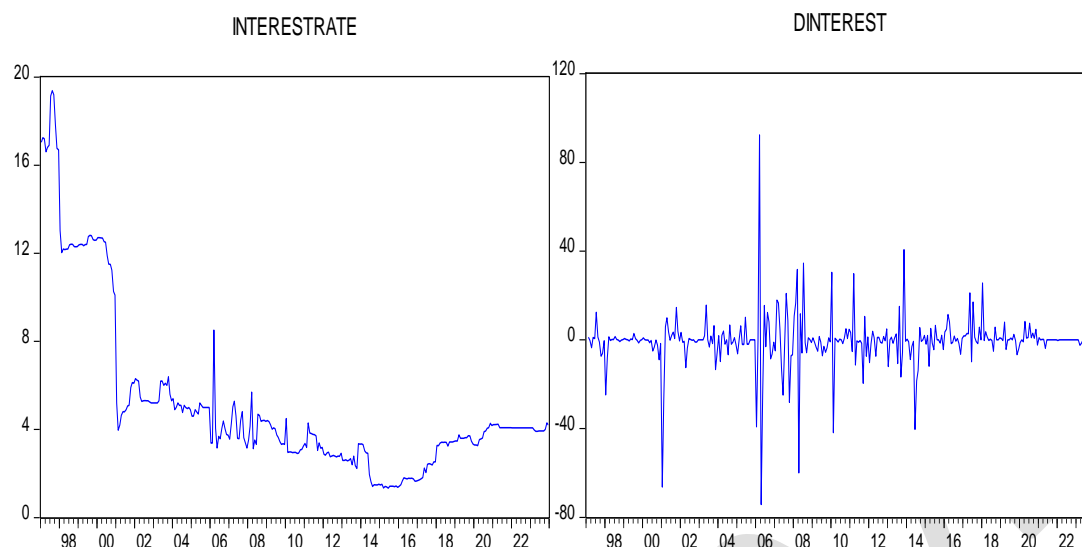


Figure 3.1 The time plot at levels and of Returns on Interest Rate

The presence of a trend in a series will make it not to remain stationary. The first difference's time plot demonstrates a stationary process and provides indications of clustering in the interest rate returns series time plot.

Table 3.1: Summary of Descriptive Statistic for Interest Rate

Mean	Medium	Maximum	mini	Std. dev	skewness	kurtosis	prob	sum	observe
5.16	4.03	19.38	1.33	3.8	3.8	1.8	0.00	1672.3	324

Similarly, the descriptive test was applied to the study's variable, interest rate returns. The descriptive statistics for the series are shown in Table (4.1), along with the mean, median, maximum, minimum, standard deviation, skewness, kurtosis, and Jarque-Bera test statistic. The variable's sample mean is positive, indicating a positive mean return; nevertheless, the return series' matching standard deviation is considerable. Furthermore, the series' skewness statistics are positive. This indicates a high increase (rightward tilt) in the series. One of the prevalent features of financial time series data is an indication that is skewed to the right. Additionally, the 5.88 kurtosis values indicate the existence of a fat tail. At 306.4977, the Jarque-Bera (J-B) test statistic demonstrates statistical significance.

Table 3.2: STATIONARITY TEST

	Raw series	Return series
Critical value	ADF test statistics (-3.382)	ADF test statistics (-16.2)
1%	-3.45	-3.45
5%	-2.867	-2.867

10%	-2.57	-2.57
-----	-------	-------

Stationarity tests were performed on the data to prevent false regression. The unit root on variables was tested using the Augmented Dickey-Fuller (ADF) test; Table 4.2 displays the outcome of Since the series is non-stationary, the results of the Augmented Dickey-Fuller (ADF) test at the level and first differences and probability values in brackets indicate the presence of a unit root. The probability values (p-values) at the level are less than 0.05 (p-values >0.05) in the variable. Gujarati (2003)

Table 3.3: Testing for the Presence of an ARCH Effect

Estimator	Lag 1
F-statistic	0.027
Prob F(1,237)	0.869
n*R ²	0.027
X ² (1,1)	0.869

3.1 Test for Heteroscedasticity

Nonetheless, the residual derived from the return series' test for heteroscedasticity (the ARCH effect) is displayed in Table (4.3). An ongoing, permanent ARCH effect was found after the return series was tested for it. The chi-squares test statistics (0.8689) correlate to a higher value than the F-statistic (0.027). This suggests that ARCH effects are present in the return series.

Table 3.4: Model Estimation for Symmetric and Asymmetric Models of Return on Interest Rate

Models/ Parameter.	Symmetric GARCH			Asymmetric GARCH		
	ARCH (1)	GARCH (1,1)	GARCH-M (1,1)	APARCH (1,1)	TS GARCH 1,1)	GARCH (1,1)
C	101.39 (0.000)	30.276 (0.000)	0.28.939 (0.000)	0.087 (0.000)	9205.931 (0.9637)	0.0019 (0.000)
α_1	0.253 (0.0042)	0.624 (0.000)	0.714 (0.000)	0.244 (0.000)	0.997 (0.000)	0.465 (0.000)
β		0.375 (0.000)	0.356 (0.000)	0.9268 (0.000)	0.4215 (0.1155)	0.616 (0.000)
γ				0.876 (0.000)	0.169 (0.323)	-0.124 (0.000)
λ				0.5278 (0.000)	0.7672 (0.000)	0.822 (0.000)

$\alpha + \beta$		1.001	1.07	1.17	1.4	1.081
log	-1227.7	-1184.89	-1184.095	-1160.21	-1184.617	-1163.325
AIC	7.61	7.36	7.35	7.22	7.366	7.234
SC	7.63	7.408	7.403	7.29	7.42	7.29
Obs	324	324	324	324	324	324

Using the information criteria and log-likelihood function, six models were estimated for the interest rate return: three symmetric models and three asymmetric models. At lag 1, every variable was stationary. The asymmetric model comprises APARCH, TS-GARCH, and EGARCH, whereas the symmetric model contains the ARCH, GARCH, and GARCH-M models. Table 4.4 indicates that all of the return series' models have a coefficient of the ARCH model that is greater than zero ($\alpha_0 > 0$), indicating extremely high volatility. The financial mark is more risky when volatility is high, and the positive news is lower when volatility is low.

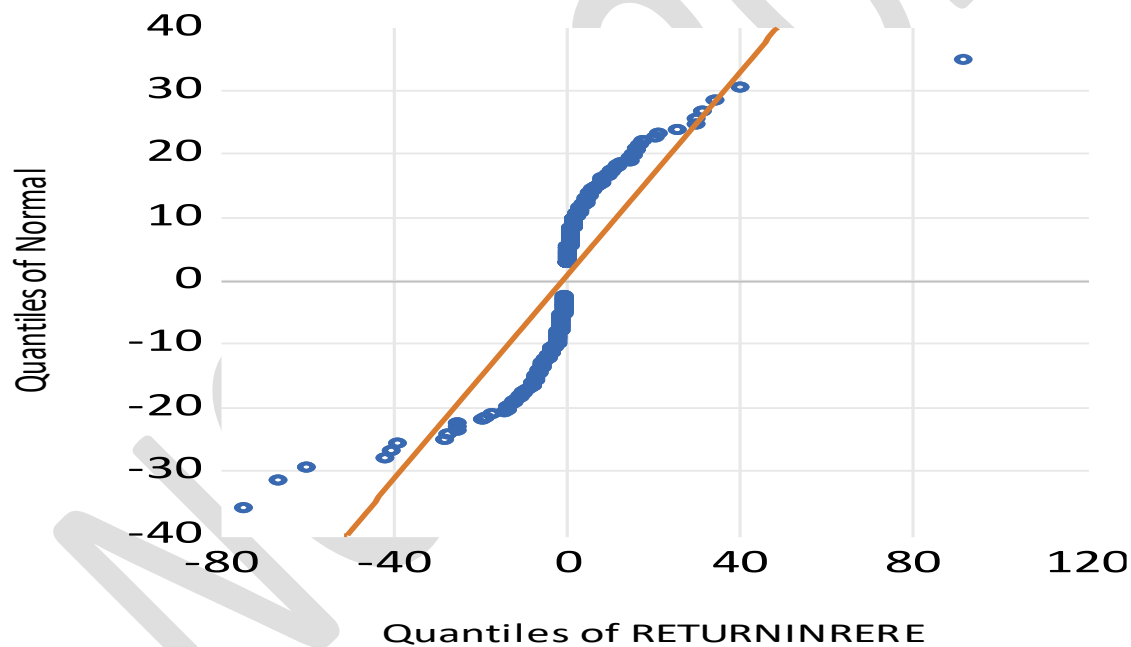


Figure 3.4 Quantile –Quantile plot of standardized residual fitted from APARCH(1,1)

Model return on interest rate

3.2 Quantile –Quantile plot of standardized residual fitted

Strong evidence of a normal distribution may be seen in the standardised residuals plot's normal quantile-quantile plots. It is seen that the lines of normalcy and quantile lie exactly on top of one another.

Included observations: 323

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.149	-0.149	7.2645	0.007
		2 -0.152	-0.179	14.849	0.001
		3 0.012	-0.044	14.893	0.002
		4 -0.011	-0.046	14.932	0.005
		5 0.053	0.041	15.875	0.007
		6 0.100	0.115	19.189	0.004
		7 -0.086	-0.033	21.628	0.003
		8 -0.068	-0.056	23.164	0.003
		9 -0.039	-0.084	23.674	0.005
		10 -0.063	-0.119	25.016	0.005
		11 0.057	-0.011	26.110	0.006
		12 0.074	0.056	27.955	0.006
		13 -0.059	-0.004	29.129	0.006
		14 -0.013	0.017	29.188	0.010
		15 -0.034	-0.036	29.589	0.013
		16 -0.019	-0.044	29.719	0.020
		17 0.094	0.036	32.736	0.012
		18 0.102	0.107	36.334	0.006
		19 -0.094	-0.022	39.366	0.004
		20 -0.053	-0.035	40.334	0.005
		21 -0.033	-0.060	40.718	0.006
		22 0.048	0.007	41.508	0.007
		23 0.027	-0.020	41.763	0.010
		24 0.123	0.142	47.057	0.003
		25 -0.175	-0.095	57.821	0.000
		26 0.014	0.037	57.895	0.000
		27 -0.031	-0.063	58.231	0.000
		28 0.034	0.008	58.640	0.001
		29 0.105	0.059	62.562	0.000
		30 -0.084	-0.068	65.102	0.000
		31 -0.037	0.017	65.584	0.000
		32 0.041	0.030	66.192	0.000
		33 0.007	0.033	66.212	0.001
		34 0.068	0.071	67.882	0.000
		35 -0.025	-0.054	68.105	0.001
		36 0.033	0.063	68.512	0.001

Figure 3.5: Correllogramme of Standardized Residuals Square

3.3 Correlogram of APARCH Model

Since the lag decreased gradually and the lags of the ACF and PACF all fell within the same level non-lag cut across the level, the correlogram of the square residual demonstrated the existence of the ARCH effect as the value of the ACF, and PACF.

4.0 Conclusion

The research model uses Nigerian commercial bank interest rate data from 1997 to 2023. The information was taken from the Nigerian Central Bank website. The application of several Garch models to model the interest rate among Nigerians was the main emphasis of this study. A rigorous analysis of the time plot showed that the series' movement was erratic. The unit root was tested for each of the variables using the Augmented Dickey-Fuller (ADF) test. Since the series is non-stationary, the unit root was found because of the ADF test at the level and first differences and probability values in brackets, the probability values (p-values) at a level are less than 0.05 (p-values >0.05) in the variables.

For interest rates, the univariate specification of generalised autoregressive conditional heteroscedasticity (GARCH) models was utilised. These models included symmetric and asymmetric models that capture features of a financial series, such as leverage impact and volatility clustering. Based on the minimal AIC and SC information criterion (AIC=-7.22 SC=7.29), the APARCH model is determined to be the best model when comparing the symmetric and asymmetric models using the information criteria. The sum of the coefficients on the lagged squared error and lagged conditional variance is one (1), and the coefficients on the lagged squared residual and lagged conditional variance term in the conditional variance equation are both highly statistically significant.

A common illustration of GARCH models for return series on financial assets is this one. This suggests that there will be very persistent shocks to the condition variance.

On the conditional variance, however, good and bad news have different effects: good news affects α , whereas bad news affects $(\alpha+\beta)$. The good news has an impact of 0.244 and the bad news has an impact of 1.17 in the APARCH model. λ has a positive and statistically significant value. This suggests that higher risk, as indicated by a higher conditional variance, raises the interest rate mean return.

REFERENCES

- Atoi, M. (2014). Testing Volatility in Nigeria Stock Marketing using Generalized Autoregressive Conditional Heteroskedastic Model. *Central Bank of Nigeria Journal of Applied Statistics*, 3(5), 2- 65.
- Bencik, M. (2009). The Equilibrium Interest Rate Theoretic Concepts and Application Research Department National Bank of Slovakia.
- Bollerslev, T. (1986). Generalized Auto-Regressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-321.
- Bollerslev *et al.* (1992). 7 GARCH Models of Volatility Handbook of Statistics 14, 209-240.
- Box, G. E. P. & Jenkins, G.M. (1976). Time series analysis: „Forecasting and control, Holden-Day, San Francisco.
- Brooks, C. (2006). Introductory Econometrics for Finance. Second Edition, Cambridge University Press.
- Central Bank of Nigeria (2024), Retrieved from <http://www.Statistics.cbn.gov.ng/statsonline> 16th
- Cont, R. (2001). Empirical Properties of Asset Returns: Stylized Facts And Statistical Issues. *Quantitative Finance*, 1(4), 223–236.
- Deebom, Z.D. & Essi, I.D. (2017). Modelling price Volatility of Nigeria Crude Oil Markets using GARCH Model: 1987–2017. *International Journal of Applied Science & Mathematical Theory*, 4(3), 23.
- Engle, R. F (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Various of UK Inflation. *Economical* 4 (50), 987-1008.
- [Heston](#) S, L. & [Nandi](#), S. (1998). A Closed-Form GARCH Option Pricing Model. *International Journal of Business and Management Review*, 1 (3), 35-55.
- Hussainatu A. & Muftaudeen O.O. (2011). Macroeconomic Policy and Agricultural Output in Nigeria Implication for Food Security. *American Journal of Economics*, 4(2), 99-113
- Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(3), 327-343.
- Taylor, S.J. (1986). *Modelling Financial Time Series*. John Wiley. Chichester
- Iulian, P. (2012), Using Garch-in-Mean Model to Investigate Volatility and Persistence at Different
- Volodymyr, K. (2002). Interest Rates And Their Role In The Economy During Transition. The Problem of High-Interest Rates. Case of Ukraine. A thesis submitted in partial fulfilment of the requirements for the degree of Master of Arts in Economics

Operators defined by Convolution on the Euclidean Motion Group

¹U. E. Edeke, ²U.N. Bassey,

¹ Department of Mathematics, University of Calabar, Calabar, Nigeria

²Department of Mathematics, University of Ibadan, Ibadan, Nigeria

ABSTRACT

Let $\delta(t)$ be the Dirac measure on $SO(2) \cong T$, the circle group, and let $G = \mathbb{R}^2 \times T$ be the Euclidean motion group. It is established that the (convolution) operator

$$A': C_c^\infty(G) \rightarrow G^\infty(G)$$

Which is defined as $A'f = f * T^n \delta(t)$ extends to a bounded linear operator on $L^2(G)$, for $f \in C_c^\infty(G)$, the space of infinitely differentiable functions on G with compact support. Furthermore, we demonstrate that the left convolution operator L_T denoted as $L_T f = T * f$ commutes with left translation, for $T \in D'(G)$. Other notable results are also obtained.

Key words: Euclidean motion group, Dirac measure, Convolution operator

AMS Subject Classification: 43A70, 43A90

Introduction

Some convolution operators on the Euclidean motion group are the subject of this study. For $G = SE(2)$ realized as $\mathbb{R}^2 \times T$, the convolution operators are defined on the space $C_c^\infty(G)$ and take values in the space $C^\infty(G)$. Let H be a Hilbert space and let $G = SE(n)$ be generated by $\mathbb{R}^n \times SO(n)$. A specified operator for the subspace made up of equivalent classes of right K -invariant functions of the H -valued L^p -space on G denoted by $L^p(G; K, H)$, $1 < p < \infty$, and its adjoint A^* is obtained. The order of the work is as follows. The major results are given in section 3, with section 2 dealing with preliminary definitions and results required for the remainder of this work

2. Preliminaries

Let us define G as $C = \mathbb{R}^2 \times T$, then the group law on G may be expressed as

$$(a, e^{i\theta})(b, e^{i\phi}) = (a + e^{i\theta}b, e^{i(\theta+\phi)})$$

For all $(a, e^{i\theta}), (b, e^{i\phi}) \in \mathbb{R}^2 \times T$. The following proposition is needed in what follows. The proposition is found in Sugiura ([5]) but the proof presented therein is a little bit sparse and as such, we present the proposition here with a more detailed proof.

2.1 Proposition. In the event that f is a rapidly decreasing function on G , then the Fourier transform of f is as follows.

$$(f(\sigma)F)(s) = \int_K K_f(\sigma; s, r)F(r)dr \quad (2.1)$$

For any $\sigma > 0$ and $F \in L^2(K)$, $K = SO(2) \cong T$, $r, s \in T$, where

$$K_f(\sigma; s, r) = \int_{\mathbb{R}^2} f(\xi, rs^{-1})e^{-i(\xi\sigma)}d\mu(\xi), \quad (2.2)$$

Proof. If F and F^1 belong to $L^2(K)$ and $g = t(\zeta)r$, then $g^{-1} = r^{-1}t(-\zeta) = t(-r^{-1}\zeta)r^{-1}$ and

$$\begin{aligned}
 (\hat{j}(\sigma)F, F^1) &= \int_{R^2} \int_K f(\zeta, r) (U_{(-r-1\zeta)r^{-1}}^\sigma F, F^1) d\mu(\zeta) dr \\
 &= \int_{R^2} \int_K \int_K f(\zeta, r) e^{-i(sr^{-1}\zeta, s\sigma)} F(rs) \bar{F}'(s) d\mu(\zeta) dr ds \\
 &= \int_{R^2} \int_K \int_K f(\zeta, r) e^{-i(sr-1\zeta)r^{-1}} F(s) \bar{F}'(s) d\mu(\zeta) dr ds \\
 &= \int_K \int_K K_f(\sigma; s, r) F(r) dr \bar{F}'(s) ds
 \end{aligned}$$

Given that F' is arbitrary in $L^2(K)$, (2.1) holds for nearly all s in K . This claim demonstrates $\hat{j}(\sigma)$ is an integral operator on $L^2(K)$ whose kernel K_f is given by (2.2) iff $f \in S(G)$.

2.2 Corollary. As a function of $\xi \in R^2$, the ordinary Fourier transform of $f(\xi, r)$ is represented by $f(\xi, r)$

$$f(\xi, r) = \int_{R^2} f(\xi, r) e^{-i(\xi, \xi)} d\mu(\xi)$$

Then the kernel $K_f(\sigma; s, r)$ is given by

$$K_f(\sigma; s, r) = \tilde{f}(r\sigma, rs^{-1}).$$

Recall that the operator $U^{(p)}(g)$ defines a representation of $SE(2)$ on $L^2(K)$ as follows

$$U^{(p)}(g) \tilde{\psi}(X) = c^{-ip(a, X)} \tilde{\psi}(A^T X)$$

for each $g = (\bar{x}, \xi) = g(x_1, x_2)$, $\xi \in SE(2)$, where $p \in R^+$ and $X.Y = x_1 y_1 + x_2 y_2$. The vector X is a unit vector ($X.X = 1$).

It is necessary to review certain definitions, that may be needed in this research, on locally convex spaces. Let E be a vector space over the field of complex numbers. Let E be the union of an increasing sequence of subspaces E_n , $n = 1, 2, \dots$, and let each E_n have a Frechet space structure such that the injection of E_n into E_{n+1} is an isomorphism, i.e., the topology induced by E_{n+1} on E_n is the same as the topology given on E_n initially. Then, we can define the Hausdorff locally convex space structure on E as follows. Assuming convexity, a subset V of E is a neighborhood of zero if and only if $V \cap E_n$ is a neighborhood of zero in the Frechet space E_n for each $n = 1, 2, \dots$. We state that E is an LF-space or, equivalently, a countable strict inductive limit of Frechet spaces when we give it this topology. We may also state that the sequence of Frechet spaces, $\{E_n\}$, is a sequence of definitions of E . If every balanced, convex, bornivorous subset of a locally convex space is a neighborhood of zero, then the space is said to be bornological. The space $C_c^\infty(G)$ is a metrizable, locally convex, complete space; on the other hand, the space $C_c^\infty(G)$ is an LF-space. Both spaces are bornological with respect to their respective topologies.

We can now present the main results of this work.

3. Main Results

For the results presented in propositions 3.2 and 3.3 we take $G = SE(n) = R^n \times SO(n)$ and $K = SO(n)$. For a Hilbert space H and for $1 \leq p < \infty$, we denote the subspace formed by equivalent classes of right K -invariant functions of the H -valued $L^p(G; K, H)$. This space is the H -valued L^p on $X = G/K$, lifted to G . We are interested in integral operators of the form

$$A_\psi(x) = \int_x S(x, y)\psi(y)dy,$$

where S is the kernel

$$S: X \times X \rightarrow L(H_1, H_2)$$

such that

$$S(gx, gy) = U_1^\sigma(g)S(x, y)U_2^\sigma(g^{-1})$$

for all $g \in G$, with some uniformly bounded representations U_1^σ, U_2^σ of G on the Hilbert spaces H_1, H_2 . By lifting all functions to G we may define an operator A by

$$Af(g) = \int_G U_2^\sigma(h)k(h^{-1}g)U_1^\sigma f(h)dh, \quad (3.1)$$

where $k(g)$ satisfies the identity

$$k(hgh') = U_1^\sigma(h)k(g)U_2^\sigma(h^{-1}) \quad (3.2)$$

for all $h, h' \in K$. If $H_2 = H_2 = C$ and U_2^σ are trivial representations, then (3.2) only means that k is left and right invariant (that is, bi- K -invariant) under K . The expression (3.1) is called a twisted convolution of f and k denoted by $f *_{1,2} k$.

Proposition 3.1 below shows L^p -boundedness of the operator A while proposition (3.2) verifies its adjoint.

3.1 Proposition. Let U_1^σ, U_2^σ be representations of G on the Hilbert spaces H_1, H_2 having uniform bounds M_1, M_2 respectively. Let $k \in L^1(G: K, L(H_1, H_2))$ satisfying (3.2). Then for all $1 \leq p \leq \infty$, A defined by (3.1) is a continuous linear map $L^p(G: K, H_1) \rightarrow L^p(G: K, H_2)$ with bound not greater than $M_1 M_2 \|k\|_1$.

3.2 Proposition. Let A be as in proposition 3.2. Then, for any $\psi \in L^p(G: K, H_2^1)$, the adjoint A^* of A is given by

$$A^*\psi(g) = \int_G U_1^\sigma(h)^{* -1} k(h^{-1}g) * U_2^\sigma(h^*)\psi(h)dh$$

where

$$k(l) = U_2^\sigma(l)(l^{-1})U_1^\sigma(h)^{-1}$$

Proof. For every $f \in L^p(G: K, H_1)$,

$$\begin{aligned} \langle f, A^*\psi \rangle &= \langle Af, \psi \rangle \\ &= \int_G \left\langle \int_G U_2^\sigma(h)k(h^{-1}g)U_1^\sigma(h)^{-1} f(h)dh, \psi(g) \right\rangle dg \\ &= \int_G \left\langle f(h), \int_G U_2^\sigma(h)^{* -1} k(h^{-1}g)U_2(h)^* \psi(g)dg \right\rangle dh \\ &= \int_G \left\langle f(h), \int_G U_2^\sigma(h)^{* -1} k(h^{-1}g)U_2(h)^* \psi(g)dg \right\rangle dh \end{aligned}$$

By Fubini's theorem, the result follows thus

$$A^*\psi(g) = \int_G U_1^\sigma(g)^{* -1} k(g^{-1}h)U_2^\sigma(g)^* \psi(h)dh.$$

3.3 Theorem. Let $T = \frac{\partial}{\partial t}$. The operator

$$A': C_c^\infty(G) \rightarrow C^\infty(G)$$

defined by $Bf = f * T^n \delta(t)$ extends to a bounded linear operator on $L^2(G)$ and is unitary up to a constant multiple.

3.4 Theorem. Let $T \in D'(G)$. The convolution map L_T is a continuous linear map from $C_c^\infty(G)$ into $C^\infty(G)$ and it commutes with translations. Conversely, if $L: C_c^\infty(G) \ni f \mapsto T * f \in C^\infty(G)$ is a continuous linear map such that

$$L \circ \tau_a = \tau_a \circ L$$

for every fixed $a \in G$, there is a unique $T \in D'(G)$ such that

$$L(f) = T * f, \quad \forall f \in C_c^\infty(G).$$

Proof. Let $T \in D'(G)$ and let a be a fixed element in G , we denote the left translation of a function f in G by T_a . Then $T_a f(g) = f(a^{-1}g)$. Now, put $g' = b^{-1}g$ for a fixed $b \in G$. Then

$$\begin{aligned} [T * (T_a f)](g') &= \langle T, (T_a f)(b^{-1}g) \rangle \\ &= \langle T, f(a^{-1}b^{-1}g) \rangle = \langle T, f(ba)^{-1}g \rangle \\ &= (T * f)((ba)^{-1}g) = [T_a(T * f)](g'), \end{aligned}$$

which verifies that the map (convolution) commutes with left translations.

Conversely, suppose that $L: C_c^\infty(G) \ni f \mapsto T * f \in C^\infty(G)$ is a continuous linear map commuting with translations. The map

$$C_c^\infty(G) \ni f \mapsto (Lf)((0,0),1) \in C^\infty(G)$$

defines a continuous linear functional on $C_c^\infty(G)$. Hence, there is a unique $T \in D'(G)$ such that

$$(Lf)((0,0),1) = (T, \check{f}),$$

where $\check{f}(g) = f(g^{-1})$. We have also that

$$(Lf)((0,0),1) = (T * \check{f})((0,0),1).$$

Since L commutes with left translation and for any $T \in D'(G)$ we have and verifies that L is a convolution map. This completes the proof.

$$\begin{aligned} (Lf)(g) &= [T_a(Lf)] = [L(T_a f)]((0,0),1) \\ &= (T * (T_a(T_a f)))((0,0),1) = \langle T, (T_a f)(g) \rangle \\ &= \langle T, f(a^{-1}g) \rangle = (T * T_a f)(g), \end{aligned}$$

Conclusion

It is established in this research (theorem 3.3) that the (convolution) operator

$$A': C_c^\infty(G) \rightarrow C^\infty(G)$$

which is defined as $A'f = f * T^n \delta(t)$ extends to a bounded linear operator on $L^2(G)$, for $f \in C_c^\infty(G)$. It is further demonstrated that the left convolution operator L_T denoted as

$L_T f = T^* f$ commutes with left translation, for $T \in D'(G)$. Lastly, the adjoint of a specified operator for the subspace made up of equivalent classes of right K-invariant functions of the H-valued L^p - space on G denoted by $L^p(G: K, H)$, $1 < p < \infty$ is obtained in proposition 3.2 and its L^p boundedness is demonstrated in proposition 3.1

References

- Astengo, F.J. Blasio, B., Ricci, F; *The Schwartz Correspondence for the Complex Motion group*. Journal of Functional Analysis. Vol. 285, issue 8, 2023.
<https://doi.org/10.1016/j.jfa.2023.110068>
- Astengo, F., Blasio, B., Ricci, F; *On The Schwartz Correspondence for Gelfand Pairs of Polynomial growth*. European Mathematical Society, Vol 32, No.1, 2021 pp 79-96.
<http://doi.org/10.4171/RLM/927>
- Christensen, J.G., Olafsson, G., Casey, S.D.; *Sampling, Amenability and the Kunze - Stein Phenomenon*. International Conference on Sampling Theory and Application (SampTA). IEEE. 2015.
- Chirikjian, G.S., Kyatkin, A.J.; *Engineering Applications of Non-Commutative Harmonic Analysis*. CRC Press, New York. 2001.
- Geller, D., Stein, E. M.; *Singular Convolution Operators on The Heisenberg Group*. Bulletin of the American Mathematical Society. Vol. 6, No.1, 1982.
- Sugiura, M.; *Unitary Representation and Harmonic Analysis: An introduction*. North- Holland, New York. 1990.
- Vilenkin, N. Ja., Klimyk, A.IL; *Special Functions and the Theory of Group Representations*. Kluwer Academic Publishers. 1991.
- Vilenkin, N. Ja., Klimyk, A.U.; *Representation of Lie Groups and Special functions Vol J*. Kluwer Academic Publishers. 1991.
- Wolf, A. Joseph; *Harmonic Analysis on Commutative Spaces*. American Mathematical Society. 1936
- Yarman, C. E., Yazici, B.; *A Wiener Filtering Approach over the Euclidean Motion group for Radon transform Inversion*. Proceeding of SPIE - The international Society for Optical Engineering. May 2003.
- Yarman, C. E., Yazici, B.; *Euclidean Motion Group Representations and the Singular Value Decomposition of the Radon Transform*. Integral Transforms and Special Functions. Vol. 00, No. 00, June 2005, 1-34

Partial survey of Fourier transform on the Euclidean Motion Group.

U.E.Edeke¹, U.N.Bassey²

¹*Department of Mathematics, University of Calabar, Calabar, Nigeria.*

²*Department of Mathematics, University of Ibadan, Ibadan, Nigeria*

Corresponding Author Email :ueedeke@gmail.com

ABSTRACT

Let $SE(2)$ be a two dimensional Euclidean motion group realized as the semi-direct product of R^2 and $SO(2)$. The structure of $SE(2)$ and its explicit spherical function are presented. Furthermore, the Fourier transform of a kind of distribution on $SE(2)$ and its boundedness are studied, and the boundedness is found to be possible in the space of square integrable functions of $SE(2)$.

KeyWords: Euclidean Motion group, invariant differential operator, distribution, universal enveloping algebra, spherical functions.

AMS Subject Classification:43-XX, 46Fxx, 16Sxx, 44-XX

1. Introduction

Special functions introduced in analysis have been shown to have tight ties to the idea of linear representations of Lie groups. The spherical functions are prominent among these functions. In the current theory of infinite dimensional linear representation of Lie groups, spherical functions are crucial. Spherical functions for $SE(2)$ are the solutions of the radial part of Laplace-Beltrami

operator $\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial \psi^2}$ on $SE(2)$. They are the Bessel functions for n in

general. For $n = 2$, they are the Bessel functions of order zero. This work is centered on the spherical transform of distribution on $SE(2)$. This distribution is realized as the product of spherical function and dirac function on $SE(2)$. The work is organized as follows. Section two deals mainly on preliminaries concerning the Euclidean motion group. Section three is concerned with the representation and Fourier transform of distribution on $SE(2)$.

Preliminaries

2.1 The Euclidean Motion Group. The group $SE(n)$ is realized as the semi-direct product of R^n with $SO(n)$. Any member of $SE(n)$ may be denoted as $g = (\bar{x}\xi)$, where $\xi \in SO(n)$ and $\bar{x} \in R^n$. For any $g_1 = (\bar{x}_1, \xi_1)$ and $g_2 = (\bar{x}_2, \xi_2) \in SE(n)$, multiplication on $SE(n)$ may be defined as

$$g_1 g_2 = (\bar{x}_1 + \xi_1 \bar{x}_2, \xi_1 \xi_2)$$

and the inverse is defined as

$$g^{-1} = (-\xi \bar{x}, \xi^1)$$

Here (ξ^t) denotes a transpose. Alternatively, $SE(n)$ may also be identified as a matrix group whose arbitrary element may be identified as $(n+1) \times (n+1)$ matrix, given in the form

$$H(g) = \begin{pmatrix} \xi & \bar{x} \\ 0^t & 1 \end{pmatrix}$$

where $\xi \in SO(n)$ and $0^t = (0, 0, \dots, 0)$. It is observed that $H(g_1)H(g_2) = H(g_1g_2)$, $H(g^{-1}) = H^{-1}(g)$ and $g \mapsto H(g)$ an isomorphism between $SE(n)$ and $H(g)$.

2.2 The Schwartz space $S(G)$. Consider the Euclidean motion group $SE(2)$ realized as $R \times T$ where $R \cong R/2\pi\mathbb{Z}$. If we choose a system of coordinates (x, y, θ) on G with $x, y \in \mathbb{R}$ and $\theta \in T$, then a complex-valued C^∞ function f on $G = SE(2)$ is called rapidly decreasing if for any $N \in \mathbb{N}$ and $\alpha \in \mathbb{N}^3$ we have

$$p_{N,\alpha}(f) = \sup_{\theta \in T, \xi \in \mathbb{R}^2} (1 + \|\xi\|^2)^N (D^\alpha f)(\xi, \theta) < +\infty,$$

where

$$D^\alpha = \left(\frac{\partial}{\partial x} \right)^{\alpha_1} \left(\frac{\partial}{\partial y} \right)^{\alpha_2} \left(\frac{\partial}{\partial \theta} \right)^{\alpha_3},$$

($\alpha = (\alpha_1, \alpha_2, \alpha_3)$; $\xi = (x, y)$). The space of all rapidly decreasing functions on G is denoted by $S = S(G)$. Then S is a Frechet space in the topology given by the family of semi-norms $\{p_{N,\alpha}; N \in \mathbb{N}, \alpha \in \mathbb{N}^3\}$.

The space $S'(G)$ of (continuous) linear functionals on $S(G)$ is referred to as the space of tempered distributions on $G = SE(2)$. This space can be topologised by strong dual topology, which is defined as the topology of uniform convergence on the bounded subsets of $S(G)$ generated by the seminorms $p_\varphi(u) = |u(\varphi)|$, where $u : S(G) \rightarrow \mathbb{R}$ and $\varphi \in S(G)$. We close this section by defining the concept of convolution on the space $S(G)$.

3. Fourier transform of Spherical function on $SE(2)$

The Fourier transform of the group $SE(2)$ is needed in what follows. The following preparations concerning the representation of $SE(2)$ is presented first.

Let $L^2([0, 2\pi], \frac{d\alpha}{2\pi})$ be the space of square integrable functions on $T \cong [0, 2\pi] \cong SO(2)$. A representation of $SE(2)$ on $L^2([0, 2\pi], \frac{d\alpha}{2\pi})$ is given as

$$U^{(p)}(g)\bar{\psi}(X) = e^{-ip(\alpha, X)}\bar{\psi}(A^T X)$$

for each $g \in SE(2)$, where $p \in \mathbb{R}^+$

3.1 Definition. The Fourier transform of $f \in S(G)$ (or $f \in L^1(G)$) is a map

$$F(f) : \mathbb{R}^+_* \rightarrow B(L^2(G))$$

defined as

$$(Ff)(\sigma) = \int_G f(g) U^\sigma(g^{-1}) d\mu$$

and the inverse is defined as

$$f(g) = \int_0^\infty \text{Tr}(U_g^\sigma f(\sigma)) \sigma d\sigma,$$

where $g = (\bar{x}, \xi)$ and $d\mu(g)$ stands for a measure on G .

The Laplace-Beltrami operator on $SE(2)$ (see [3],[5]) is given by

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial \psi^2}. \quad (1)$$

A radial solution of (1) yields

$$\begin{aligned} J_\lambda(mr) &= \Gamma(1) \left(\frac{\sqrt{\lambda}r}{2} \right)^0 I_{\frac{2-2}{2}} \left(\sqrt{\lambda}r \right) \\ &= I_0(\sqrt{\lambda}r). \quad \square \end{aligned} \quad (2)$$

Next, the Fourier transform of expression (1) is considered. The integral representation of $J_0(\lambda)$ is defined as

$$J_0(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} e^{i\lambda \cos \phi} d\phi$$

Now

$$\begin{aligned} \mathcal{F}(J_0(\lambda)) &= \int_{-\infty}^\infty J_0(\lambda) e^{it\lambda} d\lambda \\ &= \int_{-\infty}^\infty \frac{1}{2\pi} \int_0^{2\pi} e^{i\lambda \cos \phi} e^{it\lambda} d\phi d\lambda \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^\infty e^{i\lambda \cos \phi} e^{it\lambda} d\phi d\lambda \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^\infty e^{-i\lambda(-t - \cos \phi)} d\phi d\lambda \\ &= \frac{1}{2\pi} \int_0^{2\pi} \delta(-t - \cos \phi) d\phi. \end{aligned}$$

We evaluate $\delta(-t - \cos \phi)$ and substitute the result as follows. It is known that $\delta(x) = \delta(-x)$, therefore, $\delta(-t - \cos \phi) = \delta(t + \cos \phi)$. $t + \cos \phi$ is a function, and we use one of the properties of Dirac function that says

$$\delta(g(\phi)) = \sum_{i=1}^n \frac{\delta(\phi - \phi_i)}{|g'(\phi_i)|}$$

ϕ_i are the roots of $g(\phi) = 0$, $g'(\phi)$ is the first order derivative of $g(\phi)$. Now, $g(\phi) = t + \cos\phi$, therefore, $g'(\phi) = -\sin\phi$. The two possible roots of $t + \cos\phi$ are $\phi_1 = \cos^{-1}(-t)$ and $\phi_2 = 2\pi - \cos^{-1}(-t)$. So therefore,

$$\delta(t + \cos\phi) = \frac{\delta(\phi - \cos^{-1}(-t))}{|\sin(\cos^{-1}(-t))|} + \frac{\delta(\phi - (2\pi - \cos^{-1}(-t)))}{|\sin(2\pi - \cos^{-1}(-t))|}.$$

We know that $\sin^2\phi + \cos^2\phi = 1$ and $\sin\phi = \sqrt{1 - \cos^2\phi}$. Let $\phi = \cos^{-1}(-t)$,

$$\begin{aligned} \sin(\cos^{-1}(-t)) &= \sqrt{1 - \cos(\cos^{-1}(-t))\cos(\cos^{-1}(-t))} \\ &= \sqrt{1 - \cos(\cos^{-1}(-t))^2} \\ &= \sqrt{1 - t^2}. \end{aligned}$$

Therefore, $\sin(\cos^{-1}(-t)) = \sqrt{1 - t^2}$. Now,

$$\begin{aligned} \delta(t - \cos\phi) &= \frac{\delta(\phi - \cos^{-1}(-t))}{|\sin(\cos^{-1}(-t))|} + \frac{\delta(\phi - (2\pi - \cos^{-1}(-t)))}{|\sin(2\pi - \cos^{-1}(-t))|} \\ &= \frac{\delta(\phi - \cos^{-1}(-t))}{\sqrt{1 - t^2}} + \frac{\delta(\phi - (2\pi - \cos^{-1}(-t)))}{\sqrt{1 - t^2}} \\ \frac{1}{2\pi} \int_0^{2\pi} \delta(t - \cos\phi) d\phi &= \frac{1}{2\pi} \frac{\delta(\phi - \cos^{-1}(-t))}{\sqrt{1 - t^2}} + \frac{\delta(\phi - (2\pi - \cos^{-1}(-t)))}{\sqrt{1 - t^2}} d\phi \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{1 - t^2}} \int_0^{2\pi} [\delta(\phi - \cos^{-1}(-t)) + \delta(\phi - (2\pi - \cos^{-1}(-t)))] d\phi \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{1 - t^2}} \int_0^{2\pi} [\delta(\phi - \cos^{-1}(-t)) d\phi + \int_0^{2\pi} \delta(\phi - (2\pi - \cos^{-1}(-t)))] d\phi \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{1 - t^2}} (1 + 1) \\ &= \frac{2}{2\pi} \frac{1}{\sqrt{1 - t^2}} \\ &= (\pi)^{-1} \sqrt{1 - t^2} \end{aligned}$$

This is the Fourier transform of the spherical function of $SE(2)$. It is our interest to use it in studying signal analysis and a Paley Wiener type theorem in our future research.

References

- [1] Bassey, U. N., Edeke, U. E.; *Convolution operators and equation on the Euclidean motion group*. JNSPS, Vol.6, issue 4, sept., 2024
- [2] Chirikjian, G.S., Kyatkin, A.B.; *Engineering Applications of Non-Commutative Harmonic Analysis*. CRC Press, New York. 2001.
- [3] Dieudonne, Jean, *Gelfand Pairs and Spherical Functions*. Internat. J. Math. & Math. Sci. Vol. 2, No.2 153-162, 1979. <http://eudml.org/doc/44859>.
- [4] Dietsel, J., Angela, S; *Joys of Haar Measure*. American Mathematical Society. 2014. <https://doi.org/MR3186070> <https://doi.org/10.1201/9780429289385>
- [5] Dijk, Gerrit Van; *Introduction to Harmonic Analysis and Generalized Gelfand Pairs*. Walter De Gruyter, Berlin. 2009. [doi.org/ 978-3110220193](https://doi.org/978-3110220193).
- [6] Edeke, U. E., Ariyo, R. D., and Dada, O. C., A Class of Positive Definite Spherical Functions on the Euclidean Motion group. Global Journal of Pure and Applied Sciences. Vol. 30, 2024. 551-558. DOI: <https://dx.doi.org/10.4314/gipas.v30i4.13>.
- [7] Edeke, U. E., Ariyo, R. D. Alternative Approach in Computing the Haar measure of SU(2). Achievers Journal of Scientific Research. Volume 06, Issue 02, pp. 50-54, December 2024. DOI: 10.5281/zenodo.14566269
- [8] Hormander, L.; *The Analysis of Linear Partial Differential Operators I*. Springer - Verlag, Berlin Heidelberg, 1990. <https://doi.org/10.1007/978-3-642-61497-2>
- [9] Jinman Kim, Wong, M. W.; *Positive definite temperature functions on the Heisenberg group*. Applicable Analysis: An international Journal. Vol.85, No. 8, 2006.
- [10] Sugiura, M.; *Unitary Representation and Harmonic Analysis: An Introduction*. North-Holland, New York. 1990.
- [11] Vilenkin, N. Ja., Klimyk, A.U.; *Representation of Lie Groups and Special functions Vol 1*. Kluwer Academic Publishers. 1991.
- [12] Wolf, A. Joseph; *Harmonic Analysis on Commutative Spaces*. American Mathematical Society. 1936
- [13] Yarman, C. E., Yazici, B.; *A Wiener Filtering Approach over the Euclidean Motion group for Radon transform Inversion*. Proceeding of SPIE - The international Society for Optical Engineering. May 2003.

A review of Sobolev space on Locally Compact group

U. E. Edeke and M. Joy

Mathematics Department, University of Calabar, Calabar.

Introduction

Sobolev spaces are a central and fundamental concept in functional analysis, differential geometry, potential theory, approximation theory and partial differential equations (PDEs), providing a framework for understanding the regularity properties of functions and the solutions to differential equations. It is a functions spaces that accommodate functions whose derivatives existed in some generalized sense, known as “weak derivatives”. The study of this space was motivated by the need to solve boundary value problems in partial differential equations where solutions that are continuously differentiable might not exist.

Locally compact groups form a broad class of topological groups that generalize many familiar structures, such as Euclidean spaces, Lie groups, and discrete groups. A locally compact group is characterized by the property that every point has a compact neighborhood, and the group operations (multiplication and inversion) are continuous. The existence of a Haar measure on these groups, which is a unique (up to a scalar multiple) left-invariant measure, allows for the integration of functions and forms the foundation for harmonic analysis on these groups. This work focuses on basic review of Sobolev spaces in the context of locally compact groups with emphasis on the Sobolev embedding theorem on Locally compact group.

4. Preliminaries.

In this section, basis definitions that are required for proper understanding of Sobolev embedding theorem are considered.

2.1 Topological Groups. Let G be a group equipped with a Hausdorff topology τ the pair (G, τ) is a topological group provided the group operations of addition and multiplication are continuous. Examples include, but not limited to, the set of real numbers equipped with modulus topology and the General linear group.

2.2 Locally Compact group. A topological group is said to be locally compact if every point in G has a compact neighborhood. Furthermore, a topological group G is said to locally compact if;

- If G is a Hausdorff space
- Every point in G has a compact neighborhood
- The group operations of multiplication and inversion is continuous.

From the above stated definitions, this researcher defines a locally compact group as a topological group that satisfies the following listed conditions;

- The group is locally compact at its identity element e , that is, there exist a neighborhood of e that is compact.
- The group operations and inverses are continuous.
- The group is a Hausdorff topological space, that is, $\forall x, y \in G$, then $\exists \mathcal{U}_x \in \mathcal{U}(x)$ and $\mathcal{U}_y \in \mathcal{U}(y)$ such that $\mathcal{U}_x \cap \mathcal{U}_y = \emptyset$.

In essence, a topological space is said to be locally compact at a point say x if, x has compact neighborhood.

A group G is said to be locally compact if every point in G is locally compact. Examples of locally compact groups include discrete groups, Lie groups, the additive groups of rational numbers, the groups of real numbers etc.

2.3 Topological Vector Space. A topological vector space (TVS) is a Hausdorff topological space in which the vector operations of addition and multiplication are continuous.

2.4 The space $\mathcal{L}^p(\Omega)$. Let Ω be a domain in \mathbb{R}^n and let p be a positive real number we denote by $\mathcal{L}^p(\Omega)$ as the class of all measurable functions u defined on Ω for which

$$\int_{\Omega} |u(x)|^p dx < \infty.$$

The elements of $\mathcal{L}^p(\Omega)$ are thus an equivalent class of measurable functions satisfying the above integral. Also, $u \in \mathcal{L}^p(\Omega)$ if u satisfies the above integral and $u = 0$ in $\mathcal{L}^p(\Omega)$, if $u(x) = 0$ in Ω .

2.5 Sobolev Space. Let Ω be an open set in \mathbb{R}^n and let p ($1 \leq p \leq \infty$) be a real number. The Sobolev space usually denoted as $\mathcal{W}^{k,p}(\Omega)$ is a space of functions $u \in \mathcal{L}^p(\Omega)$ whose weak derivatives up to order k also belong to $\mathcal{L}^p(\Omega)$ space

$$\mathcal{W}^{k,p}(\Omega) = \{u \in \mathcal{L}^p(\Omega) \mid D^{\alpha}u \in \mathcal{L}^p(\Omega) \text{ for } |\alpha| \leq k\}$$

Where $D^{\alpha}u = \alpha$ - th order partial derivative of u

And where α is the order of the derivative. Sobolev spaces $\mathcal{W}^{k,p}(\Omega)$ consist of functions in $\mathcal{L}^p(\Omega)$ that have weak partial derivative up to order k .

The following are some properties of Sobolev space

- The Sobolev space is a Banach space, this stems from the fact that according to the property of the normed space and by consequence of this property a norm of a Sobolev space $\{\|\cdot\|_{\mathcal{W}^{k,p}(\Omega)}\}$ defines the sum of \mathcal{L}^p - norms of all weak derivatives up to order k .
- The Sobolev space $\mathcal{W}^{k,p}(\Omega)$ is uniformly convex, reflexive and separable.
- The Sobolev space is a complete inner product space that is a Hilbert space.

The concept of Sobolev space naturally extends to functions in \mathbb{R}^n by considering local charts and partitions allowing for the treatment of partial differential equation.

5. THE SOBOLEV IMBEDDING THEOREMS

The imbedding characteristics of Sobolev spaces are essential in their uses in analysis, especially in the study of differential and integral operators. The most important imbedding results for Sobolev spaces are often referred to as a single theorem called "THE SOBOLEV IMBEDDING THEOREM" although they are of several different types and can require different methods of proof. The different types of the Sobolev imbedding theorems depends on the domain, the dimension and the order of the Sobolev space involved. In this section, we shall present the different kinds of imbedding theorem on groups without proof.

3.1 SOBOLEV IMBEDDINGS INTO LEBESGUE SPACES. This version of the Sobolev imbedding give us a clue as to how Sobolev spaces $\mathcal{W}^{k,p}(\Omega)$, where k is the order of differentiability and p is the exponent of integrability can be embedded into the Lebesgue space.

Let $\Omega \subset \mathbb{R}^n$ be a domain. If $1 \leq p < n$, then there is a continuous embedding of the Sobolev space $\mathcal{W}^{1,p}(\Omega)$ into spaces $L^q(\Omega)$ for q such that:

$$\frac{1}{q} = \frac{1}{p} - \frac{1}{n}.$$

This means that functions in spaces $\mathcal{W}^{k,p}(\Omega)$ are also in $L^q(\Omega)$ for this q , and there exist a constant C such that:

$$\|u\|_{L^q(\Omega)} \leq C \|u\|_{\mathcal{W}^{1,p}(\Omega)}.$$

It is important to note however that this does not hold true for the case where $p = n$ which is the imbedding into an exponential Orlicz space. One of the major theorems of Sobolev embeddings is the Rellich-Kondrachov compact embedding theorem which was traditionally stated to embody the embedding of Sobolev spaces into the Euclidean space

3.2 THE RELlich-KONDRACHOV THEOREM OF SOBOLEV SPACES ON \mathbb{R}^n . Let $\Omega \subset \mathbb{R}^n$ be a bounded, open domain with a smooth boundary and let $1 \leq p < \infty$. Suppose $m > 0$ is an integer. Then the Sobolev space $\mathcal{W}^{m,p}(\Omega)$ is compactly embedded into $L^q(\Omega)$ for any $1 \leq q < p^*$, where

$$p^* = \frac{np}{n-p} \text{ and } p < \frac{n}{m},$$

That is,

$$\mathcal{W}^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$$

NOTE: " \hookrightarrow " denotes continuous embedding and " $\hookrightarrow\hookrightarrow$ " denotes compact embedding.

The Rellich-Kondrachov compact imbedding theorem is a recognized result in the theory of Sobolev spaces for Euclidean domains, however this theory extends to a more general setting such as the locally compact groups. As earlier determined in the study of existing literature, Sobolev spaces can be defined on locally compact groups equipped with Haar measure, which allows generalization of various classical Sobolev theory results. This result is stated as follows.

3.4 The Rellich-Kondrachov theorem on locally compact groups. Let G be a locally compact group that possesses a left-invariant Haar measure μ . Assume that:

- $K \subset G$ is a compact subset,
- $p \in [1, \infty)$, $q \in [1, \infty)$, and m is a non-negative integer,
- $\mathcal{W}^{m,p}(G)$ denotes the Sobolev space of functions on G , defined using left-invariant differential operators.

If $m > 0$, then for certain p and q , depending on the group dimension, there exist a compact embedding;

$$\mathcal{W}^{m,p}(G) \hookrightarrow L^q(G)$$

That is, any bounded sequence in $\mathcal{W}^{m,p}(G)$ has a subsequence that converges strongly in $L^q(G)$, where the condition of m, p and q depend on the dimension of the group and the regularity of the Sobolev space.

To make sense of the above theorem, it is understood that Sobolev spaces $\mathcal{W}^{m,p}(G)$ on locally compact groups are defined similarly to the Euclidean setting as earlier stated in the study and given a left invariant differential operator D^α , we define Sobolev norm as:

$$\|u\|_{\mathcal{W}^{m,p}(G)} = \sum_{|\alpha| \leq m} \|D^\alpha u\|_{L^p(G)},$$

Where α denotes a multi-index corresponding to a differential operator and $u \in \mathcal{W}^{m,p}(G)$ if u and its derivative up to order m are in $L^p(G)$.

REFERENCES

- A. Deitmar and S. Echterhoff, Principles of Harmonic Analysis, Universitext, Springer, New York, 2009.
- Adams, R.A and Fournier, J.J.F (2003), Sobolev spaces. Academic press.
- E. Hebey. (1996).Sobolev Spaces on Riemannian Manifolds, Lecture Notes in Math. 1635, Springer, Berlin.
- Gorka, Kostrzewa, Reyes (2015). Sobolev spaces on locally compact abelian groups and the string equation. Austria
- Joe, D. Spalsbury, A. (2013) The Joys Of Haar Measure. American mathematical society, USA.
- Kreyszrg, Erwin. (1978). Introductory functional analysis with application. John Wiley and sons., New York.
- R. G. Bartle, The Elements of Integration and Lebesgue Measure, Wiley Class. Library, John Wiley & Sons, New York, 1995.
- Rudin, Walter. (1991). Functional analysis, Mc Graw-Hill.

IMPROVING THE PERFORMANCE OF HYBRIDIZED RANDOM FOREST AND LONG SHORT TERM MEMORY ALGORITHMS USING SMOTE, FOR INTRUSION DETECTION.

Moses Adah AGANA¹, Emmanuel EDIM² and Ahena Isaac BASSEY³

^{1,2,3}*Department of Computer Science, University of Calabar, Calabar, Nigeria.*

Emails: ¹ganamos999@yahoo.com; ²mosesagana@unical.edu.ng; ³cubicbasyy@yahoo.com

Abstract

Most people and businesses in recent years have embraced digitalization. Most of the real-world transactions like shopping, communications, banking and education are made available via the internet. This noticeable increase in patronage enjoyed by innovation brings with it, an increase in vulnerabilities in the digital ecosystem. There are existing security measures to ensure data is safe, but this conventional security methods like encryption, use of passwords, firewalls and use of antivirus only provides first level defense. Also, most existing Intrusion Detection Systems (IDSs) struggle to detect novel attacks since present day attackers keep novelty in their tools and techniques in exploiting any kind of vulnerability. The efficiency of an IDS is determined by the learning model selected and the quality of the dataset used. Imbalanced datasets present a problem during data classification, the minority attack classes are often not properly captured due to their small size. To increase the quality of the dataset, the Synthetic Minority Oversampling Technique (SMOTE) was deployed to tackle imbalanced datasets by generating synthetic samples for the minority classes like U2R and R2L in the NSL-KDD dataset. Random Forest and Long-Short Term Memory (LSTM) models were trained independently, the trained dataset was oversampled using SMOTE, then predictions from both models were combined using the late fusion approach which in this case was a simple averaging. Result shows that the RF and LSTM individually performed less but when the resampled hybrid model was tested, it showed a very high predictive accuracy of about 94% with the minority class attacks duly captured. Due to constraint in resources, the hybrid model was deployed in a virtual network environment and its performance was evaluated using metrics like accuracy, precision, false alarm rate, detection rate, recall and F1- score. It was observed that the hybrid model outperformed every other model selected for the comparison. Therefore, it can be concluded that the research work has brought some glimmer of hope with the assurance that major network security issues could be resolved seamlessly with the minority attack classes equally captured.

Keywords: intrusion, detection, security, oversampling, accuracy, hybrid, network, dataset, attacks, model.

1. Introduction

The world is characterized with constant change and improvement. People and businesses have increasingly embraced technological developments, and most of the real-world transactions like shopping, banking and online examinations have been made available in the cyber world. However, this increase in patronage also increases the rate of vulnerability in the digital ecosystem. As security experts come up with possible remedies in trying to provide defense, the

intruders are also tirelessly improving upon their skills to beat the experts in their game. Therefore, there is urgency in addressing network security now, more than ever.

In recent years, Intrusion Detection Systems (IDSs) have emerged to improve the security of network infrastructure since traditional and conventional security methods like firewalls, encryption, password and use of antivirus only provides first level of defense and cannot completely protect the networks and systems from progressive attacks and malware [10].

IDSs can be categorized into three groups: Signature based Intrusion Detection Systems (SIDS), Anomaly-based Intrusion Detection Systems (AIDS), Hybrid Systems. SIDSs store the signatures of the malicious activities in a knowledge base and try to detect intrusions by using pattern matching techniques. AIDSs try to learn the normal behaviors of the system and classify the others as suspicious. It can identify zero-day attacks that have not been encountered previously. The integration of SIDS and AIDS to increase the detection rate of known malicious activities by reducing the false positive rate of zero-day attacks constitutes the hybrid systems[4].

The machine learning model used in training the IDS and the quality of the dataset used, directly determines the efficiency of the IDS. The NSL-KDD dataset was used in evaluating the hybrid model. However, because of the imbalance nature of the dataset, assessing User to Root (U2R), Remote to Local (R2L), Probing, and Denial-of-Service (DoS) attacks in networks simultaneously has been a major challenge in intrusion detection. Re-sampling technique was adopted to deal with these challenges and increase the detection rate of the low-sampled attack types.

The efficiency of an IDS is determined by the learning model selected and the quality of the dataset used. A good quality dataset can be defined as a dataset that improves better performance metrics in real-world transactions [5]. [1]also explained that imbalanced datasets present a problem to researchers. According to [8], a dataset is said to be imbalanced when the distribution of classes is not uniform. This is a common problem in many of the dataclassification due to the used datasets. Imbalanced dataset results the used classifier biases towards the majority class, however, in most of them, the aim is trying to detect the minority class [2]. This results in a large classification error over the minority class samples and main targets can be missed. To increase the quality of the dataset, it should be balanced according to data types.

SMOTE is an oversampling technique specifically designed to tackle imbalanced datasets by generating synthetic samples for the minority classes like U2R and R2L.

Machine Learning methods have shown excellence to achieve high detection accuracy. Still, there are some limitations of Machine Learning methods like handling raw, unlabeled or high dimensional data [6].

Research has identified Random Forest (RF) as an ensemble classifier that is best suited for classifying the intrusion attacks with higher accuracy in intrusion detection [3]. However, Deep Learning methods are known for their abilities to handle labelled or unlabeled data or solve complex problems with the help of the high-powered Graphic Processing Unit.

Despite the suitability of Random Forest as ensemble classifier for intrusion detection, and its effectiveness in solving massive intrusion data classification problems that arise within real network application environments. It is an established fact that deep learning methods have the potential to extract better representation from the data to create much better models for intrusion

detection[11].Therefore, a hybrid of RF and LSTM using SMOTE promises a very high predictive accuracy and precision.

2. Design Methodology

The machine learning development life cycle approach was adopted in designing the hybrid model. And the dataset used in evaluating the performance of the hybrid algorithms was the Neural Simulation Language- Knowledge Discovery Dataset (NSL-KDD). The following steps were taken in developing the model:

- i. Data Collection: Data was extracted from the NSL-KDD dataset.
- ii. Data Preprocessing: At this phase the dataset was processed by; Loading the dataset, exploring its structure, handling categorical features and preparing it for machine learning tasks.
- iii. Feature Extraction: The Principal Component Analysis (PCA) is a common feature extraction technique adopted in this research study, and it is available in the python Scikit library.
- iv. Implementation Stage: Ensemble methods which often enhances model performance by leveraging the strengths of multiple models was used in implementing the hybrid SMOTE-RF-LSTM model.
 - i. Evaluation Stage: Random Forest and LSTM models were trained independently, the trained dataset was oversampled, then predictions from both models were combined using the late fusion approach which in this case was a simple averaging. The final output is the predicted intrusion class. These approaches leverage the strengths of both models.

3. System Design

In the use case diagram shown in figure 1, the only action the user performs is to send network request to the server which is then intercepted by the machine learning server running the

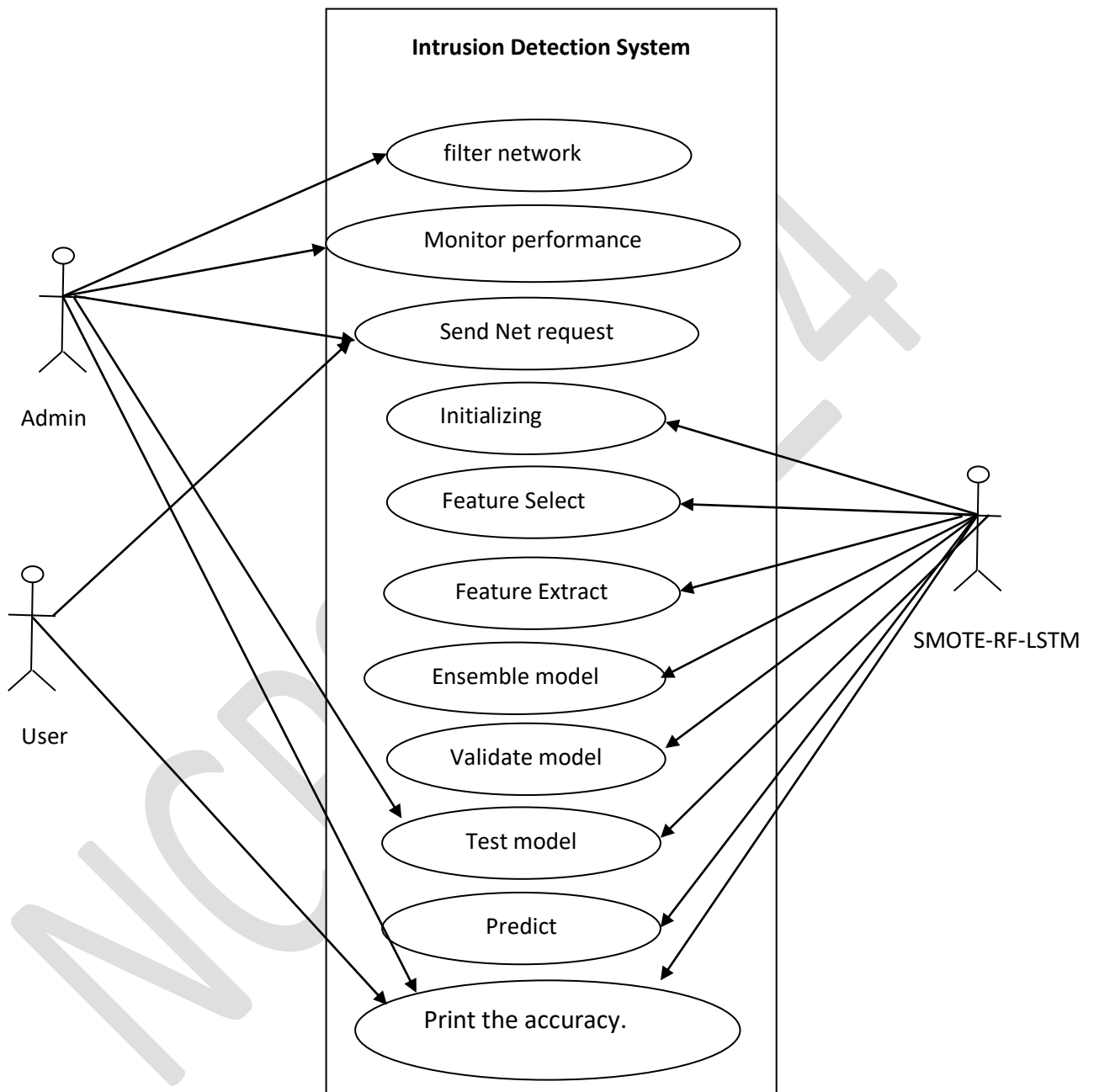


FIG 1: Use case of the hybrid system.

combined SMOTE-RF and LSTM models in an ensembled form to determine whether it is a valid request. The admin monitors the users to see where malicious request originates and can then decide to filter such a user.

3.1 Mathematical model for hybridized SMOTE- RF and LSTM algorithms

This model combines a Random Forest (RF) and Long Short-Term Memory (LSTM) network for network intrusion detection. Here's a breakdown of the mathematical representation:

1. Data Preprocessing: Let X be a network traffic dataset with n data points (samples) and d features. Each data point is represented as:

$$X_i = (x_{i1}, x_{i2}, \dots, x_{id}) \in R^d$$

where:

$i = 1, 2, \dots, n$ (index of data point)

x_{ij} = j th feature value in the i th data point

R^d = d -dimensional real number space

Hybrid Model

Late Fusion:

- Run the RF model ($Y_{RF}(X_i)$) and LSTM network (h_t) independently on the same data (X_i).

- Combine the outputs from both models using a function (ϕ) to produce the final prediction (Y_{hybrid}).

$$Y_{hybrid}(X_i) = \phi(Y_{RF}(X_i), h_t)$$

$$\phi(Y_{RF}(X_i), h_t) = w_{RF} * Y_{RF}(X_i) + w_{LSTM} * h_t$$

where ϕ is a simple weighting average trained to combine the strengths of both RF and LSTM outputs.

Training: The model is trained on labeled network traffic data. The labels indicate normal or intrusion traffic. There was an oversampling of the trained dataset, and the training process optimizes the parameters of both the RF and LSTM.

3.2 Model of the hybrid system

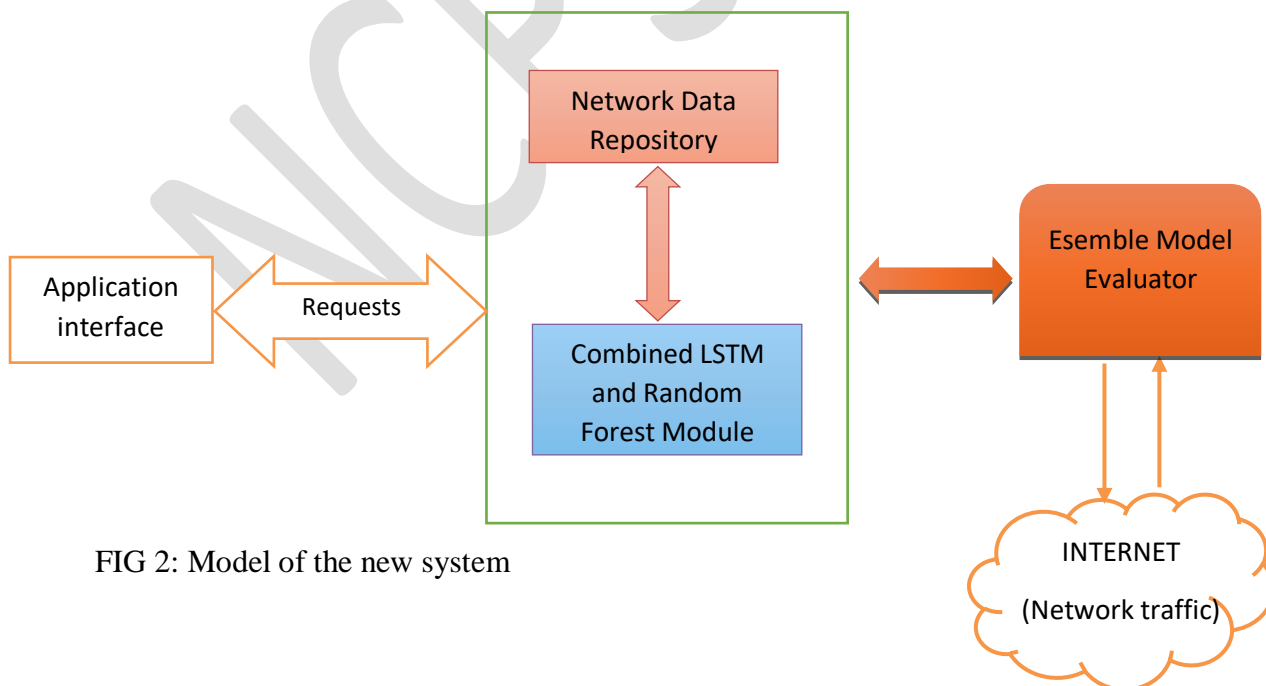


FIG 2: Model of the new system

The components of the system illustrated in figure 2 are:

Application Interface: The application interface is any GUI app that can send network requests to the Machine learning service. A good and easy example of this GUI can be a Web Browser.

Network Data Repository: The network data repository is a database of network data stored for the machine learning models to constantly learn from.

Random Forest and LSTM Module: The combined Random Forest and LSTM module houses the enhanced machine algorithm. This module is responsible for the sub - learning process of the Ensemble module.

Ensemble Evaluator: The Ensemble evaluator is the ensemble model generated from the ensemble stacking of Random Forest and LSTM. This evaluator is what produces the final result of predictions.

4. Results

Table 1 and figure 3 respectively, shows the overall attack volume across different protocols. It reveals which protocol is more susceptible to certain attack type. As seen on table 1, the TCP protocol was very susceptible to connection and communication-oriented attacks emanating from portsweep and satan. The UDP being connectionless protocol was targeted by simple Denial of Service attacks (like teardrop and nmap diagnostic tool) due to its stateless nature. The ICMP used for control message was targeted by attacks that exploit its functionality for information gathering. For example, the ipsweep attack also known as a subnet scan, is a network reconnaissance technique that targets the Internet Control Message Protocol (ICMP). It is used by attackers to identify active devices on a network subnet.

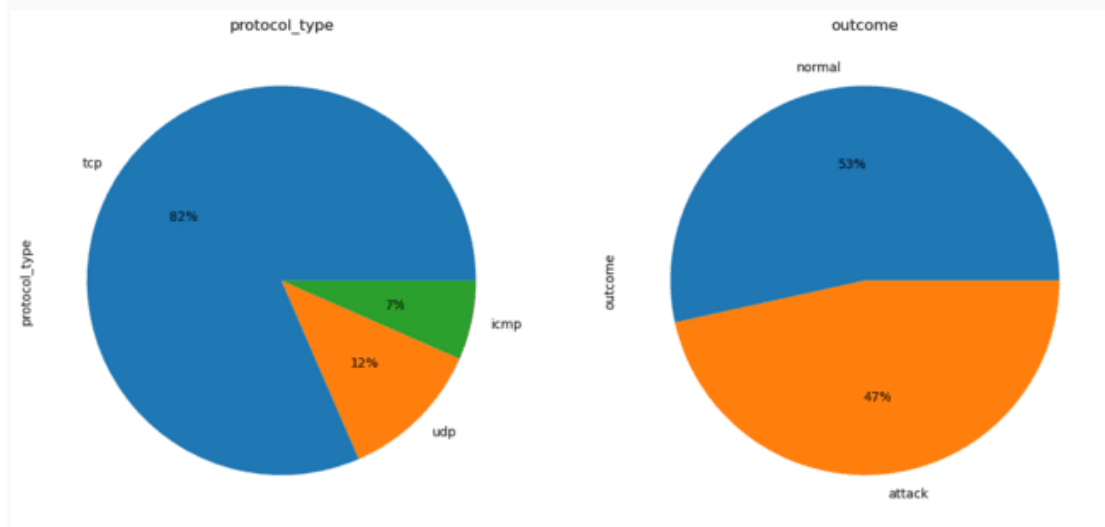


FIG. 3: Attack distribution by protocol.

TABLE 1

Protocol- types specific attack distribution

protocol_type	icmp	tcp	udp
attack			
back	0	956	0
buffer_overflow	0	30	0
ftp_write	0	8	0
guess_passwd	0	53	0
imap	0	11	0
ipsweep	3117	482	0
land	0	18	0
loadmodule	0	9	0
multihop	0	7	0
neptune	0	41214	0
nmap	981	265	247
normal	1309	53599	12434
perl	0	3	0
phf	0	4	0
pod	201	0	0
portsweep	5	2926	0
rootkit	0	7	3
satan	32	2184	1417
smurf	2646	0	0
spy	0	2	0
teardrop	0	0	892
warezclient	0	890	0
warezmaster	0	20	0

4.1 Epoch

Figure 4 shows the different epochs through the entire training dataset. The model learns and adjusts its internal parameters to improve its ability to distinguish between normal and malicious network traffic. Accuracy represents the percentage of data points where the model correctly classifies network traffic as normal or belonging to a specific attack type (e.g Dos, U2R, e.t c). higher accuracy indicates better detection capability. The result shows that accuracy increases over time as the epoch increases. This indicates better detection capability of the hybrid model.

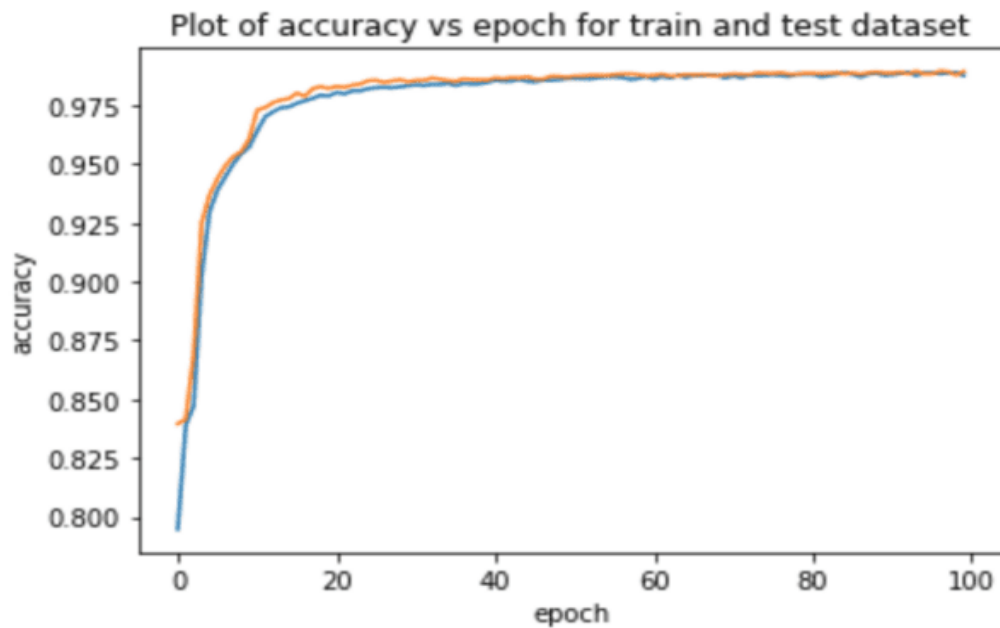


FIG 4: Plot of accuracy vs epoch

TABLE 2
Confusion Matrix

	Predicted class	
	Attack	Normal
Actual Class	Attack	Normal
	True Positive	False Negative
	False Positive	True Negative

Interpreting the Matrix:

TN (True Negative): Correctly classified normal traffic instances.

TP (True Positive): Correctly classified intrusion instances (e.g., DoS, U2R, R2L).

FP (False Positive): Normal traffic instances incorrectly classified as intrusions.

FN (False Negative): Intrusion instances incorrectly classified as normal traffic.

Here is a breakdown of a confusion matrix in this context:

- Rows represent the actual intrusion classes present in the test data.
- Columns represent the intrusion classes predicted by the model.

Each cell (i, j) of the matrix contains the number of instances.

Where:

- i represents the actual intrusion class.
- j represents the predicted intrusion class by the model.

Predicted Class	Normal	DoS	Probe	U2R	R2L
Normal	TN	FP	FN	FN	FN
DoS	FN	TP	FN	FN	FN
Probe	FN	TN	FN	TP	FN
U2R	FN	FN	TP	FN	FN
R2L	FN	FN	FN	TP	FN

Performance indicators:

True Positive Rate |TP|= $TP / (TP + FN)$

• False Positive Rate |FP|= $FP / (FP + TN)$

• True Negative Rate |TN|= $TN / (TN + FP)$

- False Negative Rate $|FN| = FN / (FN + TP)$
- Accuracy $|ACC| = (TP + TN) / (TP + TN + FP + FN)$
- Detection Rate $|DR| = (TP) / (TP + FN)$
- False Alarm Rate $|FAR| = (FP) / (FP + TN)$
- Precision $= TP / (TP + FP)$
- Recall $= TP / (TP + FN)$
- F1- score $= 2 PR / P + R$ where (P = precision and R = recall)

By analyzing these metrics, it was shown how well the hybrid SMOTE- RF-LSTM model performed in identifying different intrusion types and balancing false positives and negatives. The confusion matrix shows that the RF and LSTM individually performed less as compared to performance of the combined SMOTE-RF-LSTM model as shown in figure 5. Table 3 and Figure 6 also shows a comparison of the accuracy scores for all ML and DL considered in this research work. From the table and bar chart it is also clearly seen that the hybrid SMOTE- RF-LSTM model outperformed every other model.

The performances in summary are:

LSTM: 89%

Random Forest: 84%

SMOTE- RF-LSTM Model: 94%



FIG 5 Performance accuracy for SMOTE-RF-LSTM

TABLE 3

Comparison of the SMOTE-RF- LSTM model and other models

Model	Precision	Recall	F1 score	Accuracy score
SVM	0.7476	0.7452	75	0.7497
CNN	0.8083	0.8067	81	0.8186
K- Nearest Neighbor	0.8019	0.8014	80	0.8056
Logistic Regression	0.8494	0.8436	84	0.8463
CNN- LSTM	0.8785	0.8732	87	0.8741
KNN- SVM	0.8445	0.8401	84	0.8476
RF	0.8486	0.8458	84	0.8403
LSTM	0.8993	0.8967	89	0.8932
RF- LSTM	0.9079	0.9095	90	0.9037
SMOTE-RF-LSTM	0.9445	0.9471	94	0.9414

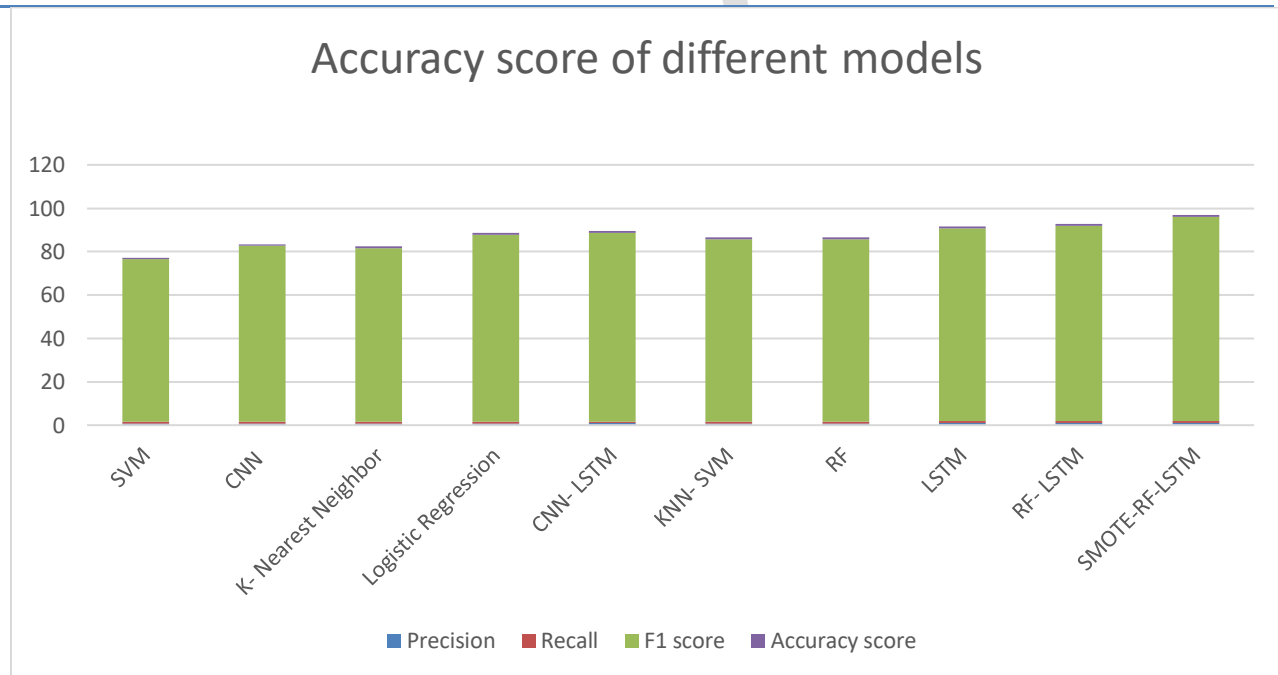


FIG 6: Bar chart representing the accuracy scores.

5. Conclusion

In view of the said security challenges, this research work has been carefully undertaken to nip the menace in the board by providing an up-to-date intrusion detection model using a hybrid of SMOTE, Random Forest (RF) and Long Short-Term Memory (LSTM) algorithms. The trained RF and LSTM are ensembled using the stacking technique which combines the predictions of the two models to get a final prediction with better performance, then the predictive accuracy of the ensembled model is evaluated on the test set. The minority attack classes were oversampled using the SMOTE technique. The ensembled method improved the overall performance of the system by leveraging the strengths of both models. Finally, the performance of the new hybrid SMOTE-RF-LSTM model was evaluated using metrics like accuracy, precision, false alarm rate, detection rate, recall and F1- score. It was observed that the hybrid model outperformed every other model selected for the comparison, with an accuracy of 94%.

Therefore, it can be concluded that the research work has brought some glimmer of hope with the assurance that major network security issues could be resolved seamlessly.

REFERENCES

- [1] Ali, A. Shamsuddin, S. M. &. Ralescu, A. L (2015). Classification with class imbalance problem. *A review, '' Int. J. Adv. Soft Comput. Appl., vol. 7, no. 3, pp. 176–204.*
- [2] Barua, S., Islam, M. M., Yao, X.& Murase, K. (2001). MWMOTE-majority weighted minority oversampling technique for imbalanced data set learning. *IEEE Trans. Knowledge. Data Eng., vol. 26, no. 2, pp. 405–425.*
- [3] Farnaaz, N. & Jabbar M. (2016). Random Forest modeling for network intrusion detection system. *Procedia computer Science, 8, 161-170.*
- [4] Gozde, K. Onder, D., & Ozgur K. S. (2020).Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset. *Digital Object Identifier: 10.1109/ACCESS.2020.2973219*
- [5] Johnson J. M. &. Khoshgoftaar, T. M (2019). Survey on deep learning with class imbalance *Journal of Big Data, vol. 6, no. 1, p. 27.*
- [6] Nguyen, T. & Reddi, V. J. (2019). Deep reinforcement learning for cyber security. *Archive preprint, 1906-5799.*
- [7] Payares, E. D. & Martinez-Santos, J. C. (2021). Quantum machine learning for intrusion detection of distributed denial of service attacks: a comparative overview. *Proc. SPIE 11699, Quantum Computing, Communication, and Simulation.Digital Object Identifier: 10.1117/12.259329*
- [8] Provost, F. (2000). Machine learning from imbalanced data sets 101. *In Proc. AAAI Workshop Imbalanced Data Sets, Menlo Park, CA, USA: AAAI Press.*
- [9] Qiang, S. Linlin, F. & Song, G. (2020). A Hybrid Method for Traffic Incident Detection Using Random Forest-Recursive Feature Elimination and Long Short-Term Memory Network with Bayesian Optimization Algorithm.*Digital Object Identifier: 10.1109/ACCESS.2020.3047340*
- [10] Srinivas, J. Das, A. K. & Kumar, N. (2019) Government regulations in cyber security: framework, standards and recommendations. *Future Generation Computer System, 92,178–188*

- [11] Yin, C. Zhu, Y. Fei, J. & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 5, 21954–21961.
- [12] Zhewei, C., WENWEN, Y. & LINYUE, Z. (2021). Proceedings of the 2021 4th International Conference on Signal Processing and Machine Learning. Pg 152-159. *Digital Object Identifier: 10.1145/3483207.3483232*.

NCPS-2024

Virtual Communication using Centralized Geolocation Network Data

Ntun Nkwam Eku, Ogar Akima Akimaand O. A. Ofem

Department of Computer Science, University of Calabar, Calabar, Nigeria

Emails: ntunnkwam@gmail.com; akimaogar@gmail.com

Abstract

In a rapidly evolving landscape of virtual communication, the efficient utilization of resources is paramount in delivering a seamless and cost-effective experience. This project introduces a novel and incremental methodology that harnesses the power of geolocation network data integration and Virtual Private Network (VPN) compression techniques. The objective is to optimize virtual communication, addressing the challenges of data costs, while simultaneously enhancing overall communication efficiency and user experience. The integration of geolocation network data stands as a pivotal aspect of the approach. Geolocation information provides valuable insights into the physical location of users, enabling more intelligent data routing and delivery mechanisms. By incorporating this geospatial data into virtual communication platforms, the project has redefined the efficiency of data transfer, paving the way for a more streamlined and responsive communication experience. The analysis of various integration models explores the potential advantages in terms of reduced latency, improved reliability, and increased overall system performance. This project adopted incremental methodology, Virtual Private Network (VPN) and Geolocation techniques were utilized using *an Open-Source Free Cross-platform Multi-protocol VPN Program, SoftEther (Software Ethernet)*. The result shows that by integrating geolocation network data and VPN compression techniques, the approach not only reduces data costs but also significantly enhances communication efficiency and user experience. As Organizations navigate the complex terrain of virtual communication, this incremental methodology will serve as a roadmap for achieving the delicate balance between cost-effectiveness and optimal performance in the digital era.

Keywords: Virtual Communication, Geolocation, Network Data, Communication Channel/Media, Network Environment, Teleconferencing/Videoconferencing, Virtual Private Network (VPN)

1. Introduction

Communication in a virtual environment encompasses both vertical and horizontal dimensions, emphasizing the significance of interaction. The ability to engage in real-time interactions is a distinct advantage of virtual communication, differentiating it from traditional communication technologies utilized by organizations and individuals across various geographical locations. The challenges associated with broadcasting information across different locales are mitigated by the technological framework of virtual communication, which stimulates responses to visual and

audiovisual stimuli, facilitates dialogue and discussions, encourages idea sharing, fosters active participation, and supports collaborative projects.

Despite the availability of numerous communication tools designed for collaboration in virtual teams, research on their overall global effectiveness remains limited. This gap presents a new challenge for organizations, necessitating the application of effective and efficient virtual communication strategies to meet evolving business demands. The pervasive nature of the internet and the rapid advancement of virtual communication networks have become increasingly important to users and organizations alike. As the scale of virtual communication expands, so too does user engagement and the associated costs of network data, standards, service methods, and protocols, leading to a substantial increase in the volume of data transmitted across networks.

Drexhage et al., (2016) describe virtual technology as a computer-simulated environment comprising software representations of real agents, objects, and processes, complemented by human-computer interfaces for interaction with these models. In contrast, George, (2014) asserts that virtual environments can significantly enhance collaborative technologies, concurrent engineering, and collaborative product design, all of which are inherently iterative. He posits that virtual environments facilitate interactive activities within distributed development contexts, accelerating communication by overcoming time and space constraints. Consequently, communication has evolved into a computer-mediated form that includes videoconferencing, which can be regarded as a modern interpretation of traditional face-to-face communication, leading to more immersive virtual interfaces.

The term "virtual" signifies the capacity to influence something without being the actual entity itself. For instance, the notion of a "Virtual Guy" suggests a figure capable of completing tasks comparable to a physical counterpart. In the realm of information technology, a virtual version of nearly every function exists. However, it is essential to note that virtual relationships often involve geographically separated individuals who interact through digital platforms, presenting themselves as a cohesive unit despite lacking a physical presence(Hyslop, 2022).The transition of service supply from cloud-based to edge computing has introduced new capabilities for meeting application latency requirements while enhancing scalability and energy efficiency. This shift alleviates network traffic burdens, positioning decentralized industrial operations as a viable solution for providing scalable services tailored to delay-tolerant applications. Platform industries have identified two critical technological enablers for advancing virtual communication:

- (1) the communication infrastructure that supports the ubiquitous connectivity of Cyber-Physical Production Systems (CPPS) (Wu, X., et al., 2019)
- (2) data management frameworks designed to facilitate efficient data distribution in future manufacturing environments. (SG Analytics, (2024).

Data communication involves the exchange of information between devices over a transmission medium, requiring both hardware and software components (Innovative Bytes, 2023). The hardware consists of sender and receiver devices, along with intermediate devices facilitating data transfer, while the software establishes the protocols governing communication processes. Effective virtual communication necessitates the establishment of direct connections between mobile devices, where users must know the translated IP address and port number of their intended communication partners. The advent of social networking services (SNS) and smartphones has revolutionized communication, introducing next-generation solutions such as platforms like Facebook, X, Zoom, and Instagram. Although these applications offer users various virtual functionalities, they do not always provide the desired outcomes (Matthew, et al., (2019). The emergence of virtual environments has empowered businesses to access a global talent pool, with knowledge-sharing platforms becoming integral to problem-solving within geographically dispersed organizations.

Technology Acceptance (TA) highlights the factors contributing to positive user attitudes toward information systems, enhancing acceptance and utilization. Key factors include perceived usefulness, ease of use, quality, and organizational support. To bolster user satisfaction and acceptance of technology, many organizations now permit employees to utilize smartphones and other virtual tools to facilitate their work and access corporate databases. Multimedia and virtual reality systems can transform the interaction between individuals and information technology, offering innovative methods for communication and creative expression. Geolocation data provides precise location information for IP addresses, facilitating accurate device tracking within networks. Location-based applications typically operate as server-based systems, maintaining user device information at location servers to allow seamless connectivity and information retrieval. The accuracy of location data, particularly through GPS technology, relies on the availability of satellites in optimal positions, enabling the monitoring of both stationary and moving objects.

2. Methodology

In the dynamic realm of virtual communication, efficiency and cost-effectiveness are paramount. This research work adopted an incremental methodology to enhance virtual communication by integrating geolocation data, VPN compression, and VPN tunneling. The aim is to reduce data costs, strengthen communication security, and optimize user experience. Geolocation data forms the foundation of this approach by refining data routing strategies to minimize latency and improve platform responsiveness. Organizations can progressively adopt this system without disrupting operations, facilitating a seamless transition to more efficient communication.

The integration of VPN compression techniques enhances data security and reduces transmission costs. By applying compression algorithms to Virtual Private Networks, data transfer sizes are minimized, communication speeds are improved, and costs are lowered. This phased deployment allows organizations to align compression levels with their communication needs. VPN tunneling, through protocols such as OpenVPN and IPsec, adds another layer of security, encapsulating data within a secure tunnel. This enhances confidentiality and data integrity, providing a robust virtual communication infrastructure. The incremental deployment strategy prioritizes user adaptation and feedback, ensuring the system meets expectations. Additionally, performance metrics and cost-benefit analyses quantify the economic and security advantages, emphasizing the practicality and scalability of the approach for long-term communication efficiency and security improvements.

3. System Schematic Diagram

Figure 1 illustrates a conceptual data flow model where switch channel technology connects individual systems to a server acting as a gateway. The server monitors network activity, assigning data usage and enhancing fault tolerance.

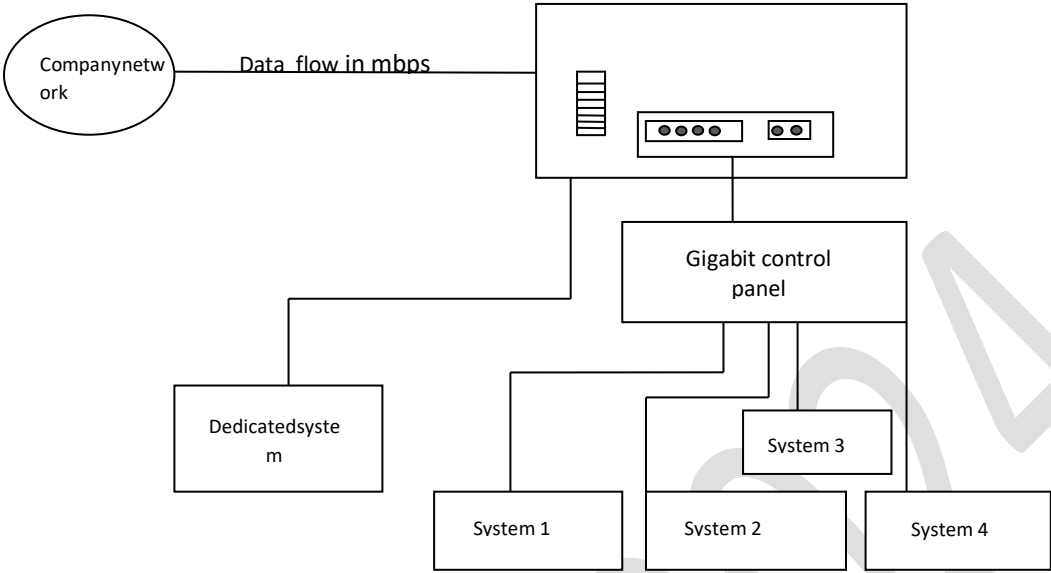


Figure 1: Conceptual data flow model

The system activity flow diagram outlines the sequence of operations. Figures 2 and 3 depict the user access flow. A new peer sends a binding request, and the server responds with an IP, port, and unique peer ID, registering system data.

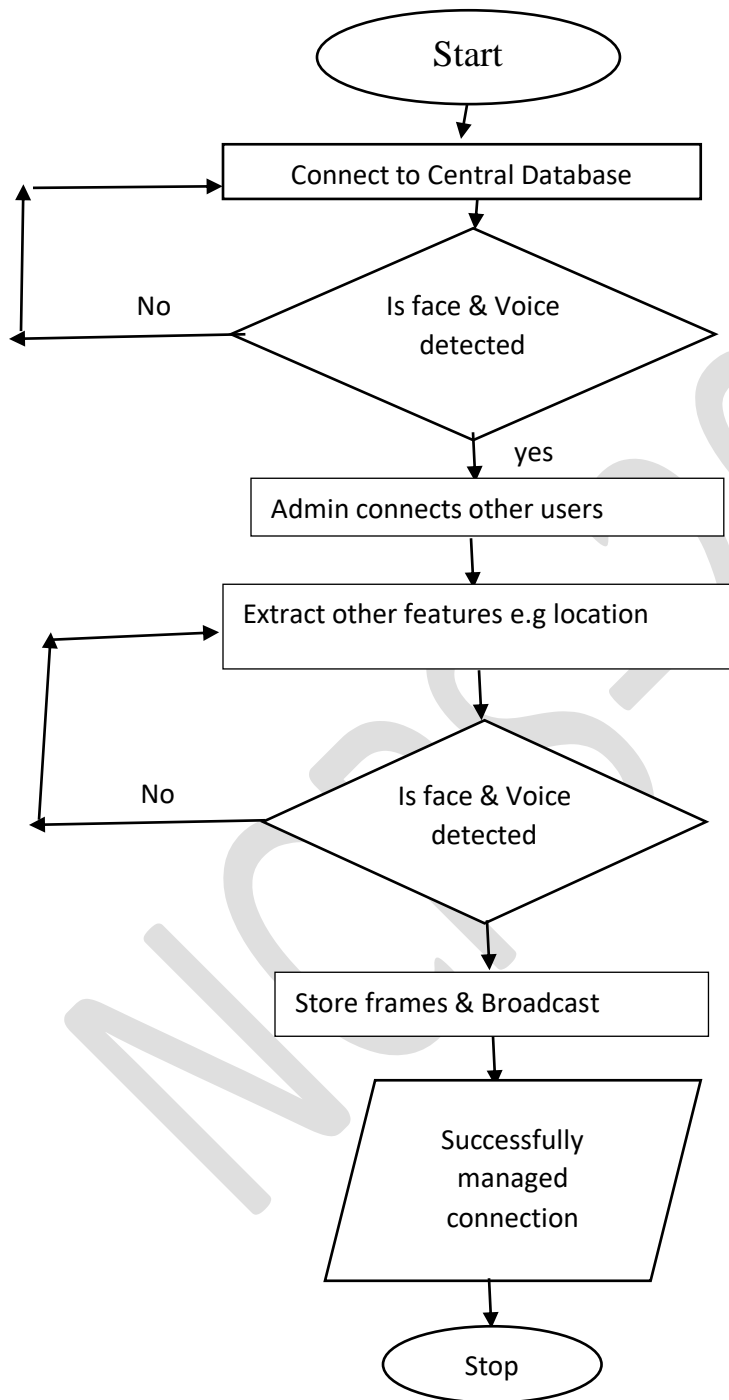


Figure 2: User System Flow Diagram

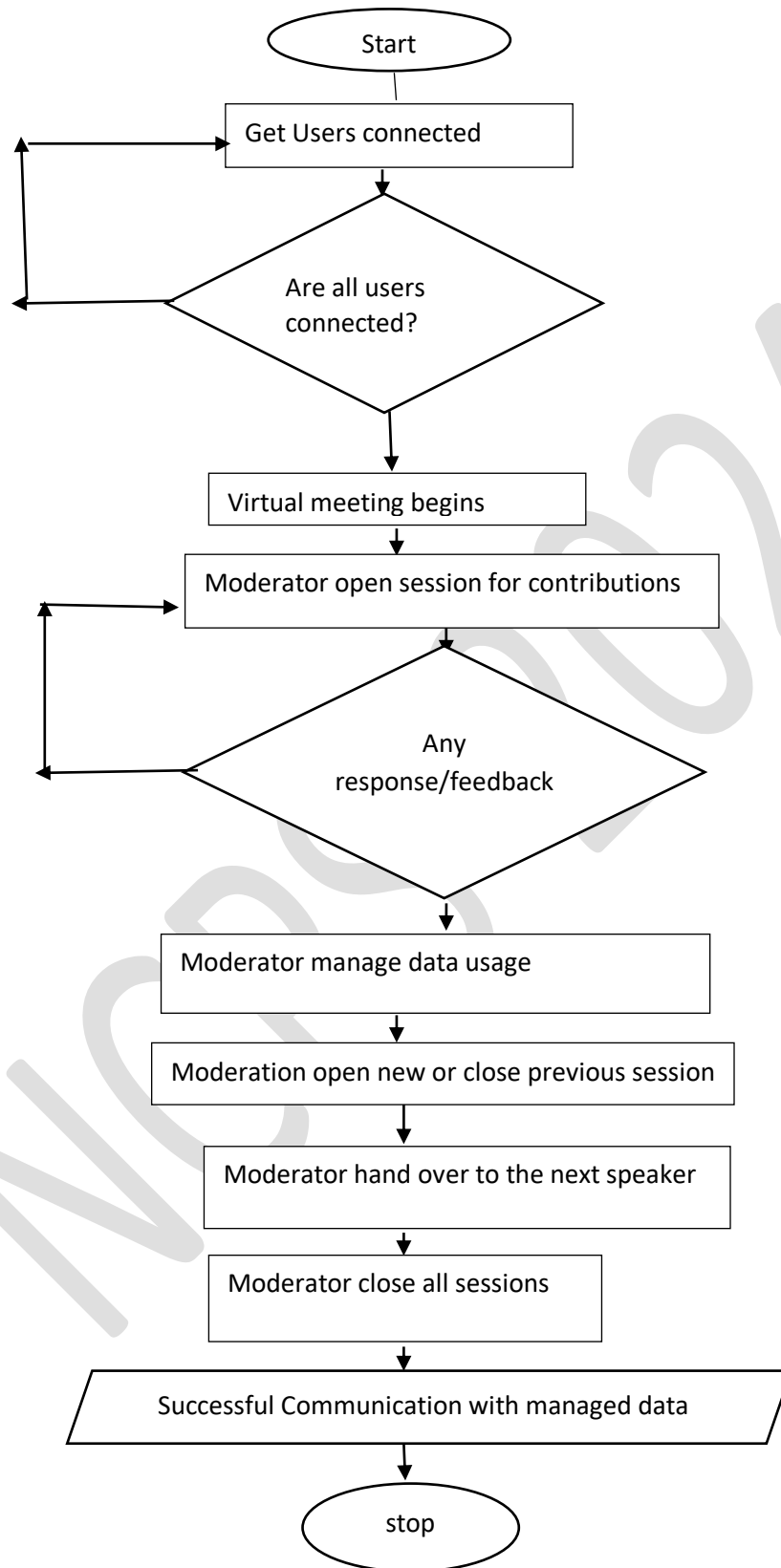
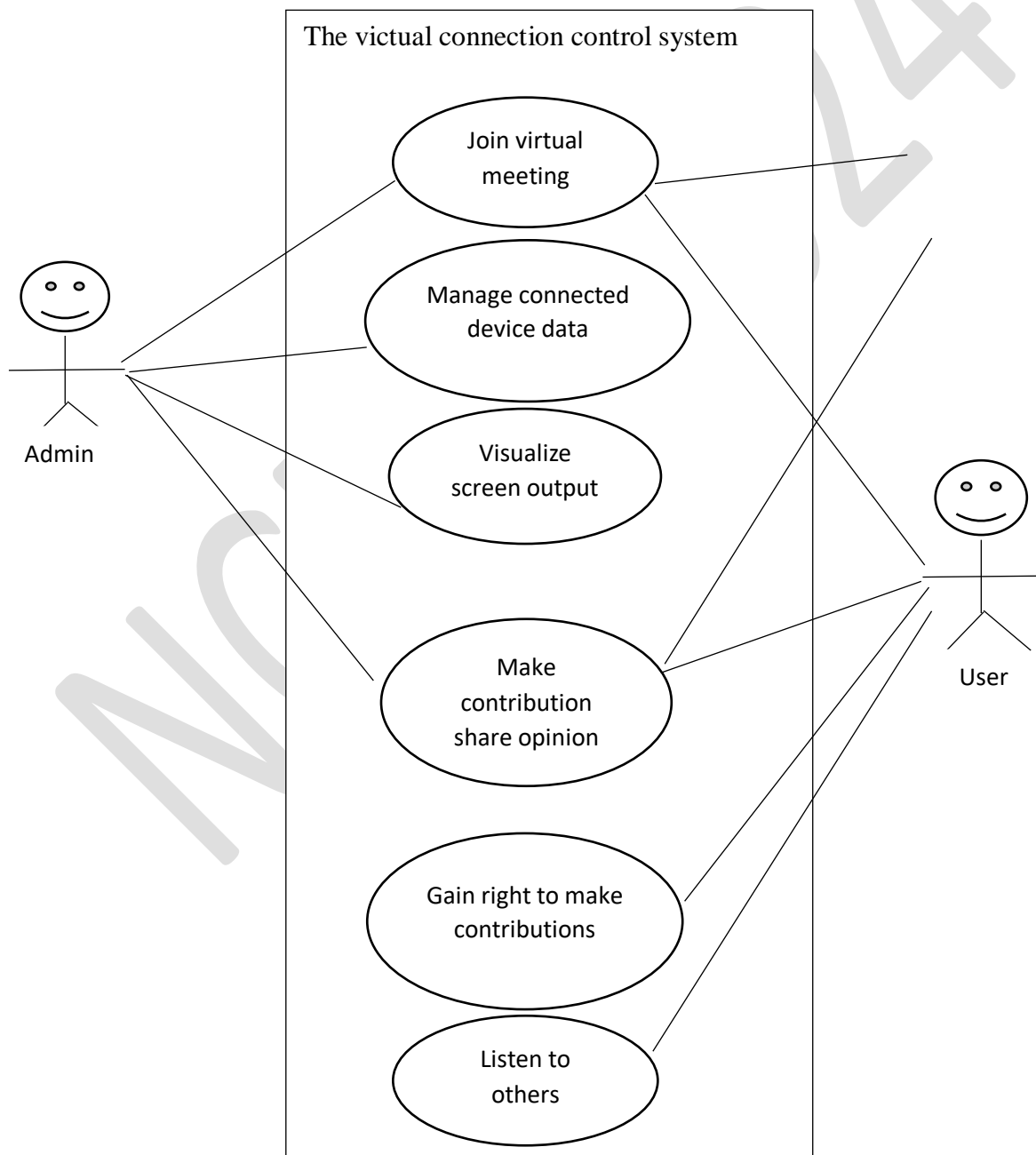


Figure 3: Admin System Activity Flowchart

Figure 4 shows a virtual system interface connected to live audio or video meetings, enabling geographically dispersed participants to communicate and collaborate. This supports voice, video, and audio graphics, facilitating communication with large groups across different locations. Traffic management may affect download speeds during peak hours, regulating network flow and distributing bandwidth equally. Light users are minimally impacted, but heavy users may experience slower speeds during these times.



The developed prototype is tested using qualitative experiments to verify if objectives are met. Users from different locations participate in a virtual process before public deployment, as shown in Figure 5, ensuring functionality prior to broader application use.

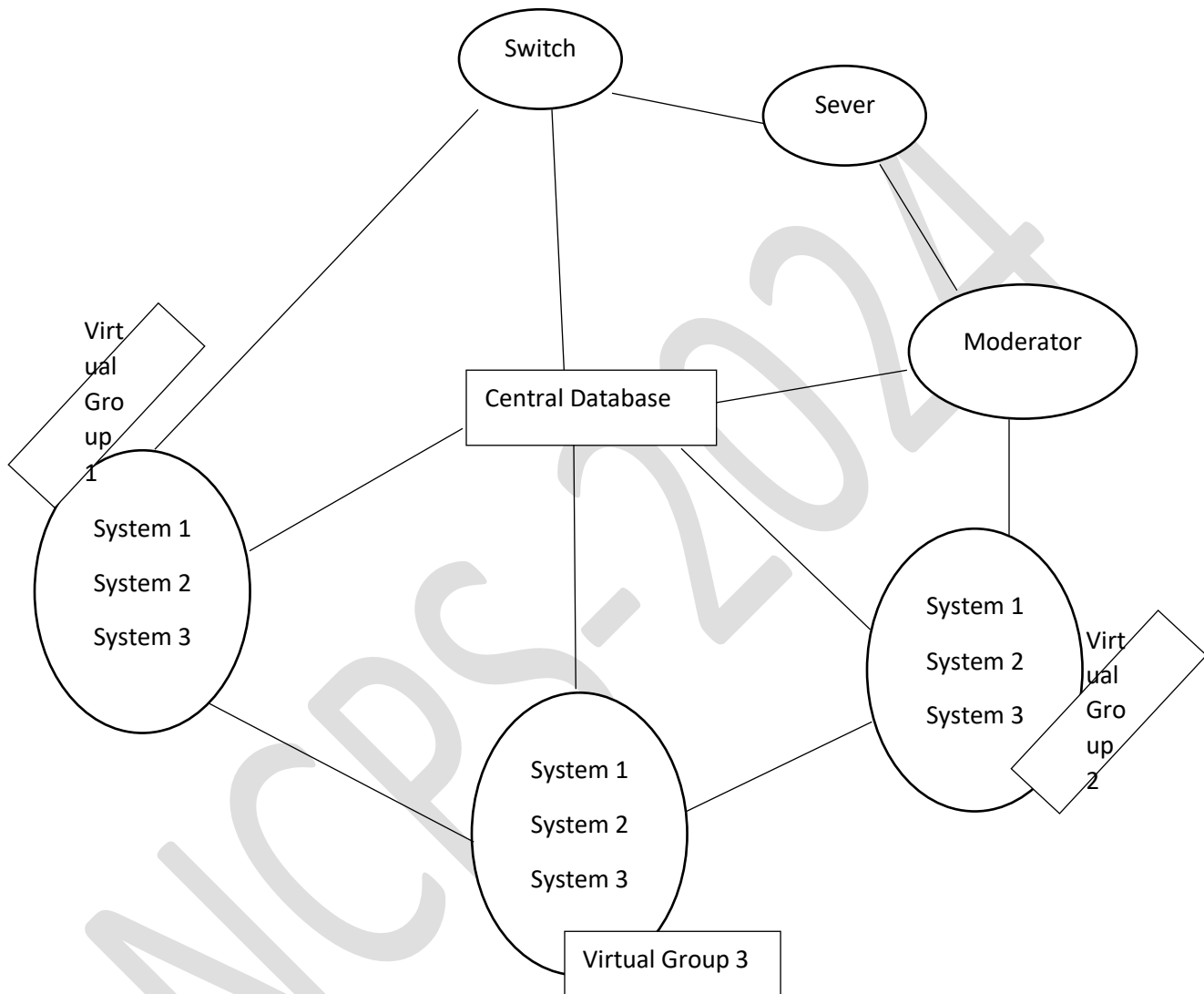


Figure 5: Diagram showing a central connection of systems to communicate

4. Results

The system employs a streamlined architecture featuring a server (admin side) and a client side (users). The server hosts a comprehensive database of connected devices, updated via the internet, ensuring real-time network status retrieval. Security is maintained through an admin username and password. The client side includes a Visual Studio application for admin monitoring and control. Client systems are configured with

the provided IP address, allowing the admin to oversee and manage activities remotely via the app, which requires an internet-connected registered device. Figure 6 shows the server application interface displays user options, status, and essential features, ensuring administrators manage configurations and maintain security effectively.

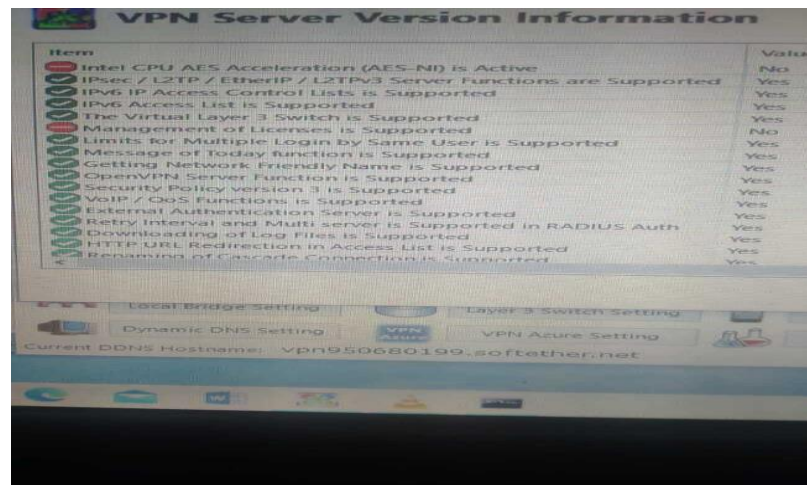


Figure 6: Server information

Figure 7 displays the server version information and connected clients, offering insights into the SoftEther VPN server's scalability and limitations for effective planning and configuration by system administrators.

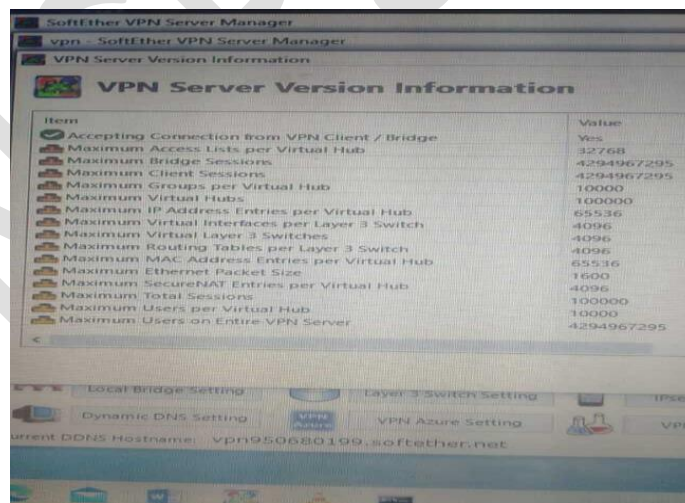


Figure 7: Server version info when actively connected

Figure 8 shows the VPN over ICMP/DNS feature in SoftEther VPN enables administrators to configure settings for establishing VPN tunnels using ICMP and DNS packets, facilitating bypassing of network restrictions and censorship.

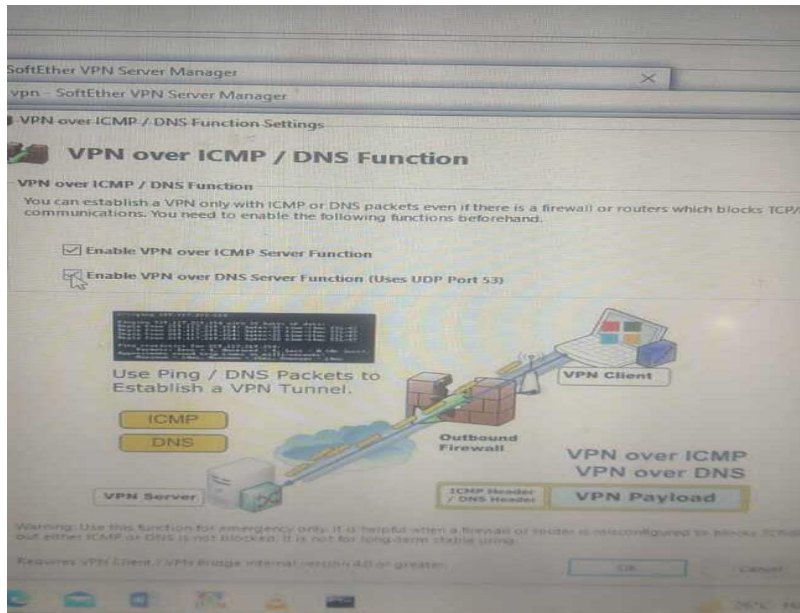


Figure 8: Enabling VPN over ICMP in Server Manager when is online

Figure 9 shows the SoftEther VPN client encompasses optimization aspects such as TCP connections, establishing intervals, encryption, and data compression options for enhanced performance.

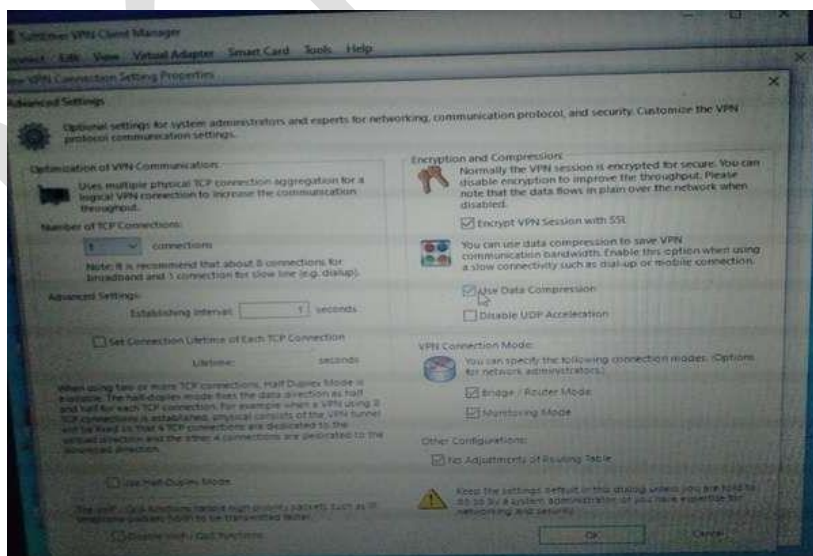


Figure 9: Client Application data type

Figure 10 illustrates the downloaded file before and after compression, demonstrating reduced output data, which consequently lowers bandwidth costs.

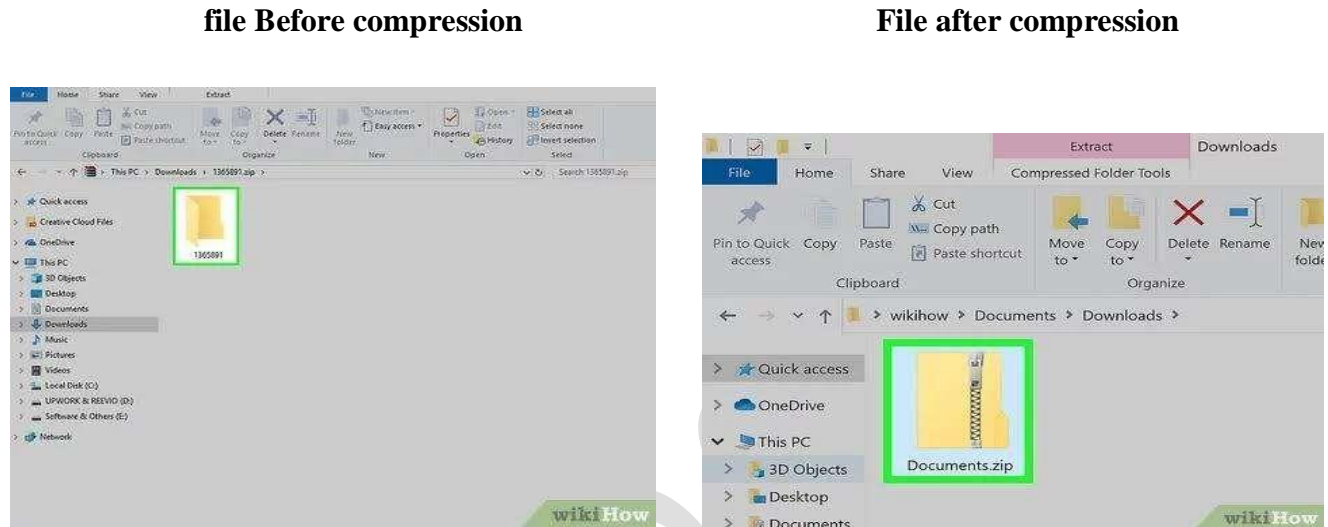


Figure 10: Downloaded file before and after compression

REFERENCES

- Adam, M. (2020). Geolocation Services and Marketing Communication from a Global Point of View; EDP Sciences; SHS Web of Conferences, *Globalization and its Socio-Economic Consequences*. pp 126-134.
- Alshahrani, A. and Al-Muhtadi B. (2020). The Impact of Digital of Digital Transformation on Business Growth. *Journals of Management and Information Technology* 15(2) 12-25
- Anuja A. and Apoorva, K. (2015) Wi-Fi Enabled Personal Computer Network Monitoring System Using Smart Phone with Enhance Security Measures, *Fourth International Conference on Eco-Friendly Computing and Communication system procedural Computer Science* 70:114-122
- Chang, F. M., & Wang, H. L. (2013). The Development of a Mobile Learning System with Provisioning for Location-based Services and Monitoring. *Journal of Science and Engineering Technology*, 9(1), 67-77.
- Chen, L., & Xu, H. (2020). Virtual communication networks using geolocation data: A systematic review. *International Journal of Distributed Sensor Networks*, 16(10), 1-14.
- Chen, C. H., Lin, M. Y. & Liu, C. C., (2018). Edge Computing Gateway of the Industrial Internet of Things Using Multiple Collaborative Microcontrollers", in *IEEE Network*, vol. 32, no. 1, pp. 24–32.
- Drexhage, J., Leiss, D., Torben, S., & Timo, E. (2016). The Connected Classroom –Using Video Conferencing Technology to Enhance Teacher Training. *Reflecting Education*, 10, 70-88
- Gao, W., & Wang, Y. (2019). Centralized geolocation network data for virtual communication: A novel approach. *IEEE Transactions on Network and ServiceManagement*, 16(3), 931-

942

Gao, W., & Wang, Y. (2019). Centralized geolocation network data for virtual communication: A novel approach. *IEEE Transactions on Network and Service Management*, 16(3), 931-942

George S. I (2014) Micro Computing & WWW Nation Open University of Nigeria
2014

Hassan, W. H., & Al-Shammari, B. (2021). Geolocation-based virtual communication networks: Challenges and opportunities. *Journal of Intelligent Information Systems*, 57(2), 297-315.

Hyslop, C., (2022). Virtual Communication: An In-Depth Guide. Retrieved October 19, from
<https://www.glasscubes.com/virtual-communication/>

Innovative Bytes, (2023). Data Communications. Retrieved March 30, from
<https://dhanvina.medium.com/data-communications-8ea102664a17>

Kim, J., & Lee, S. (2018). Virtual communication using centralized geolocation network data: A case study. *International Journal of Human-Computer Interaction*, 34(1), 43-52.

Matthew, N., Philip O., & Sarhan M. (2019). Mobile Social Media. *International Journals of Advanced Research in Computer Science and Software Engineering* ISSN: 2277-128X (Volume-8, Issue-3)

SG Analytics, (2024). Data Management Framework: Importance, Components, and Examples. Retrieved February 08, from <https://www.sganalytics.com/blog/data-management-framework/>

Wu, X., Goepp, V. & Siadat, A., (2019). Cyber Physical Production Systems: A Review of Design and Implementation Approaches. *IEEE 2019 International Conference on Industrial Engineering and Engineering Management (IEEM 2019)*, IEEE Macao Section; IEEE TEMS Singapore Chapter; IEEE TEMS Hong Kong Chapter, Dec 2019, Macao, China. pp.1588-1592,

A Binary Integer Programming Model for Detecting Arbitrage Opportunities in Currency Arbitrage Problems

¹E. O. Effanga, ²P. A. Ayuk

^{1,2}Department of Statistics, University of Calabar, Calabar, Nigeria

Abstract

This paper presents a binary integer programming (BIP) model for detecting currency arbitrage opportunities in foreign exchange markets. Currency arbitrage involves exploiting price discrepancies between different exchange rates to achieve risk-free profits. The proposed model represents the exchange rates as a directed graph, where vertices represent currencies and edges represent exchange rates. By maximizing the sum of the logarithms of selected exchange rates, the model identifies cycles in the graph that correspond to arbitrage opportunities. The BIP model includes constraints to ensure the conservation of flow, control the cycle length, and maintain consistency between vertex inclusion and edge selection. A practical example involving data from the cross-exchange rate of eight (8) most traded currencies which are the USD, EURO, JPY, GBP, CHF, CAD, AUD, and HKD downloaded from Bloomberg's Generic Composite Pricing (BGN) source used to illustrate the application of the model. The model was applied to the currency data at different cycle lengths and it was discovered that as the cycle length increases, the ROI yields improved with its 8th cycle length having the highest ROI which is \$1.7223 in 0.31sec elapsed time. However, limitations such as the computational complexity of BIP model which is generally classified as NP-hard were encountered. However, BIP model is highly recommended to be used in large-scale foreign exchange markets selectively, focusing on currencies with higher transaction volumes or volatility.

Keywords: Currency Arbitrage, Binary Integer Programming, Exchange Rates, Arbitrage Detection, Risk-Free Profit.

1. Introduction

In the global financial markets, currency arbitrage involves taking advantage of discrepancies in exchange rates to make a risk-free profit. The detection of such opportunities is crucial for traders and financial institutions, as it allows them to capitalize on price differences between different markets. This paper presents a binary integer programming (BIP) model for identifying arbitrage opportunities in a set of currencies, providing a structured and mathematical approach to detect cycles in the exchange rate graph that indicate potential arbitrage. BIP ensures that all decision variables are binary (0 or 1). This aligns with the nature of arbitrage decisions, which are inherently binary: either an arbitrage opportunity exists (1) or it does not (0). BIP also allows for the incorporation of more complex constraints and logical relationships between variables, offering a more precise and flexible approach to arbitrage detection.

Currency arbitrage occurs when a trader can buy and sell a series of currencies in such a way that they end up with more of a currency than they started with, without taking any risk. This process involves a sequence of exchanges between multiple currencies, starting and ending with the same currency, where the product of the exchange rates in the cycle exceeds one. The mathematical representation of this concept is critical in identifying arbitrage opportunities, as it involves graph theory and optimization techniques.

The aim of this research is to formulate a Binary Integer Programming (BIP) model, with the goal of improving efficiency by their ability to detect profitable paths.

2.Methodology

General Binary Integer Programming Model

$$\text{Max } Z = \sum_{j=1}^n p_j x_j$$

Subject to:

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, i = 1, 2, \dots, m$$

$$x_j = 0, 1$$

Where:

x_j := j^{th} decision variable

p_j := unit contribution of variable j to the objective function

b_i := amount of resource i available

a_{ij} := unit consumption of i by level of activity j

BIP model by Li and Zhang (2008)

Li and Zhang (2008) paper titled "Binary Integer Programming Models for Arbitrage Detection in Currency Markets" presented a binary integer programming (BIP) model designed to detect arbitrage opportunities by incorporating exchange rates and transaction costs. Here's an outline of their model:

Variables:

x_{ij} : Binary decision variable that indicates whether a transaction from currency i to currency j is part of the arbitrage cycle (1 if included, 0 otherwise).

Parameters:

r_{ij} : The exchange rate from currency i to currency j .

c_{ij} : The transaction cost associated with converting currency i to currency j .

n : The total number of currencies.

k : A large constant used to impose arbitrage detection conditions.

$$\text{maximize } Z = \sum_{i=1}^n \sum_{j=1}^n (r_{ij} - c_{ij}) \cdot x_{ij}$$

Subject to:

$$\sum_{i=1}^n \sum_{j=1}^n \log(r_{ij} - c_{ij}) \cdot x_{ij} \geq 0$$

$$\sum_{j=1}^n x_{ij} - \sum_{j=1}^n x_{ji} = 0 \quad \forall i$$

$$\sum_{j=1}^n x_{ij} \leq Ky_i \quad \forall i$$

$$x_{ij} = 0, 1 \quad \forall i, j$$

$$y_i = 0, 1 \quad \forall i$$

Model Formulation

To formulate a binary integer programming model for detecting arbitrage opportunities in a currency arbitrage problem, we represent the problem using a directed graph where the vertices represent currencies and the directed edges represent exchange rates between these currencies. The goal is to find a cycle in the graph where the product of the exchange rates along the cycle is greater than 1, indicating an arbitrage opportunity.

Variables

We define the following variables:

n : = number of currencies.

x_{ij} : = binary variable,

Where $x_{ij} = \begin{cases} 1, & \text{if the edge from currency } i \text{ to currency } j \text{ is part of the arbitrage cycle} \\ 0, & \text{otherwise} \end{cases}$,

R_{ij} : = exchange rate from currency i to currency j .

y_i : = binary variable.

Where $y_i = \begin{cases} 1, & \text{if vertex } i \text{ is included in the cycle} \\ 0, & \text{otherwise} \end{cases}$

X_0 : = Initial capital

Objective

The objective function maximizes the sum of the logarithms of exchange rates along the selected edges, which corresponds to maximizing the product of the exchange rates in the cycle. This can be transformed into a linear objective function:

$$\text{maximize } Z = \sum_{i=1}^n \sum_{j=1}^n \log(R_{ij}) \cdot x_{ij}$$

Constraints

Starting and Ending Currency constraints:

$$\sum_{j=1}^n x_{sj} = 1$$

$$\sum_{i=1}^n x_{is} = 1$$

These constraints ensure that the cycle starts and ends at the source currency s , differentiating the source currency's flow conservation from that of other currencies.

Flow Conservation Constraints:

These constraints ensure that if a currency is entered, it must be exited, maintaining the flow through the cycle. For all currencies, $i \neq j$, except the source currency s , the number of incoming edges must equal the number of outgoing edges, ensuring that the flow is balanced through these vertices.

$$\sum_{j=1}^n x_{ij} - \sum_{j=1}^n x_{ji} = 0 \quad \forall i \neq s$$

Cycle Constraint:

The cycle constraint ensures that at least one cycle is selected in the solution.

$$\sum_{i=1}^n \sum_{j=1}^n x_{ij} \geq 1$$

No-Self Loop Constraint:

To avoid trivial cycles, you may add a constraint to prevent self-loops, i.e., $x_{ii} = 0$ for all i .

Cycle Length Lower Bound:

$$\sum_{j=1}^n y_i \geq L$$

This constraint ensures that the cycle contains at least L vertices.

Cycle Length Upper Bound:

$$\sum_{j=1}^n y_i \leq U$$

This constraint ensures that the cycle contains no more than U vertices.

Vertex Inclusion and Edge Selection Relationship:

These constraints ensure that if any edge entering or leaving vertex i is selected, then vertex i must be included in the cycle.

$$\sum_{j=1}^n x_{ij} - ny_i \leq 0 \quad \forall i$$

$$\sum_{j=1}^n x_{ji} - ny_i \leq 0 \quad \forall i$$

Connecting x_{ij} and y_i :

$$y_i \geq x_{ij} \quad \forall i, j$$

$$y_j \geq x_{ij} \quad \forall i, j$$

These constraints ensure that if an edge from i to j is selected (i.e., $x_{ij} = 1$), then both vertices i and j are included in the cycle (i.e., $y_i = 1$ and $y_j = 1$).

The full model is:

$$\text{maximize } Z = \sum_{i=1}^n \sum_{j=1}^n \log(R_{ij}) \cdot x_{ij}$$

Subject to:

$$\sum_{j=1}^n x_{sj} = 1$$

$$\sum_{i=1}^n x_{is} = 1$$

$$\sum_{j=1}^n x_{ij} - \sum_{j=1}^n x_{ji} = 0 \quad \forall i \neq s$$

$$\sum_{i=1}^n \sum_{j=1}^n x_{ij} \geq 1$$

$$\begin{aligned} \sum_{j=1}^n y_i &\geq L \\ \sum_{j=1}^n y_i &\leq U \\ \sum_{j=1}^n x_{ij} - ny_i &\leq 0 \quad \forall i \\ \sum_{j=1}^n x_{ji} - ny_i &\leq 0 \quad \forall i \\ y_i - x_{ij} &\geq 0 \quad \forall i, j \\ y_j - x_{ij} &\geq 0 \quad \forall i, j \\ x_{ij} &= 0, 1 \quad \forall i, j \\ y_i &= 0, 1 \quad \forall i \end{aligned}$$

APPLICATION:

Application of the binary integer model

Table 1 below is the source data used in formulating our binary integer and table2also showed the log transformed data used also.

TABLE 1:

The overview dataas at 6th July, 2024

(i,j)	1	2	3	4	5	6	7	8
1	1	1.0840	0.0062	1.2815	1.1164	0.7331	0.6749	0.1280
2	0.9225	1	0.0057	1.1822	1.0299	0.6763	0.6226	0.1181
3	160.7500	174.2530	1	206.0011	179.4686	117.8519	108.4902	20.5752
4	0.7803	0.8459	0.0049	1	0.8712	0.5721	0.5266	0.0999
5	0.8957	0.9709	0.0056	1.1478	1	0.6567	0.6045	0.1146
6	1.3640	1.4786	0.0085	1.7480	1.5228	1	0.9206	0.1746
7	1.4817	1.6062	0.0092	1.8988	1.6542	1.0863	1	0.1897
8	7.8128	8.4691	0.0486	10.0121	8.7226	5.7279	5.2729	1

TABLE 2:
Log Transformation of Overview Data as at 6th July, 2024

(i,j)	1	2	3	4	5	6	7	8
1	0.0000	0.0350	-2.2076	0.1077	0.0478	-0.1348	-0.1708	-0.8928
2	-0.0350	0.0000	-2.2441	0.0727	0.0128	-0.1699	-0.2058	-0.9278
3	2.2062	2.2412	0.0000	2.3139	2.2540	2.0713	2.0354	1.3133
4	-0.1077	-0.0727	-2.3098	0.0000	-0.0599	-0.2425	-0.2785	-1.0004
5	-0.0478	-0.0128	-2.2518	0.0599	0.0000	-0.1826	-0.2186	-0.9408
6	0.1348	0.1699	-2.0706	0.2425	0.1826	0.0000	-0.0359	-0.7580
7	0.1708	0.2058	-2.0362	0.2785	0.2186	0.0359	0.0000	-0.7219
8	0.8928	0.9278	-1.3134	1.0005	0.9406	0.7580	0.7220	0.0000

MODEL 1: Starting/ending currency is USSD = 1

$$\begin{aligned} \text{Maximize } Z = & 0.035029282x_{12} - 2.20760831x_{13} + 0.010771861x_{14} + 0.047819828x_{15} - \\ & 0.13483678x_{16} - 0.17076057x_{17} - \underline{0.89279003}x_{18} - 0.03503363x_{21} - 2.24412514x_{23} + \\ & 0.07269096x_{24} + 0.012795058x_{25} - \underline{0.16986061}x_{26} - \underline{0.20579088}x_{27} - \underline{0.9277501}x_{28} + \\ & 2.206150982x_{31} + 2.241180264x_{32} + 2.31386954x_{34} + 2.253988475x_{35} + 2.071336589x_{36} + \\ & 2.03539051x_{37} + 1.313344065x_{38} - 0.107738399x_{41} - \underline{0.07268098}x_{42} - \underline{2.30980392}x_{43} - \\ & \underline{0.05988213}x_{45} - 0.24252805x_{46} - \underline{0.27851915}x_{47} - \underline{1.00043451}x_{48} - \underline{0.04783743}x_{51} - \\ & 0.0128255x_{52} - \underline{2.25182297}x_{53} + 0.05986622x_{54} - \underline{0.18263298}x_{56} - \underline{0.21860369}x_{57} - \\ & \underline{0.94081538}x_{58} + 0.13481437x_{61} + 0.169850702x_{62} - \underline{2.07058207}x_{63} + \underline{0.24254143}x_{64} + \\ & \underline{0.182642868}x_{65} - \underline{0.03592903}x_{67} - \underline{0.75795576}x_{68} + 0.170760281x_{71} + 0.205799622x_{72} - \\ & \underline{2.03621217}x_{73} + 0.27847922x_{74} + 0.218588016x_{75} + 0.03594978x_{76} - \underline{0.7219326}x_{78} \\ & + \underline{0.892806707}x_{81} + 0.92783726x_{82} - 1.31336373 + 1.00052518x_{84} + \underline{0.940645957}x_{85} + \\ & 0.757995427x_{86} + 0.722049535x_{87} \end{aligned}$$

Sensitivity analysis:

A sensitivity analysis was conducted to assess the impact of changes in model parameters (e.g., exchange rates and cycle length) on the optimal solution. Data from table 3 and table 4 below were used to carry out the sensitivity analysis of our binary integer model.

TABLE 3
Overview dataas at 9th October, 2024:

(i,j)	1	2	3	4	5	6	7	8
1	1	<u>1.0955</u>	<u>0.0067</u>	<u>1.3086</u>	<u>1.1650</u>	<u>0.7312</u>	<u>0.6727</u>	<u>0.1286</u>
2	<u>0.9128</u>	1	<u>0.0061</u>	<u>1.1945</u>	<u>1.0634</u>	<u>0.6674</u>	<u>0.6141</u>	<u>0.1174</u>
3	<u>148.8300</u>	<u>163.0433</u>	1	<u>194.7589</u>	<u>173.3807</u>	<u>108.8177</u>	<u>100.1179</u>	<u>19.1453</u>
4	<u>0.7642</u>	<u>0.8372</u>	<u>0.0051</u>	1	<u>0.8902</u>	<u>0.5587</u>	<u>0.5141</u>	<u>0.0983</u>
5	<u>0.8584</u>	<u>0.9404</u>	<u>0.0058</u>	<u>1.1233</u>	1	<u>0.6276</u>	<u>0.5774</u>	<u>0.1104</u>
6	<u>1.3677</u>	<u>1.4983</u>	<u>0.0092</u>	<u>1.7898</u>	<u>1.5933</u>	1	<u>0.9201</u>	<u>0.1759</u>
7	<u>1.4865</u>	<u>1.6285</u>	<u>0.0100</u>	<u>1.9453</u>	<u>1.7318</u>	<u>1.0869</u>	1	<u>0.1912</u>
8	<u>7.7737</u>	<u>8.5161</u>	<u>0.0522</u>	<u>10.1727</u>	<u>9.0560</u>	<u>5.6838</u>	<u>5.2294</u>	1

TABLE 4
Log Transformation of Overview Data as at 9th October, 2024

(i,j)	1	2	3	4	5	6	7	8
1	0	0.0396	-2.1739	0.1168	0.0663	-0.1359	-0.1722	-0.8908
2	-0.0396	0	-2.2147	0.0772	0.0267	-0.1756	-0.2118	-0.9303
3	2.1727	2.2123	0	2.2895	2.2390	2.0367	2.0005	1.2821
4	-0.1168	-0.0772	-2.2924	0	-0.0505	-0.2528	-0.2890	-1.0074
5	-0.0663	-0.0267	-2.2366	0.0505	0	-0.2023	-0.2385	-0.9570
6	0.1360	0.1756	-2.0362	0.2528	0.2023	0	-0.0362	-0.7547
7	0.1722	0.2118	-2.0000	0.2890	0.2385	0.0362	0	-0.7185
8	0.8906	0.9302	-1.2823	1.0074	0.9569	0.7546	0.7185	0

MODEL 2: Starting/Ending currency is US\$ = 1

Maximize $Z = 0.0396x_{12} - 2.1739x_{13} + 0.1168x_{14} + 0.0663x_{15} - 0.1359x_{16} - 0.1722x_{17} - 0.8908x_{18} - 0.0396x_{21} - 2.2147x_{23} + 0.0772x_{24} + 0.0267x_{25} - 0.1756x_{26} - 0.2118x_{27} - 0.9303x_{28} + 2.1727x_{31} + 2.2123x_{32} + 2.2895x_{34} + 2.2390x_{35} + 2.0367x_{36} + 2.0005x_{37} + 1.2821x_{38} - 0.1168x_{41} - 0.0772x_{42} - 2.2924x_{43} - 0.0505x_{45} - 0.2528x_{46} - 0.2890x_{47} - 1.0074x_{48} - 0.0663x_{51} - 0.0267x_{52} - 2.2366x_{53} + 0.0505x_{54} - 0.2023x_{56} - 0.2385x_{57} - 0.9570x_{58} + 0.1360x_{61} + 0.1756x_{62} - 2.0362x_{63} + 0.2528x_{64} + 0.2023x_{65} - 0.0362x_{67} - 0.7547x_{68} + 0.1722x_{71} + 0.2118x_{72} - 2.0000x_{73} + 0.2890x_{74} + 0.2385x_{75} + 0.2385x_{76} - 0.7185x_{78} + 0.8906x_{81} + 0.9302x_{82} - 1.2823x_{83} + 1.0074x_{84} + 0.9569x_{85} + 0.7546x_{86} + 0.7185x_{87}$

Results of empirical study

LINGO Optimization software was used to solve our BIP model using ASUS Laptop which has a processor, Intel core i7, RAM 8GB, Windows 10 operating system and storage capacity of 512GB

The solutions for model 1 and model 2 are summarized in Table 5 And Table 6 respectively.

TABLE 5 Summary of Model 1 solution:

Cycle length	Elapsed Time (s)	Max Z	InvLog(Max Z)	Initial Capital	Final Capital	ROI
3	0.36	0.002608292	1.006024	\$100	\$100.6024	\$0.6024
4	0.35	0.006213523	1.01441	\$100	\$101.441	\$1.441
5	0.39	0.007000242	1.016249	\$100	\$101.6249	\$1.6249
6	0.36	0.007166678	1.016639	\$100	\$101.6639	\$1.6639
7	0.32	0.007304576	1.016962	\$100	\$101.6962	\$1.6962
8	0.31	0.00741621	1.017223	\$100	\$101.7223	\$1.7223

TABLE 6 Summary of Model 2 solution:

Cycle length	Elapsed Time (s)	Max Z	InvLog(Max Z)	Initial Capital	Final Capital	ROI
3	0.7	0.2024000	1.593676	\$100	\$159.3676	\$59.3676
4	1.74	0.2034000	1.59735	\$100	\$159.735	\$59.735
5	0.52	0.2058000	1.606201	\$100	\$160.6201	\$60.6201
6	0.35	0.2059000	1.606571	\$100	\$160.6571	\$60.6571
7	0.35	0.2059000	1.606571	\$100	\$160.6571	\$60.6571
8	0.32	0.2059000	1.606571	\$100	\$160.6571	\$60.6571

Discussion of finding

The BIP model provides a more profitable allocation of capital in arbitrage detection as the cycle length increase. The solving time (elapsed time) for this model also decreases as the length of the cycle increase as showed in tables 5 and 6. BIP approaches typically focus on exact optimization, ensuring that the capital is allocated in the most efficient manner, which explains the superior ROI.

BIP is computationally intensive, especially for large datasets. Solving integer programming problems often requires a lot of computational resources (CPU, memory) and time, as the solution space grows exponentially with the number of variables and constraints. This makes BIP less suitable for real-time or large-scale applications.

Dependence on solver efficiency: The performance of BIP is highly dependent on the optimization solvers used (e.g., LINGO, Gurobi, CPLEX because these solvers were developed with advanced techniques like **branch-and-bound**, **cutting planes**, **heuristics**, and **pre-processing** to handle BIPs efficiently.). The quality of solutions and time to convergence vary with the solver's efficiency, parameter settings, and hardware used.

Summary

A BIP model was formulated for the optimal allocation of capital among detected arbitrage opportunities. The BIP approach in this study provided a mathematically rigorous solution to maximizing ROI by selecting the most profitable arbitrage cycles while considering constraints on available capital and market conditions.

Conclusion

The binary integer programming model provides a robust framework for detecting arbitrage opportunities in currency markets. By leveraging the structure of exchange rates as a directed graph and employing logarithmic transformations, the model simplifies the problem and enables efficient detection of profitable cycles. This method is not only applicable to currency arbitrage but can also be extended to other financial instruments and markets where similar opportunities arise. The continuous advancement in optimization techniques and computational power enhances the practical applicability of this model, making it an essential tool in the arsenal of traders and financial analysts.

References

- Bertsimas, D., & Tsitsiklis, J. N. (1997). *"Introduction to Linear Optimization."* Athena Scientific. P147-268
- Kendall, D. G., & Park, Y. S. (2011). Binary integer programming for high-frequency arbitrage detection. *Quantitative Finance*, 11(6), 1011-1022.
- Konno, H., & Yamazaki, H. (1991). Portfolio arbitrage and binary integer programming. *Operations Research*, 39(4), 625-639.
- Li, D., & Yorke-Smith, N. (2012). "Binary Integer Programming for Arbitrage Detection." *Operations Research Letters*, 40(3), 185-190.
- Li, X., & Zhang, Y. (2008). Binary integer programming models for arbitrage detection in currency markets. *Journal of Financial Markets*, 11(1), 41-64.
- Nemhauser, G. L., & Wolsey, L. A. (1988). *"Integer and Combinatorial Optimization."* John Wiley & Sons. P86-165

Advanced Techniques for Solving Specific Classes of Third-Order Nonlinear Partial Differential Equations Using the Balance Method: Theoretical Developments and Practical Applications

¹Chikwe, C.F ²Ezeorah, J N ³Essang, S.O

¹Department of Mathematics, University of Calabar, Calabar, Nigeria

²Department of Mathematics, University of Calabar, Calabar, Nigeria

³Department of Mathematics and Computer Science, Arthur Jarvis University, Akpabuyo, Nigeria

Abstract

Nonlinear partial differential equations (NLPDEs) are fundamental in describing a wide array of physical phenomena across various fields such as fluid dynamics, optics, and plasma physics. These equations are inherently challenging to solve due to their complex, non-linear nature. Consequently, numerous analytical and numerical methods have been developed to address these challenges. This study specifically addresses the modified Korteweg-de Vries (KdV) third-order equation, which is pivotal in modeling shallow water waves and other related physical systems. In this research, we employed the balance method to derive solutions for the modified KdV equation. The balance method, known for its efficacy in handling higher-order nonlinear terms, was utilized to construct exact soliton solutions. We successfully derived both first and second kind soliton solutions, which are critical in understanding the dynamics of nonlinear wave propagation. To further validate our approach, we visualized these solutions using advanced mathematical software, providing a graphical representation of the soliton behavior. These visual models offer significant insights into the shallow water wave phenomena and confirm the practical applicability of our theoretical results. Through a series of illustrative examples, we demonstrated the robustness and effectiveness of the balance method in solving the modified KdV equation. The results underscore the method's potential as a standard, reliable, and computationally efficient tool for tackling complex nonlinear equations. Our findings contribute to a deeper understanding of soliton dynamics and offer a solid foundation for future research in the field of nonlinear wave equations.

Keywords: Nonlinear Partial Differential Equations (NLPDEs), Modified Korteweg-de Vries (KdV) Equation, Balance Method, Soliton Solutions, Shallow Water Waves, Nonlinear Wave Propagation

1. Introduction

The study of nonlinear partial differential equations (NLPDEs) is significant place in mathematical physics and engineering. Among the myriad classes of NLPDEs, third-order nonlinear partial differential equations (NLPDEs) stand out due to their relevance in describing various wave propagation and soliton behaviors. Traditional methods for solving these equations often fall short in providing exact solutions, necessitating the development and application of advanced techniques. One such technique is the balance method, which has shown remarkable efficacy in addressing the challenges posed by specific classes of third-order NLPDEs (Wang, Zhou, & Li, 2022; Fan & Zhang, 2023).

Traditional methods for solving NLPDEs include the inverse scattering transform, Hirota's direct method, and similarity reductions, each with its own set of advantages and limitations (Ablowitz & Clarkson, 2023; Hirota, 2004; Pucci, 2024). However, these methods can become cumbersome when dealing with

higher-order or highly nonlinear terms. The need for more versatile and straightforward techniques has led to the exploration of the balance method, which simplifies the solution process by balancing the highest-order derivatives and nonlinear terms in the equations (Wang et al., 2022; Fan & Zhang, 2023).

2. Preliminaries and Definitions

2.1 Non-Linear PDEs

Nonlinearity in PDEs arises when the equation involves products or powers of the unknown function and its derivatives or when it includes transcendental functions like sine, cosine, or exponential functions of the dependent variable (Logan, 2015). A general form of a second-order nonlinear PDE could be written as:

$$A(u) \frac{\partial^2 u}{\partial x^2} + B(u) \frac{\partial^2 u}{\partial y^2} + C(u, \nabla u) = 0 \quad 1$$

2.6 General Non-Linear PDE

$$P(u, u_x, u_t, u_{xx}, u_{xy}, u_{yy}, \dots) = 0 \quad 2$$

u= wave equation, t=time, x= spatial coordinate, P= polynomial functions of the arguments

3. Materials and Methods

4.1 Hirota Bilinear Operators

The Hirota bilinear method introduces specific differential operators known as Hirota derivatives or bilinear differential operators. These operators are defined to act bilinearly on functions and are instrumental in expressing nonlinear equations in a bilinear form. For this research we will assume the

solution to the NLPDE as; $u = a_{pd} ((\ln w)_{pd} + \sum a_{kj} (\ln w)_{kj} + a_{00})^q$

$$u = u(x, t), w = w(x, t), \text{ and } \frac{\partial^{k+j}}{\partial x^k \partial t^j} \ln w(x, t) = ((\ln w))_{k,j} \quad 3$$

And we take $a_{k,j} (k = 0, 1, 2, \dots, p, j = 0, 1, 2, \dots, d)$, which are the balance coefficients that are also the constants. We evaluated them, Elaski transform, fractional derivatives, and the balance method, and obtained

$$a_{20} = 12\beta \text{ and } a_{10} = 0$$

Then we do a substitution $2a_{10} = 12\beta a_{10}$ and obtain the main equation $u = a_{20} \ln(w)_{xx} + a_{10} \ln(w)_x + a_{00}$

we so that $u = 12 \ln(w)_{xx} + 0 \ln(w)_x + a_{00} = 12 \ln(w)_{xx} + a_{00}$, and

$$12 \ln(w)_{xx} + a_{00} = 12\beta \left[\frac{w_{xx}}{w} - \frac{w_x^2}{w^2} \right] + a_{00}. \text{ where } a_{00} \text{ represents any arbitrary constant,}$$

$$u = 12 \ln(w)_{xx} + a_{00} = 12\beta \left[\frac{w_{xx}}{w} - \frac{w_x^2}{w^2} \right] + a_{00}.$$

4

Connecting the latter into the PDE, we produce the following outcome where a_{20} is given as;

$a_{20} = 12\beta$ and $a_{10} = 0$, so that

$$12\beta \left(\frac{w_{xxt}}{w} - \frac{w_{xx} w_t}{w^2} - \frac{2w_x^2 w_t}{w^3} + \frac{2w_x^2 w_t}{w^3} \right) + 12\beta \left(a_{00} \frac{w_{xxx}}{w} - 3a_{00} \frac{w_{xx} w_x}{w^2} + 2a_{00} \frac{w_x^3}{w^3} \right) \\ 12\beta \left(\beta \left[\frac{w_{xxxx}}{w} + \left(\frac{2w_{xxx} w_{xx} - 5w_{xxx} w_x}{w^2} \right) + \left(\frac{8w_{xxx} w_x^2 - 6w_{xx}^2 w_x}{w^3} \right) \right] \right) = 0$$

5

This simplifies to;

$$12(\beta[C_1 + C_2 + C_3]) = 0$$

6

Where;

$$C_1 = \frac{w_{xxt}}{w} - \frac{w_{xx} w_t}{w^2} - \frac{2w_x^2 w_t}{w^3} + \frac{2w_x^2 w_t}{w^3}$$

7

$$C_2 = a_{00} \frac{w_{xxx}}{w} - 3a_{00} \frac{w_{xx} w_x}{w^2} + 2a_{00} \frac{w_x^3}{w^3}$$

8

and

$$C_3 = \frac{w_{xxxx}}{w} + \left(\frac{2w_{xx}w_{xx} - 5w_{xxx}w_x}{w^2} \right) + \left(\frac{8w_{xx}w_x^2 - 6w_{xx}^2w_x}{w^3} \right)$$

9

So that;

$$12(\beta[C_1 + C_2 + C_3]) = 0.$$

10

Introducing integral, we have

$$\int 12\beta \left(\frac{w_{xx}}{w} - \frac{w_{xx}w_t}{w^2} - \frac{2w_x^2w_t}{w^3} + \frac{2w_x^2w_t}{w^3} \right) + 12\beta \left(a_{00} \frac{w_{xxx}}{w} - 3a_{00} \frac{w_{xx}w_x}{w^2} + 2a_{00} \frac{w_x^3}{w^3} \right) dw$$

$$\int 12\beta \left(\beta \left[\frac{w_{xxxx}}{w} + \left(\frac{2w_{xx}w_{xx} - 5w_{xxx}w_x}{w^2} \right) + \left(\frac{8w_{xx}w_x^2 - 6w_{xx}^2w_x}{w^3} \right) \right] \right) dw = 0$$

11

These yields

$$\frac{\partial}{\partial x} \left[\frac{(w_{xt}w - w_xw_t) + \beta(w_{xxxx}w - 4w_xw_{xxx} + 3w_{xx}^2) + a_{00}(w_{xt}w - w_x^2)}{w_2} \right] = 0$$

12

If we compare this with the equation;

$$u_{xxx} = a_{20} \left[\frac{w_{xxxx}}{w} - 5 \frac{w_{xxx}w_x}{w^2} - 10 \frac{w_{xxx}w_{xx}}{w^2} - 20 \frac{w_{xxx}w_{xx}^2}{w^2} + 30 \frac{w_{xx}^2w_x}{w^3} - 60 \frac{w_{xx}w_x^3}{w^4} - 24 \frac{w_x^5}{w^5} \right] \dots$$

$$\dots + a_{10} \left[\frac{w_{xxxx}}{w} - 4 \frac{w_{xxx}w_x}{w^2} - 3 \frac{w_{xx}^2}{w^2} - 12 \frac{w_{xx}w_x^2}{w^3} - 6 \frac{w_x^4}{w^4} \right]$$

13

We observe the identical terms and factors with;

$$\left[\frac{(w_{xt}w - w_xw_t) + \beta(w_{xxxx}w - 4w_xw_{xxx} + 3w_{xx}^2) + a_{00}(w_{xt}w - w_x^2)}{w_2} \right] - k(t)w^2 = 0$$

14

Where $k(t)$ is a function that is arbitrary with respect to t , likewise a_{00} .

Certainly, putting $k(t)=0$ in the equation above, we get the bilinear equation;

$$\frac{(w_{xt}w - w_x w_t) + \beta(w_{xxxx}w - 4w_x w_{xxx} + 3w_{xx}^2) + a_{00}(w_{xx}w - w_x^2)}{w_2} = 0$$

15

This equation can be written with the help of a differential operator D as;

$$(D_x D_t + \beta D_x^4 + a_{00} D_x^2)(w, w) = 0$$

16

Where;

$$D_x^m D_t^n (a, b) = (\partial_x - \partial_x')^m (\partial_t - \partial_t')^n a(x, t) b(x', t')$$

17

$$D_x^m D_t^n (a, b) = (\partial_x - \partial_x')^m (\partial_t - \partial_t')^n a(x, t) b(x', t') \big|_{x=x', t=t'}$$

18

We set $a_{00}=0$ in equation the equation above then Applying the Hiruta's Method we simplify the bilinear equation into;

$$(D_x D_t + \beta D_x^4)(w, w) = 0$$

19

This method yields a more general and applicable bilinear equation of the modified KdV equation is obtained by using Hirota's Method.

4. Results

Let us take into consideration the fractional-order nonlinear KdV system;

$$\begin{aligned} \frac{\partial^\rho \lambda}{\partial t^\rho} &= -a \frac{\partial^3 \lambda}{\partial^3 r} - 6a\lambda \frac{\partial \lambda}{\partial r} + 6v \frac{\partial v}{\partial r} \\ \frac{\partial^\rho v}{\partial t^\rho} &= -a \frac{\partial^3 v}{\partial^3 r} - 3a\lambda \frac{\partial v}{\partial r}, \quad 0 < \rho < 1 \end{aligned}$$

20

With the initial condition

$$\lambda(r, 0) = \eta^2 \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right)$$

$$v(r, 0) = \sqrt{\frac{\beta}{2}} \eta^2 \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right)$$

21

In the case of $\rho=1$, the exact KdV scheme is given by,

$$\lambda(r, t) = \eta^2 \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} - \frac{\beta \eta^3 t}{2} \right)$$

$$v(r, t) = \sqrt{\frac{\beta}{2}} \eta^2 \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} - \frac{\beta \eta^3 t}{2} \right)$$

22

We the Elzaki transform to the equations above, we get;

$$\lambda_1(r, t) = \eta^5 \beta \tanh \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) \frac{t^\rho}{\Gamma(1+\rho)}$$

$$v_1(r, t) = \eta^5 \sqrt{\frac{\beta^3}{2}} \eta^2 \tanh \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) \frac{t^\rho}{\Gamma(1+\rho)}$$

23

Furthermore,

$$[\lambda_2(r, t)] = E^{-1} \left[Es^\rho \left[-a \frac{\partial^3 \lambda}{\partial^3 r} - 6a\lambda \frac{\partial \lambda}{\partial r} + 6v \frac{\partial v}{\partial r} \right] \right]$$

$$[v_2(r, t)] = E^{-1} \left[Es^\rho \left[-a \frac{\partial^3 v}{\partial^3 r} - 3a\lambda \frac{\partial v}{\partial r} \right] \right], \quad 0 < \rho < 1$$

24

With the analytical methods emphasized in this work, we obtain;

$$\lambda_2(r, t) = \frac{\eta^8 a^2}{2} \left(2 \cosh^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) - 3 \right) \operatorname{sech}^4 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) \frac{t^{2\rho}}{\Gamma(1+2\rho)}$$

$$v_2(r, t) = \eta^5 \sqrt{\frac{2a^5}{16}} \eta^2 \left(2 \cosh^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) - 3 \right) \operatorname{sech}^4 \left(\frac{\beta}{2} + \frac{\lambda r}{2} \right) \frac{t^{2\rho}}{\Gamma(1+2\rho)}$$

25

Hence, generally, we can extend the results to n for further studies

$$\begin{aligned} [\lambda_n(r, t)] &= E^{-1} \left[Es^\rho \left[-a \frac{\partial^3 \lambda}{\partial^3 r} - 6a\lambda \frac{\partial \lambda}{\partial r} + 6v \frac{\partial v}{\partial r} \right] \right] \\ [v_n(r, t)] &= E^{-1} \left[Es^\rho \left[-a \frac{\partial^3 v}{\partial^3 r} - 3a\lambda \frac{\partial v}{\partial r} \right] \right], \quad n \geq 0 \end{aligned} \quad 26$$

The result will be expressed in the series form;

$$\begin{aligned} \lambda(r, t) &= \lambda_0(r, t) + \lambda_1(r, t) + \lambda_2(r, t) + \lambda_3(r, t) + \dots + \lambda_n(r, t) \\ v(r, t) &= v_0(r, t) + v_1(r, t) + v_2(r, t) + v_3(r, t) + \dots + v_n(r, t) \quad n \geq 0 \end{aligned} \quad 27$$

Specifically, the exact results of the KdV scheme are given as;

$$\begin{aligned} \lambda(r, t) &= \eta^2 \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} - \frac{a\eta^3 t}{2} \right) \\ v(r, t) &= \sqrt{\frac{\beta}{2}} \eta^2 \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{\lambda r}{2} - \frac{a\eta^3 t}{2} \right) \end{aligned} \quad 28$$

Let us view the above solutions in a plot;

Surface plot of $\lambda(r, t) = \operatorname{sech}^2 \left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2} \right)$, $\beta = 2$

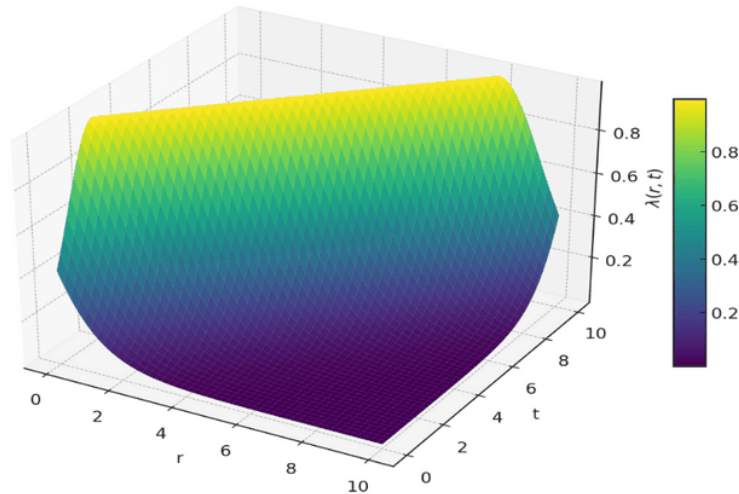


FIG 1 With $\beta=2$ and ranges for r and t from 0 to 10. This graph demonstrates the behavior of the squared hyperbolic secant function, which decays more sharply, highlighting the function's sensitivity to changes in r and t .

Surface plot of $\lambda(r, t) = \text{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = 50$

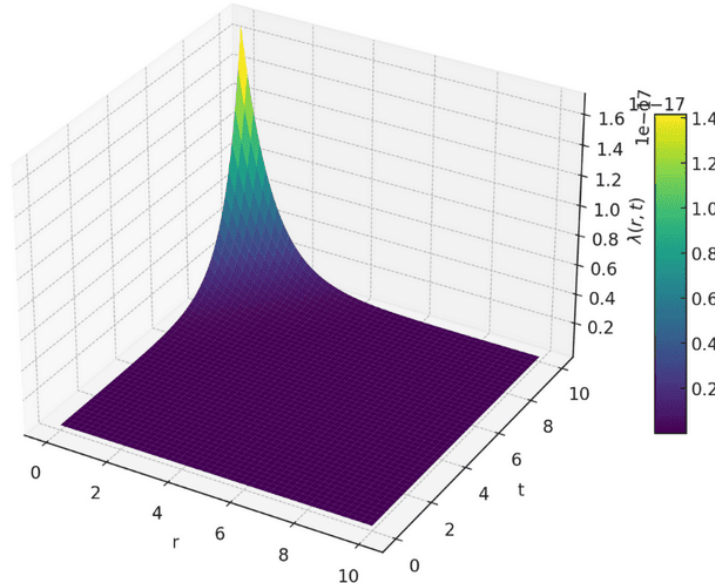


FIG 2 With $\beta=50$ increases significantly, the peak becomes very sharp and narrow, focusing the impact of the function near the center. This behavior showcases how increasing β drastically changes the distribution of λ

Surface plot of $\lambda(r, t) = \text{sech}^2\left(\frac{r}{2} - \frac{t}{2}\right)$, $\beta = 0$

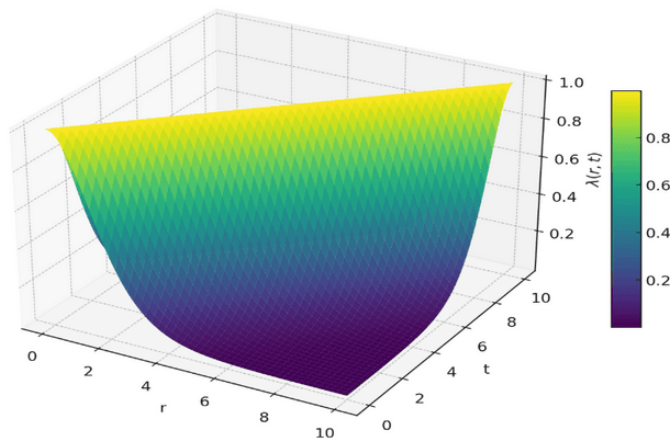


FIG 3 With $\beta=0$. This plot shows the symmetric behavior of the function around the line $r=t$, with no shift in the peak since β does not contribute any additional bias to the function's argument. The shape is more evenly distributed compared to non-zero β values

Surface plot of $\lambda(r, t) = \text{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = -20$

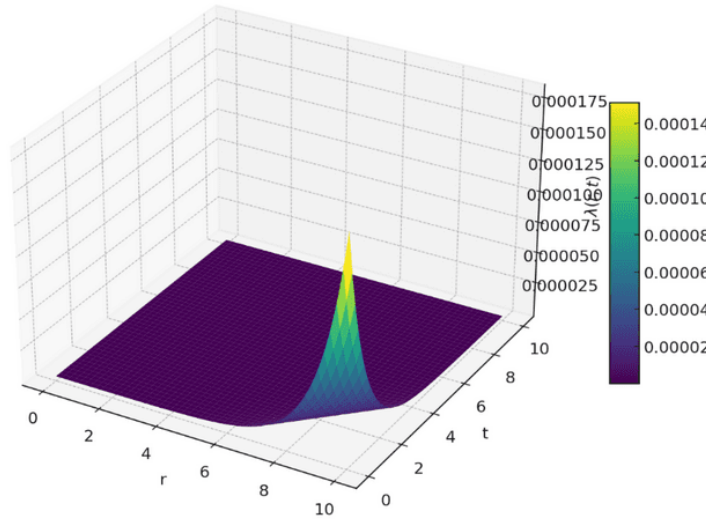


FIG 4 With $\beta=-20$. This plot illustrates how a significantly negative β shifts the peak toward higher r and t values, reflecting the influence of the negative bias in the function's argument. The peak is narrow, highlighting the localized impact of the function in this configuration

Surface plot of $\lambda(r, t) = \text{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = -50$

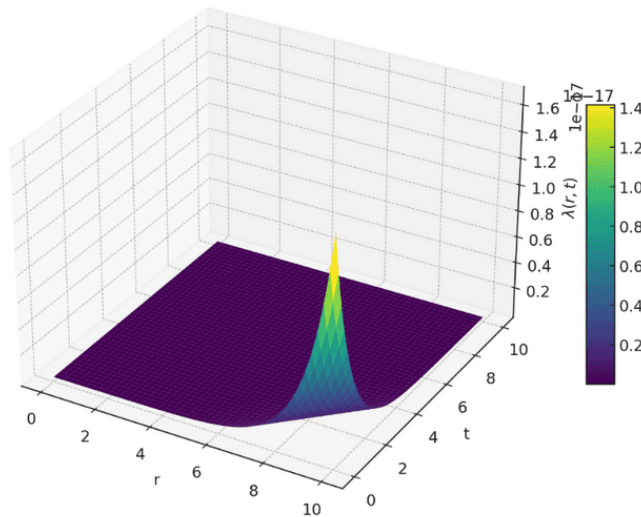


FIG 5 With $\beta=-50$. The significant negative value of β causes an even more pronounced peak shift towards higher r and t values, with a very narrow peak that emphasizes the function's steep decay. This

graph clearly shows the effect of large negative β values on the function's distribution and behavior over the given range.

Surface plot of $\lambda(r, t) = \text{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = -100$

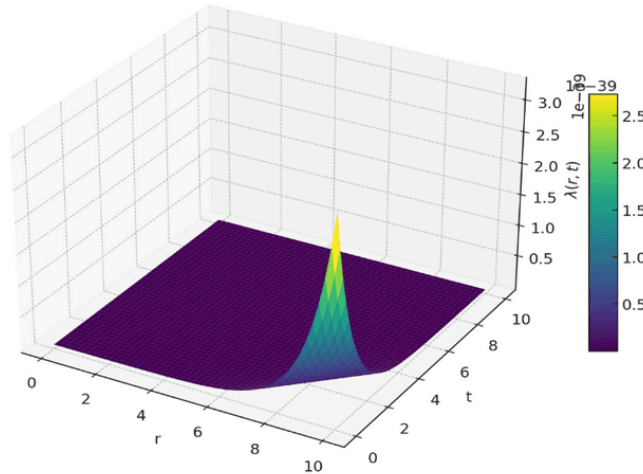


FIG 6 With $\beta=-100$. This extreme negative value of β causes the function's peak to shift further towards the higher end of both r and t values, resulting in a very localized and sharp peak. The graph emphasizes how dramatic the function's sensitivity to β is, particularly with large negative values.

Surface plot of $\lambda(r, t) = \text{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = 500$

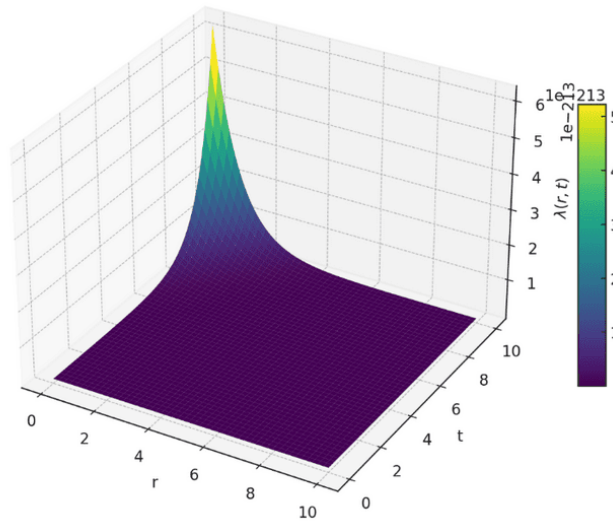


FIG 7 With $\beta=500$. With such a large positive β value, the function's peak is sharply focused near the origin and quickly decays, becoming almost negligible outside a very narrow region around the peak. This illustrates the extreme sensitivity of the function to high β values, effectively concentrating the function's effects on a tiny area of the plot.

Next, we create a comparison plot that includes all the varying values of β we've explored for the function. we'll consist of $\beta = -100, -50, -20, 0, 2, 50, 100$, and 500 in this series of plots to illustrate how the function changes with these different values. We'll place these plots side by side for a straightforward comparison.

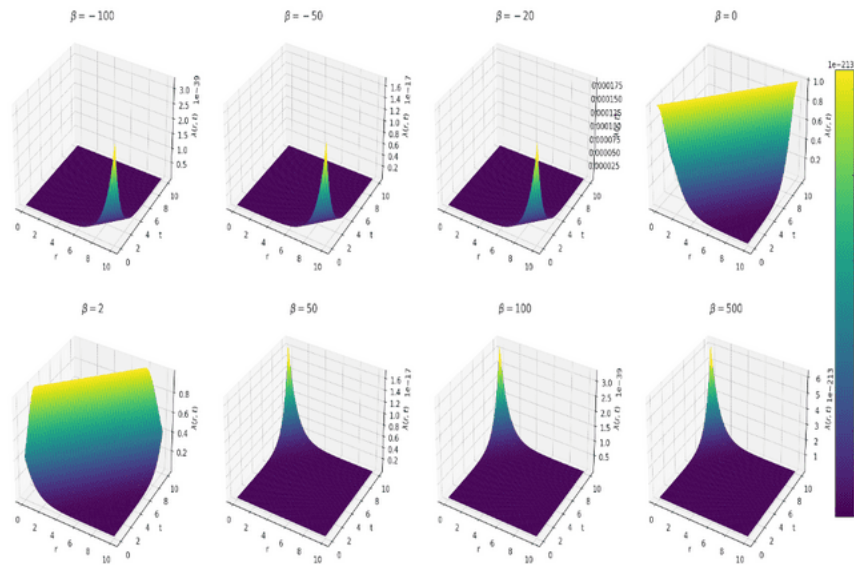


FIG 8 Comparison plot that includes all the varying values of β we've explored for the function. we'll include $\beta = -100, -50, -20, 0, 2, 50, 100$, and 500

Surface plot of $v(r, t) = \sqrt{0.5\beta} \operatorname{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = 2$

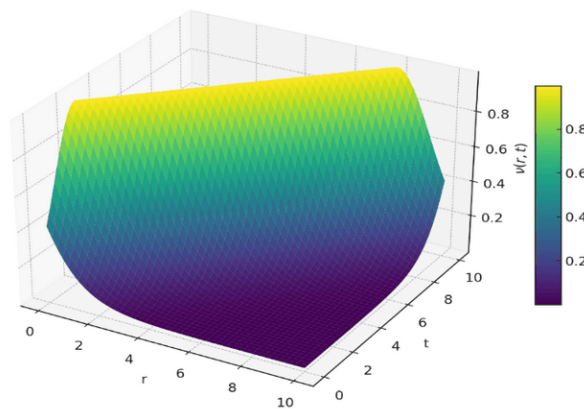


FIG 9 The surface plot for the function with $\beta=2$ and ranges for r and t from 0 to 10. This graph incorporates a scaling factor $\sqrt{0.5\beta}$, modifying the amplitude of the function. The plot shows how the function behaves over the specified ranges, illustrating the effect of the scaling on the function's shape and intensity.

Surface plot of $v(r, t) = \sqrt{0.5\beta} \operatorname{sech}^2\left(\frac{\beta}{2} + \frac{r}{2} - \frac{t}{2}\right)$, $\beta = 1000$

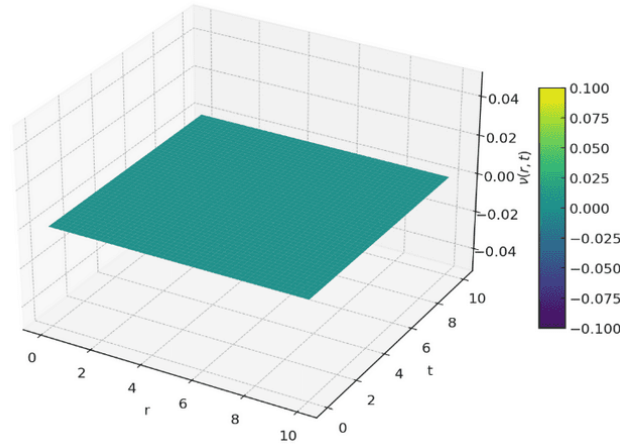


FIG 10 With $\beta=1000$. This plot demonstrates the dramatic effect of a very high β value, where the function's peak becomes extremely sharp and localized around the origin, due to the intense scaling of the function combined with the large β value. This significantly narrows the effective area where the function has substantial values, illustrating the sensitivity of the function to changes in β .

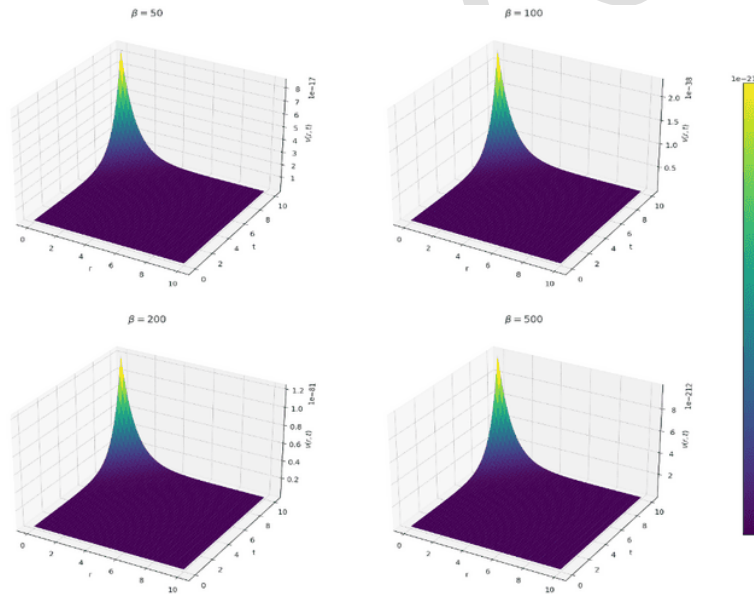


FIG 11 With β values of 50, 100, 200, and 500. Each subplot demonstrates the impact of increasing β on the function:

Higher β values result in sharper, more localized peaks, with the effect of the function being concentrated closer to the origin due to the increased influence of the function.

Scaling factor $\sqrt{0.5\beta}$ increases the amplitude of the function as β increases, enhancing the peak height while maintaining the function's narrow effective range.

This series of plots clearly shows how the scaling and shape of the function change significantly with larger β values, focusing more intensely on a smaller area.

References

- Ablowitz, M. J., & Clarkson, P. A. (2023). Solitons, nonlinear evolution equations and inverse scattering. [No DOI provided]
- Bekir, A., & Guner, O. (2014). Application of the (G'/G)-expansion method to space-time fractional differential equations. [No DOI provided]
- Drazin, P. G., & Johnson, R. S. (2023). Solitons: an introduction. [No DOI provided]
- Fan, E., & Zhang, J. (2023). On the homogeneous balance method. [No DOI provided]
- Gusu, C., & Bulo, T. R. (2022). Simplifying nonlinear PDEs using the G'/G method. [No DOI provided]
- Gusu, C., & Danu, A. (2022). Fixed point methods for boundary value problems in nonlinear wave equations. [No DOI provided]
- Gusu, C., & Diro, E. (2022). Hirota bilinear method for 2+1 dimensional NIPDEs. [No DOI provided]
- Hereman, W. (2013). Exact solutions of nonlinear partial differential equations by symbolic computation. [No DOI provided]
- Hirota, R. (2004). The direct method in soliton theory. Cambridge University Press. [No DOI provided]
- Khaled, K., Taher, S., & Ali, M. (2013). Hirota direct methods and soliton solutions for nonlinear PDEs. [No DOI provided]
- Kadomtsev, B. B., & Petviashvili, V. I. (2024). On the stability of solitary waves in weakly dispersing media. [No DOI provided]
- Korteweg, D. J., & de Vries, G. (2023). On the motion of waves in a rectangular canal, and on a new type of long stationary waves. [No DOI provided]
- Logan, J. D. (2015). Applied mathematics. Wiley. [No DOI provided]
- Ostrovsky, L., & Stepanyants, Y. (2023). Dynamics of internal solitons in the ocean. [No DOI provided]
- Pucci, E. (2024). Similarity reductions of nonlinear partial differential equations. [No DOI provided]
- Rogers, C., & Schief, W. K. (2002). Bäcklund and Darboux transformations: Geometry and modern applications in soliton theory. Cambridge University Press. [No DOI provided]
- Wang, L., Zhou, Y., & Li, Z. (2022). Application of the homogeneous balance method to nonlinear PDEs. [No DOI provided]
- Whitham, G. B. (2011). Linear and nonlinear waves. Wiley. [No DOI provided]

RADIATION SHIELDING PROPERTIES OF CLAM SHELL MORTAR FOR USE IN DIAGNOSTIC RADIOLOGY FACILITY

Abek Jonathan U, and Samuel O. Inyang.

Department of Physics, Faculty of Physical Sciences, University of Calabar. Calabar.

Email: jonathanabek@gmail.com,

Abstract

This study explores the radiation shielding properties of clamshell mortar for application in diagnostic radiology. Clamshells, rich in calcium carbonate, are evaluated for their effectiveness as a partial cement replacement. The clam shells collected from the local clam vendors were first washed, sun-dried, cracked into smaller pieces, and burned to ash at a temperature of 800°C. The crushed clamshells were sieved using a 300 µm sieve mesh, and the fine powder (pozzolan) was stored in a sealed container. A standard ratio of cement to sand was used in preparing a mortar of density 1608 kg/m³. Using this ratio, the weight of cement was replaced by 8 w/w, 12 w/w, 16 w/w, 20 w/w, 24 w/w and 28 w/w with clam shell ash, the mixtures were used to Prepare samples of mortar and label them for easy identification. X-ray irradiation test on the mortar samples was carried out at the x-ray room of the Diagnostic Radiology facility. The result shows that 16% replacement of cement with clam shell ash provides a suitable shielding material for a diagnostic x-ray facility. A thickness of 1.68 cm is found to absorb half the incident dose of 141.2 µGy, and TVL of 5.59 cm could be considered to improve radiation protection. This Research will enlighten the public on the possibility of using clamshell ash. As an alternative pozzolan for cement in the construction of mortar for shielding diagnostic radiology facilities. It will further educate the public on the need to properly shield radiation.

Keywords: clam shell, x-ray, shielding, mortar, concrete, properties.

Introduction

Diagnostic radiology is an essential tool in modern medicine, aiding healthcare professionals to detect and monitor several health conditions using both ionizing and non-ionizing radiation (Bushberg *et al.*, 2012). Despite its benefits, ionizing radiation is associated with potential risks, such as increased cancer rates and genetic damage for both patients and medical personnel (UNSCEAR, 2020; Vaiserman *et al.*, 2018). Therefore, radiation protection is of serious concern in diagnostic radiology practice.

Radiation shielding is an indispensable component of radiation protection strategy used in reducing radiation doses by attenuating the beam intensity emitted from diagnostic equipment (NCRPM, 2005). Shielding materials, such as lead and concrete, are widely used due to their high attenuation properties. However, these materials pose environmental disposal and health concerns and are costly (EPA, 2017). Thus, researchers have developed more interest in exploring alternative, eco-friendly, and maintainable shielding materials. The effectiveness of a material as a radiation shielding depends on its ability to attenuate, or absorb, the penetrating radiation. This can be characterized by the material's linear attenuation coefficient, which indicates how much the radiation is reduced per unit distance traveled through the material. In diagnostic radiology, the most commonly used radiations are x-rays and gamma rays, which have a relatively high energy. Therefore, the shielding material needs to have high density and atomic number to effectively absorb the radiation. The use of ionizing radiation in medical imaging is found in conventional radiography, fluorography, mammography, computed tomography, and nuclear medicine. Ionizing radiation can penetrate matter and tissues, interact with atoms and molecules, and cause the removal or addition of electrons (ionization) in the material. The interaction of ionizing radiation with matter produces an image that can reveal the internal structure and function of organs and tissues. However, exposure to ionizing radiation can also cause damage to deoxyribonucleic acid (DNA) and other cellular structures, which can lead to genetic mutations, cancer, and other health problems.

Non-ionizing radiation has lower energy and frequency than ionizing radiation and does not cause ionization in matter. They interact with the physical properties of tissues, such as density, conductivity, or magnetization, to produce an image or effect. Non-ionizing radiation, such as ultrasound, uses high-frequency sound waves that bounce off tissues to create an image of the internal structures. Magnetic resonance imaging uses a strong magnetic field and radio waves to align and excite protons in the body, which can produce a signal that can be reconstructed into an image. Non-ionizing radiation is generally considered safer than ionizing radiation because it does not cause direct DNA damage.

Clam Shell Mortar (CSM)

Clams are marine bivalve mollusks called “Bivalvia” scientifically, and belonging to the family of Veneridae. They are a type of shellfish with two shells hinged together, which they use to filter feed (Gökoğlu, 2021).

A clam shell mortar is a specialized construction material made by blending powdered shells of marine mollusk (clams) with traditional cementitious binders. The shells are cleaned, processed, and pulverized into fine particles before being mixed with cement, sand and water. The resulting mortar exhibits unique properties and can be used for various shielding applications. One of the most significant applications of Clam Shell mortar in building construction is its effectiveness as a radiation shielding material (Kim, 2022). Research in this area has focused on evaluating the mortar’s ability to attenuate gamma radiation. Studies suggest that the addition of powdered mollusk shells to cementitious materials enhances their radiation shielding properties, making clam shell mortar a potential material like lead or concrete (Smith *et al.*, 2018).

Clam shell mortar has been used for centuries in coastal regions where seashells are readily available. Clam shell has several advantages over modern cement-based mortars, as it is more environmentally friendly and sustainable, and it has better breathability and flexibility (Popović, 2023). Its breathability makes it a good option for buildings in areas with high humidity or where moisture infiltration is a concern. Clam shell mortar is a type of mortar that is made from crushed clam shells or other mollusk shells. It is a natural and sustainable alternative to cement, which is a major source of greenhouse gas emissions and environmental degradation, clam shell mortar has been used for centuries by various civilizations near the coastlines, where shells were abundant and easily accessible (Wang *et al.*, 2023). Clam shell mortar can be made by heating the shells to a high temperature, which causes them to release carbon dioxide and form calcium oxide (Table 1). The calcium oxide is then mixed with water, which triggers a chemical reaction that produces heat and converts the mixture into a dry, powdery material that can be used as a binder in mortar (Cahoon *et al.*, 2011). Clam shell mortar can be mixed with sand, gravel, or other fillers to create concrete for various construction purposes (Bamigboye et al 2021).

Clam shell mortar have some advantages over cement, such as lower carbon footprint, higher compressive strength, and better resistance to salt water and acid rain for future applications in coastal areas or regions with abundant shellfish farming (Wang *et al.*, 2023).

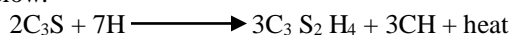
Table 1: Chemical composition of OPC and clam shell

OXIDES	OPC (%)	CLAM SHELL (%)
SiO ₂ (Silica)	20 - 25	1.60
Al ₂ O ₃ (Alumina)	5 – 10	0.92
Fe ₂ O ₃ (Iron oxide)	3 – 5	0.06
CaO (Calcium Oxide)	60 -65	51.56
MgO (Magnesium Oxide)	0.5 - 4	1.43
SO ₃ (Sulfur)	1 - 3	0.06
Na ₂ O (Sodium Oxide)	0.5	0.04
K ₂ O (Potassuim Oxide)	0.5	0.05

(Neville *et al.*, 1987; Peceño *et al.*, 2022).

Typically clam shell contains 90% to 98% of calcium carbonate (CaCO₃) which translate to a high calcium oxide (CaO) content calcinated in comparism to OPC with approximately 60% to 67% CaO, which means clam shells have a much greater calcium content than OPC. Clam shell composed of CaCO₃ which contribute to its hydration process in concrete, leading to improve compressive strength and density. This optimizes concrete properties with 6% replacement, but excessive replacement can hinder early strength due to slower hydration rates (de Freitas *et al.*, 2024).

Calcium Silicate Hydrate (C-S-H) is produced through the hydration of tricalcium silicate (C₃S) in the present of water, as shown in the equation below.



Here C-S-H acts as the primary binding phase in cement and can also interact with materials like clam shell, which contain calcium carbonate and silicates contributes to the formation of additional C-S-H, enhancing the binding characteristics of both cement and clam shell aggregates (Madadi *et al.*, 2022).

Materials and methods

Ordinary Portland Cement (OPC) (Lafarge Nigeria PLC, with 30% reduction in CO₂) was used as the major binder, sharp sand of particle fineness 0.85 mm, Clam shells were sourced from dealers at Watt market who gather them up as waste after removing the flesh of the clam. Water was sourced from nearby public water supply while moulds were designed and constructed at wood work workshop in University of Cross River State, Calabar. The cement, gravel and sharp sand were used as sourced without further purification. The following equipment's were used in this study: x-ray source fitted with lead collimator and radiation dose detector.

Clam Shell Preparation

The clam shells, were first washed with water to remove dirt such as soil and remaining clam flesh (Fig. 1). The washed clam shells were dried under the sun for three days until no water was observed on them. The dried shells were divided into two parts, one part was cracked into smaller pieces using hand held hammer and stored in a polyethene bag while the second part was



Figure: 1 (a & b): Clam shells waste

Preparation of Clam Shell Ash

The pieces of clam shells obtained (Fig. 2), were then packed into a crucible container and placed into a furnace, and the temperature set at 800°C for 2 hours to form ash (Ketebu and Farrow., 2017). The shells from the furnace were allowed to cold under normal room temperature before they were crushed into powder form using a grinding machine. The crushed shell samples were sieved using a 300 µm sieve mesh to separate the ash particles from none ash particles(Nduka *et al.*, 2023). The fine powder (pozzolan) of the clam shells were stored in a sealed container, prevented from dust, moist and other contaminants.



Figure 2: Clam shell ash.

Preparation of clam shell mortar

Moulds of thicknesses 2.5 cm, 5 cm, 7.5 cm, 10.0 cm, 12.5 cm and 15.0 cm were constructed for moulding of the samples. The moulds were cleaned and oiled with engine oil to enhance ease of removal of the blocks after setting and to prevent damage of the test blocks. The ratio of cement to sand used in preparing a standard mortar of density 1608 kg/m^3 is 1:4. This was adopted due to the sand particle size (Ketebu and Farrow., 2017). Using this ratio, the weight of cement was replaced by 0%w, 8%w, 12%w, 16%w, 20%w, 24%w and 28%w of clam shell ash and the mixtures used to prepare samples of mortar and labelled as m_0 , m_1 , m_2 , m_3 , m_4 , m_5 and m_6 respectively.

8% of the weight of cement was weighed out and replaced with the same weight of clam shell ash. This was repeated for 12%w, 16%w, 20%w, 24%w and 28%w of the clam shell ash. The control sample of mortar was prepared using the known standard ratio that offers 1608 kg/m^3 density and identified as m_0 (Ketebu *et al.*, 2017). Each sample of mortar was made by putting together the mixture of cement/clam shell ash and sand in a container before adding sufficient water to form a good paste. The paste was stirred using a stirrer that rotate at the speed of 140-180 rpm to ensure that the mortar was evenly mixed (Fig.3).



Figure 3: Clam shell mortar

X-Ray Transmission Test

X-ray irradiation test on the mortar samples was carried out using the diagnostic x-ray facility. Quality control measurements were previously conducted on the x-ray machine by the facility medical physicist and the results were found to be within set tolerances. Since the energy level (voltage tube) used in a general radiography is always in the range of 60-120 kVp (Aljabal *et al.*, 2019), it was therefore necessary to determine the attenuation level of the mortar

under irradiation of x-ray having operating potential of 80 kVp. The mortar sample was placed midway between the detector and the x-ray source to minimize scattered radiation reaching the detector. The detector source was fixed at 80 cm throughout the experiment. This procedure was repeated for all mortar samples. The dose result taken when there was no mortar was recorded as D_0 while doses transmitted through different thicknesses of mortar were also recorded as D_{m1} , D_{m2} , D_{m3} , D_{m4} , D_{m5} , D_{m6} .

Determination of linear attenuation coefficient (LAC), Half-value layer (HVL) and Tenth-value layer (TVL)

The (LAC) is a measure of how strongly a material can attenuates or absorbs a beam of x-rays photons using the equation

$$\ln D = -\mu x + \ln D_0 \quad (1)$$

where D_0 is the initial dose, D is the transmitted dose, while μ is evaluated by plotting a graph of $\ln D$ as a function of the thickness (x).

The HVL is the thickness of a material required to reduce the x-ray intensity to half its original value. HVL can be determined using the equation

$$HVL = \frac{0.693}{\mu} \quad (2)$$

The TVL is the thickness of a material required to reduce the x-ray intensity to one-tenth its original value. This can be determined using the equation\

$$TVL = \frac{2.30}{\mu} \quad (3)$$

X-ray transmission through mortar samples

At the radiology room, a transmission of x-ray photons with the initial dose D_0 of 141.2 μGy was recorded without any sample as presented (Table 2). The control samples were used as normal mortar without any replacement of cement with CSA. Transmission results for the different mortar thicknesses are given in (Table 2). There was no dose reading at 12.5 cm and 15.0 cm thicknesses as the transmitted dose was below detection limit. The transmitted doses through the control samples are observed to be less than all other sample types except for the mortar sample with 8% CSA. Mortar sample with 8% CSA had the greatest radiation absorption efficiency, as it allowed the least dose of 24.2 μGy , 16.0 μGy , 12.4 μGy and 3.3 μGy to be transmitted through its 2.5 cm, 5.0 cm, 7.5 cm and 10.0 cm thicknesses respectively, while the detector could not detect or read the doses for 12.5 cm and 15.0 cm thicknesses. Other mortar samples types displayed various radiation absorption efficiency, which cannot be compared to those of the control sample and 8% CSA mortar.

From the result shown in Table 2, it is observed that the normal mortar produced a better shielding at the initial dose of 141.2 μGy than mortars having 12% CSA, 16% CSA, 20% CSA, 24% CSA and 28% CSA, mortar sample with 8% CSA produced a better shielding property than the control sample at the same initial dose of 141.2 μGy .

Table 2: Calculated values for all doses ($\ln D$) of mortar samples

X (cm)	Dose (μGy)						
	M_0	M_8	M_{12}	M_{16}	M_{20}	M_{24}	M_{28}
0.00	4.95	4.95	4.95	4.95	4.95	4.95	4.95
2.50	3.19	3.19	4.04	4.22	3.92	3.87	3.70
5.00	2.87	2.77	2.83	2.83	2.87	3.04	2.48
7.50	2.53	2.52	2.04	2.04	1.87	2.83	2.42
10.00	1.87	1.19	BDL	BDL	BDL	BDL	1.87
12.50	BDL	BDL	BDL	BDL	BDL	BDL	BDL
15.00	BDL	BDL	BDL	BDL	BDL	BDL	BDL

(BDL=Below Detection Limit)

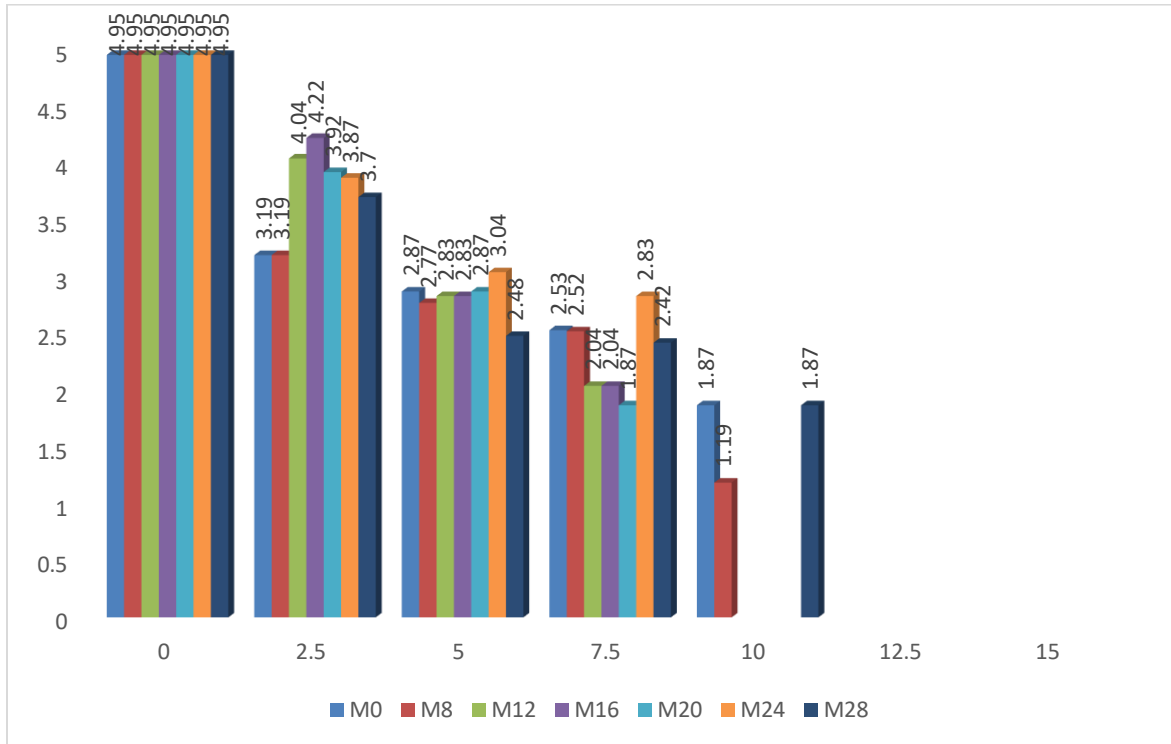
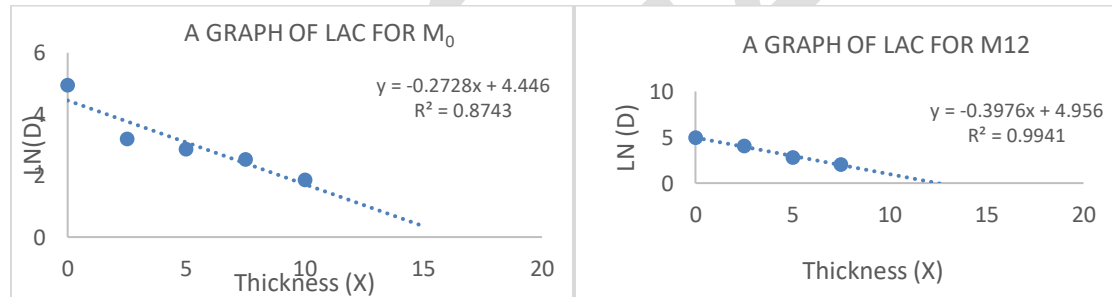
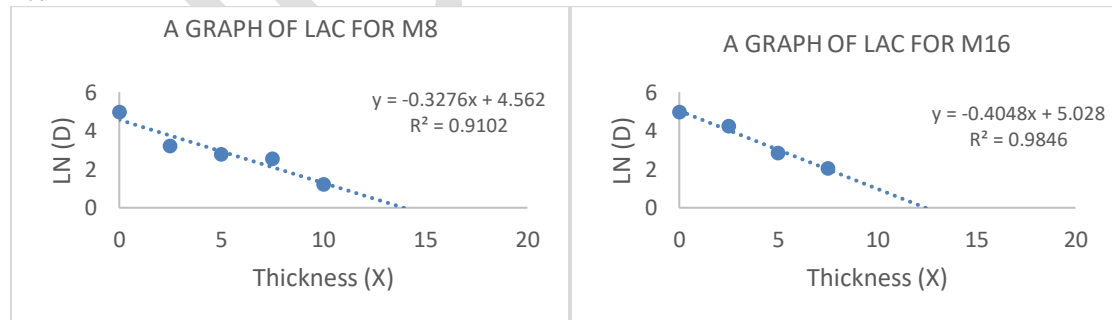


FIG 1: A chart of lnD against the thickness (x)

ab



cd



e

f

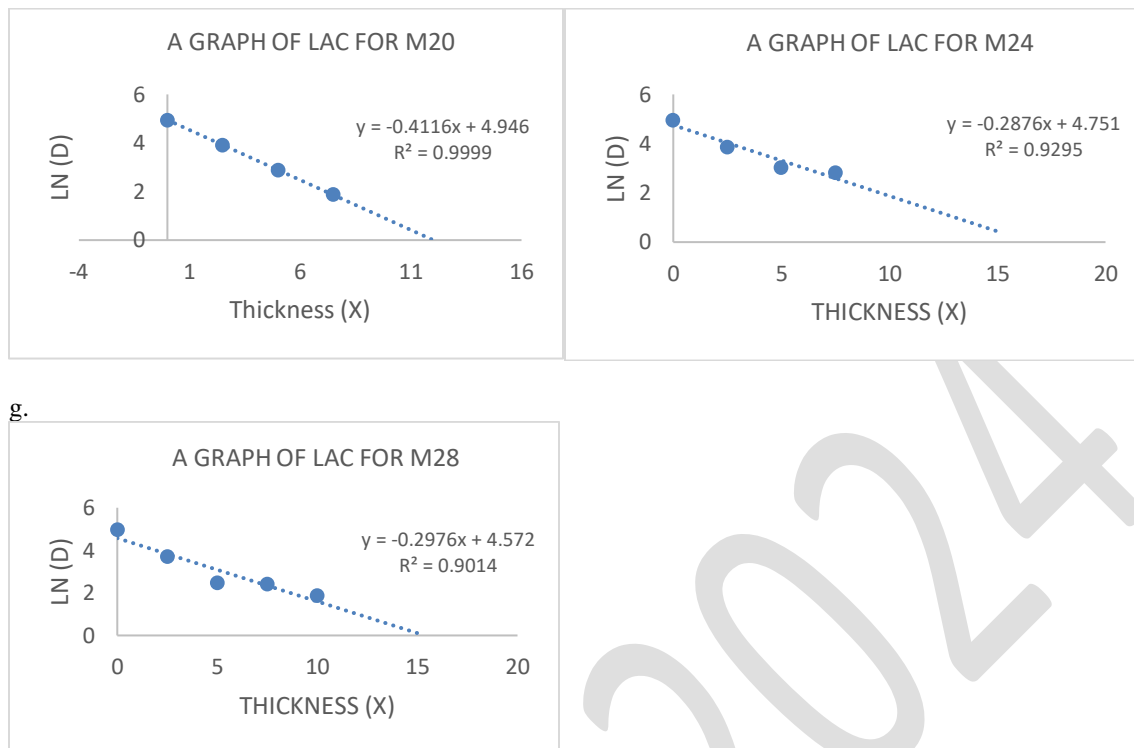


FIG. 2 (a to g): Graphs of $\ln D$ as a function of thickness (X)

Linear attenuation coefficient (μ) for mortar samples

The graphs (Fig. 6(a to h)) represent the μ for all the mortar samples at the initial dose of $141.2 \mu\text{Gy}$, this was obtained by plotting a graph of $\ln D$ against the thickness (x) of the samples. The linear attenuation coefficient represents the fraction of radiation that is absorbed or scattered per unit distance as it travels through a material expressed in cm^{-1} . The μ depends on the energy of the incoming radiation; higher-energy photons may penetrate more deeply into a material, resulting in a lower attenuation coefficient. It is also a fact that this depends on the additive of the mortar samples. The control sample had a μ of $0.27 \pm 0.02 \text{ cm}^{-1}$, Mortar sample with 20% CSA had the highest μ of $0.41 \pm 0.01 \text{ cm}^{-1}$, while other samples gave decreasing μ values of 0.33 ± 0.02 , 0.40 ± 0.01 , 0.40 ± 0.02 , 0.29 ± 0.01 and $0.30 \pm 0.03 \text{ cm}^{-1}$ for 8% CSA, 12% CSA, 16% CSA, 24% CSA and 28% CSA respectively. The mortar sample with 20% CSA showing the highest LAC can be considered to have greatest shielding properties among the mortar samples being studied in this research.

Table 2: Calculated values of μ , HVL and TVL for all mortar samples

MORTAR TYPE	$\mu \text{ (cm}^{-1}\text{)}$	HVL (cm)	TVL (cm)
M ₀	0.27 ± 0.02	2.54	8.44
M ₈	0.33 ± 0.02	2.12	7.03
M ₁₂	0.40 ± 0.01	1.74	5.79
M ₁₆	0.40 ± 0.02	1.71	5.69
M ₂₀	0.41 ± 0.01	1.68	5.59
M ₂₄	0.29 ± 0.01	2.41	8.01
M ₂₈	0.30 ± 0.03	2.33	7.74

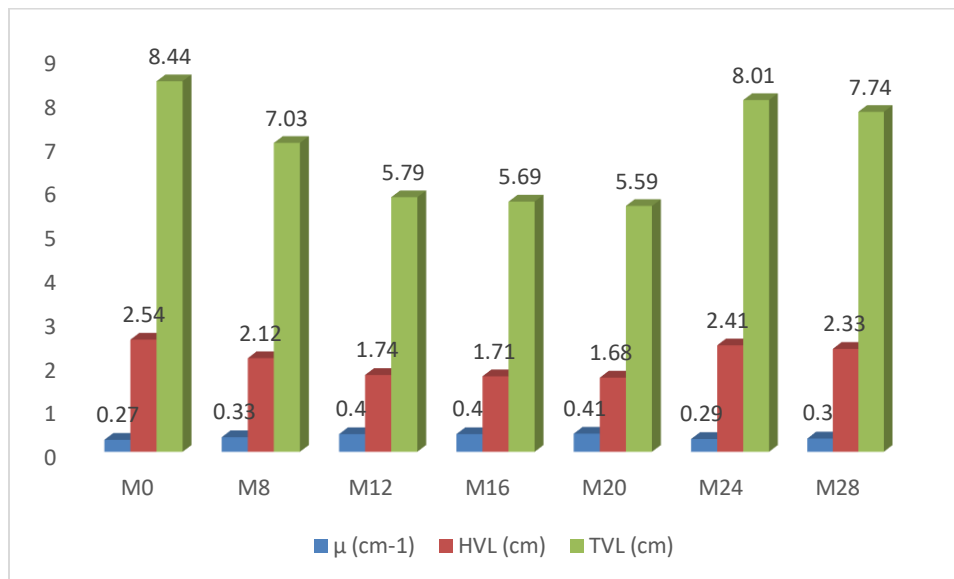


Figure 2: Comparison of μ , HVL and TVL for mortar control and other samples

4.4 Half-value layer (HVL) and Tenth-value layer (TVL) for mortar samples

The analysis below seeks to explain the HVL and TVL obtained for all mortar sample types presented in table 3 above. The HVL value in the table above represent the difference thicknesses of mortar type that will be needed to attenuate half the initial dose of 141.2 μ Gy, this was calculated following equation (18) and the result presented in Table 3. The analysis compares a control sample with the known HVL values against several percentages of CSA replacement, to evaluate their effectiveness in radiation protection. The result of HVL shown in table 3 above, indicates that the normal mortar will require the thickness of 2.54 cm to absorb half the initial dose of 141.2 μ Gy, which have a higher HVL compare to all other sample types, indicating that addition of CSA to replace the conventional cement for mortar construction are more effective at attenuating radiation. This also indicate that mortar sample constructed with the replacement of 20% of cement with CSA will require a thickness of 1.68 cm for the same dose, making it the lowest HVL suggesting it provides the highest level of radiation attenuation. Also 16% CSA replacement followed closely with HVL of 1.71 cm, while 12% CSA replacement showed moderate effectiveness with HVL of 1.74 cm higher than others. Others are 8% CSA, 24% CSA and 28% CSA with HVL of 2.12, 2.41 and 2.33 cm respectively are less effective compared to others but still provide better shielding properties than the control sample. From the HVL data, it indicates that 20% of CSA represent the optimal replacement percentage. This analysis indicate that the tested sample types significantly improve radiation attenuation when compared to the control sample.

The TVL was again calculated using equation (19) and result presented in Table 3 above. From the data presented in table 3 above, the thicknesses of each mortar sample are expected to reduce the intensity of a specified radiation by a factor of ten, as lower TVL indicates a higher attenuation capability, which is crucial in radiation shielding application. From the data presented in table 3, 12% CSA, 16% CSA and 20% CSA have significantly lower TVL values of 5.79 cm, 5.69 cm and 5.59 cm respectively, compared to the control sample with TVL value of 8.44 cm, indicating they provide better attenuation of radiation. The information presented suggest that among the best three samples, 20% CSA replacement has the lowest TVL of 5.59 cm, making it the most effective replacement among those tested in this study. The sample types 8% CSA, 24% CSA and 28% CSA have TVL values of 7.03 cm, 8.01 cm and 7.74 cm respectively, which are better than the control sample but not as effective as 12%, 16% and 20% CSA so tested. The results indicate that 12% CSA, 16% CSA and 20% CSA could be considered for approval as they provide improved protection compared to the control sample.

CONCLUSION

The outcome of the holistic investigation carried out in this research work has proven that replacing cement with CSA in different percentages during the construction of mortarsamples will also improve or reduce x-ray radiation shielding efficiency of the mortar. Built on the experimental processes and calculated result gotten in this study, it can be concluded without any doubt that 12%, 16% and 20% replacement of CSA with cement for mortar

construction improves the shielding efficiency of the mortar by reducing the porosity. It was discovered also that, increasing the percentage of CSA beyond 20% results in increasing the number of photons transmitted through the mortar samples. By utilizing CSA in the construction of mortar for shielding x-ray in diagnostic radiology facility will meaningfully reduce the cost of construction by reducing the quantity of cement used in construction. Based on the experimental observations carried out in this research, the following concluding remarks are made:

1. Not only that mortar is the least expensive used in x-ray facility as shielding material, mortar command other advantages making it more economical choice for large areas that require radiation shielding over lead. such as providing structural support, much lighter than lead, resistance to environmental factors like moisture and corrosion, less toxic, easily molded in various shapes, fire resistance and effective shielding with greater thicknesses allowing for the control of radiation exposure by simply adjusting depth of the mortar (Ling et al., 2013). These facts have been proven in this research work with 12%, 16% and 20% replacement of cement with CSA providing a suitable shielding material for diagnostic x-ray facility.
2. From the analysis gotten in the experimental work of this research on half-value layer (HVL), it has proven that, mortar sample with 20% CSA requires the least thickness of 1.68 cm to absorb half the incident dose of 141.2 μ Gy. This analysis indicate that the tested sample types significantly improve radiation attenuation when compared to the control sample.
3. The calculated results for the tenth value layer (TVL), shows that the mortar samples with 12% CSA, 16% CSA and 20% CSA, having TVL of 5.79, 5.69 and 5.59 cm respectively, could be considered to provide improved radiation protection compared to the control sample with TVL of 8.44 cm.

References

- Aljabal, A. F., Wargo, R. R. & Lin, P. P. (2019). Evaluation of lead equivalence of radiation protection apparatuses as a function of tube potential and spectral shaping filter. *Journal of applied clinical medical physics/ vol 20*. <https://doi.org/10.1002/acm2.12768>
- Ketebu, O. & Farrow, S. T. (2017). Comparative study on cementitious content of ground mollusk snail and clam shell and their mixture as an alternative to cement. <http://doi.org/10.14445/22315381/IJETT-V50P202>
- Nduka, D. O., Akanbi, E. T., Ojo, D. O., Babayemi, T. E., & Jolayemi, K. J. (2023). Investigation of the mechanical and microstructural properties of masonry mortar made with seashell particles. *Materials*, 16(6), 2471.

PRODUCTION AND CHARACTERISATION OF MICROALGAE BIO-OIL: AN ENERGY TRANSITION APPROACH AND ALTERNATIVE FOR FOSSIL FUEL

EBRI, Robert Utum

Email: ebrirobertjay@gmail.com

Abstract

Bio-oil from microalgae has attracted a lot of attention for its extraordinary raw material utilization advantages and environmentally friendly characteristics. In order to increase the renewable energy utilization efficiency of microalgae biomass, research on the production line of microalgae cultivation and pyrolysis is essential. In this study the production of bio-oil from microalgae was carried out through a pyrolysis process. The method used in the production of bio-oil from Microalgae in this study is the pyrolysis method where a fix-bed reactor was connected to a condensers and heat was applied from a gas burner and the temperature was monitored using a thermocouple connected to the reactor and time monitored by a stop watch until the last drop of product was observed from the system. The result shows that both oils produced from microalgae are heavy oils with API gravity of 8.59, heavy oil with density of 1.079 g/cm^3 at 22°C , high viscosities of $8.24 \text{ mm}^2\cdot\text{s}^{-1}$, specific gravity of 1.01, high flashpoint of 96°C , pour point of 16°C . GC/MS analysis carried out on the produced bio-oil Samples shows that the bio-oil produced from Microalgae contain several compounds. The bio-oils of microalgae contain carbon chain compounds (C8-C36). It was concluded that bio-oil produce microalgae is a potential alternative for fossil fuel. It was recommended in this study that effort should be made to convert the laboratory scale production of microalgae to industrial scale as it will be a good and possible alternative for fossil fuel.

Keywords: Bio-Oil, Biomass, Duckweed, Micro-Algae

Introduction

Energy security is a significant challenge for a sustainable economy, necessitating the development of alternative renewable energy sources capable of meeting future demands (Asif, 2007). Since the discovery of non-renewable fuels such as petroleum-based fuel over four thousand years ago, all of humanity has relied on them (Dunn, 2002). Scientists are looking for clean and renewable fuel to replace petroleum-based fuel due to rising global demand, limited supply, and environmental impact. Bio-oil has loomed as alternative energy and a potential hydrocarbon substitute. Sustainable bioenergy production is gaining popularity as a way to minimize dependency on fossil fuels and reduce greenhouse gas emissions. This study is mainly to produce and characterize bio-oil from microalgae.

Materials and Method

Materials

Microalgae was harvested (Fig. 1) from stagnant water in Diobu area, Port Harcourt, Rivers State. The materials used for the study are pyrolytic system, heater, Thermocouple, condenser, water for cooling, measuring cylinder to receive the oil, weighing balance, retort stand, beaker, separating funnel, round bottom flask, gas burner and a single-shot pyrolyzer connected to a GC-MS (Agilent 7890A/5975C, USA). The heat of reaction was supplied from a gas burner. The cooling water used was at 26°C . The condensed oil was received in a measuring cylinder to monitor the rate of product formation. A thermocouple was connected to the system to monitor the temperature in the system and a stop watch was used to measure the reaction time. The pyrolysis set-up, harvested microalgae is hereby presented (Fig.2).



Figure 1: Harvested microalgae sample



Figure 2: Pyrolysis Setup

Method

The harvested microalgae was air dried (Fig. 3) for 14 days to remove the moisture, the dried microalgae of 5.5kg was weighed and passed through the hopper of the reactor. The hopper was then properly covered with thread seal tape to avoid leakage. Adequate precautions were put in place to make sure there is no leakage before the start of the experiment. A glass condenser was connected tightly to the reactor to cool the condensing vapor from the reactor. Water at 26°C, connected counter currently was used to cool the vapor. No catalyst was used. The gas burner was then put on and the pyrolysis continued until the last drop of oil was noticed in the measuring cylinder. The volume of bio-oil produced was monitored with time and temperature. The bio-oil produced was immediately analyzed and the result was recorded. The pyrolysis was also carried out at various temperatures to determine the effect of temperature and time on the products formed.



Figure 3: Dried microalgae sample

Bio-oil analysis

The following analysis was carried out on the bio-oil; oil density, specific gravity, API gravity, pH measurement, viscosity, flash point, pour point.

Chemical Composition

The chemical compositions of the engendered bio-oil were analyzed using gas chromatography (GC) (CHEM32) and Mass Spectrometer (MS). The carrier gas was helium at a flow rate 1.0 mL min⁻¹, a temperature of the injection port of 270°C, the injection volume is 1 μ L. The starting temperature of the analysis was be at 40°C, the heating rate was kept to 1.5°C min⁻¹ to 46°C, and the heating rate will subsequently increase to 4°C min⁻¹, until the temperature reached 209°C. Real-time chemical composition of pyrolysis sample was analyzed using a single-shot pyrolyzer connected to a GC-MS (Agilent 7890A/5975C, USA), which was equipped with an inert XL mass spectrum detector and a capillary column (30 m in length, 0.25 μ m in internal diameter, HP-5 MS). Concentration level and percentage of different compounds in the pyrolysis products of the micro algae was determined by online Py-GC/MS under a ramping temperature gradient from 350°C to 400°C at a temperature interval of 100°C.

The total ion count (TIC) diagrams of micro algae pyrolysis products under different temperature conditions were obtained. Results were analyzed using Agilent MSD Productivity Chem Station for GC and GC/MS System Data Analysis application software. Retention time and peak area percentages of different compounds in pyrolysis products were determined by comparing with NIST 2011 Database. The concentrations of each compound were of right proportion to its corresponding peak area percentage.

Results and Discussion

Pyrolysis of bio-oil produced from microalgae

Result of bio-oil obtained from pyrolysis of microalgae is hereby presented (Fig. 4). At 20°C-80 °C, no product was formed from the pyrolysis process, the first drop of product was observed at temperature of 100 °C. It was observed from the experiment that as the temperature of the reaction increased from 100 °C-200 °C, the volume of the bio-oil produced also increased, also as the heating time was increasing, from 10 minutes to 150 minutes, the product yield also increased. The bio-oil yield of temperature 100-400 °C was analyzed because at this temperature, all the bio-oil has been extracted from the biomass. The mass of microalgae char formed by the pyrolysis process is 1.2kg.

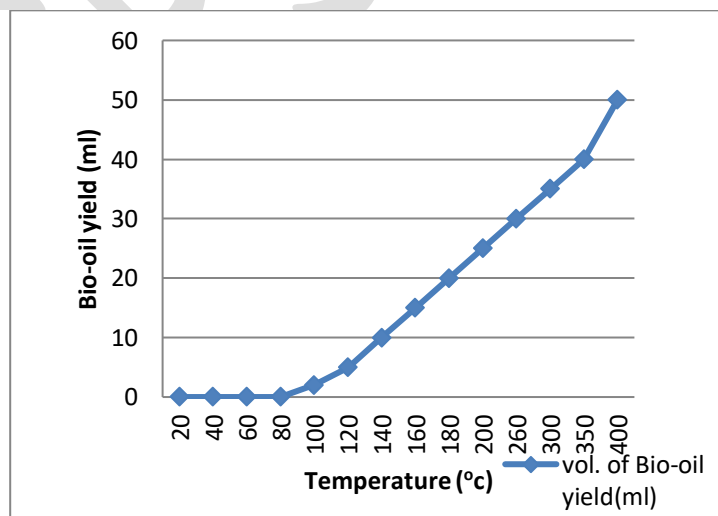


Figure 4: Effect of temperature on bio-oil yield from microalgae.

The heating time of the reaction was observed during the production of bio-oil from pyrolysis of microalgae. It was discovered that from time 0mins – 100mins heating time, no product was formed from the pyrolysis process, the

first drop of product was seen at heating time 150min, 0the volume of product formed from the pyrolysis process 5 mins after the first drop was 2ml (Fig. 5).

The bio-oil yield also increases when the system is heated for a long time, more bio-oil product is formed from the experiment. It was seen from the experiment that as the heating time of the reaction increased from 150 mins - 420 mins, the volume of the bio-oil produced also increased. The bio-oil yield of heating time 420 mins was analyzed because at this time, all the bio-oil has been extracted from the microalgae.

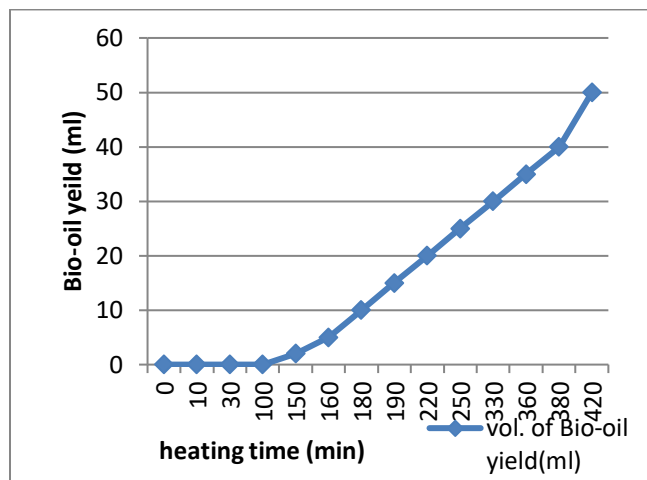


Figure 5: Effect of Heating Time on Bio-Oil Yield from Microalgae.

Characteristics of Bio-oil

Characteristics of Bio-oil produced from Microalgae

Table 1 showed the physical properties of the bio-oil produced, which were determined using American standard of testing methods and standard charts for fuel oils. The API gravity at 22°C of the produced bio-oil was 8.59 which indicated that the produced bio-oil from microalgae is a very heavy oil. pH value of the bio-oil from microalgae is 5.3. pH exceeding a value of 3.0 was found only for few biomass feedstocks. Bio-oil viscosity ($\text{mm}^2 \cdot \text{s}^{-1}$; at 25°C) of the produced bio-oil was 8.24, which are relatively higher compared to other biomass sources. The pour point was 16°C which were in range with the pour point of waxy crude oil. A high pour point is generally associated with high paraffin content, typically found in crude deriving from a larger proportion of plant material according to [ASTM Standard D5949 \(2003\)](#). The flash point of the bio-oil was 32°C which is the range of flash point for heavy crude oil. Volatile matter and the fixed carbon of the produced bio-oil are 0.6 and 86.56 respectively.

Table 1: Physical Properties of the Produced Bio-Oil from Microalgae

Physical properties	Quantity	Unit
Density at 22°C	1.079	g/cm^3
Specific gravity	1.01	-
API gravity at 22°C	8.59	-
Viscosity	8.24	$\text{mm}^2 \cdot \text{s}^{-1}$
pH	5.3	-
Pour Point	16	°C
Flash Point	96	°C
Volatile matter	0.6	%
Fixed carbon	86.56	%

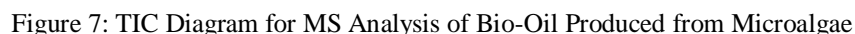
Chemical Composition

GC/MS Analysis of bio-oil produced from Micro-algae

The result of the GC/MS analysis is hereby presented (Table 2), and it showed that the retention time, and chemical composition of the samples. GC-MS analysis was carried out to determine the percentage of major compounds present in the pyrolysis products derived from the algae sample, which would provide important references for biomass thermochemical conversion and utilization in industry. TIC diagram of micro algaepyrolytic bio-oil obtained is also presented (Fig. 6& 7).

Table 2: GC results for retention time and component of bio-oil produced from microalgae

Rettime (min)	Type	Area (pA*s)	Amt/Area	Amount (ppm)	Names
2.933	HT	374.2985	9.99326e-2	748.09291	n-C8
4.243	HT	1373.77136	2.29512e-1	6305.94972	n-C9
5.621	HT	522.63556	1.21694e-1	1272.03222	n-C10
6.810	HT	794.83447	7.27325e-1	1156.20646	n-C11
7.635	HT	49.24685	1.37204e-1	135.13773	n-C12
8.243	HT	31.42998	1.1295e-1	71.00215	n-C13
9.205	HT	64.33518	3.04704e-1	392.06406	n-C14
10.074	HT	43.94883	2.75023e-1	241.73856	n-C15
11.052	HT	23.76081	2.60529e-1	123.80768	n-C16
11.780	HT	574.99908	4.16898e-1	4794.32195	n-C17
11.832	HT	89.41660	1.67242e-1	299.08366	Pristine
12.520	HT	24.44940	2.63666e-1	128.92949	n-C18
12.559	HT	68.54385	2.03913e-1	279.54033	Pristine
13.176	HT	5.70851	0.00000	0.0000	n-C19
13.245	-	-	-	-	o-Terphenyl
13.780	HH	189.41678	2.65504e-1	1005.81877	n-C20
14.403	BB	314.24121	2.74162e-1	1723.06103	n-C21
14.987	BB	440.72604	2.79295e-1	2461.85574	n-C22
15.555	HB	105.64922	2.71756e-1	574.21616	n-C23
16.096	HB	110.84991	3.14571e-1	697.40394	n-C24
16.621	BB	85.73927	2.92271e-1	501.18194	n-C25
17.203	BB	1.39754	0.00000	0.00000	n-C26
17.613	HH	89.23038	3.30521e-1	589.85076	n-C27
18.248	BB	9.68134	2.04437e-1	39.58449	n-C28
1.531	BB	43.83989	3.62930e-1	318.21661	n-C29
19.053	BB	8.464458	3.44720e-1	58.35820	n-C30
19.481	BB	3.29964	0.00000	0.00000	n-C31
19.904	BH	1.57982	6.49904e-2	2.05347	n-C32
20.200	BB	29.81846	2.10833e-1	125.73458	n-C33
20.590	HH	15.18063	1.70292e-1	51.70275	n-C34
20.978	BB	10.61378	7.08162e-2	15.03256	n-C35
21.409	BB	5.19605	1.09260e-1	11.35437	n-C36
21.570	BB	6.98663	1.18251e-1	16.52350	n-C37
22.563					n-C38
23.234					n-C39
24.013					n-C40
Total				2.41399e4	



After the MS analysis carried out on the sample, it can be concluded that the pyrolyzed products of microalgae at 350°C mainly contained 17 kinds of substances such as acids, hydrocarbons, alcohols and other organic substances. Therein, the substances whose content was found are the system monitoring compound.

Table 3: MS results for retention time and component of bio-oil produced from Microalgae

S/n	Compound	R.T.	Qion	Response	Conc	Units	Dev. (Min)
7	System Monitoring Compounds	6.822	230	1186	0.17	ppm	-0.04
1	Target Compounds						Qvalue
1	Naphthalene	3.097	128	13141281	9933.23	ppm	100
2	Acenaphthylene	4.599	152	24456	255.76	ppm	100
3	Acenaphthene	5.770	153	301263	636.43	ppm	100
4	Fluorine	5.334	166	341295	886.48	ppm	100
5	Phenanthrene	6.464	178	190959	615493	ppm	100
6	Anthracene	6.547	178	95489	298.98	ppm	100
8	Fluoranthene	7.793	202	13052	62.78	ppm	100
9	Pyrene	8.185	202	51047	228.84	ppm	100
10	Benzo (a) anthracene	9.625	228	1868	7.99	ppm	100
11	Chrysene	9.625	228	1999	8.37	ppm	100
12	Benzo (b) fluoranthene	10.738	252	2628	13.46	ppm	100
13	Benzo (k) fluoranthene	10.836	252	2032	10.09	ppm	100
14	Benz (a) pyrene	11.135	252	21	0.17	ppm	100
15	Dibenz (a,h) anthracene	12.309	276	876	13.13	ppm	100
16	Indeno (1,2,3-cd) pyrene	12.309	276	874	15.37	ppm	100
17	Benzo (g,h,i) pyrene	12.612	276	279	4.32	ppm	100

CONCLUSION

It was concluded that Microalgae are potential useful biomass for large scale bio-oil production. GC/MS analysis conducted on the produced oil samples shows that the bio-oil produced from microalgae contain 50 different compounds such as hydrocarbons, alcohol and acids.

NOMENCLATURE

GC- Gas Chromatography

MS- Mass Spectroscopy

R.T- Retention Time

TIC- Total Ionic Carbon

PAH- Polyaromatic Hydrocarbon

REFERENCES

- Asif, M. & Muneer, T. (2007). Energy supply, its demand and security issues for developed and emerging economies. *Renewable and sustainable energy reviews*, 11(7), 1388-1413.
- Dunn, R. O. (2002). Effect of oxidation under accelerated conditions on fuel properties of methyl soyate (biodiesel). *Journal of the American Oil Chemists' Society*, 79, 915-920.
- Zhou, Y., Xing, X., Lang, J., Chen, D., Cheng, S., Wei, L., & Liu, C. (2017). A comprehensive biomass burning emission inventory with high spatial and temporal resolution in China. *Atmospheric Chemistry and Physics*, 17(4), 2839-2864.

AN INTELLIGENT NETWORK INTRUSION DETECTION SYSTEM USING MULTILAYERED TECHNIQUE

B. I. Ele, O. E. Ofem & D. O. Egete

Abstract

Network users, administrators, and security professionals continue to express grave concern over the slowness, accuracy, and high false alarm rates of current network security systems. As a result, immediate action is required to address these issues. Therefore, the goal of this study is to develop a more effective model of a network intrusion detection system for virtual local area networks using a multi-layered technique with radial basis functions (RBF) and support vector machines (SVM). The structured system analysis and design methodology (SSADM) was used in this study to design the security system. The designed system was successfully implemented using JavaScript and Python programming languages with the vue.js, node.js, and express.js frameworks and tested with the NSL-KDD dataset with virtual local area network. The developed system has a detection accuracy of 99% as compared to the existing individual RBF and SVM systems with 76% or 87% respectively. This system's ability to accurately and quickly identify network-based attacks will be crucial in limiting the actions of intruders. This study's outcome was an improved NIDS that would proactively address potential security vulnerabilities by reliably and effectively detecting attacks and security policy violations in virtual local area networks. As such, its use in the 21st-century security-conscious environment is inevitable.

Keywords: Intelligent, Network, Intrusion Detection System, Radial Basis, Function, Support Vector Machine, Multilayered technique

Introduction

Network Intrusion Detection Systems (NIDS) are assigned the critical role of monitoring the security state of the network (Eleet *et al.*, 2016); therefore, the NIDS itself is a primary target of attack. The NIDS must be able to operate in a hostile computing environment and exhibit a high degree of fault-tolerance which allows for a graceful degradation. The strength of an intrusion detection system (IDS) is dependent on the robustness of the feature selection method (Joseph *et al.*, 2023). IDS can be host or network-based subject to the source of data for attack detection and also signature or behavior-based subject to the method of intrusion identification (Ele&Mbam, 2014). Humans walking out with data on a memory stick or sharing proprietary information with social engineering hackers are also prone to insecurity, whereas network security, being a subset of cybersecurity, covers what that user does on the network itself. According to Ele&Mbam (2014), various sources of security threats for any network system are authenticity, access control, confidentiality, integrity, availability, and non-repudiation. Intrusion detection techniques are continuously evolving, with the goal of improving the security and protection of networks and computer infrastructures. Despite the promising nature of multi-layered based intrusion detection system, as well as its relatively long existence, there still exist several open issues regarding security.

Network intrusion detection systems (NIDS) exist in varying forms and have several limitations on their performance and effectiveness. These limitations arise from problems associated with the traditional placement of the Network intrusion Detection Systems within any given network infrastructure. Despite the deployment of advanced security methodologies, there are still several loopholes that have not been filled. Data theft in industries is growing all over the world. Cybercrime is depressing trade and investor confidence on companies. Apart from economic loss, internet fraud, credit card fraud, insurance fraud, tax evasion, financial fraud,

securities fraud, insider attacks, money laundering and embezzlement, as well copyright and trade secret theft are constituting major problems to network systems. While it is not possible to keep industries completely from data theft, steps can be taken to protect industries from these potential harm and attacks. However, transmission of information over such networks can be compromised and security breaches such as virus, denial of service, unauthorized access prevail. This research proposes an effective multi-layered technique for augmenting the functionalities of network security technologies due to the fact that, irrespective of the type of access control being employed, attacks are still bound to occur.

Several researches with respect to intrusion detection systems have used machine learning, data mining, decision tree, neural networks, clustering, and Bayesian parameter estimation techniques to detect any intrusive performances in the computer network (Kruegelet al., 2013; Wu et al., 2013; Bhavsar&Waghmare, 2013; Amor et al., 2014; Panda &Patra, 2017; Ektefaet al., 2020).

The aim of this research is to develop an Intelligent Network Intrusion Detection System using a multi-layered approach that can detect and prevent intrusions and to integrate same on networks to secure them. This will provide a solution that can make the networked systems of industries secured from any potential threats and attacks.

Methodology

Both the prototyping methodology and the structured system analysis and design methodology (SSADM) were employed. Both methodologies were employed to bring out detailed description of the system as well as providing avenue for easy modification of the system as the need may arise in future and produce effective and efficient system. SSADM is suitable for analysing and designing large systems like the Multi-Layered Network Intrusion Detection System (MLNIDS) as it gives out a clearer view and representation of the modules, procedures, and functions with their respective relationships, and the representation of the objects (data and processes) as contained in the MLNIDS, as such giving the designers a complete analysis for the development of efficient system that meet specifications as contained in the specification documents. Prototyping methodology was adopted because of its suitability for building fast, better, cost effective, more reliable, and better quality system such as MLNIDS.

Design of the New System

The new system is designed using following tools: algorithm, flowchart, use case diagram, architectural framework, and dataflow diagram.

Algorithm for the RBF + SVM model

1. RBF Network Preprocessing:

- i) Use an RBF network as a preprocessing step to transform the input data into a higher-dimensional space where it is more linearly separable; and
- ii) Train the RBF network using the data and adjust its parameters such as number of neurons and spread to effectively capture the underlying patterns in the data.

2. Feature Extraction:

- i) Extract features from the output of the RBF network. These features represent the transformed data in the higher-dimensional space.

3. SVM Training:

- i) Use the extracted features as input to train an SVM classifier; and
 - ii) Choose an appropriate kernel for the SVM, such as linear, polynomial, or RBF kernel, depending on the nature of the problem and the characteristics of the data.
- 4. Model Evaluation:
 - i) Evaluate the combined RBF-SVM model using standard evaluation metrics to assess its performance on a validation dataset.
 - ii) Fine-tune hyper-parameters as needed to optimize performance.
- 5. Detection:
 - i) Given new input data, preprocess it using the trained RBF network.
 - ii) Extract features from the RBF-transformed data.
 - iii) Use the SVM classifier to predict the class label or outcome for the input data.
- 6. Post-processing:
 - i) Apply any necessary post-processing steps, such as threshold adjustment or probability calibration, to refine the model's predictions.

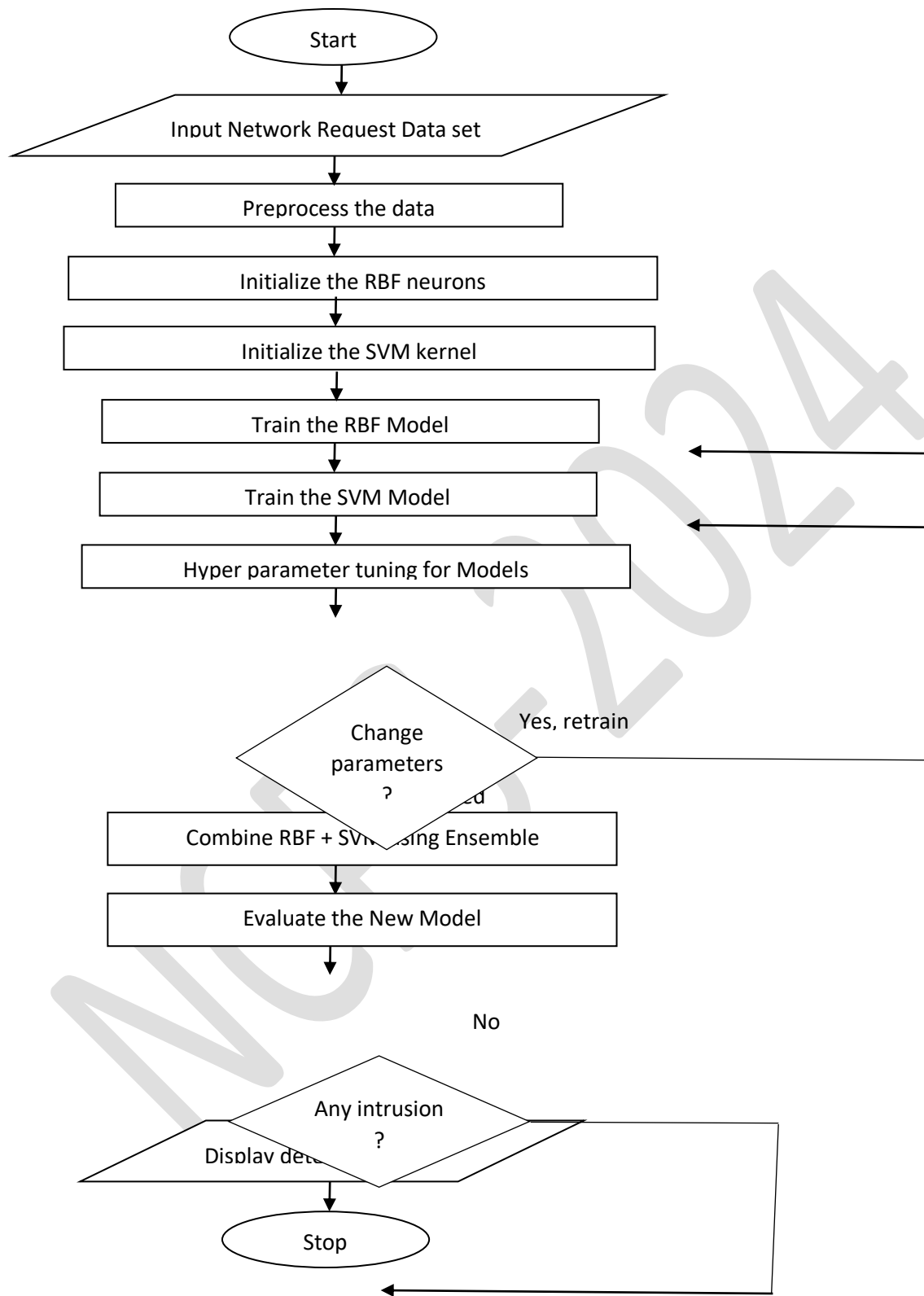


Figure 1: Flowchart of the New Security System

The use of case diagram (Fig.2) indicate that the user (attacker)monitor the network performance and send network request to the server which is then intercepted by the proxy-server to determine whether or not it is a valid request. The admin monitors the users to see where malicious request originates and can then decide to block or filter such a user.

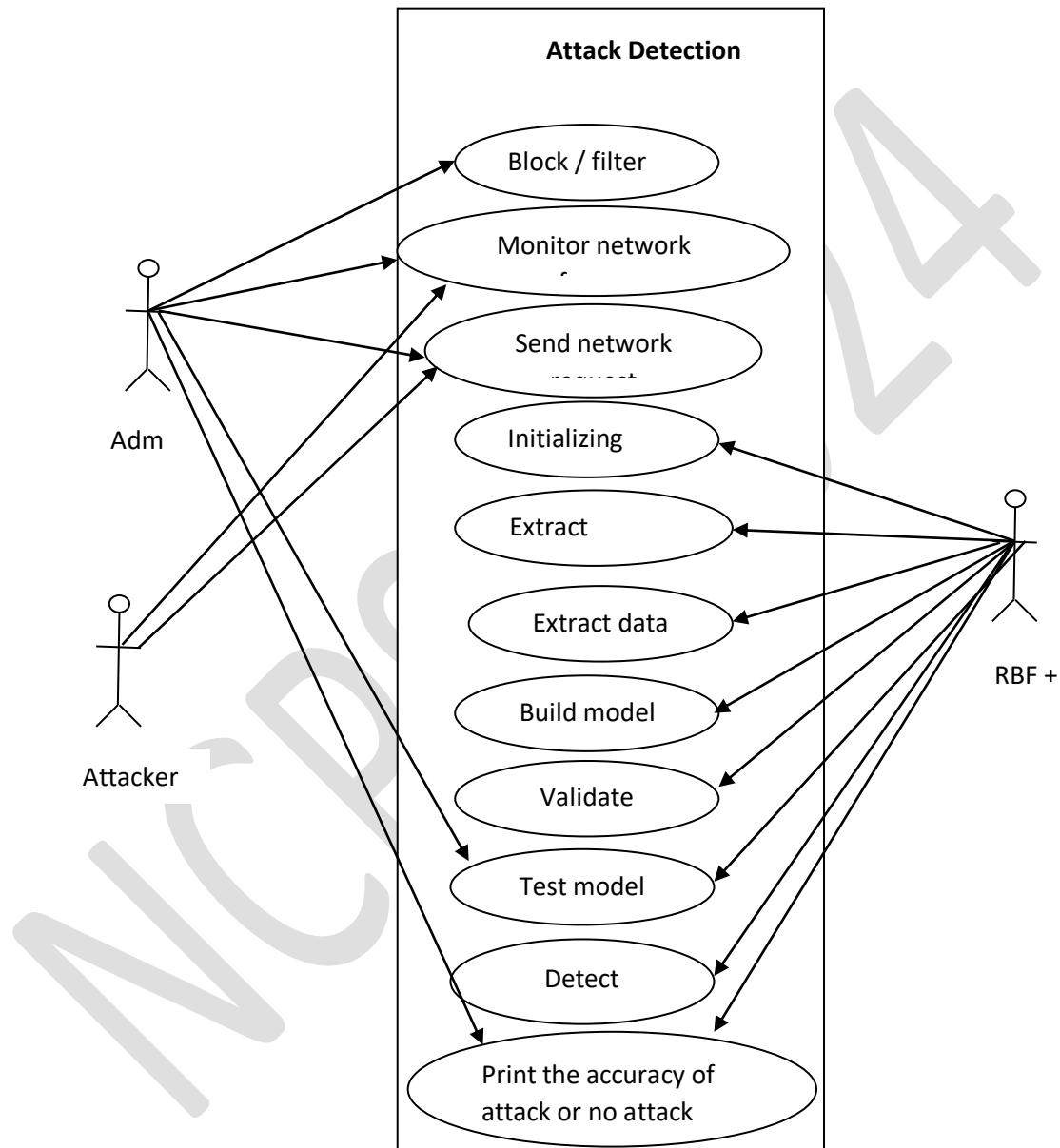


Figure 2: Use Case Diagram of the New System

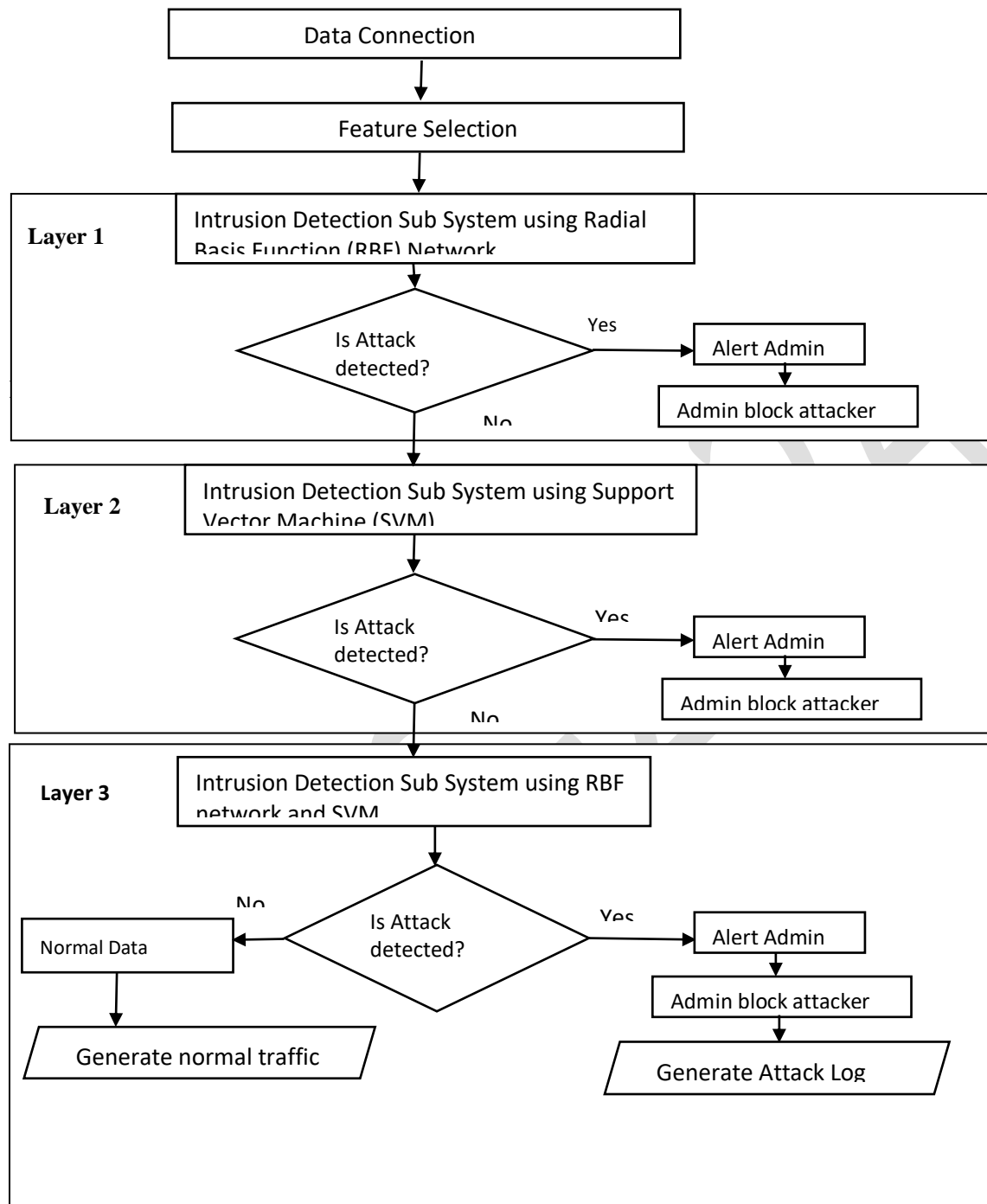


Figure 3: Architecture of the New System

Figure 3 represents a “3” layer system where each layer in itself is a sub intrusion detection module using different machine learning technique which is specially trained to identify intrusions, for instance the denial of service (DoS)attacks remote to user, user to root and probing attacks. These sub-systems are then deployed consecutively in sequence, which will help to detect attacks effectively.

Figure 3: Architectural Framework of the New Security System

web application starts by showing a list of all the network cards present in the target device and allows one to monitor network traffic based on specific network device (Fig. 7). In order to test out different attacks for the purpose of experimentation and evaluation, there is a provision in the code that enables a network administrator launch an attack on the target system. Observe that in displaying the attacks, the system also displays the time, IP address and type of attacks.



Figure 5: Main Interface of the New System

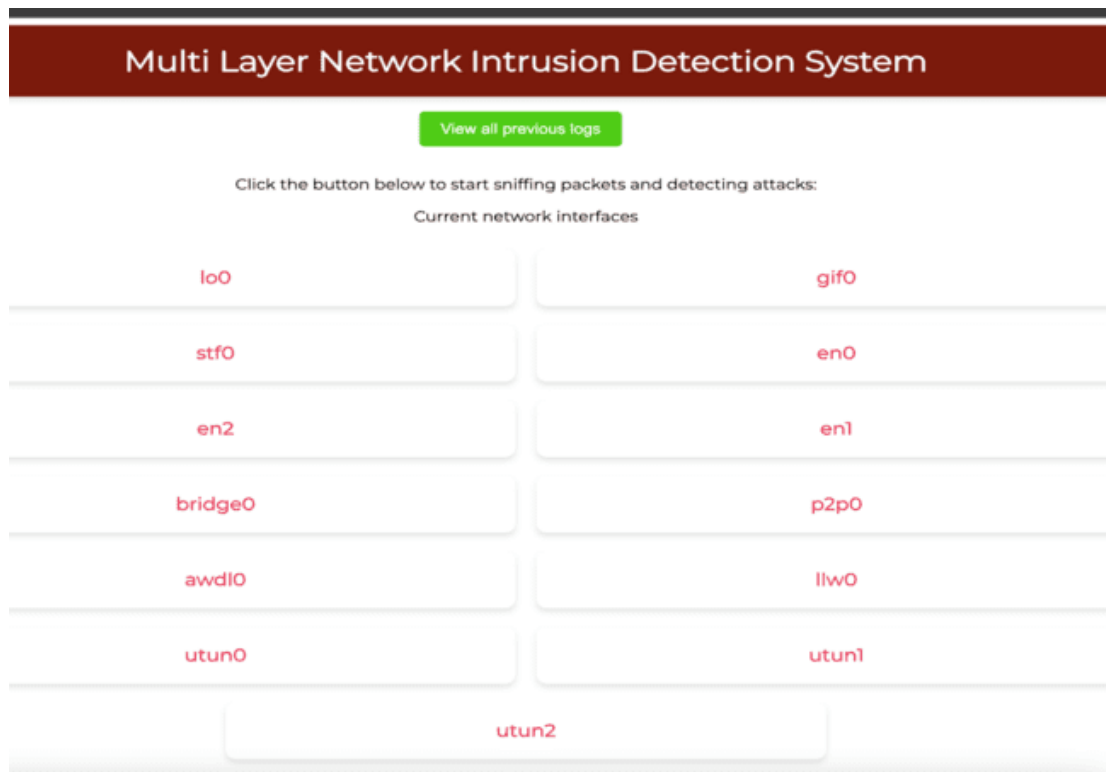


Figure 6: Interface showing list of all network cards

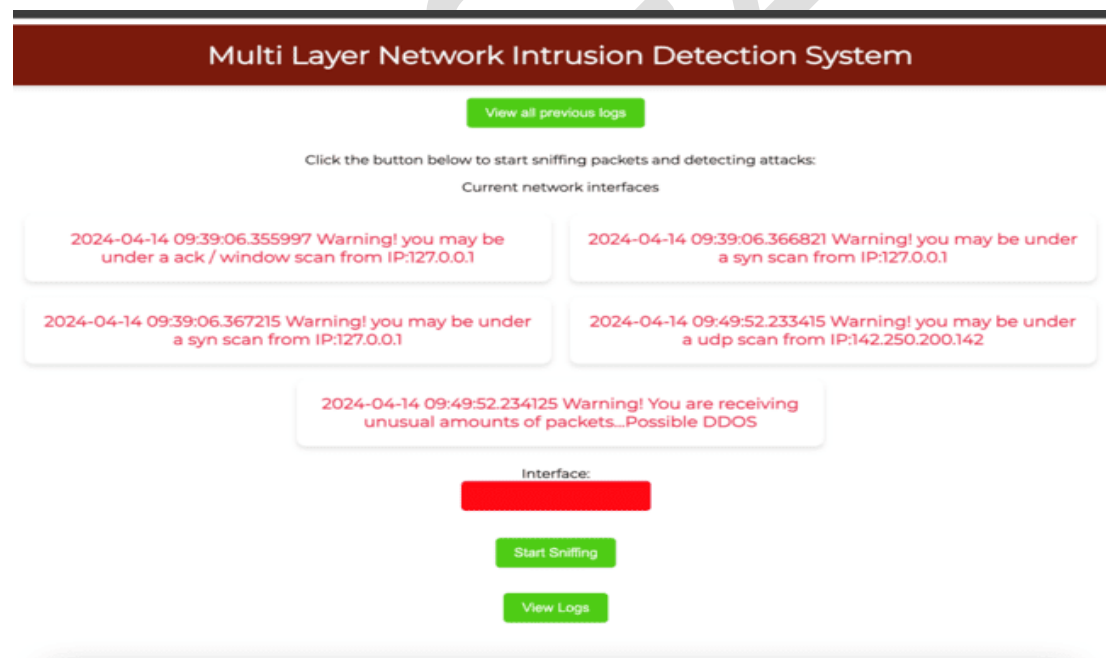


Figure 7: An interface showing the date/time, type of attack and IP address

The Result of the machine learning run (as well as the Web Graphical Interface output) can be seen in Figs. 5, 6, 7, and 8.

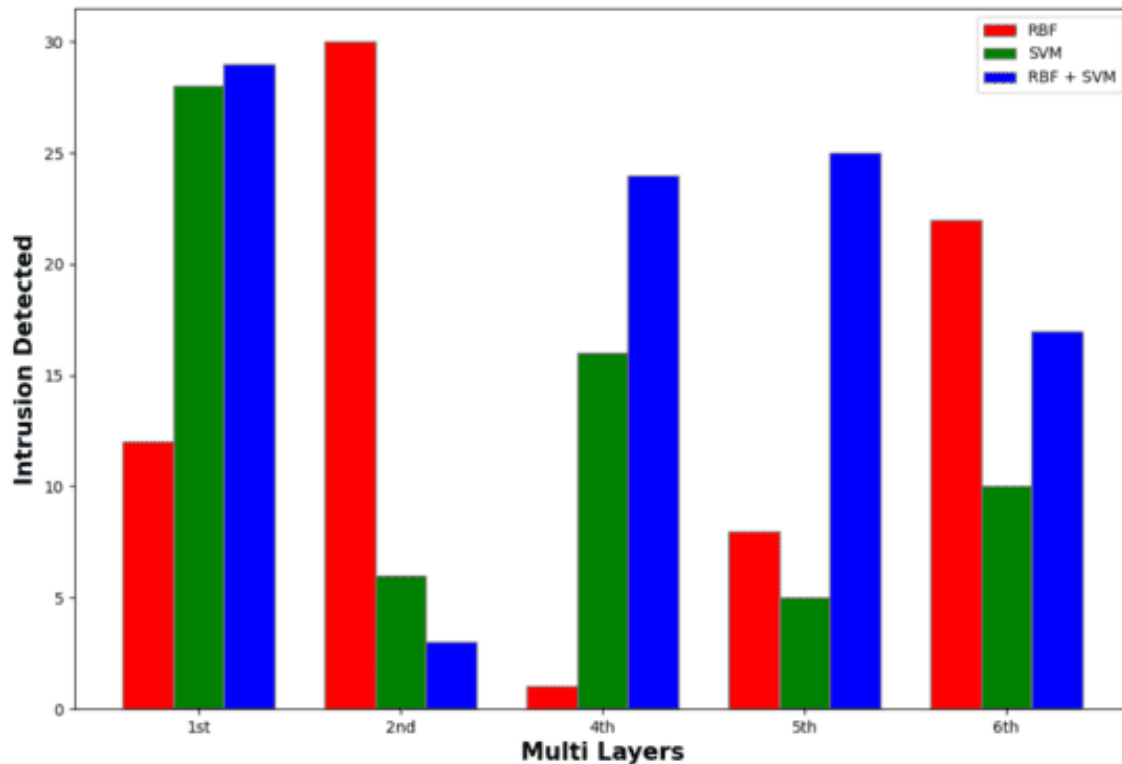


Figure 8: A bar chart showing detection level of the various techniques

The bar chart in figure 8 shows in groups of bars how the RBF, SVM, and RBF + SVM algorithms detect network intrusion. As showed in each bar group, the Blue represents the combined algorithm while the Red and Green represents RBF and SVM respectively. The bar chart shows that the combined algorithm is more effective in detecting network intrusions compared to individual RBF and SVM algorithms.

9.2 Discussion

Based on the network services included in the NSL dataset, the results in this study was calculated and examined using the new multi-layer model (RBF+SVM). The Bar charts show that the bulk of daily network usage uses the http(s) network protocol, which contributes to a significant amount of malware infections. A close examination of the charts also reveals that assaults are also being made against other types of network traffic. This suggests that several avenues are being looked for by attackers to access a user's system (some of which succeed, some of which fail).

From a network administrator's standpoint, the combination of protocol, flag, and service ought to provide a wealth of information about the type of traffic we have. However, we can better understand how the new multi-layer machine learning model anticipated the assaults across the various types of network traffics if we additionally take into account the length of a connection and the volume of data in that connection. In the course of the analysis, it was discovered that the multi-layer approach to network intrusion detection makes it easier to identify and isolate various kind of network intrusions.

Based on the evaluation of the system, the radial basis function (RBF) has a detection accuracy of 87%, the support vector machine (SVM) has a detection accuracy of 76% and the hybrid model (RBF+SVM) has a detection accuracy of 99%, which implies that the new system perform better compared to the individual RBF and SVM.

10. Conclusion

This study focused on the development of a multi-layered network intrusion detection system (MLNIDS) for virtual local area networks. In this study, the suitability of radial basis functions (RBF) network, support vector machine (SVM) and layered framework for building robust and efficient model of intrusion detection system for virtual local area networks was examined. In particular, layered framework was introduced and a multi-layered network intrusion detection system was developed and implemented which addresses the critical problems identified in section 2 that severely affect the large scale deployment of present intrusion detection systems in virtual local area networks.

The study observed that layered framework can be used to build efficient intrusion detection systems. In addition, the framework offers ease of scalability for detecting different variety of attacks as well as ease of customization by incorporating domain specific knowledge. The framework also identifies the type of attack, hence, specific intrusion response mechanism can be initiated which helps to minimize the impact of the attack.

In this study, the multi-layered approach was compared with some well-known methods and found that most of the present methods for intrusion detection fail to reliably detect denial of service attacks, root to local attacks and user to root attacks, while the integrated system developed in this study can effectively and efficiently detect such attacks. The developed system can help in identifying an attack once it is detected at a particular layer, which expedites the intrusion response mechanism, thus minimizing the impact of an attack. Finally, the developed system has the advantage that the number of layers can be increased or decreased depending upon the environment in which the system is deployed, giving flexibility to the network administrators and security professionals.

References

- Amor, N. B. Benferhat, S. and Elouedi, Z. (2014). Naive Bayes vs. Decision Trees in Intrusion Detection Systems. *Proceedings of ACM Symposium on Applied Computing (SAC '14)*, 420-424.
- Bhavsar, Y. B., &Waghmare, K. C. (2013). Intrusion detection system using data mining technique: Support vector machine. *International Journal of Emerging Technology and Advanced Engineering*, 3(3), 581-586.
- Ektefa, M., Memar, S., Sidi, F., &Affendey, L. S. (2020). Intrusion detection using data mining techniques. In:*Information Retrieval & Knowledge Management, (CAMP), 2010 International Conference on IEEE*, 200-203.
- Ele, B. I., Alo, U. R., Mbam, B. C. E. and Ofem, A. O. (2016). A Model of a Pragmatic Secure Intrusion detection System for Local Area Networks. *British Journal of Mathematics and Computer Science*. 13(2), 1 – 15.

- Ele, B. I. and Mbam, B. C. E. (2014). Development of a Layered Conditional Random Field Based Network Intrusion Detection System. *West African Journal of Industrial and Academic Research*, 12(1), 3-20.
- Joseph, B. A., Femi, E. A., Ranjit, P., Amik, G., Akash, K. B. & Paolo, B. (2023). A Multi-level Random Forest Model-Based Intrusion Detection Using Fuzzy Inference System for Internet of Things Networks. *International Journal of Computational Intelligence Systems*. 16-31, <https://doi.org/10.1007/s44196-023-00205-w>.
- Kruegel, C. Mutz, D. Robertson, W. and Valeur, F. (2013). Bayesian Event Classification for Intrusion Detection. *Proceedings of 19th Annual Computer Security Applications Conference (ACSAC '03)*, pp. 14-23.
- Panda, M., & Patra, M. R. (2017). Network intrusion detection using naive Bayes. *International journal of computer science and network security (IJCSNS)*, 7(12), 258-263.
- Wu, Y. S., Foo, B. Mei, Y. and Bagchi, S. (2013). "Collaborative Intrusion Detection System (CIDS): A Framework for Accurate and Efficient IDS. *Proceedings of the 19th Annual Computer Security Applications Conference (ACSAC '03)*, 234-244.

The effects of plant extracts on the proximate composition of Bonga Shad (*Ethmalosa fimbriata*) in Marian Market, Calabar, Nigeria.

Inyang- Etoh, A.P. and Asinde, J.

Corresponding author: aniemainyangetoh@yahoo.com

Abstract

This study investigates the impacts of garlic and Moringa extracts on the proximate composition of dry Bonga Shad. Dry Bonga Shad is widely consumed in many regions of West Africa, including Nigeria. The proximate composition of fish, including carbohydrate, moisture, protein, ash and fat, plays a significant role in determining its nutritional quality and value. Various plant extracts are known for their potential to influence the composition of food products due to their bioactive compounds. In this research, moringa and garlic extracts were applied to dry bonga shad to assess its effects on its proximate composition. Proximate analysis revealed that Garlic and Moringa extracts induced a significant reduction in moisture content, with mean values of $9.77 \pm 0.21\%$ and $8.50 \pm 0.10\%$, respectively, compared to the control group's $10.33 \pm 0.15\%$. This reduction aligns seamlessly with practical observations of diminished mold infestation, particularly pronounced in the Moringa group. Protein, ash, and fat contents significantly increased ($p < 0.05$) in treatment groups compared to the control group. The study suggests that Moringa and garlic extracts can be used in dry fish preservation as well as in enhancing the nutritional composition, thereby extending shelf life.

Introduction

In a country with a diverse aquatic ecosystem, fish has long been a staple food, and the practice of drying fish has served as a traditional preservation method for centuries. The popularity of dried fish in Nigeria stems not only from its cultural significance but also its practicality. The drying process allows fish to be stored for extended periods without the need for refrigeration, making it a convenient food choice in areas where reliable electricity and refrigeration facilities are limited (Okafor, 2017). The quality of dry fish can be affected by storage and processing method. Plants extracts have shown to have preservation properties, but their impacts on dry fish composition is unknown.

The study area

Calabar Metropolis is located in south-south Nigeria near the coast of the Gulf of Guinea. Its coordinates are latitudes $4^{\circ}54'30''\text{N}$ & $4^{\circ}95'00''\text{N}$ and longitudes $8^{\circ}19'30''\text{E}$ & $8^{\circ}21'00''\text{E}$ (Fig. 1). The city lies along the Calabar River, (8km) upstream from that river entrance into the Cross River estuary; it is watered by the Calabar River, the Great Kwa River and the Creeks of the Cross River (World Atlas, 2023).

Materials and Method

Materials

Samples of dry Bonga shad were bought from fish sellers at Marian. Samples were immediately put into plastic containers. Other materials used for this experiments include; extracts of Moringa and garlic.

Laboratory studies

The study took place within 30 days of administering the plant extracts on the Bonga Shad specimens. The application of different concentrations (0 g/ml, 20 g/ml, 40 g/ml, 60 g/ml, 80 g/ml, and 100 g/ml) of plant extracts was carried out using a plastic spray bottle. The treated Bonga Shad specimens were subjected to sun drying and subsequently placed in individual petri-dishes. Thereafter, a subset of the control and treated Bonga Shad specimens was homogenized by grinding them into a fine powder (Plate 1).

Determination of moisture content

The moisture content determination involved several sequential steps. Initially, the empty dish and lid were subjected to a 3-hour drying period in an oven set at 105°C . Following this, the dried dish and lid were carefully transferred to a desiccator to cool, and their combined weight was recorded. Subsequently, approximately 3 grams of the sample were weighed and evenly spread within the dish. The dish containing the sample was then placed in the oven and subjected to a 3-hour drying process at 105°C . After completion of the drying phase, the dish with its lid partially covered was promptly moved to the desiccator to cool. Upon cooling, the dish, along with its now-dried

sample, was reweighed to determine the final weight. This procedure allowed for the accurate determination of the moisture content in the sample, following the established AOAC (2000) guidelines.



Plate 1: Ground fish samples for proximate analysis. (Source: This study)

$$\text{Moisture (\%)} = \left(\frac{W_1 - W_2}{W_1} \right) \times 100$$

1

where:

W1 = weight (g) of sample before drying

W2 = weight (g) of sample after drying

Determination of protein content

The analysis procedure involved the following steps. Initially, a sample weighing between 0.5 to 1.0 grams was carefully placed into a digestion flask. Subsequently, 5 grams of Kjeldahl catalyst and 200 milliliters of concentrated sulfuric acid (H_2SO_4) were added to the flask. A tube containing the same chemicals as above, excluding the sample, was prepared to serve as a blank reference. The flasks were positioned in an inclined manner, and gentle heating ensued until frothing ceased. The solution was then brought to a brisk boil, continuing until clarity was achieved. Following this, the solution was allowed to cool, and 60 milliliters of distilled water were cautiously added.

The next step involved connecting the flask to the digestion bulb on the condenser, with the condenser tip immersed in standard acid and 5-7 drops of mixed indicator in the receiver. The flask was rotated to thoroughly mix the contents, and heating was applied until all ammonia (NH_3) was distilled. Upon completion of the distillation process, the receiver was removed, the condenser tip was washed, and the excess standard acid distilled was titrated using a standard solution of sodium hydroxide (NaOH). This series of steps, executed meticulously, formed the basis for the analysis, providing valuable insights into the composition of the sample.

Determination of ash content (AOAC, 2000)

The ash content determination involved a systematic set of procedures. Initially, the crucible and its lid were placed in a furnace and subjected to an overnight heating at 550°C to eliminate any impurities on their surfaces. Subsequently, the heated crucible was allowed to cool in a desiccator for 30 minutes, ensuring a stable temperature. Following cooling, the crucible and its lid were precisely weighed to three decimal places. Approximately 5 grams of the sample were then carefully measured and placed into the crucible. The crucible, with the lid partially covered,

was heated over a low Bunsen flame until fumes ceased to be produced. Once this stage was reached, the crucible and its lid were transferred back to the furnace.

The subsequent step involved heating the crucible at 550°C overnight, with a crucial note not to cover the lid during this process. After complete heating, the lid was placed to prevent the loss of fluffy ash. The cooled crucible, along with the lid, was then weighed when the sample had transformed into a gray ash. If this transformation did not occur, the crucible and lid were returned to the furnace for further ashing. This methodical approach ensured accurate determination of the ash content in the sample, adhering to the established AOAC (2000) guidelines.

$$\text{Ash (\%)} = \frac{\text{Weight of ash}}{\text{Weight of sample}} \times 100 \quad 2$$

Determination of fat content

The extraction process was carried out following a meticulous set of steps. Firstly, the bottle and its lid were placed in an incubator at 105°C overnight to ensure stability in the bottle's weight. After this preparatory step, approximately 3-5 grams of the sample were carefully weighed and wrapped in a paper filter. The weighed sample was then transferred into an extraction thimble and subsequently placed into a soxhlet apparatus. About 250 milliliters of petroleum ether were added to the bottle, which was positioned on a heating mantle. The soxhlet apparatus was connected, and water was turned on to cool the system. Following this, the heating mantle was switched on, and the sample was heated for approximately 14 hours, maintaining a heat rate of 150 drops per minute.

Once the heating process was complete, the solvent was evaporated using a vacuum condenser. The bottle containing the sample was then incubated at 80-90°C until the solvent was entirely evaporated, ensuring the bottles' complete dryness. After this drying phase, the bottle, with its lid partially covered, was transferred to a desiccator to cool. Upon cooling, the bottle and its dried content were reweighed to determine the final weight. This systematic approach ensured a thorough extraction process, providing reliable results for further analysis.

$$\text{Fat (\%)} = \frac{\text{Weight of fat}}{\text{Weight of sample}} \times 100 \quad 3$$

Determination of crude fiber

To determine the crude fiber content of the sample, a procedure was followed with precision and care. Initially, 5.0 grams of the sample was combined with 200 milliliters of 0.25M sulfuric acid, and the mixture was heated for 30 minutes. Subsequently, the resulting solution was filtered using a Buchner funnel. The residue obtained was thoroughly washed with distilled water until it was free from acidity.

Following this, a solution of 0.3M hydrochloric acid (HCl) was utilized to boil the residue for an additional 30 minutes. After filtration, the residue was washed with distilled water until it became free from alkalinity. A subsequent rinse with 10% HCl and two rinses with ethanol were carried out. To conclude, the residue underwent three rinses with petroleum ether. The purified residue was then placed in a crucible and dried at 105 °C in an oven overnight.

After cooling in a desiccator, the residue was subjected to ignition in a muffle furnace at 550 °C for 90 minutes, resulting in the formation of ash. This ash was carefully washed. The percentage of crude fiber was determined using the following formula:

$$\text{Percentage Crude Fiber} = \left(\frac{\text{Loss in weight on ignition}}{\text{Weight of sample}} \right) \times 100 \quad 4$$

Statistical analysis

The results of the experiment were analyzed using descriptive statistics. ANOVA and means were separated using Duncan's Multiple Range Test (DMRT) at $p < 0.05$. All analysis was carried out using Statistical Product for Service Solution (SPSS ver. 20).

Results

Impact of plant extracts on the proximate composition of Bonga Shad

The investigation into the proximate composition of Bonga Shad (Table 1), underscores the distinct effects of Moringa and Garlic extracts on nutritional quality. Including specific result values helps highlight how much these extracts change the nutritional quality. Specifically, the application of plant extracts resulted in substantial reductions in moisture content compared to the control group. The mean values with different letters denote significant differences between columns ($P < 0.05$).

Notably, Garlic (9.77%) and Moringa (8.50%) demonstrated a significant decrease in moisture content compared to the control (10.33%; $P < 0.05$), indicative of effective moisture reduction. This reduction in moisture content is further supported by practical observations, notably the substantial decrease in mold infestation observed in the Moringa and Garlic groups compared to others (Plate 8). More so, Moringa and Garlic extracts exhibited a substantial impact by significantly increasing the protein content compared to the control and Neem groups. For instance, Garlic (56.30%) showed a significant increase compared to the control (46.63%; $P < 0.05$).

The application of Moringa and Garlic extracts contributed significantly to higher ash content in Bonga Shad compared to the control. The statistical significance between columns ($P < 0.05$) emphasizes the distinctive influence of these extracts on the ash composition. Notably, Garlic (3.98%) and Moringa (3.96%) showed a significant increase compared to the control (3.03%; $P < 0.05$).

TABLE 1
Proximate Composition of Bonga Shad treated with plant extracts

Plant Extract Indices	Control	Moringa extract	Garlic extract
Moisture	10.33±0.15 ^{ab}	8.50±0.10 ^c	9.77±0.21 ^b
Protein	46.63±0.15 ^a	54.10±0.20 ^b	56.30±0.26 ^c
Ash	3.03±0.02 ^a	3.96±0.01 ^b	3.98±0.01 ^b
Fat	10.90±0.26 ^a	12.83±0.12 ^c	11.87±0.25 ^d
Carbohydrate	29.47±0.21 ^a	20.72±0.10 ^c	17.57±0.15 ^d

Values are means±SD; mean values with different letters between columns are significantly different at $P < 0.05$

DISCUSSION

The study demonstrates that Moringa, and garlic extracts can be used to modify the proximate composition of dried Bonga Shad. Significant reductions in moisture content were observed across all extract groups compared to the control, in line with previous studies highlighting the efficacy of various plant extracts, including Moringa, in reducing moisture and preventing mold growth in dried fish. Notably, Moringa (8.50%) and Garlic (9.77%) exhibited superior moisture reduction compared to the control (10.33%), aligning with earlier research by Rasulet *al.* (2022) showcasing Moringa's efficacy in this regard. Practical observations of reduced mold infestation in the Moringa group further support the link between moisture reduction and mold prevention, underscoring the potential practical applications of Moringa extract in dried fish preservation.

Moringa and Garlic extracts contributed to a significant increase in protein content compared to the control group. This aligns with previous studies reporting on the protein-preserving properties of certain plant extracts. Notably, Garlic exhibited the highest protein content (56.30%), surpassing Moringa (55.22%). Moringa and Garlic extracts led to a significant increase in ash content compared to the control group. While an increase in ash content may not directly affect nutritional value, it could indicate potential mineral enrichment from the extracts, suggesting an avenue for further investigation. Both Moringa and Garlic extracts resulted in a significant increase in fat content compared to the control group. The mechanism behind this increase could involve the extracts inhibiting fat oxidation or promoting fat retention during drying. While moderate fat content is desirable for taste and texture, excessive increase may impact shelf life, necessitating further exploration.

Moringa extracts led to a significant reduction in carbohydrate content compared to the control and Garlic groups. This reduction could be attributed to the extracts promoting carbohydrate utilization during drying or interacting with specific carbohydrate components in the fish muscle. Lower carbohydrate content may contribute to extended shelf life by reducing sugar availability for microbial growth.

The present study illuminates the significant impact of plant extracts on the proximate composition of Bonga Shad, with each extract exerting distinct effects on various components. These findings align with and extend existing knowledge on plant extract use in dried fish preservation. Future research avenues should prioritize optimizing extract concentrations and combinations to achieve a balanced effect on moisture reduction, protein preservation, and minimal fat increase. Investigating the underlying mechanisms of different extracts on specific proximate components and evaluating the long-term impact on sensory properties, shelf life, and nutritional value will further refine the application of plant-based extracts for effective and sustainable dried fish preservation, ultimately contributing to improved food security and healthy food choices.

CONCLUSION

The impact of plant extracts on the proximate composition of Bonga Shad indicates significant alterations in moisture, protein, ash, fat, and carbohydrate content. Practical observations, such as the reduced mold infestation in the Moringa group further support the quantitative findings.

Moringa and garlic extracts can be used as natural preservations to improve the quality of dry fish, it also revealed that plant extracts can positively influence the nutritional composition of preserved fish. Further research can optimize extract's industrial application.



Plate 2: Jelly-like plant extracts from Moringa and Garlic

Source: This study



100 g/mil

80 g/mil

60 g/mil

40 g/mil

20 g/mil

Control

Plate 3: Experimental setup for both the experimental groups and the control in triplicates (Source: This study)

Acknowledgements

I will like to acknowledge the Department of Fisheries and Aquaculture University of Calabar for the opportunity given me to carry out this research during my time with them.

References

- AOAC (2000). *Official Methods of Analysis*. 17th Edition, The Association of Official Analytical Chemists, Gaithersburg, MD, USA.
<https://www.worldatlas.com/>. World Atlas
 Okafor, C. (2017). Cultural significance of dried fish in Nigerian cuisine. *Food and Culture Review*, 25(2), 112-126.
 Rasul, M.G., Yuan, C., Yu, K., Takaki, K. & Shah, A.K.M.A (2022). Factors influencing the nutritional composition, quality and safety of dried fishery products. *Food Research* 6 (5): 444 – 466

An Improved Estimator for Estimating the Current Population Mean in a Two Occasion Successive Sampling Using One Auxiliary Variable

¹Agbebia, Catherine Shike, ¹Eric Monfung Egomo, ²Etorti, Imoke John

¹Department of Statistics, University of Cross River State (UNICROSS), Calabar, Nigeria; ²Department of Mathematics, Cross River College of Education, Akamkpa-Nigeria

Corresponding Author's E-mail: agbebiacatherine6@gmail.com

Abstract

In this work, the problem of estimating a finite population mean on the current occasion based on samples drawn over two occasions has been considered. A new estimator was proposed to estimate the population mean on the current occasion, using only the matched part and only one auxiliary variable, which is available on both occasions. The bias and mean square error was obtained. Another estimator, which is a linear combination of the means of the match and the unmatched portions on the second occasion, was also proposed. The optimum mean square error of this combined estimator was compared with several other estimators.

Keywords: Successive sampling, auxiliary variable, optimum replacement policy, mean square error, bias.

1. Introduction

In repetitive surveys for estimating the same characteristic at different points in time, it is possible to use the information collected on the previous occasion to improve upon the conventional estimator for the current period. The method of sampling units on successive occasions consists in selecting samples of the sample sizes on different occasions such that some units are common to samples selected on previous occasions. The main objective of successive sampling is to estimate the population parameter for the most recent occasion as well as changes in the parameters from one occasion to the next. The theory of successive sampling was initiated by Jessen (1942), followed by Patterson (1950) who extended the work from two occasions to more. Furthermore, Okafor and Arnab (1987), Tracy and Singh (1999), Singh (2005), Sharma and Kumar (2021), Ralte and Das (2015) and Shabbiret al.(2005), among others, have suggested several estimators by using the auxiliary information for estimating the population mean on the current occasion.

2. Notations and Some Existing Estimators

Suppose we have a finite population with size N . For simplicity, we assume that N is so large that finite population correction can be ignored. Let the characteristic under study on the first and second occasions be denoted by X and Y respectively. Also, let's assume that the information on an auxiliary variable Z is available on both occasions. Let a simple random sample of size n be drawn without replacement on the first occasion of which a subsample of size $m = n\lambda$, ($0 < \lambda < 1$), is retained (matched) for use on the second occasion. Again, at the current occasion, we draw a simple random sample (without replacement) of size $u = n - m = n\mu$ ($0 < \mu < 1$) units from the remaining $(N - n)$ units of the population so that the sample size on the second occasion is also n . λ and μ are the fractions of matched and fresh samples at the current occasion respectively, such that $\mu + \lambda = 1$. Thus, the size on the second occasion is also $n = n\lambda + n\mu$. Let:

\bar{Z} = Population mean of the auxiliary variable Z

\bar{z}_n = Sample mean of the auxiliary variable Z based on a sample of size n drawn on the first occasion.

\bar{y}_m = Sample mean of the study variable Y based on m units common to both occasions and observed on the second occasion

\bar{x}_m = Sample mean of the study variable X based on m units common to both occasions and observed on the first occasion

\bar{z}_u = Sample mean of the auxiliary variable Z based on u units drawn on the second occasion

\bar{x}_n = Sample mean of the study variable X based on n units drawn on the first occasion

\bar{y}_u = Sample mean of the study variable Y based on u units drawn on the second occasion

ρ_{xy} = The correlation coefficient between the variables X and Y

ρ_{xz} = The correlation coefficient between the variables X and Z

ρ_{yz} = The correlation coefficient between the variables Y and Z

S_x , S_y and S_z are variances of X, Y and Z respectively and

C_z , C_x and C_y are the coefficients of variation of the respective variables.

Consider the following estimators under successive sampling:

- (i) The usual mean estimator

$$T_1 = \bar{y}$$

$$Var(T_1) = \frac{S_y^2}{n} \quad (1)$$

- (ii) Cochran (1977) Estimator

$$T_2 = \Theta_2 T_{2(m)} + (1 - \Theta_2) T_{2(u)}$$

Where $T_{2(m)} = \bar{y}_m + b(\bar{X} - \bar{x}_m)$, $T_{2(u)} = \bar{y}_{2u}$ and b is the sample regression coefficient of y on x

$$Var(T_2) = \frac{S_y^2}{2n} \left(1 + \sqrt{1 - \rho_{xy}^2} \right) \quad (2)$$

- (iii) Singh (2005) Estimator

$$T_3 = \Theta_3 T_{3(m)} + (1 - \Theta_3) T_{3(u)}$$

Where $T_{3(m)} = \frac{\bar{y}_m \bar{x}_n}{\bar{x}_m \bar{z}_n} \bar{Z}$, and $T_{3(u)} = \frac{\bar{y}_u}{\bar{z}_u} \bar{Z}$

$$Var_{opt}(T_3) = \frac{(A + C)^2 + (A + C)(B - C)\mu_{30}}{n[(A + C)(B - C)\mu_{30}^2]}, \quad (3)$$

$$\mu_{30} = \frac{-(1 - \rho_{YZ}) \pm \sqrt{(1 - \rho_{YZ})(1 - \rho_{xy})}}{(\rho_{yz} - \rho_{xy})}, \quad A = \bar{Y}^2 C_y^2, \quad B = \bar{Y}^2 (C_x^2 - 2\rho_{xy} C_x C_y),$$

$$C = \bar{Y}^2 (C_z^2 - 2\rho_{yz} C_y C_z)$$

- (iv) Ralte and Das (2015) Estimator

$$T_4 = \Theta_4 T_{4(m)} + (1 - \Theta_4) T_{4(u)}$$

Where $T_{4(m)} = \frac{\bar{y}_m}{\bar{x}_m} \left[\bar{x}_m + b_{xz} (\bar{z}_n - \bar{z}_m) \right] \frac{\bar{Z}}{\bar{z}_n}$, $T_{4(u)} = \frac{\bar{y}_u}{\bar{z}_u} \bar{Z}$ and

$$Var_{opt}(T_4) = \frac{(\alpha_1^1)^2 + \alpha_1^1 \alpha_2^1 \mu_{40}}{n[\alpha_1^1 + \alpha_2^1 (\mu_{40})^2]}, \quad (4)$$

$$, \quad \alpha_1^1 = A^1 + C^1, \alpha_2^1 = B^1 - C^1, \quad A^1 = \bar{Y}^2 C_Y^2, \quad B^1 = \bar{Y}^2 (\rho_{XZ}^2 C_X^2 - 2\rho_{XZ} \rho_{YZ} C_X C_Y),$$

$$C^1 = \bar{Y}^2 (C_Z^2 - 2\rho_{YZ} C_Y C_Z)$$

$$\mu_{40} = \frac{-2(1 - \rho_{XZ}) \pm \sqrt{2(1 - \rho_{YZ})\{2(1 - \rho_{YZ}) + (1 - \rho_{XZ})(2\rho_{YZ} - \rho_{XZ} - 1)\}}}{(1 - \rho_{XZ})(2\rho_{YZ} - \rho_{XZ} - 1)}$$

- (v) Tracy and Singh (1999) Estimator

$$T_5 = \Theta_5 T_{5(m)} + (1 - \Theta_5) T_{5(u)} \text{ Where, } T_{5(m)} = h_0 \bar{y}_2 + h_1 \bar{y}_2 \left(\frac{\bar{x}}{\bar{x}_m} \right) + h_2 \bar{y}_2 \left(\frac{\bar{z}}{\bar{z}_m} \right) + h_3 \bar{y}_2 \left(\frac{\bar{x}}{\bar{x}_m} \right) \left(\frac{\bar{z}}{\bar{z}_m} \right),$$

$$T_{5(u)} = \bar{y}_{2u} \text{ and } \sum_{i=1}^3 h_i = 1$$

$$Var_{opt}(T_5) = \frac{S_y^2}{2n} \left[1 + \sqrt{1 - R_{y.zx}^2} \right] \quad (5)$$

$$R_{y.zx}^2 = \frac{\rho_{yx}^2 + \rho_{yz}^2 - 2\rho_{zx}\rho_{yz}\rho_{yx}}{(1 - \rho_{zx}^2)}$$

(vi) Shabbiret *al.* (2005) Estimator

$$T_6 = \Theta_6 T_{6(m)} + (1 - \Theta_6) T_{6(u)} \text{ Where}$$

$$T_{6(m)} = w_1 \left(\bar{y}_m + b_1 (\bar{x} - \bar{x}_m) \frac{\bar{x}}{\bar{X}} \right) + w_2 \left(\bar{y}_m + b_2 (\bar{z} - \bar{z}_m) \frac{\bar{z}}{\bar{Z}} \right) \text{ and } T_{6(u)} = \bar{y}_u + b(\bar{Z} - \bar{z}_u).$$

$w_1 + w_2 = 1$, b_1 and b_2 are sample regression coefficients of y on x and y and z respectively for matched portions.

$$Var_{opt}(T_6) = \frac{S_y^2 R_{y.zx}^2 (1 - \rho_{yz}^2)}{2n \left(1 - \frac{\rho_{yz}^2}{2} - \sqrt{(1 - \rho_{yz}^2)(1 - R_{y.zx}^2)} \right)} \quad (6)$$

Θ_i , $i=2, 3, 4, 5, 6$ are unknown constants that are to be estimated later and μ_{oi} , $i=3, 4$, are optimum values of μ (fraction of a sample to be taken afresh at the second occasion)

3. Proposed Estimator

Motivated by Ralte and Das (2015) and Shabbiret *al.* (2005), we propose a ratio-to -regression-to-exponential estimator for the population mean \bar{Y} on the second occasion which is based on a sample of size m common to both occasions given by:

$$T_{P(m)} = \left[\alpha_1 \frac{\bar{y}_m}{\bar{x}_m} \bar{x}_n + \alpha_2 (\bar{x}_n - \bar{x}_m) \right] \exp \left(\frac{\bar{Z} - \bar{z}_n}{\bar{Z} + \bar{z}_n} \right) \quad (7).$$

The bias and mean square error of the proposed estimator $T_{P(m)}$ up to the second order of approximation is obtained as follow: Let

$$e_{ym} = \frac{\bar{y}_m - \bar{Y}}{\bar{Y}}, \quad e_{xm} = \frac{\bar{x}_m - \bar{X}}{\bar{X}}, \quad e_{xn} = \frac{\bar{x}_n - \bar{X}}{\bar{X}}, \quad e_{zn} = \frac{\bar{z}_n - \bar{Z}}{\bar{Z}}, \quad e_{yu} = \frac{\bar{y}_u - \bar{Y}}{\bar{Y}}$$

$$e_{zu} = \frac{\bar{z}_u - \bar{Z}}{\bar{Z}}, \text{ such that } \bar{y}_m = \bar{Y}(1 + e_{ym}), \quad \bar{x}_n = \bar{X}(1 + e_{xn}), \quad \bar{x}_m = \bar{X}(1 + e_{xm}),$$

$$\bar{y}_u = \bar{Y}(1 + e_{yu}) \quad \bar{z}_n = \bar{Z}(1 + e_{zn}) \quad \text{and}$$

$$E(e_{ym}) = E(e_{xm}) = E(e_{xn}) = E(e_{zn}) = E(e_{yu}) = E(e_{zu}) = 0$$

$$\begin{aligned}
E(e_{ym}^2) &= \left(\frac{1}{m} - \frac{1}{N}\right) C_y^2, \quad E(e_{xm}^2) = \left(\frac{1}{m} - \frac{1}{N}\right) C_x^2, \quad E(e_{xn}^2) = \left(\frac{1}{n} - \frac{1}{N}\right) C_x^2, \quad E(e_{zn}^2) = \left(\frac{1}{n} - \frac{1}{N}\right) C_z^2 \\
E(e_{yu}^2) &= \left(\frac{1}{u} - \frac{1}{N}\right) C_y^2, \quad E(e_{zu}^2) = \left(\frac{1}{u} - \frac{1}{N}\right) C_z^2, \quad E(e_{xm}e_{zn}) = \left(\frac{1}{n} - \frac{1}{N}\right) C_{xz}, \\
E(e_{xm}e_{zn}) &= \left(\frac{1}{n} - \frac{1}{N}\right) C_{xz}, \quad E(e_{xm}e_{xn}) = \left(\frac{1}{n} - \frac{1}{N}\right) C_x^2, \quad E(e_{ym}e_{xm}) = \left(\frac{1}{m} - \frac{1}{N}\right) C_{xy}, \\
E(e_{ym}e_{xn}) &= \left(\frac{1}{n} - \frac{1}{N}\right) C_{xy}, \quad E(e_{ym}e_{zn}) = \left(\frac{1}{n} - \frac{1}{N}\right) C_{yz}
\end{aligned}$$

Suppose terms having order higher than two in e are negligible, then we express equation (7) as:

$$\begin{aligned}
T_{p(m)} &= \left\{ \alpha_1 \left(\bar{Y}(1 + e_{ym})(1 + e_{xn})(1 + e_{xm})^{-1} \right) + \alpha_2 \left(\bar{X}(e_{xm} - e_{xn}) \right) \right\} \exp \left(\frac{-e_{zn}}{\bar{Z} + e_{zn}} \right) \\
T_{p(m)} &= \left\{ \alpha_1 \left[\bar{Y} \left(1 - e_{xm} + e_{xm}^2 + e_{xn} - e_{xm}e_{xn} + e_{ym} - e_{ym}e_{xm} + e_{ym}e_{xn} \right) \right] + \alpha_2 \left[\bar{X}(e_{xm} - e_{xn}) \right] \right\} \\
&\quad \exp \left(-\frac{e_{zn}}{2} \left(1 + \frac{e_{zn}}{2} \right)^{-1} \right) \\
T_{p(m)} &= \left\{ \begin{aligned} &\alpha_1 \bar{y} + \alpha_1 \bar{y}e_{xm} - \alpha_1 \bar{y}e_{xm}^2 + \alpha_1 \bar{y}e_{xn}^2 - \alpha_1 \bar{y}e_{xm}e_{xn} + \alpha_1 \bar{y}e_{ym} - \alpha_1 \bar{y}e_{ym}e_{xm} + \alpha_1 \bar{y}e_{ym}e_{xn} + \alpha_1 \bar{y}e_{ym}e_{xn} \\ &+ \alpha_2 \bar{X}e_{xm} - \alpha_2 \bar{X}e_{xn} \end{aligned} \right\} \\
&\quad \left[1 - \frac{e_{zn}}{2} + \frac{e_{zn}^2}{4} + \frac{e_{zn}^2}{8} \right] \\
T_{p(m)} &= \left\{ \begin{aligned} &\bar{Y} + \bar{Y}(\alpha_1 - 1) + (\alpha_1 \bar{Y} + \alpha_2 \bar{X})e_{xm} - (\alpha_1 \bar{Y} + \alpha_2 \bar{X})e_{xm} + \alpha_1 \bar{Y}e_{xm}^2 - \alpha_1 \bar{Y}e_{xm}e_{xn} \\ &+ \alpha_1 \bar{Y}e_{ym} - \alpha_1 \bar{Y}e_{ym}e_{xm} + \alpha_1 \bar{Y}e_{ym}e_{xn} - \alpha_1 \bar{Y} \frac{e_{zn}}{2} - \left(\frac{\alpha_1 \bar{Y}}{2} + \frac{\alpha_2 \bar{X}}{2} \right) e_{xm}e_{zn} + \\ &\left(\frac{\alpha_1 \bar{Y}}{2} + \frac{\alpha_2 \bar{X}}{2} \right) e_{xm}e_{zn} - \frac{\alpha_1 \bar{Y}e_{ym}e_{zn}}{2} + \frac{\alpha_1 \bar{Y}e_{zn}^2}{4} + \frac{\alpha_1 \bar{Y}e_{zn}^2}{8} \end{aligned} \right\} \\
Bias(T_{p(m)}) &= E(T_{p(m)} - \bar{Y}) = \bar{Y}(\alpha_1 - 1) - \left(\frac{\alpha_1 \bar{Y}}{2} - \frac{\alpha_2 \bar{X}}{2} \right) E(e_{xm}e_{zn}) + \left(\frac{\alpha_1 \bar{Y}}{2} + \frac{\alpha_2 \bar{X}}{2} \right) E(e_{xm}e_{zn}) + \\
&\quad \bar{Y} \alpha_1 \left(E(e_{xm}^2) - E(e_{xm}e_{xm}) - E(e_{ym}e_{xm}) + E(e_{ym}e_{xn}) - E\left(\frac{e_{ym}e_{zn}}{2}\right) + E\left(\frac{e_{zn}^2}{4}\right) + E\left(\frac{e_{zn}^2}{8}\right) \right) \\
Bias(T_{p(m)}) &= \bar{Y}(\alpha_1 - 1) - \left(\frac{\alpha_1 \bar{Y}}{2} + \frac{\alpha_2 \bar{X}}{2} \right) \left(\frac{1}{n} - \frac{1}{N} \right) C_{xz} + \left(\frac{\alpha_1 \bar{Y}}{2} + \frac{\alpha_2 \bar{X}}{2} \right) \left(\frac{1}{n} - \frac{1}{N} \right) C_{xz} + \\
&\quad \alpha_1 \bar{Y} \left[\left(\frac{1}{m} - \frac{1}{N} \right) C_x^2 - \left(\frac{1}{n} - \frac{1}{N} \right) C_x^2 - \left(\frac{1}{m} - \frac{1}{N} \right) C_{xy} + \left(\frac{1}{n} - \frac{1}{N} \right) C_{xy} - \left(\frac{1}{n} - \frac{1}{N} \right) \frac{C_{yz}}{2} + \right. \\
&\quad \left. \left(\frac{1}{n} - \frac{1}{N} \right) \frac{C_z^2}{4} + \left(\frac{1}{n} - \frac{1}{N} \right) \frac{C_z^2}{8} \right]
\end{aligned}$$

Thus, assuming N is large that the finite population correction is ignored, we have

$$Bias(T_{p(m)}) = \bar{Y}(\alpha_1 - 1) + \bar{Y}\alpha_1 \left[\left(\frac{1}{m} - \frac{1}{n} \right) (C_x^2 - C_{xy}) + \frac{1}{n} \left(C_{xy} - C_x^2 - \frac{C_{yz}}{2} + \frac{C_z^2}{4} + \frac{C_z^2}{8} \right) \right] \quad (8)$$

The mean square error to the first order of approximation is given by:

$$MSE(T_{p(m)}) = E(T_{p(m)} - \bar{Y})^2 =$$

$$\begin{aligned} & \left\{ \bar{Y}^2(\alpha_1 - 1)^2 + 2\bar{Y}(\alpha_1 - 1)(\alpha_1\bar{Y} + \alpha_2\bar{X})E(e_{xm}) + (\alpha_1\bar{Y} + \alpha_2\bar{X})^2 E(e_{xm}^2) + (\alpha_1\bar{Y} + \alpha_2\bar{X})^2 E(e_{xm}^2) - \right. \\ & 2\alpha_1^2\bar{Y}^2 E(e_{ym}e_{xm}) + 2\alpha_1^2\bar{Y}^2 \frac{e_{zn}}{2} E(e_{xm}) - 2\alpha_1\alpha_2\bar{Y}\bar{X}E(e_{ym}e_{xm}) + \alpha_1\alpha_2\bar{Y}\bar{X}E(e_{zn}e_{xm}) + \bar{Y}^2\alpha_1^2 E(e_{ym}^2) - \\ & \left. \bar{Y}^2\alpha_1^2 E(e_{ym}e_{zn}) + \bar{Y}^2\alpha_1^2 E\left(\frac{e_{zn}^2}{4}\right) - 2\bar{Y}(\alpha_1 - 1)(\alpha_1\bar{Y} + \alpha_2\bar{X})E(e_{xm}) + 2\bar{Y}^2\alpha_1(\alpha_1 - 1)E(e_{ym}) \right. \\ & \left. - 2\bar{Y}^2\alpha_1(\alpha_1 - 1)E\left(\frac{e_{zn}}{2}\right) - 2(\alpha_1\bar{Y} + \alpha_2\bar{X})^2 E(e_{xm}e_{xm}) + 2\bar{Y}\alpha_1(\alpha_1\bar{Y} + \alpha_2\bar{X})E(e_{ym}e_{xm}) - \right. \\ & \left. \bar{Y}\alpha_1(\alpha_1\bar{Y} + \alpha_2\bar{X})E(e_{xm}e_{zn}) \right\} \\ & = \left\{ \bar{Y}^2 + \alpha_1^2\bar{Y}^2 \left[1 + \left(\frac{1}{m} - \frac{1}{n} \right) C_x^2 - 2\left(\frac{1}{n} - \frac{1}{N} \right) C_x^2 - 2\left(\frac{1}{m} - \frac{1}{n} \right) C_{xy} + \left(\frac{1}{m} - \frac{1}{n} \right) C_y^2 - \left(\frac{1}{n} - \frac{1}{N} \right) C_{yz} + \right. \right. \\ & \left. \left. \left(\frac{1}{n} - \frac{1}{N} \right) \frac{C_z^2}{4} \right] \right\} \\ & = \left\{ -2\bar{Y}^2\alpha_1 + 2\alpha_1\alpha_2\bar{Y}\bar{X} \left[\left(\frac{1}{m} - \frac{1}{n} \right) C_x^2 - 2\left(\frac{1}{n} - \frac{1}{N} \right) C_x^2 - \left(\frac{1}{m} + \frac{1}{N} \right) C_{xy} \right] + \right. \\ & \left. \alpha_2^2\bar{X}^2 \left[\left(\frac{1}{m} - \frac{1}{N} \right) C_x^2 - 2\left(\frac{1}{n} - \frac{1}{N} \right) C_x^2 \right] \right\} \\ & = \left\{ \bar{Y}^2 + \alpha_1^2\bar{Y}^2 \left[1 + \left(\frac{1}{m} - \frac{1}{N} \right) (C_y^2 + C_x^2 - 2C_{xy}) + \left(\frac{1}{n} - \frac{1}{N} \right) \left(\frac{C_z^2}{4} - C_x^2 - C_{xz} + 2C_{xy} \right) \right] - 2\bar{Y}^2\alpha_1 + \right. \\ & \left. 2\alpha_1\alpha_2\bar{Y}\bar{X} \left(\left(\frac{1}{n} - \frac{1}{N} \right) (C_{xy} - C_x^2) - \left(\frac{1}{m} - \frac{1}{N} \right) (C_{xy} - C_x^2) \right) + \alpha_2^2\bar{X}^2 \left(\frac{1}{m} - \frac{1}{n} \right) C_x^2 \right\} \end{aligned}$$

Therefore, the mean square error of $T_{p(m)}$ is given as:

$$MSE(T_{p(m)}) = \left\{ \bar{Y}^2(\alpha_1 - 1)^2 + \bar{Y}^2\alpha_1^2 \frac{C_y^2}{m} + \alpha_1^2\bar{Y}^2 \frac{1}{n} \left(\frac{C_z^2}{4} - C_{yz} \right) + \alpha_1^2\bar{Y}^2 \left(\frac{1}{m} - \frac{1}{n} \right) (C_x^2 - 2C_{xy}) \right. \\ \left. - 2\alpha_1\alpha_2\bar{Y}\bar{X} \left(\frac{1}{m} - \frac{1}{n} \right) (C_{xy} - C_x^2) + \alpha_2^2\bar{Y}\bar{X} \left(\frac{1}{m} - \frac{1}{n} \right) C_x^2 \right\} \quad (9)$$

Now setting

$$\frac{\partial Var(MSE(T_{p(m)}))}{\partial \alpha_1} = 0 \quad \text{and} \quad \frac{\partial MSE(T_{p(m)})}{\alpha_2} = 0, \text{ we have}$$

$$\alpha_1 = \frac{1}{1 + \frac{C_y^2}{m}(1 - \rho_{xy}^2) + \frac{1}{n}\left(\frac{C_z^2}{4} - \rho_{yz}C_yC_z + \rho_{xy}^2C_y^2\right)} \quad (10)$$

$$\alpha_2 = \frac{\bar{Y}\left(\frac{\rho C_y}{C_x} - 1\right)}{\bar{X}\left[1 + \frac{C_y^2}{m}(1 - \rho_{xy}^2) + \frac{1}{n}\left(\frac{C_z^2}{4} - \rho_{yz}C_yC_z + \rho_{xy}^2C_y^2\right)\right]} \quad (11)$$

Substituting the values of α_1 and α_2 into equation (9) and simplifying, we get the minimum mean square error of $T_{p(m)}$ as:

$$MSE_{\min}(T_{p(m)}) = \bar{Y}^2 \left\{ 1 + \frac{nC_y^2(1 - \rho_{xy}^2) + m(C_z^2 - C_{yz} + \rho_{xy}^2C_y^2)}{mn\left[1 + \frac{C_y^2}{m}(1 - \rho_{xy}^2) + \frac{1}{n}\left(\frac{C_z^2}{4} + \rho_{xy}^2C_y^2 - C_{yz}\right)\right]^2} \right\} \quad (12)$$

Following Shabbiret al. (2005), another estimator based on the unmatched portion of the sample is:

$$T_{6(u)} = \bar{y}_u + b(\bar{Z} - \bar{z}_u) \quad (13)$$

With bias and variance given as

Bias $T_{6(u)} = 0$

$$Var(T_{6(u)}) = \frac{S_y^2}{u}(1 - \rho_{yz}^2) \quad (14)$$

Combining the estimators $T_{p(m)}$ and $T_{6(u)}$, the resultant estimator of the population mean \bar{Y} is

$$T_{6p} = \Theta_p T_{p(m)} + (1 - \Theta_p) T_{6(u)} \quad (15)$$

Where Θ_p is an unknown constant to be determined under certain criterion.

The bias and mean square error up to the first order of approximation are given by:

$$Bias(T_{6p}) = \Theta_p Bias(T_{p(m)}) + (1 - \Theta_p) Bias(T_{6(u)}) \quad (16)$$

$$MSE(T_{6p}) = \Theta_p^2 MSE(T_{p(m)}) + (1 - \Theta_p)^2 MSE(T_{6(u)}) \quad (17)$$

Differentiating (17) with respect to Θ_p , we get the optimum value of Θ_p as:

$$\Theta_p = \frac{MSE(T_{p(m)})}{MSE(T_{6(u)}) + MSE(T_{p(m)})} \quad (18)$$

Then, substituting Θ_p into equation (17), the minimum mean square error of T_{6p} is:

$$MSE_{\min}(T_{6p}) = \frac{MSE(T_{6(u)})MSE(T_{p(m)})}{MSE(T_{6(u)}) + MSE(T_{p(m)})} \quad (19)$$

Substituting equations (12) and (14) in (19) and putting $\lambda = \frac{m}{n}$, $\mu = \frac{u}{n}$, $m = n(1 - \mu)$, in (19), after simplifying, we get the minimum mean square error of T_{6p} as

$$MSE_{\min}(T_{6p}) = \frac{\bar{Y}^2 D \left\{ n(1 - \mu) \left[1 + \frac{A^o}{n(1 - \mu)} + \frac{B^o}{n} \right]^2 + A^o + (1 - \mu)C^o \right\}}{n \left\{ \bar{Y} \mu [A + (1 - \mu)C^o] + (n\bar{Y}^2 \mu + D^o)(1 - \mu) \left[1 + \frac{A^o}{n(1 - \mu)} + \frac{B^o}{n} \right]^2 \right\}} \quad (20)$$

Where

$$A^o = C_y^2(1 - \rho_{xy}^2), \quad B^o = \frac{C_z^2}{4} + \rho_{xy}^2 C_y^2 - \rho_{zy} C_y C_z, \quad C^o = (C_z^2 - C_{yz}) + \rho_{xy}^2 C_y^2, \\ D^o = S_y^2(1 - \rho_{yz}^2)$$

For simplification, equation (20) can also be expressed as:

$$MSE_{\min}(T_{6p}) = \frac{\alpha_1 \mu^2 - \alpha_2 \mu + \alpha_3}{\alpha_4 \mu^3 - \alpha_5 \mu^2 + \alpha_6 \mu + \alpha_7} \quad (21)$$

Where:

$$\alpha_1 = n^2 \bar{Y}^2 D^o + 2n \bar{Y}^2 B^o D^o + \bar{Y}^2 B^{o2} D^o + n \bar{Y}^2 C^o D^o \\ \alpha_2 = 2n^2 \bar{Y}^2 D^o + 3n \bar{Y}^2 A^o D^o + 4n \bar{Y}^2 B^o D^o + 2\bar{Y}^2 A^o B^o D^o + 2B^{o2} D^o + 2n \bar{Y}^2 C^o D^o \\ \alpha_3 = n^2 \bar{Y}^2 D^o + 2n \bar{Y}^2 A^o D^o + \bar{Y}^2 A^{o2} D^o + 2n \bar{Y}^2 B^o D^o + 2\bar{Y}^2 A^o B^o D^o + \bar{Y}^2 B^{o2} D^o + \\ n \bar{Y}^2 A^o D^o + n \bar{Y}^2 C^o D^o \\ \alpha_4 = n^3 \bar{Y}^2 + 2n^2 \bar{Y}^2 B^o + n \bar{Y}^2 B^{o2} + n^2 \bar{Y}^2 C^o \\ \alpha_5 = 2n^3 \bar{Y}^2 + 3n^2 \bar{Y}^2 A^o + 4n^2 \bar{Y}^2 B^o + 2n \bar{Y}^2 A^o B^o + 2n \bar{Y}^2 B^{o2} + 2n^2 \bar{Y}^2 C^o - 2n B^o D^o \\ - n^2 D^o - B^{o2} D^o \\ \alpha_6 = n^3 \bar{Y}^2 + 3n^2 \bar{Y}^2 A^o + n \bar{Y}^2 A^2 + 2n^2 \bar{Y}^2 B^o + 2n \bar{Y}^2 A^o B^o + n \bar{Y}^2 B^{o2} + n^2 \bar{Y}^2 C^o \\ - n^2 D^o - 2n A^o D^o - 4n B^o D^o \\ \alpha_7 = n^2 D^o + 2n A^o D^o + A^{o2} D^o + 2n B^o D^o + 2A^o B^o D^o + B^{o2} D^o$$

4. Optimum Replacement

Now, we will determine the optimum fraction of the sample to be taken afresh in the second occasion that is μ , such that the population mean is determined with maximum precision. To achieve this, we minimize (21) with respect to μ . This gives us a polynomial of degree four as:

$$\delta_1 \mu^4 - \delta_2 \mu^3 - \delta_3 \mu^2 - \delta_4 \mu + \delta_5 = 0 \quad (22)$$

Where:

$$\delta_1 = \alpha_1 \alpha_4, \quad \delta_2 = 2\alpha_2 \alpha_4, \quad \delta_3 = \alpha_1 \alpha_6 - \alpha_2 \alpha_5 - 3\alpha_3 \alpha_4, \quad \delta_4 = 2\alpha_1 \alpha_7 + 2\alpha_3 \alpha_5, \quad \delta_5 = \alpha_3 \alpha_6$$

Solving (22) for μ we get

$$\hat{\mu} = 1, \quad \text{or} \quad \hat{\mu} = -1 \quad \text{or} \quad \hat{\mu} = \frac{-\beta_2 \pm \sqrt{\beta_2^2 - 4\beta_1\beta_3}}{2\beta_1}$$

$$\text{Where: } \beta_1 = \delta_1, \quad \beta_2 = 2\delta_1 - \delta_2, \quad \beta_3 = 3\delta_1 - 2\delta_2 - \delta_3.$$

μ exists in the quadratic equation $-\beta_2 \pm \sqrt{\beta_2^2 - 4\beta_1\beta_3}$, only if $\beta_2^2 \geq 4\beta_1\beta_3$. Although, there are four possible

values of μ , the admissible value must be the lowest value of μ that satisfies the condition: $0 \leq \hat{\mu} \leq 1$.

Substituting the admissible value of μ say $\hat{\mu}_o$ into equation (21), we get the optimum value of the mean square error as:

$$MSE_{opt}(T_{6P}) = \frac{\alpha_1 \mu_o^2 - \alpha_2 \mu_o + \alpha_3}{\alpha_4 \mu_o^3 - \alpha_5 \mu_o^2 + \alpha_6 \mu_o + \alpha_7} \quad (23)$$

5. Efficiency Comparison

In order to compare the efficiency of the proposed estimator with respect to: (i) the usual mean estimator, T_1 (ii) Cochran's (1977) estimator T_2 (iii) Singh's (2005) estimator T_3 (iv) Ralte and Das (2015) estimator T_4 (v) Tracy and Singh's (1999) estimator T_5 and (vi) Shabbiret *et al.* (2005) estimator T_6 , using a real life approach, a data from United Nations (UN) 2003 and 2004 was considered.

Data 1: We define:

Y= Country wise calories average supply per capita per day in 1999

X= Country wise calories average supply per capita per day in 1989

Z= Country wise proteins average supply per capita per day in 1999

Data 2: Let

Y= Country wise cellular mobile telephone subscribers in 2002

X= Country wise cellular mobile telephone subscribers in 2001

Z= Country wise population in 2002

Table 1:Input data

Notation	Data 1	Data 2	Notation	Data 1	Data 2
N	151	150	S_y^2	254320.50	4.7606 E+14
n	50	50	S_z^2	479.37	1.7973 +16
m	25	25	C_x	2.3028 E+00	6.3523E+00
u	25	25	C_y	1.8630E-01	3.9983E+00
\bar{X}	2651.53	4375893.4	C_z	2.9292E-01	3.9466E+00
\bar{Y}	2706.98	2706.98	ρ_{xy}	0.9679	0.9904
\bar{Z}	74.75	33969160.7	ρ_{yz}	0.8986	0.7149
S_x^2	372813.80	3.0012 E+14	ρ_{xz}	0.7006	0.6515

Table 2: Mean Square Error of Estimators

Estimator	MSE	
	Data 1	Data 2
T ₁	5086.41	9.5212E+12
T ₂	4172.18	5.4198E+12
T ₃	1280.14	9.51018 +12
T ₄	2599.82	9.51511E+12
T ₅	3544.05	5.2535E+12
T ₆	976.51	3.4259E+12
T_{6p}	471.73	7332532.4

By using Table 2, the percentage relative efficiencies of the various estimators are defined by:

Table 3: Percentage Relative Efficiencies of Estimators

Estimator	% Relative Efficiency (%RE)	
	Data 1	Data 2
T ₁	100	100
T ₂	121.9125253	175.6743791
T ₃	397.332667	100.1159142
T ₄	195.644723	100.063974
T ₅	143.519702	181.2353669
T ₆	520.7697269	277.9182113
T _{6p}	1078.25614	129848727.3

6. Conclusion

In this work, we proposed a new estimator in two occasion successive sampling for estimating the population mean of the study variable on current occasion using a single auxiliary variable. The properties (the bias and mean square error) of this estimator was obtained. Using a real life data, from table 2 and 3, the optimum mean square error of the proposed estimator was compared with that of (i) the usual mean estimator (ii) Cochran 1977 estimator (iii) Singh (2005) estimator (iv) Tracy and Singh (1999) estimator (v) Shabbir *et al.* (2005) Estimator and (vi) Ralte and Das (2015) Estimator. From this empirical results, it was concluded that the proposed estimator is more efficient in the estimation of the population mean than the other six already existing estimators.

References

- Cochran, W.G (1977). *Sampling Techniques*. Third Edition, John Wiley and Sons, New York
- Jessen, R. (1942). Statistical Investigation of a sample survey for obtaining farm facts. Iowa *Agricultural Experiment Station Road Bulletin no 204*, Ames, USA
- Okafor, F.C. and Arnab, R., (1987). Some Strategies of Two-Stage Sampling for Estimating Population Ratio over two occasions. *Austrian journal of statistics* 29(2):128-142
- Patterson, H.D. (1950). Sampling on Successive Occasions with partial Replacement of Units. *Journal of the Royal Statistical Society B*, 12, 241-255.
- Ralte, Z. and Das, G. (2015). Ratio - to-Regression Estimator in Successive Sampling Using one Auxilliary Variable. *Statistics in Transition New Series*, 16(2), 183-202
- Shabbir, J., Azam, M., and Gupta, S. (2005). Estimation of Current Population Mean in Successive Sampling. *Proc. ICCSC-VIII, Lahore, Pakistan*, 19(22), 77-85
- Sharma, V. and Kumar, S. (2021). Improved Estimators for Estimating the Population Mean in Two Occasion Successive Sampling. *Journal of Turkish Statistical Association*, 13(2):39-51
- Singh, G.N. (2005). On the use of chain-type ratio estimator in successive sampling. *Stat.Trans.* 7, 1, 21-26

Tracy. D.S and Singh, H.P. (1999). Efficient use of Two Auxiliary Variables in Two Phase Sampling as well as in Successive Sampling. *Pak. J. Statist.* 15(1), 27-39

United Nations, UN. (2003): *Industrial Commodity Statistics Yearbook*. New York, USA

United Nations, UN.(2004): *Statistical Yearbook*. New York, USA

NCPS-2024

Theoretical investigation of metals (Ni, Zn, Fe) doped zeolites for adsorption of volatile gases (Toluene, Xylene and Formaldehyde)

Bassey E. Inah, & Stanley J. Oduma

¹Department of Chemistry, Faculty of Physical Sciences, University of Calabar, Calabar, Nigeria

*Corresponding author's email: basseyinah058@gmail.com

Abstract

Given the influence of volatile organic compounds on human health, there is an urgent need for the removal of such chemicals from the environment, which calls for the design of materials for the removal of such substances from the environment. This study provides a detailed analysis on the sensitivity, selectivity, conductivity, and adsorption abilities of metal-doped zeolites (Ni, Zn, Fe) towards volatile organic compounds such as toluene, xylene, and formaldehyde. Employing density functional theory (DFT) computational techniques, the electronic and structural properties of these systems were thoroughly investigated. Frontier molecular orbital (FMO) analysis revealed that Fe-Zif-HCHO and Zn-Zif-HCHO, with the lowest energy gaps of 2.6643 eV and 3.2537 eV respectively, are likely to be more reactive in the case of formaldehyde, whereas Ni-Zif-Tol (3.5247 eV) and Ni-Zif-Xy (3.5054 eV) are likely to be more reactive towards toluene and xylene. Zn-Zif-Xy, on the other hand, has the highest total energy gap of 5.3381 eV, indicating greater stability and reduced reactivity. Natural bond orbital (NBO) analysis showed that Zn-Zif-HCHO had the highest stabilization energy for formaldehyde (240.23 kcal/mol), while Zn-Zif-Tol had the highest stabilization energy for toluene (212.00 kcal/mol). Ni-Zif-Xy had the highest total perturbation energy, measuring 363.80 kcal/mol, suggesting excellent stabilization and high reactivity within the system. The adsorption study revealed that Ni-Zif-HCHO is the best sensor material for formaldehyde (HCHO) detection, Fe-ZIF-Tol was the best at sensing toluene gas, and Zn-ZIF-Xy was the most effective at sensing xylene. Notably, the zinc-doped zeolite (Zn-Zif) demonstrated outstanding sensing capabilities for two out of the three VOCs tested, making it a highly versatile sensor material. This research highlights the potential of metal-doped zeolites in environmental applications, particularly for the adsorption and detection of VOCs. The findings contribute to the understanding of the electronic and structural factors governing the adsorption properties of zeolite-based materials, paving the way for the development of more efficient and selective adsorbents for environmental remediation.

Keywords: Metal-doped zeolites, DFT, Adsorption, Volatile organic compounds (VOCs)

1.0 Introduction:

Volatile organic compounds (VOCs) are carbon-based compounds that evaporate easily at room temperature, originating from both natural sources—like plant emissions (Khatib & Haick, 2022), marine environments, and wildfires—and human activities, including industrial processes, transportation, and residential habits. (Hoque, 2023) VOCs are particularly concentrated in urban and industrial areas, leading to significant health impacts based on exposure levels and individual susceptibility. Short-term exposure to high VOC concentrations can cause headaches, dizziness, and respiratory issues, while prolonged exposure (Zolghadri et al., 2023) is linked to severe health problems such as immune dysfunction, neurological disorders, and cancer, particularly in vulnerable groups like children and the elderly.

Zeolites, crystalline aluminosilicates with unique adsorption, ion-exchange, and catalytic properties, are frequently used in environmental applications, including the removal of heavy metals and organic pollutants from water and soil (Rathiet *al.*, 2024). This study employs density functional theory (DFT) to model metal-doped zeolite structures (Zn, Ni, Fe), examining how metal doping affects their structural and adsorption properties. Key analyses include frontier molecular orbitals (HOMO-LUMO), electronic transport, and Bader's QTAIM to understand sensor sensitivity and conductivity shifts upon gas adsorption

(Javed *et al.*, 2024). These insights advance our understanding of doped zeolites' electronic properties and their potential as VOC sensors.

2.0 Computational details:

The computational calculations were conducted using GaussView 6.0.16 and Gaussian 16. Geometry optimization and energy calculations utilized Density Functional Theory (DFT) with the B3LYP method and LanL2DZ basis set. Bond lengths before and after adsorption were analyzed to understand changes in bond strength, key to assessing sensor behavior. Frontier molecular orbital analysis was performed on the optimized geometry using Koopman's approximation (Linscott *et al.*, 2023). Additionally, Non-Covalent Interactions (NCI) and Quantum Theory of Atoms in Molecules (QTAIM) analyses were applied at the same theoretical level to enhance non-covalent interaction accuracy and explore inter-atomic interactions, using the NBO 3.0 integration and Multiwfn 3.7. HOMO-LUMO visualizations were generated in Chemcraft, while adsorption energies were calculated as:

$$E_{\text{ADS}} = E_{\text{COMPLEX}} - (E_{\text{ADSORBATE}} + E_{\text{ADSORBENT}})$$

1

3.0 Results and discussion

3.1 Geometry optimization:

Assessing the adsorption capabilities of metal-doped zeolites, the optimized structures of these zeolites with gases are presented in Figure 1 (supporting information). Bond lengths before and after adsorption (Table 1) reveal notable changes, indicating structural shifts upon gas interaction. For Fe-ZIF-8-HCHO, bond lengths varied, with Fe25-O9 increasing from 1.826 Å to 1.832 Å and Ga1-H19 from 1.554 Å to 1.555 Å, while P₆-O₈ decreased from 1.633 Å to 1.624 Å. Similar trends were observed in Ni-ZIF-8-HCHO, where Ni25-O9 increased from 1.841 Å to 1.886 Å, indicating bond stretching and stronger adsorption interactions.

In Zn-ZIF-8, all selected bonds lengthened slightly, such as Zn25-O7 from 1.917 Å to 1.932 Å. Fe-ZIF-8-tol and Zn-ZIF-8-tol showed minimal bond length changes, suggesting stable adsorption. In Ni-ZIF-8-xy, increments in bonds like Ni25-O9 (1.841 Å to 1.863 Å) suggest stronger adsorption interactions, while Zn-ZIF-8-xy showed variability, with Zn25-O7 increasing to 1.949 Å, hinting at enhanced gas-zeolite interaction and minor structural adjustments.

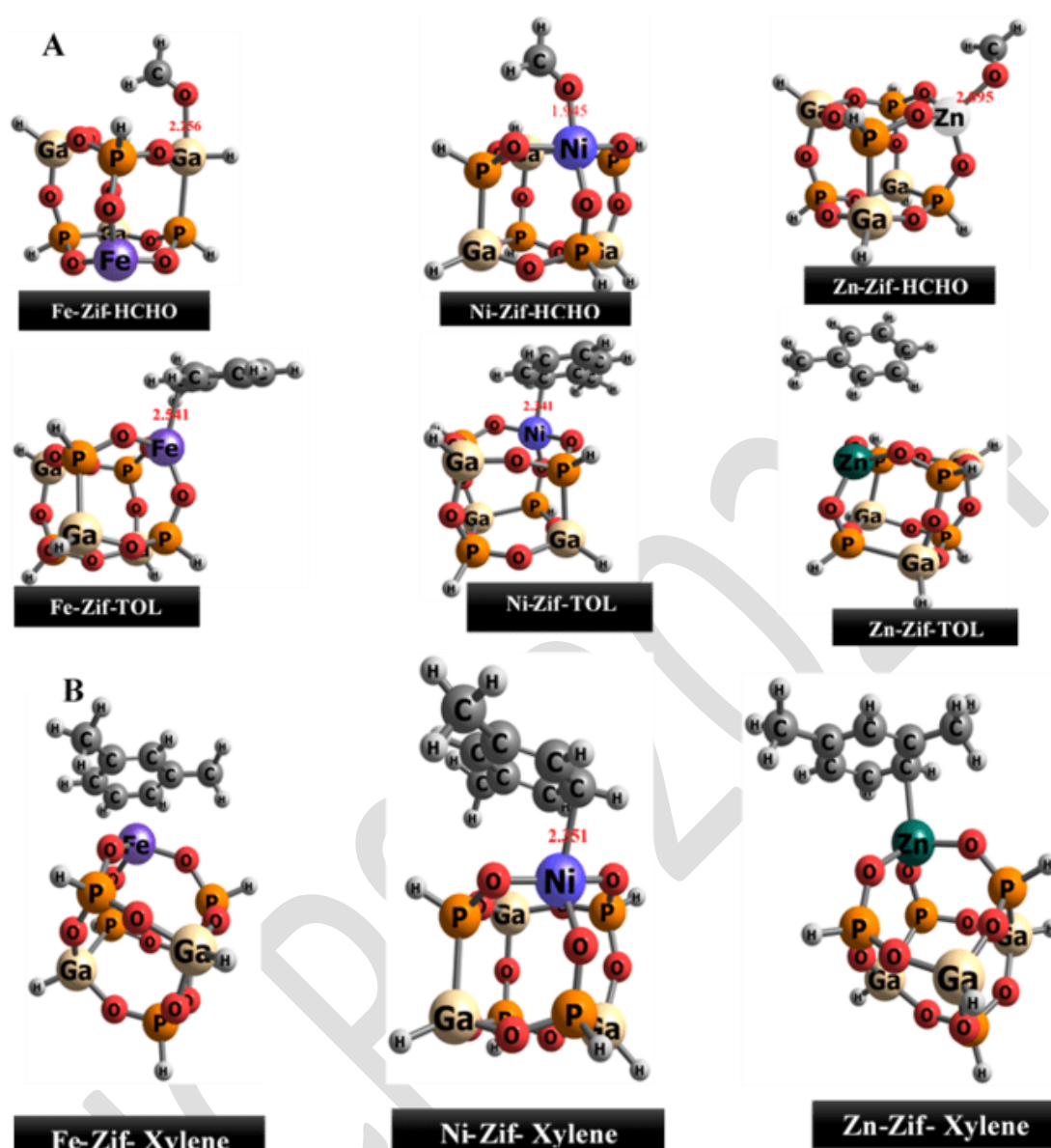


Figure 1A, B: Optimized structures of the doped metal zeolites and the gases calculated at DFT/B3LYP/LanL2DZ basis set

Table 1: Observed bond lengths of the chosen bonds in doped metal zeolites both before and after formaldehyde, toluene, and xylene were adsorbed.

BEFORE			AFTER		
Systems	Bonds	Bond lengths(Å)	Systems	Bonds	Bond lengths(Å)
	Fe25-O9	1.82628		FE25-O9	1.83201
FE-ZIF-8	P6-O8	1.63337	FE-ZIF-8-HCHO	P6-O8	1.6248
	P13-O9	1.6883		P13-O9	1.68397

	Ga1-H19	1.55478		Ga1-H19	1.55574
				Ga16-O29	2.5568
	FE25-O9	1.82628		FE40-O9	1.87512
	GA14-O15	1.79836		FE40-H34	2.55312
FE-ZIF-8	P6-O17	1.64015	FE-ZIF-8-TOL	GA14-O15	1.79583
	GA1-O3	1.78484		P6-O17	1.63946
				GA1-O3	1.78724
	Fe25-O9	1.82628		FE25-O9	1.8843
	P13-O15	1.66385		FE25-C27	2.69895
FE-ZIF-8	Ga16-H21	1.55473	FE-ZIF-8-XY	P13-O15	1.66898
	P12-H24	1.43504		Ga16-H21	1.55758
				P12-H24	1.43664
	Ni25-O9	1.84138		NI25-O9	1.88692
	Ga16-H21	1.55373		Ga16-H21	1.55555
NI-ZIF-8	P13-O9	1.70616	NI-ZIF-8-HCHO	P13-O9	1.6951
	P2-O4	1.63549		P2-O4	1.63361
				Ni25-O29	1.9448
	NI25-O9	1.84138		NI25-O9	1.859
	P2-O11	1.63381		NI25-C27	2.61107
NI-ZIF-8	Ga14-O10	1.84882	NI-ZIF-8-TOL	P2-O11	1.63332
	P6-O17	1.64214		Ga14-O10	1.84715
				P6-O17	1.63882
	Ni25-O9	1.84138		NI25-O9	1.86375
	P2-H20	1.40823		NI25-C27	2.3506
NI-ZIF-8	P13-H18	1.43289	NI-ZIF-8-XY	P2-H20	1.41138
	Ga1-O3	1.79422		P13-H18	1.4405
				Ga1-O3	1.78639
	Zn25-O7	1.91709		Zn25-O7	1.9329
	Ga14-H22	1.55311		Ga14-H22	1.55457
Zn-ZIF-8	P12-O3	1.67973	Zn-ZIF-8-HCHO	P12-O3	1.68072
	P6-H23	1.41304		P6-H23	1.41584
				Zn25-O29	2.09461
ZN-ZIF-8	Zn25-O7	1.91709	ZN-ZIF-8-TOL	Zn25-O7	1.91636
	P12-O3	1.67973		Zn25-H34	3.75767

	Zn25-O8	1.9446		P12-O3	1.67969
	Ga14-O5	1.87651		Zn25-O8	1.94365
				Ga14-O5	1.87498
	Zn25-O7	1.91709		Zn25-O7	1.94979
	P6-O17	1.64189		Zn25-C28	2.39263
Zn-ZIF-8	P13-O9	1.67324	Zn-ZIF-8-Xy	P6-O17	1.64289
	Ga1-H19	1.55624		P13-O9	1.66509
				Ga1-H19	1.55837

3.2 HOMO-LUMO Analysis:

HOMO-LUMO analysis (Table 2) provides insights into the reactivity of metal-doped zeolites in sensing formaldehyde, toluene, and xylene by examining the energy gap, EG (Katre *et al.*, 2024), ionization potential, IP, and electron affinity, EA. The HOMO signifies electron donation potential, while the LUMO represents electron acceptance capacity. A smaller Eg suggests higher reactivity due to easier electron transition, as seen in Fe-Zif-HCHO (2.6643 eV) and Zn-Zif-HCHO (3.2537 eV), making them particularly reactive toward formaldehyde. In contrast, Zn-Zif-Xy with the largest Eg (5.3381 eV) indicates higher stability and lower reactivity.

IP values suggest reactivity potential, with lower values implying easier electron removal; EA values highlight electron-accepting ability (Hassan *et al.*, 2023). Chemical softness (σ), indicating sensitivity, and electrophilicity (ω), signaling electron-accepting propensity, further characterize reactivity. Fe-Zif-HCHO and Zn-Zif-HCHO show high electrophilicity and chemical softness, suggesting strong formaldehyde sensitivity, while Ni-Zif-Tol and Ni-Zif-Xy perform well for toluene and xylene. High electrophilicity correlates with smaller Eg, enhancing suitability for electron-donating analytes, ideal for sensor applications.

Table 2: Quantum Descriptors and HOMO-LUMO Values of Metal-Doped Zeolites

Systems	Homo	lumo	Eg	IP	EA	σ	η	μ	Ω	χ
Fe-Zif	-6.3794	-1.9415	4.4379	6.3794	1.9415	0.2253	2.2189	-4.1604	3.9004	4.1604
Ni-Zif	-6.6233	-3.0991	3.5241	6.6233	3.0991	0.2838	1.7621	-4.8612	6.7055	4.8612
Zn-Zif	-6.9539	-1.7785	5.1753	6.9539	1.7785	0.1932	2.5877	-4.3662	3.6836	4.3662
Fe-Zif-HCHO	-5.9498	-3.2855	2.6643	5.9497	3.2855	0.3753	1.3321	-4.6176	8.0031	4.6176
Ni-Zif-HCHO	-6.7792	-3.0605	3.7187	6.7791	3.0605	0.2689	1.8594	-4.9198	6.5088	4.9198
Zn-Zif-HCHO	-6.5863	-3.3325	3.2537	6.5862	3.3326	0.3073	1.6268	-4.9594	7.5594	4.9594
Fe-Zif-Tol	-5.9359	-1.7976	4.1383	5.9359	1.7976	0.2416	2.0692	-3.8667	3.6129	3.8667
Ni-Zif-Tol	-6.6238	-3.0991	3.5247	6.6238	3.0991	0.2837	1.7623	-4.8615	6.7052	4.8615
Zn-Zif-Tol	-6.3299	-1.8221	4.5078	6.3299	1.8221	0.2218	2.2539	-4.076	3.6855	-4.076
Fe-Zif-Xy	-5.8077	-1.4607	4.3470	5.8077	1.4607	0.2300	2.1735	-3.6342	3.0383	-3.6342

Ni-Zif-Xy	-6.5745	-3.0692	3.5054	6.5745	3.0692	0.2853	1.7527	-4.8219	6.6328	-4.8219
Zn-Zif-Xy	-6.6042	-1.2661	5.3381	6.6042	1.2661	0.1873	2.6690	-3.9352	2.9009	-3.9352

3.3 Natural Bond Orbital (NBO) analysis:

Natural Bond Orbital (NBO) analysis was conducted to explore the sensitivity and selectivity of doped metal zeolites toward various gases by examining intra- and intermolecular charge transfer interactions (Timothy *et al.*, 2023). This evaluation, performed at the DFT/B3LYP/LanL2DZ level, assesses donor-acceptor orbital interactions using stabilization energy ($E^{(2)}$). Where higher ($E^{(2)}$) values indicate stronger adsorption interactions. This analysis is suggestive of the mathematical relation:

$$E^{(2)} = \Delta E_{ij} = -q_i \frac{F^2(i,j)}{\epsilon_i - \epsilon_j} \quad 2$$

Where q_i is the donor orbital occupancy, ϵ_i and ϵ_j represent the diagonal elements and $F_{(i,j)}$ stand for the off-diagonal elements of Fock matrix

The NBO analysis identified key interactions in Fe-Zif, Ni-Zif, and Zn-Zif systems. Fe-Zif-HCHO, Ni-Zif-HCHO, and Zn-Zif-HCHO exhibited stabilization energies of 183.90, 141.11, and 240.23 kcal/mol, respectively, indicating strong formaldehyde reactivity. The highest stabilization for toluene and xylene was observed in Zn-Zif-Tol (212.00 kcal/mol) and Ni-Zif-Xy (363.80 kcal/mol), respectively, with Ni-Zif-Xy showing the overall highest perturbation energy, signifying robust stabilization and reactivity. These findings underscore Zn-Zif-HCHO and Ni-Zif-Xy as particularly promising for gas adsorption.

3.4 Non-Covalent Interaction (NCI) Analysis:

Non-Covalent Interaction (NCI) analysis was performed to examine and quantify weak interactions, including van der Waals forces, hydrogen bonds, π - π interactions, and electrostatic interactions between the sensor materials and target gas molecules (Siddique *et al.*, 2024). Such interactions impact the sensor's sensitivity and selectivity. RDG analysis, through clusters, Reduced Density Gradient spikes, and λ_2 values, revealed the interaction types and strengths.

The analysis showed that van der Waals interactions ($\rho > 0$, $\lambda_2 < 0$) have close-to-zero density, while hydrogen bonds have more negative λ_2 and higher density values. Regions of strong repulsion ($\rho > 0$, $\lambda_2 > 0$) were also noted. Using Visual Molecular Dynamics (VMD) software, color-filled iso-surface structures (Figure 2) display these interactions. Green iso-surfaces indicate van der Waals forces in the zeolite center and at the zeolite-gas interface, while red iso-surfaces, representing steric repulsion, are observed around GA-O bonds. Additionally, steric repulsion and hydrogen bonding are observed across all studied systems.

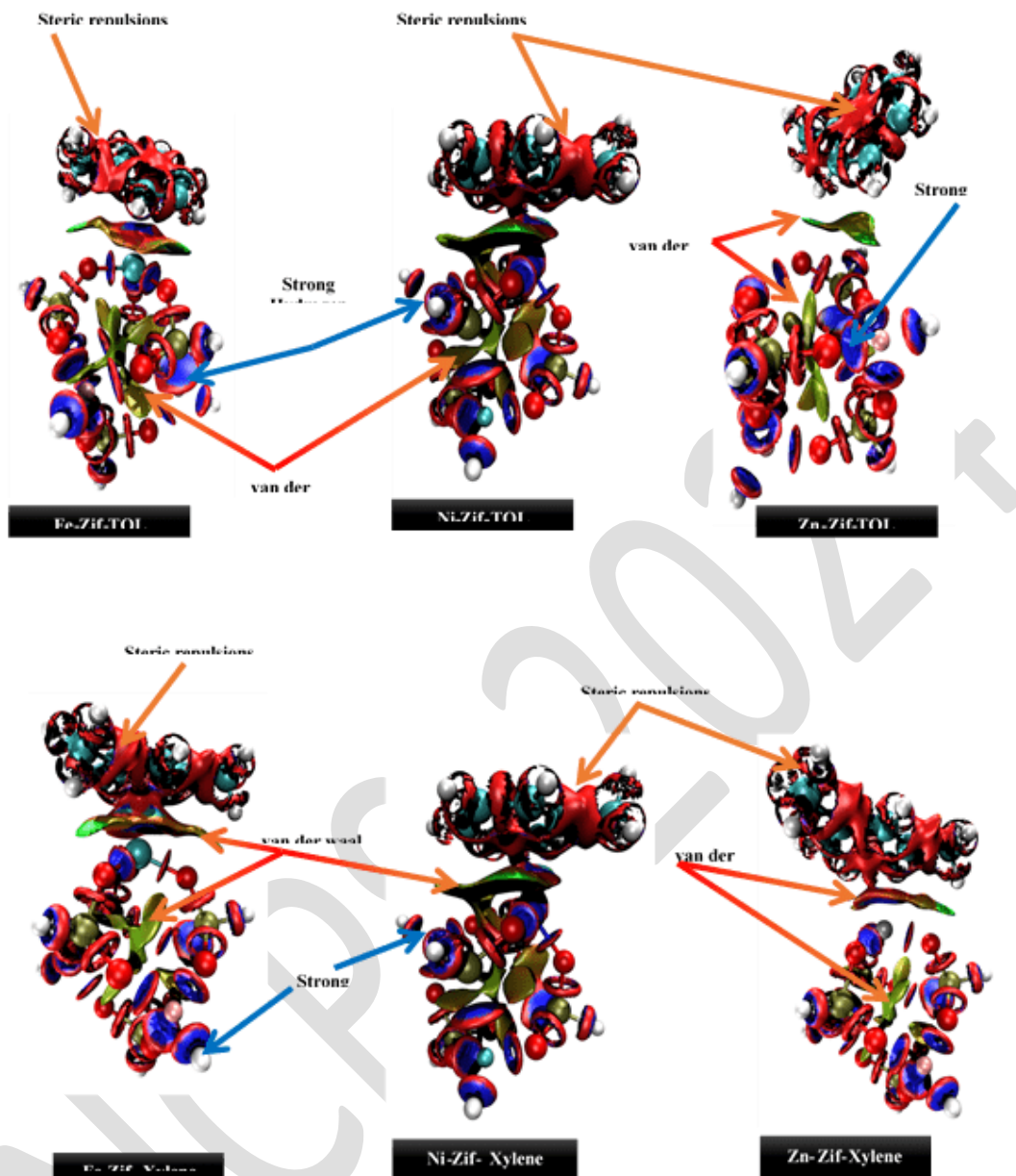


Figure 2: Visual depiction of the non-covalent interaction analysis

3.5 Quantum Theory of Atoms in Molecules (QTAIM) Analysis:

QTAIM analysis also known as the Bader's analysis provides insights into molecular electronic structure by examining electron density (ρ) and its Laplacian ($\nabla^2\rho$) across bond critical points (BCPs). A negative $\nabla^2\rho$ indicates shared-shell (covalent) bonds, while a positive value suggests closed-shell (ionic or van der Waals) interactions (Zhang, 2024). High electron density (ρ) at BCPs correlates with stronger interactions and stability.

The analysis uses parameters like the energy density (H), kinetic energy (G), and potential energy (V) to distinguish bond types. A G/V ratio ≥ 1 implies non-covalent interaction, while $G/V < 0.5$ indicates covalent bonding. The Electron Localization Function (ELF) further differentiates interaction types; values closer to 1 suggest localized bonds, while values near zero indicate delocalization (Steinmann *et al.*, 2011).

Results show predominantly positive $\nabla^2\rho$ in BCPs, signaling closed-shell interactions, except in Fe-Zif-Xy (C40-H4) and Zn-Zif-Tol (P13-H18), where negative $\nabla^2\rho$ confirms covalent bonding. These findings align with calculated G/V values and are illustrated in Figure 2 of the supporting information.

3.6 Adsorption Studies:

The relative adsorption study captured in Table 3 evaluated doped metals' adsorption capabilities, distinguishing physisorption weak adsorption type which is often associated with endothermic reaction. Chemisorption strong adsorption, is associated with exothermic. Positive adsorption energy indicates an endothermic process with weak adsorbent-adsorbate interactions, while negative adsorption energy reflects exothermic chemisorption with stronger interactions (Djioko *et al.*, 2024).

Adsorption energies include: Fe-Zif-HCHO (-0.02), Ni-Zif-HCHO (0.44), Ni-Zif-HCHO (-1.09), Fe-Zif-Tol (-0.29), Ni-Zif-Tol (0.77), Zn-Zif-Tol (0.09), Fe-Zif-Xy (-0.35), Ni-Zif-Xy (0.72), and Zn-Zif-Xy (-0.68). Ni-Zif-HCHO exhibited the strongest affinity for formaldehyde (HCHO), Fe-Zif-Tol was most effective for toluene, and Zn-Zif-Xy showed high sensitivity to xylene. Zinc-doped zeolite (Zn-Zif) proved versatile, effectively detecting two of the three gases, marking it as a promising sensor material.

Table 3: Calculated adsorption energies, Fermi energies, charge transfer, fraction of electron transfer, back donation, recovery time and work function of the studied systems

Systems	E_{ads}	Fermi energy(EFL)	Dipole Moment	Charge transfer investigation(Q t)	Fraction of electron transfers(ΔN)	Back donation (ΔE)	Recovery time(t)	Work function(Φ)
Fe-Zif-HCHO	-0.02	-4.6177	6.5505	1.4459	-0.2578	-0.3330	1.0081E-12	4.6177
Ni-Zif-HCHO	0.44	-4.9198	2.3197	1.4288	-1.8863	-0.4648	8.3735E-13	4.9198
Zn-Zif-HCHO	-1.09	-4.9594	4.7669	1.9540	-0.3087	-0.4067	1.5523E-12	4.9594
Fe-Zif-Tol	-0.29	-3.8668	2.5529	-0.7651	0.9809	0.5173	8.9244E-11	3.8668
Ni-Zif-Tol	0.77	-4.8615	0.5728	-1.1258	0.075	0.4406	1.1831E-30	4.8615
Zn-Zif-Tol	0.09	-4.0760	3.2234	-1.2716	0.0791	0.5635	5.2484E-18	4.0760
Fe-Zif-Xy	-0.35	-3.6342	3.3320	-1.1626	0.1726	0.5434	1.0674E-09	3.6342
Ni-Zif-Xy	0.72	-4.8219	0.9054	-1.1402	0.0019	0.4382	1.1059E-29	4.8219
Zn-Zif-Xy	-0.68	-3.9352	3.8984	-1.4803	-0.0196	0.6673	0.0011	3.9352

4.0 Conclusion

This study analyzed the sensitivity, selectivity, and adsorption properties of metal-doped zeolites (Ni, Zn, Fe) for volatile organic compounds (VOCs) like toluene, xylene, and formaldehyde, using density functional theory (DFT) calculations. HOMO-LUMO analysis showed Fe-Zif-HCHO and Zn-Zif-HCHO with the lowest energy gaps, indicating high reactivity with formaldehyde, while Ni-Zif-Tol and Ni-Zif-Xy were more reactive toward toluene and xylene, respectively. Zn-Zif-Xy, with the highest energy gap (5.3381 eV), displayed the most stability.

Natural bond orbital analysis revealed strong stabilization energies in Zn-Zif-HCHO for formaldehyde and Zn-Zif-Tol for toluene, while Ni-Zif-Xy had the highest perturbation energy (363.80 kcal/mol), indicating excellent stabilization and reactivity. QTAIM and NCI analysis confirmed non-covalent interactions. Relative adsorption

studies identified Ni-Zif-HCHO as the best for formaldehyde detection, Fe-Zif-Tol for toluene, and Zn-Zif-Xy for xylene. Notably, Zn-Zif effectively detected two VOCs, showcasing its adaptability.

This research highlights the potential of metal-doped zeolites in environmental applications, advancing understanding of the electronic and structural factors governing zeolite-based VOC adsorption, and supporting the development of selective adsorbents for environmental remediation

References:

- Djioko, F. H. K., Fotsop, C. G., Youbi, G. K., Nwanonenyi, S. C., Madu, C. A., & Oguzie, E. E. (2024). Unraveling the sorption mechanisms of ciprofloxacin on the surface of zeolite 4A (001) in aqueous medium by DFT and MC approaches. *Applied Surface Science Advances*, 19, 100542.
- Hassan, A. U., Sumrra, S. H., Li, Y., Mustafa, G., Zafar, W., & Noreen, S. (2023). Enriching the electronic populations of D- π -A type ON-OFF solar switches with sp² hybridization: A TD-DFT possibility to create pull-push effect. *Computational and Theoretical Chemistry*, 1228, 114295.
- Hoque, M. (2023). Unveiling the ripple effect: how human activities reshape ecosystems
- Javed, M., Khan, M. U., Hussain, R., Abbas, F., & Ahamad, T. (2024). Deciphering the detection and electrochemical sensing of the environmental pollutant CO gas with Ga₁₂As₁₂ and Al₁₂As₁₂ nanostructured materials: an insight from first-principle calculations. *Journal of Materials Science*, 59(11), 4548-4570.
- Katre, S., Baghmare, P., & Giri, A. S. (2024). Photocatalytic nanomaterials and their implications towards biomass conversion for renewable chemical and fuel production. *Nanoscale Advances*.
- Khatib, M., & Haick, H. (2022). Sensors for volatile organic compounds. *ACS nano*, 16(5), 7080-7115.
- Linscott, E. B., Colonna, N., De Gennaro, R., Nguyen, N. L., Borghi, G., Ferretti, A., ... & Marzari, N. (2023). koopmans: An open-source package for accurately and efficiently predicting spectral properties with Koopmans functionals. *Journal of Chemical Theory and Computation*, 19(20), 7097-7111.
- Rathi, B. S., Kumar, P. S., Susana, J. N. I., Virgin, J. F., Dharani, R., Sanjay, S., & Rangasamy, G. (2024). Recent research progress on the removal of heavy metals from wastewater using modified zeolites: A critical review. *Desalination and Water Treatment*, 100573
- Siddique, S. A., Siddique, M. B. A., Ahmed, E., Ullah, A., Rauf, A., ALI, M. A., ... & Arshad, M. (2024). Quantum Chemical Investigation for Enhanced Electrochemical Sensing of Toxic Gases by HexaazaphenH₂. *New Journal of Chemistry*.
- Steinmann, S. N., Mo, Y., & Corminboeuf, C. (2011). How do electron localization functions describe π -electron delocalization?. *Physical chemistry chemical physics*, 13(46), 20584-20592.
- Timothy, R. A., Louis, H., Adindu, E. A., Gber, T. E., Agwamba, E. C., Offiong, O. E., & Pembere, A. M. (2023). Elucidation of collagen amino acid interactions with metals (B, Ni) encapsulated graphene/PEDOT material: insight from DFT calculations and MD simulation. *Journal of Molecular Liquids*, 390, 122950.
- Zhang, X. (2024). Theoretical studies of the nature of actinide bonds.

Zolghadri, S., Rahimpour, E., & Rahimpour, M. R. (2023). Short-term and long-term health problems in exposure to chemicals. In *Crises in Oil, Gas and Petrochemical Industries* (pp. 239-253). Elsevier.

NCPS-2024

Physicochemical properties of Nigerian-grown *Pachira aquatica* seed oil and its potentials for national socio-economic recovery

Bassey Sam-Uket Okori¹ and Michael Akomaye Akpe²

1). Department of Chemical Sciences, Clifford University, Abia State, 2). Department of Pure & Applied Chemistry, University of Calabar, Calabar.

Corresponding Email: sambassey28@gmail.com

Abstract

Edible vegetable oils are prone to quality deterioration through oxidation and microbial degradation resulting in nutritional loss and off-flavors. The aim of this study is to explore the physicochemical parameters of *pachira aquatica* seed oil and determine if it compares favorably with common seed oils with a view to recommending its large-scale cultivation to boost socio-economic growth. Standard methods of analysis by the Association of Official Analytical Chemists (AOAC) were employed. The mean value for: % Yield was (46.00±2.00), moisture content (0.2000 ± 0.00011), specific gravity at 30 °C (0.9500 ± 0.00001), peroxide value (0.5700 ± 0.09504 meq/kg), acid value (0.0233 ± 0.00050 mg KOH/g oil), iodine value (13.5700 ± 0.13856 g I₂/100 g oil), and saponification value (4.5800 ± 0.30600 KOH/g). The results highlight that all tested physicochemical parameters of the oil samples were within the allowable standards for edible oils. The oil has shown great potential in replacing common vegetable oils such as groundnuts and olive, to mention but a few, and as such can be employed for use both domestically and commercially to boost our socio-economic growth only if harnessed.

Keywords: Properties, *Pachira aquatica*, Seed oil, Potentials, Socio-economic, Recovery.

Introduction

The need for non-conventional sources of oil to supplement the ones that are available has grown globally. This may or may not be related to the requirement to close the gap between the supply and demand for commercially available oils and fats (Dansie et al., 2012). Due to their high nutritive and calorific content, seeds are highly sought after for both industrial and human use. They are also excellent sources of edible fats and oils. As a result, numerous studies on alternative vegetable oil sources have been conducted (Odoemelam et al., 2005).

Nigeria produces a wide variety of agricultural products because of its excellent soil, conducive climate, and large amount of arable land. Edible oils are primarily derived from oil seeds, one of the main agricultural products. Nonetheless, some plant species are neglected or underutilized, particularly those that are not grown for food (Kyari, 2008). Oils and fats naturally contain lipids and triacylglycerol. Glycerides and saturated and unsaturated fatty acids make up their chemical composition. According to Afolayan et al. (2019), edible oils are essential components of our daily diet that give us energy, essential fatty acids, and fat-soluble vitamins.

The tree species *Pachira aquatica* is indigenous to tropical wetlands that stretch from northeastern Brazil and Guyana to southern Mexico. This species is a member of the subfamily Bombacaceae. Nonetheless, it is a member of the Malvaceae family as per the Angiosperm phylogeny II group classification. It is marketed as a "money tree" (Silver et al. 2015, Cheng et al. 2017) and goes by several names, including Malabao chestnut, French peanut, Guyana chestnut, monguba, false cacao, or pumpo (guatamela). It is said to bring good fortune in business. In Nigeria, it is not common, and a few who planted it see it as ornamental because of little or no knowledge about its use or potential. Thus, it does not even have native names like other trees.

When the *Pachira aquatica* plant is younger than a year old, it resembles the cassava plant, particularly in its leaves. It has smooth green stems with a distinctively enlarged and swollen base and bright green palm leaves with five to nine lanceolate leaflets. The petals are amazing, featuring long, narrow flowers and hair that resembles orange and yellow stems. *Pachira aquatica* can grow to a height of six to eighteen meters in its natural habitat, which receives full sun or partial shade (Oliveria et al. 2000). Oval-shaped and encased in a brown, woody shell, *Pachira aquatica* fruits have large seeds that are rich in lipids, proteins, and carbohydrates, as well as calcium, magnesium, and potassium. These characteristics make them an intriguing subject for research on food and biodiesel production. (Polizi et al., 2008; Azevedo, 2008; Silva, 2010).

The aim of this study is to explore the physicochemical parameters of *Pachira aquatic* (a plant not native and common to Nigeria) and see if it compares favorably with common seed oils with a view to recommending its large-scale cultivation to boost socio-economic growth.

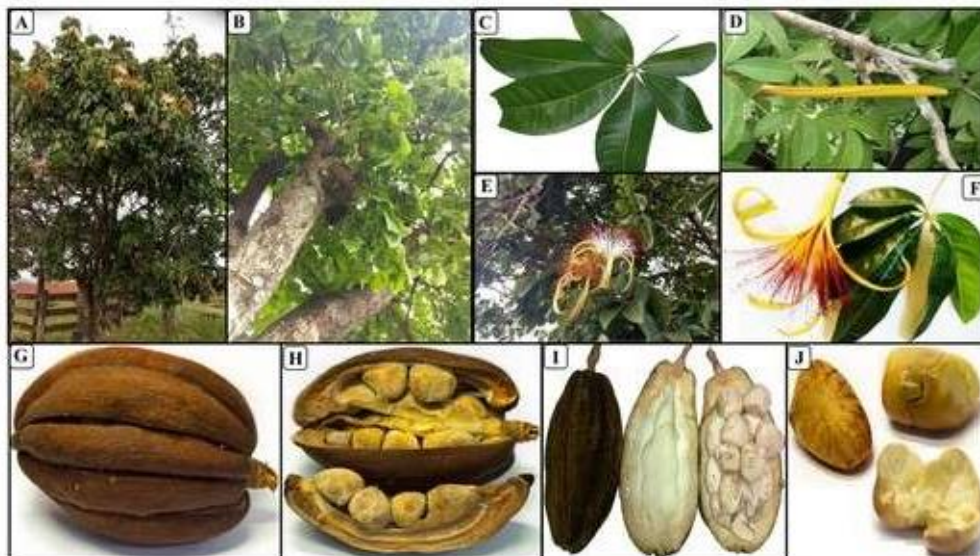


Fig 2: *Pachira aquatic* Aubl. a. Tree and flowers; b. Trunk and canopy; c. Leaves; d. Flower and bud; e and f. Leaves and flowers; g. Ripe fruit; h. Open ripe fruit and seeds; i. Immature fruit (left and right: unripe open fruit and seeds); j. Seeds (left: seed with endocarp, right-side up: seed without endocarp and right-side below: open seed).



Fig 2: Photograph of *Pachira aquatic*: (a) *Pachira aquatic* pod; (b) *Pachira aquatic* leaves. Source: Image captured by the researcher

Materials and methods

Materials

The plant material used for this work is *Pachira aquatic* seed oil extract

Chemicals and Reagents

Chemicals and reagents used for this work was of analytical grade and were products of Sigma Chemical Company, Mayer and Baker Ltd, Fluker and BDH. The following chemicals and reagents were used for this study;

hexane, chloroform, potassium hydroxide, hydrochloric acid, phenolphthalein, sodium hydroxide, diethyl ether, ethanol, iodine, sodium thiosulphate, potassium iodide, glacial acetic acid, starch.

Equipment and Apparatus

Soxhlet extractor, Rotary evaporator, Hot air oven, Heating mantle, Porcelain crucibles, digital weighing balance, extraction thimble, Buckner funnel, Whatman filter paper No 1, filtration crucible, Muffle furnace, Desiccator, Dryer, Beakers, Conical flask, Measuring cylinders.

Methods

Sample Collection / Identification

Pachira aquatica leaves and seeds were obtained from Ihie, Isialangwa North Local Government Area of Abia State, Nigeria and transported to Chemistry Laboratory of Clifford University, Owerinta, Abia State, Nigeria. Taxonomic identification/authentication of the leaves and seeds was first done using Google imaging before further identification at the Department of Plant Science and Biotechnology, Micheal Okpara University of Agriculture Umudike. The plant was assigned herbarium number and deposited at the herbarium. Further analysis was carried out at the Chemistry Laboratory of Clifford University, Abia State-Nigeria.

Sample Preparation

The mature seeds were first soaked in clean water for 12 hours to ease the removal of the seed from its capsule after which the seeds were sun dried for seven consecutive days. The dried seeds were then pulverized into nearly smooth granules with a blender and kept in a sealed container prior to extraction.

Sample Extraction

The Soxhlet extractor was employed for the extraction of the seed oil. An empty Soxhlet flat bottom flask was weighed and recorded as W. 250 ml of hexane was poured into the Soxhlet flat bottom flask and placed on a heating mantle. 50 g of the sample was weighed and transferred into an extraction thimble, which was then placed in the Soxhlet extraction chamber. The condenser was fitted above the extraction chamber, and the water inlet and outlet were connected to ensure cooling and condensation. The heating mantle temperature was then set at 60 °C. When the solvent was boiling, the vapor rose through the vertical tube into the extraction chamber and condensed at the top. The hot liquid condensate drips into the extraction thimble in the chamber, which contains the sample to be extracted. The hot condensate extracts the oil from the sample, which seeps through the pores of the thimble into the extraction chamber. When the chamber becomes full of the extract, it is refluxed into the flat bottom flask through the siphon tube. This process is allowed to continue repeatedly until the sample extracts in the chamber become clearly and free from any color. Afterwards, the process was stopped. The oil extract was then concentrated and the solvent recovered using a rotary evaporator at a temperature of 40 °C. The oil extract was then stored in a sterile plastic container for further analysis.

Percentage Yield: This indicates the amount of oil in the seed. It was determined using the equation below:

$$\% \text{ Yield} = \frac{\text{Weight of oil extract} \times 100\%}{\text{Weight of Sample}} \quad (1)$$

Determination of Saponification Value (SV)

Two grammes (2g) of the oil was weighed into a conical flask and 5ml of chloroform, 50ml of 0.5M alcoholic KOH was added and mixed. A condenser was fitted to a conical flask and heated to about 70°C for one hour. The hot mixture was titrated against 0.5N HCl using phenolphthalein as indicator until the pink color became colorless. The Saponification value (SV) was calculated thus:

$$SV = \frac{(B-S) \times N \times 56.1 \times 100}{\text{weight of sample (g)}} \quad (2)$$

Where: B = Blank titre value, S = Sample titre value and N = Normality of HCl.

Determination of Moisture content (MC)

A covered crucible was dried for about one hour in an oven at 100°C. The weight of the dried crucible was taken and recorded as W₁. About 5g of the oil sample was now poured into the already weighed crucible and placed in an oven at a temperature of 105°C for about one hour. The new weight of the crucible and dried oil was taken and recorded as W₂. The moisture content was calculated thus:

$$\% \text{ Moisture} = \frac{\text{initial weight} - \text{final weight} \times 100}{\text{weight of sample}} \quad (3)$$

Where: W₁ = Initial weight of crucible and W₂ = final weight of crucible

Determination of Specific Gravity (SG)

A density bottle was weighed and the weight recorded as W_0 . The bottle was then filled with the oil sample and the new weight recorded as W_1 . This procedure was then repeated using distilled water after washing and drying the density bottle. The weight was recorded as W_2 . The Specific Gravity was calculated thus:

$$SG = \frac{W_1 - W_0}{W_2 - W_0} \quad (4)$$

Where: W_0 = weight of empty density bottle, W_1 = weight of density bottle + Oil sample and W_2 = weight of density bottle + distilled water.

Determination of Acid Value (AV)

Approximately 1g of oil sample was weighed into a conical flask and 5ml of chloroform added and mixed. 25ml of diethyl ether and ethanol (1:1v/v) was again added and mixed. A few drops of phenolphthalein indicator were added and the mixture titrated against 0.1M KOH solution until a faint pink color which persisted for about 30 seconds appeared. The acid value was calculated as follows:

$$AV = \frac{\text{Titre value} \times N \times 56.1}{\text{weight of sample (g)}} \quad (5)$$

Where: N = Normality of KOH

Determination of Iodine Value (IV)

Two grammes (2g) of the oil sample was weighed into a flask and 10ml of chloroform added and mixed. 25 ml of Wij's iodine solution was then added and the mixture was shaken vigorously and allowed to stand in the dark for about 30 minutes. 20ml of 10% KI solution was added and the mixture titrated against 0.1M $\text{Na}_2\text{S}_2\text{O}_3$ (Sodium thiosulphate) with a few drops of starch as indicator. A colorless solution confirmed the end-point. The iodine value was calculated thus:

$$IV = \frac{(B - S) \times N \times 126.9 \times 100}{\text{weight of sample (g)}} \quad (6)$$

Where: B = Blank, S = Sample and N = Normality of sodium thiosulphate

Determination of Peroxide Value (PV)

Approximately 1g of the oil sample was weighed into a conical flask. 12ml of chloroform and 18ml of glacial acetic acid was added and mixed. 0.5ml of aqueous KI was then added and the mixture was shaken and allow to stand for about one minute. 30ml of distilled water was then added and the solution titrated against 0.1M $\text{Na}_2\text{S}_2\text{O}_3$ until the yellow color almost faded. Then about 0.5ml starch solution was added, this turned the solution blue and the titration continued until the blue color disappeared. This procedure was also carried out using water as blank. The peroxide value was calculated as follows:

$$PV = \frac{F \times (A - B) \times 10}{\text{weight of sample (g)}} \quad (8)$$

Where: F = factor of 0.1N $\text{Na}_2\text{S}_2\text{O}_3$, A = Sample titre value and B = Blank titre value

Results and Discussion

Results

The results of the physicochemical analysis are as presented in Table 1 below.

Table 1: Results of physicochemical parameters of n-hexane extract of *Pachira aquatica* seed oil

Parameter	Obtained value	SON 2000 maximum allowable value
% Yield (%)	46.00±2.00	Nil
Iodine value (g I ₂ /100g)	13.5700 ± 0.13856	45-53
Acid value (mg KOH/g)	0.0233 ± 0.00050	0.7 max
Peroxide value (mEq/kg)	0.5700 ± 0.09504	10 max

Saponification Value(KOH/g)	4.5800 ± 0.30600	195-205
Specific Gravity (30°C)	0.9500 ± 0.00001	Nil
Moisture Content (%) (105°C)	0.2000 ± 0.00011	0.29

All values are presented as mean ± standard deviation, n =3.

For other physical properties of the oil, it is liquid at room temperature (25°C), yellowish in colour and non-offensive in odour.

Discussion

Peroxide Value (PV)

The peroxide value (PV) of fats and oils can serve as an indicator of their quality and stability as well as a gauge of the extent to which rancidity reactions have occurred during storage (Ekwu and Nwagu, 2004). It was also discovered that the peroxide value increased as the oil samples were stored, warmed, and exposed to air. The calculated PV value is in Table 1. The findings reveal that the peroxide values for *Pachira aquatica* seed oil are 0.5700 ± 0.09504 , suggesting a reasonably high quality of oil that will remain relatively stable in terms of rancidity. This value agrees with the PV of soybean oil (0.53 ± 0.02) as reported by Wazed et al (2023). However, the PV of palm oil and mustard oil as reported by the same authors were 2.21 ± 0.05 and 4.33 ± 0.02 , respectively. This reported PV is below the maximum value of 10 mEq/kg. How rancid the oil has become is determined by the peroxide value. Peroxide value ranges are far below the standard value of 10 mEq/kg specified by SON (2000), suggesting that the oil is of high quality.

Saponification Value (SV)

An indicator of the average molecular mass of fatty acids in an oil sample is called the saponification value (SV). The SV value of 4.5800 ± 0.30600 KOH/g was found for the oil samples listed in Table 1. The values fall short of the 195–205 mg KOH/g of edible oil standard range given by SON (2000). The SV obtained for soybean oil, palm oil, mustard oil and bran oil were 187.11, 202.39, 191.38 and 181.60, respectively. The SV for bran oil was found to be less than 185–195, as reported by Ramachandran (2001). Reduced saponification values indicate a decrease in either the number of ester bonds or the mean molecular weight of fatty acids. This could suggest that there was no interaction between the fat molecules (Denniston et al., 2004). It also indicates that the oil may not be suitable for soap making (Akpe and Inezi, 2018).

Iodine Value (IV)

The level of unsaturation in a fat or vegetable oil is determined by its iodine value (IV). It establishes the oils' resistance to oxidation and makes it possible to qualitatively assess the total unsaturation of the fat (AOAC, 1990; Gungshik et al., 2023). The measured iodine value for *Pachira aquatica* seed oil was found to be 13.5700 ± 0.13856 g/100 g. This IV is far below the value obtained by Wazed et al (2023) for soybean oil (133.17 mg/100 g). Its higher oxidative storage stability could have been attributed to these low iodine levels. An increase in free fatty acid contents and a decrease in the total unsaturation of oils are characteristics of the oxidative and chemical changes that occur in oils during storage (Afolayan et al., 2019).

Acid Value (AV)

Vegetable oil quality can be determined in part by looking at the Acid Value (AV). Amir et al. (2014) state that acid values are used to gauge how much lipase and other physical elements, like heat and light, have broken down the glycerides in the oil. The findings demonstrate that the acid value of *Pachira aquatica* seed oil, at 0.0233 ± 0.0005 mg KOH/g, is noticeably lower than that of other vegetable oils, indicating good lubricating properties. Wazed et al (2023) have reported an acid value of 3.05 mg KOH/g and 0.304 mg KOH/g for mustard oil and soybean oil, respectively. The acid value for refined rice bran oil, as reported by Ramachandran (2001), was found to be 0.5 mg KOH/g. High acid values imply low-quality oil and a high rancid propensity (Quader et al 2018). Low acidity stops oil from oxidizing, thereby reducing the risk of corrosion and the formation of gum and sludge. The findings show that the samples' acid values, which are an indicator of the amount of free fatty acids due to enzymatic activity, were found to be extremely low and below the minimum allowable value of 0.7 max.

Specific Gravity (SG)

A substance's specific gravity, which is measured in grams per unit of water, is its weight in relation to water. Measured at 30 °C, the specific gravity of the *Pachira aquatica* seed oil was 0.9500 ± 0.0001 . It is implied that this oil is less dense than water.

Moisture Content (MC)

Rancidity is likely to occur in edible oils when the moisture content falls between 0.05 and 0.3 (Yonnas *et al.*, 2019). According to the Australian Oil Seed Federation, the highest permitted moisture content for edible oils is 0.2%, which is in line with the results of the experiment. According to earlier research, oils made with less sophisticated technology had a higher moisture content (Abass *et al.*, 2010, Okechalu *et al.*, 2011). Consequently, the oil's moisture content suggested that rancidity would probably occur, albeit very slowly. All these properties make the oil suitable for consumption and other applications.

CONCLUSION

Scientific data and understanding of lesser-known or underutilized oils promote the use of their potential for industrial and nutritional purposes. The extracted oil's suitability for industrial and culinary uses was further validated by the results of its physico-chemical properties. Nonetheless, because of its high saturation content, *Pachira aquatica* seed oil has demonstrated stability against oxidation and is thus likely appropriate for use in deep-frying, food preparation, and industrial settings. The seed oil's lower peroxide value suggests that it will be more resistant to rancidity.

REFERENCES

- Afolayan S. Sunday¹, Igbum O. Gillian and Igoli O. John. (2019). Fatty Acid Composition of Seed Oil from *Pachira aquatica* Grown in Nigeria. *Journal of Agriculture and Ecology Research International* 18(4):1-9.
- Akpe MA and Inezi FP. (2018). Physicochemical properties of *Alchorea cordifolia*, *Cyperus esculentum* and *Irvingia gabonensis* seed oil and their applications. *Asian J. Applied Chem. Res.*, 1(3), 1-7.
- Amir, P.O., Babalola, O.O and Oyediran, A.M., (2014). Physicochemical Properties of Palm Kernel Oil, *Curr. Res. J. Bio. Sci.*, 6(5): 205-207, DOI:10.19026/crjbs.5194
- Azevedo, C. C. (2008). Modificação química das proteínas de amêndoas da munguba (*Pachira aquatica* Aubl): propriedades funcionais. Master Thesis. João Pessoa, Brazil: Universidade Federal da Paraíba.
- Cheng, L. Y., Liao, H. R., Chen, L. C., Wang, S. W., Kuo, Y. H., Chung, M. I., & Chen, J. J. (2017). Naphthofuranone derivatives and other constituents from *Pachira aquatica* with inhibitory activity on superoxide anion generation by neutrophils. *Fitoterapia*, 117, 16–21.
- Dansi, A., Voduhe, R., Yedomonhan, H., Assogba, P., Adjatin, A., Loko, Y. L., DossouAminon, I., & Akpagana, K., (2012). Diversity of the neglected and underutilized crop species of importance in Benin. *The Scientific World Journal*, 2012:932947
- Ekwu, F. and Nwagu, A. (2004). Effect of processing on the quality of cashew nut oils. *J Sci Agric Food Tech Environ.*, 4:105–10.
- Gungshik, J. R., Ibrahim, M., Onyemowo, O. A. and Rabiati, L. A. (2023). Comparative Analysis of the Physico-Chemical Properties and Trace Metal Content of Palm Oil, from Selected Markets in Jos South and Jos North LGA, Plateau State, Nigeria. *Nigerian Journal of Chemical Research*, 28(2): 100-111.
- Kyari, M. Z. (2008). Extraction and characterization of seed oils. *International Journal of Agrophysics*. 22, 139-142.
- Odoemelam, S. A. (2005). Proximate composition and selected physicochemical properties of the seeds of African oil bean (*Pentacle thramarcrophylla*). *Pakistan Journal of Nutrition*. 4, 382-383.
- Okechalu JN, Dashen MM, Lar PM, Okechalu B, Gushop T. (2011). Microbiological quality and chemical characteristics of palm oil sold within Jos Metropolis. Nigeria: Plateau State; *J. Microbiol. Biotech. Res.*, 1 (2): 107-112.
- Oliveira, J.T.A., I.M. Vasconcelos, L.C.N.M. Bezerra, S.B. Silveira, A.C.O. Monteiro, and R.A. Moreira. (2000). Composition and nutritional properties of seeds from *Pachira aquatica* Aubl, *Sterculia striata* St Hil et Naud and *Terminalia catappa* Linn. *Food Chem*. 70:185–191.
- Quader M.F.B., Ali, M.S., Ahmed, S., Zim, A.F.M.I.U., Roy, S. and Ahmed, S. (2018). Qualitative assessment of common edible oils available in Bangladesh. *Asian-Australasian Journal of Bioscience and Biotechnology*, 3(2), 156-161
- Ramachandran, H.D. (2001). Effect of n-3 fatty acids on oxidative stress in liver and brain of rats. India: University of Mysore, PhD thesis.
- Silva B. de L. de A., Bora P.S. and De Azevedo C.C. (2010). Caracterização química parcial das Proteínas das Amêndoas da Munguba (*Pachira aquatica* Aubl). Instituto Adolfo Lutz.
- Standard Organization of Nigeria (SON, 2000). Standards for edible palm oil and tallow Eds

- Wazed, M.A., Yasmin, S., BasakP., Hossain, A., Rahman, M.M., Hasan, M.R., Khhair, M.M. and Khatun, M.N. (2023). Evaluation of physicochemical parameters of edible oils at room temperature and after heating at high temperature. *Food Research*, 7(4), 91-100
- Yonnas A. N., Dagnachew E. A., Bikes D. B. and Henok D.. (2019).Assessment of quality of edible vegetable oils accessed in Gondar City, Northwest Ethiopia.*BMC Res Notes* (2019) 12:793

NCPS-2024

Hydrogeochemistry and Water Assessment of Shallow Groundwater in Parts of Benue State, North Central Nigeria

Iorfa Adamu¹, Innocent Akadi², Ebenezer Kudamnya¹, Msoo Asen³, Fina Adamu³

¹ Department of Geology, University of Calabar; ² ExxonMobil Nigeria Plc; ³ Department of Public Health, University of Calabar

Email of corresponding author: adamuchristopher@yahoo.com

Abstract

This study evaluated the chemistry and quality of the shallow groundwater resource in Benue State, North central Nigeria. A total of 24 groundwater samples were collected from three shallow aquifers (Basement, Fractured Shale and Sandstone aquifers) and analyzed for selected physicochemical and bacteriological parameters using standard techniques. The total hardness and CO_3^{2-} are higher for the shale aquifer; NO_3^- , Na^+ , and K^+ are more concentrated in the basement aquifer, while the TDS, EC, Ca^{2+} , Mg^{2+} , HCO_3^- , SO_4^{2-} and Cl^- are greater in the sandstone aquifer. Hydro-chemical parameters in the piper's trilinear diagram, as well as the Gibbs plot, respectively suggest that the ground water is of Ca-Mg- HCO_3 , Ca-Mg- SO_4 and Na- HCO_3 facies, and dissolved solids are mostly products of weathering. Application of R-mode factor analysis suggests that the main components of the water characteristics are related to hydrological and lithological controls aided by environmental and anthropogenic activities. On the whole, the water is slightly acidic and soft to very hard with low dissolved solids. It also satisfies the World Health Organization (WHO) standards for drinking, domestic and industrial uses (except for pH, HT, DO and Fe) requiring usual treatments such as filtration and disinfection.

Keywords: Hydrogeochemistry, groundwater, shallow aquifers, basement, weathering, anthropogenic

Introduction

Water is the basic need without which life does not exist in the earth. Springs, streams and ground water are some of the sources widely used for drinking, domestic and agricultural purposes. In recent times, it has become very difficult, seemingly impossible, to meet increased demands and provide sufficient quality water to the population due to pollution (Sridharan and Senthil, 2017; Ukah *et al.*, 2019). Most often, these water sources could be polluted by natural processes and human activities, thereby making their usability disputable (Akaamaa *et al.*, 2023).

In Benue State, only a small percentage of the rural community is provided with pipe-borne water. The greater majority relies on supplies from groundwater, rivers, streams and channels. Due to inadequate supply and deteriorating quality of surface water to meet the requirement of human activities in Benue state, groundwater remains the only option to supplement the ever-increasing demand for water. Few shallow boreholes have been drilled by the Benue State Water Board (BNSWB), United Nations International Children's Emergency Fund (UNICEF), Rural Water Supply and Sanitation Project (RUWASSAN) and private individuals to augment the aforementioned water sources. Majority of the people depend on poorly constructed hand-dug wells, tapping the shallow unconfined aquifers, for their water supply for domestic, agricultural and other purposes.

The previous hydrogeochemical and groundwater quality studies of the shallow aquifers in the study area were mostly disjointed and focused on anthropogenic impact (such as urbanisation, waste management, mining, and agriculture). Attention was placed on water quality in urban areas with little or no attention on rural areas and processes controlling ground water chemistry (Adamu and Nyiatagher, 2005; Ishaku *et al.*, 2010; Maitera and Ismaila 2011; Mile *et al.*, 2012; Ornguga, 2014). Hence the need to evaluate the hydrogeochemistry, water quality and health risk of the shallow groundwater across the state. The information from the study would be useful in mapping sources and extent of contamination, predicting future evolution of the ground water chemistry, finding remedial solution to any identified problem and in revealing the suitability of the water in the study area for drinking, domestic, industrial and irrigation purposes.

The main aim of this study is to assess the hydrogeochemical characteristics and water quality of shallow ground water in parts of Benue State, north central Nigeria. In order to achieve this, the data generated by this study will also augment existing data base, guide further investigations and provide a framework needed to guide in the design of a long-term monitoring programmes to properly manage and protect the shallow aquifers within Benue State, North Central Nigeria

The study area

The study area lies within the lower Benue trough in the middle belt region of Nigeria and it is bounded by latitudes 6°25' and 8°08'N and longitudes 7°47' and 10°00'E. The land is generally low lying and gently undulating with occasional inselbergs, knolls, lateritic capped mesas and butts. It is situated within the subequatorial climatic region of Nigeria and is characterized by two main seasons; the wet and the dry seasons. The rainy period is confined to between April and October with a total monthly rainfall ranging from 150 to ≥ 200 cm; while the latter, characterized by relatively hot and dusty conditions, covers the duration of November to March. The mean annual minimum and maximum temperatures correspondingly vary from 20.6 to 22.6°C and 28.2 to 31.2°C, while the mean monthly humidity is between 65 to 85% (Iloeje, 2001).

The different lithologic units in the study area have varying hydrogeological characteristics. The Basement rocks form local aquifers where the degree of weathering and/or fracturing is sufficient. Crystalline and coarse-grain rocks, such as gneisses and migmatites, become sandy on weathering, thus preferentially forming aquifers. Argillaceous meta-sedimentary rocks tend to become clayey with low permeability when weathered, forming aquitards (Adamu *et al.*, 2014). Overall, basement aquifers tend not to be high yielding. They typically vary in thickness from 10 to 25m, with water table depths varying from about 5 to 15m. Boreholes tend to be drilled to depths of between 10 and 70 m, depending on local conditions. Yields of 1.8 to 2.5 l/s have been reported from boreholes drilled into the basement complex rocks of varying degree of decomposition (Offodile, 2002).

The ARG is not known to have good aquifers as it consists mainly of tightly folded shales and siltstones. Although in highly fractured areas small quantities of water can be obtained from it. The major feature of the aquifer is the occurrence of a thin shallow but extensive unconfined aquifer. The aquifer is formed by the top weathered horizon within the fractured shale and sandy horizons. The fractured shales of the ARG are characterized by high transmissivity which is related to depth of burial and degree of metamorphism of host rock (Offodile, 2002).

Materials and Methods

Sample collection

Global positioning system (GPS) was used in taking co-ordinates and elevations of each sample locations. At locations with hand-dug wells (HDW), water samples were collected and stored in well labeled clean sterilized polythene bottles. Three sets of samples were collected from each of the 24 HDWs. The first set was used in measuring the major anions; the second set was acidified to pH ~ 2 with ultra-pure HNO₃ and used in measuring major cations and heavy metals; and the third set collected in 250ml bottles was used in measuring the biological parameters. Altogether, 80 samples were collected comprising 72 representative groundwater samples and 8 duplicate samples. All the samples were filtered *in-situ* using Whatman 0.45 µm cellulose acetate membrane filters. The collected samples were stored at 4°C in a cooler and transported to the laboratory where they were preserved in the refrigerator prior to instrumental analysis. Parameters such as electrical conductivity (EC), pH, total dissolved solids (TDS), temperature and dissolved oxygen (DO) were determined in the field.

Laboratory analysis

In the laboratory, Total alkalinity (TA) was measured using sulphuric acid with a digital titrator. The total hardness (HT) was determined by titration with a standard solution of EDTA which is a complexing agent. A Hach spectrophotometer (model DR 3000) was used for the determination of colour. The analyses for anions, cations and trace metals were carried out using spectrophotometric (Wagtech 7100 Model) technique at the Benue State Rural Water Supply and Sanitation Agency (BERWASSA) in Makurdi (Table 2).

Microbiological qualities of the samples were determined using membrane filtration technique. The working environment was cleansed with methylated spirit and much talking was avoided while conducting the test to prevent saliva from falling on the samples, as saliva may increase contaminants (bacteria) in the water samples. The petri-dishes were sterilized by flaming with methylated spirit while 100ml of each of the samples was filtered using the filtration chamber. The filter paper carrying the residue was picked and gently dropped on sterilized dish and saturated with broth (feeds for bacteria) and covered with the counting chamber. The broth increased bacteria growth thereby enhancing their prominence for ease of counting. The petri-dish was then covered and placed in the incubator for 8hrs at 44°C. After which the dish was removed and allowed to cool before counting with hand lens.

Univariate and multivariate analysis

Excel (version, 2020) was used to compute descriptive statistics and to produce charts showing the distribution of ions and heavy metals in the groundwater. AquaChem software (version 2014) was used to obtain the Piper Plot in understanding the hydrogeochemical facies of the groundwater samples, while Gibbs' Plot was used to determine the sources of the ion present therein. Multivariate statistics such as factor analysis (FA) were performed with the use of the statistical software package; Statistical Package for the Social Sciences (SPSS, version 22). The ion-exchange between the ground water and its host environment during residence or travel can be understood by studying the chloroalkaline indices. The chloroalkaline indices are calculated by using the equations:

$$\text{Chloroalkaline index I (CAI-I)} = [\text{Cl}^- - (\text{Na}^+ + \text{K}^+)]/\text{Cl}^- \quad (1)$$

$$\text{Chloro alkaline index II (CAI-II)} = [\text{Cl}^- - (\text{Na}^+ + \text{K}^+)]/[\text{SO}_4^{2-} + \text{HCO}_3^- + \text{CO}_3^{2-} + \text{NO}_3^-] \quad (2)$$

When CAI is negative then there is base-exchange between sodium and potassium (Na + K) in water with calcium and magnesium (Ca + Mg) in rocks/soils, if the ratio is positive, then it is reverse ion-exchange.

RESULTS AND DISCUSSION

Physicochemical parameters

A statistical summary of the data obtained from analysis of the sampled groundwater during the study is hereby presented (Table 1). The total dissolved solids (TDS) range from 58.30 to 900mg/l with an average value of 294.27mg/l. The average TDS values for groups B, W, and S are 168.09mg/l, 270.60mg/l and 441.13mg/l respectively indicating that the water is generally of low to moderate mineralization. This is borne out of the fact that the TDS concentration falls within the range for fresh water class (<1000mg/l) of Langenegger (1990). The ranges of dissolved oxygen (DO) and water temperature are 1.4 to 4.3mg/l and 27.4 to 31.3°C respectively. The low DO values recorded in the study area may be attributed to the high temperature of the ground water. This is because with higher temperatures, natural waters release their gaseous content leading to lower concentration (Jia *et al.*, 2018). The pH values range from 5.7 to 8.5 with an average of 6.88 (Table 6) indicating slightly acidic to slightly alkaline waters. The pH values are not restricted to a particular rock type; however, the basement aquifer appears to be more acidic (mean pH, 6.43) and shale aquifer more alkaline (mean pH, 7.24). This condition may be attributable to both natural and human activities. For instance, the interaction of the different aquifer materials with the ground water and/or biotic activities that result in the production of CO₂ in the water column can all cause variation in ground water pH (Hem, 1985). EC values ranged from 92.0 to 1,102µS/cm with an average of 248.69 µS/cm. The colour values range from 4 to 530Pt/Co with a mean value of 182.29Pt/Co. It is observed that the colour values are higher in the basement aquifer and lower in the sandstone aquifer. This may be associated with the concentration of dissolved iron and manganese in the waters (Hem, 1985). The colour values in the ground waters follow the same trend as the concentration of iron. The values of total hardness and alkalinity range from 14 to 562mg/l and 1.00 to 46.00 with average values of 221.08mg/l and 15.95 respectively. Higher values of hardness in the shale aquifer (average 249.13mg/l) are expected because the shale is intercalated with limestone. The water in the basement aquifer falls into the soft to very hard water groupings. While the water from the shale and sandstone aquifers fit into the hard and very hard groupings (Hem 1985).

Table 1: Descriptive statistics overview hydrochemical data set and WHO (2011) Standards for Drinking water

Aquifer	Stati	Temp	EC	TDS	pH	TA	Color	HT	DO	HCO ₃ ⁻	SO ₄ ²⁻	Cl ⁻	NO ₂ ⁻	NO ₃ ⁻	CO ₃ ²⁻	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	T/Coli
Basement	Min	27.4	92	58.3	5.7	2.6	98	14	1.4	34	2	12	0.01	4.4	3.3	1.76	2.24	12	2	12
	Max	30.00	568	281	7.1	15	550	562	3.1	64	24	24	0.08	36.1	12	78.6	96	52	16	84
	Mean	28.39	310.8	168	6.4	8.6	335	202	2.2	44	11	16	0.05	12.7	7.9	31.2	39.16	31	8	51.75
Shale	Min	29.00	117	58.8	6.2	1	81	137	3.1	18	10	3.6	0.01	2.6	4.2	16	21	8	2.4	20
	Max	31.3	830	473	8.5	4	392	352	4.3	55	112	7.2	0.06	16	78	38	57	32	6.4	174
	Mean	29.96	551	271	7.2	2.3	202	249	3.6	36	65	5.5	0.03	7.47	23	26.7	37.76	20	4.2	88.25
Sandstone	Min	29.8	420	210	6.4	28	4	162	2	22	15	11	BDL	2	2	10	30	8	1.8	10
	Max	31	1102	900	8	46	16	255	4	354	198	35	0.28	16	56	41	47	30	8.2	230
	Mean	30.29	658.6	444	7	37	9.5	212	2.6	78	72	24	0.08	6.76	10	27.9	37.25	19	4.7	101
Overall	Min	27.4	92	58.3	5.7	1	44	14	1.4	18	2	3.6	BDL	2	2	1.76	2.24	8	1.8	10
	Max	31.3	1102	900	8.5	46	550	562	4.3	354	198	35	0.28	36.1	78	78.6	96	52	16	230
	Mean	29.55	506.8	294	6.9	16	182	221	2.8	53	49	15	0.05	8.97	14	28.6	38.06	23	5.7	76.52
MPL		-	1400	1000	6.5-8.5	200	-	150	>5.0	250	250	250	-	50	-	200	100	250	12	<1.00

Maximum permissible limit (WHO, 2008), Units mg/l except pH (standard units), EC (µS/cm), Colour (Pt/CO), Temp. (°C), Charge Balance Error (CBE %); HT= Total hardness (mg/l as CaCO₃), TA = Total Alkalinity, BDL = Below Detection Limit, NA= Not available

The major dissolve components of the ground waters include Na^+ , K^+ , Ca^{2+} , Mg^{2+} , Cl^- , HCO_3^- , CO_3^{2-} and SO_4^{2-} . The dominant cation is Ca^{2+} with a mean value of 22.18mg/l, followed by Mg^{2+} with a mean value of 21.21mg/l then Na^+ and K^+ with mean concentrations of 15.50mg/l and 5.06mg/l respectively. In 12 (50%) of the ground water samples Ca^{2+} is the dominant cation, while Mg^{2+} and Na^+ dominate in 10 (41.67%) and 2 (8.33%) of the samples respectively. The dominant anion in the ground water samples is HCO_3^- (mean, 66.50mg/l) followed by SO_4^{2-} (mean 60.00mg/l), Cl^- (mean, 16.72mg/l), NO_3^- (mean, 12.45mg/l) and CO_3^{2-} (mean 10.20mg/l). In 16 (66.67%) of the ground water samples HCO_3^- is the dominant anion while SO_4^{2-} and NO_3^- dominate in 7(29.17%) and 1(4.17%) of the samples respectively.

Hydrochemical facies and sources of ions

Piper Plot (Fig. 1) indicate that the sampled groundwater can be classified into three water types after Elueze *et al.* (2001) namely: CaMgHCO_3 (basement and sandstone aquifers), CaMgSO_4 (shale aquifer), and minor NaHCO_3 hydrochemical facies according to their order of dominance. The CaMgHCO_3 and NaHCO_3 water types express mineral during recharge of freshwater. The SO_4^{2-} is most probably from sulphide oxidation, atmospheric precipitation or dissolution of sulphate minerals in the study area (Batayneh *et al.*, 2014). According to Jia *et al.* (2018), the calcium-magnesium bicarbonate water type is an indication of reverse ion exchange and recharge processes, and such water is referred to as ‘exchange water’. This implies that the bicarbonate (HCO_3^-) anion is in deficiency of the total alkaline earths ($\text{Ca}^{2+} + \text{Mg}^{2+}$) concentration. The former is therefore not available to act as scavenger in the release of cations, notably alkalis ($\text{Na}^+ + \text{K}^+$), through exchange reaction with clay minerals. Such activity may account for the nature of the water of the study area. Hem (1985) reported that meteoric water dissolving Ca, Mg and Na/K from silicates would produce CaHCO_3 and CaMgHCO_3 water types, which might evolve to CaNaHCO_3 water type due to cation exchange.

Higher conductivity and TDS values were observed in the sandstone and shale aquifers in that order. This may be attributed to the ease of dissolution of aquifer materials in these aquifers relative to the basement aquifer. Based on the plots of TDS against the $\text{Na}/(\text{Na} + \text{Ca})$ ratio (Fig. 2), the water showed that it acquired its chemistry from weathering of the host rock or materials from the associated aquifer materials.

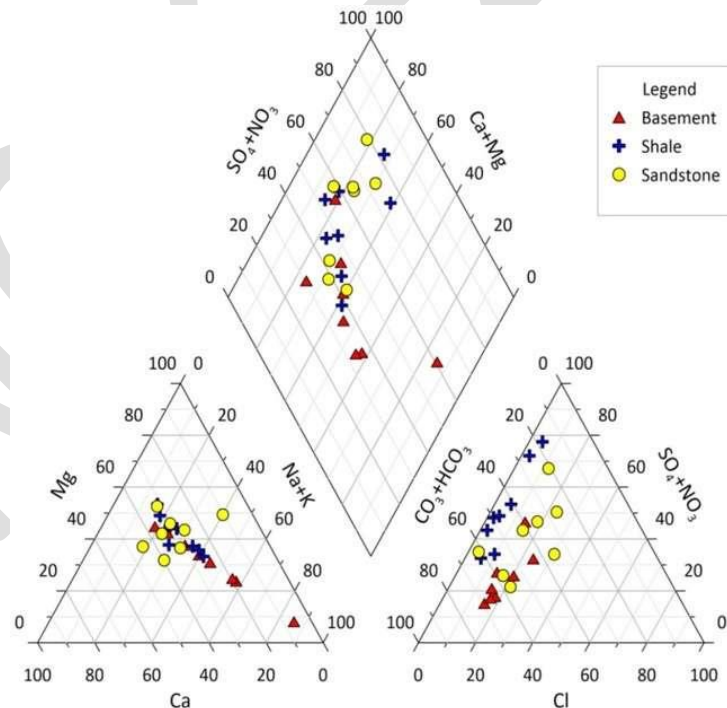


Fig. 1: Trilinear Piper diagram of ground water composition in the study area

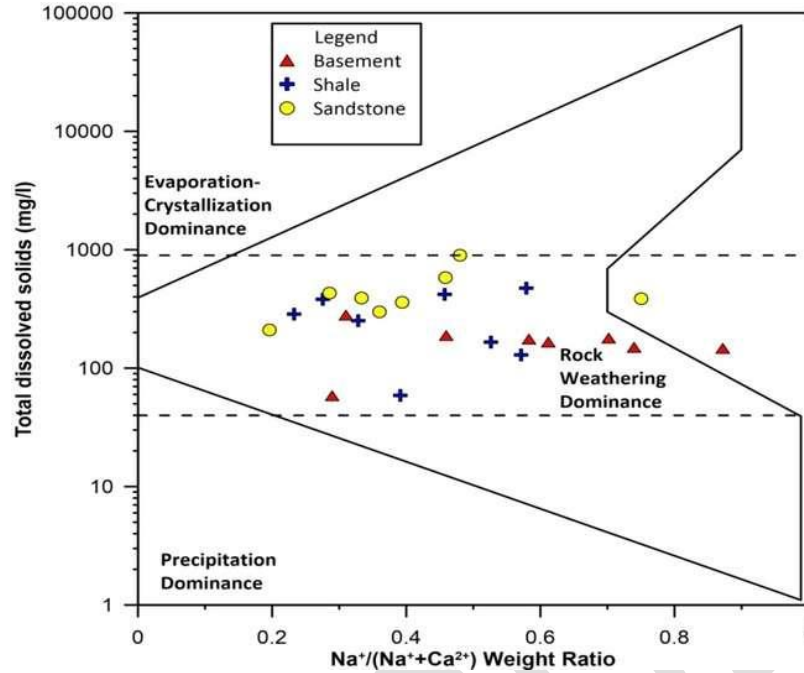


Fig.2: Gibb's plot of variation of TDSvs $\text{Na}^+ / (\text{Na}^+ + \text{Ca}^{2+})$, after Gibbs (1970).

Factor analysis

In this study, a three-factor model that accounted for 87.3% of the data variance was considered to be the most consistent with known hydrological, geological, environmental and anthropogenic processes and appropriate to the limits of analytical precision. Only a variable with a loading ≥ 0.60 was considered significant members of a particular factor. These different factors are summarized in a matrix (Table 2) and discussed below:

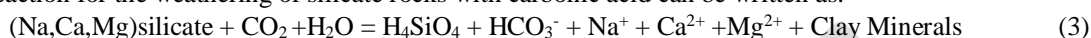
Factor 1 (EC , Mg^{2+} , TDS , HCO_3^- , Ca^{2+} , Na^+) accounts for 47.3% of total data variance. The high loading of Mg^{2+} , Na^+ , Ca^{2+} , and HCO_3^- reflect the influence of rock weathering and/or geochemical interactions of ground water with aquifer material. Factor 1 may be treated as a bicarbonate factor and explains carbonate dissolution and silicate weathering as the dominant chemical processes in the study area. In ground water of the study area, Mg^{2+} , Ca^{2+} and Na^+ constitute more than 90% of the total cations and HCO_3^- is the dominant anion. Carbonate and silicate weathering and evaporate dissolution can supply these cations in water. Since Cl^- has low loading on factor 1, evaporate dissolution is not considered a dominant source of ions. The $(\text{Mg}^{2+}, \text{Ca}^{2+})/\text{HCO}_3^-$ ratio marks the upper limit of bicarbonate input from weathering of carbonate rocks (Table 4). The high values of $(\text{Mg}^{2+} + \text{Ca}^{2+})/\text{HCO}_3^-$ (>1.00 in 87.5% of samples) indicate that $\text{Ca}^{2+} + \text{Mg}^{2+}$ is in excess of alkalinity, the magnitude of excess being $W > S > B$. This excess positive charge has to be balanced by other anions such as SO_4^{2-} , Cl^- or NO_3^- . The primary silicates

Table 2: R-Mode varimax factor matrix

Variable	Factor 1	Factor 2	Factor 3
pH	0.40977	-0.732112	-0.088429
EC	-0.18614	-0.945767	-0.103003
TDS	-0.23926	-0.81278	-0.09434
TH	0.83137	-0.462558	-0.005364
Turbidity	0.98925	0.050489	0.044021
Na	0.88675	-0.058262	-0.179765
K	0.38734	-0.30252	0.019735
Ca	-0.26126	-0.832955	0.200362
Mg	0.96539	-0.110731	-0.051555
SO_4	0.9712	0.129765	0.005884
HCO_3	0.86409	-0.383777	-0.143256
CO_3	0.85077	-0.361762	-0.164324

Cl	0.97747	0.054821	0.092481
NO ₃	0.9103	0.245933	0.081387
PO ₄	0.74983	-0.099838	-0.390657
Eigen value	13.1961	4.68533	1.384
% Total variance	54.98383	19.52221	5.76666
Cumm. EigenValue	13.1961	17.88143	19.26543
% Cumm. Variance	54.98373	74.50594	80.27261

associated with the crystalline rocks of the area include Ca-plagioclase, orthoclase, amphiboles and biotite and the general reaction for the weathering of silicate rocks with carbonic acid can be written as:



Factor 2 (Temp., K⁺ and SO₄²⁻) accounts for 24.6% of data variance. The level of K⁺ is much less than Na⁺ in the analysed samples. The excess of Na⁺ over K⁺ is expected because during continental weathering, Na⁺ turns out to be more mobile than K⁺ and dominates in natural waters. The loading of SO₄²⁻ in the group may suggest environmental and hydrological controls by the oxidation of sulphide minerals such as pyrite. It is because there are reported occurrence of pyrites and other sulphide minerals within the Benue trough (Petters, 1982). The SO₄²⁻ in the ground waters could also be sourced from associated gypsum minerals especially in the shale area. Temperature is very important factor in chemical weathering processes and its inclusion in this factor grouping supports the interpretation.

Factor 3: (NO₃⁻, pH, DO) accounts for 15.4% of the data variance. The high loading of NO₃⁻ and pH may suggest anthropogenic inputs into the ground waters from fertilizer application, animal waste and domestic sewage. The NO₃⁻ has no apparent lithological source in the study area and it may be associated with surface run-off of nitrate fertilizer from farmlands or atmospheric inputs. The NO₃⁻ is responsible for the relatively low pH in some samples in the presence of DO. Indications are that most if not all of the high NO₃⁻ concentrations are derived from anthropogenic and atmosphere precipitation.

Conclusions

The study evaluated the hydrochemical characteristics of surface water and groundwater in parts of the Middle Benue Trough, to know their suitability for various uses. Therefore, the concentration levels of some selected physico-chemical parameters (Total hardness, temperature, EC, pH, turbidity, Ca²⁺, Mg²⁺, Na⁺, K⁺, Cl⁻, SO₄²⁻, HCO₃⁻, CO₃²⁻, NO₃⁻, and PO₄³⁻) and E-coli counts of the water resources in the area were determined. Hydrochemical and statistical reduction methods have been used to investigate the ground water resource of Benue State, North Central Nigeria. Results show that the water is slightly acidic, of low mineralization and dissolved solids and generally dominated by CaMgHCO₃ and CaMgSO₄ types with a little mixture NaHCO₃ water type. Gibbs plots show that the dissolved solids are mainly products of weathering. Application of R-mode factor analysis indicate a 3 factor model showing that the main components of the water characteristics are related to hydrological, lithological, anthropogenic and environmental controls. Generally, the hydrochemical parameters satisfy WHO standard for domestic, agriculture and other industrial uses (except for pH, TH) requiring usual treatments such as filtration and disinfection. It recommended that this investigations would be most meaningful if an isotopic composition of the ground water is incorporated during future studies.

REFERENCES

- Adamu, C.I., Edet, A. and Njanje, T. (2014), Hydrochemical assessment of pond and stream water near abandoned barite mine sites in parts of Oban massif and Mamfe Embayment, southeastern Nigeria, *Envi. Earth Sci.* <https://doi.org/10.1007/s12665-013-2757-5>
- Adamu, C.I. and Nyiatagher, T. (2005), Physicochemical and Bacteriological Quality of Surface water of River Benue, Nigeria, *Global Journal of Geological Sciences*, 3, 1, 45-48.
- Akaamaa W. W., Onoja, S. B., Nwakonobi T .U. and Udochukwu, M. (2023). Hydro-geochemical and quality assessment of groundwater from sedimentary formation in the Middle Benue Trough, Nigeria. *International Journal of Agriculture, Environment and Bioresearch* 08(01):213-229. <https://doi.org/10.35410/ijaeb.2023.5811>
- Batayneh, A. (2014). Toxic (Aluminum, Beryllium, Boron, Chromium and Zinc) in groundwater: Health risk assessment. *Int. J. Environ. Sci. Technol*, 9, 153–162.

- Elueze, A. A., Ephraim, B. E, and Nton, M. E. (2001). Hydrochemical assessment of surface water in part of southeastern Nigeria. *Hydrology*, 119: 45-58
- Gibbs R.J (1970). Mechanisms controlling world water chemistry. *Science*, 170, 1088-1090
- Hem, J. B. (1985), Study and Interpretation of the Chemical Characteristic of Natural Water, pp.2254, *United State Geological Survey, (USGS), Water Supply Paper*.245-263
- Iloeje, N. P. (2001). *A New Geography of Nigeria*. 4th ed. Longman Ltd, Benin City
- Ishaku, T. H., Hussain, A.M., Dama, M. F., Zemba, A.A. and Peters, A.A. (2010), Planning for sustainable water supply through partnership approach in Wukari, Taraba State, Nigeria. *Journal of Water Resource and Protection*, 2, 916-922.
- Jia, Y., Xi, B., Jiang, Y., Guo, H., Yang, Y., Lian, X. and Han, S. (2018). Distribution, formation and human-induced evolution of geogenic contaminated groundwater in China: A review. *Sci. Total Environ*, 643: 967–993
- Langenegger O. (1990). Groundwater quality in rural areas of western Africa. UNDP Project INT/81/026, 10 pp.
- Maitera, O. N. and Ismiala, Y. S. (2011). An assessment of total coliform levels of some portions of River Gongola Adamawa State, Nigeria, *Pelagia Research library* 1, 3,191-107.
- Mile, I. I., Jande, J. A. and Dada, B .L. (2012), Bacteriological Contamination of Well in Markurdi, Benue State, Nigeria, *Parkistan Journals of Biological Science*. 15:1048-1057.
- Offodile, M. E. (2002). *Groundwater supply and development in Nigeria*. Mecon Geology and Eng. Serv. Ltd., Jos, Nigeria. 453pp
- Ornguga, T.T. (2014). Groundwater quality assessment and sanitary surveillance of boreholes in rural areas in Benue State, Nigeria, *Academic Journals of Interdisciplinary studies* 3,5,2281-4612.
- Petters S. W. & Ekweozor, C. O. (1982). Petroleum Geology of Benue trough and Chad Basin, Nigeria. *American Association of Petroleum Geology Bulletin*, 66, 1141-1149.
- Sridharan M. and Senthil N. D. (2017). Groundwater quality assessment for domestic and agriculture purposes in Puducherry region Appl Water Sci <https://doi.org/10.1007/s13201-017-0556-7>: 4037-4053
- Ukah, B.U., Egbueri, J. C., Unigwe, C.O. and Ubido, O. E. (2019). Extent of heavy metals pollution and health risk assessment of groundwater in a densely populated industrial area, Lagos, Nigeria. *Int J Energ Water Res.*, 3, 2910303. <https://doi.org/10.1007/s42108-019-00039-3>

Geospatial assessment of landslide vulnerability mapping using remotely sensed data and fuzzy-AHP in urban and rural regions of Orumba North, southeastern Nigeria.

¹G. I. Joshua, ¹E. A. Kudamnya, ¹A. S. Ekwere, ¹B. O. Onwanyi, ²A. E. Edet,

1. Department of Geology, University of Calabar, Calabar. 2. Department of Earth science, University of Uyo, Uyo.

Corresponding email: ebenezerkudamnya@unical.edu.ng

Abstract

This study attempts to assess and compare the vulnerability to landslides in Orumba North by integrating remote sensing, geographic information system (GIS), and the fuzzy analytical hierarchy process (Fuzzy AHP). The conditioning factors that predetermine landslide movements include: slope, rainfall, elevation, curvature, land use and land cover (LULC), normalized difference vegetation index (NDVI), and proximity to streams and roads, which were used for the assessment. Through fuzzy-AHP, these factors were assigned weighted importance based on expert judgement and environmental conditions, the defuzzification process converted the fuzzy weights into crisp value for precise evaluation these crisp value was used to rank the factors based on their relative importance. The defuzzified weights, was used to produce a landslide vulnerable map by classifying the study area into four susceptibility levels; low, moderate, high, very high. From the result, it was observed that urban areas (78%) such as Oko, Nanka, and Ajalli exhibit higher vulnerability to landslide due to intense human activities which includes construction, deforestation, and poor drainage management. The rural areas(22%) like Ufuma and Omogho are largely affected by natural factors such as steep slope, agricultural practices, leading to significant but less human induced susceptibility. Geospatial assessment of the susceptibility map produced offers valuable knowledge for local planner and decision makers to identify high risk area, particularly those in urban centers with higher developmental pressure. Targeted efforted can be made to improve land use planning, drainage system and slope stabilization.

Keywords: Geospatial, Remote Sensing, Fuzzy- AHP, Hazard, Urban and Rural.

Introduction

Landslide is one of the destructive geological processes which pose significant threats to lives, infrastructure, and the environment, particularly in regions with steep slopes and high rainfall. In Nigeria, southeastern states are particularly vulnerable due to their unique geomorphological, climatic, and human activities (Okoro et al., 2023) it's a prominent type of mass wasting and prevalent environmental hazards. Orumba North Local Government Area, located in southeastern Nigeria, has witnessed recurrent landslides, leading to substantial economic losses and environmental degradation. Integrating remotely sensed data ensures accurate and timely spatial assessments, while the Fuzzy-AHP technique provides a robust decision-making framework by incorporating expert judgments with uncertainty management (Adebayo et al., 2022). These geo-hazard often result from the interplay of various natural and anthropogenic factors, including geological conditions, slope stability, rainfall intensity, land use changes and seismic activities. The combination of these geological features and anthropogenic activities, such as deforestation and improper land use, contributes to frequent landslides in the area.

The complexity and variability of these factors make landslide susceptibility assessment crucial for effective risk mitigation and land use planning. Geographic information systems (GIS) and remote sensing technologies have emerged as indispensable tools in this regard, enabling researchers to analyze spatial datasets, identify potential hazard zones and develop predictive models. GIS serves as a platform for integrating diverse datasets such as topography, geology, hydrology, land use to produce landslide susceptibility maps. These maps are often derived through heuristic, statistical methods that analyze the relationships between landslides and their causative factors (Lee and Pradhan, 2007). Remote sensing complements GIS by providing up to date, high resolution imagery and terrain data, which are essential for detecting land cover changes, identifying geomorphic features and deriving critical parameter such as slope, aspect, curvature, elevation, lineament (Chen et al., 2022). The integration of geographic information systems (GIS), remote sensing technologies, and advanced multi-criteria decision-making frameworks fuzzy-AHP offers a robust approach to identifying and mapping landslide-prone areas. This study adopts geospatial technologies and Fuzzy-AHP to develop a landslide vulnerability map for the urban and rural regions of Orumba North.

Description of the Study Area

Orumba North is situated in Anambra State, southeastern Nigeria, and lies between latitudes 6°38'N and 6°53'N and longitudes 7°12'E and 7°28'E. The region covers an approximate area of 580 square kilometers and exhibits a mix of urban and rural settlements. The area is characterized by undulating terrain, with elevations ranging from 100 to 500 meters above sea level.

Geologically, Orumba North is part of the Anambra Basin, a sedimentary basin in southeastern Nigeria. The basin, formed during the Late Cretaceous, is dominated by sequences of shale, sandstone, and siltstone. The geological formations and their ages, along with their influence on landslide susceptibility, are summarized in Table 1.

Table 1: Stratigraphic formations, age and their influence on landslide

Stratigraphic Formation	Geological Age	Influence on Landslide Susceptibility
Nanka Sands	Paleocene-Eocene	Unconsolidated sands and clay layers reduce slope stability, increasing landslide risks during heavy rainfall (Ezenwa et al., 2021).
Imo Shale	Paleocene	Fine-grained shales are prone to erosion and water retention, weakening slopes (Nwankwo et al., 2023).
Ogwashi-Asaba Formation	Oligocene-Miocene	Alternating sands and clays can lead to differential erosion, affecting slope stability (Adeoye & Balogun, 2024).

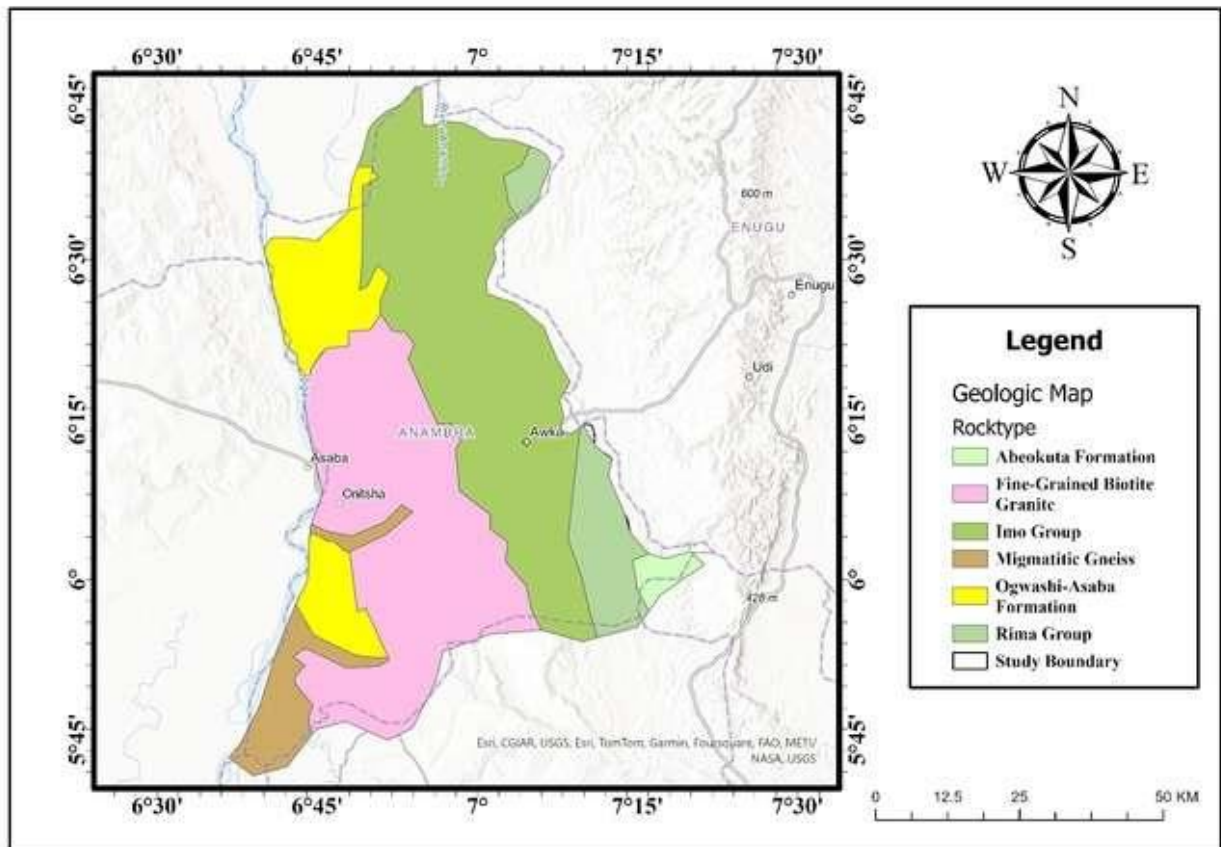


Fig.1. Map of the study location

Methodology

Thematic maps were produced in for elevation, slope, rainfall, aspect, relief, curvature, and proximity to roads, proximity to streams, land-use/land-cover, and NDVI (Normalized Difference Vegetation Index). Also, the work-flow presented (Fig. 2) was used to perform GIS processes that produced the final landslide vulnerability map in the study area.

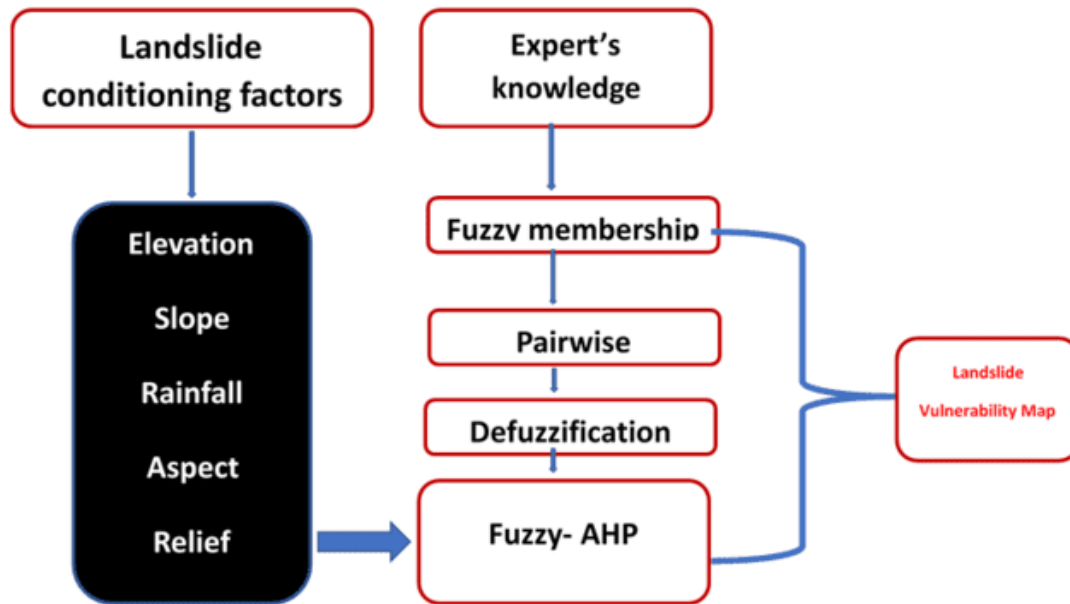


Fig. 2: Work-flow of vulnerability mapping of Orumba North

Discussion

This study comprehensively assessed landslide vulnerability in Orumba North, by adopting GIS techniques using multiple influencing factors. These factors include slope, aspect, relief, curvature, elevation, and proximity to roads, proximity to streams, land use/land cover, and NDVI (Normalized Difference Vegetation Index). Each factor was analyzed to determine its role in landslide vulnerability and its spatial distribution across the study area.

The elevation and slope analyses revealed that higher elevations and steep slopes, primarily in the northern and eastern parts of Orumba North, are more susceptible to landslides. This finding is consistent with the literature, which identifies steep slopes as critical contributors to slope instability. Proximity to streams demonstrated that areas closer to water bodies, particularly in the southern and central regions, face heightened risks due to potential undercutting and erosion. Similarly, areas near roads were found to be vulnerable, primarily due to anthropogenic activities such as road construction and deforestation, which destabilize the soil. Land use and NDVI analyses further emphasized the role of vegetation in mitigating landslides. Non-vegetated and sparsely vegetated areas, especially in the central and southern parts, exhibited higher susceptibility due to reduced root reinforcement. Conversely, moderately vegetated zones were found to be less prone to landslides. The integration of these variables through weighted overlay techniques facilitated the development of a detailed landslide vulnerability map, which categorized the region into very high, high, moderate, and low-risk zones.

The final landslide vulnerability map (Fig. 3) highlights specific areas of concern. The very high-risk zones, covering approximately 28% of the study area, are concentrated in regions such as Oko, Nanka, and Ekwulobia. These areas are characterized by steep slopes, proximity to streams, and sparse vegetation. High-risk zones, accounting for 32% of the area, are located in the central parts, including Ufuma and Ndikelionwu, where moderate slopes and sparse vegetation are prevalent. Moderate-risk zones (25% of the area) and low-risk zones (15% of the area) are distributed across the southern and northern parts, where terrain and vegetation conditions are relatively stable.

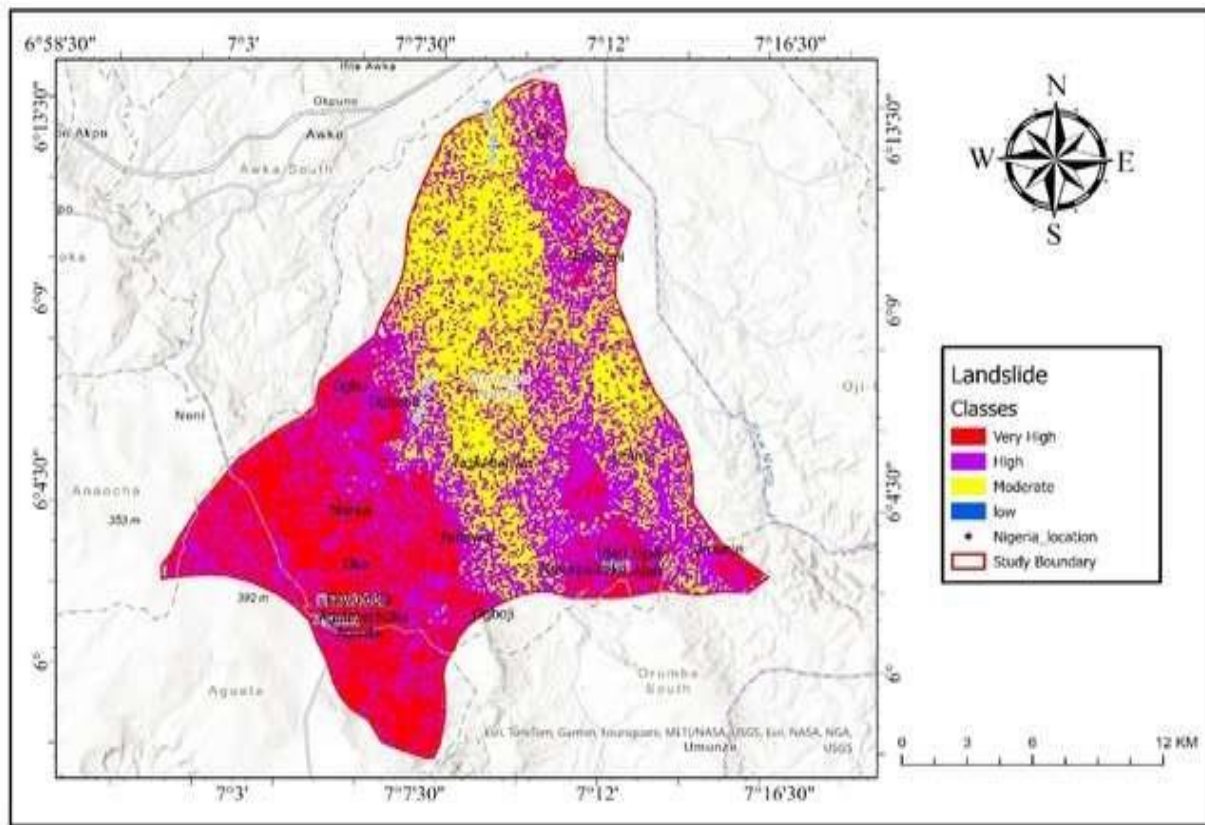


Fig. 3: Landslide vulnerability map of Orumba North

Conclusion

The study demonstrates the effectiveness of integrating geospatial techniques and multi-criteria analysis to assess landslide vulnerability. The findings provide a valuable framework for land-use planning, disaster management, and environmental conservation in Orumba North. By identifying high-risk zones and their associated factors, this study offers actionable insights for mitigating landslide risks, emphasizing the importance of preserving vegetation, regulating anthropogenic activities, and improving drainage systems.

References

- Adeoye, A., & Balogun, T. (2024). Application of Fuzzy-AHP for geohazard mapping: A case study of southeastern Nigeria. *Journal of Environmental Management*, 345, 124532. <https://doi.org/10.1016/j.jenvman.2024.124532>
- Adebayo, S. O., Adeola, T. O., & Akinyele, F. A. (2022). Application of Fuzzy-AHP in geospatial disaster riskassessment. *International Journal of Remote Sensing and GIS Applications*, 8(4), 112-127. <https://doi.org/10.1080/ijrs.2022.0804112>
- Chen, W., Pourghasemi, H. R., & Zhang, X. (2022). Landslide susceptibility modeling using GIS and machine learning. *Journal of Environmental Management*, 301(Part A), 113769. <https://doi.org/10.1016/j.jenvman.2021.113769>
- Ezenwa, U. P., Nwosu, M. O., & Agbo, C. J. (2021). Geology and landslide occurrences in the Nanka Sands Formation, southeastern Nigeria. *African Geoscience Review*, 29(2), 89-103.

- Nwankwo, J. I., Ume, A. U., & Adeoye, K. O. (2023). Landslide susceptibility mapping using GIS and multi-criteria decision analysis: Case study of southeastern Nigeria. *Applied Geography*, **152**, 102769. <https://doi.org/10.1016/j.apgeog.2023.102769>
- Okoro, E. E., Adeoye, K. O., & Balogun, T. (2023). Risk assessment of landslides in urban areas of southeastern Nigeria. *Natural Hazards and Earth System Sciences*, **23**, 3345–3360. <https://doi.org/10.5194/nhess-23-3345-2023>
- Lee, S., and Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia, using frequency ratio and logistic regression models. *Landslides*, 4(1), 33-41. <https://doi.org/10.1007/s10346-006-0067-y>

Measurement of seasonal variation of radon concentration levels associate with building types in some parts of Cross River State, Nigeria

W.E. Azogor; Inyang, I. O. Akpan

Department of Physics, University of Calabar, Calabar.

Corresponding email: azogorwilliam4@gmail.com

Abstract

Measurements of indoor radon concentration levels during wet and dry seasons in some Local Government Areas of Cross River State Calabar Municipality, Calabar South and Yakurr were carried out. A digital radon detector was used for a tile, earth and concrete building floor. The average radon concentration readings for the dry and wet seasons for Calabar Municipality Calabar South and Yakurr were measured to be 0.310pCi/l and 0.26pCi/l, 0.300pCi/l and 0.270pCi/l respectively. Also the average radon concentration levels for tile floor during dry and wet seasons were 0.230pCi/l and 0.130pCi/l respectively, while in concrete floor the average concentrations levels for dry and wet seasons were 0.273pCi/l and 0.300 pCi/l respectively. For the earth floor, the average concentration levels for dry and wet seasons were 0.462pCi/l and 0.418pCi/l respectively. Hence, the average radon concentration levels for the dry and wet seasons in the entire locations were 0.320pCi/l and 0.280pCi/l respectively. A mathematical model was designed using the maple software and standard curves were obtained which were used to predict the indoor radon concentration level for a particular location of tile, concrete and earth building types for wet and dry seasons. The concentration levels of radon in these locations were far below the minimum permissible limit or threshold of 4 pCi/l as recommended by International commission of Radiological Protection.

Keywords: Concentration threshold variations, permissible and digital

Introduction

The human environment is continually exposed to ionizing radiations such as radon resulting from natural radionuclides found within the earth's crust and cosmic rays originating from outside the earth's atmosphere (Inyanget *al.*, 2016). These continuous exposure constitute risk to human health. Radon is a colourless radioactive gaseous element belonging to group 18 of the periodic table (that noble gases); it atomic number is 86; its relative atomic mass is 222 while its density is 9.73gdm^{-3} . It has a melting and boiling points of -71°C and -61.8°C respectively. It is formed by decay of radium – 226 and undergoes alpha decay. It is used in radiotherapy. Radon occurs naturally, particularly in areas underlain by gamete, where it is thought to be a health hazard. As a noble gas, radon is practically merit, although a few compounds, For example, radon fluoride, can be made. It was first isolated by William Ransey and Robert Whytlav- Gray (1877-1958) in 1908. When radon gas is inhaled, densely ionizing alpha particles emitted by deposited short lived decay products of radon (^{218}Po and ^{214}Po) can interact with biological tissue in the lung leading to Deoxyrbonucleic Acid (DNA) damage cancer is generally though to require the occurrence at least one mutation and proliferation of intermediate calls that have sustained some degree of DNA damage available for the development of cancer (Lawrence, 2005).

Radon is present in nature as the only gas found in the natural ^{238}U radioactive decay series. It is a direct progeny of radium – 226 with a half –life of 3.82days. Two other isotopes of radon exist in nature and are members of the ^{232}Th and ^{235}U natural decay series, they are thoron (^{220}Rn) and action (^{219}Rn) are far shorter than that of ^{222}Rn (radon) and as such these isotopes are not as useful for environmental studies (Nezralet *al.*, 2005). It has been recognized as a major cause of lung cancer and it was identified as a human lung carcinogen. Also available literatures about radon limits are based on foreign measurement and monitoring. There is therefore great need to research on indoor radon emanation, exhalation and inhalation in Nigeria based on the geology and climate of Nigeria.

The aim of this study is to measure seasonal variation of radon concentration levels in tiles, concretes and earth building types in some parts of Cross River State, Nigeria.

Study Area

Three major locations namely Calabar South, Calabar Municipality and Yakurr. In each location 42 samples sites were identified. A seven-day reading per sites was obtained using the digital radon meter and this amounted to about 294 day-readings in each location in both dry and wet season.

Materials and Methods

Materials

A digital radon gas detector S/N 23022001580 produced by coventium Home Arthing, Oslo and assembled in Tunisia was used for insitu measurement of the indoor gas concentration levels. This digital device is self-calibrating each time measurement is to be taken for long and short term measurement.

Sample sites

The reading were obtained from Cross River State which lie between 428'N and 6°55'N of the equator and longitude 7°50'E and 6°18'E of the Greewich meridian at different locations such as Calabar Municipality, Calabar South and Yakurr. These geological formations are characterized by unique structural imprints. For a work of this nature before an accurate result could be obtained, monitoring has to be for seven days before reading will be displaced on the detector dashboard. Before the detector is set for use, the used button is pressed, the device revaluated itself to ensure that the results so obtained would be the same elsewhere under the same physical, environment, geological and meteorological conditions for standardization.

Radon emanation

Emanation of radon atoms is defined as the ejection of radon from the source material sometimes after the radioactivity decay of the parental radium isotope. Studies and modes of radon emanation have been performed by a number of research on the physical factors which affect its radon emanation such as soil moisture and porosity, pore size and number, radium concentration and distribution, grain size and shape, atmosphere and soil temperature pressure (Makkanen and Arrela, 2010, Baxeraset *al.*, 2014, Morawska and Philip, 2014).

Environmental radioactivity

Environmental radioactivity is produced by radioactive materials in the environment and aquatic system since the formation of the earth, (Kraveret *al.*, 2004). Radioactivity is the process by which an atomic nucleus of an unstable atom losses energy by emitting particle. It is stochastic process.

Radon diffusion theory

Fuck's law of diffusion of gas through porous media covers radon migration through unstaturated, undisturbed rocks and soils (Nazaroff and Nero, 2008).

$$\frac{d^2c}{dx^2} - \frac{\lambda C}{D} + \frac{E\lambda eR}{DE} = 0 \quad 1$$

Where: C-Interstitial concentration of radon isotope in Bqm⁻³; X-Soil depth measured from the soil air interface in; λ-Radon isotope disintegration constant in per second (S⁻¹); D-Diffusion coefficient of radon isotope in m²S⁻¹; R-Percentage radium isotope's activity concentration of the material in Bkg⁻¹; P -Bulk density of material in kg m⁻³; E-Porosity of the material

The resultant radon exhalation rate at the soil air interface can then be determined by

$$J = DE \left(\frac{dc}{dx} \right) \quad 2$$

Where

$\frac{dc}{dx}$ is determined as the solution from equation 1, equation 2 was becomes

$$J = E_{\text{exp}} PR\lambda^L \tan \frac{H}{L} \quad 3$$

Where: H- Sample thickness in m; L- Diffusion length in m given by

$$L = \sqrt{D/\lambda} \quad 4$$

Radon Exhalation Flux Density: The radon exhalation flux density coming from the earth surface can be described by the equation below

$$\text{div} J + \frac{\partial c}{\partial t} = 0 \quad 5$$

Where C is the radon concentration

$$J = -D \frac{\partial c}{\partial y} \quad 6$$

Where D is the coefficient of Eddy diffusivity putting equ 6 into 5 we have

$$D \frac{\partial^2 c}{\partial y^2} - DC = 0 \quad 7$$

Where D is the radon decay constant integrating equation 7, we obtain the radon concentration.

$$C(y) = A_{\text{exp}} \left(-Z \sqrt{\frac{\lambda}{D}} \right) + B_{\text{exp}} \left(Z \sqrt{\frac{\lambda}{D}} \right) \quad 8$$

Where: A and B are constants whose value will depend on the boundary conditions describing the real physical condition. It is possible to correlate the radon concentration of the soil – air interface, putting a condition on the radon flux itself, $j(0) = \phi$, equivalent to

$$\frac{\partial c(0)}{\partial y} = -\frac{\phi}{D} \quad 9$$

$$\frac{\partial c}{\partial y}(h) = 0 \quad 10$$

$$D \left(-A \sqrt{\frac{\lambda}{D}} + B \sqrt{\frac{\lambda}{D}} \right) = \phi \quad 11$$

$$-A \sqrt{\frac{\lambda}{D_{\text{exp}}}} + \left(-h \sqrt{\frac{\lambda}{D}} \right) + B \sqrt{\frac{\lambda}{D}} + B \sqrt{\frac{\lambda}{D_{\text{exp}}}} \left(h \sqrt{\frac{\lambda}{D}} \right) = 0 \quad 12$$

$$c(y) = \frac{\phi \left\{ e^{-Z \sqrt{\frac{A}{D}}} + e^{-(Zh-Z) \sqrt{\frac{A}{D}}} \right\}}{\sqrt{\lambda D} \left(1 - e^{-Zh \sqrt{\frac{A}{D}}} \right)} \quad 13$$

Equation 13 is able to predict the increasing concentration level while the mixing height decreases. At ground level, $y = 0$

$$C(0) = \frac{\phi \left(1 + e^{-zh\sqrt{\frac{A}{D}}} \right)}{\sqrt{\lambda D} \left(1 - e^{-zh\sqrt{\frac{A}{D}}} \right)} > \frac{\phi}{\sqrt{\lambda D}} \quad 14$$

For $h \rightarrow \infty$

$$C(\infty) = \frac{\phi}{\sqrt{\lambda D}} \quad 15$$

Results

The tables shows the radon concentration levels for Calabar Municipality, Calabar South and Yakurr for wet and dry season. The average radon concentration for Calabar Municipality and Calabar South in wet season was 0.11 pCi/l for a tile room, 0.27 pCi/l for a concrete room and 0.40 pCi/l for a earth room and in the dry season (Table 1 and 2). It was obtained that the average radon concentration for a tile room was 1.64 pCi/l, for concrete room it was 2.56 pCi/l, while for the earth room it was 3.28 pCi/l. From Table 3, the results for the radon concentration level for Yakurr and it envious shows that for a tile room 0.139 pCi/l, for a concrete room 0.318 pCi/l and for an earth room 0.427 pCi/l.

Table 1

Radon concentration levels of different locations in Calabar Municipality and Calabar South in wet season as at May, 2016-October, 2016

s/n	Location	Building type	Radon Concentration Level (pCi/l)	Room Volume (m ³)
1	A	Tiled	0.08	48.47
2	B	Concreted	0.29	48.47
3	C	Tiled	0.1	47.13
4	D	Earthened	0.39	40.2
5	E	Concreted	0.3	100.15
6	F	Tiled	0.12	136.68
7	G	Tiled	0.15	136.68
8	H	Concreted	0.4	107.13
9	I	Concreted	0.35	50.6
10	J	Earthened	0.5	51.72
11	K	Earthened	0.52	48.6
12	L	Earthened	0.39	41.22
13	M	Tiled	0.1	48.27
14	N	Concreted	0.2	48.2
15	O	Tiled	0.09	136.68
16	P	Earthened	0.3	100.15
17	Q	Earthened	0.4	53
18	R	Tiled	0.13	60.73
19	S	Concreted	0.18	65.71
20	T	Concreted	0.18	72.3
21	U	Tiled	0.15	75.4

Discussion

Table 1 and 2 showed that radon exhalation is highest in the earth room and smallest in the tile room because the particles of soil is most compacted in the tile room and least compacted in the earth room. It always observed that radon exhalation is highest in dry season because of the conditions of diffusion. During wet season the exhalation and diffusion of the gas is poor. This is because probably the gas is dissolved by moisture and the conditions for diffusion of the gas is poor. The average radon concentration for dry and wet season for Calabar Municipality Calabar South and Yakurr were obtained. Also the average radon concentration levels for tile floor, concrete floor and earth floor were obtained for wet and dry season.

Table 2

Radon concentration levels of different locations in Calabar Municipality and Calabar South in dry season as at November, 2016-April, 2017

s/n	Location	Building Type	Radon Concentration Level (pCi/l)	Room Volume (m ³)
1	A	TILED	0.2	52.71
2	B	Concreted	0.37	60.7
3	C	Tiled	0.35	60.72
4	D	Earthened	0.53	49.75
5	E	Concreted	0.6	60.71
6	F	Tiled	0.23	131.32
7	G	Tiled	0.4	85.75
8	H	Concreted	0.53	81.32
9	I	Concreted	0.21	138.4
10	J	Earthened	0.22	70.11
11	K	Earthened	0.48	80.32
12	L	Earthened	0.2	150.11
13	M	Tiled	0.15	155.1
14	N	Concreted	0.14	80.13
15	O	Tiled	0.39	70.1
16	P	Earthened	0.37	151.33
17	Q	Earthened	0.14	140.22
18	R	Tiled	0.16	67
19	S	Concreted	0.35	87.4
20	T	Concreted	0.17	160
21	U	Tiled	0.1	160

Conclusion

Radon concentration level for different building types showed that the exhalation rate in earthen building was the highest followed by concrete building, then the tile building. Hence, in terms of safety, the tile building is safer to dwell in because radon is highly radioactive.

REFERENCES

- Baxieras, C. A., Erlandson, B. R. & Font, L. I. (2014). Radon emanation from soils samples, *Radiation Measurements*, 34, 441-443.
- Inyang, S. O., Essien, I. E. & Egbe, N. O. (2016). Exposure level within building materials shop in Calabar, Cross River State, Nigeria. *Journal of Applied Physical Science International*, 7(1); 12-17.
- Kraver, E. M., Amin, N. O. & Paul, T. O. (2004). Characteristics of the German Uranium river cohort study. *Health Physics* 83, 26-34.

- Lawrence, C. A. (2005). A device to distinguish between radon (R_n -222), thoron (R_n -220) and antinon (R_n -219) as the basics of alpha pair counting in centre for medical, Health and Environmental Physics Bribane, Queensland University of Health. P. 120.
- Markkanen, M., & Arrela, H., (2010). Radon Emanation from soils. *Radiation Protection Dosimetry*, 45 (1), 269-272.
- Morawska, L. & Philips, C., (2014). Dependence of the radon emanation coefficient of radium distribution and internal structure of the material. *Geochimica et Cosmochimica*. 51, 1783-1997.
- Nazaroff, W. W. & Nero, A. V. (2008). *Radon and its decay products in indoor air*, New York, John Wiley and Sons, Inc.
- Nezral, M. N. Pye, P. W. & Ganks, T. F. (2005). *The new methods and techniques for the assessment of the radon risk of building sites*, (Zech Geological survey, Prague).

CONGLOMERATE FACIES OF THE AWI FORMATION, CALABAR FLANK, SOUTHEAST NIGERIA: IMPLICATIONS FOR PALEOENVIRONMENTS AND ECONOMIC GEOLOGY

Okon E. E.¹, Kudamnya, E. A.¹, Essien, N. U.^{1,2}, Inyang, D. O.¹, Nton, M. E.³ and Ugbaja, A. N.¹

1. University of Calabar, Calabar – Nigeria
2. AkwaIbom State University, Mkpai-Enin – Uyo, Nigeria
3. University of Ibadan, Ibadan – Nigeria

Abstract

Pebble morphometry of unbroken quartz pebbles collected from the basal section of Awi Formation exposed within the Calabar Flank, southeastern Nigeria was studied for paleoenvironmental reconstruction. The pebbles were selected from 4 locations across the Awi Formation. Morphometric studies carried out involved the measurement of the three orthogonal axes (long, intermediate and short), determination of their corresponding roundness, flatness indices and elongation indices for no less than 200 pebbles. The depositional processes responsible for shaping the pebbles and the environment that prevailed during past geological times were characterised. The results show that the pebbles are sub-rounded to sub-angular and predominantly compact-bladed. The mean values for the following morphometric parameters: Flatness index, elongation ratio, maximum projection sphericity index and oblate-prolate (OP) index are 0.57, 0.78, 0.74 and 15.65 respectively. The results were integrated with deductions from bivariate plots of roundness against elongation ratio and sphericity against OP index. Both plots indicated paleo-depositional environment of the conglomeratic sandstones in a fluvial setting with subordinate transitional marine setting (littoral influence). With the integration of these observations (e.g. pebble imbrication, sediment stratification, grading style), a typical fluvial paleoenvironmental setting is suggested. The geology of the hinterland areas where sediments are sourced contributes largely to the materials at the depocenter. It is possible that the jointing, faulting, sheeting and/or exfoliation of the rocks of the Precambrian Oban Massif, which is believed to be the principal source of the sediments (provenance), also accounts for the abundance of vein quartz in the chosen area of this study. More importantly, besides the significance of pebble morphometry in deciphering paleoenvironments, it also gives clues for potential sites of ore bodies, concentrating them as placer deposits.

Keywords: Morphometric parameters; bivariate, conglomerates; paleoenvironmental reconstruction; fluvial setting; elongation ratio.

1. Introduction

The significance of textural characteristics of sediments as an invaluable tool for characterizing both depositional processes and environment of deposition is well established (Visser 1965; Miall 1985; Essien et al 2016; Miall 2016; Okon and Ojong 2019; Okon et al 2021). The conditions of deposition are well recorded in sedimentary packages and in cases where erosion is not of immense significance, information regarding sedimentation is easily reconstructed using facies analysis. These reflect the assemblages of individual lithofacies elements and other associated geomorphic elements such as channels and bars. The log motif depicted in the vertical profile is often used in concert with the morphology of the channels to carry out interpretation of characteristic depositional environment. The Awi Formation, the oldest known sedimentary deposit in the Calabar Flank, is composed principally of para-conglomerates, cross bedded arkosic sandstones and mudrocks belonging to the basal section of the sedimentary succession of the Calabar Flank, southeastern Nigeria (Okon, 2015).

The study of the sedimentological attributes of the facies succession of the Awi Formation has aided paleoenvironmental interpretation while the integration of geochemistry of the sediments has facilitated provenance determination (Boboye and Okon 2014; Okon 2015). Among the series of approaches used in carrying out paleoenvironmental analyses of any given sedimentary sequence, textural analyses (Awasthi, 1970; Friedman 1979; Inyang and Enang, 2002; Essien and Okon, 2016), analysis of fossil contents (faunal and floral) in the rock unit (Nyong and Ramanathan 1985; Akpan and Ntekim 2004; Itam et al 2016), analysis of traces made by organisms that lived on/within the sediments (Akpan and Nyong 1987; Pemberton et al 2012), and geochemical proxies (Amajor 1987; Tijani et al 2010; Okon, 2015; Ibe and Okon 2021) are important, to mention a few. Paleoenvironmental analysis refers to the study or use of ancient geological materials (rocks) to unravel the environment of depositional which they were deposited and the results have greatly aided interpretations from basin analysis to identification of valuable placer deposits (Okon et al 2018). This study focuses on the conglomerate

facies of the Awi Formation and its significance in characterization of the energy of depositing medium, possible environment of deposition and indication of where placer deposits may be investigated further.

2. Study area description

The study was carried out within the Calabar Flank, southeastern part of the Benue Trough. The basin is underlain by horsts and graben structures and has a NW-SE trending structural orientation (Nyong and Ramanathan, 1985). Fluvial sediments within the Calabar Flank are restricted to the Awi Formation, predominantly clastics, ranging in size from coarse grains to mud class. The area under investigation is delimited to the west by the Ikpe platform and to the east by the Cameroon Volcanic Line. To the south, the Calabar hinge line separates it from the north-eastern portion of the Niger Delta (Fig. 1). Its origin is closely associated with the breakup and subsequent separation of Africa and South America about 120-130Ma ago (Murat, 1972). Sedimentation in the Calabar Flank commenced with the deposition of fluvio-deltaic clastics (Awi Formation) of Neocomian-Aptian age unconformably on the Precambrian crystalline basement complex rocks of the Oban massif. This is succeeded by the shallow marine Mfamosing Limestone, the deeper marine Ekenkpon Shale, New Netim Formation and the Nkporo Shale (Fig 2).

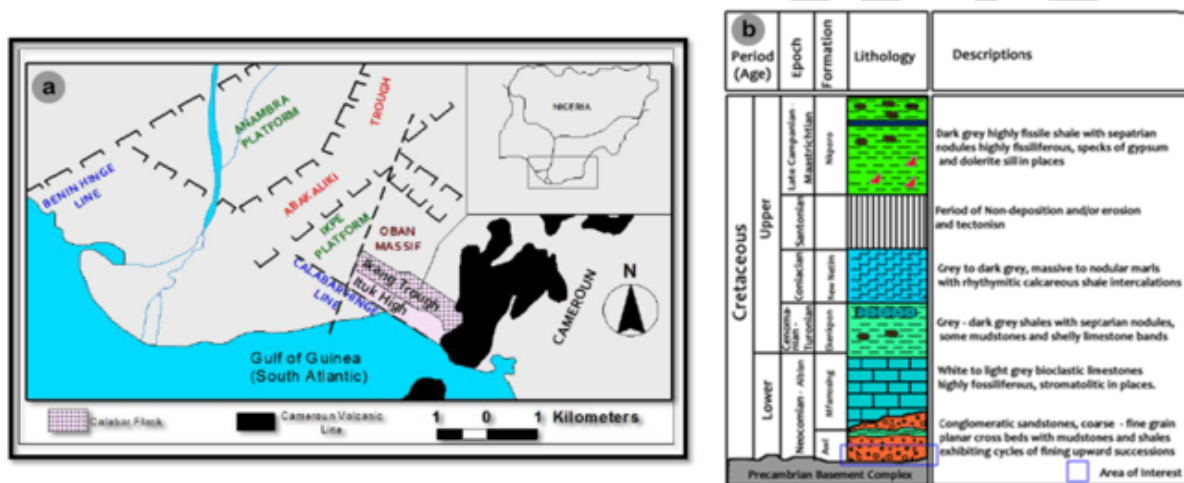


Fig 1. (a) Sketch of southern Nigeria showing the major tectonic elements and geographic location of Calabar Flank; (redrawn from Nyong and Ramanathan 1985) (b) Stratigraphic subdivision of the Calabar Flank (after Okon et al 2017).

3. Materials and Method

Road cut sections of the Awi Formation exposed along Calabar-Itu Highway, Okoyong-Usang-Abasi-Njagachang road, Abiati area and near Ewen community were instrumental to sample collection and analyses for this study. The conglomerates constitute a significant member the basal section of the Awi Formation (Fig 3a-b) and present a non-conformity between the basement rocks of the Oban Massif in the Calabar Flank (Okon et al 2017). The visited outcrops were properly logged and described (Fig. 3c). At each location, 50 unbroken quartz pebbles were collected in 5 batches of 10 each. The analysis was carried out with the mean form of at least 10 pebbles taken from each sampling station. In each case, 5 sets per sample location representing 50 pebbles for the four locations visited. The measured parameters were subjected to statistical and mathematical treatments as presented in the ratios in table 1.

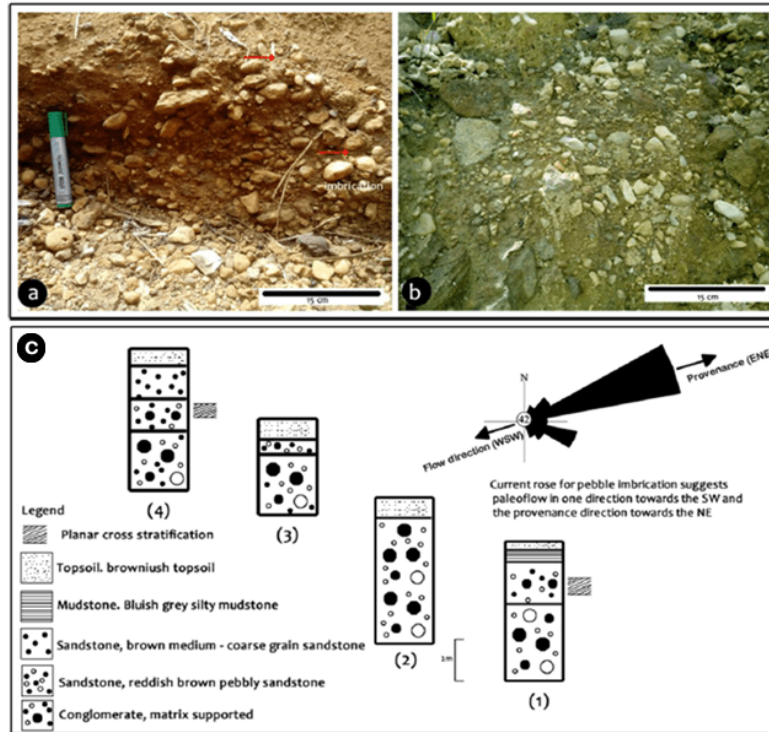


Figure 3. (a) Matrix – supported conglomerates showing clast imbrication; (b) admixture of brecciated rock units with sub-rounded pebbles; (c) Lithologic log for the studied samples

Table 1. Formulae used for statistical computation of pebble morphometric parameters

S/No	Formula	Reference
1	Maximum Projection Sphericity Index (MPSI) = $\{S^2/Li\}^{1/3}$	Sneed and Folk, 1958
2	Elongation Ratio, (ER) = l/L	Lutig, 1962; Sames, 1966
3	Flatness Ratio $FR = S/L$	Lutig, 1962
4	Flatness Index = $(L - l + S) / L$	Illenberger 1992
5	Oblate – Prolate Index, $OPI = \frac{10 \left(\frac{L-l}{L-S} - 0.50 \right)}{S/L}$	Dobkins and Folk 1970
6	Roundness = $[(\sum r)/nR]$	Waddell 1932

4. Results and discussion

The result for the mean pebble morphometric parameters is presented in Table 2. The pebbles are massive, matrix supported and crudely bedded, with clasts diameter ranging from 2.63 – 3.40 cm. The sorting is poor and pebble grains are weakly imbricated, with brecciated ferruginized layer admixed with sub-rounded pebbles (see Figs 3a-b). These features suggest lag deposits and conform to Miall (1978) facies classification “Gm”. Regarding the clast sphericity, roundness and “Oblate – Prolate” Indexes, the parametric values of an average of 10 pebbles was used in the analysis (Sneed and Folk 1958). This formula was adopted because it was established comparing the volume of the particle with its maximum projection area which naturally opposes the direction of motion. The results show that the pebbles are sub-rounded to sub-angular and predominantly compact-bladed (Fig 4). Shape indices as paleoenvironmental indicators of quartzite rich rocks have been the subject of considerable discussions among experts (Els, 1988; Lorang and Komar, 1990; Illenberger and Reading, 1993), and the result have greatly aided interpretations from basin analysis to identification of valuable placer deposits. The mean values for the following morphometric parameters: Flatness index, elongation ratio, maximum projection sphericity index and oblate-prolate (OP) index are 0.57, 0.78, 0.74 and 15.65 respectively. The various parameters obtained in the pebble morphometric analysis together, with the different bivariate scatter plots have shown that the depositional environment of the basal of the section of the Awi Formation is predominantly influence by fluvial processes with beach / littoral influence (Figs 5a-b; Figs 6a-b) (Lutig, 1962; Okon et al 2019).

Table 2. Result for the mean values of 20 batches of pebble morphometric parameters for AwiFormation

S/N	L	S	I	S/L	Elongation (I/L)	L-I	L-S	L-I	S-I	OPI	MPS	FI	Roundness	Form name
L1/B1	2.80	1.63	2.31	0.58	0.83	0.49	1.17	6.47	2.66	-0.74	0.74	76.04	0.38	CB
L1/B2	2.67	1.44	2.26	0.54	0.85	0.41	1.23	6.03	2.06	-1.95	0.70	69.22	0.4	CB
L1/B3	3.04	1.82	2.30	0.60	0.76	0.74	1.22	6.97	3.29	0.69	0.78	84.39	0.43	CB
L1/B4	3.17	1.71	2.47	0.54	0.78	0.70	1.46	7.83	2.92	-0.46	0.72	75.74	0.41	CB
L1/B5	2.63	1.39	2.02	0.53	0.77	0.61	1.24	5.29	1.93	-0.24	0.71	76.28	0.34	CB
L2/B6	2.82	1.60	2.12	0.57	0.75	0.70	1.23	5.98	2.54	0.43	0.75	81.54	0.46	CB
L2/B7	2.68	1.69	2.04	0.63	0.76	0.64	0.99	5.46	2.86	2.21	0.81	87.77	0.45	CB
L2/B8	3.40	1.63	2.43	0.48	0.71	0.97	1.77	8.23	2.64	-0.03	0.68	75.99	0.47	CB
L2/B9	2.68	1.58	2.26	0.59	0.84	0.43	1.11	6.04	2.48	0.02	0.74	77.55	0.42	CB
L2/B10	2.63	1.66	2.04	0.63	0.78	0.59	0.97	5.34	2.76	0.50	0.80	85.88	0.39	CB
L3/B11	2.69	1.57	2.00	0.58	0.74	0.69	1.12	5.38	2.46	0.97	0.77	82.54	0.41	CB
L3/B12	2.97	1.52	2.06	0.51	0.69	0.92	1.46	6.10	2.30	0.14	0.72	80.89	0.41	CB
L3/B13	2.75	1.45	2.23	0.53	0.81	0.52	1.31	6.13	2.09	-0.69	0.70	71.91	0.44	CB
L3/B14	2.76	1.63	2.19	0.59	0.79	0.57	1.13	6.02	2.66	0.15	0.76	80.43	0.44	CB
L3/B15	2.97	1.53	2.32	0.52	0.78	0.65	1.44	6.89	2.34	-0.21	0.70	72.98	0.47	CB
L4/B16	2.38	1.31	1.85	0.55	0.78	0.53	1.07	4.40	1.72	-0.04	0.73	76.95	0.38	CB
L4/B17	2.49	1.45	1.97	0.58	0.79	0.52	1.04	4.91	2.10	0.003	0.75	79.66	0.44	CB
L4/B15	2.49	1.47	1.88	0.59	0.76	0.61	1.02	4.68	2.16	0.53	0.77	83.62	0.39	CB
L4/B19	2.55	1.56	2.05	0.61	0.80	0.50	0.99	5.23	2.43	0.64	0.78	82.88	0.44	CB
L4/B20	2.55	1.48	1.90	0.58	0.75	0.65	1.07	4.85	2.19	0.80	0.77	84.30	0.42	CB
Mean	2.75	1.55	2.13	0.57	0.78	0.62	1.20	5.91	2.43	-0.74	0.74	76.04	0.42	-

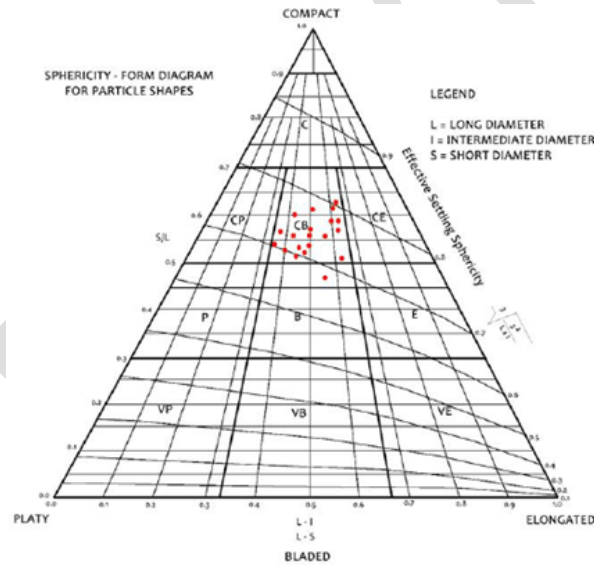


Figure4. Sphericity-form diagram for particle shapes (after Sneed and Folk 1958) – majority of the samples plot under compact bladed pebbles

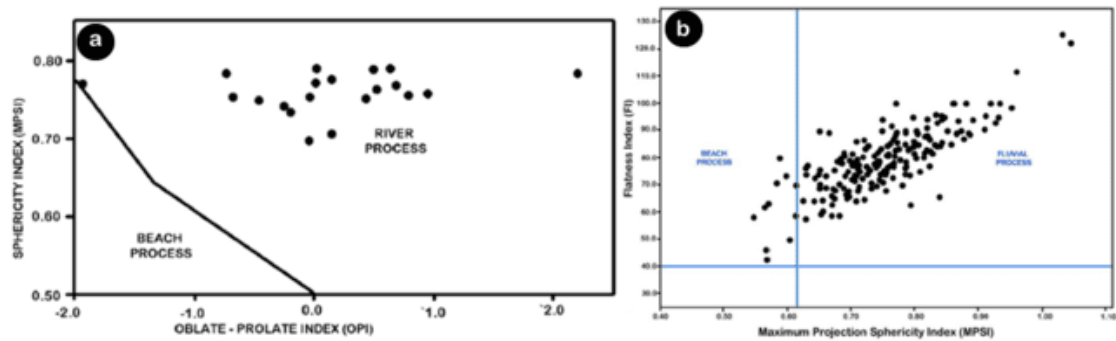


Fig 5. (a) Plot of sphericity against OPI (fields after Dobkins and Folk, 1970); (b) Plot of flatness index (FI) against maximum projection sphericity index (fields after Stratten 1974)

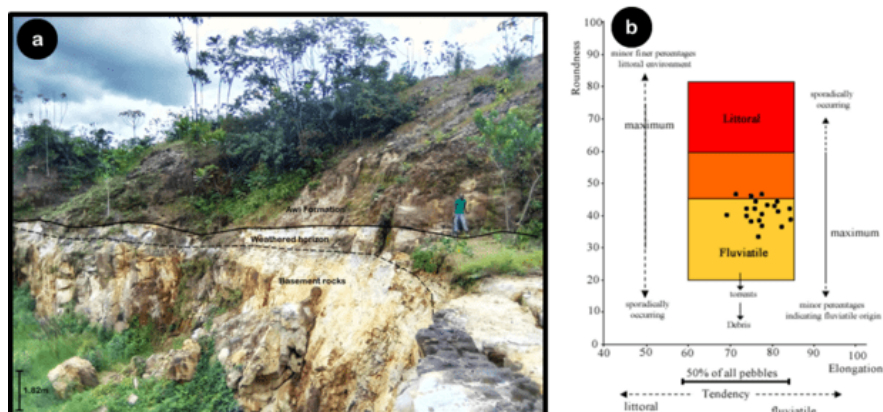


Figure 6. (a) Field photo showing fining upwards succession in the Awi Formation with the conglomeratic facies of the Awi Formation at the base; (b) Environmental determination chart (modified after Sames 1966).

The sediments along the transect show energy setting associated with a mix of high and relaxed energy at different times during depositional process, establishing conditions for selective gravitational sorting of the sediments, allowing for the possible concentration of placers of economic interests at various depocenter. In such cases, transient placer deposits (alluvial-colluvial and placers and fluvial placers) are the most likely targets for mineralization. These sediments are typically deposited by traction currents as bedloads and as bed contact loads and this interpretation is justified by their accumulation in different bedforms.

Conclusion

In this study, pebble morphometric analysis has aided the determination of paleoenvironments prevalent during the deposition of the basal conglomerates of Awi Formation. The depositional processes (abrasion conditions) responsible for shaping the pebbles were characterized from the study of the clasts morphology and their imbrication azimuth. Fluvatile process with some overlapping littoral influence has been shown to be responsible for the variation in clasts morphology of the para-conglomerates (matrix-supported) of Awi Formation. It is possible that the jointing, faulting, sheeting and/or exfoliation of the rocks of the Oban Massif, which is believed to be the principal source of the sediments (provenance), also accounts for the abundance of vein quartz in the area which was eventually adapted for this study. The recent mining activities at the southern fringes of the Oban Massif, involving gold extraction from active recent drainage channels suggests that some paleo-channels may form targets for gold exploration if proper attention is given to studies of this nature. This hypothesis is yet to be tested but is drawn from the fact that the rocks from where the gold grains are mined pre-date the Awi Formation and studies have shown that the provenance of the Awi Formation is the nearby Oban Massif rocks. This forms a template for further research in the area, especially geared towards testing the outlined hypothesis.

References

- Akpan E. B. and Ntekim E. E. (2004). Cretaceous bivalves and paleoenvironments of the Calabar Flank, SE Nigeria. *Global Journal of Geological Sciences*, 2(1), 15-36.
- Akpan E. B. and Nyong E. E. (1987). Trace fossils assemblage and depositional environment of Turonian calcareous sandstones in the southern Benue Trough, Nigeria. *Journal of African Earth Sciences*, 6(2), 175-180.
- Amajor L. C. (1987). Major and trace element geochemistry of Albian and Turonian shales from the Southern Benue trough, Nigeria. *Journal of African Earth Sciences*, 6(5), 633-641.
- Awasthi A. K. (1970). Skewness as an environmental indicator in the Solani river system, Roorkee (India). *Sedimentary Geology*, 4, 177-183.
- Boboye O. A. and Okon E. E. (2014). Sedimentological and geochemical characterization of the Cretaceous strata of the Calabar Flank, southeastern Nigeria. *Journal of African Earth Sciences*, 99, 427-441.
- Dobkins J. E. and Folk R. L. (1970). Shape development of Tahiti-Nui. *Journal of Sedimentary Petrology*, 40, 1167-1203.
- Els B. G. (1988). Pebble morphology of an ancient conglomerate: The middelvlei gold placer, Witwatersrand, South Africa. *Journal of Sedimentary Petrology*, 58(5), 894-901.

- Essien N.U. and Okon E. E. (2016). Sedimentary processes and depositional environments of part of the Benin Formation in the Calabar Flank, southeastern Nigeria. *British Journal of Applied Science and Technology*, 18(5), 1–10.
- Essien N. U., Ilori A. O., Okon E. E. and Njoku S. N. (2016). Textural characteristics and depositional processes of sediments from a 47 km transect in the Niger Delta, Southern Nigeria. *Journal of Geography, Environment and Earth Science International*. 7(1), 1-11.
- Friedman, G. (1979). Differences in size distributions of populations of particles among sands of various origins. *Sedimentology*, 26, 3-32
- Ibe C. U. and Okon E. E. (2021). Provenance and tectonic settings of the Eze-Aku Sandstone (Turonian) in Awajir and adjoining areas, Southern Benue Trough, Nigeria: evidence from petrography and geochemistry. *Journal of Sedimentary Environments*, 6, 237–254
- Illenberger W. K. (1992). Pebble shape (and size). *Journal of Sedimentary Petrology*. 61(5), 756–767.
- Illenberger W.K. and Reddering J. S. V. (1993). An evaluation of shape indices as paleoenvironmental indicators using quartzites and metavolcanicsclasts in Upper Cretaceous to Paleogene beach, river and submarine fan conglomerates (Discussion). *Sedimentology*, 40, 1019–1020
- Inyang D. O. and Enang E. I. (2002). Use of statistical parameters in the sedimentological study of conglomerates in the north-eastern part of the Niger Delta, Nigeria. *Journal of Mining and Geology*, 38(1), 21-26.
- Itam A. E., Inyang D. O., Umana U. S., Akpan E. B., Udimwene E. (2016). Palynostratigraphy of Lemna Road Transect of Benin Formation, Calabar Flank, Nigeria. *Asian Journal of Physical and Chemical Sciences*, 1(1), 1-10.
- Lorang M. S. and Komar P. D. (1990). Pebble shape. *Nature*, 347, 433–434.
- Lutig G. (1962). The shapes of pebbles in continental, fluvial and marine facies. International Association of Geology and Geophysics. *International Association of Hydrogeologists*. 59, 253-258.
- Miall, A. D. (2016). Facies Analysis. In: *Stratigraphy: A Modern Synthesis* (pp 77–159). Springer, Cham. https://doi.org/10.1007/978-3-319-24304-7_3
- Miall A. D. (1978). Lithofacies types and vertical profile models in braided river deposits: A summary. In: A. D. Miall (Ed), *Fluvial Sedimentology* (pp 597-604). Canadian Society Petroleum Geologists Memoir, 5.
- Miall, A. D., (1985). Architectural-element analysis: A new method of facies analysis applied to fluvial deposits. *Earth Science Reviews*, 22, 261-308.
- Murat R. C. (1972). Stratigraphy and paleogeography of the Cretaceous and Lower Tertiary in Southern Nigeria. In: T.F.J.Dessauvagie and A.J. Whiteman (Eds). *African Geology* (pp 251-266). University of Ibadan Press, Nigeria
- Nyong E. E. and Ramanathan R. M. (1985). A record of oxygen deficient paleoenvironments in the Cretaceous of the Calabar Flank, southeast Nigeria. *Journal of African Earth Science*, 3(4), 455–460.
- Okon, E. E. (2015). Sedimentologic characteristics and provenance studies of AWI Formation, Calabar Flank, SE Nigeria. *Journal of Mining and Geology*, 51(2), 121–138.
- Okon E. E., Essien N. U. and Adeyemi G. O. (2017). Geochemistry of sandstones of Awi Formation, Southeastern Nigeria: Implications for weathering, provenance and tectonics settings. *International Journal of Science and Technology*, 6(4), 742–755.
- Okon E. E., Asi M. O. and Ojong R. A. (2018). Morphometric studies of pebbles from Ewen area, Calabar Flank, Southeastern Nigeria: Implications for paleoenvironmental reconstruction. *Physical Science International Journal*, 17(4), 1-12.
- Okon, E. E. and Ojong, R. A. (2019). Paleoenvironmental Analysis and Its Significance in Sedimentology: Case Study of the Conglomerate Facies of the Awi Formation, Calabar Flank, Southeast Nigeria. In: M. Rafatullah, (Ed). *Advances and Trends in Physical Science Research*. 2. (pp84 – 98).
- Pemberton S. G., MacEachern J. A., Dashtgard S. E., Bann K. L., Gingras M. K. and Zonneveld J. P. (2012). Shorefaces. In: D. Knaust and R. G. Bromley (Eds.), *Trace Fossils as Indicators of Sedimentary Environments* (pp 563–604). Developments in Sedimentology. 64.
- Sames C. W. (1966). Morphometric data of some recent pebble associations and their application to ancient deposits. *Journal of Sedimentary Petrology*, 36, 126–142.

- Sneed E. D. and Folk R. L. (1958). Pebbles in the lower Colorado river, Texas. A study in particle morphogenesis. *Journal of Geology*, 66, 114–150.
- Stratten T. (1974). Notes on the application of shape parameters to differentiate between beach and river deposits in southern Africa. *Transactions of the Geological Society of South Africa*, 77, 59-64
- Tijani M. N., Nton M. E. and Kitagawa R. (2010). Textural and geochemical characteristics of the Ajali Sandstone, Anambra Basin, SE Nigeria: Implication for its provenance. *Comptes Rendus Geoscience* 342, 136–150
- Visher, G. S. (1965). Use of vertical profile in environmental reconstruction; American Association of Petroleum Geologists Bulletin, 49, 41-61.

NCPS-2024

Yttrium-encapsulated Tin-decorated carbon-based fullerene (Sn-Y@C₈₀) nanostructured as a sensor for AsH₃, NH₃, PH₃ and SbH₃: Insight from DFT and QTAIM

Chioma G. Apebende^{1*}, Ismail O. Amodu², Ubua U. P¹, Simon James³, Marshal A. Igbilishi³, Mercy A. Sunday³ and Stephen A. Adalikwu⁴

¹Department of Pure and Applied Chemistry, University of Calabar, Calabar, Nigeria; ²Department of Mathematics and Statistics, University of Calabar, Calabar, Nigeria; ³Department of Science Laboratory Technology, University of Calabar, Calabar Nigeria; ⁴Department of Chemistry, Cross River University of Technology

*Corresponding author: gloriaapebende@gmail.com

Abstract

The increasing release of toxic gases (AsH₃, NH₃, PH₃, and SbH₃) from industrialization and fossil fuel combustion poses serious health risks. This study explores the adsorption capability of Tin-doped Yttrium-encapsulated fullerene (Sn-Y@C₈₀) using DFT/M06-2X/GEN/Auto/LanL2DZ/Def2-SVP computational methods. Structural changes were observed upon adsorption, the energy gap falls within a close range of (2.067–2.069 eV), making it suitable for semiconductor applications. The adsorption followed a chemisorption mechanism, with adsorption energy in the order PH₃ < AsH₃ < SbH₃ < NH₃, where NH₃ had the highest (-3.1690 eV). PH₃ showed the shortest recovery time (182.72s), making Sn-Y@C₈₀ highly efficient for PH₃ detection. The material's moderate recovery time for other gases suggests its potential for future gas sensor applications.

Keywords: Fullerene, adsorption, QTAIM, sensor mechanism, recovery time, DFT

1.0 Introduction

Gas pollutants such as PH₃, NH₃, SbH₃, and AsH₃ originate from industrial activities, chemical processes, and mining, posing significant risks to human health (Ai-Taai and M ahammed, 2022), the environment, and industrial operations. These toxic gases contribute to air pollution, leading to respiratory issues, poisoning, and climate change effects. Therefore, effective detection and management of these gases are crucial for environmental safety.

Nanomaterials, including carbon nanotubes, metal oxides, and graphene, have gained attention for gas sensing due to their high surface area, selectivity, and sensitivity. Nano-based sensors, which interact with target gas molecules and produce measurable signals, have been developed for air quality monitoring and gas leakage detection (Agache *et al.*, 2022). However, challenges such as high costs and equipment size hinder their widespread adoption. Advances in doped nanotube materials, UV light activation, and gas adsorption mechanisms (chemisorption and physisorption) offer promising solutions for improving gas sensing performance. Fullerenes, a class of carbon allotropes, exhibit unique electronic, chemical, and thermal properties, making them ideal for gas sensing applications. Studies have demonstrated the effectiveness of C₆₀, C₇₀, and C₈₀ fullerenes in adsorbing toxic gases.

This study explores the Tin-doped Yttrium-encapsulated fullerene (Sn-Y@C₈₀) nanostructure for gas adsorption. Yttrium was chosen for its large atomic radius and metallic characteristics, while Tin was used for its high electronegativity and catalytic properties. The combination of these elements is expected to create diverse adsorption sites, enhancing selectivity and catalytic activity. The study employs Density Functional Theory (DFT) to analyze the structural geometry, electronic properties, adsorption energy, and sensor mechanisms of this modified fullerene for toxic gas detection.

2.0 Computational Methodology

All computational calculations were performed under the framework of the density functional theory (DFT). Employing the Gaussian 16 program (Dennington *et al.*, 2016) together with the GaussView 6.0.16 graphical user interface (Walker *et al.*, 2013), sketching and optimizations of molecular structures were performed. The M06X-2X functional (Chiodo *et al.*, 2006) has been known for accuracy and precision among the hybrid DFT functionals. Hence it is used in the present investigation. Due to the presence of heavy metals (Sn and Y), the LANL2DZ (Los Alamos National Laboratory 2 Double-Zeta), developed by Hay and Wadt was used specifically for the metal atoms (Mohammedi *et al.*, 2023). For the prediction of adsorption phenomena, this M06X-2X/Gen/Auto/LanL2DZ

computational method has been known for delivering to the nearest accuracy. All molecular structures have been sketched and visualized prior to optimization using GaussView 6.0.16 software. To investigate the influence of adsorption on the morphology of the tailored molecule, structural geometry optimization was carried out. The adsorption energy, structural geometry, quantum description and weak interaction studies have been carried out via geometry optimization processes. The energy of the complex formed were computed using the **Equation (1)**.

$$E_{ad} = E_{Complex} - (E_{Surface} + E_{Molecule}) \quad (1)$$

Where E_{ad} is the adsorption energy, $E_{Molecule}$ is the energy of the adsorbate, $E_{Surface}$ energy of the molecule, and $E_{Complex}$ energy of the complex respectfully. The multiwfn 3.7 program was employed for the calculation of the non-covalent interactions (NCI) and quantum theory of atoms-in-molecules (QTAIM), then visualized using the visual molecular dynamic (VMD) software package. The Chemcraft 1.6 Software Program was used to generate the HOMO-LUMO plots and visualized the molecular structures of the gases. The NBO was computed using the NBO using the NBO package embedded in the Gaussian software.

3.0 Result and Discussion

3.1 Electronic Properties

The results of an electronic property investigation are pertinent in estimating and determining the electronic property of any system. This investigation includes various parameters such as ionization potential energies, softness, electron affinity and chemical hardness amongst others (Akman *et al.*, 2022). Ionization potential (IP) is the required energy needed for an electron to move from ground state to excited state (when an atom gain an extra electron). Using the Koopman's theorem postulated by the Tjalling Koopmans, the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) values can be extracted and used in determining the charge potential and differences (Liu *et al.*, 2023). The HOMO-LUMO in this investigation was calculated using the **Equations 2 - 7** (Timothy *et al.*, 2023) and the results are shown in **Table 3**.

$$IP = -E_{HOMO} \quad (2)$$

$$EA = -E_{LUMO} \quad (3)$$

$$\eta = \frac{1}{2} (IP - EA) = \frac{E_{LUMO} - E_{HOMO}}{2} \quad (4)$$

$$\mu = -\frac{1}{2} (E_{HOMO} + E_{LUMO}) \quad (5)$$

$$S = \frac{1}{2\eta} = \frac{1}{IP - EA} = \frac{1}{E_{LUMO} - E_{HOMO}} \quad (6)$$

Electrophilicity index (ω), was calculated using the equation given as thus:

$$\omega = \frac{\mu^2}{2\eta} \quad (7)$$

Where the ionization potential, electron affinity, chemical hardness, chemical potential, chemical softness, and electrophilicity index are denoted by IP, EA, η , μ , S, and ω respectively. Also, it is important to know that the E_{HOMO} and E_{LUMO} are the energies of HOMO and LUMO respectively. For the complexes Sn-Y@C₈₀, AsH₃-Sn-Y@C₈₀, NH₃-Sn-Y@C₈₀, PH₃-Sn-Y@C₈₀, and SbH₃-Sn-Y@C₈₀, the energy gap values are 2.025 eV, 2.069 eV, 2.068 eV, 2.067 eV and 2.068 eV, respectively. When compared to its counterparts, AsH₃-Sn-Y@C₈₀ complex has the largest energy gap, which tends to indicate low conductivity and chemical stability. In contrast, SbH₃-Sn-Y@C₈₀ complex was observed to hold the lowest energy gap, with a difference of 0.002 eV. When Sn-Y@C₈₀ interacted with SbH₃ at the Stibine site, there was a discernible decrease in the energy gap. This could have been caused by the interaction's dual effects of a simultaneous rise in the LUMO energy and a decrease in the HOMO energy. A system tends to be more reactive and less stable when its chemical hardness value is lower, and vice versa. **Table 3** shows that SbH₃-Sn-Y@C₈₀ has the lowest chemical hardness (1.033 eV) and the highest electrophilicity (28.038 eV), indicating a more reactive and sensitive as compared to their studied counterparts. The HOMO-LUMO plots before and after adsorption were shown in **Figure 4**.

Table 3. The quantum chemical descriptors (all units in electron Volt (eV) except chemical softness (S), which has a value of eV⁻¹), calculated at the DFT/MO62X/Gen/AUTO/LanL2DZ computational method.

Systems	E _{HOMO}	E _{LOMO}	E _{gap}	IP	EA	μ	η	S	ω	ELF
Sn_Y@C ₈₀	-6.434	-4.409	2.025	6.434	4.409	-5.421	1.013	0.493	14.505	-5.421
AsH ₃ _Sn_Y@C ₈₀	-8.681	-6.612	2.069	8.681	6.612	-7.647	1.035	0.483	28.250	-7.467
NH ₃ _Sn_Y@C ₈₀	-8.722	-6.654	2.068	8.722	6.654	-7.688	1.034	0.483	28.580	-7.688
PH ₃ _Sn_Y@C ₈₀	-8.687	-6.619	2.068	8.687	6.619	-7.653	1.034	0.483	21.854	-7.653
SbH ₃ _Sn_Y@C ₈₀	-8.644	-6.577	2.067	8.644	6.577	-7.611	1.033	0.483	28.038	-7.611

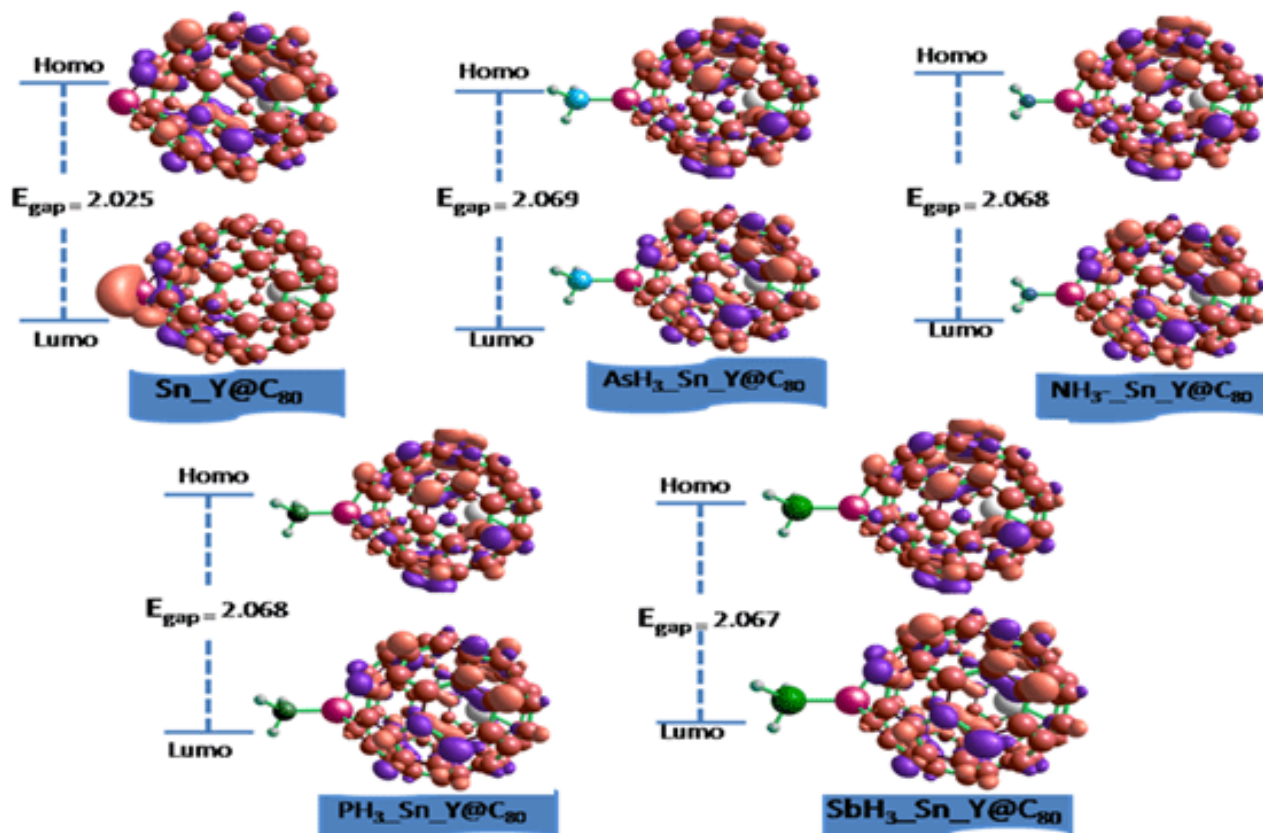


Figure 4. The generated HOMO-LUMO plots of the complexes formed after adsorptions

3.2 Gases adsorption on the Sn-Y@C₈₀ molecule

In order to ascertain the most stable adsorption configuration, energy minimization was first considered (Lu *et al.*, 2022). To determine the minimum energy configuration, the coordinated molecule have been optimized. The ability of Sn-Y@C₈₀ molecule to detect Arsine (ASH₃), Ammonia (NH₃), Phosphine (PH₃) and Stibine (SbH₃) gas molecules was examined in this present section by looking thoroughly into the intricacy of gas adsorption from the adsorption energy. It is noteworthy to mention that the degree of the strength of the interaction existing between adsorbate and adsorbent can be understood via the adsorption energy (Zhao *et al.*, 2022). Adsorption energy is also required in order to determine how closely an adsorbate is adsorbed by an adsorbent. During adsorption, four distinct stable complexes were formed following: AsH₃-Sn-Y@C₈₀, NH₃-Sn-Y@C₈₀, PH₃-Sn-Y@C₈₀ and SbH₃-Sn-Y@C₈₀ respectively. The high negative value of E_{ads} shows a strong adsorption/interaction between the gas molecules (adsorbate) and the molecule (adsorbent). Mathematically, the adsorption energy has been calculated using the expression in **Equation (1)**. The greater the magnitude of E_{ads}, the stronger the adsorption, thereby leading

to a relatively better adsorbing property (Odeyet *et al.*, 2023). **Table 5** presents the results of the adsorption energy calculated for all the complexes, with all units in electron volt energy (eV). As a result of the negative adsorption energy, the gases adsorption is said to have chemisorbed on the tailored molecule [18]. Generally, the adsorption follows an increasing trend of: $\text{PH}_3\text{-Sn-Y@C}_{80} < \text{AsH}_3\text{-Sn-Y@C}_{80} < \text{SbH}_3\text{-Sn-Y@C}_{80} < \text{NH}_3\text{-Sn-Y@C}_{80}$. Greatest adsorption energy value of -3.1690 eV exhibited by $\text{NH}_3\text{-Sn-Y@C}_{80}$, whereas $\text{PH}_3\text{-Sn-Y@C}_{80}$ reflects the least energy of -0.7728 eV, indicating the strongest and weakest adsorption among the studied systems. The $\text{SbH}_3\text{-Sn-Y@C}_{80}$ showcased a relatively higher energy than $\text{PH}_3\text{-Sn-Y@C}_{80}$ and $\text{AsH}_3\text{-Sn-Y@C}_{80}$, depicting a relatively stronger adsorption than the aforementioned complexes. In all cases, the studied molecule showcased a good adsorbent behavior towards the adsorption of the labelled gases.

Table 5. Computed adsorption energy results for the four complexes calculated at DFT/MO62X/Gen/AUTO/LanL2DZ computational method.

Complex	E_{Complex}	E_{Molecule}	E_{Gas}	Eads (H)	Eads (eV)
AsH₃-Sn-Y@C₈₀	-3056.012993	-3048.374689	-7.569567	-0.068737	-1.8704
NH₃-Sn-Y@C₈₀	-3104.662600	-3048.374689	-56.171449	-0.116462	-3.1690
PH₃-Sn-Y@C₈₀	-3391.163719	-3048.374689	-342.760631	-0.028399	-0.7728
SbH₃-Sn-Y@C₈₀	-3055.280844	-3048.374689	-6.803633	-0.102522	-2.7898

3.3 Visual Study

3.3.1 Reduced density gradient (RDG)

The non-covalent interaction (NCI) study (Sukanya *et al.*, 2022) to effectively analyze and distinguish the interactions identified between the nano-materials and to comprehend the kinds of non-covalent interaction once the complexes have been architecturally imaged. The NCI makes use of the second eigenvalue of the Hessian matrix (λ_2) and the electron density and its gradient (RDG), which are plotted and visualized in the **Figure 5**. The NCI investigation consists of compelling forces such as hydrogen bonds, electrostatic forces, and Van der Waals (repulsive forces), which offer important insights into the type and intensity of interactions within complexes and surfactant behavior (Lohith *et al.*, 2022). In order to determine the non-covalent contact and how it impacts the capacity of the labelled gases to sense, the interaction between the $\text{Sn}_Y\text{@C}_{80}$ molecules was examined in this work. The NCI was envisioned and the 3D RDG plots were included in the assessment of the readings acquired by the VMD application. This type of contact is indicated by the colored region that resembles a loop around the molecules of AsH_3 , PH_3 and SbH_3 . Moreover, the region colored blue denotes a strong attraction where it is seen that the second Eigen function ($(\lambda_2)\rho(r) < 0$) is observed to be negative (Po, 2023). According to Bakhsh *et al.* (2023), the presence of repulsive forces of attraction is indicated by the red hue where the second Eigen function ($(\lambda_2)\rho(r) < 0$) is positive, and has been displayed in green color where the second Eigen function is equal to zero ($(\lambda_2)\rho(r) = 0$). As shown in the picture below the iso-surface of $\text{AsH}_3\text{-Sn-Y@C}_{80}$, $\text{PH}_3\text{-Sn-Y@C}_{80}$ and $\text{SbH}_3\text{-Sn-Y@C}_{80}$ is dominated by blue and red colors respectively, suggesting a stronger electrostatic force and less steric interaction. Furthermore, a stronger presence of attraction force is noticed in the complex $\text{NH}_3\text{-Sn-Y@C}_{80}$ due to the blue color dominated in the loop, while $\text{Sn}_Y\text{@C}_{80}$ has no color around the loop region. From this result, $\text{AsH}_3\text{-Sn-Y@C}_{80}$, $\text{PH}_3\text{-Sn-Y@C}_{80}$ and $\text{SbH}_3\text{-Sn-Y@C}_{80}$ show the presence of strong attraction and repulsive force of attraction and are also with the potential to be used as an adsorbent material.

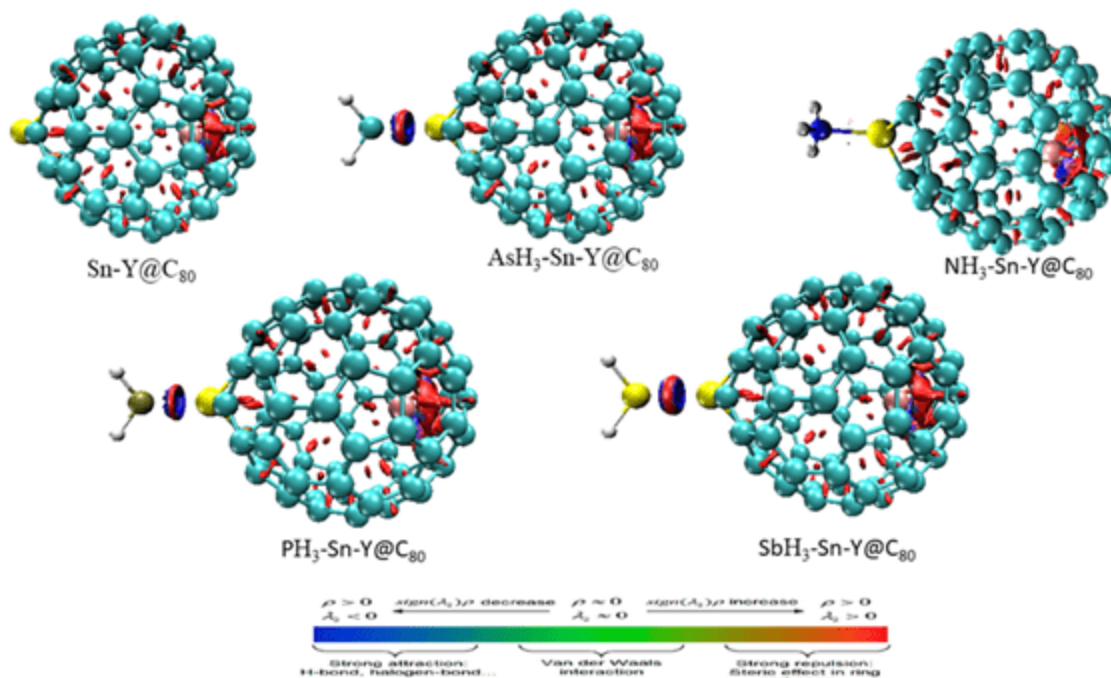


Figure 5. The 3D RDG plots obtained from VMD visualizing the NCI among the studied systems

3.3.2 Recovery Time

Desorption or recovery time can be seen as a “reversible process” of adsorption. After an adsorbate has been adsorbed by an adsorbent, the recovery of the adsorbate becomes essential as the reusability of a sensor device depends on its recovery time. Oftentimes, shorter recovery time depicts that the adsorbent materials in question has the ability to desorb the adsorbate in the shortest possible time, and this is required for a good sensor material. Among the studied systems, $\text{PH}_3\text{-Sn-Y@C}_{80}$ has the least recovery time of 182.72 s, depicting that the adsorption of PH_3 gas molecules on the Sn-Y@C_{80} molecule will take about 3 minutes to desorb the gas from the surface. This is an indicative of excellent adsorption as it will make recovery effortless. Other systems show a relatively moderate recovery time of which desorption of gases upon adsorption will take couple of hours.

Table 7. Summary of the results obtained for the charge transfer mechanism, FET, and E-back donation calculated at DFT/MO62X/Gen/AUTO/LanL2DZ computational method.

Systems	Q_t (eV)	ΔN (eV)	% ΔN percentage	ΔE Back-donation	Recovery time (τ)
$\text{AsH}_3\text{-Sn-Y@C}_{80}$	1.567	-0.024	-2.400	-0.25875	3.29×10^{22}
$\text{NH}_3\text{-Sn-Y@C}_{80}$	3.265	-0.024	-2.400	-0.2585	3.04×10^{46}
$\text{PH}_3\text{-Sn-Y@C}_{80}$	1.859	-0.023	-2.300	-0.2585	182.72

SbH₃-SnY@C₈₀	0.983	-0.022	-2.200	-0.25825	3.05 x 10 ³⁹
---	-------	--------	--------	----------	-------------------------

Note: The unit of Q_t is in electron (e), ΔN is dimensionless, and the $\% \Delta N$ has its unit in percentage. The $\Delta E_{\text{back-donation}}$ has a unit of electron volt (eV). The recovery time has its unit in seconds (s).

4.0 Conclusions

In this study, fullerene C₈₀ was modified, forming a newly tailored Sn-Y@C₈₀ molecular surface. This newly formed material has been examined via the density functional theory to inquire into the adsorption capability of the molecule. Various computational techniques such as the geometry optimization; the FMO and NBO analyses for the electronic properties; the AIM and NCI analyses for visual study; adsorption energy for the strength and phenomena of adsorption; and lastly the mechanism of sensing ranging from the charge transfer mechanism, electrical conductivity, fraction of electron transfer, electrical back-donation to the recovery time. Finally, after thorough analyses the following conclusions were drawn:

1. The changes in the morphology of the molecule suggests that there is a stretch in the molecule upon adsorption. Also, the interaction distance lies in the narrow range of 2.616 to 2.870 Å.
2. The energy gap falls within a close range of 2.067 to 2.069 eV, showing that the studied Sn-Y@C₈₀ molecule is in considerable energy gap range for semiconductor towards the adsorption of AsH₃, NH₃, PH₃, and SbH₃. It was also noted that energy gap slightly increased upon adsorption.
3. Generally, the adsorption follows an increasing trend of: PH₃-Sn-Y@C₈₀ < AsH₃-Sn-Y@C₈₀ < SbH₃-Sn-Y@C₈₀ < NH₃-Sn-Y@C₈₀, with the greatest adsorption energy of -3.1690 eV in NH₃-Sn-Y@C₈₀, and the least energy of -0.7728 eV in PH₃-Sn-Y@C₈₀, showcasing the strongest and weakest adsorption among the studied systems.
4. All systems indicate a partial covalent bond type. Also, the λ_1/λ_3 values indicates the existence of a weak intermolecular force between the interactions.
5. Least recovery time of 182.72 s is attributed to PH₃-Sn-Y@C₈₀, depicting that the adsorption of PH₃ gas molecules on the Sn-Y@C₈₀ molecule will desorb adsorbate from the surface in no time. This is an indicative of excellent adsorption as it will make recovery effortless. A relatively moderate recovery time (within a couple of hours) was observed among other systems.

References

1. Agache, I., Sampath, V., Aguilera, J., Akdis, C. A., Akdis, M., Barry, M., & Nadeau, K. C. (2022). Climate change and global health: a call to more research and more action. *Allergy*, 77(5), 1389-1407.
2. Akman, F., Kazachenko, A. S., & Issaoui, N. (2022). DFT calculations of some important radicals used in the nitroxide-mediated polymerization and their HOMO–LUMO, natural bond orbital, and molecular electrostatic potential comparative analysis. *Polymer Science, Series B*, 64(5), 765-777.
3. Al-Taai, S. H. H., & Mohammed al-Dulaimi, W. A. (2022). Air Pollution: A Study of Its Concept, Causes, Sources and Effects. *Asian Journal of Water, Environment and Pollution*, 19(1), 17-22.
4. Bakhsh, S. A., Khan, M. A., Ahmed, F., Alshehri, M. S., Ali, H., & Ahmad, J. (2023). Enhancing IoT network security through deep learning-powered Intrusion Detection System. *Internet of Things*, 24, 100936.
5. Chiodo, S., Russo, N., & Sicilia, E. (2006). LANL2DZ basis sets recontracted in the framework of density functional theory. *The Journal of chemical physics*, 125(10).
6. Dennington, R., Keith, T. A., & Millam, J. M. (2016). GaussView, version 6.0. 16. *Semichem Inc Shawnee Mission KS*.

7. Lin, J., Li, Y., & Ke, Z. (2023). Feature Analysis in High-Dimensional Data: Structure–Activity Relationships of Lewis Acid–Transition-Metal Complex-Catalyzed H₂ Activation. *The Journal of Physical Chemistry A*, 127(20), 4375-4387.
8. Lohith, T. N., Hema, M. K., Karthik, C. S., Sandeep, S., Mallesha, L., Mallu, P., & Lokanath, N. K. (2022). N-[2-(5-bromo-2-chloro-pyrimidin-4-yl) thio]-4-methoxy-phenyl]-4-chlorobenzenesulfonamide: The existence of H-bond and halogen bond interactions assisted supramolecular architecture—A quantum chemical investigation. *Journal of Molecular Structure*, 1267, 133476.
9. Lu, J., Zhang, W., Zhao, X., & Xu, X. (2022). Comparison of the interfacial properties of native and refolded myofibrillar proteins subjected to pH-shifting. *Food Chemistry*, 380, 131734.
10. Mati, I. K., & Cockroft, S. L. (2010). Molecular balances for quantifying non-covalent interactions. *Chemical Society Reviews*, 39(11), 4195-4205.
11. Mohammadi, M. D., Abdullah, H. Y., Biskos, G., & Bhowmick, S. (2023). Adsorbing CNCl on pristine, C-, and Al-doped boron nitride nanotubes: a density functional theory study. *Computational and Theoretical Chemistry*, 1220, 113980.
12. Odey, D. O., Edet, H. O., Louis, H., Gber, T. E., Nwagu, A. D., Adalikwu, S. A., & Adeyinka, A. S. (2023). Heteroatoms (B, N, and P) doped on nickel-doped graphene for phosgene (COCl₂) adsorption: insight from theoretical calculations. *Materials Today Sustainability*, 21, 100294.
13. Po, J. H. (2023). *Synthesis and characterizations of 2D nanoparticles with controlled folding for circularly polarized optical properties* (Doctoral dissertation, Université Paris sciences et lettres).
14. Sukanya, R., Aruldas, D., Joe, I. H., & Balachandran, S. (2022). Spectroscopic and quantum chemical computation on molecular structure, AIM, ELF, RDG, NCI, and NLO activity of 4-VINYL benzoic acid: a DFT approach. *Journal of Molecular Structure*, 1253, 132273.
15. Timothy, R. A., Okon, E. E., Gber, T. E., Onwuabusim, N. A., Obianuju, N. A., Offiong, O. E., ... & Louis, H. (2023). Theoretical Investigation of Single-Atoms Encapsulated by Fullerenes (C₅₉X: X= As, Ga, Ge) as Biosensors for Uric Acid (UA). *ChemistrySelect*, 8(42), e202303442.
16. Walker, M., Harvey, A. J., Sen, A., & Dessent, C. E. (2013). Performance of M06, M06-2X, and M06-HF density functionals for conformationally flexible anionic clusters: M06 functionals perform better than B3LYP for a model system with dispersion and ionic hydrogen-bonding interactions. *The Journal of Physical Chemistry A*, 117(47), 12590-12600.
17. Zhao, Q., Xie, T., Hong, X., Zhou, Y., Fan, L., Liu, Y., & Li, J. (2022). Modification of functional properties of perilla protein isolate by high-intensity ultrasonic treatment and the stability of o/w emulsion. *Food Chemistry*, 368, 130848.

Geochemistry of the Awi sandstone, Calabar Flank, Southeastern (SE) Nigeria: Constraints on the metal enrichment, provenance, and tectonic setting.

BENJAMIN ODEY OMANG, ARIKPO TEMPLE OKAH, LATIFA TIJANI OJIKOTU, GODSWILL
EYONG ABANG and GODWIN TERWASE KAVE

Department of Geology University of Calabar, Cross River State Nigeria

Introduction

The global demand for essential metals continues to rise, driven by increasing industrialization and population growth. This necessitates the exploration of new mineral deposits and the implementation of sustainable extraction practices. This study focuses on the geochemical analysis of Awi sandstone in the Calabar Flank, Southeastern Nigeria, examining its metal enrichment, provenance, and tectonic setting. The Awi sandstone is primarily composed of sand-sized silicate grains, with quartz as the predominant mineral due to its high weathering resistance. Geochemical analysis is crucial for understanding the sediment's origin, depositional conditions, and potential metal concentration (Ekpo et al., 2012; Boboye & Okon, 2014; Goswami & Deopa, 2018). The study employs advanced geochemical techniques to characterize the sandstones and assess their mineral potential, emphasizing the importance of their mineral composition for determining the types and quantities of economically viable minerals.

Description of the Area

The study area, the Calabar Flank, is located between latitude 5°0'0"N and 5°15'0"N, and longitude 8°15'0"E to 8°30'0"E, surrounded by significant geological features including the Oban Massif and the Niger Delta (Figure 1). The Awi Formation, which comprises the Awi sandstone, is the oldest sedimentary deposit in this basin, characterized by a composition of immature arkosic sandstones and conglomerates with mudstone and shale interbeds (Nton, 1999). The Awi sandstone exhibits a thickness of approximately 50 meters and is underlain by the Basement Complex primarily composed of amphibolites. The geological history of the area indicates that the Calabar Flank formed during the Early Cretaceous rifting period, marked by fault systems that influenced sediment deposition.

Materials and Methods

Field materials utilized for this study included GPS devices, compasses, hand augers, sample bags, and markers. The laboratory work involved equipment such as mortars, pestles, sieves, and balances for sample preparation and analysis. Fifteen fresh sandstone samples were collected from various outcrops within the Awi Formation. Samples weighing between 30 and 50 kg were collected from a depth of approximately 1 meter. After drying, the samples were crushed and sieved, with the powdered samples sent to Activation Laboratories in Canada for geochemical

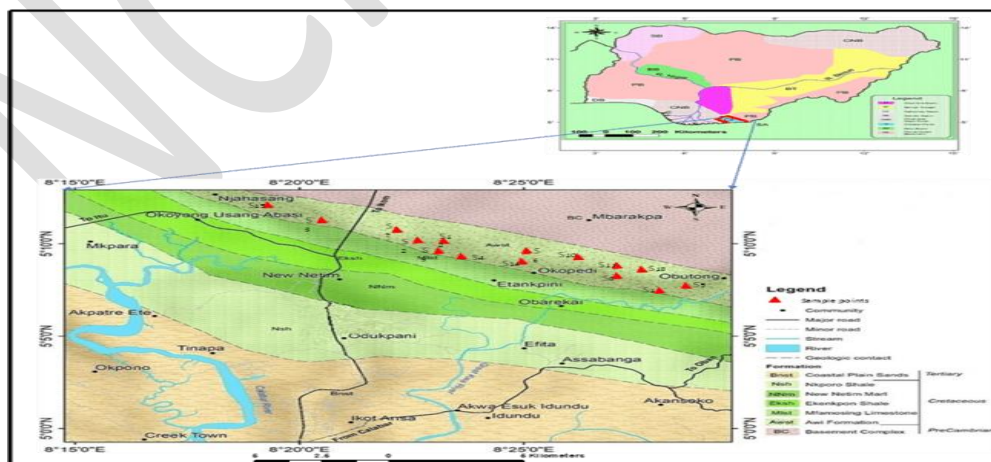


FIG. 1: Geology and location of the Study Area

analysis using inductively coupled plasma mass spectrometry (ICP-MS). Data interpretation involved various discrimination plots and statistical analyses using software like Microsoft Excel and GCDkit to assess provenance and tectonic settings.

Results and Discussion

Major Oxide Geochemistry

The geochemical analysis revealed that the Awi sandstones exhibit high levels of silicon dioxide (SiO_2), ranging from 53.24% to 80.26%, with aluminum oxide (Al_2O_3) ranging from 12.03% to 17.56%. The low concentrations of iron oxide, magnesium oxide, and calcium oxide suggest that these sandstones primarily consist of quartz and feldspar. The analysis also highlighted a negative correlation between SiO_2 and other oxides, indicating mineral crystallization processes (Manning, 2010; Hoogsteen et al., 2015)

Metal Enrichment

Trace metal analysis classified several elements as highly enriched, including barium (Ba), chromium (Cr), rubidium (Rb), and strontium (Sr) (Figure 2a,b,c, &d). The average concentrations of these elements indicate potential for hydrothermal mineralization processes, with spatial distribution maps revealing trends in metal concentration across the study area. The presence of enriched transition metals suggests contributions from mafic or ultramafic rocks.

Rare Earth Element (REE) Geochemistry

The Awi sandstones displayed a higher concentration of light rare earth elements (LREE) compared to heavy rare earth elements (HREE) (Figure 3). This distribution suggests preferential retention of HREE in heavy minerals during sedimentary processes, impacting the metal enrichment potential of the sandstone.

Provenance and Tectonic Setting

Geochemical plots, including TiO_2 versus Zr and various discrimination diagrams, indicate that the Awi sandstones are derived from intermediate to felsic igneous rocks, specifically granodiorite (Figure 4a &b). The tectonic setting is characterized by a passive to active continental margin, reflecting the complex geological history of the region involving both extensional and compressional tectonics (figure 4).

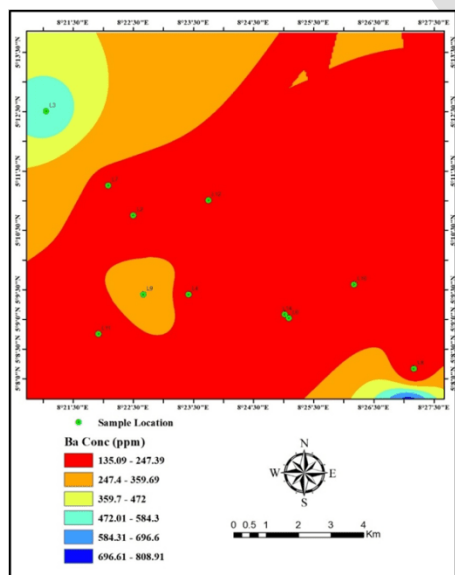


FIG. 2a Geochemical map of Ba concentration (ppm) in the study area

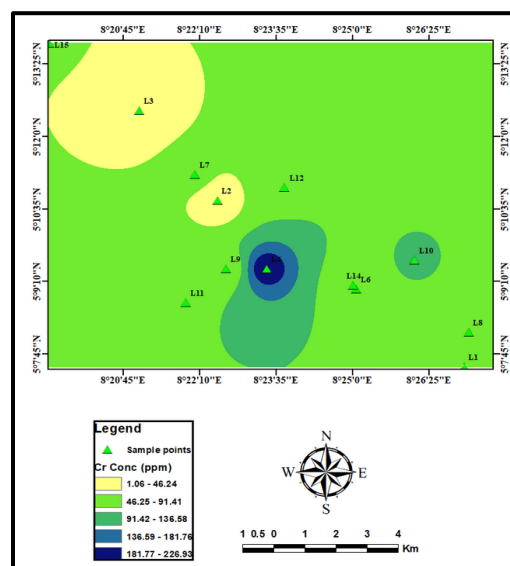


FIG. 2b Geochemical map of Cr concentration (ppm) in the study area

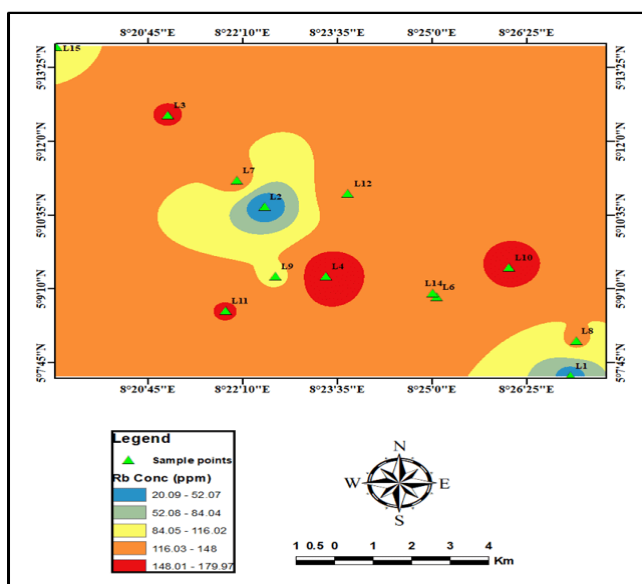


FIG. 2c Geochemical map of Rb concentration (ppm) in the study area

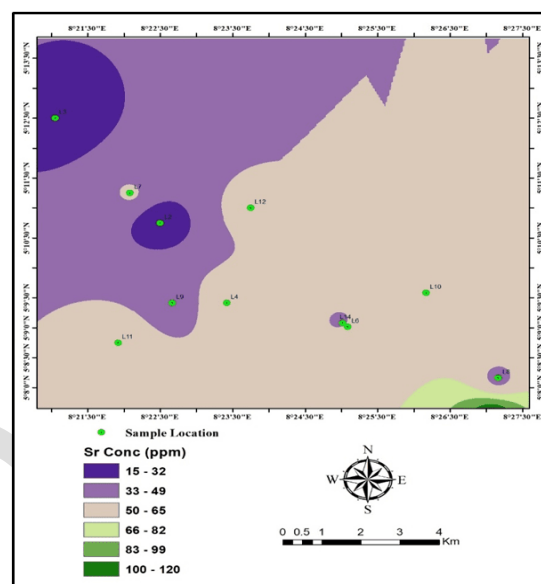


FIG. 2d Geochemical map of Sr concentration (ppm) in the study area

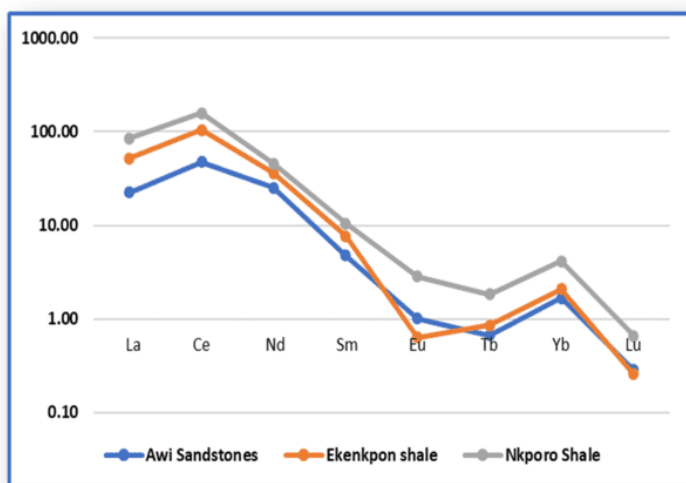


FIG. 3. Spatial distribution patterns of REEs (ppm) from Awi Sandstones compared to Nkporo and Ekenkpon Shale (Adamu *et al.*, 2020)

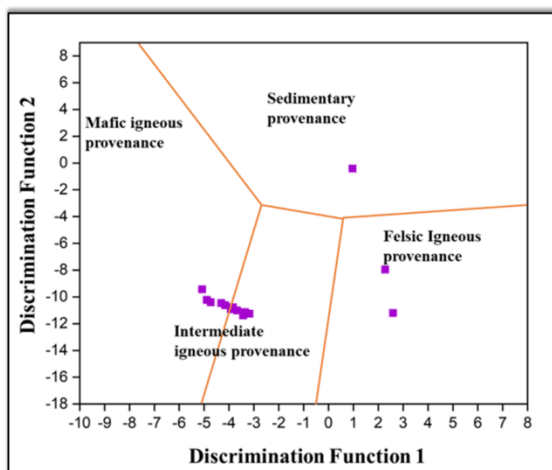


FIG. 4a. Major element Discriminant Function diagram for provenance (fields after Murali *et al.* 1983)

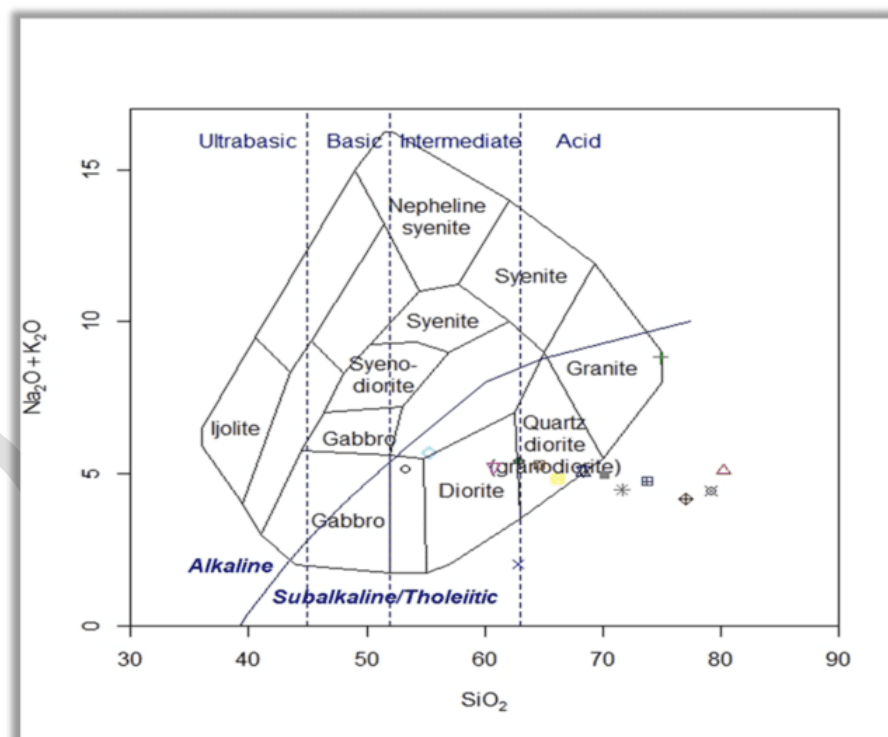


FIG. 4b. $\text{Na}_2\text{O}+\text{K}_2\text{O}$ vs SiO_2 classification plot of plutonic igneous rocks (fields after Cox, 2013)

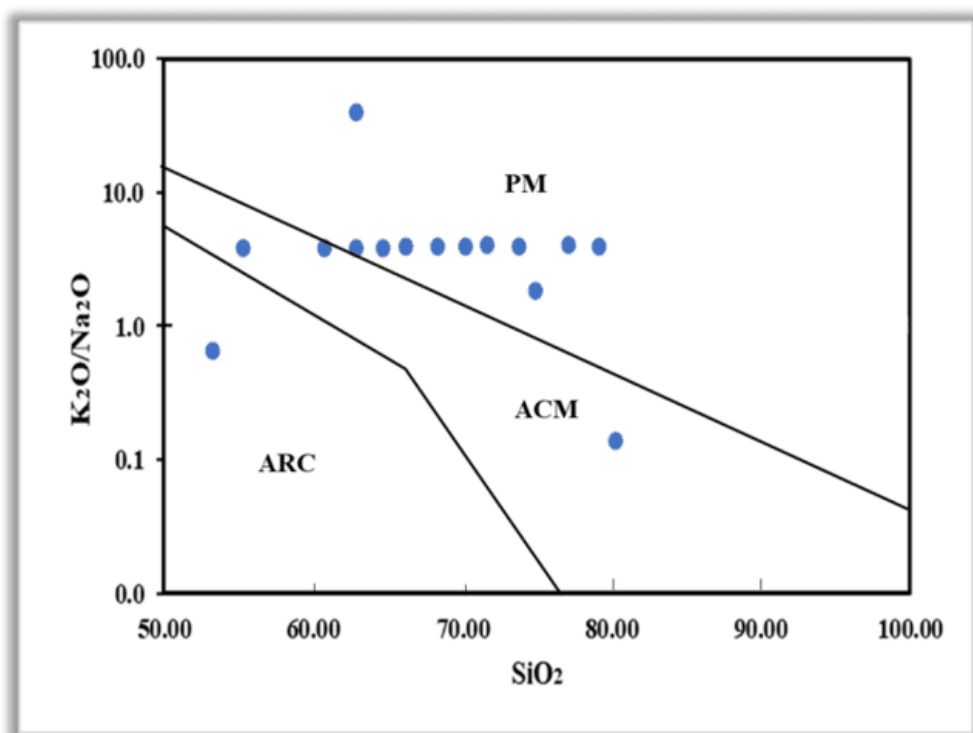


FIG. 5. Discrimination plot of log ratio of (K_2O/Na_2O) against SiO_2 (Roser and Korsch, 1988)

Conclusion

The geochemical analysis of Awi sandstone reveals significant metal enrichment, particularly in barium, rubidium, and strontium. The sandstones primarily originate from granodiorite protoliths and are associated with both passive and active continental margins. This study underscores the potential of Awi sandstone as a source of economically viable minerals and highlights the need for further exploration in the Calabar Flank.

Recommendations

It is recommended to conduct detailed geochemical studies across different stratigraphic units in the Calabar Flank to better understand the distribution and concentration of metals.

References

- Adamu, C. I., Omang, B. O., Oyetade, O. P., Johnson, O., & Nganje, T. N. (2021). Trace and rare earth element geochemistry of the black and grey shales of the Calabar flank, Southeastern Nigeria: constraints on the depositional environment and the degree of metal enrichment. *Acta Geochimica*, 40, 312-324.
- Boboye, O. A., & Okon, E. E. (2014). Sedimentological and geochemical characterization of the Cretaceous strata of Calabar Flank, southeastern Nigeria. *Journal of African Earth Sciences*, 99, 427–441. <https://doi.org/10.1016/j.jafrearsci.2014.04.035>
- Cox, K. G. (Ed.). (2013). The interpretation of igneous rocks. *Springer Science & Business Media*.
- Ekpo, B. O., Ibok, U. J., Essien, N., & Wehner, H. (2012). Geochemistry and organic petrography of Cretaceous sediments of the Calabar Flank, southeastern Nigeria. *Marine and Petroleum Geology*, 35(1), 252–268. <https://doi.org/10.1016/j.marpetgeo.2012.03.010>
- Ekwok, S. E., Akpan, A. E., Kudamnya, E. A., & Ebong, E. D. (2020). Assessment of groundwater potential using geophysical data: A case study in parts of Cross River State, south-eastern Nigeria. *Applied Water Science*, 10(6), 144. <https://doi.org/10.1007/s13201-020-01224-0>
- Goswami, P. K., & Deopa, T. (2018). Petrotectonic setting of the provenance of Lower Siwalik sandstones of the Himalayan foreland basin, southeastern Kumaun Himalaya, India. *Island Arc*, 27(2), e12242. <https://doi.org/10.1111/iar.12242>
- Hoogsteen, M. J., Lantinga, E. A., Bakker, E. J., Groot, J. C., & Tittonell, P. A. (2015). Estimating soil organic carbon through loss on ignition: effects of ignition conditions and structural water loss. *European Journal of Soil Science*, 66(2), 320-328.
- Manning, D. A. (2010). Mineral sources of potassium for plant nutrition. A review. *Agronomy for Sustainable Development*, 30, 281-294.
- Murali, A. V., Parthasarathy, R., Mahadevan, T. M., & Das, M. S. (1983). Trace element characteristics, REE patterns and partition coefficients of zircons from different geological environments—a case study on Indian zircons. *Geochimica et Cosmochimica Acta*, 47(11), 2047-2052.
- Nton, M. E. (1999). Sedimentology and depositional environment of Awi formation Calabar flank, southeastern Nigeria.
- Roser, B. P., & Korsch, R. J. (1988). Provenance signatures of sandstone-mudstone suites determined using discriminant function analysis of major-element data. *Chemical geology*, 67(1-2), 119-139.
- Sylvester, P. J. (1989). Post-collisional alkaline granites. *The Journal of Geology*, 97(3), 261-280.

Compositional Signatures of Gold from two Pan-African settings: Igarra Schist Belt, Southwest Nigeria and Lom Series Schist Belt, East Cameroon.

Godwin Terwase Kave and Benjamin Odey Omang

Department of Geology University of Calabar, Cross River State, Nigeria

ABSTRACT

This study evaluates morphology, surface features, mineral inclusions, microbial weathering, and alloy compositions of panned alluvial gold grains recovered from (1) the Uneme-Dangbala drainage systems, and (2) the Lom river basin, both located within the Pan-African metasedimentary belts that define the mobile belt zone between the West African Craton and the Congo Craton, to infer their source rocks and the evolution of the gold grains in their depositional surficial environment. The Lom grains lack inclusions and they possess secondary gold precipitates, as well as evidence of amalgamation and bacterial growth in crevices. Gold grains from the Igarra schist belt (Uneme-Dangbala drainage) contain Bi-Te-bearing mineral inclusions such as hedleyite, sulphotsumoite, and aleskite with variable contents of Pb, S, and Au. Their well-defined shapes and sizes preclude a formation through exsolution from a gold alloy. Compositionally, the inclusions vary from Bi-Te-Pb-S, Pb-S-Bi-Te, Bi-Te-Au-Pb-S, Pb-S-Bi, Pb-S, Bi-Te, Bi-Au to Au-Bi. The common occurrence of Bi-tellurides and Bi-in-gold particles provides evidence for gold formation through (1) a Bi-Te-collector process, and (2) Au scavenged by liquid Bi-bearing phases. The characteristic Bi-Pb-Te-S cluster revealed by the inclusions indicates an oxidized calc-alkaline magmatic-hydrothermal origin. Thus, granitoids that intrude the metasediments or schists are the potential source of the hypogene gold mineralization. Gold from both schist belts, although derived from granitoids, has a different composition. The detrital gold particles from both areas correspond primarily to an Au-Ag alloy, although grains from the Lom basin reveal traces of Cu. They show gold finenesses that range from 684 to 995 (av. 884) in the Igarra schist belt and from 861 to 1000 (av. 925) in the Lom basin. This average fineness is lower than that of gold from studied orogenic deposits.

Keywords: Lom, Igarra, Fineness, Craton, Metasedimentary

1. Introduction

The Igarra (SW-Nigeria) and the Lom Series (E-Cameroon) schist belts are essentially Pan-African in age (Toteu et al., 2006) and have been the objects of sustained primary gold exploration for over a decade now. Primary exploration models have traditionally targeted the aureoles of the granitic bodies that intrude the country rocks in this area. Artisanal gold exploitation is rampant in both belts, although efforts to identify a primary target have not yet been successful. This has inspired this research study in which the nature/geochemistry of gold grain compositions across this vast mobile belt is in the focus in the hope of finding subtle but informative differences that might warrant re-orientation of the current exploration plan. Common gold exploration strategies usually involve the determination of the gold particle compositions and the mineral inclusions in them (Omang et al., 2015).

2. Geology

2.1 Igarra schist belt

The Uneme-Dangbala area is constrained within latitude $7^{\circ}20'N$ and $7^{\circ}30'N$ and longitude $6^{\circ}10'E$ and $6^{\circ}20'E$, and covers an area of about 100 sq. km. Detrital gold grains from the drainage systems were recovered from a tributary draining a catchment dominated by rocks of the so-called Igarra schist belt. The Igarra schist belt is part of the southwestern corner of the Nigerian basement complex which is part of the Neoproterozoic (650 ± 200 Ma) Trans-Saharan Pan-African mobile belt (Caby, 1988). The N-S/NNW-SSE oriented Igarra schist belt is one of over thirteen schist belts that make up the Trans-Saharan Pan-African mobile belt (Dada, 2008).

2.2 Lom Series schist belt

The Lom Series - Longitude. $14^{\circ}22'E$ and $14^{\circ}44'E$ and Latitude. $6^{\circ}9'N$ and $6^{\circ}29'N$. Gold grains from the Lom basin were recovered from a tributary draining a catchment dominated by rocks of the Lom Series (Fig. 2c, d). The Lom Series is part of the Central Cameroon Shear Zone (CCSZ), a SW extension of the Central African Shear Zone (Toteu et al., 2001, 2006). The basin is located within the North Equatorial Central African Fold Belt at the northern edge of the Congo Craton. The fold belt is composed of Palaeoproterozoic basement that can be traced from the Atlantic through Cameroon into NE-Brazil and the Trans-Sahara Belt.

3. Materials and Methods

Detrital gold grains used in this study were recovered by panning of stream sediments collected above stream confluences from the Uneme-Dangbala drainage systems and the Lom basin. The heavy mineral fraction was then air-dried and the gold grains subsequently handpicked under a binocular microscope.

3.2 Morphology, surface texture and microchemical characterization

The gold grains recovered by handpicking under the binocular microscope from both sites were divided into two sets and were embedded in epoxy resin, polished using standard procedures, and viewed under reflected polarized light for their morphology at the Technische Universität Clausthal (TUC; Germany), and at the University of Queensland Australia at the University of Queensland (Australia) using a JEOL JSM-7100F field emission scanning electron microscope (SEM). Also, Electron Micro-Probe Analysis (EMPA) was engaged in order to characterize the gold grade variations in the locality, as well as identify the mineral associations, morphology, surface features, microbial weathering and alloy composition of the placer gold grains. The following elements Ag, As, Au, Bi, Cu, Cd, Fe, Hg, Pb, Pd, Pt, S, Sb, Se, Sn, Te, and Zn were checked with the Inductively Coupled Mass Spectrometer (ICP-MS), although only Ag and Au were detected in the gold matrix from the Igarra schist belt and Au, Ag, and Cu from the Lom Series schist belt. The fineness for Au was calculated using the formula $Au*1000/Ag + Au$

4. Results

4.1 Gold grain morphology and surface textures

Gold particles from the Igarra schist belt reveal a wide variation in shape and size. The gold grains vary in shape from irregular, elongated, funnel to sub-rounded with characteristic smooth edges (Fig 1.). The majority of the gold grains has sizes $<200 \mu m$. A characteristic continuous and well-developed core-rim zonation is discernable by some grains. Ag-depleted zones along solution fissures within gold grains are common. The surfaces of the grains have cavities and etch pits. Gold grains recovered from streams draining the Lom Series schist belt vary in morphology from anhedral, irregular, oblong to elongated. They are generally $<200 \mu m$. The grains show variable textures including indented surfaces, etched pits, fractures, scratches, cavities and rounded to micro-folded rims. Many grains show an accumulation of secondary gold. Rod-shaped bacteria and filamentous microorganisms.

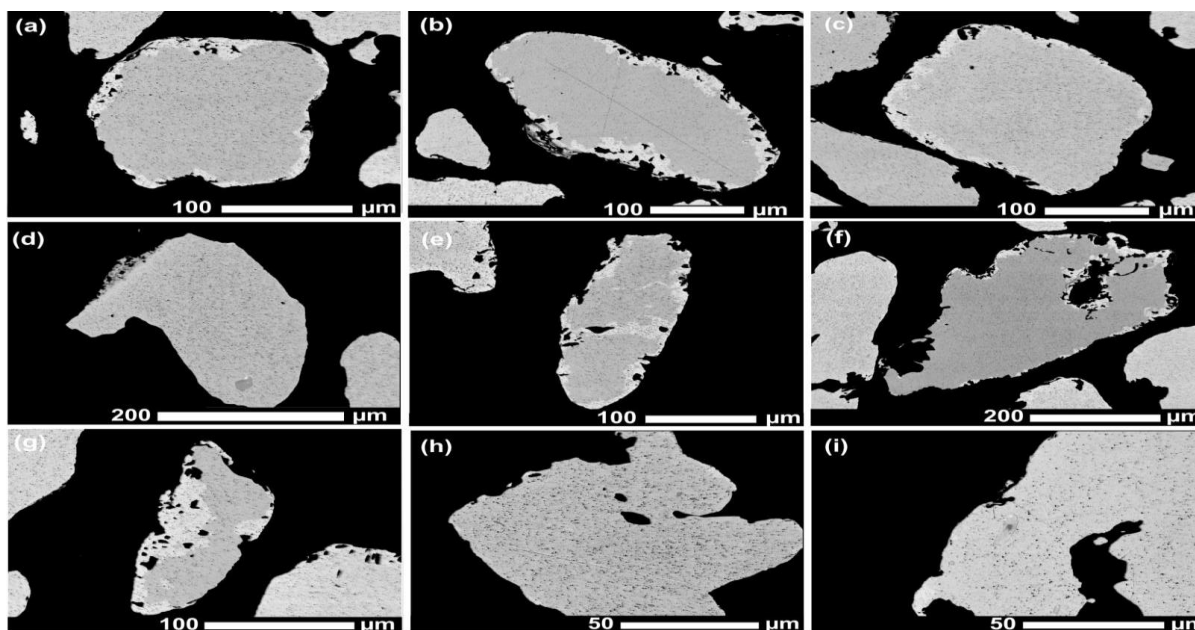


Fig. 1. BSE images of representative gold grains from the Umene-Dangbala drainage systems, Igarraschist belt, showing variations in morphology.

4.2 Characterization of inclusion signatures

The inclusions vary in shape from spherical to elongated and reach a maximum size of ~20 µm. While some of the inclusions are intact, some are deformed. Some of the inclusions show distinct replacement textures. Compositionally, the mineral inclusions are composed of different Bi-Te-bearing minerals such as hedleyite, sulphotsumoite (largely tetradymite group), and alesskite with variable contents of Pb, S, and Au-S-Bi-Te,

4.3 Gold microchemistry

Gold grains from the Umene-Dangbala drainage systems are alloyed mainly with Ag with concentrations that vary between ~0.46 and 3.43 wt% in the rim and from ~13.22 to 31.59 wt% in the core. Correspondingly, the Au concentrations range between ~68.24 and 87.37 wt% in the core and between ~96.81 and 100 wt% (pure gold) in the rims. The gold grains show an average purity value of 884 with core fineness that ranges from 684 to 867 (and rim fineness that varies between 966 and 995). Gold grains from the Lom Series schist belt are mainly alloyed with Ag (≤ 13.64 wt%), although some grains show minor amounts of Cu with contents that are generally <0.03 wt%, between ~86.09 and 99.12 wt% Table 1. The purity of the gold grains varies from 862 to 1000 with an average of 925.

5. Discussion

5.1 Origin of inclusions

Mineral inclusions within gold grains/particles are important tools used to predict the hypogene depositional conditions (Chapman et al., 2022; Omang et al., 2015).

In this study the Lom gold grains are devoid of inclusions while gold grains from the Igarraschist belt reveal contained inclusions

. According to Chapman et al. (2021), inclusion suites reflect the ore mineral assemblage(s) which underlines their use as indicators of the style of the primary mineralization of the detrital gold grains. The signature of gold derived from magmatic-hydrothermal systems reveals a Bi-Pb-Te-S signature.

In both areas, Ag is ubiquitous as an alloy component. In the Igarraschist belt the concentration of Ag is up to 3.42 wt% in the rims and 31.59 wt% in the cores. This is associated with maximum Au contents of 87.37 wt% in the cores and pure Au rims. Combined, this results in an average gold fineness of 884 which is lower than the averages for gold fineness from studied orogenic deposits. Gold grains from the Lom basin show Ag contents that reach a maximum of 13.87 wt% in the core and 2.19 wt% in the rims, which corresponds to finenesses reaching a maximum

of 997 along solution fissures in the cores and 1000 in the rims. The average gold fineness is 925 which is similar to that of gold formed in orogenic settings.

5.2 Post depositional modification

Gold grain morphology, surface texture, Ag content, and crystal imprints inherited in the surficial environment have been used worldwide to determine the mechanism of transportation, nature, and location of primary gold mineralization sources as well as modification inherited by the gold grains during the process of transportation (e.g., Alam et al., 2019).

5.3 Secondary gold precipitation

In this study, SEM images of the surfaces of gold grains recovered from streams draining the Lom Series schist belt are characterized by their bacteriomorphic structures formed by amalgamation and bioaccumulation of secondary precipitated gold particles.

6. Conclusions

The following conclusions can be drawn from this study.

1. Gold grains recovered from both areas represent single-source native gold grains with gold grains that lack evidence of epigenetic origin in the secondary environment.
2. The individual native gold grains from the Igarra schist belt have Bi-Te-bearing minerals such as hedleyite, sulphotsumoite and alesskite with varying proportions of Pb, S, and Au. The characteristic Bi-Pb-Te-S signature of the inclusions indicates an oxidized calc-alkaline magmatic-hydrothermal system for primary mineralization with the granitoids as the potential source of gold mineralization.
3. Gold grains from both areas are alloyed primarily with Ag, although grains from the Lom Series schist belts show minor quantities of Cu. The lack of an internal compositional variation in the core composition suggests that the gold grains formed during a single gold precipitation event.
4. Bacteria metabolic activities observed on the surfaces, pits and cavities of gold grains from the Lom basin are responsible for secondary gold precipitation and enrichment processes.

6.2 Recommendations

1. Subsequent exploration studies should include sampling of surface outcrops to complement stream sediments.
2. Regional and detailed geochemical surveys including sampling of more stream sediments, soil, rock, water, plants, etc. to identify possible targets and geochemical patterns associated with the mineralized zones, and resource estimation.

REFERENCES

- Alam, M., Li, S.R., Santosh, M., Yuan, M.W., 2019. Morphology and chemistry of placer gold in the Bagrote and Daintier streams, northern Pakistan: implications for provenance and exploration. *Geol. J.* 54, 1672–1687.
- Chapman, R.J., Banks, D., Styles, M.T., Walshaw, R.D., Sandra Piazzolo, S., Morgan, D.J., Grimshaw, M.R., Spence-Jones, C.P., Matthews, T.J., Borovinskaya, O., 2021. Chemical and physical heterogeneity within native gold: implications for the design of gold particle studies. *Miner. Depos.* 56, 1563–1588.
- Chapman, R.J., Mortensen, J.K., Allan, M.M., Walshaw, R.D., Bond, J., MacWilliam, K., 2022. A New Approach to Characterizing Deposit Type Using Mineral Inclusion Assemblages in Gold Particles. *Econ. Geol.* 117, 361–381.
- Dada, S.S., 2008. Proterozoic evolution of the Nigeria–Borema province. In *West Gondwana: Pre-Cenozoic Correlations Across the South Atlantic Region*; Pankhurst, R.J., Trouw, R.A.J., BritoNeves, B.B., De Wit, M.J., Eds.; Geological Society of London, Special Publications: London, UK, 2008; Volume 294, pp. 121–136.
- Omang, B.O., Suh, C.E., Lehmann, B., Vishiti, A., Chombong, N.N., Fon, A.N., Egbe, J.A., Shemang, E.M., 2015. Microchemical signature of alluvial gold from two contrasting terrains in Cameroon. *J. Afr. Earth Sci.* 112, 1–4.
- Toteu, S.F., Penaye, J., Deloule, E., Van Schmus, W.R., Tchameni, R., 2006. Diachronous evolution of volcano-sedimentary basins north of the Congo craton: insights from U-Pb ion microprobe dating of zircons from the Poli, Lom and Yaoundé Groups (Cameroon). *J. Afr. Earth Sci.* 44, 428–442.

Assessment of Groundwater Suitability for Irrigation in Uburu and its Environs, SE Nigeria

NNODIM, Chigozie J., KUDAMNYA Ebenezer A., OCHELEBE Ibu, EYONG Godswill A., AKAMBENDE Emmanuel A.

Department of Geology, University of Calabar, Calabar - Nigeria

Email of corresponding author: obeydelaw2013@gmail.com

ABSTRACT

The study was carried out within and around Uburu in Ebonyi State, southeast Nigeria, to assess the suitability of groundwater for irrigation purposes using irrigation water quality indices such as Sodium Absorption Ratio (SAR), Kelly Ratio (KR), Magnesium Hazard (MH), Total Hardness (TH), Permeability index (PI), Percentage Sodium (% Na), Residual sodium Carbonate (RSC) and Potential Salinity (PS). A total of twenty (26) groundwater samples were collected from boreholes in the study area and taken to the laboratory for analysis. The concentration of the major cations (Ca^{2+} , Mg^{2+} , Na^{+} , K^{+}) and major anions (Cl^{-} , SO_4^{2-} , and HCO_3^{-}) were determined, and used to compute values for the irrigation water quality indices. Results obtained revealed that the range of values evaluated for the suitability of the groundwater in the study for agricultural purposes include; TH (45.91-595.53 mg/l), %Na (6.16-71.07 %), SAR (1.24-14.36 meq/l), RSC (-153.70-567.50 meq/l), PI (13.18-98.14 meq/l), MH (21.11-47.02 %), KR (0.05-1.67 meq/l), and PS (1.31-16.34 meq/l). Furthermore, groundwater from the study area was found to be adequate for irrigation with respect to %Na, SAR, PI, MH, KR and PS, while in-terms of RSC, majority were doubtful. A plot of %Na against conductivity showed that majority of the groundwater samples were “*excellent to good*” for irrigation, and only at two locations there were unsuitable. Similarly, the plot of SAR against Salinity Hazard showed that majority of the samples were classified within the *low-medium* salinity hazard to *low to medium* SAR, and is an indication that groundwater in the study area is good for irrigation on all types of soils. Generally, the groundwater sampled from the study area was found suitable for irrigation purposes. However, it is *hard water*, suggesting that the water contain excess chloride and sulphate salts of alkaline earth metals, as this may affects its suitability for agricultural purposes.

Keyword: Groundwater, Irrigation, Major cations, Major anions, indices

1 INTRODUCTION

The study area covers Uburu and its environs. It is located in Ohaozara Local Government Area of Ebonyi State. Geographically, the area is located between latitudes $60^{\circ} 00'N$ and $60^{\circ} 10'N$, and longitudes $70^{\circ} 42' 50''E$ and $70^{\circ} 52' 50''E$. Uburu and its environs lies within the southeastern Nigeria and are underlain by the Southern Benue Trough (Reyment 1965). The area is underlain by the Asu River Group and Ezeaku Formation, predominantly underlain by shale, sandstone, sandy shale and thin lenses of highly calcareous limestone lithologies. The area is part of the tropical hinterland climate (Illoeje, 1979), with an average monthly rainfall of about 222mm, and mean annual temperature of about $29^{\circ}C$. It is characterized by two seasonal variations, rainy and dry seasons. Irrigation is being carried out within the region for all year round vegetable production and water quality is mostly affected in dry seasons. The knowledge of irrigation water quality is critical to soil conservation and optimal agricultural

productivity (Udom et al 2019). Water for irrigation depends on the mineral constituents, as salts exceeding the permissible limit in irrigation water can affect soil permeability, soil structure as well as crop growth and production. (Badmus et al., 2020; Murty & Jha, 2011). Therefore this study was carried out to determine the suitability of groundwater for agricultural purposes in this area.



Fig 1: Map of ebonyi state with the location of the study area (Modified from Onyinye et al. 2019)

2 MATERIALS AND METHOD

A total of twenty-six (26) groundwater samples were collected from different groundwater sources. Plastic bottles were used for the collection of the water sample, and at each sampling point the bottle was rinsed thoroughly with the water to be sampled before the water was collected. Two water samples from each location was taken, one acidified and the other non- acidified was carefully collected, packaged and transported to the laboratory for analysis. The suitability of the groundwater quality for irrigation purposes were assessed using the following irrigation indices: total hardness (TH), sodium adsorption ratio (SAR), residual sodium carbonates (RSCs), magnesium hazard (MH), percent sodium (%Na), permeability index (PI), potential salinity (PS) and Kelly's ratio (KR). These irrigation chemical parameters were computed using the following equations and taking the ions in milliequivalents per liter (meq/L), while the units of TH are expressed in mg/l.

TH: Total hardness was determined according to Sawyer and McCarthy (1967) by the following equation: $TH \text{ as } CaCO_3, \frac{mg}{l} = 2.5Ca^{2+} + 4.1Mg^{2+}$

%Na: The amount of sodium expressed in percentage, were computed using the equation after Todd (1980) as follows: $\%Na = \frac{Na^{+} + K^{+}}{Ca^{2+} + Mg^{2+} + Na^{+} + K^{+}} \times 100$

SAR: SAR is used to estimate alkali hazards in irrigation water, and it is associated with the absorption of Na^{+} by soil. The SAR was computed using the equation according to Richards (1954) as follows: $SAR = \frac{Na^{+}}{\sqrt{\frac{Ca^{2+} + Mg^{2+}}{2}}}$

RSC: Eaton (1950) developed an equation to quantify RSC in water with high HCO_3^- , because they tend to precipitate as carbonates of Ca^{2+} and Mg^{2+} . The equation is as follows: $\text{RSC} = (\text{HCO}_3^- + \text{CO}_3^{2-}) - (\text{Ca}^{2+} + \text{Mg}^{2+})$

PI: Doneen (1964) developed the concept of PI which is used to evaluate water quality for irrigation purposes. It is computed using: $\text{PI} = \frac{\text{Na}^+ + \sqrt{\text{HCO}_3^-}}{\text{Ca}^{2+} + \text{Mg}^{2+} + \text{Na}^+} \times 100$

MH: Excess Mg^{2+} and Ca^{2+} will adversely impact the soil by making it more alkaline, thus decreasing crop output. This can be estimated in terms of the MH by the following equation: $\text{MH} = \frac{\text{Mg}^{2+}}{\text{Ca}^{2+} + \text{Mg}^{2+}} \times 100$

PS: Doneen(1961) developed potential salinity. Groundwater Potential Salinity is estimated based on the concentration of Cl^- and SO_4^{2-} , it is measured in Meq/L. The potential salinity for the present study was computed using the equation below: $\text{PS} = \text{Cl}^- + \frac{1}{2} \text{SO}_4^{2-}$

KR: Kelly ratio as suggested by Kelley (1963), is computed using the equation: $\text{KR} = \frac{\text{Na}^+}{\text{Ca}^{2+} + \text{Mg}^{2+}}$

3 RESULTS

The result of the major ions (Na^+ , Ca^{2+} , Mg^{2+} and K^+ , Cl^- , HCO_3^- , SO_4^{2-} , NO_3^- and PO_4^{3-}) were obtained from the laboratory analysis.

4 DISCUSSIONS

The groundwater suitability in the area is summarized in table 1. The results of the calculated irrigation indices reveals that, TH varies between 45.91-595.53 mg/l, %Na has a value that ranges from 6.16-71.07 %. SAR values are in the range of 1.24-14.36 meq/l, RSC values range from -153.70-567.50 meq/l, PI has the value of the range 13.18-98.14 meq/l, MH value range is from 21.11-47.02%, while KR and PS values ranges from 0.05-1.67 meq/l and 1.31-16.34 meq/l respectively. For TH, based on their range of values for groundwater investigated showed that 11.50% of the water sample were classified as soft water, only 3.80% were moderately hard, were 61.50% were found to be hard, while 23.10% is very hard. The outcome suggests that water resources from this study are predominantly hard water. %Na reveals that 23.10% of the water samples were excellent, 61.50% were good, and 11.50% were permissible, while 3.80% were doubtful for irrigation purposes. A plot of %Na against conductivity (fig 3), showed that majority of the groundwater samples were “*excellent to good*” for irrigation. 25 groundwater samples falls below <10 as shown in table 1 for SAR which indicates excellent water quality, a plot of SAR against Salinity Hazard in fig 2, shows that majority of the groundwater samples were within the “*low-medium salinity hazard to low to medium*” SAR, and is an indication that groundwater in the study area is good for irrigation on all types of soils. MH values obtained showed that 100% of the water sampled is suitable for irrigation purposes. The result of the study in table (1) reveals that 96.20% of groundwater is suitable for agriculture based on Kelly Ratio. The PS of groundwater sample obtained from the study area was classified as excellent to good for irrigation purposes.

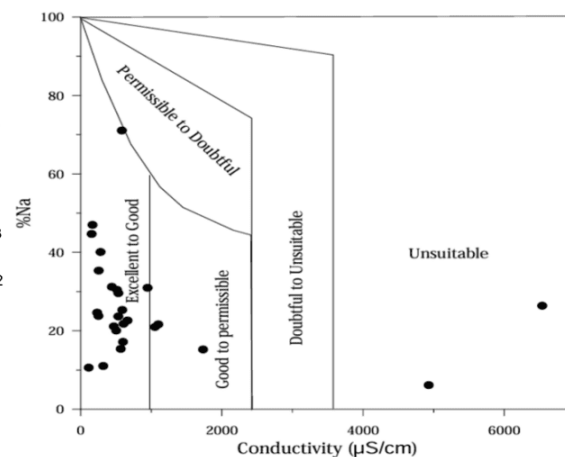
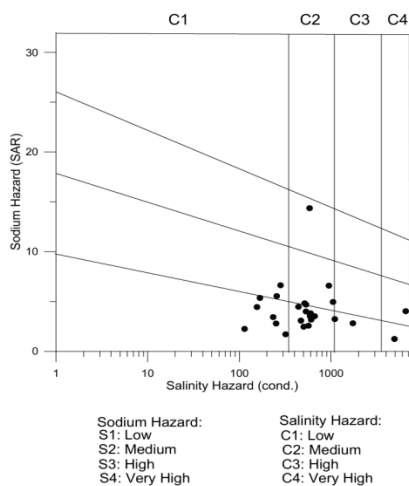


Fig 2: Plot of SAR Versus Salinity Hazard (after USSS 1954)

Fig 3: plot of %Na versus EC after Wilcox (1955)

TABLE 1: Irrigation indices assessment for water within the study area

sources	Range	indices	Class	No. of sample	Percentage%
Sawyer and McCarthy (1967)	<75	Total hardness (TH)	Soft	3	11.50
	75-150		Moderately hard	1	3.80
	150-300		Hard	16	61.50
	>300		Very hard	6	23.10
Todd (1980)	<20	Percent sodium (%Na)	Excellent	6	23.10
	20-40		Good	16	61.50
	40-60		Permissible	3	11.50
	60-80		Doubtful	1	3.80
	>80		Unsuitable	0	0.00
Richard (1954)	<10	Sodium adsorption ratio (SAR)	Excellent	25	96.20
	10-18		Good	1	3.80
	18-26		Doubtful	0	0.00
	>26		Unsuitable	0	0.00
Eaton(1950)	<1.25	Residual sodium carbonate(RSC)	Good	11	42.30
	1.25-2.5		Doubtful	3	11.50
	>2.5		Unsuitable	12	46.20
Doneen(1964)	<25	permeability index(PI)	Not suitable	6	23.10
	25-75		Moderate	19	73.10
	>75		suitable	1	3.80
Ravikumar et al. (2011)	<50	Magnesium hazard(MH)	Suitable	26	100
	>50		unsuitable	0	0.00
Kelley (1963)	<1	Kelly ratio (KR)	Suitable	25	96.20
	>1		Unsuitable	1	3.80
Doneen(1961)	<5	Potential salinity (PS)	Excellent to good	9	34.60
	5-10		Good to injurious	13	50.00
	>10		Injurious to unsatisfactory	4	15.40

5: CONCLUSION

The groundwater suitability for irrigation purposes was evaluated in this present study using the Irrigation water Quality Indices such as Sodium Adsorption Ratio (SAR), Percentage Sodium (%Na), Permeability Index (PI), Kelly Ratio (KR), Magnesium Hazard (MH), Total Hardness (TH), Residual Sodium carbonate (RSC) and Potential Salinity (PS). The values of the indices obtained were compared against known standards in order to categorize the samples based on these standards. It was observed that almost all the indices had most of the values within the recommended range suggesting that the condition of water from the study area is generally good and can be considered fit for irrigation purposes. Though majority of RSC, were unsuitable and TH reveals hard water.

REFERENCES

1. Badmus, G. O., Akinyemi, O. D., Gbadebo, A. M., & Oyedepo, J. A. (2020). Hydrochemical Analysis of Groundwater Quality along the Coastal Aquifers in part of Ogun Waterside, Ogun State, Southwestern Nigeria. *Heliyon*, 6, 1-12. <https://doi.org/10.1016/j.heliyon.2020.e05661>
2. Doneen, L.D. (1964) Notes on water quality in agriculture. Published as a water science and engineering paper 4001. Department of Water Sciences and Engineering, University of California
3. Donnen, L. D., (1961) The Influence of Crops and Soil on Percolating Water. Proceeding of Ground Water Recharge Conference, California, U.S.A
4. Eaton, F.M. (1950) Significance of carbonates irrigation waters. *Soil Science* 39:123–133
5. Iloje, N.P. (1979) A New Geography of West Africa, Longman Group Limited, Nigeria.
6. Kelly, W.P. (1963) Use of saline irrigation water. *Soil Science* 95:355–391. <https://doi.org/10.1097/00010694-196306000-00003>
7. Murty, V. N., & Jha, M. K. (2011). Land and Water Management Engineering. 6th edition. KALYANI Publishers.
8. Onyinye P. C. Laura S.O. Robert U.O. Stella N.C. Lenis S.O. Liverpool-Tasie, & Louie Rivers (2019) A Resilience Approach to Community-Scale Climate Adaptation, Sustainability 2019, 11, 3100; doi:10.3390/su11113100
9. Reyment R. A (1965) Aspects of Geology of Nigeria. *Ibadan university press*, p 145
10. Sawyer, G.N, & McCarthy, D.L. (1967) Chemistry of sanitary engineers, second edition. McGraw Hill, New York
11. Todd, D. K (1980). Groundwater Hydrology. second edition. John Willey and sons. New York.
12. UDOM, E.A., EKPO, A.E., AND OLUKA, S.I. (2019) Assessment of Irrigation Water Qualities of Abak River, Abak, AkwaIbom state, Nigeria. *Global journal of pure and applied sciences* vol. 25, 2019: 135-14. DOI: <https://dx.doi.org/10.4314/gjpas.v25i2.2>
13. USSL (1954) Diagnosis and improvement of saline and alkali soil. USDA handbook no. 60, Washington
14. Wilcox LV (1955) Classification and use of irrigation waters. US Department of Agriculture Circular, p 969

ASSESSMENT OF GROUNDWATER POTENTIALS USING DAR-ZARROUK PARAMETERS IN PARTS OF AKAMKPA, SOUTHEAST NIGERIA

¹Akambende E. A., ¹Asinya E. A., ¹Kudamnya E. A., ¹Okon, E. E., ²Ochelebe I. ¹Nnodim C. J.

1. Department of Geology, University of Calabar, Nigeria; 2. National Bureau of Statistics, Calabar

Corresponding email: obeydelaw2013@gmail.com

Abstract

An investigation of the groundwater potential of Akamkpa Local Government Area of Cross River State Nigeria, was evaluated using Vertical Electrical Sounding (VES). This study is aimed at determining the electrical resistivity of the subsurface formations, evaluation of the Dar-Zarrouk parameters to examine the aquifer protective capacity of the potential layers and finally delineates groundwater potential zones of the area. Thirty (30) vertical electrical sounding was conducted with an average maximum spacing of 150m. Field data was acquired using the ABEM SAS 1000 terrameter and modeled using the Interpex1XD software. The result, reveal majorly a three to six geo-electrical earth model of the subsurface formations exception of VES location one which shows a seven earth layer model. The fifth geo-electrical model being the most dominant occupying a total of twelve(12) VES location points followed by the sixth layer and covering a total of 10 VES points. The groundwater potential zones were assessed based on; longitudinal conductance (S_L), transverse resistance (T_r), coefficient of anisotropy (λ), formation resistivity (ρ_m), reflection coefficient (R_c) and Resistivity contrast (F_c). Resistivity values of the aquifer zones ranges from 32.052 Ω m (VES 12) to 1034.7 Ω m (VES 5). Depth to the aquifer ranges from 0.74m (VES 8) to 124.22m (VES 7). The study results and the formations evaluated shows good prospect for groundwater localization and they are hereby recommended for groundwater development.

Keywords: Aquifer, Geo-electrical models; Groundwater potential; Resistivity; Subsurface formations

Introduction:

Groundwater is one of the most valued and widespread natural earth resource and has a wide range of uses including; domestic, agricultural and industrial application (Pathak, Awasthi, Shaema, Hadaha and Nema, 2018; Diancoumba, 2020). Due to the complexities inherent within the basement terrain, groundwater may not be available everywhere in the subsurface. Hence, detailed geo-electrical survey are necessary to identify potential occurrences of groundwater. Elsewhere, several authors have used hydrogeological methods (Ifedigbo and Aghomiche, 2012; Edet and Okereke 2022); geo-electrical techniques (Edet, 1990; Edet, Kudamnya and Ekwere, 2023) and remote sensing methods (Kudamnya, Sylvanus, Essien, Adamu and Omang, 2019; Machiwal, Jha and Mal, 2011; Bera and Bandyopadhyay, 2012), for groundwater prospecting within the study area. The study area lies within the Southeast Basement Complex of Nigeria and composed of rocks suits of Precambrian age and situated at about forty-five kilometers from the capital city Calabar. It is bounded between latitudes 5° 0' - 5° 40' N and longitudes 8° 10' - 8° 50' E (Figure 1). There is a constant report on poor yield and failure of some drilled borehole within the study. This research however is focused on using the Dar-Zarrouk parameters to assess the groundwater potentials of the area to determine the aquifer protective capacity based on the evaluated longitudinal conductance values S_L .

Materials and Methods

Vertical electrical sounding survey was conducted to generate the geo-electrical resistivity data from thirty (30) different location points within the study area. At each location, geographical coordinates and elevation values were taken and recorded using a hand-held Global Positioning System (GPS) with ± 5 accuracy. The situation of the locations points was done bearing in mind the space limitations due to topography and settlement patterns constraining maximum current electrode separation (AB) to be 300m. The method adopted in this study was dependent on the introduction of an artificial current source to the ground using points electrodes (Telford, Gel dart, Sheriff and Keys, 1978). By increasing the electrode spacing, there is a corresponding increase in current penetration. The electrical resistivity technique employed is the Schlumberger electrode configuration.

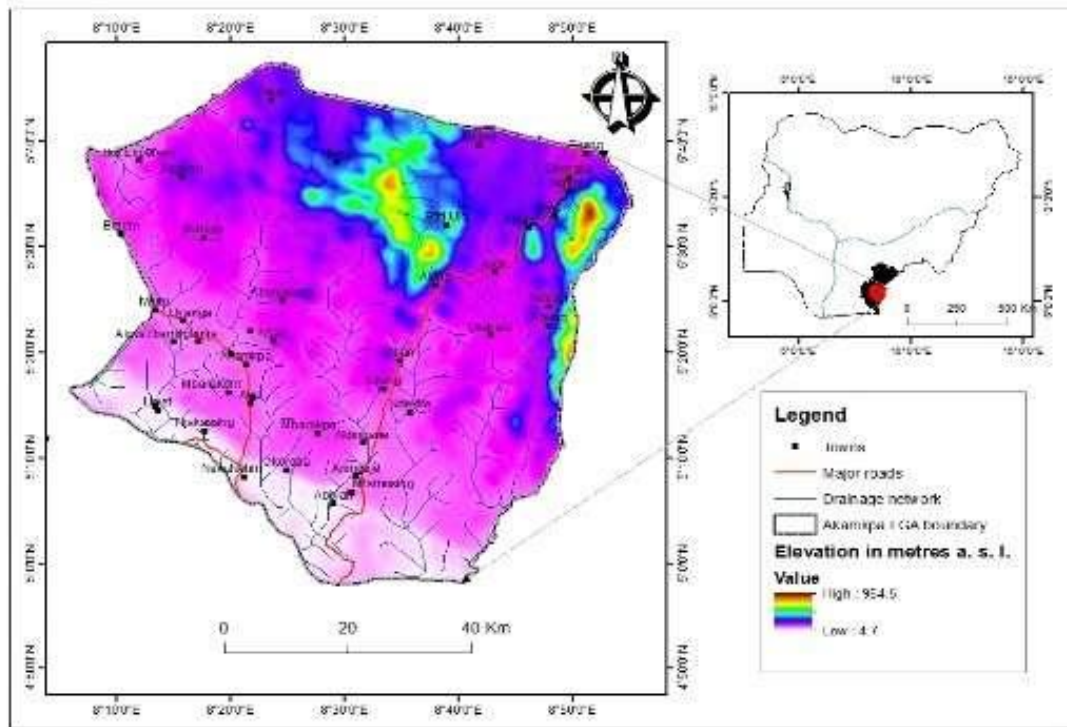


Figure 1: Map showing the location of the study area

However, the equipment used in acquiring the data was the ABEM SAS 1000 terrameter. The corresponding apparent resistivity values were computed using the formular below;

$$\rho_a = K \frac{V}{I} \dots\dots\dots 1$$

Where K, is the geometric factor ($K = \pi MN[(AB/2)^2 - 1/4]$)2

The resistivity field data generated were then plotted against half current electrode spacing on a log-log graph using a modelling software known as INTERPEX 1-XD. The ambiguity (noise) influences and lateral in-homogeneities on the field curves were removed by using the smoothening tool, and the model curves were produced for all the thirty VES points. The geo-electrical resistivity data were further evaluated using the Dar-Zarouk parameters of longitudinal conductance (S_L), transverse resistance (T_r), coefficient of anisotropy (λ), formation resistivity (ρ_m), longitudinal resistivity (P_L), transverse resistivity (P_t), reflection coefficient (R_c) and Resistivity contrast (F_c). To reflect the characteristic of subsurface layers. These parameters are computed for a known layer thickness, h and resistance p. By definition:

$$S_L = \sum_{i=1}^n \frac{h_i}{\rho_i} \dots\dots\dots 3$$

$$\text{and } T_r = \sum_{i=1}^n h_i \rho_i \dots\dots\dots 4$$

Any increment in the value of S_L between the survey stations will indicates a corresponding increase in thickness of subsurface layers encountered, a decrease in the average longitudinal resistivity P_L or both. Nevertheless,

$$P_L = \sum_{i=1}^n \frac{h_i}{S_i} \dots\dots\dots 5$$

Also average transverse resistivity P_t is given by;

$$P_I = \sum_{i=1}^n \frac{T_i}{h_i} \dots\dots\dots 6$$

As result, the will be a geo-electrical anisotropy with regards to resistivity. Coefficient of resistivity, λ is computed as;

$$\lambda = \sqrt{\frac{P_I}{P_L}} \dots\dots\dots 7$$

Where λ is real and greater than 1

Formation resistivity (ρ_m) can be defined as;

$$\rho_m = \sqrt{P_L P_I} \dots\dots\dots 8$$

The reflection coefficient (R_c) and resistivity contrast (F_c) of fresh basement rocks of the study area can be computed after the method of Oladunjoye and Jekayinfa (2015), given by:

$$R_c = \frac{p_n - p_{n-1}}{p_n + p_{n-1}} \dots\dots\dots 9$$

and

$$F_c = \frac{p_n}{p_{n-1}} \dots\dots\dots 10$$

Where, p_n = resistivity of the n^{th} layer and p_{n-1} = resistivity of the overlying n^{th} layer

Results and Discussion

The Dar-Zarrouk parameters of the potential layers were computed and presented below (Table 1). The VES results revealed majorly a three to six (3-6) geo-electric earth layer models with an exception of VES location one (1) which shows a seven (7) layers geo-electric model. From the assessment of the groundwater potential layers using the Dar-Zarrouk parameters, it is observed that, the area is majorly characterized by very low longitudinal conductance values far less than $0.1 \Omega\text{m}$ which is an indication of an area with very poor aquifer protective capacity according to the rating by Olusegun *et al.*, 2016. But VES locations 7, 11, 23 and 26, have longitudinal conductance values in the range of 0.1156 - 0.1846 being within the weak aquifer protective capacity rating and VES locations 18, 24 and 30 have values in the range of 0.2788 – 1.3684 falling within the moderate aquifer protective capacity rating. Generally, the aquifer protective capacity of the study area ranges from very poor to moderate. Resistivity values of the aquifer zones ranges from $32.052 \Omega\text{m}$ (VES 12) to $1034.7 \Omega\text{m}$ (VES 5). Depth to the aquifer ranges from 0.74m (VES 8) to 124.22m (VES 7). The study area also showed high values of transverse resistance which is an indication of high resistivity formations found within the subsurface. And this is in conformity with the area being a basement terrain.

Table 1: Evaluated parameters from the vertical electrical sounding (VES) undertaking in the study area

VES no.	P_a (Ωm)	h (m)	d (m)	S_L (Ωm)	T_r (Ωm)	R_c (Ωm)	F_c (Ωm)	P_L (Ωm)	P_I (Ωm)	λ (Ωm)	P_m (Ωm)
1	759.72	28.303	31.983	0.0373	21502.36	-0.9896	0.0052	634.8092	7792.7023	3.5037	2224.158
2	443.23	108.55	116.68	0.2449	48122.6165	-0.6643	0.2017	466.9712	553.4882	1.0887	508.3926
3	549.81	15.489	16.54	0.0282	8515.8522	0.7182	6.0969	651.5562	770.972	1.0878	708.7535
4	547.49	11.7332	15.467	0.0214	6423.8097	-0.9901	0.005	409.4114	3832.4684	3.0596	1252.6198
5	1034.7	22.279	27.988	0.0215	23052.0813	-0.7456	0.1458	1363.1649	31961.2118	4.8421	6600.6364
6	371.62	10.805	20.922	0.0291	4015.3541	0.3341	2.0033	269.7493	1320.9108	2.2129	596.921
7	568.18	65.698	124.22	0.1156	37328.2896	-0.5481	0.2919	838.6982	1292.8969	1.2416	1041.3214
8	134.25	0.50202	0.7359	0.0037	67396.185	-0.9846	0.0078	2231.0926	16208.8858	2.6954	6013.6116
9	451.23	22.481	23.44	0.0498	10144.1016	-0.9248	0.0391	922.4481	21456.6607	4.8229	4448.8937

10	881.1	72.489	74.514	0.0823	63870.0579	0.8237	10.3464	847.1602	931.4923	1.0486	888.3261
11	93.648	18.043	42.19	0.1927	1689.6909	-0.865	0.0724	199.6414	7806321806	6253.1373	1248384.842
12	32.052	1.5493	2.6339	0.0483	49.6646	-0.9948	0.0026	130.1045	1927.4381	3.8489	500.7678
13	691.1	2.8524	5.3402	0.0041	1971.2936	-0.8909	0.0577	2247.6804	11810.3076	2.2923	5152.2613
14	1364.7	28.971	29.82	0.0212	39536.7237	0.9467	36.5558	2453.5502	3171.2202	1.1369	2789.3992
15	171.5	2.039	6.4592	0.0119	349.6885	0.4326	2.5248	27.8732	4031.0435	12.0258	335.1987
16	554.72	24.15	26.286	0.0435	13396.488	-0.9553	0.0229	2095.4181	2363.9208	1.0621	2225.6241
17	886.72	15.243	21.554	0.0172	13516.2729	0.5556	3.5001	1507.2151	2163.6776	1.1981	1805.8592
18	65.53	29.75	42.263	0.454	1949.5175	-0.9456	0.0278	79.5657	760.279	3.0912	245.9514
19	1014.3	6.7245	9.7665	0.0066	6820.6604	0.995	397.952	28.5931	1032.4445	6.009	171.8161
20	687.76	5.984	11.324	0.0087	4115.5558	-0.7804	0.1233	1743.4029	7200.4231	2.0323	3543.055
21	135.69	4.5635	7.538	0.0336	619.2213	-0.9565	0.0222	188.2311	2215.2285	3.4305	645.736
22	535.8	26.467	44.778	0.0494	14181.0186	0.9614	50.8639	51.6694	431.6869	2.8905	149.3486
23	277.32	36.931	43.995	0.1332	10241.7049	-0.5708	0.2732	395.5121	7765.2149	4.4309	1752.4944
24	33.281	45.541	46.195	1.3684	1515.65	-0.2946	0.5449	33.4969	33.6748	1.0027	33.5857
25	1024.8	46.971	51.299	0.0458	48135.8808	0.4318	2.5198	908.323	972.6477	1.0348	939.9353
26	30.534	5.6366	8.9924	0.1846	172.1079	-0.9917	0.0042	46.6432	15641.159	18.3123	854.1391
27	99.511	1.8335	4.4336	0.0184	182.4534	-0.7639	0.1338	202.2472	477.2451	1.5361	310.6791
28	304.45	18.932	33.673	0.0622	5763.8474	0.9398	32.2118	33.8833	182.9364	2.3236	78.7305
29	378.99	1.6469	4.9008	0.0043	624.1586	-0.4691	0.3614	893.5042	6705.109	2.7394	2447.6607
30	86.537	19.781	45.133	0.2788	1711.7884	-0.9425	0.0296	161.8275	785.8979	2.2037	356.6229

Conclusion

The groundwater potential zones were assessed based on; longitudinal conductance (S_L), transverse resistance (T_r), coefficient of anisotropy (λ), formation resistivity (ρ_m), longitudinal resistivity (P_L), transverse resistivity (P_t), reflection coefficient (R_c) and Resistivity contrast (F_c). Resistivity values of the aquifer zones ranges from 32.052 Ω m (VES 12) to 1034.7 Ω m (VES 5). Depth to the aquifer ranges from 0.74m (VES 8) to 124.22m (VES 7). Aquifer thickness ranges from 1m (VES 8) to 109m (VES 2). The aquifer protective capacity evaluation based on longitudinal conductance values ranges from poor to moderate. The study results and the formations evaluated shows good prospect for groundwater localization and they are hereby recommended for groundwater development.

REFERENCES

- Adelana, S. M. A., Olasehinde, P. I., Bale, R. B., and Vrbka, P. (2008): An overview of geology and hydrogeology of Nigeria. In. S.M.A. Adelana, & A. M. MacDonald (Eds), *Applied groundwater studies in Africa* (IAH Selected papers on hydrogeology, pp. 171-197). London: C R C press.
- Anudu, G. K., Onuba, L. N., and Ufodu, L. S. (2011): Geoelectric sounding for groundwater exploration in the crystalline Basement Complex terrain around Onipe and adjoining areas, southwestern Nigeria. *Journal of Applied Technology in Environ Sanitation*, 1(4), 343-354.
- Ayodele, M. O., Omoboriowo, A. O. and Adiola, U. P. (2017): A geologic field report on the Geology of part of the Oban Massif and Calabar Flank, South Eastern Nigeria. *International Journal os Science*, PP 29-34.
- Badrinarayanan T. S. and Thiyagarajan G. (2018): Software applications for interpreting geoelectric data in groundwater exploration techniques. <https://www.researchgate.net/publication/327161>.
- Badrinarayanan, T. S. and Venkatesan, R. (2007): Inverse slope method of interpretation of geoelectrical sounding curves and its application for ground water exploration in crystalline rock terrains- Mineral exploration: Recent strategies. *Annamalai University, Chidambaram, Tamilnadu*.
- Chenini, I., Mammon, A. B., and El-May, M. (2010): Groundwater recharge zone mapping using GIS-based multi-criteria analysis: A case study in central Tunisia (Maknassy Basin). *Water Resources Management*, 24(5), 921-939.
- Diancoumba, O. (2020): Assessment of the Effects of Climate and Land Use/Land Cover Change on Groundwater Resources in Koda Catchment, Mali, West Africa. Doctor of Philosophy (PhD), *The University of Abomey-Calavi 9Benin Republic*.
- Edet, A. (2016): Hydrology and groundwater evaluation of shallow coastal aquifer, southern Akwa Ibom State (Nigeria). *Applied Water Science*, 7, 2397-2412.
- Edet, A. and Okereke, C. (2022): Investigation of hydrogeological condition of fractured shale aquifer in Yala Area. (SE Nigeria) characterized by saline groundwater. *Applied Water Science* <https://doi.org/10.1007/s 1321-022-01715-2>
- Edet, A. E. (1993): Hydrogeology of parts of Cross River State, Nigeria: Evidence from aero geological and surface resistivity studies. *PhD dissertation*, University of Calabar, Nigeria.
- Edet, A. E., and Okereke, C. S. (1997): Assesment of hydrogeological conditions in basement aquifers of the Precambrian Oban Massif and Obudu Plateau, southern Nigeria. *Journal of Applied Geophysics*, 36, 195-204.
- Edet, A. E., Teme, S. C., Okereke, C. S., and Esu, E. O. (1994): Lineament analysis for groundwater exploration in Precambrian Oban Massif and Obudu Plateau, SE Nigeria. *Journal of Mining and geology*, 30(1), 87-95.
- Edet, A., Kudamnya, E. A. and Ekwere, S. A. (2023): Geoelectrical and hydrogeological investigations of the University of Calabar campus (SE Nigeria): Implication of sustainable groundwater development. *Solid earth sciences* <https://doi.org/10.1016/J.sesci.2022.07.003>
- Kudamnya, E. A. and Andongma, W. T. (2017): Predictive mapping of groundwater within Sokoto Basin, North Western Nigeria. *Geography, Environment & Earth Science International*, 10(2): 1-14. <https://doi.org/10.9734/JGEEES1.2017.32440>
- Pathak, R., Awasthi, M. K., Shaema, S. K., Hadaha, M. K., and Nema, R. K. (2018): Ground water flow Modelling Using MODFLOW- A Review. *International Journal of Current Microbiology an Applied Sciences*, 7, 83-88. <https://doi.org/10.20546/ijcmas.2018.702.011>

AN INTERGRATED STUDY OF BIOGEOCHEMICAL CONTROLS ON HEAVY METAL TRANSFER FROM SOIL TO FOOD CHAIN USING CASSAVA TUBERS AROUND MAJOR ROCK QUARRYING SITES IN WESTERN OBAN MASSIF, SOUTH-EASTERN NIGERIA

A. S. Ekwere, B.B Edet and A.N. Ugbaja

Department of Geology, University of Calabar, E-mail: zerrata77@yahoo.com

ABSTRACT

An assessment of biogeochemical controls on heavy metals transfer from soil to food chain using tubers of cassava (*Manihot esculenta*) around major quarrying vicinities in Western Oban Massif was carried out. The aim of this study was to evaluate the effect of organic carbon, organic matter fractions and enzymatic activities on the distribution of heavy metals in soils and subsequent transfer into food chain (cassava tubers). Seven randomly selected sampling locations around quarrying sites including a control point distance away were investigated and 22 samples (soil and cassava tubers) were obtained in rainy and dry seasons. Soil temperature, pH and EC were made in-situ and confirmed in the laboratory. Soil and cassava samples were examined for ten (10) heavy metals; Cu, Cd, Cr, Cu, Pb, Zn, Fe, Mn, Ni, As by Flame AAS using UNICAM Model 969 and REYLEIGH-AAS using WFX320 Model, respectively. Soil moisture was determined by gravimetric method, organic carbon by the Walkley-Black density fractionation method, organic matter and fractions were determined using the method described by Sohi, et al., (2001). TDN by the nitrogen Kjeldahl method, Na_2O , K_2O , CaO , MgO , PO_4^{3-} , SO_4^{2-} , Al_2O_3 , and Cl^- were determined by Atomic Absorption spectrophotometry using Perkin Elmer 2380 model, while dehydrogenase and urease activities were determined by the method described in Von Mersi and Schinner, (1991) and the procedure for the non-buffer method by Paulson and Kurt, (1969). From assessment, in 85% of the samples, the concentration of measured heavy metals in soil and cassava samples were below the permissible limits given by FEPA (1991), WHO (2004) and SON (2002). Also, soils in the area were generally uncontaminated except Fe with severe contamination in locations OMS6 and OMS7, and there was no pollution, no ecological and potential ecological risk with respect to all measured heavy metals. Transfer factors of all measured metals were <1 , which indicates that the cassava plant was of the excluder's category with poor response towards metal absorption except in samples OMP1 and OMP4 with build-up of Mn and Ni respectively in rainy season, and in sample OMP3 with build-up of As in dry season. In soil samples from both seasons, strong and moderate correlations existed between; temperature, pH, EC, organic carbon, organic matter, organic matter fractions, enzymatic activities, and heavy metals indicating a relationship between these parameters. In factors 1-3 for both seasons, the elements that contributed to the total data variance were temperature, pH, EC, organic carbon, organic matter, organic matter fractions and heavy metals, and their sources were related to geogenic processes of weathering, minerals dissolution and anthropogenic quarrying / associated activities which affected heavy metals bioavailability suggesting a relationship between these parameters.

Keywords: Heavy Metals Transfer, metals Build-up, Bioavailability, Enzymatic Activities, Organic Matter Fractions, Food Chain and Transfer Factors.

1. Introduction

Fergusson (1990) described heavy metals as metallic elements and metalloids characterized by high atomic weight and relative high density compared to water. They occur naturally in the Earth's crust, induce toxicity even at very low exposure levels, and not easily degraded once they enter the environment (Duffus, 2002; Tchounwou et al., 2014; Lenart-Boron and Boron, 2014). Sources of heavy metals may be both geogenic including weathering, atmospheric, volcanic eruptions, and anthropogenic including domestic, industrial, agricultural, and pharmaceutical (Alloway and Jackson, 1999; Anguelov and Anguelova, 2009; Wang et al., 2016).

According to Sracek and Zeman (2004), the principal metals released from quarrying include mercury, lead, cadmium, arsenic, aluminium, chromium, zinc, selenium, and titanium, ions such as Na^+ , K^+ , Ca^{2+} , Mg^{2+} , Cl^- , oxides (SO_4^{2-} , PO_4^{2-} , Al_2O_3 , Na_2O , K_2O , CaO , MgO), and other species. The soil is a long-term reservoir and the medium for which growing plants, soil organic matter, enzymes, heavy metals, and other chemical compounds released from different human activities interacts to alter soil properties and the food chains. The bioavailability and absorption of these species in the soil is a dynamic process influenced by specific combinations of chemical,

biological, and environmental factors. (Dube et al., 2001; Pająk et al., 2016; Kuzniar *et al.*, 2018). Furthermore, soil biotic factors, including soil organic matter, humus content, and enzymatic activities, contribute to this intricate interplay which potentially establishes pathways for heavy metals infiltration into the environment, the soils and growing plants(Lasota et al., 2020; Khan et al., 2008). This study aims to evaluate; the concentration of heavy metals in soils, soils-to-plant transfer of heavy metals into the food chain (cassava tubers) and the probable control of heavy metal distribution by biogeochemical species.

1.1 Description of the study area

The study area is located between Latitudes 5° 00' and 5° 45' N and Longitudes 8° 00' and 8° 55' E, Figure 1.Geologically, the area consists of rocks such as granites, gneisses and schists with varying grades of weathering (Ekwueme, 2003; Edet and Okereke 1997), which are intruded by pegmatite, granodiorites, diorites, tonalites, monzonites, charnokites and dolerites (Ekwueme *et al.*, 1995; Ekwere and Edet, 2023) - Figure 2. Soils in the areas are mostly lateritic in nature due to weathering of acidic crystalline rocks (Abua and Eyo, 2013; Ekwere and Edet, 2021),they are loamy sand to sandy loam at the surface characterized by coarse texture, weakly-structured, leached, and with a pH range of 4.9 - 6.8. These lateritic soils are susceptible to rapid decomposition by organic matter due to high temperature and humidity which enhances bacterial activity and high organic matter content. Soils have low contents of exchangeable bases, medium organic carbon contents, rich in phosphorous and contains clay in the surface (Abua& Eyo, 2013; Abua and Edet, 2007).

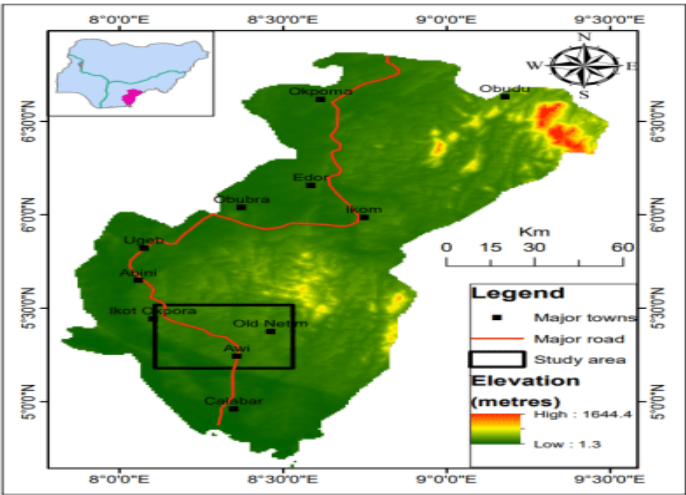


FIG. 1: Cross River State map showing the Study Area & an inset map of Nigeria. Adapted from Edet and Okereke (2007).

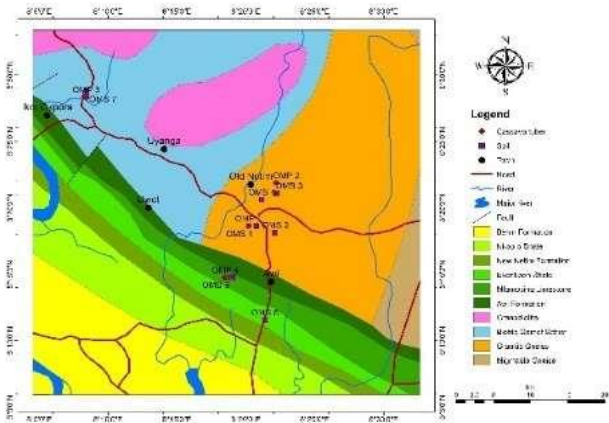


FIG. 2: Geologic map of the study area showing samples locations.

2. Materials and methodology

This research work was conducted around quarry locations in Western Oban Massif in rainy and dry season. Seven (7) sampling sites were chosen randomly (Fig. 2), and 22 samples (soil and cassava tubers) were obtained in both seasons. Ten(10) heavy metals; Pb, Zn, Cu, Cd, Cr, Co, Fe, Mn, Ni, and As were analysed by means of Flame AAS using UNICAM Model 969 and REYLEIGH AAS using model WFX320 for soil and cassava samples respectively. Moisture content was determined by Gravimetric method, organic carbon by Walkley-Black (1934) method, organic matter and its fractions by method described in Sohi et al.,(2001).Na₂O, K₂O, CaO, MgO, Al₂O₃,P₂O₅,Cl and SO₄²⁻ by AAS via Perkin Elmer 2380Model, TDN by the nitrogen Kjeldahl method, Dehydrogenase activity by the method used in Von Mersi and Schinner (1991) and Urease activity by the procedure for the non-buffer method.

3. Results and discussions

Table 1 Statistical summary of measured parameters in soils from the study area

Parameters	Wet season				Dry season				Avg. Shale Comp	WHO (2004)	FEPA (1991)
	Min.	Max.	Mean	Sdev.	Min.	Max.	Mean	Sdev.			
Temp (°C)	27.4	28.4	27.98	0.334	28.4	33.2	31.76	1.826	-	-	-
pH	5.6	6.6	6.043	0.306	4.89	6.00	5.463	0.354	-	6.5-8.5	5.5-9.5
EC (µs/cm)	44.7	108.1	70.21	19.56	33.4	142.3	75.81	31.79	-	1400	-
MC (%)	15.9	41.21	24.873	7.712	3.16	14.92	7.137	4.018	-	-	-
U A (µ/mg protein)	49.4	165.9	109.96	46.158	11.80	107.85	66.771	29.398	-	-	-
DA (µ/mg protein)	42.00	409.75	184.43	138.14	18.75	76.55	54.739	17.617	-	-	-
OC (%)	1.00	3.17	2.043	0.644	0.96	2.69	1.554	0.549	1.00	-	-
HC (%)	0.01	0.16	0.033	0.052	0.04	0.269	0.103	0.072	-	-	-
HA (%)	0.068	0.169	0.095	0.032	0.04	0.24	0.114	0.063	-	-	-
FA (%)	0.00	0.00	0.00	0.00	0.01	0.18	0.095	0.085	-	-	-
SOM (%)	2.97	9.40	6.06	1.907	1.94	5.44	3.546	1.286	-	-	-
TDN (mg/kg)	20.00	40.00	28.571	9.897	1.40	5.60	2.95	1.373	-	-	-
Al ₂ O ₃ (mg/kg)	0.01	0.22	0.071	0.076	0.02	0.22	0.141	0.063	18.50	-	-
CaO (mg/kg)	36.22	68.11	58.153	10.195	1.12	2.41	1.947	0.413	2.40	-	-
MgO (mg/kg)	18.11	43.21	29.23	7.470	2.79	8.42	5.494	1.910	2.90	-	-
K ₂ O (mg/kg)	0.01	0.51	0.192	0.159	0.01	0.02	0.014	0.005	4.00	-	-
Na ₂ O (mg/kg)	0.01	0.22	0.127	0.085	0.01	0.26	0.16	0.078	1.00	-	-
Cl ⁻ (mg/kg)	2510	7100	5891.43	1920.58	0.82	1.70	1.476	0.350	-	-	-
SO ₄ ²⁻ (mg/kg)	0.11	0.21	0.143	0.043	0.51	2.11	1.389	0.470	-	-	-
PO ₄ ³⁻ (mg/kg)	0.001	0.021	0.01	0.008	0.05	1.80	1.016	0.606	0.20	-	-
Fe	1.614	2.452	2.073	0.276	2.843	4.135	3.197	0.405	4.70	1.0-3.0	5.0
Ni	0.012	0.064	0.041	0.015	0.013	0.132	0.074	0.040	68.00	10.00	0.03
Mn	0.048	0.149	0.069	0.033	0.02	0.206	0.075	0.055	900.00	0.40	0.48

Cd	0.024	0.321	0.234	0.097	0.103	0.35	0.273	0.135	40.00	0.003	0.01
Cu	0.276	0.612	0.389	0.107	0.385	0.83	0.482	0.148	39.00	0.05-2.0	0.05
Cr	0.012	0.232	0.032	0.028	0.009	0.041	0.026	0.020	90.00	1.60	0.03
Pb	0.062	1.026	0.207	0.335	0.219	1.236	0.865	0.332	23.00	2.00	0.05
As	0.003	0.014	0.009	0.004	0.001	0.006	0.004	0.002	13.00	0.01	-
Zn	1.128	2.186	1.485	0.440	2.093	2.752	2.313	0.203	120.00	1.0-3.0	0.75
Co	0.01	0.037	0.023	0.009	0.009	0.037	0.019	0.009	19.00	0.001	0.10

Comparing all measured values of temperature, pH and EC in soil samples from the study area with the Federal Environmental Protection Agency FEPA (1991) and the WHO(2004) standards revealed that they were within permissible limits indicating that, the soils are good for agricultural practices. Cation oxides across seasons were also below their average shale compositions given by Wedepohl (1971).

Reduced levels of enzymatic activities within the study area were probably due to heavy metals presence caused by the ongoing mining and quarrying activities. According to Wang *et al.* (2006), enzymatic activities are seriously depressed by heavy metals occurrence due to anthropogenic activities such as mining and quarrying. Generally, enzymatic activities were high in dry season samples than in rainy season samples. Similarly, organic matter, humified carbon, humic acid and fulvic acid were observed to be higher in samples obtained during dry season than in samples obtained in rainy season except organic matter. According to Cao et al.,(2022)organic carbon in soils may probably come from the use of organic fertilizer by farmers.

Heavy metals in soil samples for both seasons were below their average shale composition. They were also below the WHO (2004) standard except Fein samplesOMS1, OMS2, OMS4,

OMS6, OMS7in dry season. Mean concentrations of all measured heavy metals in soil samples across seasons were below the FEPA (1991) permissible limits for soils except Ni, Cd, Cr. They were also below 0.03 mg/kg FEPA (1991) permissible limit for soils except in samples OMS1-4, OMS6, OMS7in rainy season but all above the FEPA (1991) limit except in samples OMS2, OMS7in dry season.

Similarly, mean concentrations of heavy metals in cassava samples were below the permissible limits given by FEPA (1991) except Cd in samples OMP1-2 and OMP1-4 in rainy and dry seasons respectively and Co in OMP3 in

Terrain/ Standards	Statistics	Fe (mg/kg)	Ni (mg/kg)	Mn (mg/kg)	Cd (mg/kg)	Cu (mg/kg)	Cr (mg/kg)	Pb (mg/kg)	As (mg/kg)	Zn (mg/kg)	Co (mg/kg)
Wet season											
Oban Massif	Min	0.02	0.02	0.01	0.1	0	0.1	0.01	0.002	0.11	0
	Max	0.12	0.11	0.1	0.12	0.01	0.11	0.11	0.11	0.11	0.01
	Mean	0.083333	0.08	0.07	0.106667	0.003333	0.106667	0.076667	0.074	0.11	0.003333
	Sdev.	0.044969	0.042426	0.042426	0.009428	0.004714	0.004714	0.04714	0.050912	0	0.004714
Dry season											
Oban Massif	Min	0.001	0.01	0.002	0.1	0.02	0.002	0.02	0.001	0.1	0.002
	Max	0.2	0.12	0.11	0.21	0.12	0.21	0.12	0.11	0.3	0.11
	Mean	0.107	0.046667	0.070667	0.14	0.08	0.107333	0.08	0.073667	0.206667	0.074
	Sdev.	0.08176	0.051854	0.048726	0.049666	0.043205	0.084937	0.043205	0.051383	0.082192	0.050912
WHO (2004)	-	1.0-3.0	10.00	0.4-6.0	0.003-0.03	0.05-2.0	1.6	2.00	0.01	1.0-3.0	0.001
FEPA (1991)	-	5.0	0.03	0.48	0.05	0.5	0.3	0.42	-	0.75	0.10
SON(2002)		0.03	0.02	0.05	0.003	1.00	0.05	0.01	-	3.0	-

and OMP1-2 in dry seasons, respectively. They were also below the WHO (2004) standards except Cd in all samples but OMP3, As in OMP1-2 in both seasons and Co in OMP3- in rainy and OMP1-3 in dry season samples.

The observed trends were possibly due to selective absorption by the cassava plant, pH and enzymatic activities in soil. Generally, it was also observed that the concentrations of measured heavy metals in 97% of the soil samples were greater than their concentrations in cassava samples obtained in the same location and during the same season.

3.1 Assessment of soil contamination and pollution by measured heavy metals

3.1.1 Contamination factor (CF): CF values < 0.1 = uncontaminated soil, 0.10-0.25 = slightly contaminated, 0.26-0.50 = moderately contaminated, 0.51-0.75 = severely contaminated, 0.76-1.00 = very severely contaminated soil, 1.10-2.00 = slightly polluted soils, 2.10- 4.00 = moderately polluted soils, 4.10-8.00 = severely polluted soils, 8.10-16.00 = very severely polluted soils, > 16.0 = excessively polluted soils (Qingjie et al., 2008; Ripin et al., 2014). Soils from the study area were uncontaminated with respect to all measured heavy metals except Fe with severe contamination in locations OMS6 and OMS7.

3.1.2 Degree of contamination (Cd): $C_d < 8$ = low degree of contamination, 8 – 16 = moderate degree of contamination, 17 – 32 = considerable degree of contamination, > 32 = very high degree of contamination. The results revealed that all soil samples from the study area exhibited low degree of contamination in respect to the measured heavy.

3.1.3 Pollution load index (PLI): According to Harikumar and Jisha (2010), PLI < 1 means no pollution and PLI > 1 means pollution. It was observed that PLI values in all soil samples in both seasons were < 1, indicating no pollution in respect to the tested heavy metals.

3.1.4 Ecological Risk Index Factor (ERIF): ERIF < 40 = low ecological risk, 40 – 80 = moderate probable ecological risk, 80 – 160 = considerable possible ecological risk, 160 – 320 = high potential ecological risk, ERIF ≥ 320 = very high potential ecological risk. Computed ERIF for all measured heavy metals in samples were < 40. Fe poses the highest ecological risk followed by Cd then Cu while the metals with the least ecological risk are Mn and Cr in soils.

3.1.5 Potential ecological risk index factor (PERIF): PERIF < 150 indicates low potential ecological risk, 150 - 300 = moderate potential ecological risk, 300 - 600 indicates significant potential ecological risk. Computed PERIF of all the tested heavy metals in samples were < 150 indicating low potential ecological risk by these metals across the area.

3.2 Assessment of biogeochemical relationship and transfer of measured heavy metals

3.2.1 Correlation matrix: From correlation analysis, positive correlations existed between temperature, EC, pH, organic carbon, organic matter, organic matter fractions, enzymatic activities and heavy metals as an indication of the influence of these parameters on heavy metal accumulation and distribution in soils.

3.2.2 Factor analysis: In factors 1-3 for both seasons, the elements that contributed to the total data variance were temperature, pH, EC, organic carbon, organic matter, organic matter fractions and heavy metals. These factor loadings from the three factors analysed can be interpreted to be contributed by geogenic activities of enzymes and their chelation effects on metals, organic matter fractions, biogeochemical releases from weathering, mineral dissolution as well as aggregate quarrying and associated activities which affected heavy metals bioavailability (Mohamed and Folorunsho, 2015; Olatunde and Onisoya, 2017).

3.3.3 Transfer factor (TF): According to Mganga et al., (2011), values of TF > 1 indicates that there is a build-up of a heavy metal concentration in the plant, whereas, TF < 1 indicates that the plant is an excluder. Results of computed TF indicated that heavy metals absorption by cassava from soils across the study area in both seasons was all low and the cassava was of the excluder's category for all metals except in locations OMP1 and OMP4 with build-up of Mn and Ni respectively. While in the dry season, cassava samples locations OMP3 with build-up of As. The build-up of As, Mn and Ni in cassava samples in these locations can probably be due to the natural occurrence of these metals in soils and tailings from intense mining, quarrying and associated processes resulting to a long residence time of these metals in soils and subsequent transfer to the cassava tubers (Ekwere and Edet, 2021; Adeleke et al., 2021; Akankpo, et al., 2021).

4. Conclusion

Soil enzymes, humus content and organic matter fractions interact with heavy metals thereby influencing metals speciation, mobility, toxicity, soluble organic complex formation, metal binding, leaching as well create pathways for heavy metals accumulation and distribution in soils. Soil pH influences the bioaccumulation of heavy metals because metals leaching in soils depend on the soil's pH and temperature. There was a potential biogeochemical link revealed by strong correlations between temperature, pH, EC, soil organic matter fractions, enzymatic activities, Pb, Cd, Ni, Zn, Fe, As, Mn-Cr, Co. This revealed link emphasizes the essential role of the intricate interplay between soil enzymes and heavy metals which highlight their significance in understanding the complex dynamics of heavy metal distribution in soils and subsequent transfer to the food chain.

5. Recommendations

- (1). A routine geochemical assessment of soils and other crops grown in the study area to ascertain their soil to plant transfer factor which are certainly different from that of the examined cassava is quite essential and highly recommended.
- (2). The observed build-up of heavy metals in soils and cassava tubers in the study area may become a future risk, it is important to introduce crop with high resistance or selective absorption to avoid futuristic heavy metals contamination of the food chain.
- (3). The application of remediation methods such as phyto-extraction, phyto-stabilization and rhizo-filtration is also recommended in areas of threat.

REFERENCES

- Abua, M. & Eyo, E. (2013). Assessment of soils around quarry terrain in Akamkpa Local Government Area, Cross River State-Nigeria. *Merit Research Journal of Agricultural Science and Soil Sciences*, 1(1), 1-5.
- Abua, M. & Edet, E. O. (2007). Morphological and physico-chemical characteristics of coastal plain soils of Southern Cross River State, Nigeria. *Nigerian Geographical Journal*, 5(1), 109-114.
- Adeleke, O., Akinlabi, S. A., Jen, T. C., & Dunmade, I. (2021). Application of artificial neural networks for predicting the physical composition of municipal solid waste: An assessment of the impact of seasonal variation. *Waste Management & Research*, 39(8), 1058-1068.
- Akankpo, A., Essien I., Nyong, A. & Inyang, E. (2021). Soil to cassava transfer factors of natural radionuclides in farms in Ini Local Government Area, Akwa Ibom State, Nigeria. *Journal of Scientific Research & Reports*, 27(9): 131-138, Article no. JSRR.74322
- Alloway, B. J. & Jackson AP. (1999). Behaviour of trace metals in sludge-amended soils. *Sci Total Environ.*, 100, 151–176.
- Anguelov, G. & Anguelova, I. (2009). Assessment of land-use effect on trace elements concentrations in soils solution from Udisols in North Florida. *Agric. Ecosyst and Env.*, 130, 59–66.
- Cao H, Wang Z, Liao X, Li Y and Zhu Y (2022) Research on the Regional Environmental Impact and Risk Assessment Affected by Mineral Resource Development: A Case Study of the Taojia River Watershed in Hunan. *Front. Ecol. Evol.* 10:918446. doi: 10.3389/fevo.2022.918446.
- Dube, A., Zbytniewski, R., Kowalkowski, T., Cukrowska, E., & Buszewski, B. (2001). Adsorption and migration of heavy metals in soils. *Polish Journal of Environmental Studies*, 10 (1), 1-10.
- Duffus J.H. (2002). "Heavy metals" a meaningless term? *Pure Appl Chem* 2002; 74(5): 793-807. doi: 10.1351/pac200274050793.
- Edet, A. E & Okereke, C.S. (1997). Assessment of hydrogeological conditions in basement aquifers of Precambrian Oban Massif, South-eastern Nigeria. *Journal of Applied Geophysics*, 36, 195–204.
- Ekwere, A.S. & Edet, B. B. (2021). [Temporal variations of heavy metals in sediment, soil and dust particulates across the rock quarrying districts of the Oban Massif, South-eastern Nigeria](#). *Environmental Nanotechnology, Monitoring & Management*; 15: 100431, Elsevier.
- Ekwere, A.S & Edet, B. B., (2023). Hydrogeochemistry of groundwaters within the rock [quarrying districts of the Oban Massif, South-eastern Nigeria](#). *Indian J. Environmental protection*, 43 (2): 108-118.
- Ekwueme, B. N. (2003). *The Precambrian geology and evolution of the Southeastern Nigerian*
- Ekwueme, B. N., Nyong, E. E., & Petters, S. W. (1995). *Geological excursion guidebook to Oban Massif, Calabar Flank and Ikom Mamfe Embayment, Southeastern Nigeria*. Calabar, Nigeria: Dec-Ford Publishers, 36pp.
- Fergusson J. E. (1990). *The Heavy Elements: Chemistry, Environmental Impact and Health Effects*. Oxford: Pergamon Press.
- Federal Environmental Protection Agency (FEPA) (1991), Guidelines and Standards for Environmental Pollution Control in Nigeria.
- Harikumar, P. S. & Jisha, T. S. (2010). Distribution pattern of trace metal pollutants in the sediments of an urban wetland in the Southwest Coast of India. *International Journal of Engineering Science and Technology*, 2(5), 840-850.
- Khan S, Cao Q, & Zheng Y.M., (2008) Health risk of heavy metals in contaminated soils and food crops irrigated with wastewater in Beijing, China. *Environ Pollut* 152:686–692.
- Kuźniar, A., Banach, A., Stepniowska, Z., Frąc, M., Oszust, K., Gryta, A., Kłos, M., & Wolińska, A. (2018). Community-level physiological profiles of microorganisms inhabiting soil contaminated with heavy metals. *International Agrophysics*, 32, 101–109.
- Lasota, J., Błońska, E., Łyszczarz, S., & Tibbett, M. (2020). Forest humus type governs heavy activity of forest soils that are heavily contaminated by lead and zinc. *Water, Air, & Soil Pollution*, 227, 348.
- Lenart-Boron A. and Boron P. (2014). The Effect of Industrial Heavy Metal Pollution On Microbial Abundance and Diversity in Soils — A Review. *Environmental Risk Assessment of Soil Contamination*. DOI: 10.5772/57406.
- Mganga, N. Manoko, M. Rulangeranga, Z. (2011). Classification of plants according to their heavy metal content around North Mara gold mine, Tanzania: implication for phytoremediation, *Tanzania J. Sci.* 37.

- Mohammed, S.A. & Folorunsho, J.O. (2015). Heavy metal concentration in soil and *Amaranthus retroflexus* grown on irrigated farmlands in the Makera Area, Kaduna, Nigeria. *Journal of Geography and Regional Planning*, 8 (8): 210-217
- Olatunde, E., & Onisoya, M. O. (2017). Assessment of Heavy Metal Concentrations in Pawpaw (*Carica Papaya* Linn.) around Automobile Workshops in Port Harcourt Metropolis, Rivers State, Nigeria. *Journal of Health and Pollution*, 7, 48-61. <https://doi.org/10.5696/2156-9614-7.14.48>
- Pajak, M., Błońska, E., Frac, M., & Oszust, K. (2016). Functional diversity and microbial activity of forest soils that are heavily contaminated by lead and zinc. *Water, Air, and Soil Pollution*, 227, 348.
- Qingjie, G., Jun, D., Yunchuan, X., Qingfei, W., & Liqiang, Y. (2008). Calculating pollution indices by heavy metals in ecological geochemistry assessment and a case study in Parks of Beijing. *Journal of China University of Geosciences*, 19(3), 230–241.
- Ripin, S. N. M., Hasan, S., Kamal, M. L., & Hashim, N. M. (2014). Analysis and pollution assessment of heavy metal in soil, Perlis. *The Malaysian Journal of Analytical Sciences*, 18, 155–161.
- Sohi, S. P., Mahieu, N., Arah, J. R., Powlson, D. S., Madari, B., & Gaunt, J. L. (2001). A procedure for isolating soil organic matter fractions suitable for modeling. *Soil Science Society of America Journal*, 65(4), 1121-1128.
- Sracek, O. & Zeman, J. (2004). *Introduction to environmental hydrogeochemistry*. Brno: Masaryk University.
- Standard Organization of Nigeria (SON), (2002). Nigerian Standards for Drinking Water Quality, Revised edition 2007.
- [Tchounwou, P. B.](#), [Yedjou, C. G.](#), [Patlolla, A. K.](#) and [Sutton, D. J.](#) (2014). Heavy Metals Toxicity and the Environment. HHS PUBLIC ACCESS. [EXS. 2012; 101: 133–164.](#)
- Von Mersi W. and Schinner F. (1991) An improved and accurate method for determining the dehydrogenase activity of soils with iodonitrotetrazolium chloride. *Biology and Fertility of Soils* 11, 216-220.
- Walkley A. & Black, I. A (1934). Methods of soil analysis, Part 2, Amer Soc. of Agronomy, Madison, pp 539 – 579.
- Wang F, Wang Z, Kou C, Ma Z and Zhao D. (2016). Responses of wheat yield, macro- and micronutrients, and heavy metals in soil and wheat following the application of manure compost on the North China plain. *PLoS ONE*. 2016;11(1):e0146453.
- Wedepohl, K. H. (1971). Environmental influences on the chemical composition of shale and clays, in Ahrens, L. H., Press, F., Runcom, S. K. and Urey, H. C. eds. *Physics and chemistry of the earth*. Oxford Pergamon press, 8: 305-335.
- World Health Organization. (2004). *Guideline for drinking water quality*. (3rd Ed). New York, Geneva, Switzerland.

CONGLOMERATE FACIES OF THE AWI FORMATION, CALABAR FLANK, SOUTHEAST NIGERIA: IMPLICATIONS FOR PALEOENVIRONMENTS AND ECONOMIC GEOLOGY

Okon E. E.¹, Kudamnya, E. A.¹, Essien, N. U.^{1,2}, Inyang, D. O.¹, Nton, M. E.³ and Ugbaja, A. N.¹

4. University of Calabar, Calabar – Nigeria
5. Akwa Ibom State University, Mkpatt-Enin – Uyo, Nigeria
6. University of Ibadan, Ibadan – Nigeria

Abstract

Pebble morphometry of unbroken quartz pebbles collected from the basal section of Awi Formation exposed within the Calabar Flank, southeastern Nigeria was studied for paleoenvironmental reconstruction. The pebbles were selected from 4 locations across the Awi Formation. Morphometric studies carried out involved the measurement of the three orthogonal axes (long, intermediate and short), determination of their corresponding roundness, flatness indices and elongation indices for no less than 200 pebbles. The depositional processes responsible for shaping the pebbles and the environment that prevailed during past geological times were characterised. The results show that the pebbles are sub-rounded to sub-angular and predominantly compact-bladed. The mean values for the following morphometric parameters: Flatness index, elongation ratio, maximum projection sphericity index and oblate-prolate (OP) index are 0.57, 0.78, 0.74 and 15.65 respectively. The results were integrated with deductions from bivariate plots of roundness against elongation ratio and sphericity against OP index. Both plots indicated paleo-depositional environment of the conglomeratic sandstones in a fluvial setting with subordinate transitional marine setting (littoral influence). With the integration of these observations (e.g. pebble imbrication, sediment stratification, grading style), a typical fluvial paleoenvironmental setting is suggested. The geology of the hinterland areas where sediments are sourced contributes largely to the materials at the depocenter. It is possible that the jointing, faulting, sheeting and/or exfoliation of the rocks of the Precambrian Oban Massif, which is believed to be the principal source of the sediments (provenance), also accounts for the abundance of vein quartz in the chosen area of this study. More importantly, besides the significance of pebble morphometry in deciphering paleoenvironments, it also gives clues for potential sites of ore bodies, concentrating them as placer deposits.

Keywords: *Morphometric parameters; bivariate, conglomerates; paleoenvironmental reconstruction; fluvial setting; elongation ratio.*

5. Introduction

The significance of textural characteristics of sediments as an invaluable tool for characterizing both depositional processes and environment of deposition is well established (Visher 1965; Miall 1985; Essien et al 2016; Miall 2016; Okon and Ojong 2019; Okon et al 2021). The conditions of deposition are well recorded in sedimentary packages and in cases where erosion is not of immense significance, information regarding sedimentation is easily reconstructed using facies analysis. These reflect the assemblages of individual lithofacies elements and other associated geomorphic elements such as channels and bars. The log motif depicted in the vertical profile is often used in concert with the morphology of the channels to carry out interpretation of characteristic depositional environment. The Awi Formation, the oldest known sedimentary deposit in the Calabar Flank, is composed principally of para-conglomerates, cross bedded arkosic sandstones and mudrocks belonging to the basal section of the sedimentary succession of the Calabar Flank, southeastern Nigeria (Okon, 2015).

The study of the sedimentological attributes of the facies succession of the Awi Formation has aided paleoenvironmental interpretation while the integration of geochemistry of the sediments has facilitated provenance determination (Boboye and Okon 2014; Okon 2015). Among the series of approaches used in carrying out paleoenvironmental analyses of any given sedimentary sequence, textural analyses (Awasthi, 1970; Friedman 1979; Inyang and Enang, 2002; Essien and Okon, 2016), analysis of fossil contents (faunal and floral) in the rock unit (Nyong and Ramanathan 1985; Akpan and Ntekim 2004; Itam et al 2016), analysis of traces made by organisms that lived on/within the sediments (Akpan and Nyong 1987; Pemberton et al 2012), and geochemical proxies (Amajor 1987; Tijani et al 2010; Okon, 2015; Ibe and Okon 2021) are important, to mention a few. Paleoenvironmental analysis refers to the study or use of ancient geological materials (rocks) to unravel the environment of depositional which they were deposited and the results have greatly aided interpretations from basin analysis to identification of valuable placer deposits (Okon et al 2018). This study focuses on the conglomerate facies of the Awi Formation and its significance in characterization of the energy of depositing medium, possible environment of deposition and indication of where placer deposits may be investigated further.

6. Study area description

The study was carried out within the Calabar Flank, southeastern part of the Benue Trough. The basin is underlain by horsts and graben structures and has a NW-SE trending structural orientation (Nyong and Ramanathan, 1985). Fluvial sediments within the Calabar Flank are restricted to the Awi Formation, predominantly clastics, ranging in size from coarse grains to mud class. The area under investigation is delimited to the west by the Ikpe platform and to the east by the Cameroon Volcanic Line. To the south, the Calabar hinge line separates it from the north-eastern portion of the Niger Delta (Fig. 1). Its origin is closely associated with the breakup and subsequent separation of Africa and South America about 120-130Ma ago (Murat, 1972). Sedimentation in the Calabar Flank commenced with the deposition of fluvio-deltaic clastics (Awi Formation) of Neocomian-Aptian age unconformably on the Precambrian crystalline basement complex rocks of the Oban massif. This is succeeded by the shallow marine Mfamosing Limestone, the deeper marine Ekenkpon Shale, New Netim Formation and the Nkporo Shale (Fig 2).

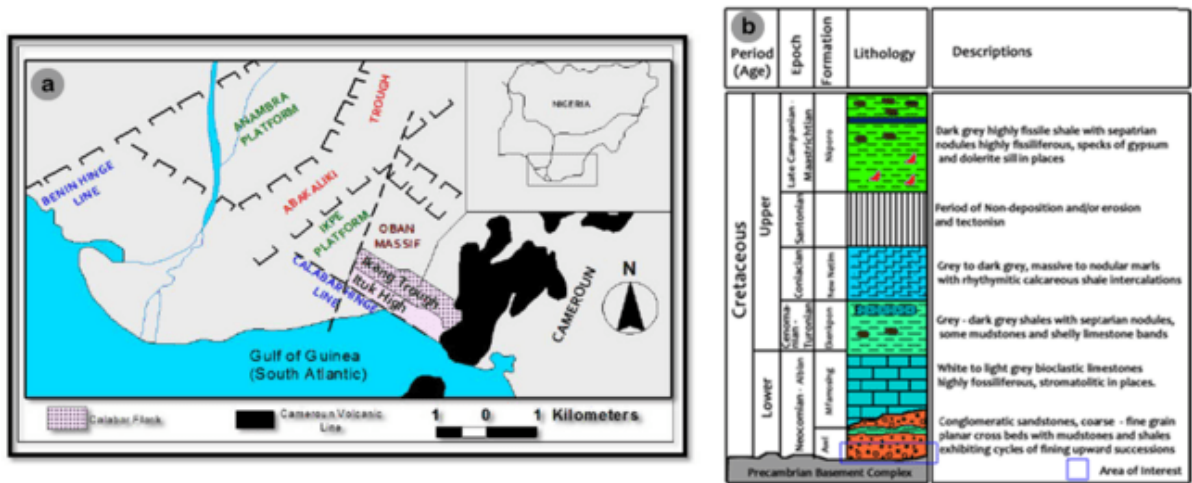


Fig 1. (a) Sketch of southern Nigeria showing the major tectonic elements and geographic location of Calabar Flank; (redrawn from Nyong and Ramanathan 1985) (b) Stratigraphic subdivision of the Calabar Flank (after Okon et al 2017).

7. Materials and Method

Road cut sections of the Awi Formation exposed along Calabar Itu Highway, Okoyong Usang Abasi-Njagachang road, Abiati area and near Ewen community were instrumental to sample collection and analyses for this study. The conglomerates constitute a significant member the basal section of the Awi Formation (Fig 3a-b) and present a non-conformity between the basement rocks of the Oban Massif in the Calabar Flank (Okon et al 2017). The visited outcrops were properly logged and described (Fig. 3c). At each location, 50 unbroken quartz pebbles were collected in 5 batches of 10 each. The analysis was carried out with the mean form of at least 10 pebbles taken from each sampling station. In each case, 5 sets per sample location representing 50 pebbles for the four locations visited. The measured parameters were subjected to statistical and mathematical treatments as presented in the ratios in table 1.

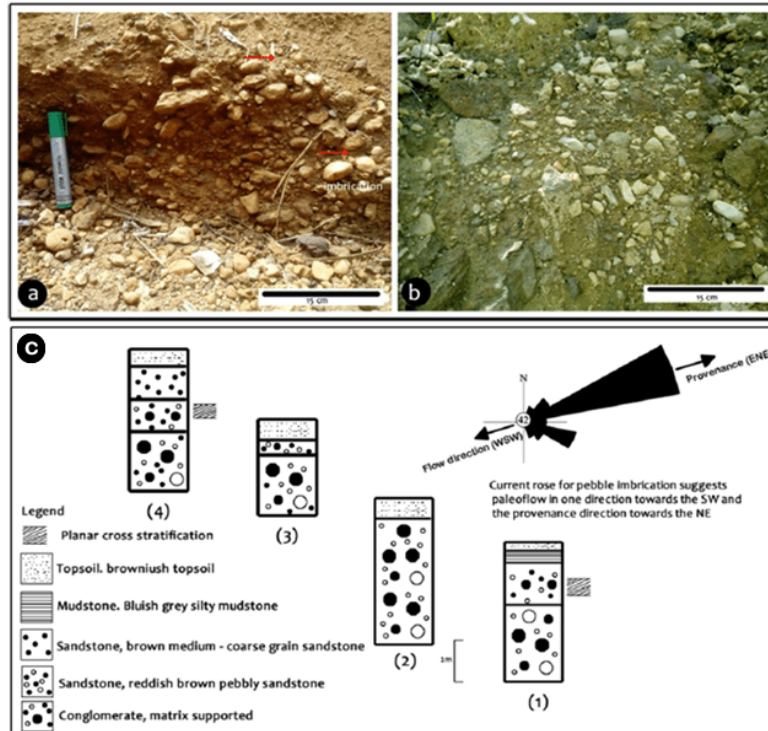


Figure 3. (a) Matrix – supported conglomerates showing clast imbrication; (b) admixture of brecciated rock units with sub-rounded pebbles; (c) Lithologic log for the studied samples

Table 1. Formulae used for statistical computation of pebble morphometric parameters

S/No	Formula	Reference
1	Maximum Projection Sphericity Index (MPSI) = $\{S^2/LI\}^{1/3}$	Sneed and Folk, 1958
2	Elongation Ratio, (ER) = I/L	Lutig, 1962; Sames, 1966
3	Flatness Ratio $FR = S/L$	Lutig, 1962
4	Flatness Index = $(L - I + S) / L$	Illenberger 1992
5	Oblate – Prolate Index, $OPI = \frac{10 \left(\frac{L-I}{L+S} - 0.50 \right)}{S/L}$	Dobkins and Folk 1970
6	Roundness = $[(\sum r)/nR]$	Waddell 1932

8. Results and discussion

The result for the mean pebble morphometric parameters is presented in Table 2. The pebbles are massive, matrix supported and crudely bedded, with clasts diameter ranging from 2.63 – 3.40 cm. The sorting is poor and pebble grains are weakly imbricated, with brecciated ferruginized layer admixed with sub-rounded pebbles (see Figs 3a-b). These features suggest lag deposits and conform to Miall (1978) facies classification “Gm”. Regarding the clasts sphericity, roundness and “Oblate – Prolate” Indexes, the parametric values of an average of 10 pebbles was used in the analysis (Sneed and Folk 1958). This formula was adopted because it was established comparing the volume of the particle with its maximum projection area which naturally opposes the direction of motion. The results show that the pebbles are sub-rounded to sub-angular and predominantly compact-bladed (Fig 4). Shape indices as paleoenvironmental indicators of quartzite rich rocks have been the subject of considerable discussions among experts (Els, 1988; Lorang and Komar, 1990; Illenberger and Reading, 1993), and the result have greatly aided interpretations from basin analysis to identification of valuable placer deposits. The mean values for the following morphometric parameters: Flatness index, elongation ratio, maximum projection sphericity index and oblate-prolate (OP) index are 0.57, 0.78, 0.74 and 15.65 respectively. The various parameters obtained in the pebble morphometric analysis together, with the different bivariate scatter plots have shown that the depositional environment of the basal of the section of the Awi Formation is predominantly influence by fluvial processes with beach / littoral influence (Figs 5a-b; Figs 6a-b) (Lutig, 1962; Okon et al 2019).

Table 2. Result for the mean values of 20 batches of pebble morphometric parameters for Awi Formation

S/N	L	S	I	S/L	Elongation (I/L)	L-I	L-S	L-I	S-I	OPI	MPS	FI	Roundness	Form name
L1/B1	2.80	1.63	2.31	0.58	0.83	0.49	1.17	6.47	2.66	-0.74	0.74	76.04	0.38	CB
L1/B2	2.67	1.44	2.26	0.54	0.85	0.41	1.23	6.03	2.06	-1.95	0.70	69.22	0.4	CB
L1/B3	3.04	1.82	2.30	0.60	0.76	0.74	1.22	6.97	3.29	0.69	0.78	84.39	0.43	CB
L1/B4	3.17	1.71	2.47	0.54	0.78	0.70	1.46	7.83	2.92	-0.46	0.72	75.74	0.41	CB
L1/B5	2.63	1.39	2.02	0.53	0.77	0.61	1.24	5.29	1.93	-0.24	0.71	76.28	0.34	CB
L2/B6	2.82	1.60	2.12	0.57	0.75	0.70	1.23	5.98	2.54	0.43	0.75	81.54	0.46	CB
L2/B7	2.68	1.69	2.04	0.63	0.76	0.64	0.99	5.46	2.86	2.21	0.81	87.77	0.45	CB
L2/B8	3.40	1.63	2.43	0.48	0.71	0.97	1.77	8.23	2.64	-0.03	0.68	75.99	0.47	CB
L2/B9	2.68	1.58	2.26	0.59	0.84	0.43	1.11	6.04	2.48	0.02	0.74	77.55	0.42	CB
L2/B10	2.63	1.66	2.04	0.63	0.78	0.59	0.97	5.34	2.76	0.50	0.80	85.88	0.39	CB
L3/B11	2.69	1.57	2.00	0.58	0.74	0.69	1.12	5.38	2.46	0.97	0.77	82.54	0.41	CB
L3/B12	2.97	1.52	2.06	0.51	0.69	0.92	1.46	6.10	2.30	0.14	0.72	80.89	0.41	CB
L3/B13	2.75	1.45	2.23	0.53	0.81	0.52	1.31	6.13	2.09	-0.69	0.70	71.91	0.44	CB
L3/B14	2.76	1.63	2.19	0.59	0.79	0.57	1.13	6.02	2.66	0.15	0.76	80.43	0.44	CB
L3/B15	2.97	1.53	2.32	0.52	0.78	0.65	1.44	6.89	2.34	-0.21	0.70	72.98	0.47	CB
L4/B16	2.38	1.31	1.85	0.55	0.78	0.53	1.07	4.40	1.72	-0.04	0.73	76.95	0.38	CB
L4/B17	2.49	1.45	1.97	0.58	0.79	0.52	1.04	4.91	2.10	0.003	0.75	79.66	0.44	CB
L4/B15	2.49	1.47	1.88	0.59	0.76	0.61	1.02	4.68	2.16	0.53	0.77	83.62	0.39	CB
L4/B19	2.55	1.56	2.05	0.61	0.80	0.50	0.99	5.23	2.43	0.64	0.78	82.88	0.44	CB
L4/B20	2.55	1.48	1.90	0.58	0.75	0.65	1.07	4.85	2.19	0.80	0.77	84.30	0.42	CB
Mean	2.75	1.55	2.13	0.57	0.78	0.62	1.20	5.91	2.43	-0.74	0.74	76.04	0.42	-

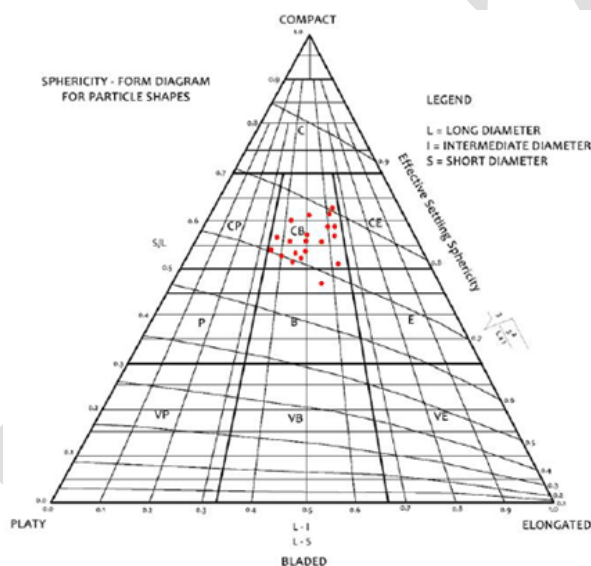


Figure4. Sphericity-form diagram for particle shapes (after Sneed and Folk 1958) – majority of the samples plot under compact bladed pebbles

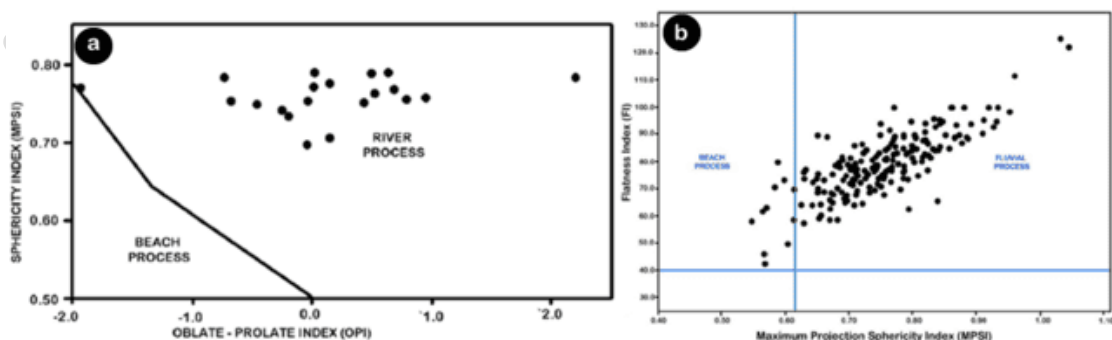


Fig 5. (a) Plot of sphericity against OPI (fields after Dobkins and Folk, 1970); (b) Plot of flatness index (FI) against maximum projection sphericity index (fields after Stratten 1974)

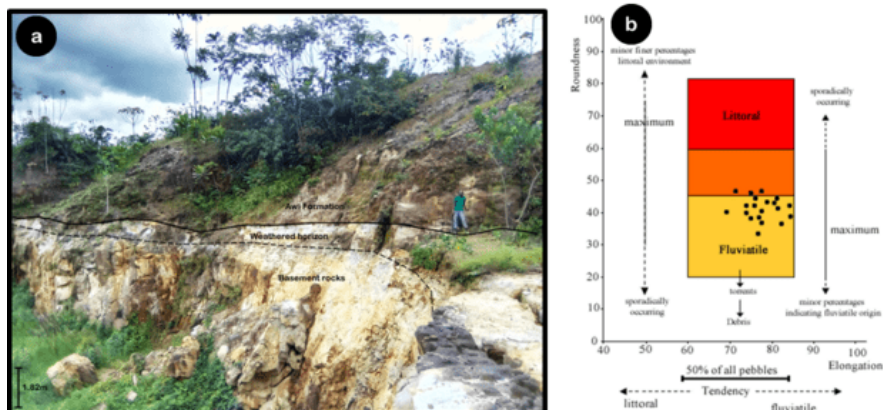


Figure 6. (a) Field photo showing fining upwards succession in the Awi Formation with the conglomeratic facies of the Awi Formation at the base; (b) Environmental determination chart (modified after Sames 1966).

The sediments along the transect show energy setting associated with a mix of high and relaxed energy at different times during depositional process, establishing conditions for selective gravitational sorting of the sediments, allowing for the possible concentration of placers of economic interests at various depocenter. In such cases, transient placer deposits (alluvial-colluvial and placers and fluvial placers) are the most likely targets for mineralization. These sediments are typically deposited by traction currents as bedloads and as bed contact loads and this interpretation is justified by their accumulation in different bedforms.

Conclusion

In this study, pebble morphometric analysis has aided the determination of paleoenvironments prevalent during the deposition of the basal conglomerates of Awi Formation. The depositional processes (abrasion conditions) responsible for shaping the pebbles were characterized from the study of the clasts morphology and their imbrication azimuth. Fluvatile process with some overlapping littoral influence has been shown to be responsible for the variation in clasts morphology of the para-conglomerates (matrix-supported) of Awi Formation. It is possible that the jointing, faulting, sheeting and/or exfoliation of the rocks of the Oban Massif, which is believed to be the principal source of the sediments (provenance), also accounts for the abundance of vein quartz in the area which was eventually adapted for this study. The recent mining activities at the southern fringes of the Oban Massif, involving gold extraction from active recent drainage channels suggests that some paleo-channels may form targets for gold exploration if proper attention is given to studies of this nature. This hypothesis is yet to be tested but is drawn from the fact that the rocks from where the gold grains are mined pre-date the Awi Formation and studies have shown that the provenance of the Awi Formation is the nearby Oban Massif rocks. This forms a template for further research in the area, especially geared towards testing the outlined hypothesis.

References

- Akpan E. B. and Ntekim E. E. (2004). Cretaceous bivalves and paleoenvironments of the Calabar Flank, SE Nigeria. *Global Journal of Geological Sciences*, 2(1), 15-36.
- Akpan E. B. and Nyong E. E. (1987). Trace fossils assemblage and depositional environment of Turonian calcareous sandstones in the southern Benue Trough, Nigeria. *Journal of African Earth Sciences*, 6(2), 175-180.
- Amajor L. C. (1987). Major and trace element geochemistry of Albian and Turonian shales from the Southern Benue trough, Nigeria. *Journal of African Earth Sciences*, 6(5), 633-641.
- Awasthi A. K. (1970). Skewness as an environmental indicator in the Solani river system, Roorkee (India). *Sedimentary Geology*, 4, 177-183.
- Boboye O. A. and Okon E. E. (2014). Sedimentological and geochemical characterization of the Cretaceous strata of the Calabar Flank, southeastern Nigeria. *Journal of African Earth Sciences*, 99, 427-441.
- Dobkins J. E. and Folk R. L. (1970). Shape development of Tahiti-Nui: *Journal of Sedimentary Petrology*, 40, 1167-1203.
- Els B. G. (1988). Pebble morphology of an ancient conglomerate: The middelvlei gold placer, Witwatersrand, South Africa. *Journal of Sedimentary Petrology*, 58(5), 894-901.
- Essien N. U. and Okon E. E. (2016). Sedimentary processes and depositional environments of part of the Benin Formation in the Calabar Flank, southeastern Nigeria. *British Journal of Applied Science and Technology*, 18(5), 1-10.
- Essien N. U., Ilori A. O., Okon E. E. and Njoku S. N. (2016). Textural characteristics and depositional processes of sediments from a 47 km transect in the Niger Delta, Southern Nigeria. *Journal of Geography, Environment and Earth Science International*, 7(1), 1-11.
- Friedman, G. (1979). Differences in size distributions of populations of particles among sands of various origins. *Sedimentology*, 26, 3-32.

- Ibe C. U. and Okon E. E. (2021).** Provenance and tectonic settings of the Eze-Aku Sandstone (Turonian) in Awajir and adjoining areas, Southern Benue Trough, Nigeria: evidence from petrography and geochemistry. *Journal of Sedimentary Environments*, 6, 237–254
- Illenberger W. K. (1992).** Pebble shape (and size). *Journal of Sedimentary Petrology*, 61(5), 756–767.
- Illenberger W.K. and Reddering J. S. V. (1993).** An evaluation of shape indices as paleoenvironmental indicators using quartzites and metavolcanics clasts in Upper Cretaceous to Paleogene beach, river and submarine fan conglomerates (Discussion). *Sedimentology*, 40, 1019–1020
- Inyang D. O. and Enang E. I. (2002).** Use of statistical parameters in the sedimentological study of conglomerates in the north-eastern part of the Niger Delta, Nigeria. *Journal of Mining and Geology*, 38(1), 21–26.
- Itam A. E., Inyang D. O., Umana U. S., Akpan E. B., Udinmwun E. (2016).** Palynostratigraphy of Lemna Road Transect of Benin Formation, Calabar Flank, Nigeria. *Asian Journal of Physical and Chemical Sciences*, 1(1), 1–10.
- Lorang M. S. and Komar P. D. (1990).** Pebble shape. *Nature*, 347, 433–434.
- Lutig G. (1962).** The shapes of pebbles in continental, fluvial and marine facies. International Association of Geology and Geophysics. *International Association of Hydrogeologists*, 59, 253–258.
- Miall, A. D. (2016).** Facies Analysis. In: *Stratigraphy: A Modern Synthesis* (pp 77–159). Springer, Cham. https://doi.org/10.1007/978-3-319-24304-7_3
- Miall A. D. (1978).** Lithofacies types and vertical profile models in braided river deposits: A summary. In: A. D. Miall (Ed), *Fluvial Sedimentology* (pp 597–604). Canadian Society Petroleum Geologists Memoir, 5.
- Miall, A. D., (1985).** Architectural-element analysis: A new method of facies analysis applied to fluvial deposits. *Earth Science Reviews*, 22, 261–308.
- Murat R. C. (1972).** Stratigraphy and paleogeography of the Cretaceous and Lower Tertiary in Southern Nigeria. In: T.F.J. Dessauvage and A.J. Whiteman (Eds). *African Geology* (pp 251–266). University of Ibadan Press, Nigeria
- Nyong E. E. and Ramanathan R. M. (1985).** A record of oxygen deficient paleoenvironments in the Cretaceous of the Calabar Flank, southeast Nigeria. *Journal of African Earth Science*, 3(4), 455–460.
- Okon, E. E. (2015).** Sedimentologic characteristics and provenance studies of AWI Formation, Calabar Flank, SE Nigeria. *Journal of Mining and Geology*, 51(2), 121–138.
- Okon E. E., Essien N. U. and Adeyemi G. O. (2017).** Geochemistry of sandstones of Awi Formation, Southeastern Nigeria: Implications for weathering, provenance and tectonics settings. *International Journal of Science and Technology*, 6(4), 742–755.
- Okon E. E., Asi M. O. and Ojong R. A. (2018).** Morphometric studies of pebbles from Ewen area, Calabar Flank, Southeastern Nigeria: Implications for paleoenvironmental reconstruction. *Physical Science International Journal*, 17(4), 1–12.
- Okon, E. E. and Ojong, R. A. (2019).** Paleoenvironmental Analysis and Its Significance in Sedimentology: Case Study of the Conglomerate Facies of the Awi Formation, Calabar Flank, Southeast Nigeria. In: M. Rafatullah, (Ed). *Advances and Trends in Physical Science Research*. 2. (pp84 – 98).
- Pemberton S. G., MacEachern J. A., Dashtgard S. E., Bann K. L., Gingras M. K. and Zonneveld J. P. (2012).** Shorefaces. In: D. Knaust and R. G. Bromley (Eds.), *Trace Fossils as Indicators of Sedimentary Environments* (pp 563–604). Developments in Sedimentology. 64.
- Sames C. W. (1966).** Morphometric data of some recent pebble associations and their application to ancient deposits. *Journal of Sedimentary Petrology*, 36, 126–142.
- Sneed E. D. and Folk R. L. (1958).** Pebbles in the lower Colorado river, Texas. A study in particle morphogenesis. *Journal of Geology*, 66, 114–150.
- Stratten T. (1974).** Notes on the application of shape parameters to differentiate between beach and river deposits in southern Africa. *Transactions of the Geological Society of South Africa*, 77, 59–64
- Tijani M. N., Nton M. E. and Kitagawa R. (2010).** Textural and geochemical characteristics of the Ajali Sandstone, Anambra Basin, SE Nigeria: Implication for its provenance. *Comptes Rendus Geoscience* 342, 136–150
- Visher, G. S. (1965).** Use of vertical profile in environmental reconstruction; American Association of Petroleum Geologists Bulletin, 49, 41–61.

SUITABILITY ASSESSMENT OF WATER RESOURCES AROUND AKAMKPA AND ENVIRONS FOR IRRIGATION PURPOSES: AN INTEGRATED APPROACH.

B.B. Edet and A. S. Ekwere

Department of Geology, University of Calabar
Calabar, CRS, Nigeria.

E-mail: edetbernard8@gmail.com; zerrata77@yahoo.com

ABSTRACT

The quality of ground and surface waters within Akamkpa and environs were assessed and characterised for irrigational suitability. From 18 randomly selected sampling locations (streams, rivers, hand dug wells and boreholes), 35 water samples were collected for investigation in both rainy and dry seasons using the APHA standard methods. Temperature, pH, and EC were determined in-situ using appropriate equipment and methods. Samples were analysed for Na^+ and K^+ by flame photometric method, Mg^{2+} and Ca^{2+} by titrimetric method using EDTA, SO_4^{2-} via turbidimetric method, HCO_3^- and CO_3^{2-} by titrimetric method. NO_3^- via the Brucine colorimetric method described in Ademoroti (1996) using UNICAM UV2 spectrophotometer. Dissolved Oxygen was determined by method of precipitation and Cl^- via volumetrically using standard AgNO_3 solution. Comparing parameters such temperature, pH, EC, total hardness and total alkalinity with standards showed that they were all below the permissible limits for irrigation waters. Computed values of SAR, %Na, RSC, Kelly's Ratio, MR, PI, Ion Exchange as chloro-alkaline indices, and Wilcox Plot indicated that both surface and ground water resources from the area were all below the permissible limits making the waters fit for irrigational purpose.

Keywords: Suitability, Physicochemical Parameters; Total Hardness; Alkalinity; Irrigational Indices; Salinity, Ratio.

INTRODUCTION: Water is very essential for crop production. Narayanan, (2007) posited that, the quality of natural water is generally controlled by factors such as; geology, recharge, discharge, human activities, residence time and the prevalent environmental conditions through which the water travels thus increasing its ionic content and total dissolved solids. To evaluate water qualitatively for irrigational use, there is need to have a better understanding of its prevalent chemical characteristics, i.e., the concentration of its total dissolved constituents (Raihan and Alam, 2008; Kavidha, 2012; Ibraheem and Khan, 2017). As suggested by Talukder et al., (1998), testing water quality prior to irrigation contributes to effective management and utilization of water resources since poor quality of irrigation water affects crop yields and soil physical conditions. Therefore, concerted efforts should be geared towards characterising as well as protecting all water sources and resources to ensure suitability and sustainable agricultural productivity in line with the recent agricultural revolution for food security by Federal Government of Nigeria.

Study area Description

The study area, Akamkpa town and its environs is situated in western Oban Massif, Southeast Nigeria and has an estimated area of about 4,500 km^2 . Geographically, the area is approximately located between Longitudes $8^\circ 00'$ and $8^\circ 55' \text{ E}$ and Latitudes $5^\circ 00'$ and $5^\circ 45' \text{ N}$, (Fig. 1).



FIG.1: Map of Cross River State showing geologic settings, the study location and insert of Nigerian map. (Modified from Ekwere and Edet, 2021).

Geologically, the area consists of highly deformed and weathered basement rocks of mainly igneous and metamorphic origin such as gneisses, schist and granite intruded by granodiorites, tonalites, pegmatite, diorites, charnockites and dolerites (Edet and Okereke, 1997; Ekwueme, 2003; Ekwere and Edet, 2023) (Fig. 2).

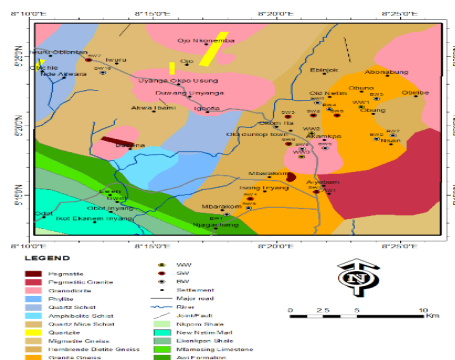


FIG. 2: Geologic map of the study area showing major drainage and sample locations

According to Abua and Edet, (2007), soils in the area are derivatives of acidic crystalline rocks mainly granites and gneisses, with loamy sand to sandy loamy surface texture, leached mineral soils, acidic in reaction, low-in contents of exchangeable bases, organic carbon, total and available phosphorous. The area has many rivers and streams Fig 2.

MATERIALS AND METHODS

Collection of water samples: Surface and ground water samples were collected in wet and dry seasons at 18 locations (Fig. 3), and 37 water samples were collected for examination. At each sampling station, two portions of each sample were collected into plastic containers.

Sample analysis: Water samples were analyzed for; Na^+ and K^+ , using flame photometric method, Mg^{2+} and Ca^{2+} were determined by titrimetric method using EDTA, SO_4^{2-} was determined using turbidimetric method, HCO_3^- and CO_3^{2-} by titrimetric method. NO_3^- was determined via the Brucine colorimetric method described in Ademoroti (1996) using UNICAM UV2 spectrophotometer. Dissolved Oxygen was determined by method of precipitation and Cl^- was determined volumetrically using standard AgNO_3 Solution.

Irrigation Indices Used.

$$\text{SAR} = \frac{\text{Na}^+}{\sqrt{\text{Ca}^{2+} + \text{Mg}^{2+}}} \dots\dots\dots(1)$$

$$\text{RSC} = [(\text{CO}_3^{2-} + \text{HCO}_3^{2-}) - (\text{Ca}^{2+} + \text{Mg}^{2+})] \dots\dots\dots(2)$$

$$\text{RSBC} = \text{HCO}_3^{2-} \text{Ca} \dots\dots\dots(3)$$

$$\text{PI} = (\text{Na}^+ + \sqrt{\text{HCO}_3^{2-}} * 100 / (\text{Ca}^{2+} + \text{Mg}^{2+} + \text{Na}^+)) \dots\dots\dots(4)$$

$$\text{MR} = \text{Mg}^{2+} * 100 / (\text{Ca}^{2+} + \text{Mg}^{2+}) \dots\dots\dots(5)$$

$$\text{Kelly's Ratio} = \text{Na}^+ / (\text{Ca}^{2+} + \text{Mg}^{2+}) \dots\dots\dots(6)$$

$$\% \text{Na}^+ = [(\text{Na}^+ + \text{K}^+) / (\text{Ca}^{2+} + \text{Mg}^{2+} + \text{Na}^+ + \text{K}^+)] * 100 \dots\dots\dots(7)$$

$$\text{CAI -I} = \text{Cl}^- - (\text{Na}^+ + \text{K}^+) / \text{Cl}^- \dots\dots\dots(8)$$

$$\text{CAI -II} = \text{Cl}^- - (\text{Na}^+ + \text{K}^+) / (\text{HCO}_3^{2-} + \text{SO}_4^{2-} + \text{CO}_3^{2-} + \text{NO}_3^-) \dots\dots\dots(9)$$

$$\text{Total Hardness (TH)} = (\text{Ca}^{2+} + \text{Mg}^{2+}) * 50 \dots\dots\dots(10)$$

$$\text{Todd's Ratio (TR)} = \text{Cl}^- / \text{HCO}_3^{2-} \dots\dots\dots(11)$$

RESULT AND DISCUSSION

Physical Parameters: Temperature, pH, EC, and TDS were below their permissible limits when compared with standards as shown in tables 1.

Permeability index (PI): $\text{PI} < 60$ is considered suitable for irrigation while $\text{PI} > 60$ indicates otherwise. During the wet season, PI values vary from 0.003 - 0.24 with mean of 0.060, and in dry season, 0.004 - 0.04 with mean of 0.020, and were all below 60 meq/L. Therefore, waters in the study area are suitable for irrigation purposes in respect to PI.

Total Hardness (TH): Computed total hardness (TH) of waters from the study area vary between 3.90 - 128.00 (mean 47.8) in wet season and between 5.00 - 297.00 with mean value of 52.20 in the dry season, respectively. Based on TH, all water samples fall under the excellent to good categories across seasons and therefore safe for irrigation.

TABLE 1. Statistical Summary of Calculated Irrigation Indices of water samples from the study area.

SEASON	STATISTIC	SAR (meq/L)	RSC (meq/L)	RSBC (meq/L)	Mg ⁺ Ratio (meq/L)	Percentage Na ⁺ (%)	KR (meq/L)	EC (μS/cm)	PI (%)	Total Hardness (meq/L)	Todd's Ratio (meq/L)	Chloro-Alkaline Indices (meq/L)	
												CAI-I	CAI-II
WET	MIN	0.03	1.07	0.24	3.2	2.5	0.07	10	0.003	3.9	0.04	1.0	-0.03
	MAX	0.69	7.17	6.19	36.8	88.24	1.76	310	0.24	128	4.28	0.99	2.34
	MEAN	0.12	3.67	2.57	22.19	26.54	0.38	124.02	0.06	47.8	1.74	0.24	1.034
	SDEV	0.151	1.958	1.748	8.615	22.658	0.409	111.19	0.0733	40.137	1.258	0.45	0.723
DRY	MIN	0.08	0.25	-1.19	-56.1	3.12	0.02	10	0.004	5	0.11	-2.39	-0.31
	MAX	0.85	6.49	3.03	380.0	79.59	3.04	448	0.04	297	24.92	0.98	5.62
	MEAN	0.39	1.88	0.729	44.14	33.10	0.547	138.51	0.020	52.2	3.80	0.42	1.719
	SDEV	0.265	1.615	1.094	88.319	21.246	0.761	133.15	0.0105	67.98	5.52	0.934	1.162

Chloro-Alkaline Indices (CAI): In wet season, most of the water samples have positive Schoeller index values while in dry season, they have negative Schoeller index. Basically, the positive index indicates that in these waters, there is a possibility of exchange of Na^+ and K^+ ions with Mg^{2+} and Ca^{2+} and are therefore safe for irrigation purposes in respect to chloride content.

SAR: Richards, (1954) suggested 5 classes; SAR values <10 = Excellent, $10-18$ = Good, $18-26$ = Fair, >26 = Poor class. Water samples from the study area in both seasons have SAR values less than 10meq/L which falls within the Excellent class meaning the waters are free from any sodium hazard and therefore suitable for irrigational purpose.

RSC: Richards, (1954) proposed that; RSC values <1.25 = Safe, RSC values ranging from $1.25-2.5$ = Marginal, RSC values >2.5 = Unsuitable. RSC values of water samples vary from 1.07 to 7.17 with mean of 3.67 in wet season, and 0.25 to 6.49 with mean of 1.88 in dry season. RSC values in samples BW1,3,7, SW1-6, and WW2-3 and BW1,3 and 4 in rainy and dry seasons respectively, were of the unsuitable category and requires treatment in respect to these parameters, while others were suitable for use.

RSBC: Values of RSBC for waters during the wet season varies from $0.24 - 6.19$ with mean of 2.57 , and $-1.19 - 3.03$ with mean of 0.73 in dry season, respectively. RSBC values indicates that all waters from the area across seasons falls under good category and are safe for irrigation. Positive residual sodium bicarbonate (RSBC) value indicates that dissolved calcium and magnesium ions are less than that of carbonate and bicarbonate contents.

%Na: Values of $\% \text{Na}^+ < 20$ = Excellent, $\% \text{Na}^+ 20-40$ = Good, $\% \text{Na}^+ 40-60$ = Permissible, $\% \text{Na}^+ 60-80$ = Doubtful, and $\% \text{Na}^+ > 80$ = Unsuitable. Computed $\% \text{Na}^+$ of waters from the study area vary between $2.50 - 88.24\%$ with mean of 26.54% in wet season and between $3.12-76.59\%$ with mean of 33.10% in the dry season. The $\% \text{Na}^+$ values of all water samples from the study location falls under the excellent to good categories across seasons.

MR: Values of calculated MR in water samples from the study area range from 3.20 to 36.8 with a mean of 22.19 in wet season and -56.10 to 380.00 with mean of 44.14 in waters in the dry season samples. All values of MR for water samples from the study area across seasons were less than 50 and falls within the permissible limit for irrigational purposes except SW3.

KR: Kelly, (1963) proposed two major categories of water based on Kelly ratio; $\text{KR} > 1$ indicates excess Na^+ , while $\text{KR} < 1$ indicates less Na^+ and the water suitable for irrigation. Computed KR ranges between $0.07 - 1.76$ with mean of 0.38 , and $0.02 - 3.04$ with mean of 0.55 in wet and dry seasons, respectively. KR values of most water samples in both seasons were <1 , hence, suitable for irrigational, others with high KR values are not safe for use and requires treatment.

Todd's Ratio: Todd's ratio ranging from $0.5-1.3$ = low salinity hazard, $1.3-2.8$ = medium salinity hazard, $2.8-6.6$ = high salinity hazard, and $6.6-15.5$ = very high salinity hazard. Across seasons, most water samples had computed Todd's ratio between the low salinity hazard and the medium salinity hazard categories indicating that waters are safe for irrigation purposes.

Wilcox Plot: In wet season, 22.2% falls within the Excellent – good zone, 66.7% are within the Good – Permissible class and 11.1% are unsuitable while in dry season, 23.5% falls within the Excellent – Good zone, 70.6% are within the Good – Permissible class while, 5.9% are unsuitable category. Generally, from Wilcox plots, greater percentage of waters fall within the Good – Permissible class indicating that water resources across seasons in the area are good and permissible for irrigation use.

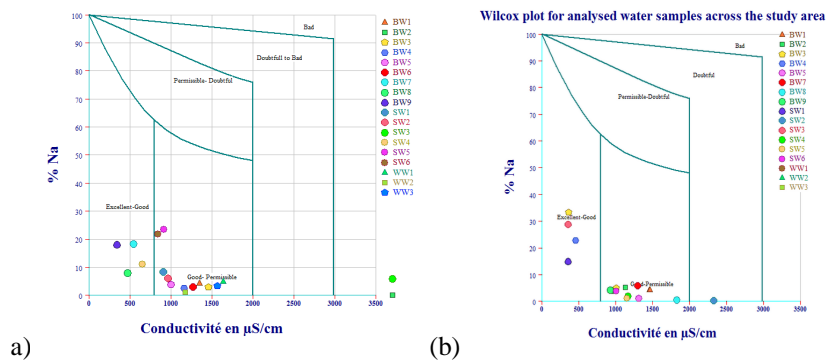


Fig. 3: Wilcox plot for analysed water samples across the study area (a) rainy and (b) dry season.

CONCLUSION: The physicochemical parameters of both ground and surface waters within the area were assessed and confirmed to be below and within the permissible limits for irrigational suitability. The study revealed that SAR, MR, PI, Kelley's ratio, TH, EC, and Ion Exchange values of all sampled waters across the seasons were of the excellent categories for irrigation standard guidelines. Indices such as %Na, RSC, RSBC, Wilcox plot +-+and Todd's ratio were of the permissible category across the seasons. Therefore, it was concluded that on a general scale, the overall quality of all sampled ground and surface waters from the study area were within the safe limit as compared to standard safe water quality norms for irrigation and therefore safe for use by farmers.

RECOMMENDATION: Suitability check should be a routine, extended beyond the study area and for areas of no compliance, appropriate prevention and treatment should be carried out.

REFERENCES

- Ademorati, C.M.A. (1996). *Environmental chemistry and toxicology*. Ibadan, Nigeria: Foludex press.
- Edet, A. E.& Okereke, C.S. (1997). Assessment of hydrogeological conditions in basement aquifers of Precambrian Oban Massif, Southeastern Nigeria. *Journal of Applied Geophysics*, 36, 195–204.
- Ekwere, A.S. & Edet, B. B. (2021). [Temporal variations of heavy metals in sediment, soil and dust particulates across the rock quarrying districts of the Oban Massif, South-eastern Nigeria](#). *Environmental Nanotechnology, Monitoring & Management*; 15: 100431, Elsevier.
- Ekwere, A.S & Edet, B. B., (2023). Hydrogeochemistry of groundwaters within the rock [quarrying districts of the Oban Massif, South-eastern Nigeria](#). *Indian J. Environmental protection*, 43 (2): 108-118.
- Ekwueme, B. N. (2003). *The Precambrian geology and evolution of the Southeastern Nigerian basement complex*. Calabar, Nigeria: University of Calabar Press, 135pp.
- Ibraheem, A. M. & Mazhar Nazeem Khan, S. M. (2017). Suitability Assessment of Groundwater for Irrigation Purpose in Veppanthattai Block, Perambalur District, Tamil Nadu., *World Scientific News WSN* 81(2) (2017) 81-93
- Kavidha R. (2012). Groundwater potential evaluation In erode district, tamilnadu, using Gis and remote sensing. Thesis, 86.
- Kelly, W. P., (1963). Use of Saline Irrigation Water. *Soil Sci.*, 95(4): 355-391
- Narayanan, P. (2007). *Environmental Pollution: Principles, Analysis and Control*. New Delhi, India: LBS Publishers & Distributors.
- Raihan, F. and Alam, J. B. (2008): Assessment of Groundwater Quality in Sunamganj of Bangladesh. *Iranian Journal of Environmental Health Science & Engineering*, , Vol. 5, No. 3, pp. 155-166.
- Richards L. A. (1954). Diagnostics and improvement of saline and alkaline soils. U.S. Dept. of Agriculture handbook no. 60. U.S. Salinity Laboratory, Washington, DC
- Todd, D.K. (1995). *Groundwater Hydrology*. John Wiley and Sons Publications, 3rd Ed, New York.
- Wilcox L.V. (1948). The quality of water for irrigation use. US Department of Agricultural Technical Bulletin 1962, Washington.

Statistical and Decadal Prediction of Temperature Trend for 121 Years (1901-2022) Using ARIMA (2,2,1) Model in Calabar Nigeria.

^aMoses. A. Okono, ^{a,b}Igwe O.Ewona, ^aEmmanuel A Awak, ^{a,b}Jonathan E. Osang

mosesokono3@gmail.com, igweewona@gmail.com, emma4xst@gmail.com,

jonathanosang@unicross.edu.ng

^aPhysics Department, University of Calabar, Calabar, 540242, Nigeria

^bPhysics Department, University of Cross River State, Calabar, 540242, Nigeria

ABSTRACT

Temperature variation has been used extensively as a major index in determining climate change. The study was conducted in Calabar and satellite data were obtained from the database of climateknowledgeportal.worldbank.org/ for temperature for 121 years (1901 to 2022). The study analyzes temperature trends using time series models. Statistical tools such as Mann-Kendall trend test are dependable method often used to identify the behavior of time series data while the ARIMA model were applied to predict the temperature trend for 20 years. Mann-Kendall trend test indicates an upward trend and with a large Z-value of 4.54 suggesting a very strong increasing trend and an extremely small p-value of $5.5e-6$ far below the common 5% significance level reflect the trend is highly statistically significant. Using ARIMA (2,2,1) modeling, the results also reveal an increasing temperature trend over the past century. Using this model, temperature projections for the next decades indicates a rise from 26.30°C in 2022 to 27.40°C by 2032 and 28.10°C by 2042. The result from ARIMA (2,2,1) model shows an increase in annual mean temperature from 26.3°C in 2022 to 27.4°C by 2032 and 28.1°C by 2042, making an overall possible increase of 1.8°C in the next two decades. This predicts a temperature increase at the rate of 6.84% by 2042. These findings emphasize the influence of anthropogenic factors, such as greenhouse gas emissions and land use changes on climate change. And these changes will definitely have severe consequences on climate change with severe repercussions on the environment, especially on agriculture, thus threatening food security and other socioeconomic implications. The knowledge gained from this study could be very important while planning and mitigating the impacts of climate variability in the region.

Keywords: Temperature; ARIMA model; Mann-Kendall trend test; Greenhouse Gases; Climate Change

1.0 Introduction

Understanding and predicting temperature trends is crucial in a region like Calabar, Nigeria, where climate variability has significant implications for agriculture, health, and socio-economic activities.[1] As climate change continues to impact weather patterns globally, there is a growing need for accurate forecasting models that can anticipate future temperature trends and help mitigate the associated risks [2]. Among others, some of the most common approaches to time series forecasting are the ARIMA and the SARIMA models, which have been found efficient in the capture of the trend in climate data.

Human-induced climate change, through activities involving the use of fossil fuels and the clearing of forests, is now causing soaring temperatures, violent weather, natural calamities, extinction of species, and reduced productivity in crops. [3] Conference of Parties (COP 21) held in Paris in 2015 propose to limit this rise in temperature by 2°C to 1.5°C by 2100 compare to the pre-industrial level as the world is heading to $3-4^{\circ}\text{C}$ by the end the century. COP 26proposes decarbonization and deforestation to maintain the temperature at 1.5°C by 2050 though, the present trend indicates a rise up to 2.4°C by 2050 [4].

General climate variability already has a discouraging impact on crop production in Sub-Saharan Africa. These could be buffered by mitigation practices like crop diversification and irrigation, Certain climate-smart agricultural practices not only enhance productivity but also reduce emissions. In Konso, Ethiopia has significantly improved crop yields amidst climate variability. [5].

Ayoade [7] analyzed SARIMA models to forecast annual minimum and maximum temperatures, and rainfall in Ijebu Ode City, southwestern Nigeria, using data from 1989 to 2018. The study identified SARIMA (4,1,1)(1,1,1)₁₂ for maximum temperature, SARIMA (1,1,1)(1,1,1)₁₂ for minimum temperature, and SARIMA (2,1,1)(0,1,1)₁₂ for rainfall as the best-fitting models. The forecasts predict rising temperature and rainfall trends from 2019 to 2028, suggesting significant implications for regional economic and agricultural productivity. The research underscores the value of SARIMA models in guiding future weather-related planning strategies in Ijebu Ode City and its surroundings.

Shitu et al. [8] developed a forecasting model using the SARIMA(4,1,6)(2,1,2)₁₂ model for monthly mean temperature in Jos city, analyzing data from 1986 to 2023. The model, validated through MAE, RMSE, MAPE, and R², was identified as the most reliable for temperature prediction in the area.

M.A. Okono et al., 2022[10] investigated the relationship of meteorological conditions with solar energy reception in southern Nigeria using data across 17 locations in three eco-climatic zones from 2005 to 2016. The analysis, through the Mann-Kendall trend test, indicates that no significant annual solar irradiation and temperature trends occur. However, a strong relationship between the two variables does exist, indicating that regional climate and vegetation exerted a major impact on their variation. It will also be presenting findings that could aid in decision-making for the adoption of solar energy technology in the region

Adams et al. [9] applied a SARIMA model to Osun State's monthly rainfall data, identifying SARIMA (1,0,1)(1,1,1) as the best fit after seasonal differencing. The analysis, supported by ACF and PACF plots, revealed seasonality. Despite an initial model showing an insignificant autoregressive parameter, the final model was used successfully for forecasting.

This study focused on analyzing the temperature variation and predicting future temperature in Calabar, Nigeria (mangrove swamp). Annual trends of temperature in Calabar will be plotted for the period of 122years (1901-2022). The Mann-Kendall test has been applied to bring out the significance of

the trend of increase/decrease of temperature while ARIMA (2,2,1) will be adopted to predict the future temperature in Calabar. Gaussian distribution and box plots have also been carried out to study the nature of distribution.

2.0 Methodology

Trend analysis is a statistical procedure that determines whether data values over time are going up, down, or staying constant. [10]. There are several types of properties possessed by the data, such as seasonal variation among others, which may make identifying the trend difficult. All these widely used trend detection methodologies have certain limitations and prerequisites for validity, such as independence of data and normality of distribution.[11]. The Mann-Kendall test, though effective with increased sample size, may be unreliable in specific cases, such as when there is a high variability within a time series or when the pre-whitening test is applied. Stationarity test is adopted to remove the seasonality from the data. [12]

The partial autocorrelation function (PACF) and autocorrelation function (ACF) is use to choose the most perfect lags to fit into the ARIMA model. [13] The Box-Jenkins ARIMA (2,2,1) Model which has been shown to have several advantages over others have been adopted.[11]Other methods used in the analysis of the data obtained include the use of box plots. Kernel density Estimation-KDE- plots have also been used to explain the distribution of the data better.[14]

2.1. Theoretical Review

Box plots depict the statistical distribution of data. Since the KDE does not provide a detailed view of data, it gives more about the skewness of the data.[14] Box plots are represented in flow;

Lower extreme/minimum ($(Q1 - 1.5 \times IQR)$), lower quartile ($Q1$), median ($Q2$), upper quartile ($Q3$) and upper extreme

($Q4 = Q3 + 1.5 \times IQR$).

KDE is one method of estimating features and distribution in a set of observations. When using KDE, an idea about the probability density function for the normal distribution is taken into consideration: [15,16]

For a univariate KDE, the PDF is given as [17].

$$f(x/\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \quad (1)$$

Where, σ is the standard deviation and μ is the mean.

In equation (1), the ‘argument’ exponential function $\frac{1}{2\sigma^2}(x-\mu)^2$ is a quadratic function of x . This is a parabola which points downward as very well seen from Gaussian/normal distributions); this is because the quadratic function has a negative coefficient [17]

The coefficient of the function from equation is a subject in front of the exponential function which does not depend on. Normalization factors are given by;

$\frac{1}{\sqrt{2\pi}\sigma^2}$ from equation (1) is the coefficient of the function which does not depend on x . The normalization factors are;

$$\frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] = 1 \quad (2)$$

Equation(2) can be simplified by assuming a standard deviation (σ) of 1 and a mean (μ) of 0.

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (3)$$

Equation (3) is the breakdown of the one-dimensional (1D) Gaussian kernel function which has been employed in this study.

2.2. The Mann-Kendall Trend Test

The Mann-Kendall test is another important non-parametric test widely used by the researchers in time-series data. The most important advantage of this test is that it does not require any particular distribution of data, and thus it can be applied for a wide range of datasets [18]. It was developed to work out linear trends in the data. The MK test, in general, is thus an important analysis technique that maps out trends and patterns of time-series data for use in underpinning various scientific inquiries. [19] This test will be applied within a given range of data. x_j agrees with the equation below;

$$x_j = f(t_i) + \varepsilon_i, \quad (4)$$

$f(t_i)$ is a function of continuous increasing or decreasing monotonically, ε_i are the residuals.

$$S = \sum_{k=1}^{n-1} \sum_{j=\lambda=1}^n \text{sgn}(x_j - x_k), \quad (5)$$

where,

$$\text{sgn}(x_j - x_k) = \begin{cases} +1; & \text{if } (x_j - x_k) > 0 \\ 0; & \text{if } (x_j - x_k) = 0 \\ -1; & \text{if } (x_j - x_k) < 0 \end{cases}$$

(6)

The number of values is represented by n. A positive S value indicates an upward variation, and a negative S value signify decreasing trend variation.

To calculate the value of the Z statistic, the variance of S' VAR(S) is used [20]

$$VAR(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t+5) \right] \quad (7)$$

g is the number of tied groups in the series (showing that the test takes the tied or equal values into account). The test statistic Z is obtained using the values of $VAR(S)$ and S [20];

$$Z = \begin{cases} \frac{S-1}{\sqrt{VAR(S)}}; & S > 0 \\ 0; & S = 0 \\ \frac{S+1}{\sqrt{VAR(S)}}; & S < 0 \end{cases} \quad (8)$$

2.3. ARIMA model

ARIMA stands for Autoregressive Integrated Moving Average and is the general class of statistical models used in time series analysis for forecasting. ARIMA bases its forecast of future values on a times series 'past values' and/ or forecast errors. A non-seasonal ARIMA model is classified as an "ARIMA (p, d, q)" model where p refers to the number of autoregressive terms, d refers to the number of non-seasonal differences and q refers to the number of lagged forecast errors (moving average) in the prediction equation. A process, is said to be ARIMA (p, d, q). [22]

2.3 The Box-Jenkins ARIMA Model

The Box-Jenkins methodology refers to the set of procedures for identifying, fitting, and checking ARIMA models with time series data. Forecasts follow directly from the form of the fitted model. By Box-Jenkins, a p th order autoregressive model. [23] AR (p);

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (9)$$

From the model above, the future values Y is dependent of previous lagged values of Y , regression of Y_t on $Y_{t-1}, Y_{t-2}, \dots, Y_{t-P}$, P = order of AR; current value of Y is dependent on how many previous lagged value of current Y . If $P = 2$, that means Y_t is dependent on Y_{t-1} and Y_{t-2} , P from *PACF* [24]

Moving Average

MA(q)

$$Y_t = c + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (10)$$

The future values of Y is dependent of previous lagged values of white noise that is the irregular component. White noise is just the error is the difference between the actual value and predicted value. So, we take into consideration the error also to predict the future value, autocorrelation between the errors, *ACF* gives q [25]

Integration

d(I)

$$Y_t = Y_t - Y_{t-1} \quad (11)$$

Integrated means number of times we difference the data when we have to integrated it back to get the original series, we differentiate to remove trend and seasonality to it stationary series as only after making a series stationary then, we can implement AR, I and MA.[26]. The combined model (ARIMA) is given as;

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (12)$$

3.0 Study location

The capital of Cross River State is Calabar, which is situated at latitude 4.970N and longitude 8.300E, in the southern part of Nigeria. The city is characterized by a tropical climate with marked rainy and dry seasons.[16] Normally, the dry season runs from November to March, while the rainy season runs from April to October. This study has procured from <https://climateknowledgeportal.worldbank.org> annual temperature data for 122 years.and this data will be analyzed using ARIMA (2,2,1) model. The aim is to investigate the trend of temperature variation and prediction of future trend of temperature. The findings from the study are expected to provide a basis for future trend of temperature.

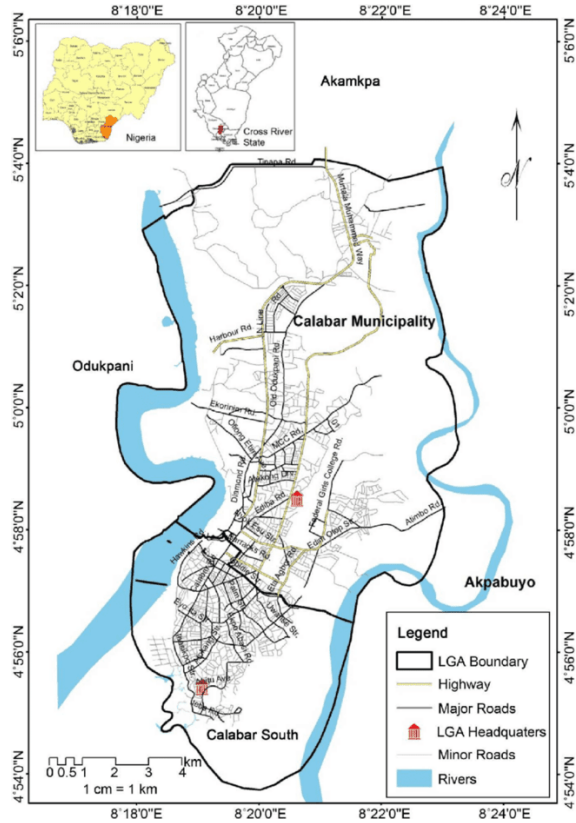


Figure 1. Map of Calabar.

Building ARIMA model

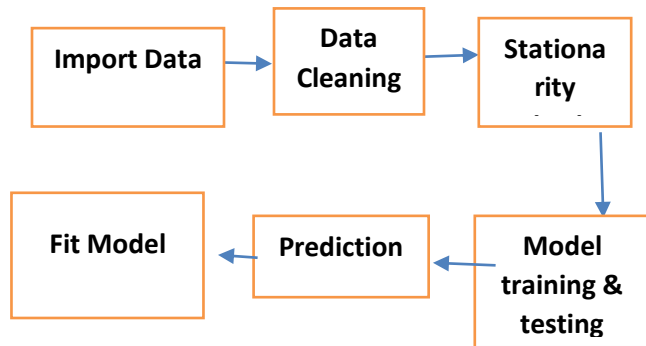


Figure 2. Building ARIMA model

4.0 Results

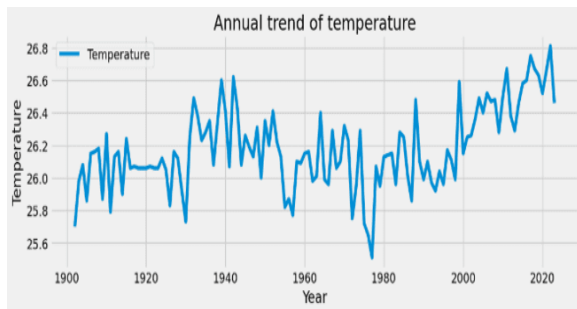


Figure 3. Annual trends of temperature

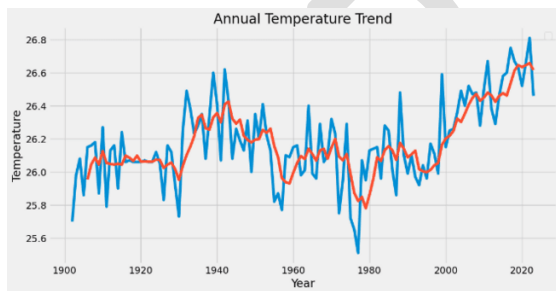


Figure 4. Rolling mean of Annual temperature trend.

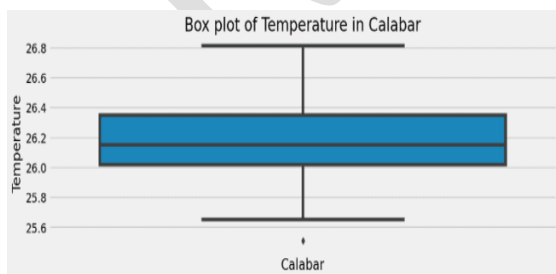


Figure 5. Box plot showing temperature distribution

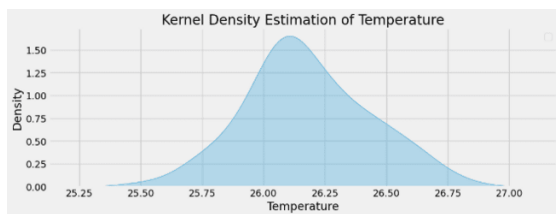


Figure6. KDE plot showing temperature distribution

Table 1

The M-K trend test

Location	Test Statistic (Z)	p-value (Two-tailed)	Sen's slope (Q)	Test Interpretation	Trend
Calabar	4.54	0.0000055	0.003	TRUE	Increasing

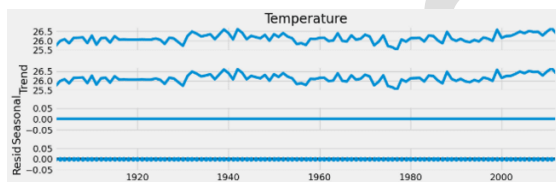


Figure 7. Decomposition of annual temperature

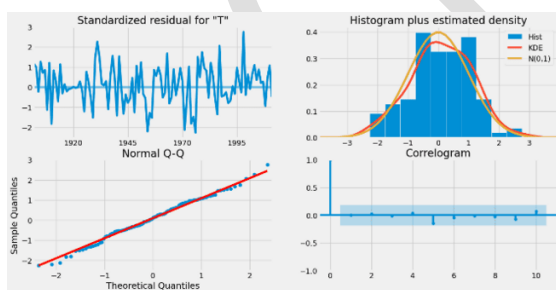


Figure 8. Diagnostics plots of Temperature

Table 2

Stationarity test of temperature

ADF test (Augmented Dickey-Fuller test)

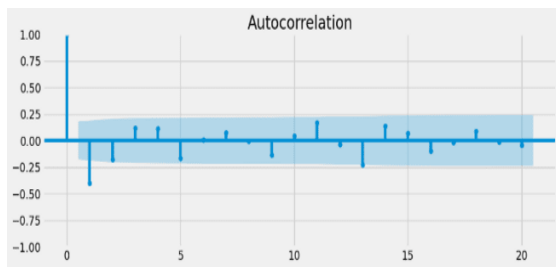


Figure 9. Autocorrelation plot

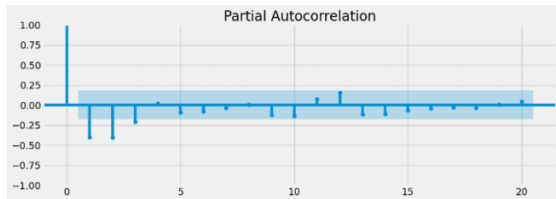


Figure 10. Partial autocorrelation plot

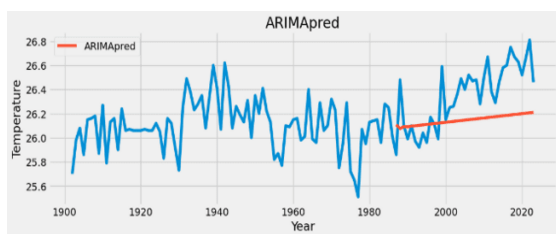


Figure 11. Fitting ARIMA model

mean_squared_error =0.09395011551873507

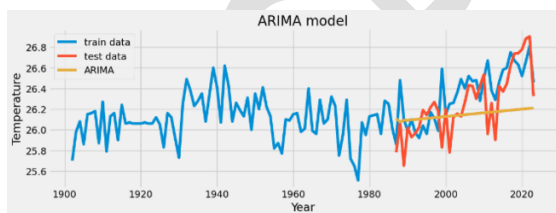


Figure 12. ARIMA model (predicted test data from train data)

Year Forecast

2023-12-31 26.733329
 2024-12-31 26.316430
 2025-12-31 26.862820
 2026-12-31 26.797346
 2027-12-31 26.868436

TEST TYPE	Test statistics	P-value
ADF	-3.819472	0.037083

2028-12-31 27.102729
 2029-12-31 27.253906
 2030-12-31 27.324608
 2031-12-31 27.333422
 2032-12-31 27.470180

2033-12-31	27.534623
2034-12-31	26.888278
2035-12-31	27.376386
2036-12-31	26.943230
2037-12-31	27.522655
2038-12-31	27.490889
2039-12-31	27.556279
2040-12-31	27.859422
2041-12-31	28.026848
2042-12-31	28.154350

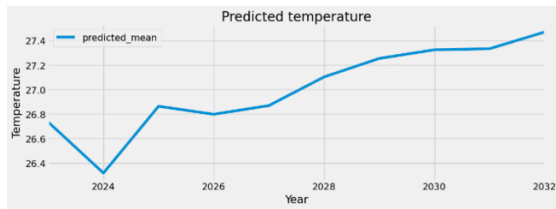


Figure 13. Decadal prediction of temperature

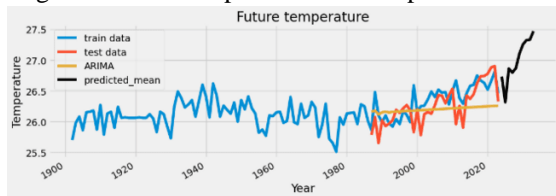


Figure 14. Fitted decadal prediction of temperature

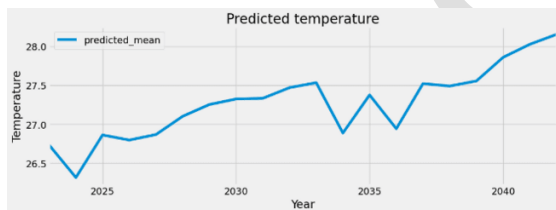


Figure 15. ARIMA model (predicted temperature)

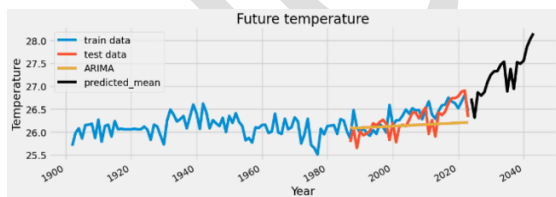


Figure 16. Fitted ARIMA model

4.0 Discussion

Figure 3 and 4 show a line plot of annual temperature trend in Calabar and the rolling mean of temperature. Temperature recorded its lowest value at 25.5°C in 1978 and highest value of 26.8°C in 2021. From 1978 to 2021, temperature lower than or equal to 25.5° have not been observed as the temperature keep increasing.

To visualize the distribution of temperature data in Calabar, the box plot present in figure 5. The result shows that 50% of the data are between the range of 26.0⁰c to 26.4⁰c while the lower part constitute 25% and 25% in the upper whisker.

Figure 6 shows a kernel density estimation (KDE) plot of temperature in Calabar. The univariate Gaussian (normal) distribution is used to show the modal peaks and skewness of the distribution. The aim of this plot is to reveal the median, modal peak and density. The unimodal distribution seems to have its modal peak at about 26.1° and density of 1.6.

The Mann-Kendall trend test

The M-K test has been used to analyze the significance of temperature trend in Calabar with Sen's slope. The result in Table 1 shows the temperature trend in Calabar with a large z-value of 4.54 suggesting a very strong increasing trend and an extremely small p-value of 5.5E-6 far below the common 5% significance level and a positive Sen's slope value of 0.003 indicates a consistent upward trend in temperature suggesting a gradual increase over the period analyzed. At 5% significance level, the temperature variation to accept the alternative hypothesis H_1 and reject the null hypothesis H_0 and after its probability value (P-value) of increasing trend was found to be less than the significance level $\alpha = 0.05$. This can be interpreted as there is an upward trend and its statistical significance.

The Augmented Dickey-Fuller (ADF) test has been adopted to check the stationarity of the data, it has a null hypothesis component and the alternative hypothesis component as shown in Table 2. The test Statistics of -1.819472 indicates stronger evidence against the null hypothesis. The p-value of 0.037083 is less than the commonly used significance level of 0.05, we can reject the null hypothesis with 95% confidence and conclude that the series is likely stationary.

Figure 7 shows the decomposition of the annual temperature into different components. From figure 7, it was observed that the trend of temperature is without seasonality since the data is distributed as the annual mean temperature.

This model consists of the following procedure which includes; model identification, model estimation, model checking and model prediction.

Model identification

The study of ACF and PACF plots are important in fitting ARIMA model, and the objective is to be sure that the plot is stationary also to identify the suitable lag to apply. The ACF and PACF are the correlation of the time series with its lags, after removing the lower-order lags between them. Figure 9 and 10 show that the shaded area are the insignificant areas, the point where line extended out of the insignificant area is what we consider. The PACF plot has a significant spike at lag 1, 2 and 3 therefore, either lag 1, 2 or 3 can be applied, the value 'zero' is not of great importance to this plot. In this study, lag 2 has been adopted which means $p=2$ for the PACF plot. Similarly, the value of ACF plot is 1, therefore, $q=1$. This suggests an ARIMA (p,d,q) model to be (2,2,1). The PACF measured the balance variance of the lag. It identifies whether such lag can be included with the autoregression (AR). Again, the ACF measured the correlation with lag, it judges the moving average model (MA). Therefore, we focus on the two models. The 'd' component of the ARIMA model is the number of times we difference the data to make stationary either one time or more.

Goodness of fit/model verification/model Diagnostics

Ljung-Box test

This approach of time series modelling is good in chosen the best model that fits the data. From Figure 8 above, the standardization residual indicates that the residuals of the model have a mean of zero and constant variance since the residuals are concentrated at -2 to 2. Ljung-Box statistical test shows a p-value of 0.05 and been flat means that the lags is between the 95% confidence interval or the standardized error bounds and all the lags are between the shaded area, it is said that the correlogram is flat and no information that have been left uncaptured. Therefore, ARIMA model is the most ideal and the prediction will be based on this model.

Normality test for residual

The Q-Q plot (quantile-quantile plot) in figure 8 shows that the observation is within the line and point close to the straight line. This imply that the residuals in the model are normal. The histogram shows a bell shape distribution plot of the residuals and it indicate the model is good for prediction.

Fitting ARIMA model using Box-jenkins Approach

This section discusses and analyzes the future temperature trend with the help of an ARIMA model. Figure 11 is a line graph showing annual mean temperature recorded from 1901 to 2022. The upward trend looks really steady from 1978 to 2022; the minimum recorded temperature in this tenure ranges from 25.50° to a maximum of 26.80°. Since 1979, no average annual temperature has been recorded below 25.50°C.

Prediction

Figure 11 show an ARIMA plot with a mean_squared_error of 0.09, the result indicate that the model is good for prediction. In other to confirm if this model can predict future temperature trend, the data were divided into 70% and 30% where the 70% is the train data and the remaining 30% is the test data. This result show that the test data fitted well into the original data, it also indicates that the model is good for prediction.

Figure 16. shows a graph of decadal prediction of annual temperature and the result indicates a minimum temperature of 26.3°C in 2024 and a maximum predicted annual temperature of 27.4°C by 2032. Similarly, prediction of another decadal variation shows that by 2042, annual mean temperature will increase to 28.1°C. These findings agree with Manga and Kalu (2024) for Jos after their result shown that ARIMA model was good for the prediction of temperature trend. Similarly, Fashae *et.al.*, (2019) compare the artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) to model River Opeki discharge (1982–2010) and to use the best predictor to forecast the discharge of the river from 2010 to 2020. The result showed that autoregressive integrated moving average performs better considering the level of root-mean-square error and higher correlation coefficient.

Conclusion

it has been shown that the annual temperature trend for Calabar, the capital of Cross River State in Nigeria, from 1901 to 2022. The study was undertaken to obtain a better view of temperature trends and distribution of data. The Mann-Kendall trend test was adopted for studying the significance of the trend. The ARIMA model of order (2,2,1) is used to predict future temperature data. The following results are outlined.

- ❖ The trend of the annual temperature of Calabar was increasing with much significance from the Mann-Kendall trend test.
- ❖ The result from the distribution plots (Box plot and KDE plot) shows that the

- density of the temperature data is between 26.0⁰c to 26.3⁰c for the period 122 years (1901-2022).
- ❖ The result from ARIMA (2,2,1) model shows an increase in annual mean temperature from 26.3⁰c in 2022 to 27.4⁰c by 2032 and 28.1⁰c by 2042, making an overall difference of 1.8⁰c for the next two decades.
 - ❖ This result implies that anthropogenic activities greatly influence the change in climate. The greenhouse gases such as CO₂, CH₄, N₂O O₃ and HCFC emitted from industries, transport system burning of fossil fuel, coal mining, and agriculture sector. These gases act like a thick blanket wrapped around the Earth, trapping the outgoing long-wave (infrared) radiation from the earth surface and thus raising the atmospheric temperature. Land use factors such as deforestation, roads construction and urbanization reduce the rate of reflectivity of solar radiation thereby increase the amount of energy budgeted to reach the earth surface.
 - ❖ Consequences of climate change in Calabar is expected to experiences increase flooding and drought in some part of the state. Northern part of the state will experience prolong seasonal drought which will decrease food production and leads to famine, population displacement, conflicts and biodiversity loss. Changes in rainfall pattern with increase temperature and increase in floods and draughts will impact substantially on food security and water availability. Increase in rainfall will lead to soil erosion and water logging of crops as some crops are vulnerable to high temperature.

References

- [1] Umar, D. U. A., Ramli, M. F., Aris, A. Z., Sulaiman, W. N. A., Zaudi, M. A., & Tukur, A. I. (2019). An overview of climate change and variability impact studies in Nigeria. *Arabian Journal of Geosciences*, 12(19), 610.
- [2] Sasanya, B. F., Ademola, A. A., & Adesogan, S. O. (2024). Spatiotemporal trend analysis of precipitation, temperature, and soil moisture in selected agrarian locations within Nigeria—a comparative study. *Arabian Journal of Geosciences*, 17(8), 1-19.
- [3] Dimri, T., Ahmad, S., & Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach. *Journal of Earth System Science*, 129, 1-16.
- [4] Shivanna, K. R. (2022). Climate change and its impact on biodiversity and human welfare. *Proceedings of the Indian National Science Academy*, 88(2), 160-171.
- [5] Gashure, S., Wana, D., & Samimi, C. (2022). Impacts of climate variability and climate-smart agricultural practices on crop production in UNESCO designated cultural landscapes of Konso, Ethiopia. *Theoretical and Applied Climatology*, 150(3), 1495-1511.
- [6] Solaraju-Murali, B., Bojovic, D., Gonzalez-Reviriego, N., Nicodemou, A., Terrado, M., Caron, L. P., & Doblas-Reyes, F. J. (2022). How decadal predictions entered the climate services arena: an example from the agriculture sector. *Climate services*, 27, 100303.
- [7] Ayoade, I. A. (2023). Statistical modelling of temperature and rainfall in Ijebu Ode Nigeria using SARIMA. *Faculty of Natural and Applied Sciences Journal of Scientific Innovations*, 5(2), 55–68. Retrieved from <https://fnasjournals.com/index.php/FNAS-JSI/article/view/220>
- [8] Shitu, D. A., Abdulkadir, A., Abbas, F. U., Gambo, A. M., Mairam, S. A. B., & Mafolasire, S. (2024). Prediction of monthly mean surface air temperature using SARIMA in Jos North, Plateau State, Nigeria. *Bima Journal of Science and Technology*, 8(1b), 358-366.

- [9] Adams, S. O., Mustapha, B., & Alumbugu, A. I. (2019). Seasonal autoregressive integrated moving average (SARIMA) model for the analysis of frequency of monthly rainfall in Osun state, Nigeria. *Physical Science International Journal*.
- [10] Ay M, Kisi O. Investigation of trend analysis of monthly total precipitation by an innovative method. *Theoretical and Applied Climatology* 2015;120:617–29.
- [11] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- [12] Hamed, K. H., & Rao, A. R. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of hydrology*, 204(1-4), 182-196.
- [13] Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). *Time series analysis and its applications* (Vol. 3, p. 4). New York: springer.
- [14] Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.
- [15] Diwekar, U., David, A., Diwekar, U., & David, A. (2015). Probability Density Functions and Kernel Density Estimation. *BONUS Algorithm for Large Scale Stochastic Nonlinear Programming Problems*, 27-34.
- [16] Okono MA, Agbo EP, Ekah BJ, Ekah UJ, Ettah EB, Edet CO. Statistical Analysis and Distribution of Global Solar Radiation and Temperature Over Southern Nigeria. *Journal of the Nigerian Society of Physical Sciences* 2022:588–588.
- [17] Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.
- [18] Phuong DND, Tram VNQ, Nhat TT, Ly TD, Loi NK. Hydro-meteorological trend analysis using the Mann-Kendall and innovative-Şen methodologies: a case study. *International Journal of Global Warming* 2020;20:145–64.
- [19] Kendall MG. Rank correlation methods. London: Griffin; 197
- [20] Mann HB. Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society* 1945:245–59. <http://dx.doi.org/10.2307/1907187>.
- [21] Albeladi, K., Zafar, B., & Mueen, A. (2023). Time Series Forecasting using LSTM and ARIMA. *International Journal of Advanced Computer Science and Applications*, 14(1), 313-320.
- [22] Mwenda, A., Kuznetsov, D., & Mirau, S. (2015). Analyzing the impact of historical data length in non-seasonal ARIMA models forecasting. *Mathematical Theory and modeling*, 5(10), 77-85.
- [23] Box, George EP, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [24] Brownlee, J. (2017). A Gentle Introduction to the Box-Jenkins Method for Time Series Forecasting. *Machine Learning Mastery*.
- [25] VerHoef, J. M., & Peterson, E. E. (2010). A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*, 105(489), 6-18.

- [26] Alfaki, M. M. A., & Masih, S. B. (2015). Modeling and forecasting by using time series ARIMA models. *International Journal of Engineering Research & Technology (IJERT)*, 4(3), 2278-0181.
- [27] Manga, E., & Kalu, D. O. Time Series Analysis and Forecasting of Average Monthly Temperature Levels in Jos, Nigeria.
- [28] Fashae, O. A., Olusola, A. O., Ndubuisi, I., & Udomboso, C. G. (2019). Comparing ANN and ARIMA model in predicting the discharge of River Opeki from 2010 to 2020. *River research and applications*, 35(2), 169-177.

NCPS-2024

CLIMATIC VARIABILITY OVER THE VEGETATION ZONES IN NIGERIA FOR 31 YEARS (1991 – 2022)

^aMoses. A. Okono, ^{a,b}Igwe O.Ewona, ^aEmmanuel A Awak, ^{a,b}Jonathan E.

Osang mosesokono3@gmail.com, igweewona@gmail.com, emma4xst@gmail.com,
jonathanosang@unicross.edu.ng

^aPhysics Department, University of Calabar, Calabar, 540242, Nigeria

^bPhysics Department, University of Cross River State, Calabar, 540242, Nigeria

ABSTRACT

Temperature and precipitation variabilities is of great important in understanding climate change. The study was conducted on fifteen locations across the five vegetation zones, ranging from Humid Forest to arid Sahel Savanna. Satellite data were obtained from the database of climateknowledgeportal.worldbank.org/ for temperature and precipitation for 31 years (1991 to 2022). Statistical tools such as the Mann-Kendall trend test have been adopted to study the trend of time series data, while the linear regression and kernel density estimation have also been adopted to understand the relationship between the two parameters. The Mann-Kendall trend test results have shown the general increase in temperature over all zones with large positive Z-values and very small p-values indicating that the trends are highly statistically significant, with most pronounced upward trend seen over all the zones except for Maiduguri in Sahel Savanna that is not statistically significant. Increasing and decreasing trend for precipitation were observed across the zones in this case the trends were not statistically significant as the p-values were seen to be greater than the common 5% (0.05) significance level. These findings also reveal a weak negative correlation between temperature and precipitation in all the vegetation zones except for Yenagoa with a moderate relationship as temperatures keep increasing along with lower rainfall. Most importantly, the findings have implications in agriculture, water management, and climate adaptation strategies, since rising temperatures combined with shifting rainfall patterns pose risks for drought-prone areas. In understanding these climatic trends, policymakers and other relevant stakeholders will be better positioned to take interventions that reduce the adverse effects of climate variability on ecosystems, agriculture, and livelihoods in Nigeria.

Keywords: Temperature; Precipitation; Mann-Kendall Test; Sen's slope; Kernel Density; Regression; Greenhouse Gases; Climate Change; Vegetation

1.0 Introduction

Climatic variability, at present times, has been a global concern because it influences most of the environmental processes and ecosystems.[1] Climatic fluctuations have become one of the important determinant factors in vegetation zones and affect agriculture, water, and biodiversity in Nigeria. The different vegetation zones extend from the mangrove swamps along the coasts to the savannah and semi-arid zones that lie farther inland (Ita, R. E., & Ogbemudia, F. O. (2023). It will be important to understand the levels in which climate variability occurs and specific impacts it causes on these vegetation zones for developing appropriate response strategies on adaptation and mitigation. [2]

Zaied et. al. 2023 For example, studies have recorded that, in Nigeria, climate variability has induced sudden changes in vegetation cover, which in turn affects agriculture, the supply of water, and public health. [3] Desertification, due to fluctuating rainfall and increased temperature, occurs in northern Nigeria, with implications for agricultural productivity. [4] In southern Nigeria, changes in temperature and rainfall patterns have affected crop and natural vegetation growth, thus affecting food security and the local economy.[5] These climatic changes are causing a shift in the distribution of crops and may therefore alter the livelihood of farmers and communities who rely on agriculture. [6]This study will be important for policymakers, environmental planners, and farmers in devising adaptive strategies to overcome adverse impacts brought about by the variability in climate, such as shifting cultivation, deforestation, and water management systems. [7] Understanding time-varying climatic changes will enable stakeholders to adopt sustainable agriculture, reforestation, and better water management practices that can reduce impacts of climate change. [8].

E.P. Agbo et al., [9] in their 2023 study, discussed the current state and variation of vegetation zones in Nigeria. They have related these changes to climate change, which is largely human-induced through activities like deforestation and greenhouse gas emissions. It covers data from 1981 to 2020, applying the innovative trend analysis after Sen in concert with the Mann-Kendall test to investigate the main environmental variables, including trends of surface temperature, photosynthetically active radiation, relative humidity, and precipitation. The results attained indicate mixed trends; while some of the variables are increasing, others decline. The trend in PAR is insignificant, according to this study. In other words, the authors of this paper conclude that such an increase in these parameters might serve as an indicator of the trend toward greening vegetation in Nigeria and, accordingly, a potential for regional re-greening.

M.A. Okono et al., 2022[10] investigated the relationship of meteorological conditions with solar energy reception in southern Nigeria using data across 17 locations in three eco-climatic zones from 2005 to 2016. The analysis, through the Mann-Kendall trend test, indicates that no significant annual solar irradiation and temperature trends occur. However, a strong relationship between the two variables does exist, indicating that regional climate and vegetation exerted a major impact on their variation. It will also be presenting findings that could aid in decision-making for the adoption of solar energy technology in the region.

Ewona I.O. presented a study in 2014 [11] with the analysis of 30 years of rainfall data from 23 weather stations across Nigeria, from 1997 to 2007. The regression equations, with constants showing trends, varying from -0.006 to 0.230 and coefficients from 35.81 to 228.5 showing the rainfall volume, were presented using ten parameters. The variation in rainfall showed a uniform increase with time. Besides, there was a marked latitudinal effect on the constant b which decreased with latitude at a rate of -18.87 per degree of latitude at a correlation of 0.92. Among the stations, Katsina recorded the lowest rainfall and Calabar had the highest. Overall, a strong latitudinal gradient in rainfall was observed with higher amounts falling in the southern parts.

This study aims to analyze the variation of temperature and precipitation in 15 locations across the vegetation zones in Nigeria (the Humid Forest zone, Derived Savanna, Guinea Savanna, Sudan Savanna and Sahel Savanna). Satellite data were obtained from the database of climateknowledgeportal.worldbank.org/ for temperature and precipitation for 31 years (1991 to 2022). Monthly trend in each vegetation will be plotted and compared with the annual trends. Mann-Kendall and Sen's slope estimator will be adopted to study the significant of increasing or decreasing of the temperature and precipitation.

2.0 Methodology

Various methods have been used in analyzing the data obtained. These include box plots describing data, linear regression plots between temperature and precipitation for better understanding of variation between the two, and Kernel density estimation plots that have also been used for better explanation of distribution of the data.

2.1. Theoretical Review

Kernel Density Estimate (KDE) is a nonparametric estimation method for the probability density function of a random variable. It smoothed the data points by placing a kernel on each data point and summing these kernels to create a continuous probability distribution to visualize the underlying distribution without assuming any particular parametric form. The KDE is given as [12].

$$f(x/\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \quad (1)$$

σ is the standard deviation and μ is the mean.

we see that the argument of the exponential function, $\frac{1}{2\sigma^2}(x-\mu)^2$, is a quadratic function of the variable x . Moreover, the parabola opens downwards, since the coefficient of the quadratic term is negative. [12]

The coefficient in front, $\frac{1}{\sqrt{2\pi\sigma^2}}$, from equation (1) does not depend of x ; therefore, we can consider it only as a "normalization factor" used to ensure that;

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] dx = 1 \quad (2)$$

Equation (1) can assume a standard deviation (σ) of 1 and a mean (μ) of 0.

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (3)$$

The equation (3) is the simplified version of one-dimensional Gaussian Kernel function, which has been used in this study due to its flexibility and capability of providing a smooth representation of distribution without assuming any particular form, which are advantageous in the process of exploratory data analysis.

Linear regression

The dependent and independent relationship between two variables is what the linear regression function demonstrates. For a simple linear regression, we will have one independent and one dependent variable [13].

It is given by

$$Y = \pm mX \pm c + e \quad (4)$$

where Y is dependent variable, X is the independent variable, m gives the value of slope, c gives the value of intercept, and e the residual error. The equation exposit a relationship-linear among the intercept and the slope. Using this equation, the coefficient of correlation, R, between the temperature and the precipitation is worked out. [14]

$$Correl(x, y) = \frac{\sum (x - \bar{X})(y - \bar{Y})}{\sqrt{\sum (x - \bar{X})^2 \sum (y - \bar{Y})^2}} \quad (5)$$

where x and y are the two variables of precipitation and temperature.

The Mann-Kendall trend test

Among several others, the Mann-Kendall trend test is one of the most frequent methods used for time-series data investigation. This is a non-parametric test that depends less on the actual distribution of data and hence is versatile in application on various data sets in general, MK test is an efficient trend and pattern analysis tool in time series data, which can be used to support various scientific studies.[15]

We can apply this test when a given range of data x_j agrees with the relation;

$$x_j = f(t_i) + \varepsilon_i, \quad (6)$$

$f(t_i)$ is a function of continuous increasing or decreasing monotonically, ε_i are the residuals.

The following Mann-Kendall test null hypothesis- H_0 implies no trend, whereas the alternate hypothesis- H_1 indicates there is a trend in series. If the test resultant obtained agrees with the null hypothesis-meaning there is no trend-it means the given data x are randomly ordered in time t, while on the other side the alternative hypothesis indicates either an increasing monotonic or decreasing monotonic trend. [16]

The Mann-Kendall test uses the statistic S, calculated using;

y are the mean of the x and y values respectively. The coefficient of determination can be easily obtained by: just squaring the above equation. An estimate of the error of the regression slope is given by

$$\partial s = \frac{W}{nR} \quad (7)$$

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(x_j - x_k), \quad (8)$$

where,

$$\text{sgn}(x_j - x_k) = \begin{cases} +1; & \text{if } (x_j - x_k) > 0 \\ 0; & \text{if } (x_j - x_k) = 0 \\ -1; & \text{if } (x_j - x_k) < 0 \end{cases}$$

To compute the value of the Z statistic, the variance of S' VAR(S) is used [17]

$$\text{VAR}(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t_p+5) \right] \quad (9)$$

g is representing the number of tied groups in the series (indicating that the test considers the tied or equal values). Using the values of VAR(S) and S the statistic test Z is obtained as

$$Z = \begin{cases} \frac{S-1}{\sqrt{VAR(S)}}; & S > 0 \\ 0; & S = 0 \\ \frac{S+1}{\sqrt{VAR(S)}}; & S < 0 \end{cases} \quad (10)$$

Sen's Slope Estimator

The nonparametric Sen Technique is utilized to determine the incline of a current linear pattern. The linear equation is represented by [18].

$$f(t) = Qt + B \quad (11)$$

The equation describes this as a time function that is continuously monotonically rising or falling, with Q denoting the slope and B denoting a regular. Calculating the slopes of each and every data item pairing yields the slope estimate in equation [19].

$$Q_i = \frac{x_j - x_k}{j - k} \quad (12)$$

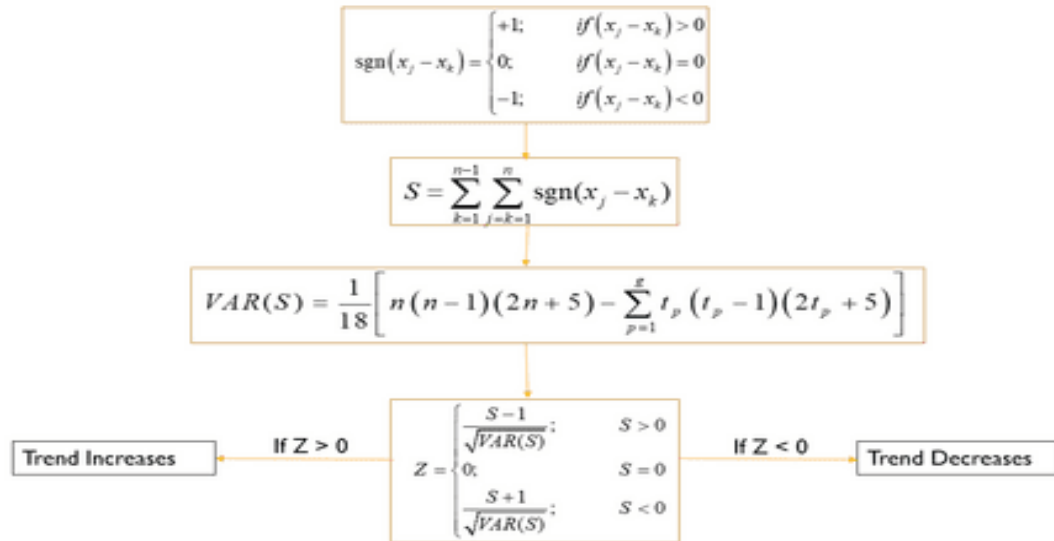


Figure 1. Mann Kendall Trend Test Methodology

Study location

Nigeria is bounded by longitude 2° E and 15° E and latitude 4° N and 14° N, when two important seasons, wet and arid, are considered. The seasons have thus influenced the adoption of different

renewable energy sources; for instance, the various seasons in the country affect hydropower stations due to water availability and variability [21].

Geographically, northern states in Nigeria tend to be hotter than southern states. As such, extreme temperature and radiation values characterize it. [22] Nigeria is situated just above the equator that by default receives the highest sunshine intensity [22]. However, due to the fact that the southern region of Nigeria is close to the Atlantic Ocean and the northern region close to the Sahara Desert, we realize our earlier described sunshine, temperature and solar radiation [23].

Data collection

Monthly and annual data is been obtained from <https://climateknowledgeportal.worldbank.org/> for temperature and precipitation for the period of thirty-two years (1991-2022).

Data visualization

Python is a programming language that has been used as the Analytical tool used in the research.

All graphical displays statistical analysis using a programming language. This Python program includes the packages sklearn, seaborn, matplotlib, pandas, and numpy used for visualizations such as box plots, linear regression, 1D kernel density estimate (KDE) plots, and Mann-Kendall trend test. \bar{x} and \bar{y} are the mean of the x and y values respectively

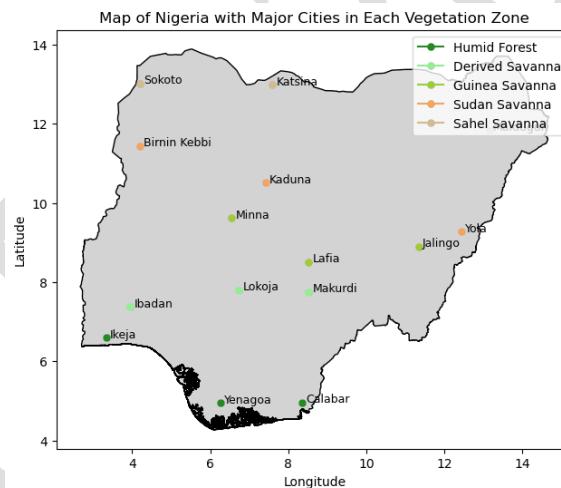


Figure 2. Map of Nigeria showing all the vegetation zones.

TABLE 1; Geographical location for vegetation

			Calabar	4.96N	8.36E
Location	Latitude	Longitude	Derive Savanna		
Humid Forest zone			Ibadan	7.35N	3.87E
			Lokoja	7.81N	6.74E
Ikeja	6.60N	3.34E	Makurdi	7.75N	8.52E
Yenagoa	4.94N	6.26E	Guinea		

Savanna

Minna 9.62N 6.53E

Lafia 8.51N 8.52E

Jalingo 8.90N 11.34E

Sudan Savanna

Birnin 12.47N 4.20E

Kebbi

Kaduna 10.55N 7.41E

Yola

9.27N

12.44E

Sahel Savanna

Sokoto 13.04N 4.20E

Katsina 13.00N 7.59E

Maiduguri 11.85N 13.14

3.0 Results

Monthly trends of temperature and precipitation

Figure 3. Humid forest zone

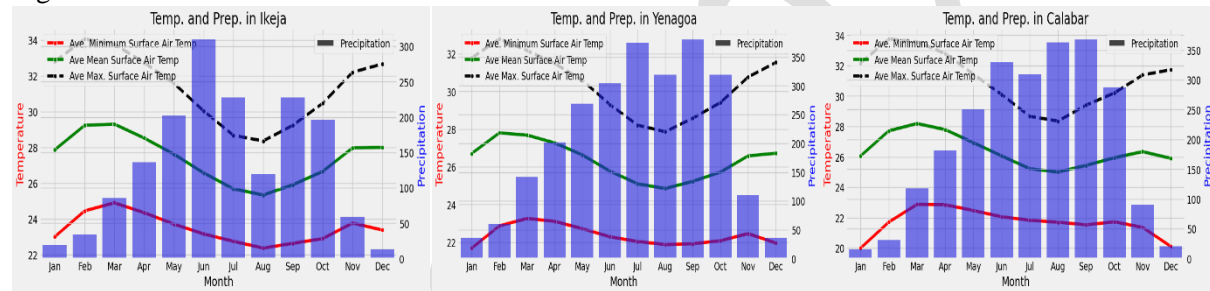


Figure 4. Derived Savanna

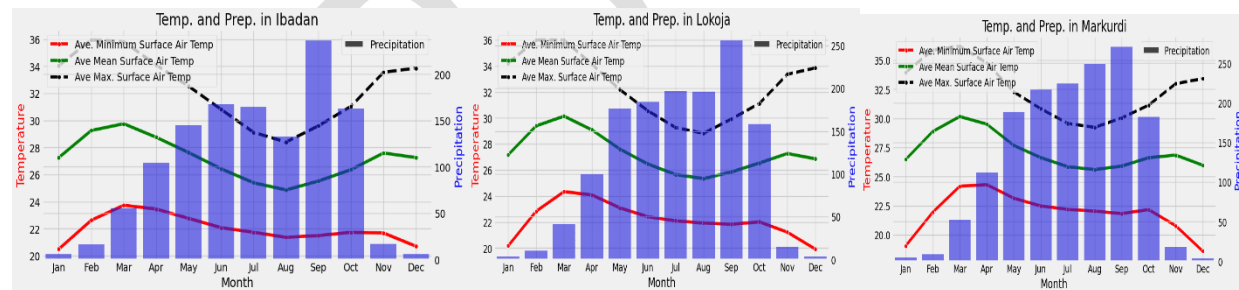


Figure 5. Guinea Savanna

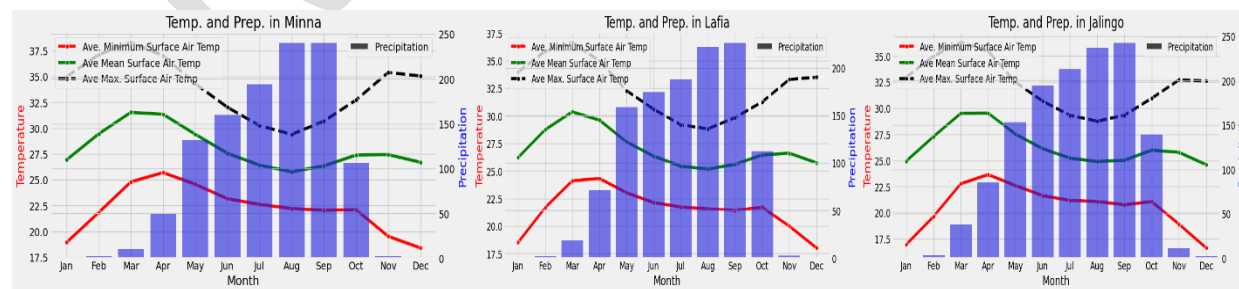


Figure 6. Sudan Savanna

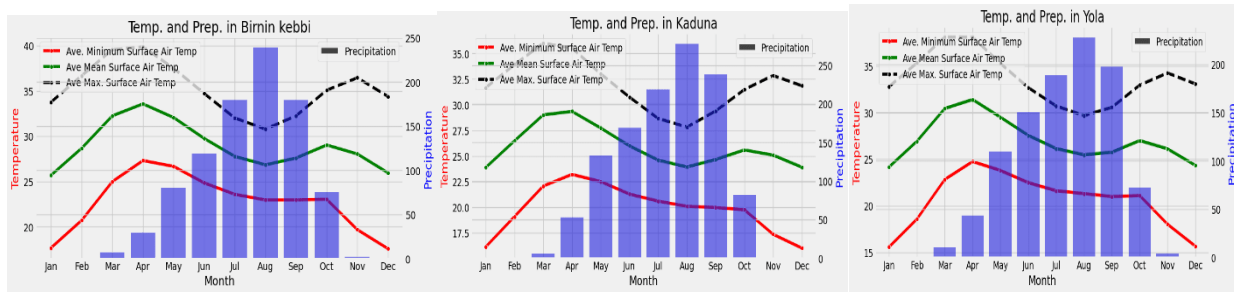


Figure 7. Sahel savanna

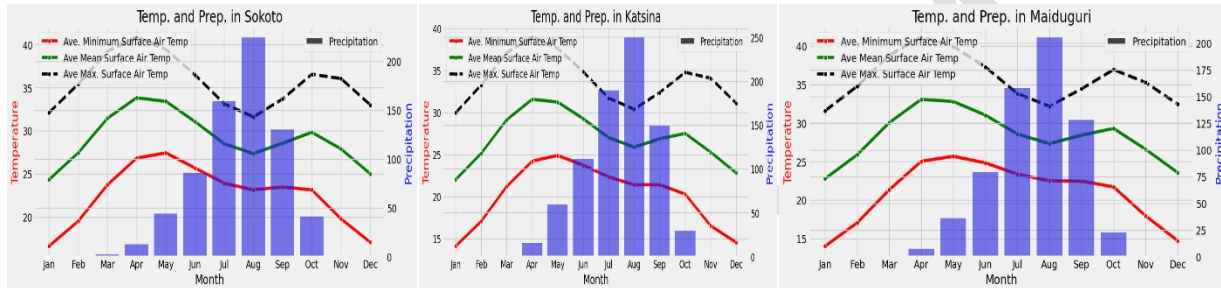
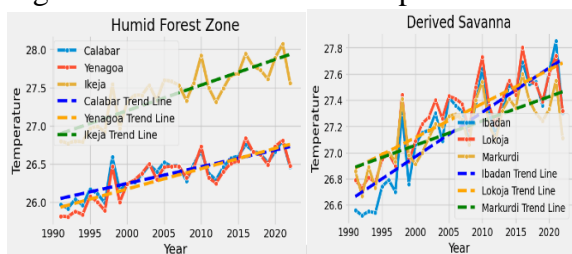


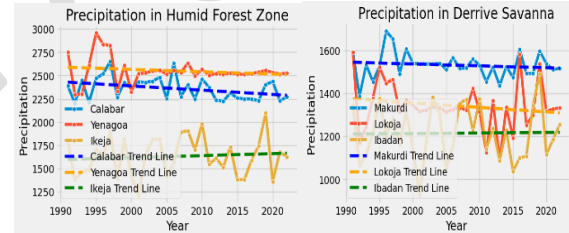
Figure 8. Annual trend for temperature

Figure 9. Annual trend for precipitation



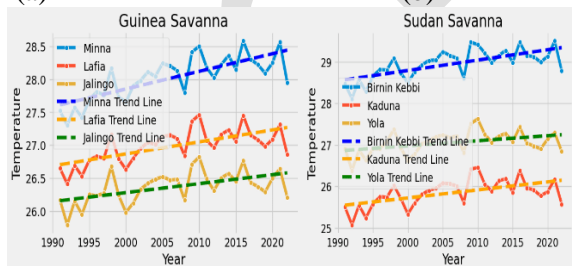
(a)

(b)



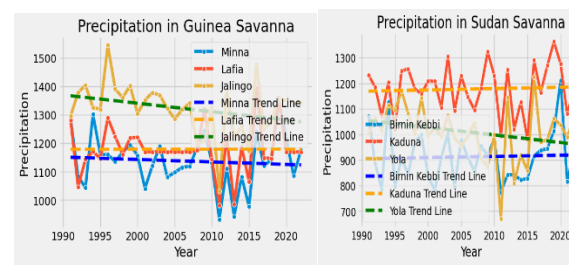
(c)

(d)



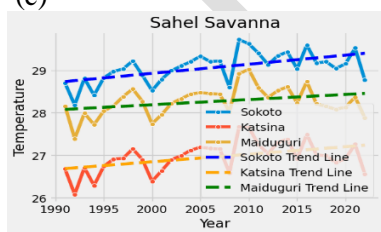
(e)

(f)

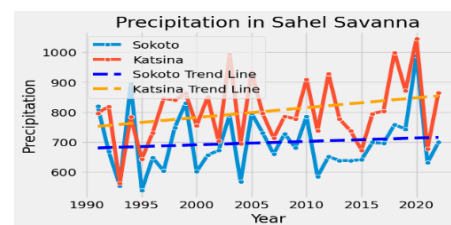


(g)

(h)



(i)



(j)

Mann-Kendall and Sen's slope Test for Temperature from the period of 1991-2022

Location	Kendall's Tau	Mann Kendall's Statistic (S)	Test Statistic (Z)	p-value (Two-tailed)	Intercept	Sen's slope (Q)	Test Interpretation	Trend
Humid Forest Zone								
Calabar	0.63	317.0	5.12	2.957e-07	26.09	0.023	TRUE	Increasing
Yenagoa	0.66	328.0	5.30	1.11e-07	26.00	0.027	TRUE	Increasing
Ikeja	0.65	327.0	5.29	1.213e-07	26.9	0.034	TRUE	Increasing
Derrive savanna								
Makurdi	0.56	280.0	4.52	5.98e-06	26.9	0.02	TRUE	Increasing
Lokoja	0.61	303.0	4.89	9.61e-07	26.9	0.02	TRUE	Increasing
Ibadan	0.67	335.0	5.42	5.81e-08	26.7	0.03	TRUE	Increasing
Guinea savanna								
Jalingo	0.38	191.0	3.08	0.002	26.1	0.01	TRUE	Increasing
Lafia	0.48	243.0	3.92	8.63e-05	26.7	0.01	TRUE	Increasing
Minna	0.57	283.0	4.57	4.79e-06	27.6	0.02	TRUE	Increasing
Sudan savanna								
Yola	0.29	144.0	2.32	0.02	26.8	0.01	TRUE	Increasing
Kaduna	0.39	196.0	3.16	0.001	25.5	0.02	TRUE	Increasing
Beni-kebbi	0.48	241.0	3.89	9.82e-05	28.6	0.02	TRUE	Increasing
Sahel Savanna								

Maiduguri	0.20	100.0	1.60	0.108	28.0	0.01	FALSE	Increasing
Katsina	0.32	160.0	2.57	0.009	26.6	0.02	TRUE	Increasing
Sokoto	0.39	197.0	3.18	0.001	28.7	0.02	TRUE	Increasing

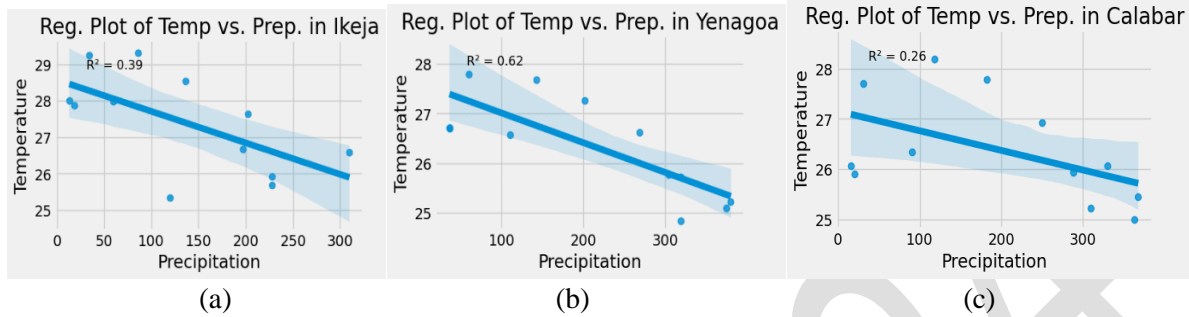
TABLE 3

Mann-Kendall and Sen's slope Test for Precipitation from the period of 1991-2022

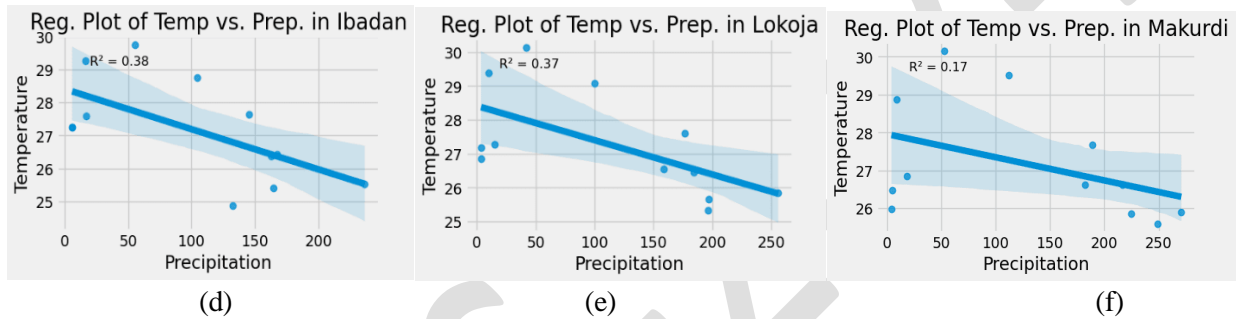
Location	Kendall's Tau	Mann Kendall's Statistic (S)	Test Statistic (Z)	P-value (Two-tailed)	Intercept	Sen's slope (Q)	Test Interpretation	Trend
Humid Forest Zone								
Calabar	-0.22	-112.0	-1.80	0.07	2429.0	-3.97	FALSE	Decreasing
Yenagoa	-0.12	-60.0	-0.95	0.33	2549.0	-1.22	FALSE	Decreasing
Ikeja	0.05	26.0	0.40	0.68	1603.3	1.75	FALSE	Increasing
Derived savanna								
Makurdi	-0.15	-76.0	-1.21	0.22	1544.6	-1.04	FALSE	Decreasing
Lokoja	-0.14	-74.0	-1.18	0.23	1358.2	-1.88	FALSE	Decreasing
Ibadan	-0.008	-4.0	-0.04	0.96	1206.2	-0.10	FALSE	Decreasing
Guinea savanna								
Jalingo	-0.22	-110.0	-1.76	0.07	1374.5	-2.66	FALSE	Decreasing
Lafia	-0.03	-16.0	-0.24	0.80	1170.3	-0.02	FALSE	Decreasing
Minna	-0.04	-24.0	-0.37	0.70	1144.8	-0.73	FALSE	Decreasing
Sudan savanna								
Yola	-0.22	-112.0	-1.80	0.07	1045.4	-2.70	FALSE	Decreasing
Kaduna	0.03	18.0	0.27	0.78	1169.6	0.89	FALSE	Increasing
Birninkebbi	0.03	19.0	0.29	0.77	898.6	0.72	FALSE	Increasing
Sahel Savanna								
Maiduguri	0.08	616.0	1.36	0.17	581.9	0.37	FALSE	Increasing
Katsina	0.15	78.0	1.24	0.21	743.7	3.33	FALSE	Increasing
Sokoto	0.05	28.0	0.43	0.66	657.3	1.23	FALSE	Increasing

Figure 10. Regression Analysis

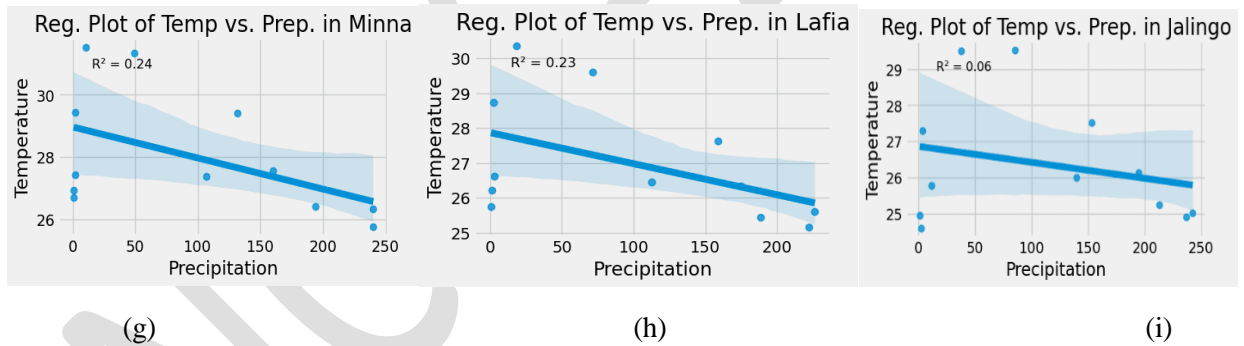
Humid Forest Zone



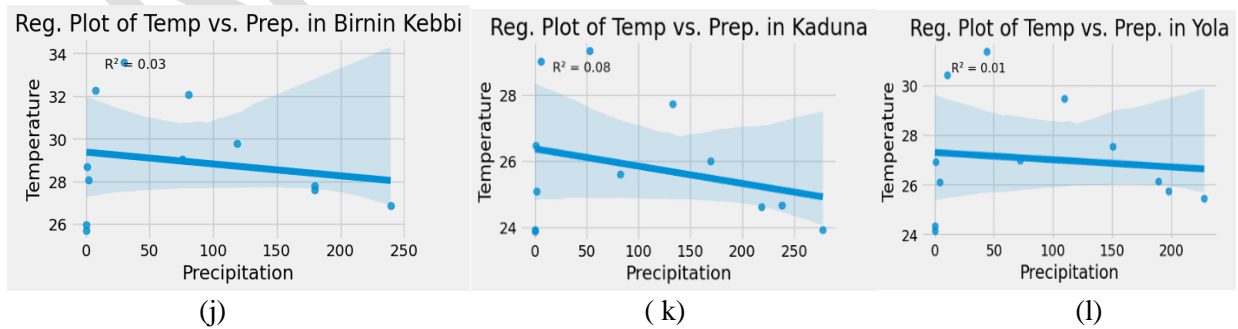
Derrive Savanna



Guinea Savanna



Sudan Savanna



Sahel Savanna

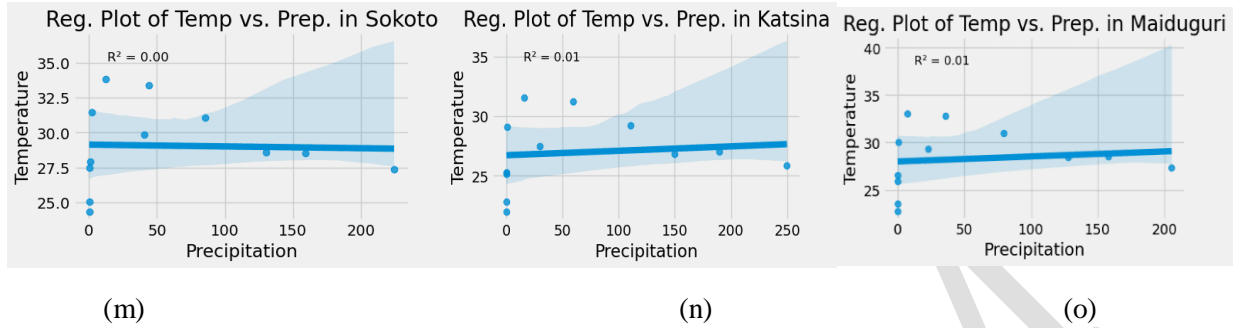
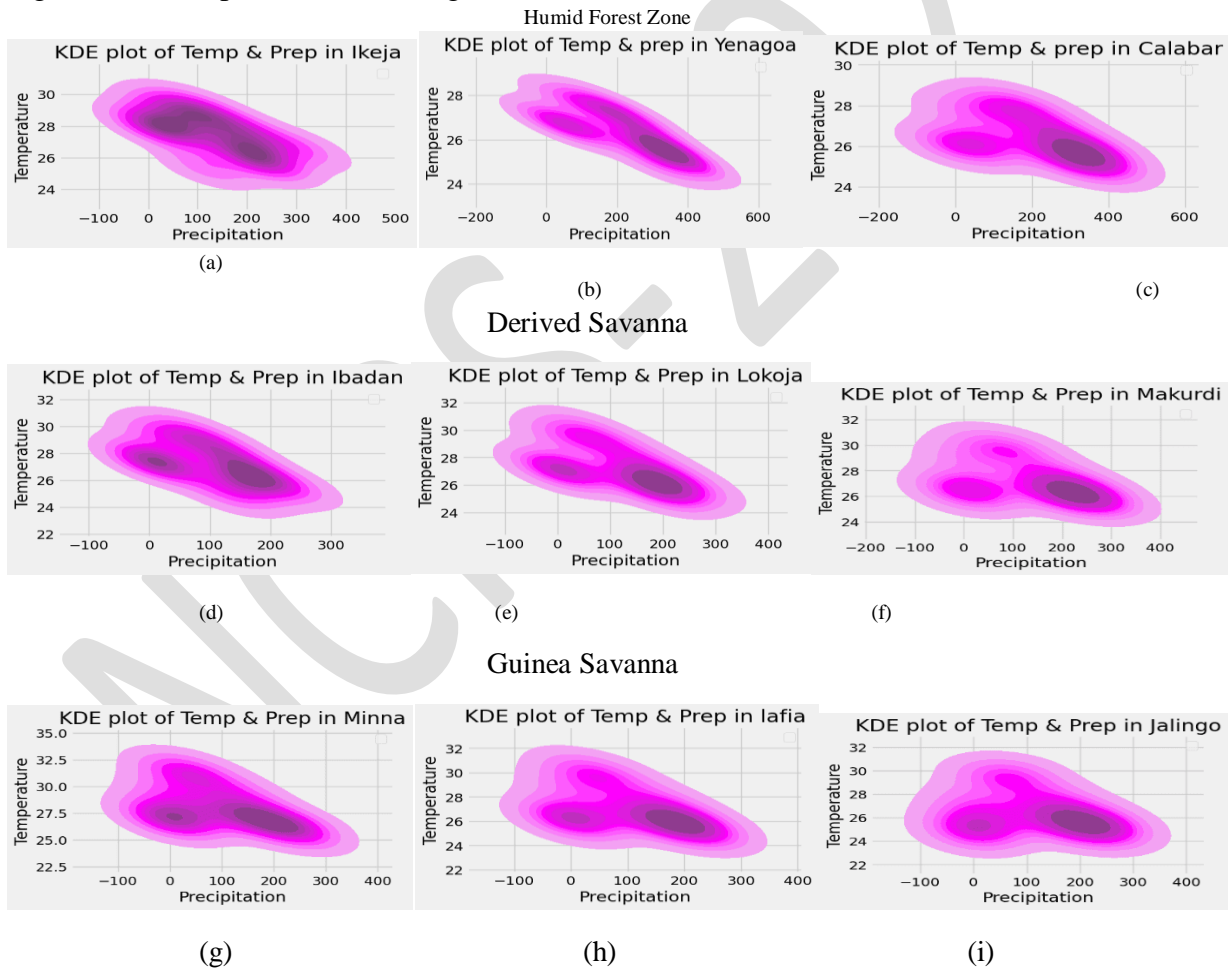
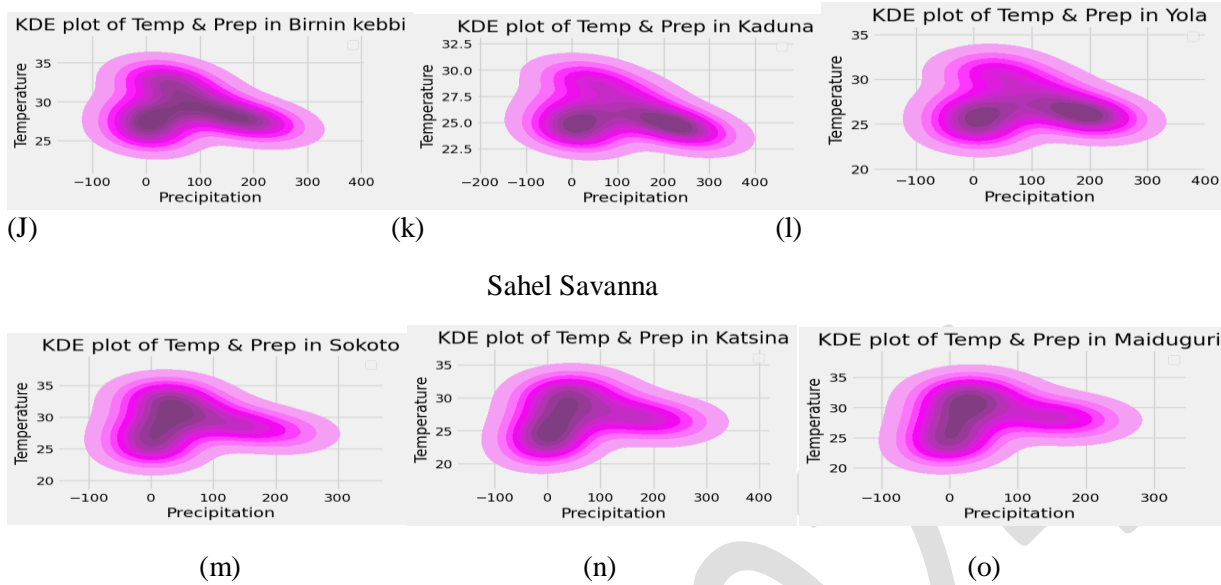


Figure 11. KDE plot for all the Vegetation Zone



Sudan Savanna



Sahel Savanna

4.0 Discussion

Monthly trends of temperature and precipitation

Figure 3 presents the monthly trends of maximum, mean, and minimum temperatures with precipitation in the humid forest zone: Calabar, Yenagoa, and Ikeja. The mean temperature ranges from an average of 25.0°C in August to 28.19°C in March throughout the period under review. Precipitation also ranged from 15.60 mm in January to 367.79 mm in September. Regarding temperatures, the minimum mean temperature is 24.85°C in Yenagoa during the month of August, and its maximum is 27.80°C during February. Similarly, the minimum and maximum precipitations stand at 35.94 mm in January and 380.44 mm in September, respectively. In Ikeja, the minimum temperatures range from 25.35°C during the month of August to 29.31°C during the month of March. Precipitation is high during the month of June at 309.81 mm while the lowest is during the month of December at 13.02 mm . These data give an indication that in this zone, temperature and precipitation are inversely related; that is, lower temperatures correspond to higher precipitation.

Figure 4 shows the Monthly variation for maximum, mean and minimum temperature and precipitation for Derived Savanna region (Makurdi, Lokoja and Ibadan) The lowest and highest mean temperatures for Makurdi vary from 25.60°C in August to 30.16°C in the month of March, while the precipitations vary from 4.49 mm in January to 270.74 mm in September. Similarly, the pattern observed in Lokoja is such that the temperatures range between 25.34°C in August and 30.14°C in February, while the precipitation ranges from 3.55 mm in January to 255.92 mm in September. In the case of Ibadan, the monthly mean temperature ranges from 25.40°C in July to 29.26°C in March, while the rainfall varies from 5.91 mm in January to 236.47 mm in September. The results show that, within the tropics, there is a tendency for temperature to increase with latitude, while precipitation tends to decrease with an increase in latitude.

Figure 5 depicts Monthly trend of maximum, mean and minimum temperatures with precipitation for Guinea Savanna (Jalingo, Lafia, and Minna). The monthly mean temperature of Jalingo has a variation from 24.60°C in December to 29.54°C in April, while precipitation fluctuates from 0.92 mm in January to 242.51 mm in September. The lowest and highest average temperature is usually recorded in Minna: that is, 25.78°C in August to a high of 31.53°C in March. Precipitation ranged from 0.70 mm recorded in the month of December to as high as 240.18 mm in September. For Lafia, the temperatures range between 25.17°C in August and 30.36°C in March, while the precipitation varied between 0.60 mm in December and 226.43 mm in September. From the data, it can be observed that with an increase in latitude, there is a shift toward a general decrease in precipitation.

Figure 6 illustrates the Monthly Maximum, mean, and minimum temperature, along with precipitation, trends for the Sudan Savanna. For Yola, the variation of mean temperature varies from 24.16°C in the month of January to 30.43°C in March, while precipitation is from 0.15 mm from the January to 227.97 mm in the month of August. The average temperatures in Kaduna range from 23.88°C in January to 29.35°C in April, while the rainfall ranges from 0.15 mm in December to 277.84 mm in August. In Kebbi State, however, the mean temperature changes from 25.71°C in January to 33.58°C in April, and the rain falls between a low of 0.42 mm in January and 239.42 mm in August. Results in the graph above show a seasonality shift in lower temperatures in January, associated with low precipitation due to the prevalence of cold and dry air.

Figure 7 shows the monthly trends in maximum, mean, and minimum temperatures and precipitations for Sahel Savanna: Maiduguri, Katsina, and Sokoto. In Maiduguri, the mean temperature ranges from 22.78°C in January to 33.07°C in March, while the precipitation ranges from 0.01 mm in January to 205.27 mm in August. The range of temperature in Katsina is from an average of 22.01°C in January to 31.56°C in April, while the rainfall ranges from 0.00 mm in January to 249.74 mm in August. For Sokoto, the temperature ranges from 24.35°C in January to 33.83°C in April, while the precipitation ranges from 0.20 mm in January to 224.34 mm in August. The temperature decline is followed by a reduction in rainfall in the entire Sahel Savanna region, while in the humid forest zone, the closeness of its area to the Atlantic Ocean is believed to be the force behind its general weather. Over the Sahel Savanna, with cold dry air, there is low humidity that contributes to low levels of precipitation.

4.2 Annual Trends

Annual trends of temperature in the different vegetation zones: Humid Forest up to the arid Sahel Savanna; figures 8a to 8e and annual trends of precipitation in the different zones, figures 9a to 9e are presented. A general increase of temperature is seen in all zones through visual examination, while a general decrease in precipitation is observed in most areas.

Figure 8 and 9 present the trends of temperature and precipitation, respectively, for these zones. Concretely speaking, the least variation in temperature is recorded for the Humid Forest zone, as might be expected from the highest quantum of precipitation received, as shown in figure 8a. In fact, figure 8a indicates that Ikeja has the highest annual temperature of 28⁰ C in 2021, considering it received relatively lower precipitation in figure 9a. On the contrary, Calabar and Yenagoa have the lowest temperatures and highest precipitation of all the vegetation zones.

The Derived Savanna shows a gradual increase in temperature while there is a decrease in precipitation, as shown in figure 9b. Makurdi has recorded the highest rainfall within this zone, whereas Ibadan has recorded the lowest precipitation within the same zone. The Guinea Savanna, as shown in figure 8c, exhibits a gradually rising temperature with a fall in precipitation, as indicated in figure 9c; Minna records the highest temperature with the lowest precipitation within this zone. It can be clearly seen that the Sudan

Savanna-mega zone exhibits an increasing temperature trend within all locations. Precipitation tends to slightly increase, as portrayed in figure 9d, with the trend in Kaduna and Birnin Kebbi showing significant trends that are worthy of further analysis when subjected to the Mann-Kendal trend test.

Figure 8e shows the Sahel Savanna to record the highest values of temperature among the mega vegetation zones of Nigeria. Temperature values are on the increase in the locations, while precipitation indicates a slight upward trend, as shown in figure 9e. The highest temperature value in this zone is recorded in Sokoto. Also, proximity to water bodies enhances temperature regulation and precipitation patterns; for example, the Humid Forest zone, because of its proximity to the Atlantic Ocean, has lower temperatures and high precipitation compared to other zones. Additionally, latitudinal variations also contribute to the differences in temperature and precipitation: with increased latitude due to atmospheric circulation, temperature generally rises and precipitation decreases. This can also be explained by the consistent rise in temperature through human-induced activities, such as industrial emissions of greenhouse gases that allow the Earth's surface to continue heating up.

4.3 Mann-Kendall and Sen's Slope trend test

Table 2 and Table 3 shows Mann-Kendall and Sen's Slope trend test for temperature and Precipitation in all the vegetations. The Humid Forest zone; Calabar, Yenagoa and Ikeja with a strong positive z-value of 5.12, 5.30 and 5.29 respectively indicates an increasing trend for temperature data over time. Their z-value are large suggesting a very strong increasing trend. The p-value of $2.9e-7$, $1.1e-7$ and $1.2e-7$ reflect the statistical significant of the trend and are extremely small and far below the common 5% significance level (0.05). This implies that the trend is highly statistical significant. The positive Sen's slope value of 0.023, 0.027 and 0.037 for Calabar, Yenagoa and Ikeja gives the magnitude of the trend. Their positive Sen' slope value indicates a consistent upward trend in temperature suggesting a gradual increase over the period analyzed. The results for Humid Forest zone strongly support the existence of a meaningful upward trend in temperature data. This suggests that the temperature has been steadily rising over the period studied. Comparatively, the precipitation trend from TABLE 2 show that the Humid Forest zone; Calabar, Yenagoa and Ikeja indicates a negative z-value of -1.80, -1.95 and 0.4 respectively. The negative value for Calabar and Yenagoa suggest a decreasing trend over time, although it indicates a moderate, not strong negative trend while Ikeja indicate a magnitude of 0.40 suggest a weak positive trend. Their p-value of 0.07, 0.33 and 0.68 at 5% significance level, their p-value are greater than 0.05 suggests that the trends are not statistically significant but Calabar is approaching significance. Their Sen's slope value confirms the direction and magnitude of the trend, showing a significant decrease over time and increase for Ikeja. The overall results for precipitation trend in Humid Forest zone is not statistically significant.

The annual trend for Derrive Savanna ; Makurdi, Lokoja and Ibadan show a strong positive z-value of 4.52, 4.89 and 5.42 suggest a very strong increasing trend . p-value of $5.9e-6$, $9.6e-7$ and $5.8e-8$ indicates that the trend is highly statistically significant. The positive Sen's slope of 0.02, 0.02 and 0.03 show the magnitude of the trend and indicate a consistent upward trend. The precipitation trend from TABLE 3 show the three locations for Derrive Savanna have a negative z-value of -1.21, -1.18 and -0.04 with p-value of 0.22, 0.23 and 0.96. This indicates decreasing trend from z-value and statistically not significant when compare with the 5% level.

Temperature and precipitation trends from Table 2 and Table 3 for Guinea Savanna (Jalingo, Lafia, Minna), all locations exhibit significant positive trends in temperature with z-values ranging from 3.08 to 4.57 and very low p-values (0.002 to 0.0000004), indicating strong upward trends. The Sen's slope values (0.01 to 0.02) suggest that temperatures are increasing at a consistent rate across these locations. In Guinea Savanna, the precipitation trends are largely non-significant, with negative z-values (-1.76 to -0.37) and p-values above 0.05, suggesting little to no significant change, although Jalingo indicates a

slight declining trend (Sen's slope = -2.66). In Sudan Savanna (Yola, Kaduna, Birnin Kebbi), the z-values (2.32 to 3.89) and low p-values (0.02 to 0.0000098) also confirm significant warming, with Sen's slope values of 0.01 to 0.02, suggesting a consistent increase in temperature. In Sudan Savanna, precipitation trends are also mostly non-significant, with mixed z-values (-1.80 to 0.29) and corresponding Sen's slope values (-2.70 to 0.89). In Sahel Savanna (Maiduguri, Katsina, Sokoto), two locations (Katsina and Sokoto) show statistically significant warming ($p < 0.05$) with z-values of 2.27 and 3.18, respectively, while Maiduguri shows a non-significant increase ($z = 1.60$, $p = 0.108$). Sen's slope values (0.01 to 0.02) indicate steady temperature rises. In Sahel Savanna, precipitation shows non-significant trends (p-values > 0.05), though Katsina has a notable Sen's slope (3.33), indicating a possible increase, while the other locations show smaller changes.

The overall result for Mann-Kendall and Sen's slope trend test depicts temperature increases are statistically significant in most areas across the region. These results concur with those of Ragatoa et al. (2019), who found significant upward trends in temperature in some part Nigeria. Precipitation changes are less consistent and largely non-significant but approaching significant in most cases.

4.4. Regression Analysis

The outputs of the regression plots provided were for various climatic zones in Nigeria, each showing the relationship between temperature and precipitation. In general, there is a negative slope in most areas-the implication being that increased temperature reflects reduced precipitation. However, how strong this relationship is would be different in every zone, represented by the R^2 values. The plot from the humid forest zones of Calabar, Yenagoa, and Ikeja from figure 10a, 10b and 10c reveals a negative correlation between temperature and precipitation as the temperature increases, the precipitation tends to decrease in those areas. The R^2 value of 0.26 for Calabar means that just 26% of the variation in precipitation is influenced by temperature, and therefore the influence of temperature on precipitation is negatively weak. Yenagoa with $R^2 = 0.62$ indicates that there is a moderate negative correlation; therefore, 62% of variation in precipitation is explained by temperature. It can be said that temperature most likely plays an important factor in the precipitation pattern over Yenagoa. Ikeja with $R^2 = 0.39$ suggest This is a weak negative correlation; 39% of the variation in precipitation is explained by temperature. Other parameters beyond temperature might drive precipitation.

In Derived Savanna (Ibadan, Lokoja, Makurdi) from figure 10d, 10e and 10f Ibadan with $R^2 = 0.38$ shows that a weak relationship is present since 38% of the variability in the amount of precipitation can be accounted for by temperature. Lokoja, with $R^2 = 0.37$ station is relatively close to Ibadan; hence, temperature accounts for 37% of the variability in precipitation amounts, which indicates a moderate relationship. Makurdi, $R^2 = 0.17$ is only a weak correlation, with a mere 17% of the precipitation variance described by temperature. This indicates that other factors are of overriding dominance in the control of precipitation.

In Guinea Savanna (Minna, Lafia, Jalingo) from figure 10g, 10h and 10i, Minna with $R^2 = 0.24$ portrays poor correlation, as only 24% of the variation in precipitation could be described by temperature. Lafia indicates $R^2 = 0.23$, Here too, a similar value is derived, with 23% of precipitation variability explained by temperature. Jalingo with $R^2 = 0.06$ is very weak correlation of 6% justifies that the temperature is never influenced by precipitation.

In Sudan Savannah (Birnin Kebbi, Kaduna, Yola) from figure 10j, 10k and 10l, Birnin Kebbi with R^2 value of 0.03 indicates a very weak correlation of 3% suggests that there is almost no influence of temperature on precipitation in Birnin Kebbi. Kaduna with R^2 value of 0.08 indicates a very weak correlation of 8% suggests that there is a very weak influence of temperature on precipitation. Yola with

R^2 of 0.01 means that temperature explains only about 1% of the variability of the precipitation, thus almost no relationship. In Sahel Savanna (Sokoto, Katsina, Maiduguri) from figure 10m 10n and 10o, Sokoto ($R^2 = 0.00$) means that no relationship whatsoever between the two variables. Katsina ($R^2 = 0.01$, positive slope), hardly any correlation, with a slightly positive slope, indicating that temperature does vary positively with precipitation. In fact, it is just about insignificant. Changes of Maiduguri $R^2 = 0.01$, slope is positive.

This is similar to Katsina with a slight positive slope, but the R^2 value of 0.01 shows that temperature has practically no impact on precipitation. This relationship is closer in the humid forest zones, especially in Yenagoa, where temperature alone explains a wide variation in precipitation. The relationship is moderate in the derived savanna areas, while it weakens in the Guinea and Sudan savannas. Within the Sahel savanna, there is virtually no meaningful relationship between temperature and precipitation, while slight positive slopes were observed in Katsina and Maiduguri with practically no explanatory power. As one moves from the humid forest zone to the Sahel savanna, the influence of temperature on precipitation progressively weakens to suggest that other climatic, geographical, or atmospheric factors are far more influential in drier zones like the Sahel.

Kernel Density Estimation (KDE)

The kernel density estimate plot bring into focus the temperature and precipitation patterns of Nigeria's ecological zones. The major density for the Humid Forest Zone in figure 11a, 11b and 11c, falls within moderate temperatures of 25-27.5°C, with substantial rainfall, especially at Calabar and Yenagoa; Ikeja is warmer and receives less rain. In the Derived Savanna, Ibadan, Lokoja, and Makurdi are expected to have moderate rainfall with temperatures oscillating between 25 and 28°C. Makurdi and Lokoja are somewhat wetter than Ibadan from figure 11d, 11e and 11f. In the Guinea Savanna, which includes areas such as Minna, being a little drier; Lafia and Jalingo from figure 11g, 11h and 11i, the temperatures move within the range of 25 to 27°C, with moderate rainfall. It is also hotter in the Sudan Savanna, that is, Birnin Kebbi, Kaduna, and Yola from figure 11j 11k and 11l, since the temperature has been within the range of 26 to 32.5°C, with low precipitation-the drylands being represented by Birnin Kebbi. While Sokoto, Katsina, and Maiduguri from figure 11m, 11n and 11o have a temperature of 34°C at the Sahel Savanna zone, the precipitation is very low, which characterizes arid conditions. From the KDE plot in each zone, one notices an increased trend in temperature, which inversely relates to the decreased precipitation level, especially in the Sudan and Sahel Savanna zones; this evidences Nigeria's great variation in climatic conditions, which changes from the wetter southern zone into the arid north.

5.0 Conclusion

This study, therefore, further emphasizes a considerable deviation from the standpoint of temperature and precipitation within the ecosystem boundary of Nigeria-from the humid forest to the arid zone of the Sahel Savanna. The results indicated an increase in temperature within all zones; most pronounced warming has taken place in the drier regions, including Sudan and Sahel Savanna. The precipitation trends are rather varied, with a significant decline in the southern zones and very slight in the northern regions. Results of the Mann-Kendall and Sen's Slope test of trend showed that rising temperature trends are statistically significant while changes in precipitations are mostly less consistent and not significant. Analysis indeed shows that, in the savannas-where rainfall is especially scanty-increasing temperatures are indeed associated with reduced rainfall. Very key implications arise here as regards agriculture, water management, and adaptation strategies in Nigeria; these data point to increasing challenges coming with warming temperatures, along with shifting patterns of precipitation, particularly in regions already prone to drought and desertification. Understanding these climatic trends is paramount in devising appropriate interventions that can be cast to reduce the impact of climate change on ecosystems, agriculture, and livelihood.

References

- [1] Morales, J. A. (2022). Climate: Climate Variability and Climate Change. In *Coastal Geology* (pp. 375-388). Cham: Springer Nature Switzerland.
- [2] Farooqi, Z. U. R., Sabir, M., Qadeer, A., Naeem, A., Murtaza, G., & Yousaf, H. (2022). Understanding the causes of climatic change in the environment. In *Climate Change: The Social and Scientific Construct* (pp. 37-47). Cham: Springer International Publishing.
- [3] Zaied, Y. B., Cheikh, N. B., & Rahmoun, M. (2023). On the effects of climate variability on agricultural crops: evidence from an in-transition economy. *Environmental Economics and Policy Studies*, 25(2), 143-159.
- [4] Joseph, O., Gbenga, A. E., & Langyit, D. G. (2018). Desertification risk analysis and assessment in Northern Nigeria. *Remote Sensing Applications: Society and Environment*, 11, 70-82.

- [5] Osuji, E. E., Olaolu, M. O., & Tim-Ashama, A. C. (2023). Impacts of Climate Change on Sustainable Crop Production in Nigeria. In *Climate Change Impacts on Nigeria: Environment and Sustainable Development*(pp. 39-56). Cham: Springer International Publishing.
- [6] Ani, K. J., Anyika, V. O., & Mutambara, E. (2021). The impact of climate change on food and human security in Nigeria. *International Journal of Climate Change Strategies and Management*, 14(2), 148-167.
- [7] Gwambene, B., Liwenga, E., & Mung'ong'o, C. (2023). Climate change and variability impacts on agricultural production and food security for the smallholder farmers in Rungwe, Tanzania. *Environmental Management*, 71(1), 3-14.
- [8] Patel, S., Dey, A., Chaturvedi, A., Sharma, A., & Singh, R. (2024). Adaptation and Mitigation Strategies Under Climate Change Scenario. In *Technological Approaches for Climate Smart Agriculture* (pp. 213-228). Cham: Springer International Publishing.
- [9] Agbo, E. P., Nkajoe, U., Edet, C. O., & Ali, N. (2023). Identification of trends of key parameters affecting vegetation over Nigeria's vegetation zones using innovative trend analysis. *Environmental Earth Sciences*, 82(20), 464.
- [10] Okono, M. A., Agbo, E. P., Ekah, B. J., Ekah, U. J., Ettah, E. B., & Edet, C. O. (2022). Statistical analysis and distribution of global solar radiation and temperature over southern Nigeria. *Journal of the Nigerian Society of Physical Sciences*, 588-588.
- [11] Ewona, I. O., Osang, J. E., & Udo, S. O. (2014). Trend analyses of rainfall patterns in Nigeria using regression parameters. *Int. J. Technol. Enhanc. Emerg. Eng. Res*, 2, 129-133.
- [12] Do, C. B. (2008). The multivariate Gaussian distribution. *Section Notes, Lecture on Machine Learning*, CS, 229.
- [13] Agbo, E. P. (2021). The role of statistical methods and tools for weather forecasting and modeling. *Weather Forecasting. IntechOpen*, 3-22.
- [14] James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Linear regression. In *An introduction to statistical learning: With applications in python* (pp. 69-134). Cham: Springer International Publishing.
- [15] Phuong, D. N. D., Tram, V. N. Q., Nhat, T. T., Ly, T. D., & Loi, N. K. (2020). Hydro-meteorological trend analysis using the Mann-Kendall and innovative-Şen methodologies: a case study. *International Journal of Global Warming*, 20(2), 145-164.
- [16] Agbo, E. P., Ekpo, C. M., & Edet, C. O. (2020). Trend Analysis of Meteorological Parameters, Tropospheric Refractivity, Equivalent Potential Temperature for a Pseudoadiabatic Process and Field Strength Variability, Using Mann Kendall Trend Test and Sens Estimate. *arXiv preprint arXiv:2010.04575*.
- [17] Agbo, E. P., Edet, C. O., Magu, T. O., Njok, A. O., Ekpo, C. M., & Louis, H. (2021). Solar energy: A panacea for the electricity generation crisis in Nigeria. *Heliyon*, 7(5).
- [18] Alhaji, U. U., Yusuf, A. S., Edet, C. O., Oche, C. O., & Agbo, E. P. (2018). Trend analysis of temperature in Gombe state using Mann Kendall trend test. *J. Sci. Res. Rep*, 20(3), 1-9.

- [19] Agarwal, S., Suchithra, A. S., & Singh, S. P. (2021). Analysis and interpretation of rainfall trend using Mann-Kendall's and Sen's slope method. *Indian Journal of Ecology*, 48(2), 453-457.
- [20] Agbo, E. P., & Ekpo, C. M. (2021). Trend analysis of the variations of ambient temperature using Mann-Kendall test and Sen's estimate in Calabar, southern Nigeria. In *Journal of Physics: Conference Series* (Vol. 1734, No. 1, p. 012016). IOP Publishing.
- [21] Agbo, E. P., Nkajoe, U., Okono, M. A., Inyang, E. P., & Edet, C. O. (2023). Temperature and solar radiation interactions in all six zones of Nigeria. *Indian Journal of Physics*, 97(3), 655-669.
- [22] Olusegun, C., Ojo, O., Olusola, A., & Ogunjo, S. (2024). Solar radiation variability across Nigeria's climatic zones: a validation and projection study with CORDEX, CMIP5, and CMIP6 models. *Modeling Earth Systems and Environment*, 10(1), 1423-1440.
- [23] Ibe, O., & Nymphas, E. F. (2010). Temperature variations and their effects on rainfall in Nigeria. *Global Warming: Engineering Solutions*, 565-578.
- [24] Ragatoa, D. S., Ogunjobi, K. O., Okhimamhe, A. A., Francis, S. D., & Adet, L. (2018). A trend analysis of temperature in selected stations in Nigeria using three different approaches. *Open Access Library Journal*, 5(2), 1-17.

MAJOR AND TRACE ELEMENT GEOCHEMISTRY OF PEGMATITES IN UMAI, OBAN MASSIF, SE, NIGERIA; CONSTRAINTS ON ITS CHARACTERIZATION AND MINERALIZATION POTENTIAL

BENJAMIN ODEY OMANG, LATIFA TIJANI OJIKOTU, ARIKPO TEMPLE OKAH, GODSWILL EYONG
ABANG and GODWIN TERWASE KAVE

Department of Geology University of Calabar

Keywords: Pegmatites, Umai, Mineralization, Granodioritic, Geochemistry

Introduction

Pegmatites are coarse-grained intrusive igneous rocks, often associated with late-stage crystallization of magmas. These rocks are particularly valued for their ability to host a wide range of economically important rare metals such as lithium (Li), beryllium (Be), tantalum (Ta), niobium (Nb), cesium (Cs), and gemstones, including topaz, tourmaline, and aquamarine. Pegmatites also serve as vital sources of industrial minerals like feldspar, quartz, and mica, which are critical to several industries such as electronics and ceramics (London, 2018; Karampelas et al., 2020). Globally, pegmatites have been classified into various families, including the Lithium-Cesium-Tantalum (LCT) and Niobium-Yttrium-Fluorine (NYF) types, based on their rare metal contents and geochemical properties (Cerny & Ercit, 2005; Dill, 2015).

In Nigeria, pegmatite deposits are well-documented, particularly within the Precambrian Basement Complex. However, most research has focused on pegmatites in the north-central and southwestern regions of the country, while the southeastern deposits remain underexplored. This study addresses the gap by investigating the geochemistry and mineralization potential of the pegmatites in the Umai area of the Oban Massif, Southeastern Nigeria. Specifically, the study aims to determine the mineralization potential of these pegmatites.

Description of the Area

The Umai area is situated within the Oban Massif, Southeastern Nigeria, a region characterized by diverse rock types including schists, gneisses, phyllites, granites, and granodiorites. Pegmatite dykes intrude these older formations, and they are believed to have formed during the Pan-African Orogeny, a significant geological event that affected much of West Africa around 600 million years ago (Ukaegbu & Ekwueme, 2006). The pegmatites in this region are known for their potential to host rare metals, but systematic studies of their mineralization potential remain scarce compared to other regions like the north-central Nigerian pegmatites associated with tin and tantalum (Adekoya, 1998).

Geographically, the Umai area lies between latitude 5°30'0''N and 5°35'0''N and longitude 8°10'0''E (Figure 1). The pegmatites observed here are typically granodioritic and have been mapped as part of broader lithological studies of the Oban Massif (Ekwueme & Matheis, 1995).

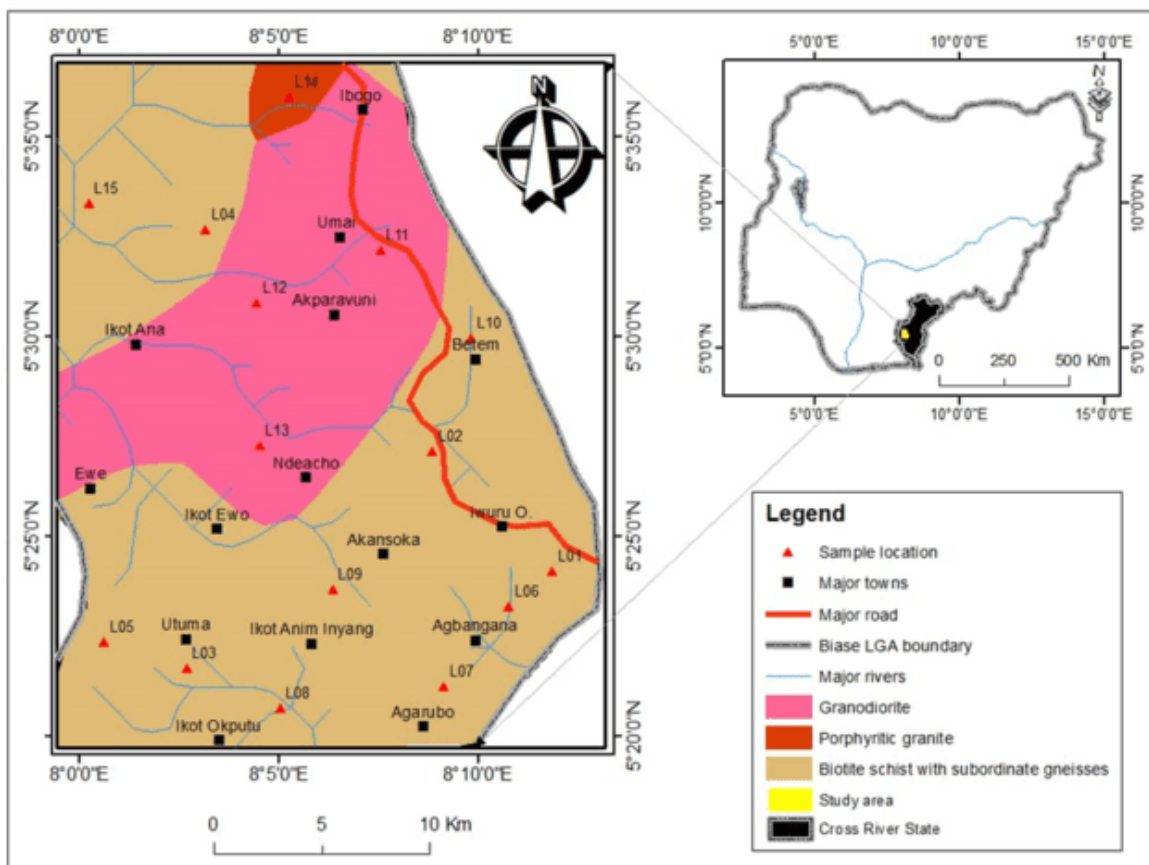


FIG. 1: Geologic map of the Umai area, Oban massif, SE, Nigeria

Materials and Methods

A total of fifteen fresh pegmatite samples were collected from various locations within the Umai area. The samples were carefully labeled, and field observations were noted for each location. These samples were pulverized into fine powder in the geology laboratory at the University of Calabar and subsequently sent to Activation Laboratories (Actlabs) in Ontario, Canada, for detailed geochemical analysis using inductively coupled plasma mass spectrometry (ICP-MS) (Okon et al., 2022).

Results and Discussion

Geochemical Characteristics

The geochemical analysis of the Umai pegmatites revealed a wide range of silica (SiO_2) content, from 43.25% to 69.10%, with an average of 57.30%. These results classify the pegmatites as intermediate felsic rocks, consistent with their granodioritic composition (Figure 2). The $\text{K}_2\text{O}/\text{Na}_2\text{O}$ ratios ranged from 5 to 51.75, indicating a dominance of K-feldspar over albite, which is common in highly fractionated pegmatites. Additionally, the aluminum oxide (Al_2O_3) content exceeded the combined content of calcium oxide (CaO), sodium oxide (Na_2O), and potassium oxide (K_2O), indicating a peraluminous nature (Cerny, 2002; Dostal & Gerel, 2022).

Mineralization Potential

The low K/Rb ratios (below 100) across all samples suggest significant mineralization potential (Figure 3), consistent with earlier findings that such ratios indicate pegmatite mineralization (Wise & Brown, 2012). Nb/Ta ratios below 1 in several samples confirmed the presence of tantalite mineralization (Figure 4), while higher Nb/Ta ratios in other samples suggested columbite mineralization (Van Lichtenvelde et al., 2007; Adetunji & Ocan, 2010).

Notably, the presence of beryllium (Figure 5), with concentrations exceeding 20 ppm in some samples, also indicates possible beryllium mineralization, which is rare in most Nigerian pegmatites (Cerny, 2002).

Comparison with Other Regions

Compared to pegmatites from other regions of Nigeria, particularly those in central and southwestern Nigeria, the Umai pegmatites exhibit lower Na_2O content but higher concentrations of elements such as Rb, Sr, and Cs. The tantalum concentrations (up to 25.07 ppm) are notably higher than those reported from other localities, further supporting the mineralization potential of the Umai pegmatites (Mafimisebi, 2023; Pagung et al., 2023). These elevated trace element concentrations indicate a high degree of fractionation, making these pegmatites significant for tantalum and columbite exploitation.

Conclusion

The Umai pegmatites are predominantly granodioritic and exhibit a high degree of fractionation, confirming their mineralization potential. Low K/Rb and Nb/Ta ratios suggest that these pegmatites are highly mineralized, particularly for tantalum and columbite. The discovery of significant beryllium concentrations further enhances the economic value of these pegmatites, indicating that the region holds potential for future exploration and sustainable exploitation (London, 2018; Van Lichtenvelde et al., 2007).

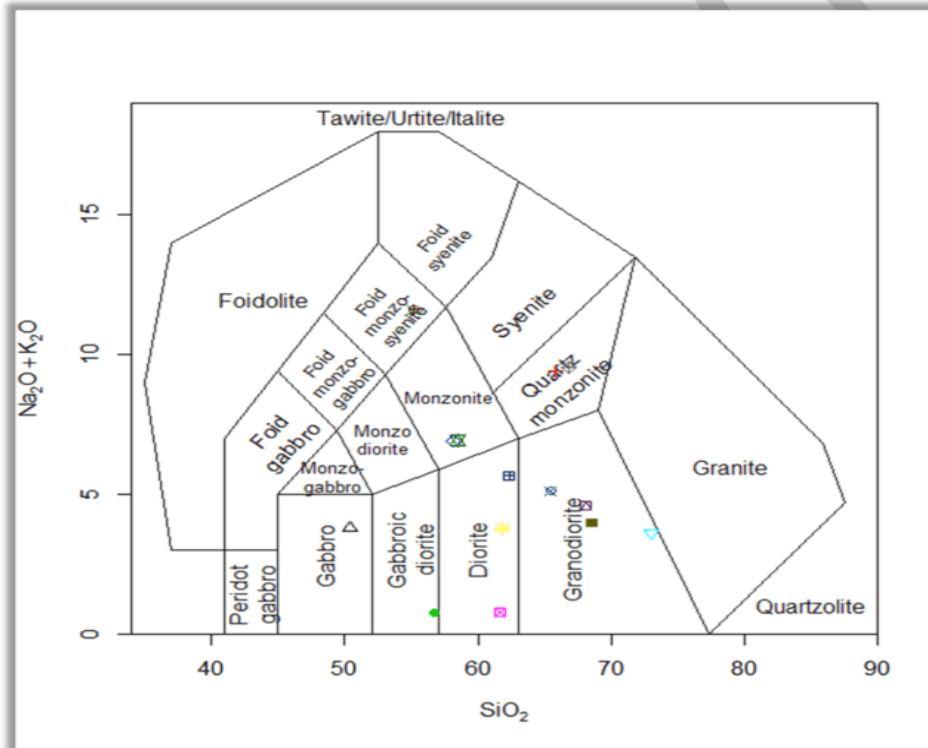


FIG.2. TAS (Middlemost 1994) classification diagram for pegmatite samples of Umai.

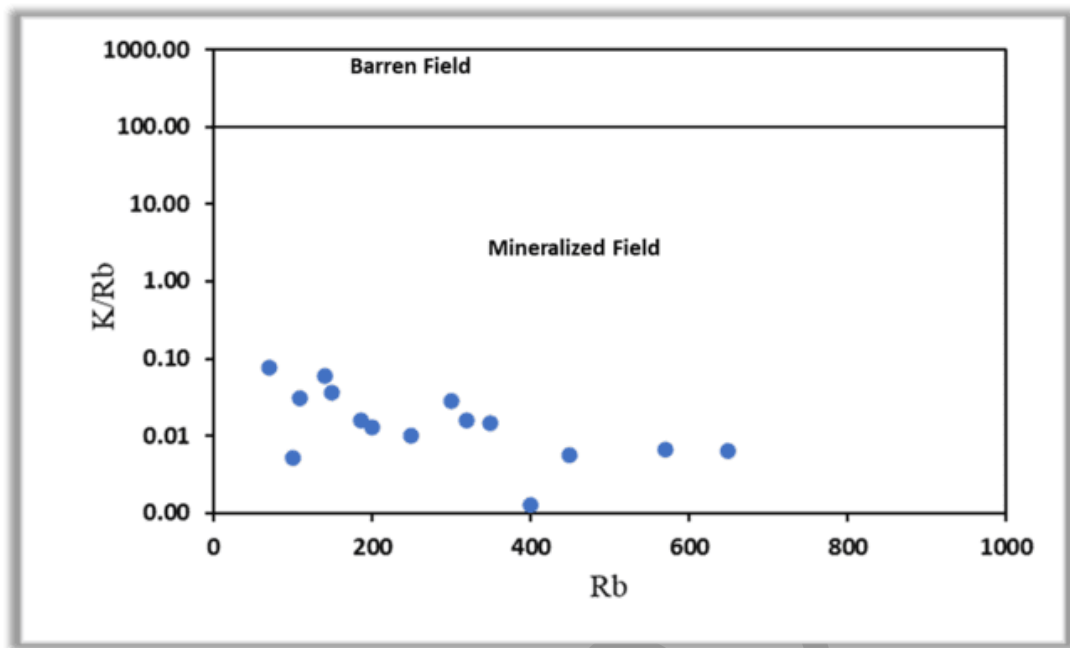


FIG. 3. K/Rb vs. Rb distribution pattern in the pegmatites of Umai, southeastern Nigeria showing barren and mineralized pegmatites (fields after Staurovet *al.*, 1969)

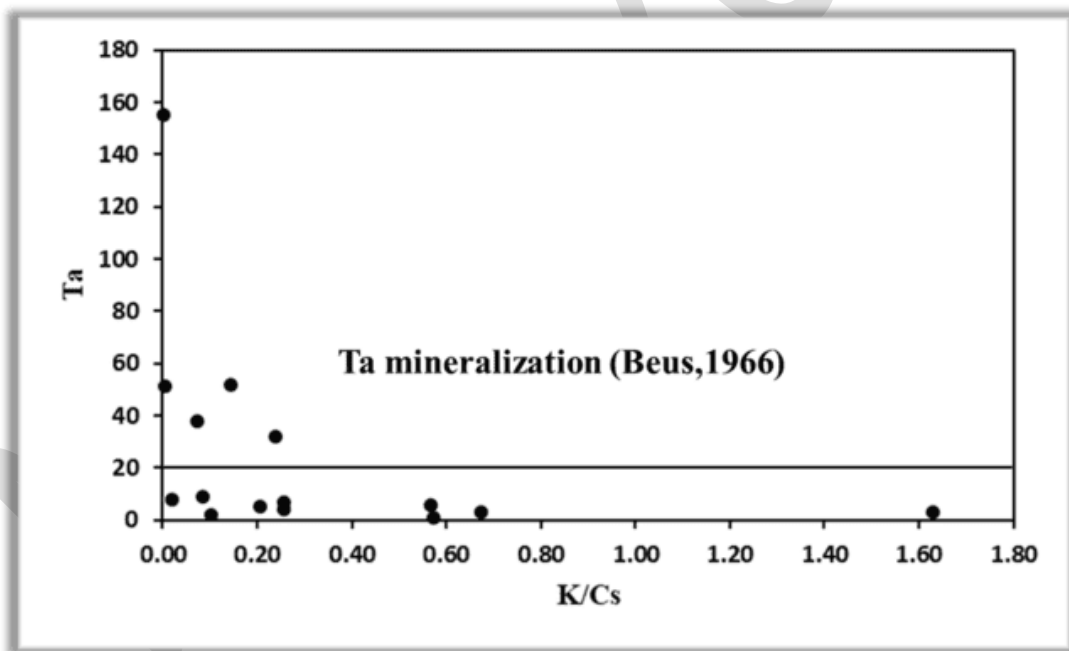


FIG. 4. Plot of Ta vs. K/Cs to determine the Tantalum potential of the Umai pegmatites (field after Beus, 1966).

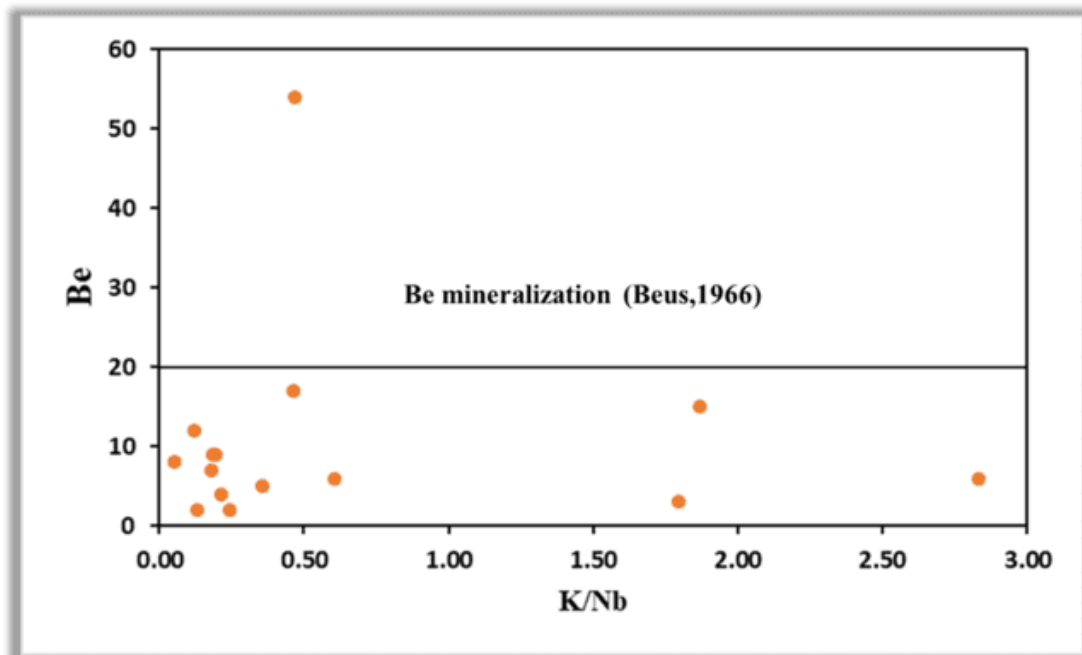


FIG. 5. Plot of Be vs. K/Nb to determine the Beryllium potential of the Umai pegmatites (field after Beus, 1966).

Recommendations

Further geophysical and geochemical studies should be conducted to better delineate the mineralization zones in the Umai area.

References

- Adekoya, J. A. (1998). The geology and geochemistry of the Maru Banded Iron-Formation, northwestern Nigeria. *Journal of African Earth Sciences*, 27(2), 241–257. [https://doi.org/10.1016/S0899-5362\(98\)00059-1](https://doi.org/10.1016/S0899-5362(98)00059-1)
- Beus, A. A. (1966). *Geochemistry of Beryllium: And Genetic Types of Beryllium Deposits*. WH Freeman.
- Cerny, P. (2002). Mineralogy of Beryllium in Granitic Pegmatites. *Reviews in Mineralogy and Geochemistry*, 50(1), 405–444. <https://doi.org/10.2138/rmg.2002.50.10>
- Cerny, P., & Ercit, T. S. (2005). The classification of granitic pegmatites revisited. *The Canadian Mineralogist*, 43(6), 2005–2026.
- Dill, H. G. (2015). Pegmatitic rocks and economic geology. In H. G. Dill, *The Hagendorf-Pleystein Province: The Center of Pegmatites in an Ensialic Orogen* (Vol. 15, pp. 1–54). Springer International Publishing. https://doi.org/10.1007/978-3-319-18806-5_1
- Dostal, J., & Gerel, O. (2022). Occurrences of niobium and tantalum mineralization in Mongolia. *Minerals*, 12(12), 1529. <https://doi.org/10.3390/min12121529>
- Ekwueme, B. N., & Matheis, G. (1995). Geochemistry and economic value of pegmatites in the Precambrian basement of Southeast Nigeria. *Journal of Geological Sciences*, 4(2), 375–392.
- Karampelas, S., Kiefert, L., Bersani, D., & Vandenabeele, P. (2020). *Gems and Gemmology: An Introduction for Archaeologists, Art-Historians and Conservators*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-35449-7>
- London, D. (2018). Ore-forming processes within granitic pegmatites. *Ore Geology Reviews*, 101, 349–383. <https://doi.org/10.1016/j.oregeorev.2018.04.020>
- Mafimisebi, P. (2023). Geochemistry, classification characteristics, and mineralization potential of Pegmatites from Oke Ode, Kwara State, Northeastern Nigeria: Implications for provenance and weathering history. *Earth Sciences Repository*. <https://doi.org/10.31223/X5S09Z>
- Middlemost, E. A. (1994). Naming materials in the magma/igneous rock system. *Earth-science reviews*, 37(3–4), 215–224.
- Okon, E. E., Kudamnya, E. A., Oyeyemi, K. D., Omang, B. O., & Ojo, O. (2022). Field observations and geophysical research applied to the detection of manganese (Mn) deposits in the eastern part of Oban

- Massif, South-Eastern Nigeria: An integrated approach. *Minerals*, 12(10), 1250. <https://doi.org/10.3390/min12101250>
- Pagung, R., Rosière, C. A., & Silva, R. C. F. (2023). The Serra da Serpentina Group: a back-arc basin related to the Paleoproterozoic Minas-Bahia orogeny? *Journal of South American Earth Sciences*, 128, 104427.
- Stavrov, D., Stolyaro, Is., & Iovcheva, e. (1969). Geochemistry and origin of Verkh-Isetgranitoid massif in central Ural. *Geochemistry International Ussr*, 6(6), 1138.
- Ukaegbu, V. U., & Ekwueme, B. N. (2006). Petrogenesis and geotectonic setting of the Pan-African basement rocks in Bamenda Massif, Obudu Plateau, Southeastern Nigeria: Evidence from trace element geochemistry. *Chinese Journal of Geochemistry*, 25(2), 122–131. <https://doi.org/10.1007/BF02872171>
- Wise, M., & Brown, C. D. (2012). Mineral chemistry, petrology, and geochemistry of the Sebago granite-pegmatite system, southern Maine, USA. *Journal of Geosciences*, 3-26. <https://doi.org/10.3190/jgeosci.061>

TREND AND LATITUDINAL VARIATION OF WIND SPEED DISTRIBUTION IN THREE SELECTED LOCATIONS IN NIGERIA FOR 41 YEARS (1981-2022). AN INSIGHT FROM ATMOSPHERIC CIRCULATION.

^aMoses. A. Okono, ^{a,b} Igwe O.Ewona, ^aEmmanuel A Awak, ^{a,b}Jonathan E. Osang
mosesokono3@gmail.com, igweewona@gmail.com, emma4xst@gmail.com,
jonathanosang@unicross.edu.ng

^aPhysics Department, University of Calabar, Calabar, 540242, Nigeria

^bPhysics Department, University of Cross River State, Calabar, 540242, Nigeria

ABSTRACT

Wind speed variation has been used majorly in determining the development of sustainable power generation. The study analyzes wind speed in Nigeria for 41 years (1981 to 2022), using time series model. The data used in this study was extracted from satellite database for three important locations namely, Katsina (13°N), Abuja (9°N), and Port Harcourt (4.78°N) at 10 meters and 50meters high. In the current study, long-term wind speed data has been used to study the spatial variation with latitude and to observe the development potential of wind energy. The Mann-Kendall trend test has been applied to identify the trend in time series data while the Weibull distribution model and kernel density estimation shows a better understanding in the distribution of the data. The wind speed at 50m were observed to be greater than that of 10m high. The Mann-Kendall trend test results indicate a decreasing trend in wind speed in Katsina with a large negative Z-value of -4.68 and -3.50 at 10m and 50m suggesting a very strong decreasing trend and an extremely p-value of 2.8e-6 and 0.0004 far below the common 5% significance level reflect the trend is highly statistically significant. In Abuja, the decrease in trend were observed to be non-statistically significant with Z-values of -0.18 and -0.66 and p-value of 0.85 and 0.50. Port Harcourt at 10 meters high shows a decrease in trend and it has been observed to be statistically significance with Z-value of -2.37 and p-value of 0.01 while at 50 meters is non-statistically significant with Z-value of -1.72 and p-value of 0.08 although is approaching significant. The result indicates the increase in the wind speed with latitude, hence Katsina is the windiest, while the lowest wind speed was recorded at Port Harcourt, owing to its closeness to the equator and influence of atmospheric circulation systems such as the ITCZ. Katsina has the highest wind power density and therefore a high potential for wind energy development, Abuja and Port Harcourt have relatively lower wind speeds; such areas are not suitable for commercial wind energy applications and hence favor solar alternatives. These findings are useful in wind energy possibilities within the nation, especially in the northern parts of the country under the influence of northeast trade winds which blows from Sahara Desert. Critical insights are developed from this research in renewable energy planning against the impacts brought about by climate change on wind resources in Nigeria.

Keywords: Wind speed; Mann-Kendall Test; Sen's slope; Kernel Density; Trade winds; Climate Change;

1.0 Introduction

Wind energy has been assessed worldwide as one of the very feasible renewable resources due to its economically viable and environmentally friendly, hence quite sustainable. (Oyedepo, Adaramola, and Paul 2012) Knowledge of the wind speed trend is significant for the optimum utilization of this type of energy, particularly over geographical territories where climatic and geographical conditions are inconstant. Nigeria,

having a lot of zones with different characteristics of climate, provides a very interesting case that may be considered for studying the distribution of wind speed and its latitudinal variation. (Danlami et al. 2019)

This research aims to examined the behavioral pattern of wind speed in Katsina at 13°N, Abuja at 9°N, and Port Harcourt at 4°N. Long-term trends and spatial variation in wind speed have been analyzed in the 41-year period from 1981

through to 2022. This analysis further helps reveal the trend of variation in wind speed with latitude, as well as influences from atmospheric circulation patterns that were very helpful in assessing the potential for wind energy development across different regions of the country.

(Fagbenle et al. 2011) evaluates the wind energy potentials of Maiduguri and Potiskum, North-East Nigeria, using 21 years of wind data. Wind speed ranges from 3.90 to 5.85 m/s in Potiskum, and from 4.35 to 6.33 m/s in Maiduguri. Both sites have some potential for stand-alone and medium-scale wind power generation.

(Adaramola and Oyewola 2011) reviewed the wind speed distribution and energy potential in Nigeria; it presents the wind speed varies from 2 to 9.5 m/s and power density between 3.40 to 520 kW/m². Wind speed increased from south to north, with the ability to identify areas fitting for electricity generation and water pumping.

(Safari 2011) modeled the suitability of Lagos and Calabar, two coastal Nigerian stations, for wind energy generation at 70 m height using power law and diabatic methods. The diabatic method performed better by considering site-specific conditions. Lagos demonstrated stronger wind speeds than Calabar, particularly during nighttime stable atmospheric conditions. (Safari 2011)

2.0 Methodology

Several methods have been adopted to analyzed the data, these include the Mann-Kendal trend to detect the trend in the data and kernel density estimation plot for better description of the data

The Mann-Kendall trend test

Mann-Kendall trend test method is a frequent method adopted for time series analysis; it is a non-parametric technique that is to say it does not require the data to conform with a particular distribution. The M-K test detect monotonic trend of a random data also identify if a particular trend is statistically significant or

not. (Okono et al. 2022). (E P Agbo and Ekpo 2021)

This method can be applied when data x_j agrees with the relation;

$$x_j = f(t_i) + \varepsilon_i,$$

$f(t_i)$ describe a function of continuous increasing or decreasing monotonically, ε_i are the residuals.

In the Mann-Kendall test, the null hypothesis H_0 assumes no trend, while the alternative hypothesis H_1 suggests that there is a trend in the series. If the result favors the null hypothesis, then this means the data points are randomly distributed over time. On the other hand, when the alternative hypothesis is accepted, it implies that the trend is either monotonically increasing or decreasing. (Emmanuel P. Agbo, Ekpo, and Edet 2020). (Danlami et al. 2019)

The Mann-Kendall test uses the statistic S , calculated using

$$S = \sum_{k=1}^{n-1} \sum_{j=\lambda=1}^n \text{sgn}(x_j - x_k), \quad (1)$$

where,

$$\text{sgn}(x_j - x_k) = \begin{cases} +1; & \text{if } (x_j - x_k) > 0 \\ 0; & \text{if } (x_j - x_k) = 0 \\ -1; & \text{if } (x_j - x_k) < 0 \end{cases}$$

The value of the Z statistic have been computed and the variance of S' $\text{VAR}(S)$ is used (Emmanuel P. Agbo, Ekpo, and Edet 2020)

$$\text{VAR}(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t+5) \right] \quad (2)$$

g be the number of tied groups in the series, it means that test take care about the tied or equal values. Using the values of $\text{VAR}(S)$ and S the statistic test Z is obtained as

$$Z = \begin{cases} \frac{S-1}{\sqrt{VAR(S)}}; & S > 0 \\ 0; & S = 0 \\ \frac{S+1}{\sqrt{VAR(S)}}; & S < 0 \end{cases} \quad (3)$$

Sen's Slope Estimator

Sen's slope is also a non-parametric technique used to identify the magnitude of a trend in a particular distribution and the direction of such trend whether increasing or decreasing per unit time.(E P Agbo and Ekpo 2021)

The linear equation is represented by

$$f(t) = Qt + B \quad (4)$$

It defines this as a time function that is always monotonic-increasing or monotonic-decreasing, Q being the slope, and B is a common. The computation of the slopes between all pairs of data items gives the slope estimate in equation.(Emmanuel P. Agbo, Nkajoe, and Edet 2023)(Oteng Mensah, Alo, and Ophori 2024)

$$Q_i = \frac{x_j - x_k}{j - k} \quad (5)$$

Where, $j > k$.

If the time series has values x_j , as many $N = n(n-1)/2$ slopes are possible estimations Q_i . The median of the N 's Q_i values represents the slope as estimated by Sen. The N values of Q_i are sorted from lowest to largest and the Sen's estimator is(Emmanuel P. Agbo, Ekpo, and Edet 2020)(Okono et al. 2022)

$$Q = \begin{cases} Q_{[(N+1)/2]}; & \text{if } N \text{ is odd} \\ \frac{1}{2}Q_{[N/2]} + Q_{[(N+2)/2]}; & \text{if } N \text{ is even} \end{cases} \quad (6)$$

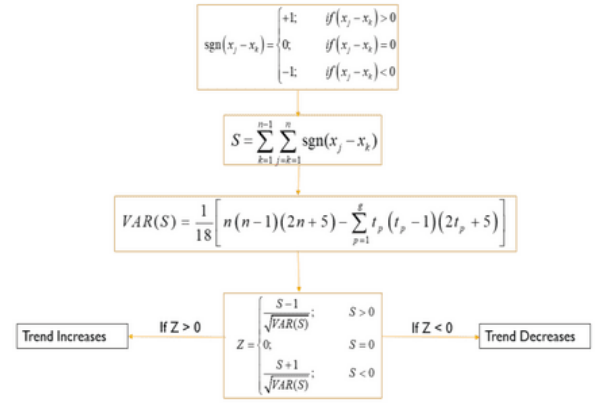


Figure 1. Methodology for Mann-Kendall trend test.

Weibull distribution

Weibull distribution is a typical distribution used to describe the variation of wind velocity, the probability density function and the corresponding cumulative distribution functions. Thus, it is significant to note that the probability density function, $f(V)$ give the probability of the wind at a given velocity, V . Correspondingly, the cumulative distribution function, $F(V)$ of the velocity, V gives the probability that the wind velocity is greater than or less than V , or within a given wind speed range.(Akdağ and Dinler 2009) The Weibull probability density function $f(V)$ is given by.(Abiodun Isaac Chukwutem and Anisiji Obiora Emeka 2024)(Tuller and Brett 1984) where k is the shape parameter and λ is the scale parameter.

$$f(v; \lambda, k) = \left\{ \left(\frac{k}{\lambda} \right) \left(\frac{v}{\lambda} \right)^{k-1} \exp \left[- \left(\frac{v}{\lambda} \right)^k \right] \right\} \quad (7)$$

From the Weibull distribution function, the wind power density is given as.(Mzough and King Ededem 2023)(Mudasiru and Mustafa 2018)

$$WPD = \frac{1}{2} \times \rho \times v^3 \times P(v) \quad (8)$$

ρ = Air density, V = wind speed, $P(v)$ = probability density.

3.0 Study Location

Nigeria lies between longitude 2° E and 15° E, and between latitude 4° N and 14° N when considering two significant seasons, wet and arid. The seasons have thereby led to the adoption of different renewable energy sources; for instance, the different seasons in the country affect hydropower stations due to the availability and its variation in water.(Chinachi, Nwankwojike, and Uche 2022)

Geographically speaking, northern states in Nigeria tend to be hotter than southern states. Hence, it is characterized by extreme temperature and radiation values. Nigeria is situated just above the equator, which, by default, receives the highest sunshine intensity.(Igbawua et al. 2016). (Emmanuel P. Agbo, Nkajoe, and Edet 2023)

However, considering the fact that the southern region of Nigeria is close to the Atlantic Ocean and northern close to Sahara Desert, the earlier described sunshine, temperature, and solar radiation are actually modified.(Fashae, Olusola, and Adedeji 2017). (Ohunakin and Akinnawonu 2012)

Data collection

Monthly and annual data is been obtained from NASA POWERDATA for wind speed for the period of forty-one 41years (1981-2022).Annual surface pressure and temperature data were also collected to calculate the specific air density for the three selected location.

Data visualization

Python programming language has been adopted as the Analytical tool used in this research. It displaysgraphical statistical analysis using a programming language. This Python program includes the packages sklearn, seaborn, matplotlib, pandas, and numpy used for data visualizations, 1D kernel density estimate (KDE) plots, and Mann-Kendall trend test.

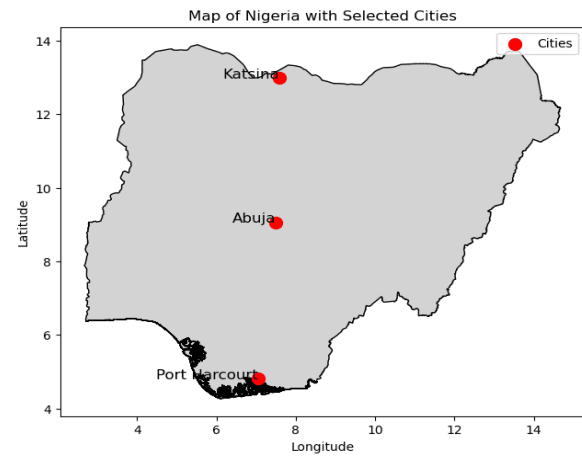


Figure 2. Map of Nigeria

TABLE 1

Locations	Latitude	Longitude
Katsina	13.00	7.60
Abuja	9.07	7.49
Port Harcourt	4.78	7.00

4.0 Result and Discussion

Monthly trend of wind speed

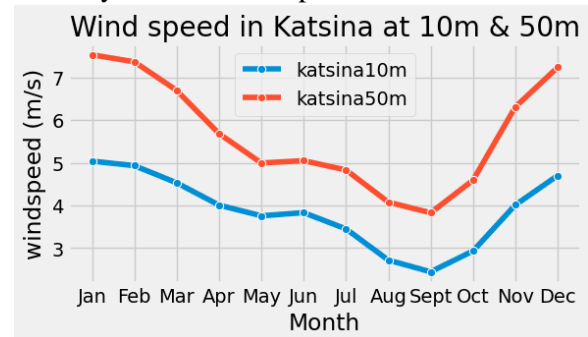


Figure 3. Monthly trend of wind speed in Katsina

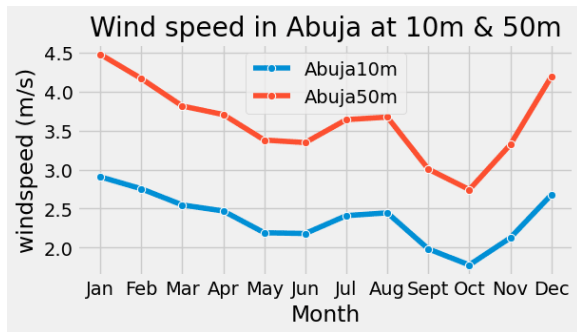


Figure 4. Monthly trend of wind speed in Abuja

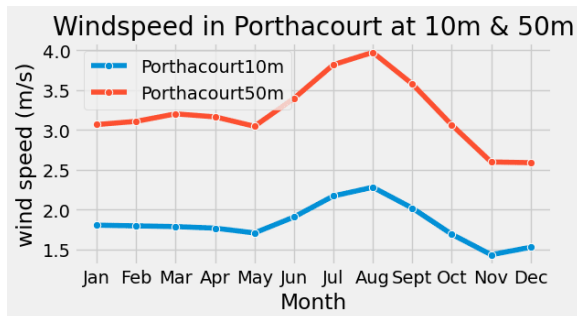


Figure 5. Monthly trend of wind speed in Port Harcourt.

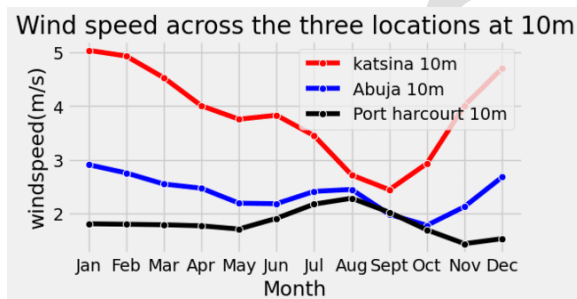


Figure 6. wind speed across the locations at 10m

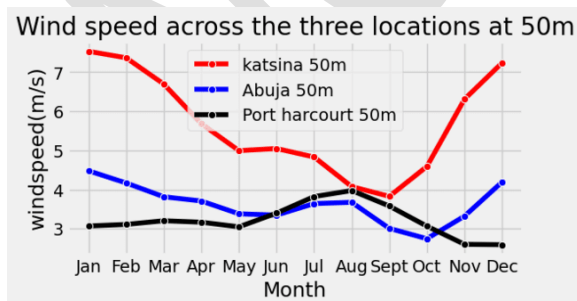


Figure 7. wind speed across the locations at 50m

Monthly trend of wind speed across the three locations.

The result of the monthly trend of wind speed in Katsina with latitude 13°N , Abuja 9°N and Port Harcourt 4.78°N shows the behavioral pattern of wind speed at the various latitudes. It has been observed that the wind speed at 50 meter is higher than that of 10-meter height across the three locations of the study. This can be attributing to the fact that tall objects obstruct wind circulation.

Figure 2 show a graph of wind speed in Katsina with highest wind speed of 7.5m/s at 50m and 5m/s at 10m recorded in January and its lowest value of 3.9m/s at 50m and 2.1m/s at 10m. Katsina is located in the northern part of Nigeria characterize by semi-arid to arid climate typically the Sahel Savannah region, these region is majorly influence by Northeast trade wind or the continental winds blowing from the Sahara Desert during the dry season especially around January as show in figure 2. In August and September, the inter-tropical convergence zone (ITCZ) (is a region where the southeast trade wind meet with the Northeast trade wind) moves Northward weaken the strength of the Northeast trade wind and lower the wind speed. This northward movement of the ITCZ is accompany with a brief raining season due to the moist air from Atlantic Ocean creating more stability and less turbulent condition. The Coriolis effect (the deflection of wind due to earth rotation) is stronger at latitude 13°N and weaken as it approaches the equator.

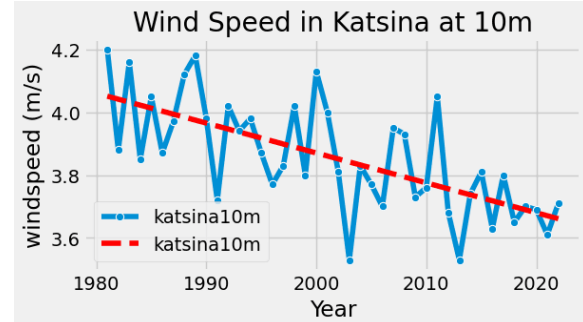
Figure 3 illustrate the distribution of wind speed in Abuja at 10m and 50m, from the graph the highest value of 4.5m/s at 50m and 2.9m/s at 10m in January and its lowest value of 2.7m/s at 50m and 1.5 at 10m in October. Abuja (9°N) has a tropical savannah climate with a balance influence of dry and raining seasons. Wind speed increase in January due to influence of Northeast trade wind originated from Sahara Desert these winds prevail because the ITCZ moves southward. In July and August, increase

in wind speed is also observed, this can be attributed to the fact that the monsoon winds from the Atlantic Ocean get to the region due to the Northward movement of the ITCZ which result to increase wind speed in July and August.

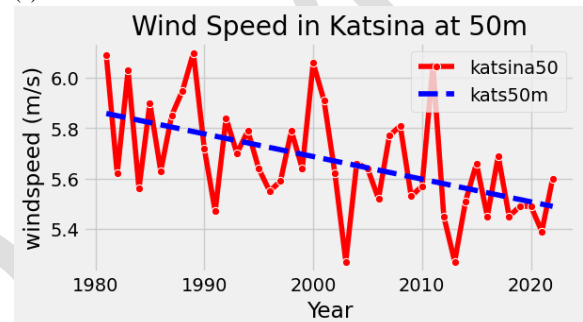
Figure 4 show the wind distribution in Port Harcourt with lat. 4.7°N . From the graph, the wind speed tends to decrease in January with the value of 3m/s at 50m and 1.8m/s at 10m , with an increase of 4m/s at 50m and 2.3m/s at 10m in August respectively. Port Harcourt experiences a tropical rain forest climate as its proximity to Atlantic Ocean. The Northeast trade wind has no influence in this due to its nearest to the equator, the southwesterly wind from the Atlantic Ocean is predominant in this region. In July and August, the ITCZ shift to the north thereby bringing the monsoon rains and increase in wind speed due to the prevailing southwesterly trade winds. In January, the ITCZ moves southward close to the equator causing the wind speed to decrease compare with Abuja and Katsina. This region is popularly called the “doldrum” because of it low wind speed. Port Harcourt is unaffected by the Coriolis effects due to its proximity to the equator. This can be attributed to the counter circulation pattern from southwesterly.

Figure 5 and 6 show the wind speed across the three locations at 10m and 50m . From the graph, the wind speed increase with increase in latitude. Katsina is predominated by the northeast trade wind or the continental wind originated from Sahara Desert but slowdown in August and September due to the shift in ITCZ. Abuja being in the middle is influence by both continental wind and the southwesterly trade winds. Port Harcourt experiences the lowest wind speed due to its proximity to the equator.

Figure 8. Annual trend of wind speed in Katsina

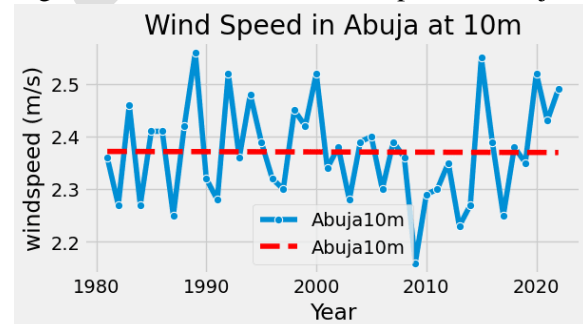


(a)

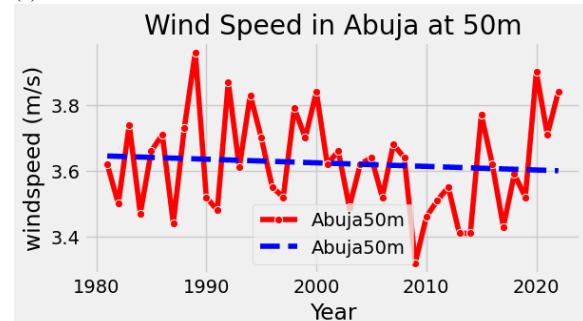


(b)

Figure 9. Annual trend of wind speed in Abuja

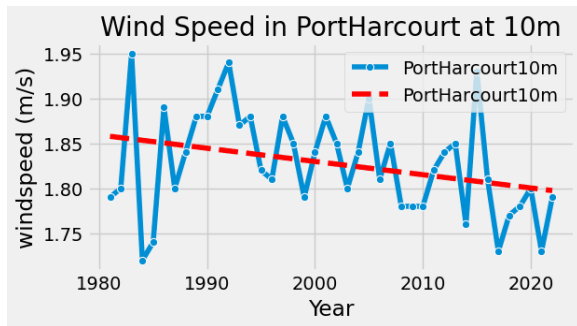


(a)

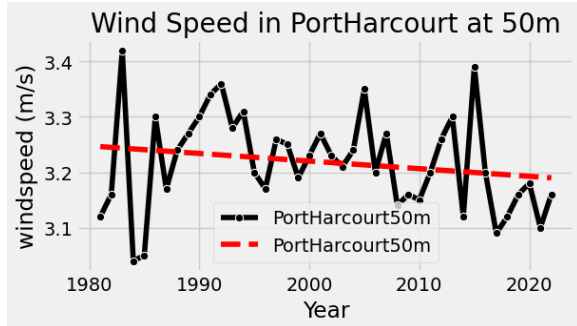


(b)

Figure 10. Annual trend of wind speed in Port Harcourt



(a)



(b)

Figure 8, 9 and 10. Annual trend of wind speed across the three locations

Figure 8a and 8b show the annual trend of wind speed in Katsina lat.130N at 10m and 50m height. The result from 8a indicates a decrease in trend with wind speed of 4.2m/s in 1981 to 3.7m/s in 2022. Similarly, at 50m height the wind speed decreases from 6.2m/s in 1981 to 5.7m/s in 2022. This significant decrease may be due to the region increasing desertification and land degradation. These affect the landscape, the loss of vegetation and increasing soil erosion including changes in surface roughness reduce wind momentum. Changes in West Africa monsoon intensity due to climate change could result in decrease in wind speed as seen in recent decades as the west Africa monsoon weaken the wind speed in the Northern part of Nigeria. These results affirm the report from

intergovernmental panel on climate change (IPCC 2021) that the average wind speed will drop by 10% by 2100.

Figure 9a and 9b show the annual wind speed in Abuja lat.9N. From 9a the wind speed at 10m recorded its highest value of 2.6m/s in 1989 and 2.6m/s in 2016 and lowest value of 2.1m/s in 2009. From 9b, the highest wind speed of 3.9m/s in 1989 and 3.8m/s in 2020, the lowest value of 3.3m/s in 2009 respectively. This region experiences no significant trend because it is influenced by both the Harmattan wind blowing from Sahara Desert and the North and south movement of the ITCZ.

Figure 10a and 10b show the annual trend of wind speed in Port Harcourt lat. 4.7N. From 10a, the maximum wind speed at 10m is 1.95m/s in 1984 and its minimum value is 1.7m/s in 2021. From 10b at 50m the maximum wind speed is 3.41m/s in 1983 and minimum of 3.0m/s in 2016. Port Harcourt is characterized with the lowest wind speed compared with Abuja and Katsina due to its proximity to the equator where it is greatly influenced by the ITCZ.

TABLE 2
Mann-Kendall and Sen's slope Test for Wind speed from the period of 1981-2022

Location	Kendall's Tau	Mann Kendall's Statistic (S)	Test Statistic (Z)	p-value (Two-tailed)	Intercept	Sen's slope (Q)	Test Interpretation	Trend
Katsina 10m	0.50	-433.0	-4.68	2.8e-6	4.02	-0.009	TRUE	decreasing
Katsina 50	0.37	-324.0	-3.5	0.0004	5.82	-0.0008	TRUE	decreasing
Abuja 10m	0.02	-18.0	-0.18	0.85	2.37	-0.00	FALSE	decreasing
Abuja 50m	0.07	-62.0	-0.66	0.50	3.64	-0.001	FALSE	decreasing
Port Harcourt 10m	0.2	-219.0	-2.37	0.01	1.86	-0.002	TRUE	decreasing
Port Harcourt 50m	0.18	-160.0	-1.72	0.08	3.25	-0.002	FALSE	decreasing

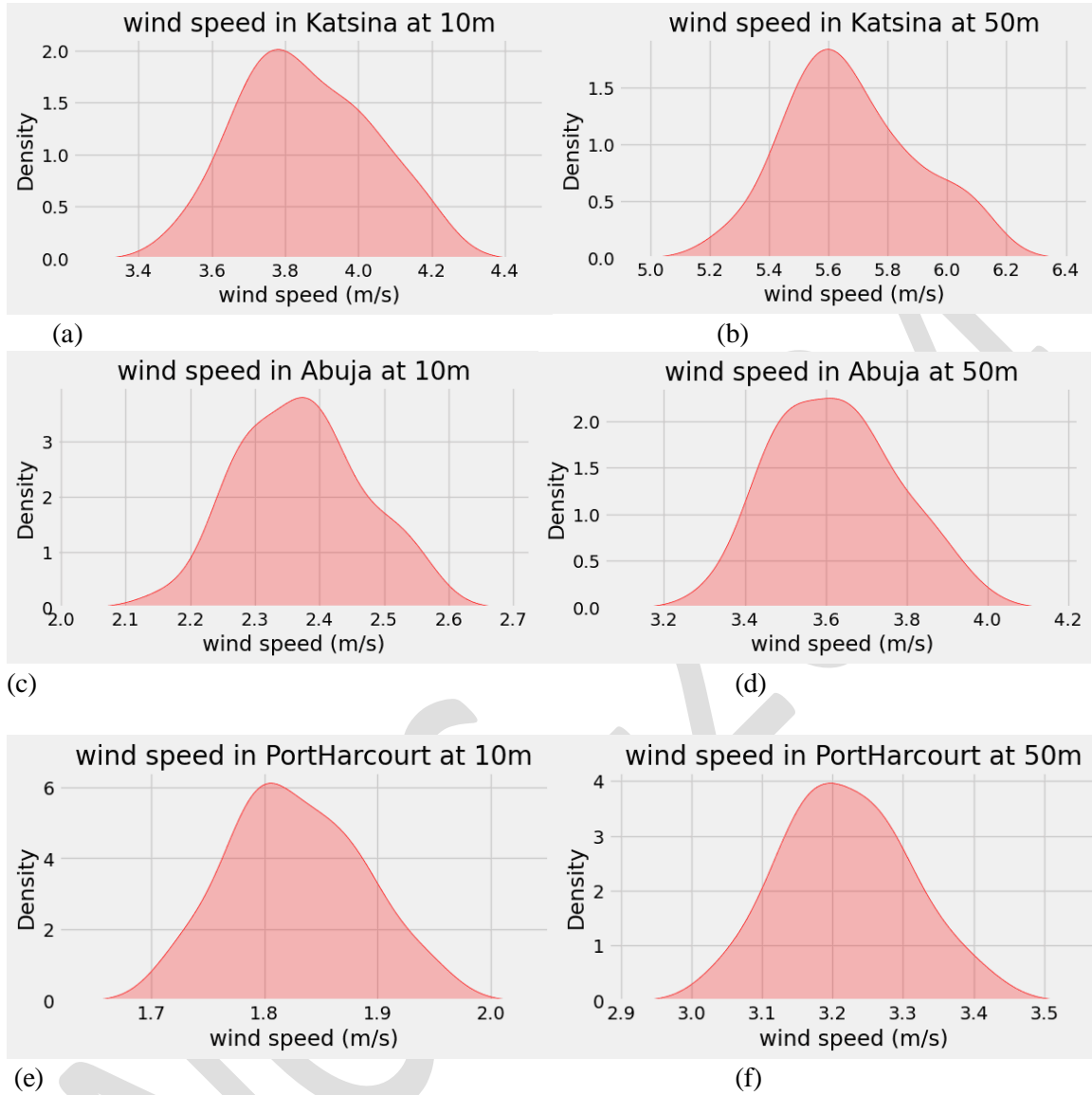
Mann-Kendall and Sen's slope trend test

TABLE 2 show the Mann-Kendall and Sen's slope trend test for wind speed in Katsina, Abuja and Port Harcourt. The Mann-Kendall trend test in Katsina at 10m has a z-statistics which measure the strenght and direction of the trend. The negative z-value of -4.68 reflecting a very strong decreasing trend of wind speed. An extremely small p-value of 2.8e-6 indicates that the wind speed trend is highly statistically significant when compared with the common 5% significance level (0.05). Sen's slope indicates the magnitude of the trend, a negative value of -0.009 suggesting that the wind speed is decreasing on an average by -0.009 unit per time. Similarly, Katsina at 50m with z-value of -3.5 indicates a strong decreasing trend and a small p-value of 0.0004 reflecting the trend is highly statistically significant and a Sen's slope value of -0.0008 indicating the wind speed is decreasing by -0.0008 unit.

Abuja at 10m and 50m with their negative z-value of -0.18 and -0.66 suggesting a very weak decreaseing trend of wind speed . the p-value of 0.85 and 0.50 indicate that the trend is not statistically significant as the p-value is greater than the common 5% significance level. The Sen's slope value of -0.00 and -0.001 reflecting a very weak magnitude per unit time.

Port Harcourt at 10m with a negative Z-value of -2.37 indicating a decreasing trend and a p-value of 0.01 suggesting that the wind speed trend is statistically significant when compare with 5% confident level.the Sen's slope value of -0.002 indicates the wind speed decrease by an average of -0.002 units per unit time. At 50m, the Z-value of -1.72 suggesting a weak decreasing trend is not statistically significant. Sen's slope value of -0.002 suggesting that the wind speed decreases by an average of -0.002 per unit time.

Figure 10. Kernel density estimation



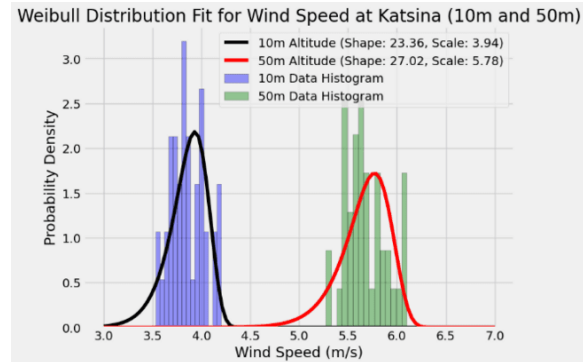
Kernel density estimation (KDE) plot have been adopted as a better way to visualize the distribution of data, it provide a smooth estimate of the probability density function(PDF) of a continuous random variable. Figure 10a and 10b show a KDE plot of Katsina with a unimodal distribution, the peaks of the curves are 3.8m/s and 5.6m/s at 10m and 50m high are the most frequent values in the distribution, the values represent the mode or the highest probability

density. Figure 10c and 10d show KDE plot in Abuja with peaks values of 2.4m/s and 3.6m/s at 10m and 50m here the values are lower compare to that of Katsina.

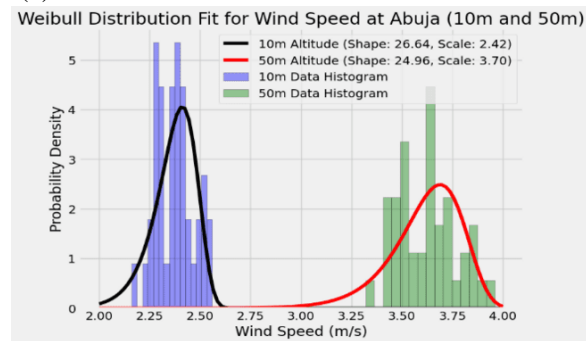
Figure 10e and 10f show KDE plot in PortHarcourt with peaks values of 1.8m/s and 3.2m/s at 10m and 50m. This result indicates that majority of the data are concentrated between 1.8 and 3.2m/s.

Weibull distribution

Figure 11. Weibull distribution across the three locations.

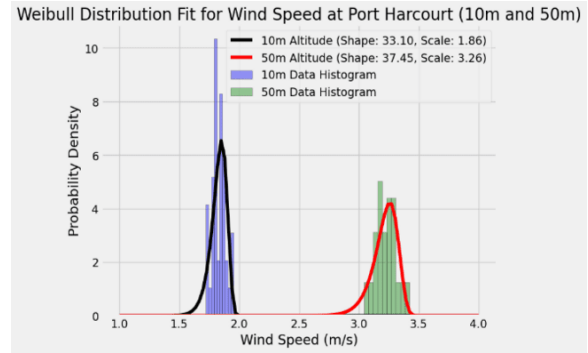


(a)



(b)

Weibull distribution is commonly used for wind energy analysis due to its effectively characterizes wind speed variations. In figure 11a, Katsina shows, the Weibull parameters provided shape parameter $k=23.36$ and scale parameter $c=3.94\text{m/s}$ at 10 meters high, while $k=27.02$, and $c=5.78\text{m/s}$ at 50 meters high. The wind speed at 10m indicates low to moderate variability this mean that there are fairly consistent near the surface. The c parameter indicates moderate average wind speed of approximately 3.94m/s . At 50 meters high, the wind speed suggest a even lower variability with more stable condition, which is advantageous for energy production. The c parameter reflect a significant increase in wind speed compared to that of 10meters, with an average of 5.78m/s . The wind speed improvement can be related with the wind shear phenomenon, where wind speeds increase with height due to reduced surface friction. The wind speed in Katsina at



(c)

10meters can e seen to be suitable for small scale application while that of 50meters are favourable for wind energy generation.

Figure 11b illustrate the weibull distribution of wind speed in Abuja at 10meters and 50meters.

The k parameter of 26.64 indicates that the wind speeds at 10 meters are highly stable with low variability, the scale parameter of 2.42m/s indicates that wind speeds are been clustered at these point. It suggest a low average wind speeds at 10 meters which is unsuitable for commercial wind energy applications. At 50 meters, the k parameter of 24.96 indicates a slight increase in wind speed variability compared to 10 meters, the scale parameter $c = 3.70\text{m/s}$ reflects a noticeable increase in wind speed compared to 10 meters due to surface friction although it still fall below the optimal range for efficient energy generation by a standard wind turbines.

Figure 11c shows the weibull distribution in Port Harcourt with shape parameter of $k=33.10$ and scale parameter of $c=1.86\text{m/s}$ at 10 meters

height the result indicate a narrowly distributed wind speed around the mean value and very high indicating consistent but weak winds. The average wind speed is 1.86 m/s, making this height unsuitable for any meaningful wind energy generation. At 50 meters, the k parameter suggests exceptional stability in wind speeds. The distribution of the wind speed is even narrower compared to 10 meters. The improvement reflects the wind shear effect, where wind speeds increase with altitude due to reduced surface friction.

However, the overall wind resource remains weak. Base on the latitudinal distribution, the wind speeds tend to increase with latitude due to climatic and geographical factors. Katsina

stability with extremely low variability,

(northern Nigeria) experiences higher wind speeds, while Port Harcourt (southern Nigeria) has weaker winds due to its coastal, equatorial location. Wind speeds increase significantly with height across all locations due to the wind shear effect, but the magnitude of increase varies. Katsina benefits the most, with a significant jump from 3.94 m/s to 5.78 m/s. the higher k-values in Port Harcourt reflect greater wind stability, while moderate k-values in Katsina suggest higher variability, which can occasionally favor stronger winds for energy generation.

Wind Power Density

Wind power density measure the amount of energy available per unit area (W/m^2). It greatly depends on the density of air, surface pressure, wind speed and cross-sectional area of the wind.

Location	Height (meters)	Wind power density
Katsina	10	74.88 W/m^2
	50	182.68 W/m^2
Abuja	10	31.94 W/m^2
	50	62.77 W/m^2
Port Harcourt	10	23.08 W/m^2
	50	83.59 W/m^2

The result indicates a general increase in wind power density at 50-meters height compared to that of 10 meters height. Katsina recorded the highest wind power density of 182.68 W/m^2 suggesting a high potential for wind energy

generation, it can be attributed to the geographical position of Katsina (Northern Nigeria) as wind speed increase with Latitude. Abuja and Port Harcourt has the lowest wind speed, less than the recommended wind speed of ($\geq 6\text{m/s}$), these suggest an alternative renewable system such as solar energy. These locations are seen to be close to the equator and its greatly influence by the intertropical convergence zone (ITCZ)

5.0 Conclusion

This research work studied the long-term trend and spatial variability in wind speed across Katsina, Abuja, and Port Harcourt in Nigeria between 1981 and 2022. The result indicates that there is considerable variation in the increase of wind speed with latitude; the highest being at Katsina, while the least is at Port Harcourt. Although both the Mann-Kendall trend test and Sen's slope estimator confirmed the statistically significant decline in wind speeds at 10 m and 50 m heights for Katsina, most of the results for Abuja and Port Harcourt show a weakly decreasing or non significant trend, except for the decline at 10 m in Port Harcourt.

The results reflect the dominance of geographical factors and atmospheric

circulation, including the ITCZ, Northeast trade winds, and Coriolis effect. A potential gradient or wind shear was witnessed, an increase in the velocity of winds with height, as all locations showed this phenomenon. The maximum wind power density at 50 m level was at Katsina, which therefore presents the most promising area for wind energy development. The low wind speed estimations over Abuja and Port Harcourt suggest that there will be limited utility for commercial wind power generation, thus favoring alternative sources of renewable energy, for instance, solar energy.

These findings are very important in formulating region-specific renewable energy policies and for responding to the impacts of climate change on available wind resources. The study lays a foundation for further investigation into viable wind energy options, especially in northern Nigeria, which bears the greatest potential for wind energy.

- Abiodun Isaac Chukwutem and Anisiji Obiora Emeka. 2024. "Statistical Analysis and Estimation of Wind Speed Characteristics and Energy Potential for Power Generation in Forcados, South Southern, Nigeria.," September. <https://doi.org/10.5281/ZENODO.13768078>.
- Adaramola, M.S., and O.M. Oyewola. 2011. "On Wind Speed Pattern and Energy Potential in Nigeria." *Energy Policy* 39 (5): 2501–6. <https://doi.org/10.1016/j.enpol.2011.02.016>.
- Agbo, E P, and C M Ekpo. 2021. "Trend Analysis of the Variations of Ambient Temperature Using Mann-Kendall Test and Sen's Estimate in Calabar, Southern Nigeria." *Journal of Physics: Conference Series* 1734 (1): 012016. <https://doi.org/10.1088/1742-6596/1734/1/012016>.
- Agbo, Emmanuel P., Chris M. Ekpo, and Collins O. Edet. 2020. "Trend Analysis of Meteorological Parameters, Tropospheric Refractivity, Equivalent Potential Temperature for a Pseudoadiabatic Process and Field Strength Variability, Using Mann Kendall Trend Test and Sens Estimate." arXiv. <https://doi.org/10.48550/ARXIV.2010.04575>.
- Agbo, Emmanuel P., Ugochukwu Nkajoe, and Collins O. Edet. 2023. "Comparison of Mann–Kendall and Şen's Innovative Trend Method for Climatic Parameters over Nigeria's Climatic Zones." *Climate Dynamics* 60 (11–12): 3385–3401. <https://doi.org/10.1007/s00382-022-06521-9>.
- Akdağ, Seyit A., and Ali Dinler. 2009. "A New Method to Estimate Weibull Parameters for Wind Energy Applications." *Energy Conversion and Management* 50 (7): 1761–66. <https://doi.org/10.1016/j.enconman.2009.03.020>.
- Chinachi, E. A., B. N. Nwankwojike, and R. Uche. 2022. "Time Series Assessment and Simulation of the Effects of Climate Change on Sustainable Wind Power Generation in Nigeria." *International Journal of Energy and Water Resources*, September. <https://doi.org/10.1007/s42108-022-00217-w>.
- Danlami, Dantata, Saidu Idris, Richard Sunday Thlakma, and Golly Sammy Gwandum. 2019. "THE SPATIO-TEMPORAL VARIATIONS OF WIND SPEED DURING HARMATTAN SEASON IN NORTHEASTERN NIGERIA." *Geosfera Indonesia* 4 (2): 105. <https://doi.org/10.19184/geosi.v4i2.11474>.
- Fagbenle, R.O., J. Katende, O.O. Ajayi, and J.O. Okeniyi. 2011. "Assessment of Wind Energy Potential of Two Sites in North-East, Nigeria." *Renewable Energy* 36 (4): 1277–83. <https://doi.org/10.1016/j.renene.2010.10.003>.
- Fashae, Olutoyin, Adeyemi Olusola, and Oluwatola Adedeji. 2017. "Geospatial Analysis of Changes in Vegetation Cover over Nigeria." *Bulletin of Geography. Physical Geography Series* 13 (1): 17–27. <https://doi.org/10.1515/bgeo-2017-0010>.
- Igbawua, Tertsea, Jiahua Zhang, Qing Chang, and Fengmei Yao. 2016. "Vegetation Dynamics in Relation with Climate over Nigeria from 1982 to 2011." *Environmental Earth Sciences* 75 (6): 518. <https://doi.org/10.1007/s12665-015-5106-z>.
- Mudasiru, Mustapha, and Mohd Wazir Mustafa. 2018. "Wind and Solar Radiation Potential Assessment in Kano, Nigeria Using Weibull and Samani Models." *ELEKTRIKA- Journal of Electrical Engineering* 17 (1): 21–27. <https://doi.org/10.11113/elektrika.v17n1.69>.
- Mzough, Oranderen Matthew, and Ewa-henshaw King Ededem. 2023. "ASSESSMENT OF WIND ENERGY RESOURCE AND POWER PRODUCTION ESTIMATES: A CASE STUDY OF CALABAR, SOUTH EASTERN NIGERIA." <https://doi.org/10.14293/PR2199.000029.v1>.

- Ohunakin, Olayinka S., and Olaolu O. Akinawonu. 2012. "Assessment of Wind Energy Potential and the Economics of Wind Power Generation in Jos, Plateau State, Nigeria." *Energy for Sustainable Development* 16 (1): 78–83. <https://doi.org/10.1016/j.esd.2011.10.004>.
- Okono, M. A., E. P. Agbo, B. J. Ekah, U. J. Ekah, E. B. Ettah, and C. O. Edet. 2022. "Statistical Analysis and Distribution of Global Solar Radiation and Temperature Over Southern Nigeria." *Journal of the Nigerian Society of Physical Sciences*, August, 588. <https://doi.org/10.46481/jnsps.2022.588>.
- Oteng Mensah, Felix, Clement Aga Alo, and Duke Ophori. 2024. "Hydroclimatic Trends and Streamflow Response to Recent Climate Change: An Application of Discrete Wavelet Transform and Hydrological Modeling in the Passaic River Basin, New Jersey, USA." *Hydrology* 11 (4): 43. <https://doi.org/10.3390/hydrology11040043>.
- Oyedepo, Sunday O, Muiyiwa S Adaramola, and Samuel S Paul. 2012. "Analysis of Wind Speed Data and Wind Energy Potential in Three Selected Locations in South-East Nigeria." *International Journal of Energy and Environmental Engineering* 3 (1): 7. <https://doi.org/10.1186/2251-6832-3-7>.
- Safari, Bonfils. 2011. "Modeling Wind Speed and Wind Power Distributions in Rwanda." *Renewable and Sustainable Energy Reviews* 15 (2): 925–35. <https://doi.org/10.1016/j.rser.2010.11.001>.
- Tuller, Stanton E., and Arthur C. Brett. 1984. "The Characteristics of Wind Velocity That Favor the Fitting of a Weibull Distribution in Wind Speed Analysis." *Journal of Climate and Applied Meteorology* 23 (1): 124–34. [https://doi.org/10.1175/1520-0450\(1984\)023<0124:TCOWVT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<0124:TCOWVT>2.0.CO;2).

Assessing the Correlation between Seasonal Atmospheric Heating and Ionospheric TEC Dynamics: Insights from Spring Equinox and Summer Solstice and Fall

Awak, E. A¹., George, A. M¹, Okono, M. A¹., and Odong, P. O².

1. Department of Physics, University of Calabar, Calabar, Cross River State, Nigeria.

2. Department of Earth Science, Arthur Jarvis University, Akpabuyo, Cross River State, Nigeria.

Abstract

This study investigates the correlation between seasonal atmospheric temperature profiles and variations in Total Electron Content (TEC) in the ionosphere during the spring equinox, summer solstice and fall at the GPS station in Mbarara, Uganda (MBAR, Lat: -0.601 and Long: 30.738). Utilizing high-resolution atmospheric data from local meteorological stations and TEC measurements derived from GPS observations for eleven years, 2008 to 2018, we analyzed temporal variations in atmospheric temperature, specific humidity and surface pressure TEC. The overall results showed significant negative correlations between increased atmospheric temperatures and reduction in TEC levels, increase in surface pressure with decrease in TEC and finally increase in specific humidity with TEC, particularly during the summer solstice, suggesting a potential influence of solar radiation on ionospheric dynamics. Temperature and TEC parameters show significant density correlation between 20 – 21°C and 10 – 12 TECU in other years but in solar maximum years, the results of the correlation moved and settled around 20 – 22°C and 25 – 35 TECU, indicating high releases of solar radiation which influences the Earth thermosphere, leading to disruption of satellite operation, also causing atmospheric expansion and create potential environmental impact. The coefficients of correlation for the eleven years varies from -0.44 to -0.91 for Temperature and TEC, -0.26 to 0.90 for surface pressure and TEC, and 0.56 to 0.99 for specific humidity. The results also showed moderate to high negative relationships between temperature and TEC with meta-analysis statistics of $i^2 = 35.1\%$, $Q = 174.21$ ($p < 0.001$) and $\tau^2 = 16.52$, for TEC and surface pressure, the statistics were $i^2 = 0\%$, $Q = 10.42$ ($p = 0.416$) and $\tau^2 = 0$ and for TEC and specific humidity; $i^2 = 42.16\%$, $Q = 143.21$ ($p < 0.001$) and $\tau^2 = 6.57$. This study enhances the understanding of ionospheric responses to seasonal atmospheric changes, providing valuable insights for geophysical research and practical applications in satellite communications and navigation in equatorial regions.

Keywords: Ionosphere, Atmosphere, TEC, Temperature, Surface pressure, Surface humidity, Solstice, Equinoxes

1. Introduction

The dynamisms of the Earth's atmosphere are influenced greatly by seasonal variations in solar heating. This, in turn has great effect on the atmosphere and its phenomena, such as Total Electron Content (TEC). In understanding the behavior of the ionosphere, TEC is a crucial parameter, as it gives the total electron density in range of atmospheric column above a specific area, especially from a satellite to a ground receiver (Bhuyan& Borah, 2007). This parameter particularly, finds important applications in satellite communications, navigation systems, commercial services, civilian services, and space weather forecasting (Adewalea et al., 2013; El-Rabbny, 2002; Ciruolo et al., 1994; Mannucci et al., 1993. Ionospheric TEC varies with events such as diurnal variations, seasonal variations, relative locations to the magnetic equator (latitudinally), etc. Most of these

events are caused by solar activities (Liu and Chen, 2009; (Klobuchar, 1987; Dabbakuti et al., 2016; Sridhar et al., 2016, Ratnam et al, 2015; 2017).

For several reasons, research in this area is crucial and significance as it contributes to the increasing knowledge of the behavior of the ionospheric, particularly in the regions where atmospheric phenomena are prominently pronounced due to high increased in solar radiations (Rishbeth, 1998). Also, satellite communication systems and navigation technologies usually suffer disruptions caused by ionospheric irregularities, so the understanding of these dynamics can lead to an improved precursor to avert the challenges. Furthermore, electricity power grid usually suffers collapses, which may be as a result of geomagnetic storm, for instance in Nigeria (equatorial region), the collapses of National electricity grid had moved from two per year to about seven per year (The Nation, 2024), the understanding of the variability of these parameters can lead to a permanent solution of this blackout menace. Finally, insights gained from this research may also help in predicting the impacts of climate change on atmospheric dynamics and ionospheric behavior, which is increasingly relevant in today's context of global warming. In this work, we assessed the correlation between seasonal atmospheric heating, as defined by local solar metrics (temperature, surface pressure and specific humidity), and the corresponding variations in ionospheric TEC during these periods. By utilizing ground-based measurement data and advanced statistical techniques, the research seeks uncover patterns and that will contribute to a more nuanced understanding of the ionospheric dynamics influenced by seasonal atmospheric conditions in the region. Through this exploration, we will attempt to provide valuable insights into the mechanisms driving ionospheric behavior in response to seasonal atmospheric heating, thereby enhancing our predictive capabilities in this critical area of atmospheric science. Zhang et al., (2017) noted that long term variation of TEC and thermospheric temperature showed fluctuation during the 24th Solar Cycle caused by solar and geomagnetic storm. Their work has brought some insights into this current research. The variations in thermospheric temperature and Ionospheric TEC are significantly seen to be impacted by energy from solar events which affected the thermal characteristics of the thermosphere, resulting in changes in emission of molecules such as nitrogen monoxide (NO) and carbon monoxide (CO) which have direct bearings with atmospheric temperature (Liu et al., 2006; Liu and Chen, 2009). Changes in temperature in specific periods within a year directly affect ionospheric electron density (Chen et al., 2020; Mendillo, 1974).

2. Data and Methodology

2.1 TEC data

The ionospheric TEC data used in this work were accessed through <https://unavco.org/pub/rinex> and <https://sonel.org>. These RINEX format data are unreadable. The algorithm by Gopi GPS-TEC, version 2.9.5 (Gopi, 2017) processed and converted the data into human readable format (Vertical TEC) which is a converted version of the measured TEC (slant TEC) (eq. 1)

$$vTEC = M(e) * sTEC \quad (1)$$

2.2 Solar metric data

The temperature was downloaded from (<https://pvgis.com>) at 2m above the mean sea level. On the other hand, surface pressure and specific humidity data at the same coordinates and location as TEC

and same height as temperature were downloaded from (<https://power.larc.nasa.gov/data-access-viewer/>). Since temperature and these parameters are directly related, they were taken at the same high as temperature. The segmentations of the data were done by considering the seasonal periods of spring equinox, summer solstice and fall.

Equations 2 and 3 show the linear relationship of the variables.

$$TEC = S_1 M + S_o \quad (2)$$

Where M is the independent variable (temperature/specific humidity) S_1 is the slope and S_o is the TEC regression line intercept.

$$sP = P_1 TEC + P_o \quad (3)$$

Where TEC is the independent variable P_1 is the slope and P_o is the intercept of surface pressure regression.

Also, Mann-Kendall (MK) test trend model shown in Eq. 4, was used to statistically assess if there is a monotonic upward or downward trend of the TEC and the solar metrics of interest over selected period of time (Kendall, 1975; Mann, 1945). A monotonic downward (upward) trend implies that the variable consistently decreases (increases) through time, though the trend may or may not be linear (Hirsch et al., 1992)

$$Z_{MK} = \begin{cases} \frac{S-1}{\sqrt{V(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sqrt{V(S)}} & \text{if } S < 0 \end{cases} \quad (4)$$

2.3 Meta-analysis statistical approach

We applied meta-analysis method to combine the linear regression equations obtained from the relationships between the parameters. The coefficients (S_1 and S_o) and standard errors (SE) from each of the 11 regression equations were identified for each of the relationship between solar metrics and TEC. After the calculation of weighted average coefficients (WAC) based on inverse of the variance of the coefficient, the equations were tested for heterogeneity to obtain. We use comprehensive meta-analysis (CMA) software V4 for this processing (Borenstein et al., 2002).

3. Result and discussion

This work aimed to establish the relationship of Ionospheric TEC with solar metrics (temperature, surface pressure and specific humidity) which contribute to the atmospheric heating during the three seasons of spring equinox, summer solstice and the fall over a decade (2008 – 2018). The work considers the seasons to span from the ending period of spring equinox (April/May), the whole of summer solstice and to the end of fall (November/December).

3.1 General picture of TEC and temperature

Fig. 1a shows the variation of TEC with temperature during the years (2008 – 2018) at the three seasons; Spring equinox (the months of March, April and May), Summer solstice (June, July and

August) and the Fall (September, October and November). The observation shows similar trend of increases in temperature with decreases in TEC from the end of spring equinox, particularly from the month of April/May. The temperature gets to the maximum and TEC gets to minimum during summer solstice. The beginning of the fall (autumn equinox) marks the season of exchange of phase where the temperature drops and TEC rises to their respective minimum and maximum at the middle of the season. The year 2008 marked the end of 23rd solar cycle and the beginning of the 24th solar cycle few solar flares and coronal mass ejections (CMEs) occurred this year. The low coefficient of correlation in 2017 could also be traced to low solar radiation which could be attributed to the natural decline of Solar Cycle 24. As this cycle progressed towards its minimum, solar activity decreased, resulting in lower solar radiation, with average sunspot number of 15(Zharkova et al., 2023).

Fig. 2 displays the bivariate KDE plot joint distribution of temperature and TEC, where the NW – SE trending direction of the plot reveals the inverse relationship of the two variables. The plot also reveals that much of the temperature data within the period concentrated between 21.5 – 22.5 °C with corresponding TEC values concentrating around 10 – 18 TECU.

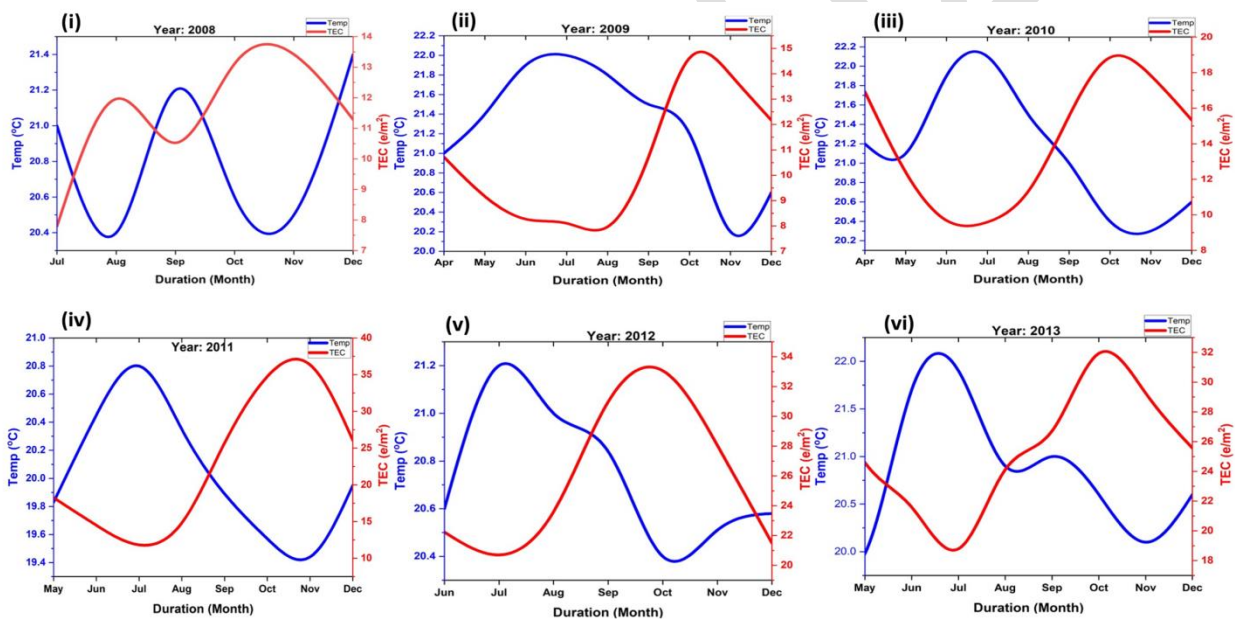


Fig. 1: Variation of TEC with atmospheric temperature during the seasons (2008 – 2013)

The densely populated regions indicate the phase changes where increase in temperatures over the periods coincides with increase in TEC Fig. 2b(i) shows that the temperature peaks at about 21.8°C indicate the mode of the values with highest probability density about 0.55, the highest frequency. The narrow spread between 19°C to 23°C indicates relatively small fluctuations in the heating of the atmosphere. The lightly skewed distribution toward the right suggests the tendency towards warmer temperature which corresponds to the left skewed of the TEC (Figure 2b(ii)) confirming strong negative correlation existing between the two variables.

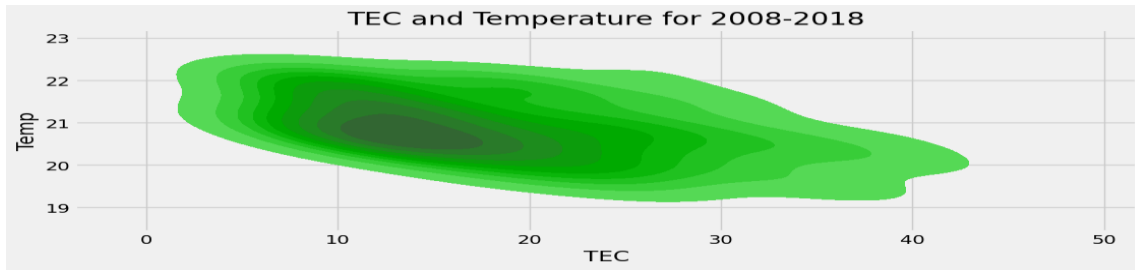


Fig. 2a: KDE plot showing inverse relationship between temperature and TEC

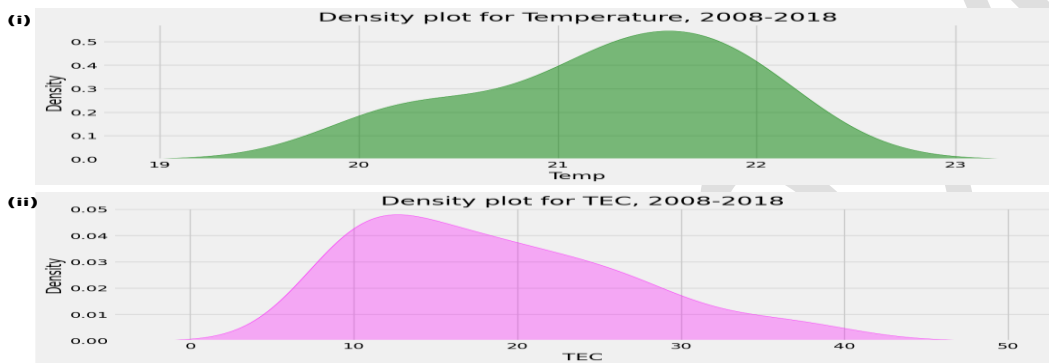


Fig. 2b: Density plots of temperature and TEC

Table 1 summarizes the relationship insights between the two variables;

Table 1
Temperature and TEC density plot insights

Plot insight	Temperature (Fig. 2b(i))	TEC (Fig. 2b(ii))
Mode/Implication	21.8°C/ Mild fluctuation atmospheric heating	15 TECU/ Quiet ionospheric conditions
Probability density	0.55	0.048
Right tail/Implication	23°C /Tending to warmer temperature	50 TECU/Ionospheric disturbance
Left tail/Implication	19°C /Mild temperature caused by spring equinox season	6 TECU/Natural boundary of periods of low ionization or ionospheric depletion.
Range/Implication	19 – 23°C/Relatively low atmospheric heating.	6 – 46 TECU/Variable electron density
Distribution type	Asymmetric	Asymmetric
Skewness	Right	Left

3.2 Correlation between TEC and surface pressure

Several parameters contribute to the heating of the atmosphere. To further assess the correlation of this heating and TEC, surface pressure data were correlated with TEC data. Fig. 3. Show the variation of surface pressure with TEC. Towards the end of spring equinox, it was discovered that

the surface pressure values increased with decrease in TEC, to maximum in the summer solstice before decreasing with increase in TEC towards the end of the season.

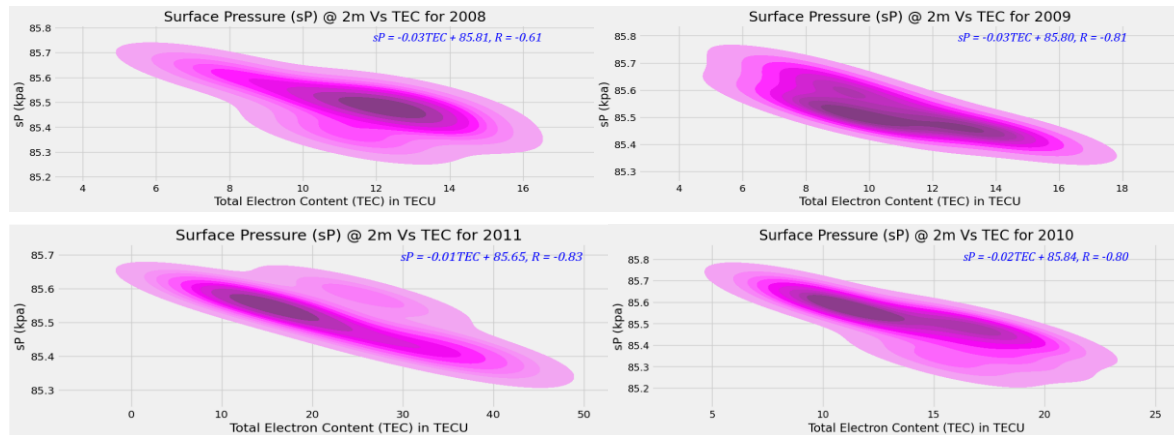


Fig. 3, Relationship Between Surface Pressure and Ionospheric TEC (2008 – 2010)

The early fall marked the exchange of the phase between the surface pressure and TEC. The NW – SE trend of the surface pressure - TEC plots shown in Figures 3 indicate the strong negative relationship between these parameters. The heating pattern contribution of surface pressure aligns with those of the temperature shown in Fig. 1.

The correlation plots also show that through Hadley cell, surface pressure influences the heating of the atmosphere of this region thermodynamically. In the tropical region, surface low pressure causes air to rise, creating convective currents, hence releasing latent heat to the atmosphere, hence, increase in temperature which leads to low TEC. A very coefficient was recorded in 2012. This year was near the peak of Solar Cycle 24 which ionospheric conditions could have altered the atmospheric condition through increased geomagnetic activity. Also, climatologically, the atmosphere in the year experienced a weak El Nino Southern Oscillation (ENSO) event could be attributed, which could have caused the weak recorded correlation. Additionally, the changes in Arctic Oscillation (AO) which has influence on surface pressure and TEC could have caused the uncorrelation. Respectively, the average surface pressure values declined from 85.6 kPa to 85.5 kPa to 85.3 kPa until it inclined increased to 85.5 kPa in 2015. These data confirm as the atmosphere is heated, the electron density in the Appleton zone (EIA) decreased.

3.3 Correlation between TEC and specific humidity

The mass of water vapor present in a unit mass of air plays a significant role in atmospheric heating thereby correlating with TEC. Figs. 4 show the correlation between TEC and specific humidity.

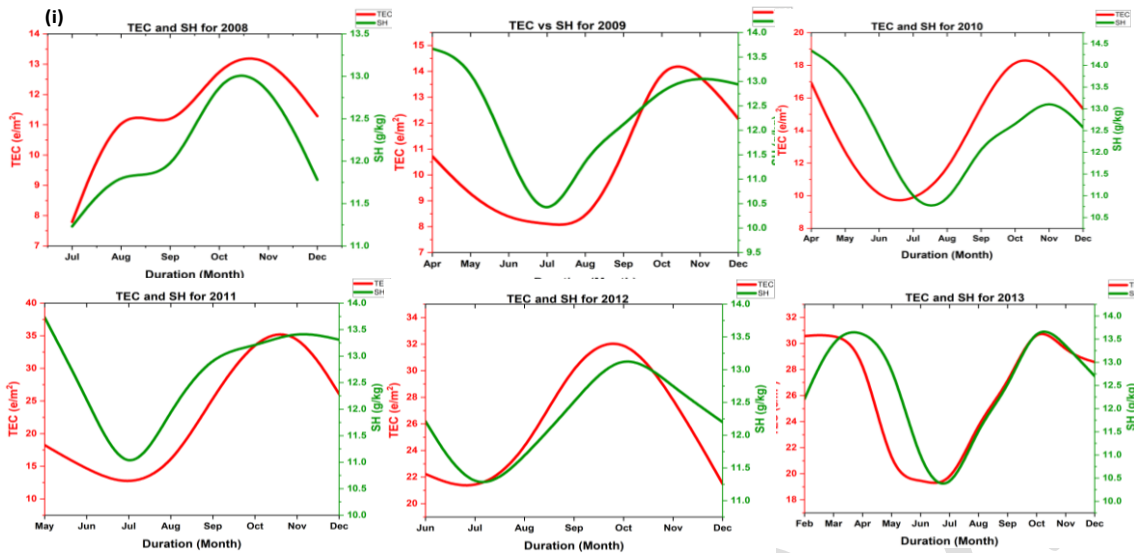


Fig. 4: Correlation plot between TEC and specific humidity

When the solar radiation reaches the Earth's surfaces, supplies it with heat through conduction. The counter infrared radiation causes the warm air to rise on the tropical regions. In the atmosphere, the air cools down and falls as dry air around latitude 30°N or S (Hadley cell). Through atmospheric electricity, as water vapour in this warm air gets ionized in the upper atmosphere and releases electrons and positive ions, which in turns influences the electron density which contribute to TEC, hence, the positive correlation (Figures 4). Through charge separation, water droplets in the atmosphere becomes electrically charged through collision with other particles and the process can enhance the release of electrons in the EIA zone leading to strong positive correlation between the parameters as shown in Fig. 5.

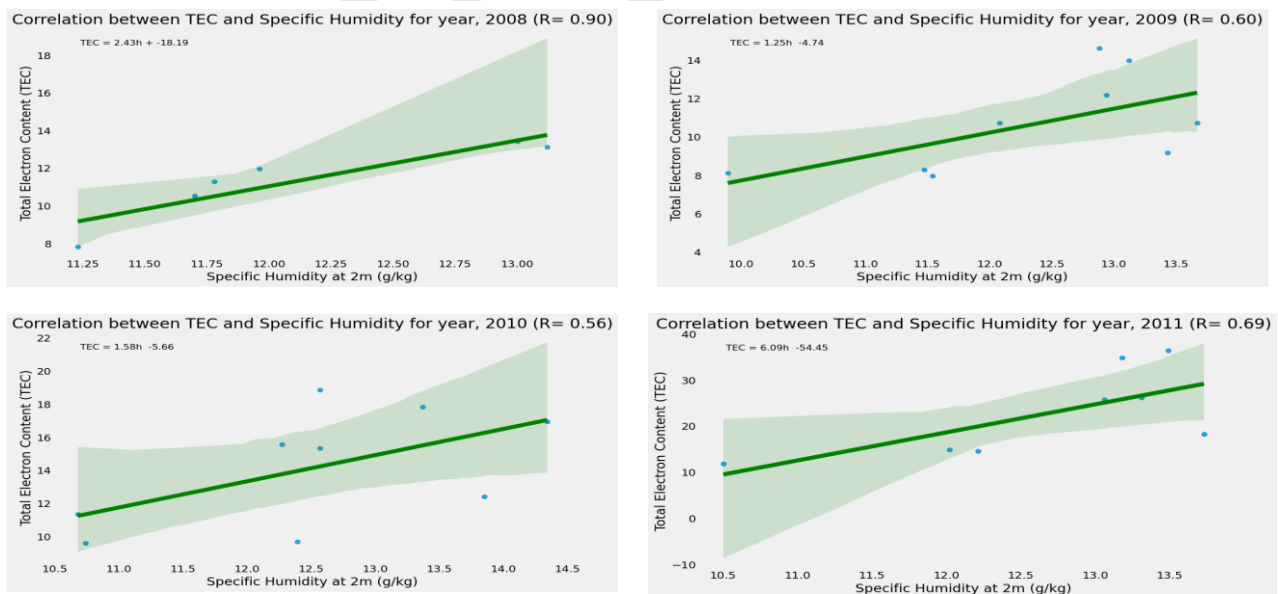


Fig. 5: Correlation plot between TEC and specific humidity

During summer solstice, there is reduction in solar radiation which leads to reduction in the ionization. Weaker geomagnetic field in summer also reduces ionospheric electron density. These two processes contribute to the reduction in TEC during summer. As stated in section 3.1, lower thermospheric temperature which is caused by high atmospheric temperature, also leads to a reduced ionization during summer solstice and contribute to the low TEC. Due to the overwhelming power of the NE trade winds, during summer, the SE trade winds become weaker in the tropic and it reduces the moisture transport, hence the drop in specific humidity.

Table 2 shows the insights from Mann-Kendall test to the monotonic trend of the analyzed result. The null hypothesis, $h = F(F - false)$ in the Mann-Kendall program which (positing that there is no trend in the data) is not rejected, indicating that the TEC data are fluctuations without evidence of particular trend of variability within the period.

Table 2:
Mann-Kandell monotonic test model results over the period

Parameter	τ	z	p-val	slope	h	Trend
TEC (TECU)	0.095	1.305	0.191	0.046	F	No
Temp (°C)	-0.043	-0.583	0.559	-0.001	F	No
sP (kPa)	0.150	2.055	0.039	0.0008	T	increasing
SH (g/kg)	0.071	0.968	0.333	0.004	F	No

The MK test model detected a statistically significant increasing trend in the variability of surface pressure (sP) data over the analyzed period of time. The null hypothesis is ($h = T$) is rejected because the p-value of 0.039 is well below the threshold of 0.05, the z-value (2.055) is greater than the standard value of 1.96, the tau (0.150) indicates a moderate positive association between the variables, suggesting a tendency for increases in its value with time. The increasing rate of surface pressure poses pollution accumulation risk as high-pressure systems can trap pollutants close to the surface, leading to poorer air quality and increased health risks, particularly in urban areas, also the region may stand a risk of wildfire due to dry condition as a result of increasing surface pressure. High increase pattern also poses heat distribution risk due to extreme temperature; weather system stands a risk of being altered by potential prolonged drought.

There is no significant trend in the variability of specific humidity (SH) as the p-value is more than 0.05. The z-value of 0.968 indicate weak increasing trend.

Table 3 gives generalized regression equations (GREs) for the relationship between the TEC and the heating solar metrics under consideration. The GREs were formed after subjecting the coefficients of model linear regression equations to CMA software. The software program calculated the weighted average coefficients (WAC) and the standard errors (SE).

The relationship between surface pressure (sP) and TEC indicates that for every unit increase in TEC, sP decreases by 0.02 units.

In table 3, the 0.99 value for R^2 , the coefficient of determination implies that approximately there is 99% of the variability in the sP and TEC, hence, indicating an excellent fit of the GRE, $i^2 = 0\%$ indicates no heterogeneity between the model equations for the 11 selected years, suggesting no significant differences between the equations.

Table 3:
Generalized parametric equations for TEC relationship with solar metrics

Relationship	WAC		SE		R^2	GRE	Meta-Analysis Statistic
	S_1	S_o	S_1	S_o			
sPvs TEC	-0.02	85.82	0.005	0.07	0.99	$sP = -0.02TEC + 85.82$	$i^2 = 0\%, Q = 10.42 (p = 0.0406), \tau^2 = 0$
TEC vs SH	3.41	-23.11	0.21	1.43	0.81	$T = 3.41h - 23.11$	$i^2 = 95.6\%, Q = 241.1 (p < 0.001, \tau^2 = 13.41)$
TEC vs Temp	-5.51	121.41	0.32	2.35	0.72	$T = -5.51m - 121.41$	$i^2 = 95.5\%, Q = 241.12 (p < 0.001) \tau^2 = 23.11$

The Cochran's Q statistic of 10.42 also indicates no significant heterogeneity between the model equation. The $\tau^2 = 0$ indicate that the variance between the equations is zero, confirming the consistency in the relationship. The implications of these statistics suggest a strong negative linear relationship between surface pressure and TEC; hence, the GRE is fit for prediction.

The relationship between TEC and SH indicates moderate to high predictive power of the GRE ($R^2 = 0.81$), as there are 3.41 units increase in TEC for every unit increase in SH. From the meta-analysis statistics, there is high percentage of heterogeneity in the model equations indicating significant differences among them which suggest that there are other factors leading to the positive linear relationship between them. These factors could probably be traced to study-level factor such as discrepancies in sampling rate, instrumental factor such as instrument calibration or environmental factors such as climate variability, natural climate fluctuations such as El Nino-Southern Oscillation (ENSO). To address the heterogeneity, the equations were subjected to Meta-Regression Analysis (MRA) and a final regression equation (Eq. 5) was obtained with a reduced meta-analysis statistic.

$$TEC = 3.21h - 20.53 + 0.05n \quad (5)$$

In Eq. 5, n represent the number of samples, and the reduced statistics are $i^2 = 41.2\%$, $Q = 143.21 (p < 0.001)$ and $\tau^2 = 6.51$. This equation is suitable for prediction.

For every unit rise in temperature, TEC is observed to drop by 5.51 units (Table 3), confirming a negative relationship existing between them. The coefficient of determination, $R^2 = 0.72$ indicates that the model explains approximately 72% of the variability in TEC, indicating a moderate to good

fit equation. The high i^2 value of 95.60% and Q value of 241.12 jointly signify significant heterogeneity between the 11 equations. The $\tau^2 = 23.11$ indicates substantial variance between the model equations. These high statistics suggest complex factors involved in relationship between TEC and temperature. This high heterogeneity was addressed using meta-regression analysis where the statistics were significantly reduced when a new GRE was obtained (Eq. 6).

$$\text{TEC} = -5.19m - 114.91 + 0.02n \quad (6)$$

To arrive at Eq. 6, where n represents the number of samples used, the reduced statistics are $i^2 = 35.1\%$, $Q = 174.21$ ($p < 0.001$) and $\tau^2 = 16.52$. Eq. 6 is suitable for the estimation of ionospheric TEC from atmospheric temperature data.

Conclusion

This study investigated the correlation between seasonal atmospheric heating and ionospheric Total Electron Content (TEC) dynamics, utilizing temperature, surface pressure, and specific humidity data during Spring Equinox, Summer Solstice, and Fall. The analysis revealed distinct relationships between these atmospheric parameters and TEC. Temperature and surface pressure exhibited strong negative correlations with TEC, indicating that increased atmospheric heating leads to decreased ionospheric electron density. The negative correlation between temperature and TEC implies that atmospheric heating during Spring Equinox and Summer Solstice may lead to decreased ionospheric electron density, potentially affecting radio communication and navigation systems. Surface pressure's negative correlation with TEC indicates that changes in atmospheric pressure patterns may influence ionospheric dynamics.

Also, specific humidity, however, showed a strong positive correlation with TEC, suggesting that increased atmospheric moisture enhances ionospheric electron density. The positive correlation between specific humidity and TEC suggests that increased atmospheric moisture during Fall may contribute to enhanced ionospheric electron density, improving communication and navigation.

As the study demonstrates the significance of atmospheric heating in shaping ionospheric TEC dynamics, highlighting the complex interplay between temperature, surface pressure, and specific humidity. These findings contribute to our understanding of the atmospheric-ionospheric interface, ultimately informing improved forecasting and modeling of ionospheric conditions. We recommend for future research, the investigation of the causal mechanisms underlying these correlations and exploration of the impact of other atmospheric parameters (e.g., wind patterns, ozone concentration) on TEC dynamics.

Reference

- Adewalea, A.O., Oyeyemia, E.O., Adeloyea, A.B., Akalaa, A.O., and Oyebolaa, O.O., (2013). Ionospheric time delay variations at an equatorial station during low solar activity. *J. Sci. Res.* 14, 85–94.
- Bhuyan, P. K., and Borah, R. R., (2007) TEC derived from GPS network in India and comparison with the IRI. *Adv. Space Res* 39: 830–840.
- Borenstein, M., Hedges, L., Higgins, J., and Rothstein, H. (2002). *Comprehensive Meta Analysis*. National Institute of Health.

- Chen, C. Y., Lin, C. H., Lo, T. W. Y., and Lin, J. T. (2020). Relationship Between TEC and Atmospheric Temperature During Equinoxes. *Journal of Atmospheric and Solar-Terrestrial Physics*, 208, 105233
- Ciraolo, L., Spalla, P. and P. Beni, P (1994). An analysis of consistency of TEC evaluated using pseudo range GPS observations, in Proceedings of the International Beacon Satellite Symposium, edited by L. Kersley, pp. 21 – 24, *Univ. of Wales, Aberystwyth*.
- Dabbakuti J.R.K.K., Ratnam D.V., and Sunda, S., (2016) Modelling of ionospheric time delays based on adjusted spherical harmonic analysis, *Aviation*, 20 (1), 1- 7, ISSN 16487788.
- El-Rabbany, A., (2002). Introduction to GPS: The Global Positioning System. Artech House, Boston, London.
- Gopi, K. S. (2017). GPS-TEC analysis application. Indian Institute of Geomagnetism (IIG), Navi Mumbai, India. Institute of Scientific Research, (2), 2 – 9.
- Helsel, D. R. (2005). Nondetects and data analysis: Statistics for censored environmental data. John Wiley & Sons, p. 191
- Hirsch, R.M., Slack, J. R., and Smith, R. A. (1982). Techniques of trend analysis for monthly water quality data, *Water Resources Research* 18(1), 107-121.
- Kendall, M.G. (1975). Rank Correlation Methods, 4th edition, Charles Griffin, London.
- Klobuchar, J.A., (1987). Ionospheric time-delay algorithm for single frequency GPS users. Institute of Electrical and Electronic Engineers Transactions on Aerospace Electron System, 23, 325– 331.
- Liu, L., and Chen, Y. (2009). Statistical analysis of solar activity variations of total electron content derived at Jet Propulsion Laboratory from GPS observations. *Journal of Geophysics Resources*, 114, A10311, doi:10.1029/2009/JA014533
- Liu, L., Wan, W., Zhang, M. L., and Ning, B. (2006). Variations of Ionospheric Total Electron Content and Thermospheric Temperature. *Journal of Geophysical Research*, 111, A07303).
- Mann, H.B. (1945). Non-parametric tests against trend, *Econometrica* 13, 163-171.
- Mannucci A. J., Wilson B. D., and Edwards C. D. (1993) A new method for monitoring the earth's ionosphere total electron content using the GPS global network. In: Proceedings of ION GPS93. *Institute of Navigation*, 1323–1332.
- Mendillo, M., Klobuchar, J. A., and Stauning, P. R. (1974). Ionospheric Total Electron Content and Thermospheric Temperature. *Journal of Geophysical Research*, 79(10), 1513–1520.
- Nibigira, J. D., VenKat, D. R. and Prasad, G. S. V. (2019). Analysis of Ionospheric GPS-TEC Variability over Uganda IGS station during descending phase of 24th Solar cycle, 2018-year, *International Journal of Scientific and Technology Research*, 8(12), 3947-3950.
- The Nation News paper (18th October, 2024): The Incessant collapse of the national grid. <https://thenationonlineng.net/incessant-collapse-of-the-national-grid>. Accessed: 18th October, 2024
- Ratnam, D. V., Dabbakuti, J. R. K. K., and Sunda, S. (2017). Modeling of ionospheric time delays based on a multishell spherical harmonics function approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(12), 5784-5790. DOI:10.1109/JSTARS.2017.2743695.
- Ratnam, D. V., Sivavaraprasad, G., and Jiyun, L. (2015). An Automatic Ionospheric scintillation detector for GNSS receivers, *IET Radar, Sonar, & Navigation (UK)*, 9(6), 702 – 711, doi: 10.1049/iet-rsn.2014.0232.

- Rishbeth, H (1998). How the thermospheric circulation affects the ionospheric F2-Layer. *J. Atmos. Sol. Terr. Phys*, 60(14), 1385-1402
- Sridhar, M., Venkata, R. D., Uday, B. T., Leela, P. A., Rohit, B., and Ramaraju, A. (2016), Analysis of fading effects due to ionospheric scintillations using modern GNSS signals observed at northern low latitude station, *ARPJ Journal of Engineering and Applied Sciences*, 11 (15), pp: 9329 - 9334. ISSN 18196608.
- Zhang, S.R., Emmert, J. T., Drob, D. P., and Forbes, J. M., (2017). TEC and Thermospheric Temperature Variations During Solar Cycle 24. *Journal of Geophysical Research: Space Physics*, 122, 9461– 9474. DOI: 10.1002/2017JA024410
- Zharkova, V., Vasilieva, I., Shepherd, S. and Popova, E. (2023) Periodicities in Solar Activity, Solar Radiation and Their Links with Terrestrial Environment. *Natural Science*, 15 (3), 111-147. doi: [10.4236/ns.2023.153010](https://doi.org/10.4236/ns.2023.153010).

Deep Learning Framework for Predicting Alternative Routes for Information Management in Intra-city Road Traffic Network

Ifreke J. Udoeka^{1*}, Moses A. Agana², Ofem A. Ofem³ and Ifioek J. Udo⁴

¹*Department of Computer Science, Akwa Ibom State University*

^{2,3}*Department of Computer Science, University of Calabar*

⁴*Department of Computer Science, University of Uyo*

**Correspondence: ifrekeudoeka@aksu.edu.ng*

Abstract

Traffic congestion associated with increased vehicular movements and coupled with poor road network often leads to loss of man hours and low productivity, thus presenting a setback to the development of major cities. This challenge can be addressed by providing accurate information about the traffic-flow within the intra-city road network to guide the road users on the choice of alternative routes to use to avert frequent traffic congestions. In this paper, we present a deep learning framework that incorporates the direct and indirect cost burdens of intra-city routes to predict possible alternative routes for informed choices of traffic flow management within the intra-city road network. The proposed deep learning framework is informed by its ability to analyze big data with high-dimensional features, spanning, weather, point of interest, nature of the roads, time of the day with accuracy. The paper utilized primary datasets obtained from South-South region (Asaba, Benin City, Calabar, Port Harcourt, Uyo, Warri and Yenegoa and its suburbs) of Nigeria using Google Maps, Open Weather maps, OpenStreetMap (OSM) and ride hailing services. The paper leverages on recurrent neural networks (RNNs) to extract cost related features to predict alternative routes with variations of cost burden for informed choice of route by the road users. The results in terms of preprocessed features of the data obtained from the area under study revealed that the features range from 0.30 to 0.69 in correlation analysis. This indicates that the features utilized for the deep learning framework are relevant features without dependency. Also the strength of the linear relationship between the FARE and the target output is 0.823014 at a significance level of 0.05. This study has the potential of enhancing urban mobility and increase productivity for businesses, hence it is recommended for intelligent transport applications.

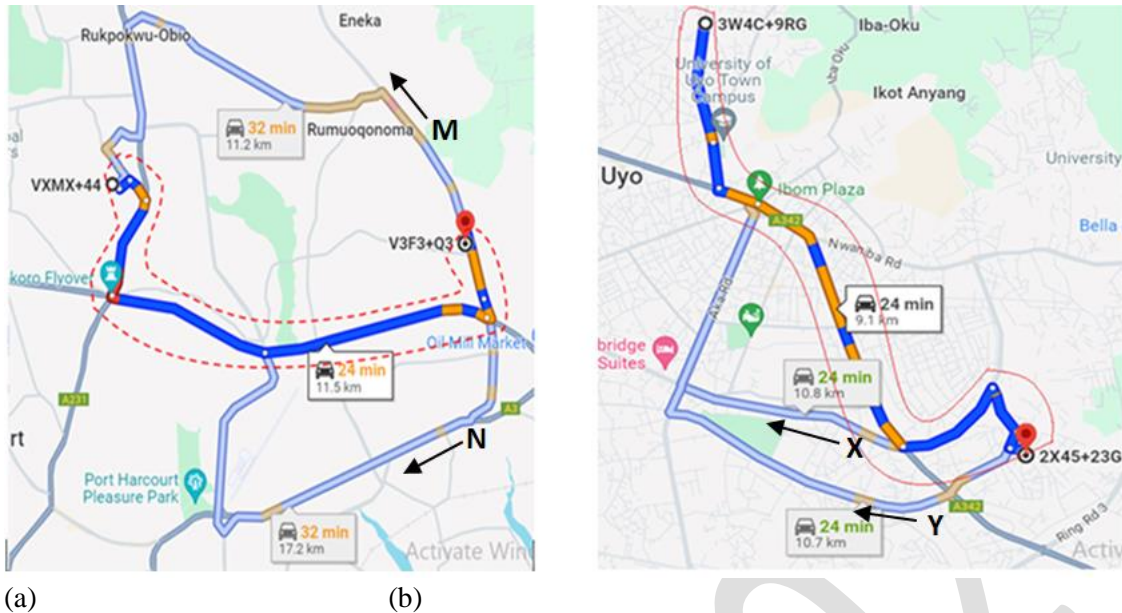
Keywords: alternative route; cost-burden; traffic flow; deep learning

1. Introduction

Road traffic congestion is a universal problem that causes so much distress to urban dwellers, hindering smooth movement of people, goods and services. This phenomenon is attributed to increasing high density of vehicles struggling to access limited road spaces particularly during the peak hours or traffic incidents such as accidents and vehicle breakdowns (Kreindler, 2024). Moreover, in developing countries, rare utilization of intelligent transport applications that can empower road users to make informed decision on choice of available alternative routes to evade congestion contribute to traffic bottleneck or gridlock. The increase in population as well as high vehicle density resulting from rapid development of rural areas into urban centres increases pressure

on urban road networks and infrastructures. The challenges of traffic congestion leads to losses of man-hour and increase fuel consumption as vehicles spend more time idling or moving slowly (ViaMichelin, 2022). In addition, congestion reduces productivity, augments stress and frustration, increase commuting delays, and generates pollution that impact on environment and add additional cost to the society (Thakre & Pawade, 2024). The persistence increase of noise and air pollutions is detrimental to public health and mental wellbeing among residents of urban dwellers (Conceição *et al.*, 2023; Wojuade&Olateju 2020, James *et. al.*, 2016). Given the negative impacts posed by poor information management of traffic congestion and air pollution, tolling has been introduced in several cities around the world to curb vehicles from entering the inner city (Zhang *et al.*, 2019, Vosough *et al.*, 2022). These schemes encourage drivers to make use of the available alternative routes based on cost considerations. However, there is inadequate digital information management framework to guide the choice of alternative route network within the intra-city road network. The cost of congestion is multi-faceted and affects various aspects of the society (Sweet, 2014). In Nigeria, there is a 2023 report by Danne Institute for Research that revealed about 4 trillion naira annual loss incurred by Lagos State due to impact of traffic congestion. This is culminated from the loss of an estimated 14.12 million hours by the residents of the state on daily basis. Apart from Lagos, the impact of traffic congestion is felt in other cities, such as: Ibadan, Federal capital territory Abuja, Kano, Kaduna Asaba, Benin-City, Calabar, Port Harcourt, Uyo, Warri, Yenegoa and more. The aim of this paper is to utilize a deep learning intelligent framework to proffer options of definite alternative routes to road users to evade traffic congestion in the area under study. Although recent advancements in deep learning has offered significant opportunities for solving problems (LeCun *et al.*, 2015), traffic information management within intra-city route network still poses a challenge to road users and its environments. Therefore, there is the need for transformative changes across various domains, including urban planning and transportation information management using modern computing infrastructures (Townsend, 2013). Based on the light of the foregoing, the specific objective of this paper is to collect and analyze traffic-specific data within intra-city routes covering the major cities within South-South region of Nigeria over a period covering January to November, 2021. The analysis of the data entails the removal of feature dependency and transformation to produce relevant data features suitable for deep learning. In addition, a conceptual framework for prediction of alternative routes to guide road users on traffic information within the intra-city road network is alsopresentedin the paper.

The study area of this research are intra-city routes in Asaba (6.2059° N, 6.6959° E), Benin City (6.3350° N, 5.6037° E), Calabar (4.9757° N, 8.3417° E) Port Harcourt (4.8472° N, 6.9746° E), Uyo (5.0377° N, 7.9128° E), Warri (5.5494° N, 5.7669° E) and Yenegoa (4.9514° N, 6.3492° E). These cities are within the south-south region of Nigeria where there is rapid growth of urbanization due to economic activities as a result of oil exploration.



(a) (b)
Figure 1 Intra-city routes with adjoining Alternative paths from (a) Rumuodomaya to Elimbu in Port Harcourt (b) UruaEkpa to Etoi 1 in Uyo.

Given road segments in Figure 1, a specific road sections shown in Figure 1 (a) represents the route from Rumuodomaya to Elimbu in Port Harcourt traversing via Rumuji-Mpakurche Rd with alternative routes *M* via Rumuowha-Eneka road and *N* via Rumuola-Rumukoroshe road. Figure 1 (b) represents the route from UruaEkpa to Etoi 1 in Uyo traversing via IkotEkpene - Calabar Rd/Oron Rd/Uyo-Oron Rd/A342 with alternative route *X* via Nsikak Eduok Avenue and *Y* via Aka Rd and Ring Rd 2/Udo Udoma Avenue. The red dotted line represents the driving path section of the road that is heavily congested during the peak hours. These congested paths pose a challenge for drivers, causing delays, slower movement, or stoppages. For drivers to evade traffic on these congested routes, Figure 1 (a) and (b) suggested alternative routes *M*, *N* and *X*, *Y* respectively. These alternatives routes can be suggested based on having the minimal features associated with cost burden, implying that they either save time, fuel, or fare compared to the congested route. This can provide drivers with an informed decision to rather opt for these routes to bypass the congestion and maintain a more efficient and cost-effective journey.

The dataset covered 210 road segments comprising of 30 roads from each of the city under study. The dataset features considered in this study include traffic flow with attributes of smooth (SMO), slightly congested (SCO), congested (CON) and highly congested (HCO); weather condition with attributes of sunny (SNY), rainy (RNY), cloudy (CDY) and clear (CLR); point of interest (POI) with attributes of densely and sparsely populated area; time of the day (TOD) with attributes of peak and off-peak period; distance (DST) with attributes of short and long distances; Fare with attributes of low and high rate; estimated time of arrival (ETA) with attributes of short and long; nature of the road (NOR) with attributes of paved and unpaved road. The study also considered four time of the day namely; morning (6am to 11:59am), afternoon (12pm to 3:59pm), evening (4pm to 8:00pm) and night (8pm to 10:00pm). Figure 2 shows the traffic status of the study area at different times of the day.

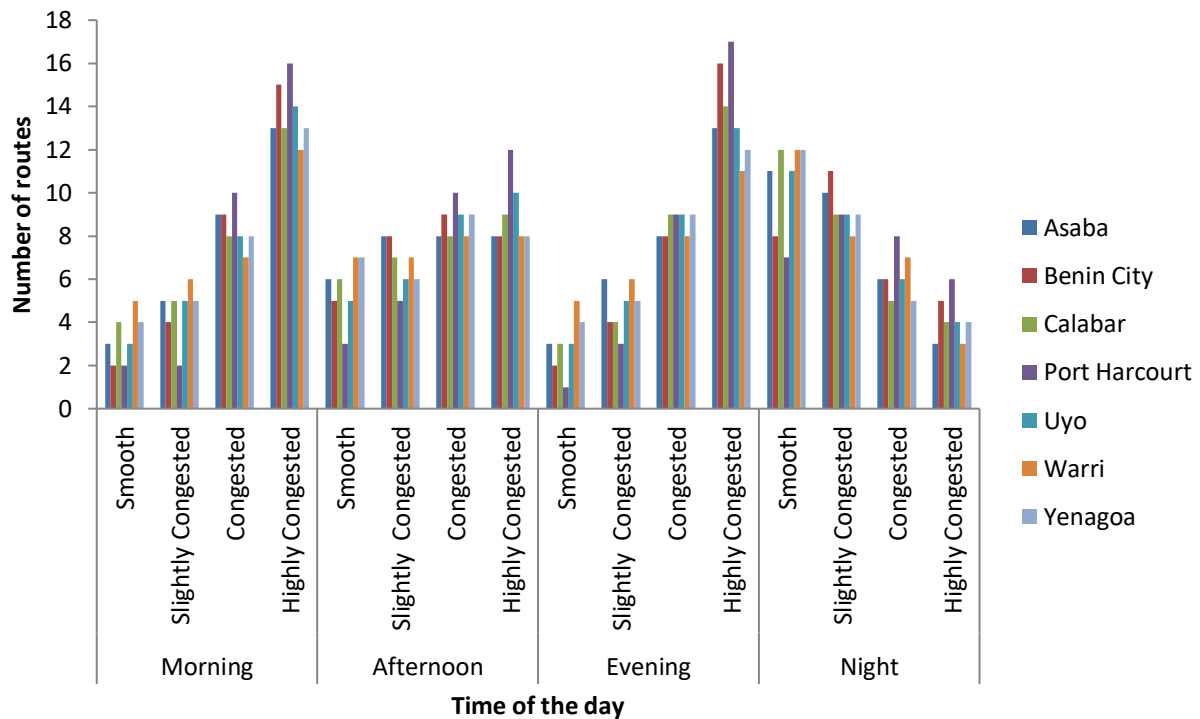


Figure 2: Traffic status of the study area.

Figure 2 shows a large numbers of highly congested traffic statuses in the morning and evening (peak) periods particularly in Port Harcourt, Benin City, Calabar, Uyo and Asaba. There is also increased number of smooth in the night (off-peak period) predominantly in Yenagoa, Warri, Uyo, Calabar and Asaba. The focus of this paper captures the development of a deep learning framework for predicting alternative routes in intra-city route network. The subsequent sections of this paper are arranged as follows. Section 2 discusses the related works on the field of traffic flow prediction. Section 3 introduces the proposed framework, including the traffic data preprocessing model. Section 4 presents the results, and finally, section 5 presents the conclusion of this paper.

2. Related Works

In recent years, the continuous increase in urban population has triggered the surge in traffic congestion mostly in intra-city road network. The need to explore advanced technologies for efficient traffic management and route optimization has become crucial for smooth navigation in urban areas. Several works have focused on using these technologies to detect traffic jam and predict alternative routes to minimize travel time and reduce congestion. Guidoni *et al.*, (2020) presented a Vehicular Traffic Management Service called the Re-RouTE. The mechanism of Re-RouTE employs the concepts of flow density model of traffic engineering theory to classify network congestion. It considers only the street density of vehicles to classify traffic conditions, detect traffic jams and suggest alternative routes to reduce traffic congestion. Musa *et al.*, (2023) utilized machine learning algorithms and cloud computing to develop models and decision making approaches. These approaches suggested alternative routes traversal within the urban routes, predict traffic conditions (including the outcomes, forecasting, and decongestion), reduce lost hours for road users, and simplifying urban transportation activities for urban dwellers. Stan *et al.*, (2023) introduces a traffic threshold mechanism to predict vehicular traffic congestion and then suggest alternative routes in order to avoid congestion in the urban area. The results of the mechanism showed a significant

reduction on time spent on traffic by 69% and carbon emissions by 61% with 50K+ scenario. Xu *et al.*, (2024) proposed a machine learning-based traffic flow prediction and intelligent traffic management using a Time-Graph Convolutional Network (T-GCN) model. The datasets from SZ-taxi in Shenzhen City and the Los-loop detectors in Los Angeles were used. The model demonstrated the ability to capture the dynamic trend of road traffic information and obtain a better prediction results. Wu *et al.*, (2024) proposed a Short-Term traffic flow prediction using Grey Relational Analysis for Data Filtering and Stacked LSTM Modeling. This study uses PeMS dataset with Occupancy data, Traffic volume and speed as predictors. The results of the experiment showed that the model achieved improved prediction accuracy by 3.6% and a reduced prediction time of 27.33%. Zhang *et al.* (2017) proposed a deep learning framework called Deep Spatio-Temporal Residual Networks (DSTRN) for predicting citywide crowd flows. The model leverages both spatial and temporal information to capture the complex dynamics of crowd movements in urban areas. The model utilized residual connections to facilitate the flow of information across different layers. These connections help in mitigating the vanishing gradient problem and enable better information propagation. The model is evaluated on real-world citywide crowd flow datasets, and the results demonstrated that the DSTRN outperforms several baseline methods in terms of crowd flow prediction accuracy, including traditional methods and other deep learning approaches. Yu *et al.* (2017) introduced a deep learning framework called Spatio-Temporal Graph Convolutional Networks (STGCN) for traffic forecasting. The model leverages both the spatial and temporal dependencies in traffic data by incorporating graph convolutions into recurrent neural network (RNN) architecture. The model also employed temporal convolutional layers to capture the temporal dependencies in traffic data. The convolutional layers capture patterns over different time scales, enabling the model to capture both short-term and long-term temporal dependencies. The experiment was conducted using two real world traffic datasets, BJER4 and PeMSD7 collected in Beijing and California respectively. The results demonstrated that the STGCN outperforms several baseline methods in terms of traffic forecasting accuracy, including traditional methods and other deep learning approaches. Khan *et al.*, (2020) proposes a smart route Internet-of-Vehicles (IoV)-based congestion detection and avoidance (IoV-based CDA) scheme to monitor traffic in real time and detect congestion patterns. The scheme integrates IoV technology with advanced rerouting algorithms using modified Evolving Graph (EG)-Dijkstra to reroute vehicles to alternative route when the required threshold is exceeded. The results of the experiment revealed that proposed system can significantly reduce travel time and improve traffic efficiency compared to traditional traffic management approaches. Xu *et al.*, (2018) proposed a hybrid framework to integrate both urban traffic flow characteristic theory and machine learning techniques. The features considered in the study were traffic volume, road speed limit, route distance, traffic light and weather conditions. The experiment was conducted using Deep Belief Network (DBN) on two real world trajectories datasets named BJTaxi and Ucar, and the results revealed that DBN method provided a better classification accuracy of 8% than the Bayesian model. Despite the enormous studies on road traffic prediction and route optimization, still, there is dearth of literature on cost-burden associated with alternative route in urban areas, hence this research.

3. Proposed Framework

In this framework, RNN is employed to model the dynamic nature of intra-city traffic flows by analyzing real-time traffic data. The network processes time-dependent traffic parameters, such as traffic density, time of the day, travel time, estimated time of arrival and non-time dependent traffic parameters such as point of interest, fare, nature of the road across different routes. These inputs allow the RNN to predict not only the likelihood of congestion but also the potential cost burdens associated with taking various alternative routes. The model then provides accurate predictions for

travel time, and cost burden for each routing option by learning from real-time traffic sequences. Figure 3 show the processing layers of the proposed framework.

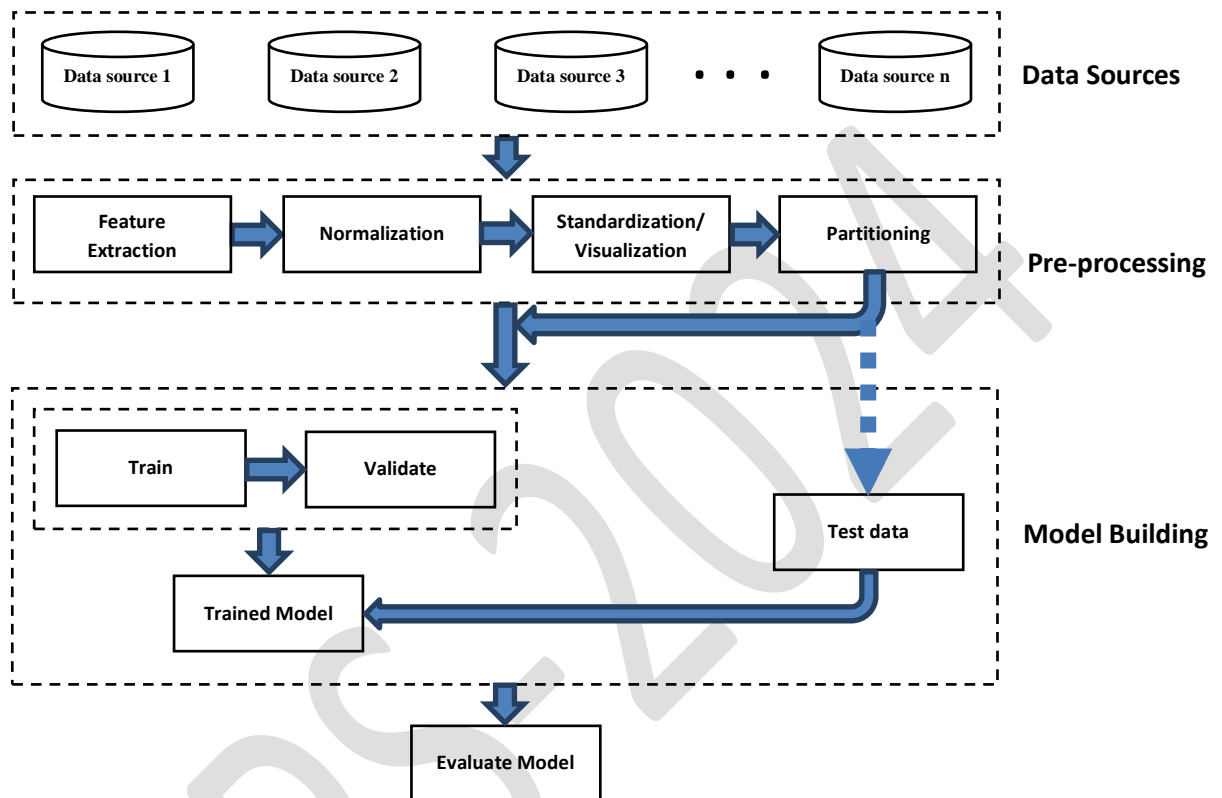


Figure 3: Processing Layers of the proposed framework

The proposed framework shown in Figure 3 is segmented into four phases. The first phase comprises of the data obtained from intra-city road in South-South, Nigeria comprising sources such as map features, weather features and so on. The second layer is the data pre-processing phase which includes feature extraction, normalization where numerical features (POI, FARE, DST and ETA) are normalized using min-max scaling technique, and categorical features such as traffic, weather, nature of the road (NOR) and time of the day (TOD) are normalized using ordinal encoder technique. Standardization/ Visualization stage is to prevent bias in algorithms and improve model performance and explore relationships between features, guiding preprocessing and modeling decisions. Then the data is partitioned into training and validation/ testing sets to ensure that the model is properly evaluated on unseen data and not overfitted to the training set. The third phase is the model building layer comprising of training, validation and testing. The final phase is the model evaluation.

4. Results

In order to adapt the dataset to deep learning model, removal of feature dependency is implemented using the Pearson correlation analysis shown in Figure 4. The high positive and negative correlation suggested that features provide similar information and are redundant to some extent. In predictive modeling, the highly correlated features add no much additional information to the model.

	SMO	SCO	CON	HCO	CDY	RNY	CLR	SNY	TOD	NOR	POI	FARE	DST	ETA
SMO	1													
SCO	-0.30496	1												
CON	-0.21728	-0.62437	1											
HCO	-0.12544	-0.36047	-0.25682	1										
CDY	0.051775	-0.03142	-0.02406	0.034396	1									
RNY	-0.02936	-0.03695	0.038973	0.027101	-0.6007	1								
CLR	0.117532	-0.02469	-0.02233	-0.03576	-0.22161	-0.03238	1							
SNY	0.018381	-0.00756	0.00214	-0.00785	-0.14353	-0.02097	-0.00774	1						
TOD	-0.1008	-0.01344	0.059844	0.026149	0.007801	0.024409	-0.06792	-0.03662	1					
NOR	-0.08134	0.010291	0.018789	0.030258	-0.01866	0.02358	-0.05541	-0.02725	-0.07966	1				
POI	-0.05625	0.045382	0.004143	-0.02382	-0.00671	0.000773	-0.01386	0.003619	0.307848	-0.04399	1			
FARE	-0.07382	-0.01795	0.044915	0.02971	0.007109	0.009437	-0.02422	-0.02792	0.329225	0.150828	0.099468	1		
DST	-0.02998	-0.00324	0.006628	0.022023	-0.00386	-0.00832	-0.0009	0.005574	0	0.161424	0.064791	0.609419	1	
ETA	-0.25062	-0.1954	0.152766	0.300523	0.001335	0.013133	-0.02832	0.017956	0.051269	0.089887	0.016893	0.24858	0.338179	1

Figure 4: Pearson Correlation Analysis

The preliminary results of correlation analysis conducted on the pre-processed features revealed correlation coefficients ranging from 0.30 to 0.69. This indicates that there are moderate to strong positive relationships between the variables. Specifically, correlations between 0.30 and 0.49 suggest moderate associations, where the variables are somewhat linked, while those between 0.50 and 0.69 represent stronger linear relationships. These results imply that, for some pairs of features, changes in one variable are accompanied by relatively predictable changes in another, though the strength of these relationships varies. These findings are consistent with expectations, given the nature of the data and are relevant features without dependency. The strength of the linear relationship between the FARE and the target output is 0.823014 as indicated in Figure 5.

	FARE	TARGET
FARE	1	
TARGET	0.823015	1

Figure 5: The Linear relationship between FARE and target output

5. Conclusion

This study utilized a primary dataset of intra-city route and traffic prediction obtained from cities within South-South Nigeria (Asaba, Benin City, Calabar, Port Harcourt, Uyo, Warri and Yenagoa), this adds to the body of knowledge for future research. Information regarding alternative routes and points-of-interest in the dataset intends to serve as a bye-pass to mitigating traffic congestion in the study area. This study has also provided additional information and cost burden associated with route and traffic prediction model available to road users. The application of deep learning models for intra-city route and traffic prediction holds significant promise in addressing the complex and dynamic nature of urban transportation systems. The utilization of advanced technique, such as recurrent neural networks (RNNs), enables these models to capture intricate temporal and spatial dependencies within traffic data. Stacked GRU, with multiple layers of recurrent units, enhance the model's capacity to understand both short-term and long-term trends in traffic patterns. As Intelligent Transportation System (ITS) continues to evolve, the collaboration between researchers, city planners, and industry stakeholders becomes crucial for the successful integration of these models into practical urban transportation systems.

References

- [1] Conceição, M. A., Monteiro, M. M., Kasraian, D., van den Berg, P., Haustein, S., Alves, I., ...& Miranda, B. (2023). The effect of transport infrastructure, congestion and reliability on mental wellbeing: a systematic review of empirical studies. *Transport reviews*, 43(2), 264-302.
- [2] Danne Institute for Research. (2023, November). 'Behavioural Causes of Traffic Congestion in Lagos'. *BusinessDay*. Retrieve from <https://businessday.ng/news/article/lagos-loses-n4trn-yearly-to-traffic-congestion-report/> accessed on 24th August, 2024.
- [3] Guidoni, D. L., Maia, G., Souza, F. S., Villas, L. A., & Loureiro, A. A. (2020). Vehicular traffic management based on traffic engineering for vehicular ad hoc networks. *IEEE Access*, 8, 45167-45183.
- [4] James, I., Ansa, G., & Udoeka, I. (2016). Conceptual Framework on Overcoming the Challenges of Multiple Vehicle Registration in Nigeria: A Mobile Application Approach. *IOSR Journal of Mobile Computing and Application*, 3(5), 1-3.
- [5] Khan, Z., Koubaa, A., & Farman, H. (2020). Smart route: Internet-of-vehicles (iov)-based congestion detection and avoidance (iov-based cda) using rerouting planning. *Applied Sciences*, 10(13), 4541.
- [6] Kreindler, G. (2024). Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications. *Econometrica*, 92(4), 1233-1268.
- [7] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [8] Musa, A. A., Malami, S. I., Alanazi, F., Ounaies, W., Alshammari, M., & Haruna, S. I. (2023). Sustainable Traffic Management for Smart Cities Using Internet-of-Things-Oriented Intelligent Transportation Systems (ITS): Challenges and Recommendations. *Sustainability*, 15(13), 9859.
- [9] Stan, I., Ghore, D. A., Dan, P. I., & Potolea, R. (2023). Urban congestion avoidance methodology based on vehicular traffic thresholding. *Applied Sciences*, 13(4), 2143.
- [10] Sweet, M. (2014). Traffic congestion's economic impacts: Evidence from US metropolitan regions. *Urban Studies*, 51(10), 2088-2110.
- [11] Thakre, Y. K., & Pawade, P. Y. (2024, June). Traffic Congestion at Urban Road-Review. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1326, No. 1, p. 012094). IOP Publishing.
- [12] Townsend, A. M. (2013). *Smart cities: Big data, civic hackers, and the quest for a new utopia*. WW Norton & Company.
- [13] ViaMichelin Magazine (2022), Traffic jams: our tips for saving fuel. Retrieved from: <https://www.viamichelin.com/magazine/article/traffic-jams-our-tips-for-saving-fuel/> Accessed on 12th July, 2023.
- [14] Vosough, S., de Palma, A., & Lindsey, R. (2022). Pricing vehicle emissions and congestion externalities using a dynamic traffic network simulator. *Transportation Research Part A: Policy and Practice*, 161, 1-24.
- [15] Wojuade, C.A., & Olateju, A.P. (2020). Assessment of Traffic Noise Pollution in Residential Neighbourhood of Lagos, Nigeria. *International Journal of Research*, 7, 396-403.
- [16] Wu, Z., Huang, M., Xing, Z., & Yang, T. (2024). Improving Short-Term Traffic Flow Prediction using Grey Relational Analysis for Data Filtering and Stacked LSTM Modeling. *International Journal of Computers Communications & Control*, 19(1).
- [17] Xu, J., Zhang, Y., & Xing, C. (2018, October). An effective selection method for vehicle alternative route under traffic congestion. In *2018 IEEE 18th International Conference on Communication Technology (ICCT)* (pp. 494-499). IEEE.

- [18] Xu, Z., Yuan, J., Yu, L., Wang, G., & Zhu, M. (2024). Machine Learning-Based Traffic Flow Prediction and Intelligent Traffic Management. *International Journal of Computer Science and Information Technology*, 2(1), 18-27.
- [19] Yu, B., Yin, H., & Zhu, Z. (2017). Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. *In Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, 3634-3640
- [20] Zhang, J., Zheng, Y., & Qi, D. (2017, February). Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. *In Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, No. 1).
- [21] Zhang, S., Campbell, A. M., & Ehmke, J. F. (2019). Impact of congestion pricing schemes on costs and emissions of commercial fleets in urban areas. *Networks*, 73(4), 466-489.

AI-Powered Geophysical Assessment of Dumpsites for Economic Recovery: A Case Study of Umusedeli and Azunze, Kwale, South-South Nigeria

Okorafor, C. Nneka

Novena University Ogume,
Delta State, Nigeria

Salami, A. Sikiru

University of Benin, Edo
State, Nigeria

Ugbe, C. Felix

Delta State University,
Abaraka, Delta State,
Nigeria.

Okorafor G. Nwaji

Federal University of
Technology, Ikot-Abasi,
Akwa Ibom State, Nigeria

Abstract: This study investigated the potential for national economic recovery through conducted geophysical assessments of dumpsites in Umusedeli and Azunze, Kwale, South-South Nigeria. Geophysical methods such as Vertical Electrical Sounding (VES), Electrical Resistivity Tomography (ERT), and Induced Polarization Tomography (IPT) were used to identify potential resources within dumpsites, such as metals, organic matter, and building materials. The identified dumpsites materials were separated using Artificial Intelligence (AI) techniques, and were evaluated based on their economic value and potential for reuse or recycling. The study also looked at the environmental impact of dumpsite remediation and resource extraction to ensure sustainable practices. Based on the findings, policymakers were given recommendations on how to promote regional economic development and environmental sustainability.

keywords: Geophysical Assessment, VES, ERT, IPT, AI Techniques, leachate, dumpsite.

1. Introduction

Waste disposal dumps are common phenomena especially in industrial and highly populated cities where dumps are generated in tons on a daily basis and thus becomes a more important and efficient way of maintaining a clean environment in urban areas. In developing countries unregulated disposal dumps are commonly located adjacent to large cities, releasing harmful contaminants into a leachate and thereby polluting underground water. Similar groundwater contamination occurs in well waters due to leachate from household sewage and from settlement around refuse dumps.

Over the years, dumpsites have been identified as one of the key threats to groundwater resources. Considering the population of developing country like Nigeria, there is massive increase in the rate of municipal waste generated on a daily basis.. Perpetually, wastes are generated continually and disposed indiscriminately in rivers and dumps without recourse to the underground environment, local geology and their proximity to the living quarters. Leachate generated by decomposed wastes happens to be part of groundwater flow system as soon as it reaches the water table. Poor management of municipal solid waste materials leads to potentially devastating environmental and health hazards. Among health hazards that have resulted from lack of an effective disposal system are periodic epidemic and communicable diseases. The challenges in solid waste dumping, handling, and management, all pose great threats to the entire environment. Humans are therefore exposed to a range of environmental hazards but particularly percolation of polluted leachate into the shallow aquifers which is the main source of drinking water in developing countries. In most cases in developing countries, refuse disposal sites are not properly planned.

DC resistivity and EM-methods have proven particularly successful for the evaluation of mixed waste landfills. Binley and Kemna, 2005, Loke, H.M. (1999), Fenning and Williams, 1997). DC resistivity and electromagnetic (EM) surveys are used for similar purposes, to assess the flow of electrical current in the subsurface; however, the measurements are made in different ways. Subsurface resistivity is a function of the soil/rock type, porosity and conductivity of the fluids that fill pore spaces. Conductivity is the reciprocal of resistivity and is therefore also dependent of the same subsurface parameters. Both conductivity and resistivity are governed by Ohm's Law which deals with the correlation between voltage and current in a conductor. The law states that across a conductor, the potential difference (voltage) is proportional to the current through it (Burger, 2006). Ohm's Law is generally written as $V=I/R$, where V is the potential difference (volts), I is the electrical current (amps) and R is resistance (Ohms).

Both resistivity and conductivity methods can be used to map natural (and anthropogenic) variations in subsurface conductivity (Nielsen, 2005; Telford, et al(1982). The success of these electrical methods at landfill sites is based on the highly conductive nature of landfill leachate when compared to the natural background values. Thus, variations in

conductivity within and surrounding a landfill can provide insight into leachate flow pathways as well as the spread of contamination plumes outside the landfill (Soupios et al., 2005; Dawson et al., 2002; Stanton and Schrader 2001; Karlik and Kaya 2001; Bernstone, et al (2000), Aristodemou, 2000; Atekwana, et al (2000),

The aim of this research work is to showcase the effectiveness of integrating geophysical methods in environmental assessment of waste disposal in order to map the contaminant plume, groundwater flow direction and buried metals at Umusedeli/Azunze Kwale. Plate 1 is the picture of the sorted waste materials from the dumpsite. The geophysical methods we introduced are electrical resistivity tomography ERT, vertical electrical sounding VES and induced polarization IP methods.



Plate1; sorted waste materials from the dumpsite.

Electrical Resistivity Tomography (ERT) is a technique for imaging the subsurface electrical structure using electrical currents. From a series of electrodes, low frequency electrical current is injected into the subsurface, and the resulting potential distribution is measured. but prototype data-collection hardware and research-grade inverse codes suitable for field scale applications soon followed. The method has been developed to detect leaks from large storage tanks, monitor underground airsparging and mapping movement of contaminant plumes.



Plate 2; picture showing the open dumpsite.

2. The Study Area

The study area (fig.1) is located within geographical co-ordinates of latitudes $5^{\circ}40'32''$ N and $5^{\circ}43'5''$ N and longitude $6^{\circ}25'20''$ E and $6^{\circ}24'39''$ E in Kwale Ndokwa-west Local Government area of Delta State, South-south Nigeria.

2.1. Geology of the study area

The study area is part of the Niger Delta sedimentary basin. The lithostratigraphic units correspond respectively with the uppermost Benin Formation aged Oligocene to recent, the Agbada Formation which lies beneath the Benin Formation aged Eocene to recent and the lowest lying Akata Formation aged Paleocene to recent [Short and Stauble. (1967). The Akata Formation is composed mainly of marine shales with sandy and silty beds which are believed to have been laid down as turbidities and continental slope fills with an estimated thickness of 7,000 m. The Agbada Formation which is the petroleum bearing unit in the Niger Delta, consists mostly of shoreface and channel sands with minor shales in the upper parts and an intercalation of sands and shales in equal proportion in the lower parts, with a thickness of over 3,700 m [Doust. and Omatsola, (1990). The Benin Formation consisting of continental sands and gravels is about 280 m thick and may be up to 2,100 m in the region of maximum subsidence Whiteman, (1982). The study area is underlain by Benin formation and its of fresh water swamp and lenses clay. The lithological units of the study area are generally composed of sand and clayey sand. The area has a flat topography and is situated by the of river Niger.it is thickly populated and of high economic value.

3. Materials and Methods

3.1. Geophysical Method

Integrated geophysical methods were used for the investigation of possible contamination of groundwater by leachate from an open dumpsite. The electrical resistivity method which includes, vertical electrical sounding (VES), 2D electrical resistivity tomography (ERT) and induced polarization tomography (IPT) methods. VES was used to determine the depth of aquifer occurrence and subsurface lithology. Eight VES were undertaken with the ABEM SAS 1000 Terrameter using the Schlumberger configuration with current electrode spacing (AB/2) that range from 1m to a maximum of 300m. A total of seven Electrical Resistivity Tomography (ERT) lines and induced polarization tomography (IPT) lines which was measured simultaneously with 2D ERT were surveyed covering the dumpsite (**Figure 1**). ERT was measured using ABEM terrameter SAS 1000, manually imaging system with 4 electrodes. Wenner-gamma electrode configuration was chosen for its relative sensitivity to vertical changes in the sub- surface resistivity below the center of the array and for its ability to resolve vertical changes. Electrode spacing of 5m @ 0.5m, 10m @ 0.5m and 15m @ 0.5m for all the profiles was maintained to attain a reading within the depth range of polluted aquifer in the area. The traverses were oriented in NW-SE direction paralleled to each other and according to the shape/size of the dumpsite, (**Figure 1**) and to ensure maximum site coverage and maximum recovery of data beneath the dumpsite. Profile three was established 80 m away from the dumpsite as a control while the other profiles were rightly located on the dumpsite.

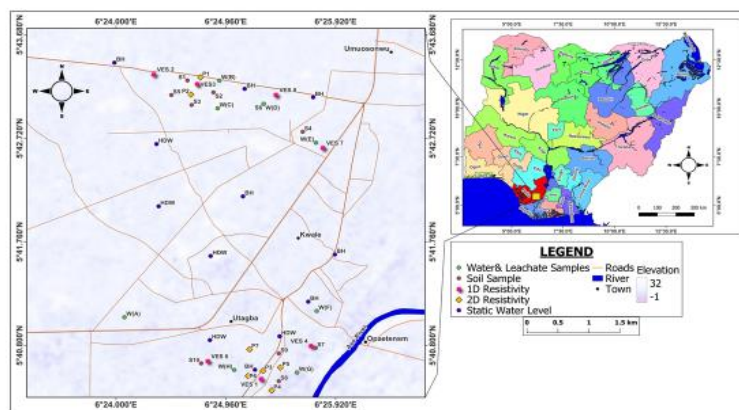


Fig. 1.1 Sample Location Map Of The Study Area.

Resistivity data were inverted using the Res2Dinverse software. The Res2dinverse computer program automatically reduce the measured resistance to apparent resistivity values, based on smoothness-constrained least-square technique. The subsurface is divided into small rectangular blocks with position and size fixed by forward modelling. The

resistivity of the block is then determined so that the calculated apparent resistivity values agree with the measured values from the field survey by adjusting the resistivity of the model blocks and consequently iterate to reduce the difference between the calculated and measured (field) apparent resistivity. These differences are expressed in form of RMS error this compares well with (Ayolabi,et al, 2013).

VES resistivity field data acquired were used to plot the sounding curves in a log-log graph. The curves from the log-log graph were interpreted by partial curve matching techniques using master curves (Koefoed, 1979) and auxiliary point diagrams (Orellana and Mooney, 1966). The resistivity and thickness of the layer obtained from the partial curve matching were then subjected to computer iteration using the Ipi2Win software (figure 2a-h), Ugbe, et al, 2021.

In addition to resistivity data, measured Induced Polarization (IP) data was also inverted with the same resistivity software in similar process of subdividing the earth to smaller models block by forward modelling, iterating to reduce the difference between the calculated and measured apparent data; and expressing the differences in RMS error. Differences in the RMS values obtained in IP compared to ERT indicates the independence of IP data inversion from its counterpart ERT though data were acquired and recorded at the same time as well as inverted by the same computer program.

4. Results, Interpretation and Discussion

4.1. ERT and IP Results

The resistivity distribution derived from 2-D inversions of ERT and IPT data are presented and discussed here with their resistivity-depth models (Figures 4a-g and 4ai-vii).

Generally, the subsurface below the dumpsite was characterized by low resistivity possibly influenced by contaminants emanating from the dumps. The maximum depth penetrated is 10.8 m with resistivity values of contaminated layers below 100 Ωm except some few cases of high resistivity in the top layer and isolated high resistivity at depth under some profiles. The chargeability value is generally low within the impacted region (0.934 – 9.35 ms, Figure 4(ai-vii) which suggest possible high

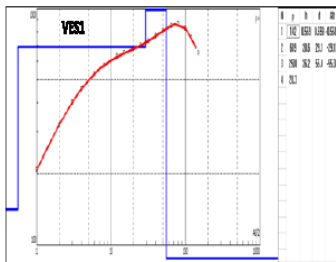


Fig.2a; VES point 1

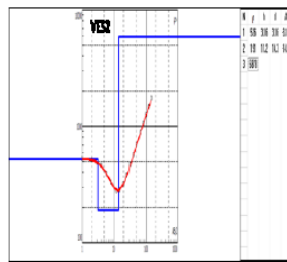


Fig.2b; VES point 2

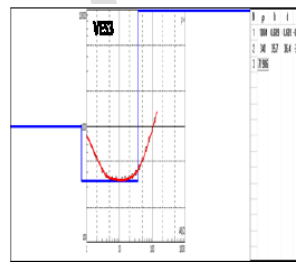


Fig.2c; VES point 3

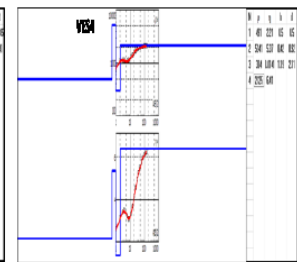


Fig. 2d; VES point 4

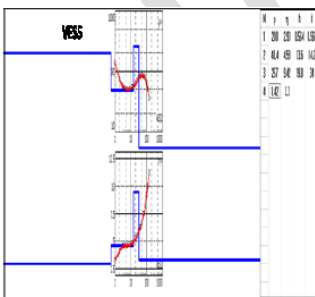


Fig.2e; VES point 5

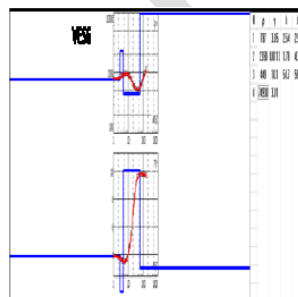


Fig.2f; VES point 6

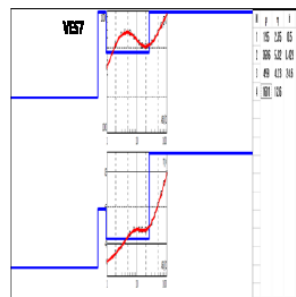


Fig.2g' VES point 7

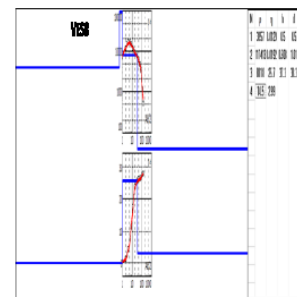


Fig.2h' VES point 8

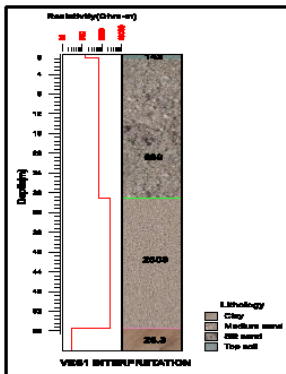


Fig.2.1a; strata plot of VES1

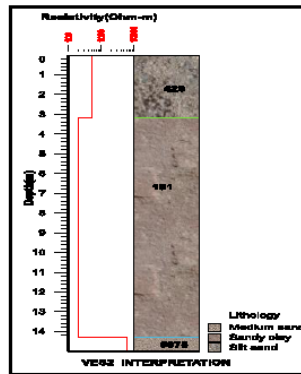


Fig. 2.1b; strata plot of VES 2

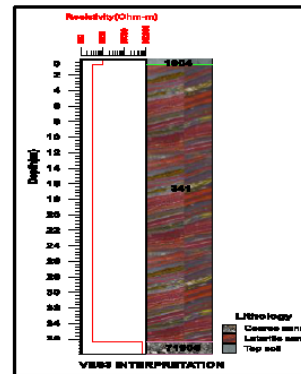


Fig.2.1c; strata plot of VES 3

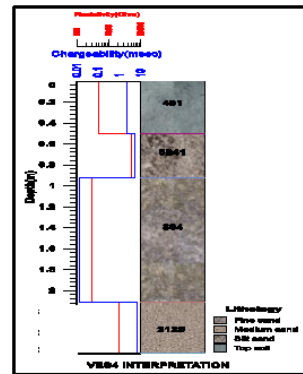


Fig.2.1d; strata plot of VES 4

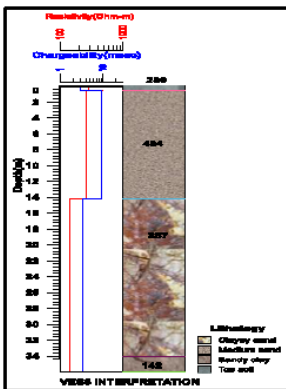


Fig.2.1e; strata plot of VES 5.

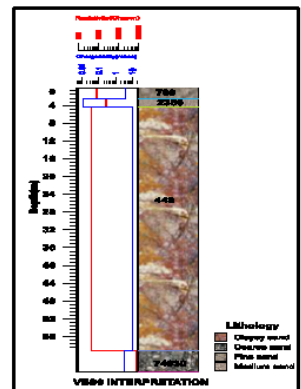


Fig.2.1f; strata plot of VES 6.

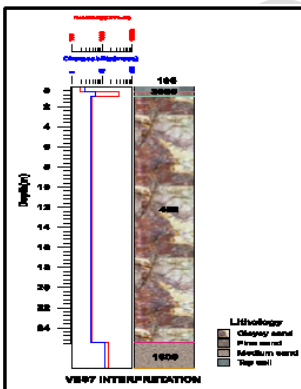


Fig.2.1g; strata plot of VES 7.

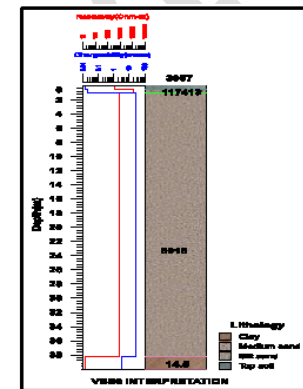


Fig.2.1h; strata plot of VES 8.

porosity and may compose of sand that has been highly impacted with leachate.

A careful look at the resistivity-depth model from ERT profile 1,3, and 4 shows high infiltration of leachate to the subsurface soil. From profile 3 and 4, subsurface under these profiles was characterized by resistivity between 3 and 30 Ωm . profile 6 show leachate impact between electrode positions 16.25 and 25 at the depth of 5m. profile 7 show impact at the beginning and ending part of the profile. Generally, the result reflects impact of leachate from the decomposed materials from most of the profiles. The depth of pollution with low resistivity values is indicative that the leachate from the decomposed refuse material has impacted the subsurface as a result polluting the groundwater. High resistivity encountered between electrode positions 10 and 30 at the depth of 0 – 10.8 m under profile 7 possibly reflects the presence of a resistive waste material and the IPT under the same profile show low chargeability values which could be as a result of highly polarizable materials in the waste. It should also be noted that waste dumped at the site were not sorted as earlier mentioned.

Similarly, models from profiles 6 - 7 reflect moderate resistivity from 44 - 100 Ωm for the polluted area but more pronounced in the middle half of profile 6 and ending part of profile 7 than in the beginning parts.

Worthy of note is the control profile (profile 5) which indicates contamination almost across the profile and the profile is 100m away from the dumpsite. The possibility could be that the leachate might have migrated to that point.

4.1.1. VES Interpretation

The results of Vertical Electrical Soundings and induced polarization obtained were interpreted. Four geoelectric sections were obtained for VES 1,4, 5,6 & 7 while VES 2 & 3 review three geoelectric section the inversion of the interpreted field resistivity data. The inversion of the interpreted field resistivity data was carried out with the aim of delineating the subsurface geologic sequence present in the study area, and determines their geoelectric parameters, layer thicknesses and resistivities (Table 1, fig.2a – h). The lithologies inferred for geoelectric layers and depths were taken to starter software for graphic display of the litho-stratigraphy, depths and thickness which are fundamental subsurface information needed for the environmental assessment of the waste of dumpsite (fig. 2.1a-2.1h).

A total of eight VES points were occupied during the field exercise using Schlumberger array which were acquired alongside with induced polarization data. (table 1). The layers include; Topsoil, clay, medium sand, coarse sand, fine sand, sandy clay and sand. Careful examination of the curves revealed that the VES point one has AK-curve type, VES point two & three have H-curve type, VES point four (4) has KA-type curve, VES point five (5) HK-type curve, VES points six (6), seven (7) KH-type curve and VES

point eight has KQ-curve type. The resistivity values of the topsoil for the VES points 1, 3,4,5,7 & 8, varied from 142 Ω m to 3957 Ω m and the thickness varied between 0.43 – 54.7m, for VES point six (6) their first layer resistivity value is 789 Ω m which is fine sand with the thickness of 2.54m, and VES point 2 the first layer resistivity is 526 Ω m, (table 1). In the second layer, VES point3 had lateritic sand with resistivity value of 341 Ω m and thickness of 35.7m, VES point 1,4&8 had silt sand with resistivity values between 689 to 117413 Ω m and their thickness is between 0.42m to 28.6m, VES points 5&6 had medium sand with resistivity values varies varied from 484 Ω m to 2350 Ω m, and their thickness varied from 1.7 to 13.6m,VES point 7 had fine sand with resistivity value of 3696 Ω m and the thickness is 0.43m.In the third layer, VES point 3 had coarse sand with the resistivity value of 71906 Ω m there was no layered thickness because current terminated at this point, VES point 4 had fine sand with resistivity value of 304 Ω m and the thickness is 1.19m, VES points 5,6,& 7 had clayey sand with the resistivity values of 267 Ω m-459 Ω m and their depth varied from 0.93m -14.2m. In the fourth layer, VES points 4 & 7 had medium sand with resistivity values varied between 1601 Ω m-2125 Ω m and the depth varied from 2.11m-25.5m, VES point 5 had sandy-clay with resistivity value of 142 Ω m and the depth is 34.0m, VES point 6 had coarse sand with resistivity value of 74930 Ω m and the depth is 59.1m. However, VES points 4,5,6 & 7 at the fourth layer does not have layered thickness because current was terminated at this point. At this layer that the bottom depth is not showed needed greater spread to go beyond the fourth layer but because it is environmental studies. For VES points 4 – 7 where there is chargeability values that was acquired alongside with resistivity, the IP values were used to differentiate clay and clean sand bearing water. Clean sand bearing water do have chargeability higher than clay. So both IP and resistivity can differentiate clayey sand from sandy clay and clay as it is displayed in table1.

The graphic display of the lithostratigraphy shows the depth to the bottom, the layered thickness and the lithologies inferred which are fundamental subsurface information needed. The essence of having this graphic plot is to have a quick look at the necessary information needed. For instance, seeing how deep or shallow a layer is and thick or thin a layer is can easily be seen since it displays both top and bottom depth (fig.2.2c -2.2g). The depth to the aquifer in the study area varied from 14.2m (VES 5) to 25.5m (VES

TABLE 1; RESULT OF VES INTERPRETATION.

VES	Layers	Curve character & type	Resistivity (Ω m)	Chargeability (msec)	Thickness (m)	Depth To Top (m)	Lithology inferred
1	1	AK	142		0.56	0.0	Top soil
	2		689		28.6	0.56	Silt sand
	3		2508		26.2	29.1	Medium sand
	4		28.3		-	55.4	Clay
2	1	H	526		3.16	0.0	Silt Sand
	2		191		11.2	3.16	Sandy clay
	3		5878		-	14.3	Medium sand
3	1	H	1004		0.69	0.0	Top soil
	2		341		35.7	0.69	Lateritic sand
	3		71906		-	36.4	Coarse sand
4	1	KA	491	2.21	0.50	0.0	Top soil
	2		5241	5.37	0.42	0.50	Silt sand
	3		304	0.0141	1.19	0.92	Fine sand
	4		2125	6.41	-	2.11	Medium sand
5	1	HK	280	2.93	0.56	0.0	Top soil
	2		484	4.59	13.6	0.56	Medium sand
	3		257	9.42	19.8	14.2	Clayey sand
	4		142	3.3	-	34.0	Sandy clay
6	1	KH	789	3.85	2.54	0.0	Fine sand
	2		2350	0.0113	1.78	2.54	Medium sand
	3		449	10.1	54.7	4.32	Clayey sand
	4		74930	3.01	-	59.1	Coarse sand
7	1	KH	195	2.65	0.5	0.0	Top soil
	2		3696	5.92	0.43	0.5	Fine sand
	3		459	4.23	24.6	0.93	Clayey sand
	4		1601	12.6	-	25.5	Medium Sand
8	1	KQ	3957	0.0129	0.5	0.0	Top soil
	2		117413	0.0192	0.51	0.5	Silt sand

3		8018	25.7	37.5	1.01	Medium sand
4		14.5	2.99	-	38.3	Clay

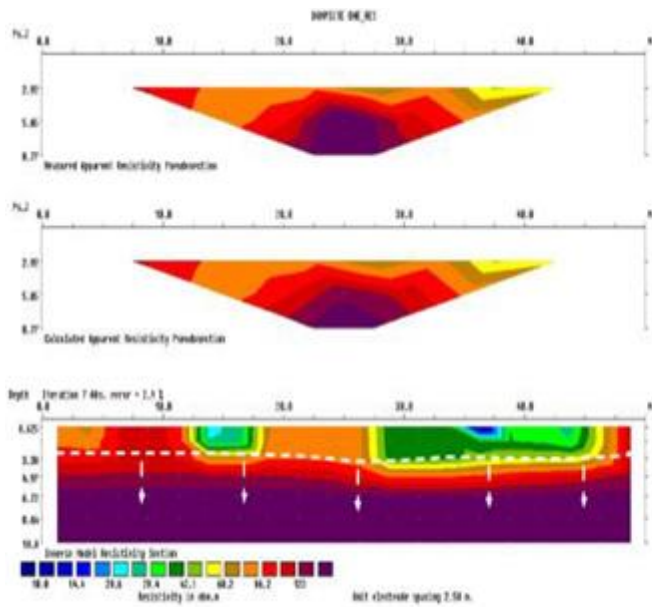


Fig.4a: profile one of ERT

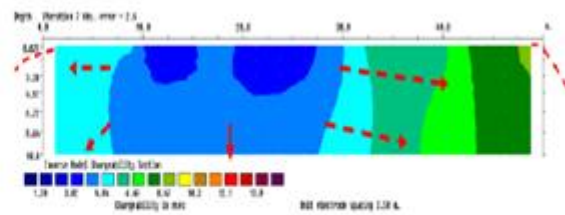


fig.4ai: profile one of IPT

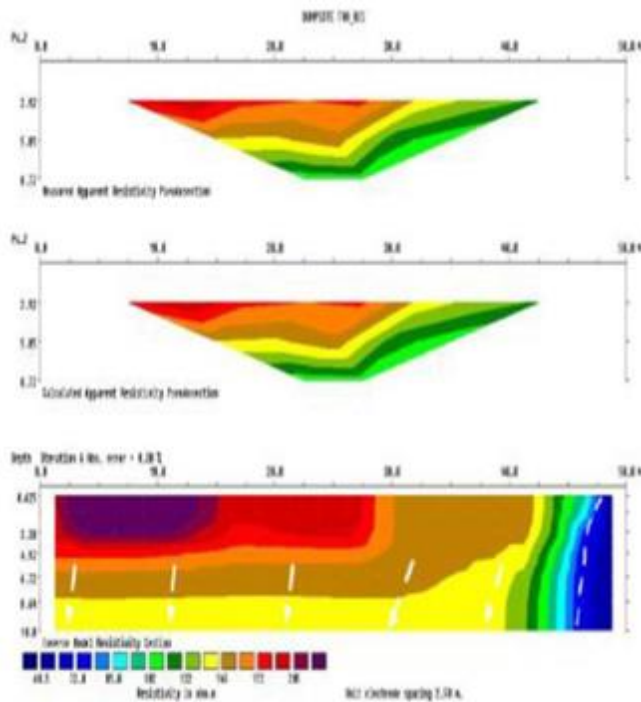


fig.4b: profile two ERT

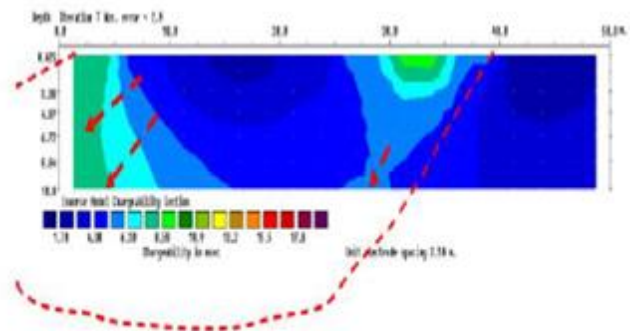


fig.4bii: profile two IPT

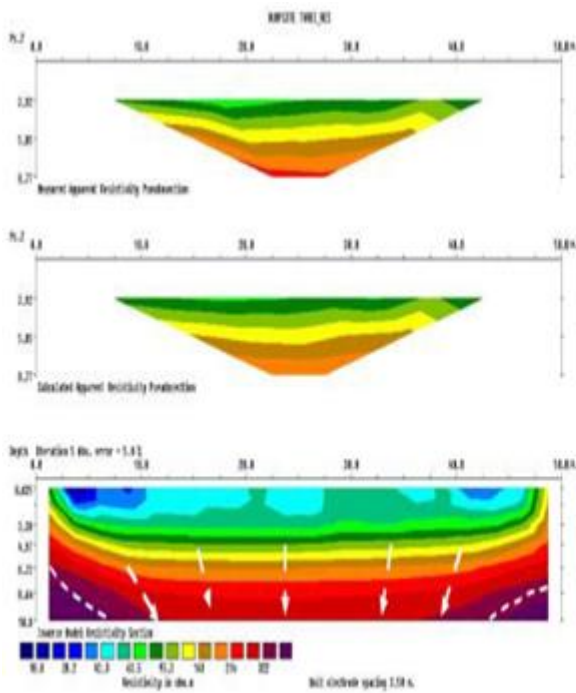


fig.4c; profile three ERT

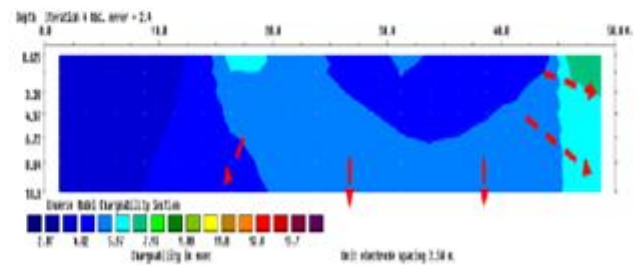


fig.4cii; profile three IPT

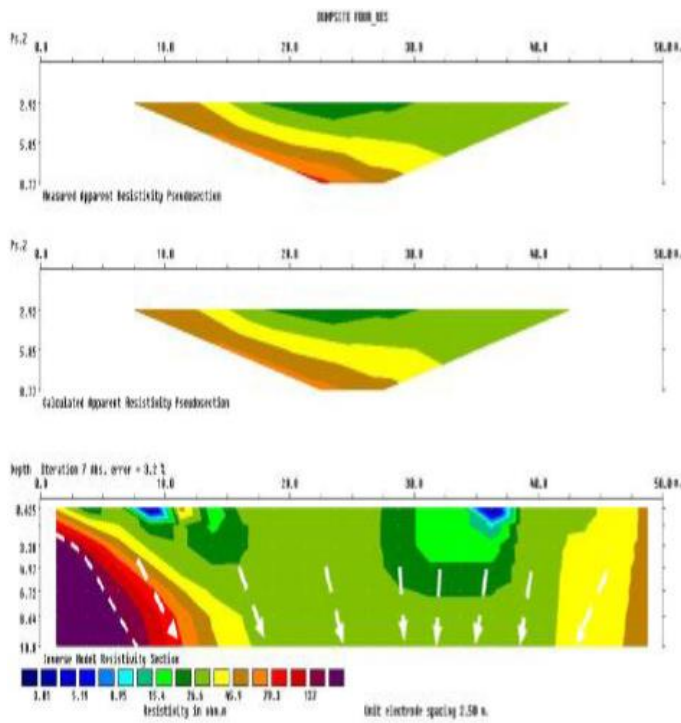


fig.4d; profile four ERT

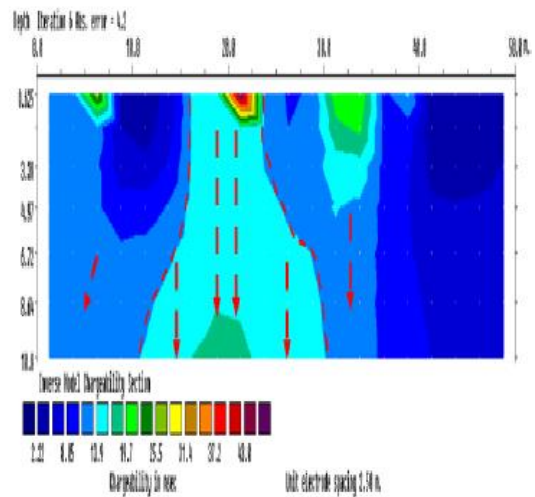


fig.4diii; profile four IPT

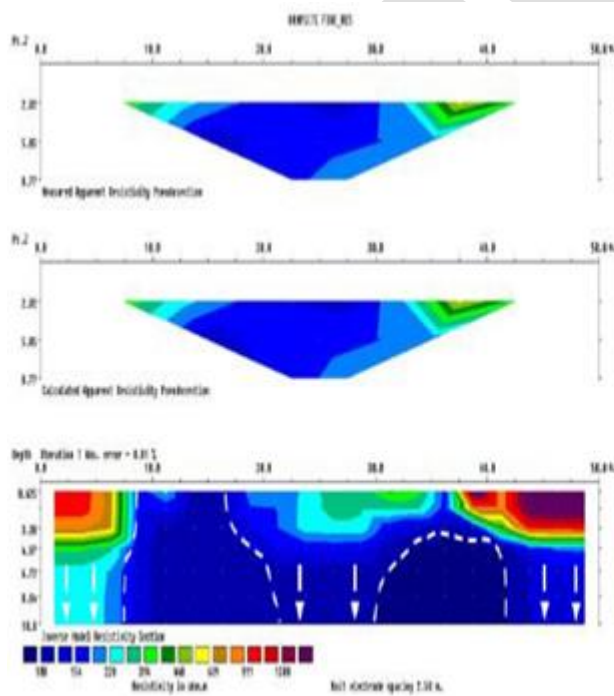


fig.4e; profile five ERT

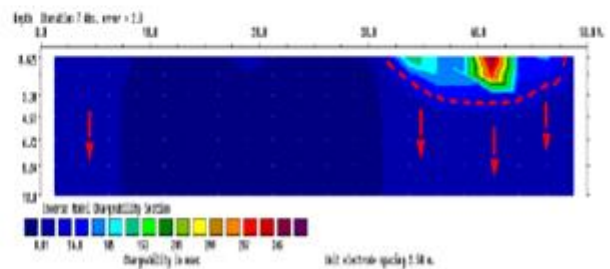


fig.4eiv; profile five IPT

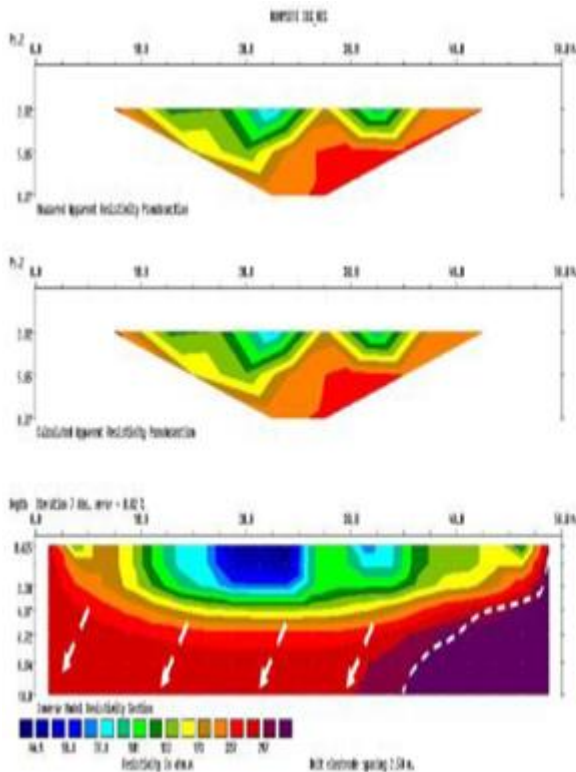


fig. 4f; profile six ERT

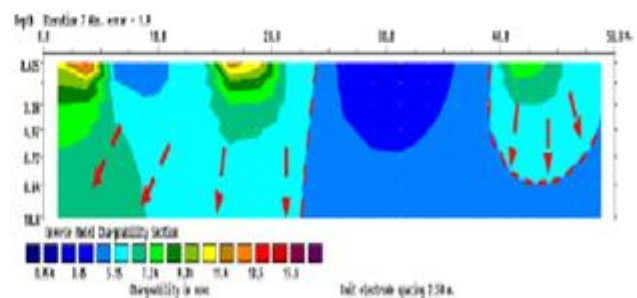


fig. 4fvi; profile six IPT

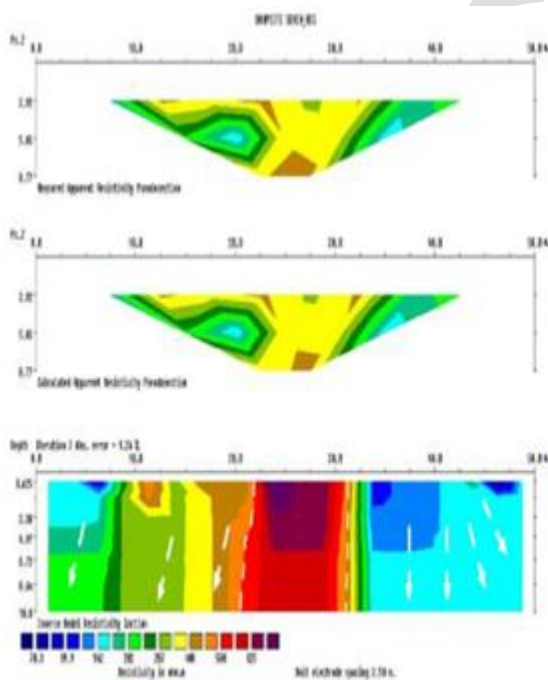


fig.4g; profile seven ERT

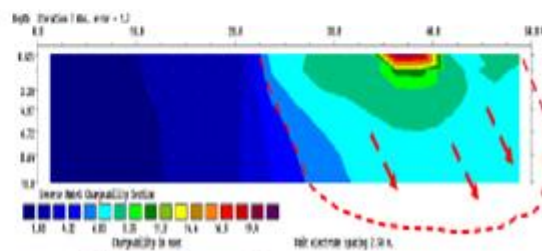


fig. 4gvii; profile seven IPT

5. Conclusion

Integration of electrical methods has been used to assess the subsurface conditions under a municipal dumpsite system in Kwale. Our geophysical method was able to map and delineate the contaminant plume (leachate) beneath the open dumping system. The integrated methods have proven to be tools for environmental assessment of waste site. The ERT, IPT and

VES indicate a polluted depth of over 10 m. The integrated use of non-destructive geophysical techniques; induced polarization and resistivity methods have been successful in characterizing the waste site in Kwale town. The integration of geology from the vertical electrical sounding, electrical resistivity tomography and induced polarization measurements yielded useful information for mapping the site. The results were able to delineate the spatial distribution of the leachate plume due to the fact that in general, for resistivity values $< 100 \Omega\text{m}$ the area is expected to have high leachate content. Values greater than that were interpreted as coarse sand. The result of ERT and IPT indicated the direction of leachate flow in the surveyed areas.

REFERENCES

- (1) Ofomola, M.O., A. F. Akpolile, Ochuko, A., Taiye O. Adeoye, Musa A. B. (2021). Detection of Trace Metal Contamination Around a Dumpsite in Iyara Area Warri Nigeria Using Geoelectrical and Geochemical Methods. *Environmental and Earth Sciences Research Journal* Vol. 8, No. 3, September, 2021, pp. 125-133.
- (2) Isua, A. I., Salami, S. A., and Osayomore, M. O. (2020). Geophysical survey for groundwater investigation around parts of University of Benin, Nigeria. *international journal of engineering applied sciences and technology*, vol. 4, ISSN no. 2455-2143, 360-364pp.
- (3) Ige, O.O., Adebawale, T. Owolabi, Oluwaseun F. Olabode, Obasaju, D.O. (2022). Groundwater quality evaluation: a case study of Igando waste dumpsite, southwestern Nigeria. *Applied Water Science* (2022) 12:79 <https://doi.org/10.1007/s13201-022-01601-x>.
- (4) Oseji, JO; Atakpo, EA; Okolie, EC (2005). Geoelectric investigation of the aquifer characteristics and groundwater potential in Kwale, Delta State, Nigeria. *J. Appl. Sci. Environ. Manage.* 9: 157-160.
- (5) Short, KC; Stauble, AJ (1967). Outline of geology of Niger Delta. *Amer. Assoc. Petrol. Geol. Bull.* 51:761-779.
- (6) Doust, H. and Omatsola, E. (1990). Niger Delta: In J.D. Edwards P.A. Santogrossi, eds divergent/passive margin basins: *American Association of Petroleum Geologists Bulletin Memoir* 48, pp 201-328.
- (7) Nwajide CS (2006). A guide for geological field trips to Anambra and related sedimentary basins in southeastern Nigeria. PTDF Chair, University of Nigeria, Nsukka.
- (8) D. Fatta, A. Papadopoulos and M. Loizidou, "A Study on the Landfill Leachate and Its Impact on the Groundwater Quality of the Greater Area," *Environmental Geochemistry and Health*, Vol. 21, No. 2, 1999, pp. 175-190. [doi:10.1023/A:1006613530137](https://doi.org/10.1023/A:1006613530137)
- (9) S. Mor, K. Ravindra, P. R. Dahiya and A. Chandra, "Leachate Characterization and Assessment of Groundwater Pollution near Municipal Solid Waste Landfill Site," *Environmental Monitoring and Assessment*, Vol. 118, No. 1-3, 2006, pp. 435-456. [doi:10.1007/s10661-006-1505-7](https://doi.org/10.1007/s10661-006-1505-7)
- (10) E. A. Ayolabi, A. F. Folorunso, L. Adeoti, S. Matthew and E. Atakpo, "2-D and 3-D Electrical Resistivity Tomography and Its Implications," *The 4th Annual Research Conference and Fair held at the University of Lagos*, Akoka, 8 January 2009, p. 189.
- (11) E. A. Ayolabi, "Geoelectric Evaluation of Olushosun Landfill Site Southwest, Nigeria and Its Implications on Groundwater," *Journal of the Geological Society of India*, Vol. 66, No. 3, 2005, pp. 318-322.
- (12) E. A. Ayolabi and J. Oyelayo, "Geophysical and Hydrochemical Assessment of Groundwater Pollution Due to a Dumpsite in Lagos, Nigeria," *Journal of the Geological Society of India*, Vol. 66, No. 5, 2005, pp. 617-622.
- (13) E. A. Ayolabi and D. Y. Peters, "Hydrochemical and Electrical Resistivity Assessment of the Impact of Solid Waste on the Groundwater at Oke-Alfa Refuse Dumpsite," *Journal of Engineering and Technology Management*, Vol. 12, No. 1, 2004, pp. 5936-5946.
- (14) J. Pastor and A. J. Hernández, "Heavy Metals, Salts and Organic Residues in Old Solid Urban Waste Landfills and Surface Waters in Their Discharge Areas: Determinants for Restoring Their Impact," *Journal of Environmental Management*, Vol. 98, 2012, pp. S42-S49. [doi:10.1016/j.jenvman.2011.06.048](https://doi.org/10.1016/j.jenvman.2011.06.048)
- (15) E. G. Akpokodje, "Principle of Applied and Environmental Geology," Paragraphic Publisher, Port Harcourt, 1999.
- (16) A. A. Odewande, "Hydro-Chemical Characteristic of the Environmental Impact of Solid Waste; Dumpsite, Orita Aperin Central Refuse Dumpsite in Ibadan. South Western, Nigeria," Ph.D. Thesis, University of Ibadan, Ibadan, 1999.
- (17) G. J. Udom and E. O. Esu, "A Preliminary Assessment of the Impact of Solid Wastes on Soil and Groundwater System in Part of Port Harcourt City and Its Environs. Rivers State Nigeria," *Global Journal of Environmental Sciences*, Vol. 4, No. 1, 2005, pp. 23-30.
- (18) S. Anon, "Hydrogeological Implications in Solid Waste Disposal," *Bull. International Association Hydrogeol*, 2009, pp. 146-158.
- (19) N. J. Bunce, "Environmental Chemistry," Winers Publishing Ltd., Winnipeg, 1990.

- (20) W. D. Daily, W. Lin and T. Buscheck, "Hydrological Properties of Topopah Spring Tuff: Laboratory Measurements," *Journal of Geophysical Research*, Vol. 92, No. B8, 1987, pp. 7854-7864. [doi:10.1029/JB092iB08p07854](https://doi.org/10.1029/JB092iB08p07854)

NCPS-2024

Mathematical Analysis of Poor Gas Mileage in Automobiles Using Cause-effect Diagram

Bassey Echeng Bassey

Department of Mathematics, University of Cross River State, 540252, Calabar, Nigeria

*Correspondence email: echengba.rexson@yahoo.com

Abstract

Credibly, with the importance of automobiles in the activities of mankind, it is eminent to note that automobiles are often confronted with a number of mechanical constraints among, which is poor gas mileages. Worth noting, is the fact that no comprehensive literature on mathematical application of Cause-effect Diagram in analysis and evaluation of poor gas mileage in automobiles. Therefore, this study seeks and explored mathematical analysis theory in the investigation of the consequential effect of the emission of poor gas mileage in automobiles and critical evaluation of root cause effect. The system was determined by a deterministic mathematical approach with a cluster unites of automobiles investigated using the 3Ms and P-methods (method, materials, machinery, and people) of Ishikawa model known for its fishbone-like structure in the presence of gasoline. Analytic predictions of the investigation evolved round a 6-step identification approach, which clearly defined the outcome (effect) been analyzed. Furthermore, the use of chart pack with horizontal spine directed to an effect box and identification of the main causes and sub-branches contributing to the effect was explored. Diagram analysis were conducted. The results of the analysis indicated that under balanced constructed system with no repeated causes, poor maintenance appears to be a cause for which development of durable measurement tools cannot be overruled and therefore, requires urgent attention. Thus, this novel approach is worthy of replication in related scientific investigations.

Keywords: *Poor-gas-mileage, automobiles, Cause-effect-diagram, effect-box, horizontal-pine, sub-branches*

1. Introduction

A Cause-effect diagram (also called the Ishikawa diagram) is a mathematical tool used for the identification of the root causes of mathematical quality problems. It was named after Kaoru Ishikawa, a Japanese quality control statistician, who pioneered the use of this chart in the 1960's [1]. Because of the structure and function of the Ishikawa diagram (Cause-effect diagram), the system is often referred to as a "Fishbone diagram" [2]. The Fishbone diagram is an analysis tool that provides a systematic representation of the relationship between the event under investigation and all possible causes influencing the giving system – the cause-effect [3]. That is, a Cause-and-Effect Diagram is a documentation of group thinking process that investigate the root cause of an event, leading to stratification and collection of data in confirmation of system relationship with counter measures[4].

The steps involved define clearly, the problem or effect or event for which the cause is to be identified. Structurally, a horizontal line with an arrow at the right-hand end and a box in front of it is drawn with the problem statement written in the effect box. The next step is identifying the main causes, which are then denoted as major categories. Brain storming is normally used to identify the possible major causes. After identifying a primary cause, the team goes into deeper identification of many secondary or tertiary causes as possible as sub-branches of the primary causes. Each of the major causes is placed in a box horizontal to the first line and connected to that line at an inclination of approximately 70° [5], [6]. After identifying the major causes, the root causes are investigated by adopting root-cause analysis techniques. The logical validity is checked for all causes identified considering the present scenario. It is important to understand the potential pit falls while using Cause-and-Effect diagram, which shouldn't be treated as a substitute for data. Rather, it should be drawn only after preliminary data has been collected to narrow down the focus of a problem [7], [8]. An overview of cause-effect diagram is as depicted in fig. 1:

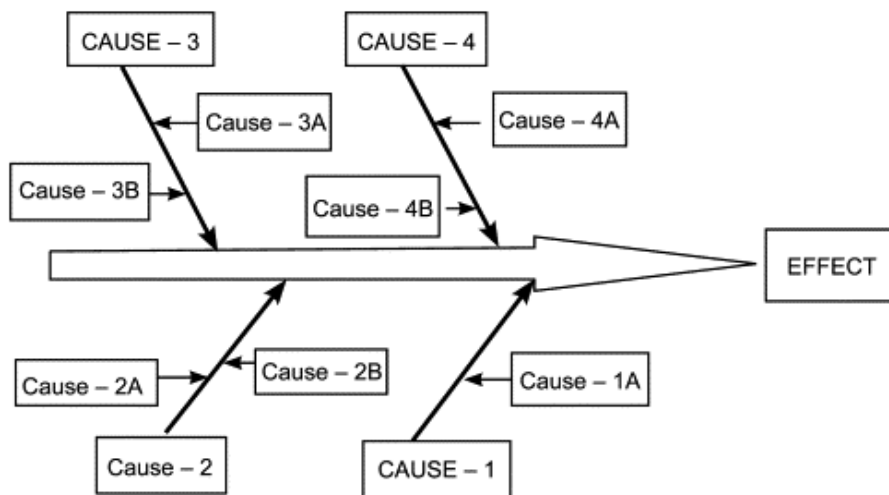


Fig. 1: Cause-effect diagram [8]

Often there have been alarming occurrence of poor gas mileage in automobile, which could be attributed to several number of causes i.e. the method of driving, the machinery, material components of the system and the set of individuals operating the system. Worth noting is the fact that no comprehensive literature on mathematical application of Cause-effect Diagram in analysis and evaluation of poor gas mileage in automobiles. Therefore, in this study, we seek to explore mathematical analysis theory to investigation the consequential effect of poor mileage emission of gas in automobiles and possible evaluation of root cause effect.

The diagram can also be used to determine the risks of the causes and sub-causes of the effect, but also of its global risk [9]. Usually, the analysis after Fishbone diagram continues with other representation and establishing treatment priorities methods as can be seen in fig. 2:

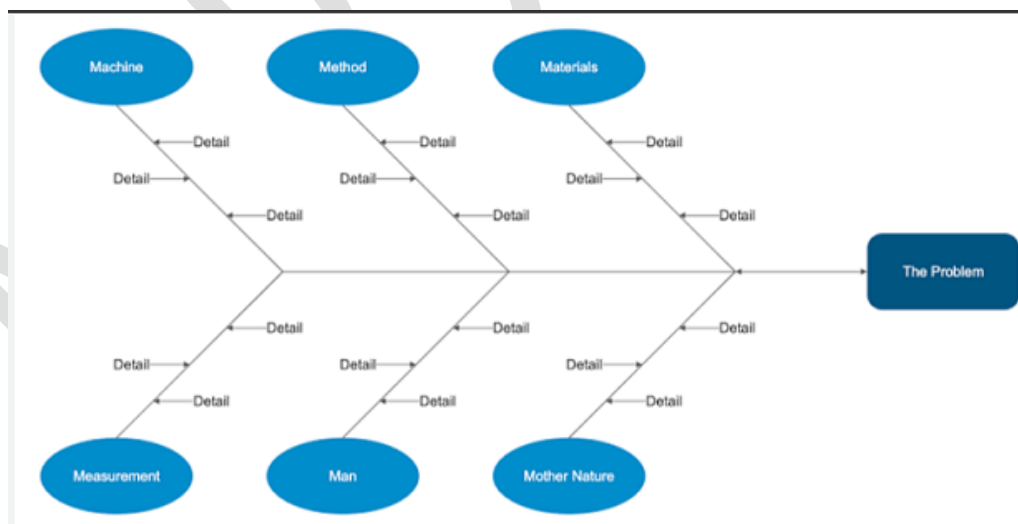


Fig. 2: Fishbone Diagram: The 6 Ms[10]

Mathematically, Cause and Effect diagram or Fishbone diagram is a graphic technique and is a good tool to find and significantly analyze affecting factors in identifying the characteristics of work output quality [10]. This fishbone diagram is known as a cause-and-effect diagram. Why is it that this Ishikawa's diagram has been called "fish bone"?

Well, when observed the diagram (see diagram 1), the fishbone, its form has a similarity to a fish, which has a head (as an effect) and a body in the form of bones, illustrated as causes of known problems [11], [12].

Remarkably, is the fact that by dissolving any one of the multiple root causes, the symptoms can be reduced or even eliminated. Positive or negative—Our successes, as well as our failures, have root causes. By studying the roots for our successes, we may find strategies that can be applied to improving all of our processes. Symptoms—In dealing with problems, symptoms are found at the surface. They are the —red flag that draws attention to the issue. A symptom is usually a noticeable gap. In this research therefore, we seek to determine critical inevitable causes and possibilities of ameliorating these causes using mathematical tool known as “cause-effect diagram”, noting that a cause-effect diagram is a visual tool used to logically organize possible causes for a specific problem or effect by graphical display of the increasing detail. Therefore, the Ishikawa model known for its fishbone-like structure, is significance in the sense that it aids in teams understand of many causes that contribute to an effect of a given problem. Moreso, it graphically not only displays the relationship of the causes to the effect and to each other but also, identification of areas for improvement. Therefore, the application of cause-effect diagram for the evaluation of poor gas mileage in automobiles is timely. The outcome of which will meaningful to all automobile usages and for the purpose of academic enhancement.

Thus the entire composition of the present investigation revolves round 5-sections with section 1, depicting the introductory aspect of the study. Section 2, is focuses on to the materials and methods as adopted in this work. Section 3, is devoted to the formulation of a cause effect diagram for poor gas mileage in automobiles. The mathematical illustrations and discussion arising from analytical predictions is domicile in section 4. Finally, in section 5, we present study incisive conclusion and recommendations. Remarkably, the investigation is targeted to exhume insight to the methodological application of Cause-effect diagram for the evaluation of consequential effect of poor gas mileage in automobiles.

2. Materials and Methods

In this section, we shall discuss the materials and methods (methodology) as applied in this investigation. The materials and methods of the present study lies in in the application of fundamental theory of differential equation in relation to mathematical modeling emanating from the gap in literature of existing models. That is, we consider the mechanism of an automobile in the presence of gasoline.

2.1. Statement of the problem

Many of the main causes for poor gas mileage are mechanical and inclusive of improper air flow, faulty injectors or bad oxygen sensors. The concept of the Cause-effect diagram is said to have been around since the 1920s, but it was popularized by Japanese professor of engineering, Kaoru Ishikawa, who, most notably, put into place the quality management processes for the Kawasaki shipyards. One of the famous uses of the fishbone diagram wasn't used to find causes of existing problems, but rather in the design phase to *prevent* problems. Mazda Motors used fishbone diagrams to design the Miata (MX5) sports car. Details down to the design of the car's doors, so drivers could rest their arm on it while driving, were considered [5].

In nearly any industry or business, variation is qualities and efficiency's kryptonite. Once an order has been placed, a predictable and standardized process should be followed through to delivery. Variation in that process usually means variation in the final product. Fishbone diagrams are a great tool to either determine potential variables in a process to mitigate defects or failures, or to figure out what's causing a defect or failure that is currently occurring. Worth noting, is the fact that no comprehensive literature on mathematical application of Cause-effect Diagram in analysis and evaluation of poor gas mileage in automobiles. Therefore, in this study, we seek to explore mathematical analysis theory to investigation the consequential effect of poor mileage emission of gas in automobiles and possible evaluation of root cause effect. That is, present investigation is geared towards:

- i. Formulating a cause-effect diagram that mathematically identify and analysis poor utilization of gas by automobiles using the Ishikawa method (cause-effect diagram)
- ii. Determine the root causes of poor gas mileage in automobiles.

2.2 Model formulation proceedings

Here, in analyzing the effect of poor gas utilization in automobile, we adopt the M6 (Man-Machine- Method-Material- Measurement- environment) techniques for the purpose of the system evaluation. For instance, *Man*—It should be checked and confirmed that the operators who are involved in the in-seam of jeans are following standard work instructions and they are possessing the necessary skills to undertake the tasks. The absence of proper training and lack of skills, and non-adherence to the instructions may result in the wrong panel attach where proper seam margins are not kept by the operators. Also, there are chances that the operator may not be following the notch marks of the leg panels.

Machine—The in-side leg seam of jeans pant is generally done using a feed-of-arm (FOA) machine, which is a double-needle chain stitch machine. During RCA, it should be ensured that all the machine settings and work-aids are done correctly. There can be problems due to faulty settings (such as stitches-per-inch settings of the machine) or not properly fitted seam folder (work-aid), and due to which the machine is not able to take the sufficient seam allowance (fabric margin) while stitching. If the fabric margin is lesser in the seam, it may come out during the washing process.

Method—The method of in-seam stitching should be checked properly. The manner fabric plies (leg panels) are aligned precisely, it is inserted in the seam folder properly, and the handling of the panels while stitching using the FOA machine is some key points to be checked while investigating the method. Also, it is important to check and confirm that all the operators are following the same method during stitching. Apart from the stitching method, the method of pattern making, cutting, and washing methods can also be the source of problems, and each of these should be investigated thoroughly. For example, if the seam margins are not properly followed while cutting or the notches are not marked properly, the alignment of the panels during sewing may get disturbed and may lead to the opening of in-seam.

Material—The reason of defect may be a faulty material as well. For example, if the fabric is having some defects like skewness where the fabric is not balanced lengthwise and getting distorted. The other issues may be related to the quality of the sewing thread used, there can be chances where the sewing thread is not strong enough to sustain the rigorous washing treatment. Sometimes, if the fabric is stretchy or elastic it may also distort the seam responsible for seam open.

Measurement—This is related to the measurement- or dimension-related issues causing errors. There are chances that the seam margin is wrongly measured and mentioned, it may be due to a wrong scale used. Also, due to wrong measurements, the settings of the seam folder may be done incorrectly causing insufficient fabric margins inside the seam.

Milieu (environment)—The environment may also be a factor that is responsible for the generation of the defect. The opening of the seam may be due to the high temperature (at the needle point) during stitching, where the sewing thread is losing its strength. Also, this problem of opening the seam, may occur due to the rise in the temperature (inside the washing machine) during washing, that may cause weakening of the seam. Further, inadequate light while sewing may act as the source of the problem, as the operator may not be able to see the fabric panels, and the seam line properly due to lower light intensity at the needlepoint.

3. Formulation of a Cause-effect Diagram for Poor Gas Mileage

For a typical poor gas mileage in automobile, developing a Cause-and-Effect Diagram, requires the construction of a structured, which is a pictorial display of a list of causes organized to show their relationship to a specific effect[13]. Essentially, fig.3 below shows the basic layout of the Cause-and-Effect Diagram.

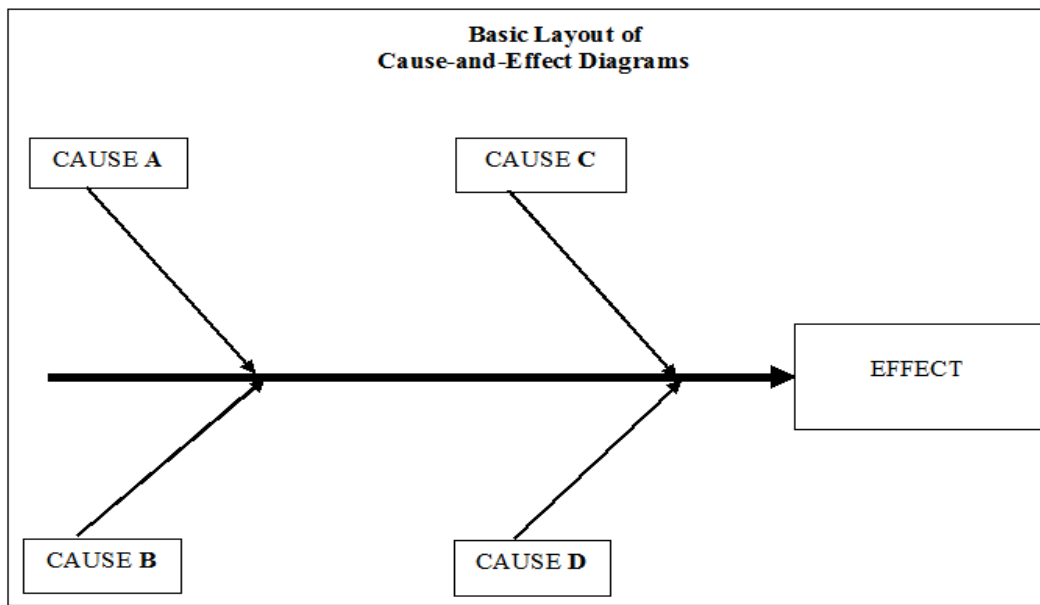


Fig. 3: Initiating a cause-effect diagram

From fig. 3, we observe that the structure is characterized by a *cause* side and an *effect* side. Thus, the steps for constructing and analyzing a Cause-and-Effect Diagram are outlined below:

3.1. Step 1 - Identify and clearly define the outcome or EFFECT to be analyzed

This step involves:

- a) Decide on the effect to be examined. Effects are stated as particular quality characteristics, problems resulting from work, planning objectives, and the like.
- b) Use Operational Definitions. Develop an Operational Definition of the effect to ensure that it is clearly understood.
- c) Remember, an effect may be positive (an objective) or negative (a problem), depending upon the issue that's being discussed.
 - i. Using a positive effect which focuses on a desired outcome tends to foster pride and ownership over productive areas. This may lead to an upbeat atmosphere that encourages the participation of the group. When possible, it is preferable to phrase the effect in positive terms.
 - ii. Focusing on a negative effect can sidetrack the team into justifying why the problem occurred and placing blame. However, it is sometimes easier for a team to focus on what causes a problem than what causes an excellent outcome. While you should be cautious about the fallout that can result from focusing on a negative effect, getting a team to concentrate on things that can go wrong may foster a more relaxed atmosphere and sometimes enhances group participation.

In a summary form for step 1, is described below:

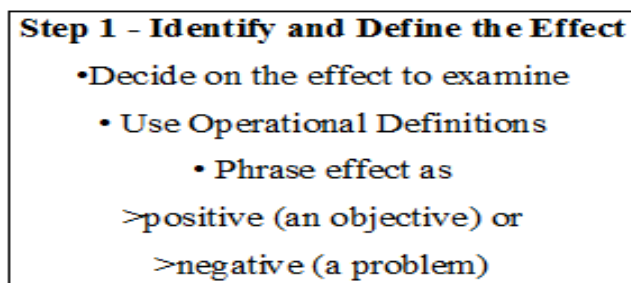


Fig. 4: Summary form for step 1

3.2. Step 2 - Use of chart pack with SPINE and EFFECT box.

This step involves:

- i. Draw a horizontal arrow pointing to the right. This is the spine.
- ii. To the right of the arrow, write a brief description of the effect or outcome which results from the process.

For instance, we illustrate this using our poor gas mileage for automobile as depicted by fig. 5 below:

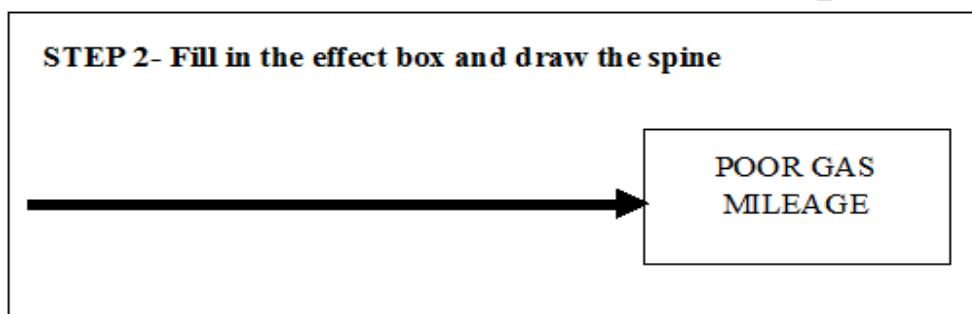


Fig. 5: Spine and effect-box for cause-effect diagram

3.2. Step 3 - Identify the main CAUSES contributing to the effect being studied

These are the labels for the major branches of your diagram and become categories under which to list the many causes related to those categories. The procedure includes:

- a) Using audible identifiable labels, we establish the main causes, or categories, under which other possible causes will be listed. Here are some commonly used categories:
 - i. 3Ms and P - methods, materials, machinery, and people
 - ii. 4Ps - policies, procedures, people, and plant
 - iii. Environment - a potentially significant fifth category
- b) Write the main categories your team has selected to the left of the effect box, some above the spine and some below it.
- c) Draw a box around each category label and use a diagonal line to form a branch connecting the box to the spine.

An illustration of step 3 is given by fig. 6, below, where we had use the 3Ms and P to start the development of the diagram incorporating step 2.

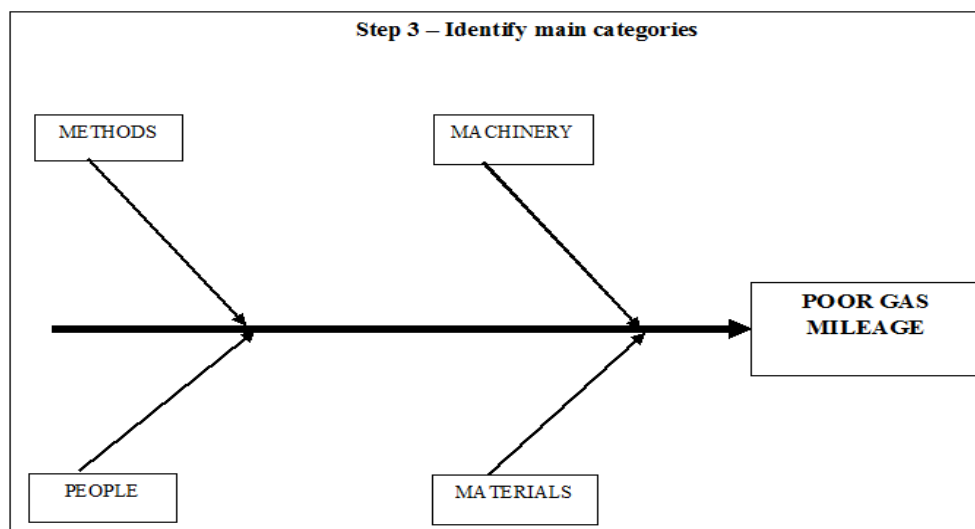


Fig. 6: Identify main categories for cause-effect diagram

3.3. Step 4 – Identification of other sub-branches

Here, we are required to induce specific causes, which are closely related to each major branch. That is, input of other specific factors which may be the CAUSES of the EFFECT. It is expected to identify as many causes or factors as possible and attach them as sub-branches of the major branches. We illustrate this step using poor gas mileage in automobile by filling the details for each cause as in fig. 7. If a minor cause applies to more than one major, list it under both.

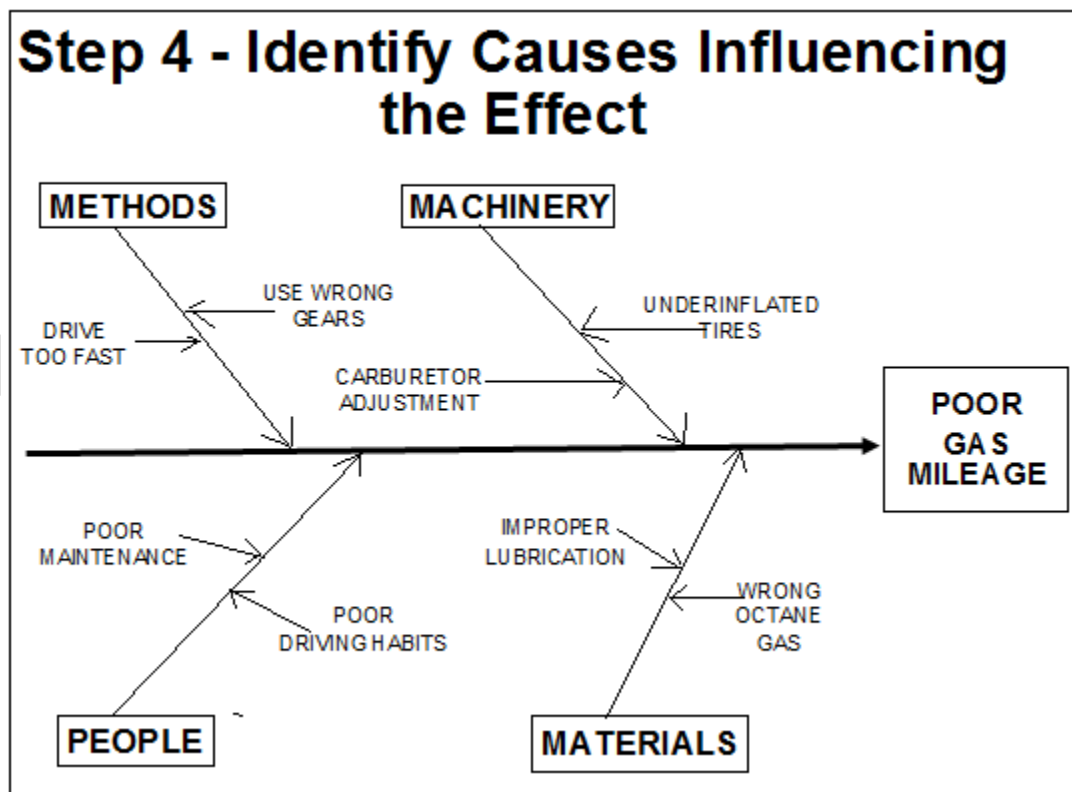


Fig. 7: Identification of Cause-effect sub-branches

3.5. Step 5 – Identification of more detailed levels

Remarkably, Cause-effect diagram could be more explicit by taking further steps in identify increasingly more detailed levels of causes and continue the organization of these sub-causes under related sub-branches or categories. This can be achieved by asking a series of “why questions.” For instance, we’ll use a series of why questions to fill in the detailed levels for one of the causes listed under each of the main categories.

Q: Why was the driver USING THE WRONG GEAR?

A: The driver couldn't hear the engine.

Q: Why couldn't the driver hear the engine?

A: The radio was too loud.

A: Poor hearing

Q: Why were the TIRES UNDERINFLATED?

A: No record of tire pressure

A: Difficult air stems

Q: Why were the air stems difficult?

A: Poor design

Q: Why was MAINTENANCE POOR?

A: Lack of money

A: No awareness

Q: Why was WRONG OCTANE GAS used?

A: Didn't know recommended octane

Q: Why wasn't recommended octane known?

A: No owner's manual

Of interest, the illustration of answers to the question shall be fully demonstrated in next our chapter, where cogent and explicit case of poor gas mileage will be considered. In that chapter, it is important to show how the diagram looks when all the contributing causes that were identified by the series of why questions have been filled in.

Remark 1

You may need to break your diagram into smaller diagrams if one branch has too many sub-branches. Any main cause (3Ms and P, 4Ps, or a category you have named) can be reworded into an *effect*.

3.6. Step 6 – Diagram analysis

Here, we consider the analyzation of our system diagram. The essence here is the fact that diagram analysis enhances the identification of causes that warrant further investigation. Since Cause-and-Effect Diagrams identify only possible causes, you may want to use a Pareto Chart to help your team determine the cause to focus on first. The essential criteria are:

- a) Look at the “balance” of system diagram, checking for comparable levels of detail for most of the categories. The following points must be noted:
 - i. A thick cluster of items in one area may indicate a need for further study.
 - ii. A main category having only a few specific causes may indicate a need for further identification of causes.
 - iii. If several major branches have only a few sub-branches, you may need to combine them under a single category.
- b) Look for causes that appear repeatedly. These *may* represent root causes.
- c) Look for what you can measure in each cause so you can quantify the effects of any changes you make.
- d) Most importantly, identify and circle the causes that *you can take action on*.

For instance, we present here, the theoretical analysis expected of a constructed cause-effect diagram i.e.

- i. The level of detail is pretty well balanced.
- ii. No causes are repeated.
- iii. *Poor Maintenance* appears to be a cause for which you could develop measurements.
- iv. Moreover, for typical case of poor gas mileage, it could be anticipated that *Poor Maintenance* could often appears to be a cause that may require immediate action on.

A clearer illustration will be given in out next section.

4. Mathematical Illustration and Discussion

In line with our set goal, we shall illustrate as a typical example, the consequential role of poor gas mileage in automobile operation. Following the outlined stages in section three, we evaluate the causes of poor gas mileage in the operation of automobiles.

4.1 Illustrative example

Example 1 (*Using poor gas mileage in automobiles*)

Construct a cause-effect diagram (fishbone) to illustrate the causes of poor gas mileage on the running of automobiles. Analyzes the diagram and mark out causes of immediate action.

Solution

The required sketch of the details for the cause-effect of a poor gas mileage in the operation of automobiles is as follows;

- i. Positioning of horizontal arrows spine to the right to an effect box clearly inscribed computer downtime.
- ii. Identification of main causes, main categories contributing to the effect and their presentation in small boxes.
- iii. Connection of main categories with diagonal arrows (to form fish bone) to the horizontal spine arrow.
- iv. Presentation of sub branches identity as specific factors as causes of the effect.
- v. Analyze the diagram and marking out causes that could require immediate action.

The completed diagram is as depicted in fig.8:

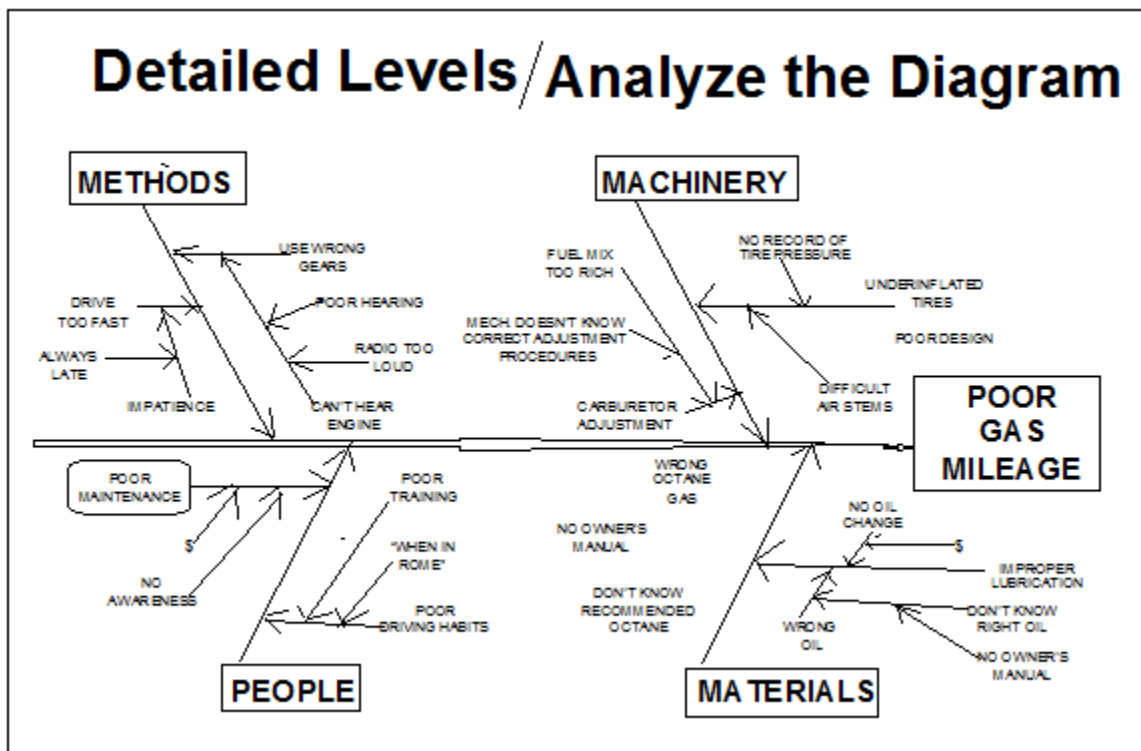


Fig.8: Diagrammatic representation of details and analysis of a Cause-effect for poor gas mileage in automobiles

4.2 Discussion

In this project work, we took into account that fact that it has been of concern to automobiles users the frequent cases of poor mileage of gas utilization. A situation that has not been given due consideration. To this effect, we have proposed a model that critically understudy the dynamics leading to the aforementioned situation with the goal of identifying these causes and possible areas that needed urgent attention.

Remarkably, the construction of the entire system emanated from the identification of the major cause effect, which was duly engraved in a box, called effect-box directed from left to right by a spine arrow. The procedure further led to the identification of four major branches (categories). The explicitness of the model requires further sub-branches from which more sub-sub branches were observed. In a nutshell, we critically analyze our system following the step-by step construction of our desired cause-effect diagram. Of note, the following have been observed:

- i. That the details of the system are explicitly balanced.
- ii. That no cause(s) were repeated
- iii. Poor maintenance appears to be a cause for which durable measurements could be developed.
- iv. Moreover, poor maintenance appears to be a cause that urgent attention is required and which has been clearly circled, earmarking it for further investigation.

5. Conclusion

Following the construction of a cause effect diagram for poor gas mileage in automobile, the study significantly developed a mathematical model that identified the causes of poor gas utilization by automobiles. Furthermore, the study not only identified the possible root causes but form a tool that sorted out and related main factors affecting poor utilization of gasoline. Remarkably, the investigation afforded an explicit template for the analyzation of the existing problems as it affects poor gasoline utilization. Specifically, the result of system analysis indicated that poor maintenance appears to be a cause for which immediate measurement tool could be developed in order to ameliorate

the continuous utilization of poor gasoline in automobile system. Thus, the model is highly recommended for possible replication in other related scientific investigations.

References

- [1] Juran J. M. (1999). *Juran's Quality Handbook*. (5th Edition). McGraw Hill.
- [2] Graham, A. (2004). Cause and effect diagram. In P. Fahlstrom & R. G. Björk (Eds.), *The international encyclopedia of ergonomics and human factors* (pp. 1577-1580). Taylor & Francis.
- [3] Kume, Y. (2017). Application of the Ishikawa diagram to identify factors that affect patients' trust in healthcare organizations. *Japanese Journal of Quality and Safety in Healthcare*, 12(2), 112-118.
- [4] Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. John Wiley & Sons
- [5] Okes D. (2019). Root Cause Analysis: The Core of Problem Solving and Corrective Action [2nd ed.]. Available online at: <https://dokumen.pub/root-cause-analysis-the-core-of-problem-solving-and-corrective-action-2ndnbsped-9780873899826-0873899822.html>
- [6] Anderson P., Bruijn A., Angus K., Gordon R. and Hastings G. (2009). Impact of Alcohol Advertising and Media Exposure on Adolescent Alcohol Use: A Systematic Review of Longitudinal Studies, *Alcohol & Alcoholism*, 44(3), 229–243.
- [7] Lewis S. (2020). Fishbone diagram (Ishikawa cause and effect). Available online at: <https://www.techtarget.com/whatis/definition/fishbone-diagram>
- [8] Purushothama B. (2013). *Work Quality Management in the Textile Industry*. Woodhead Publishing, India.
- [9] Ciocoiu, C. N. (2008). *Managementul riscului. Teorii, practici, metodologii*. Bucharest: ASE.
- [10] San, G. S., Tjitro, S., and Santoso, M. (2003). Desain Eksperimen untuk Mengoptimalkan Proses Pengecoran Saluran Keluar Teko. *JURNAL TEKNIK MESIN*, 5(1), 5–10. Available online at: <http://www.academia.edu/1071634/>
- [11] Tiann. (2012). Diagram Fishbone dari Ishikawa. Retrieved on 5th Oct., 2024. Available at: <https://tianno.wordpress.com/2012/05/>
- [12] Preuss, P. G. (2003). *School Leader's Guide to Root Cause Analysis: Using Data to dissolve Problems*. Larchmont, NY: Eye on Education.
- [13] Ishikawa, K. (1985). *What is total quality control? The Japanese way*. Prentice Hall.

Exploring patterns of breast cancer tumor characteristics using factor analysis

Ekaette Inyang Enang and Elizabeth Abednego Essien

Department of Statistics, University of Calabar, Calabar, Nigeria

Abstract

Breast cancer is one of the significant global health challenges, demanding a deeper understanding of its complex nature for improved diagnosis and treatment. This study sought to identify the underlying factors in breast cancer tumor characteristics using the factor analysis technique. The dataset which comprised of 7 tumor characteristics measurements of 569 patients was sourced from kaggle.com, <https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset>. Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were employed to justify the factor analysis assumptions thereby ensuring the suitability of the dataset for factor analysis. The principal component technique was used to estimate the factor loadings and communalities while the varimax rotation technique was employed in the interpretation of these latent factors. The result of the analysis showed that the first two components accounted for 81.7% of the total variance and that while the first component captured a set of variables related to tumor size (radius mean, perimeter mean, and area mean), the second component was associated with characteristics related to tumor shape (smoothness mean, compactness mean, and concavity mean). The identified patterns are recommended for improved diagnostic methods and tailored treatment approaches in the fight against breast cancer.

Keywords: Breast Cancer, Tumor Characteristics, Factor Analysis, Principal Component, Varimax Rotation, Latent Factors.

Introduction

Breast cancer is a health challenge characterized by the uncontrolled growth of abnormal cells within the breast tissue. Breast cancer ranks as the most frequently diagnosed cancer among women and holds the top spot as the most prevalent cancer worldwide. The year 2020 witnessed over 2.26 million new instances of breast cancer in women. It is the leading cause of cancer-related deaths globally, making it a critical public health concern. Breast cancer can affect individuals of all genders, although it is most commonly associated with women (World Cancer Research Fund International, 2022).

Tumors form when a single cell divides unchecked, leading to an unwelcome growth known as cancer. Breast cancer symptoms include an increase in breast mass, a change in breast size and form, a change in breast skin color, breast discomfort, and changes in the breast's genetic makeup. Worldwide, breast cancer has been described to be the second leading cause of death in women after heart disease and it affects more than 8% of women at some point in their lives. According to the WHO's annual report, more than 500,000 women have breast cancer every year. It is predicted that the prevalence of this disease will arise in the future due to environmental damage (World Health Organization Annual Report, 2015).

Obesity, hormone treatment therapy during menopause, family medical history of breast cancer, lack of physical activity, long-term exposure to infrared energy, having children later in life or not at all, and early age at which first menstruation occurs, race and ethnicity, being taller, some birth control methods and drinking alcohol/smoking are some of the risk factors for breast cancer in women. (Shaikh, *et al* 2021)

Many tests, including ultrasound, biopsy, and mammography, are performed on patients to determine whether they have breast cancer. This is because the symptoms of breast cancer vary widely. The biopsy, which includes the extraction of tissue or cell samples for analysis, is the most suggestive of these procedures. Humankind has struggled to understand and treat breast cancer since the earliest documentation more than 3500 years ago. The visible signs and symptoms of breast cancer and the palpability and tangibility of the lumps at later stages of the disease have enabled easy diagnosis by physicians in almost every period of recorded history. Tumor characteristics play a pivotal role in the prognosis, treatment planning, and personalized medicine for breast cancer patients.

Factor analysis is a multivariate technique designed to analyze correlations among many observed variables and to explore latent factors. Factor analysis was developed by the British statistician and psychologist Charles Spearman in the early 20th century as a technique for analyzing intelligence structure. Factor analysis plays a crucial role in data analysis and interpretation for several reasons: It simplifies complex datasets by reducing the number of variables while retaining essential information. This simplification makes data more manageable and interpretable. Factor analysis reveals hidden patterns and associations among variables that may not be apparent through direct observation. It uncovers the underlying structure that drives observed correlations. It groups related variables under common factors, providing a more comprehensive view of the data. This grouping helps researchers focus on broader constructs and simplifies subsequent analysis. Factor analysis enhances the interpretability of data by identifying the key factors that explain its variance. This understanding aids in making informed decisions and drawing meaningful conclusions (Tavakol & Wetzel, 2020).

Factor analysis has found applications in various fields, including psychology, social sciences, environmental sciences, machine learning, finance, medicine and health care, education, marketing, biostatistics and genetics, human resources, engineering, and more.

In studying diseases, Luo *et al.* (2020) applied factor analysis in the study of the clinical symptoms of coronavirus disease 2019 (COVID-19) based on a combination of factors. Factor loadings were calculated, and pairwise correlation analysis of symptoms was performed and the result of the investigation showed that the clinical symptoms of COVID-19 cases could be divided into respiratory-digestive, neurological, cough-wheezing, upper respiratory, and digestive symptoms.

Wu, *et al.* (2020) applied factor analysis in the investigation of the relationship between the gut microbiome and breast cancer risk factors and tumor characteristics. In their pilot cross-sectional study involving 37 incident breast cancer patients, they analyzed fecal samples collected before chemotherapy using a 16S ribosomal RNA gene-based sequencing protocol. Statistical analyses such as the Wilcoxon rank-sum test and a zero-inflated negative binomial regression model were employed, adjusting for total counts, age, and race/ethnicity. However, their work rightly acknowledged the need for further investigations to comprehensively understand the characteristics of the human microbiome and its intricate interplay with breast cancer hormone receptor status and established risk factors.

This study aims to identify underlining factors and relationships in breast cancer tumor characteristics by applying factor analysis to a set of seven key variables. Although there are numerous breast cancer tumor characteristics, this work restricts itself to seven of them, namely; radius mean, texture mean, area mean, perimeter mean, smoothness mean, compactness mean, and concavity mean. These variables derived from breast cancer cell images are commonly used in diagnostic processes, and hold crucial information about the nature and behavior of the tumor.

1. Methodology

2.1 Data and Data Preparation

The dataset which comprised of 7 tumor characteristics measurements of 569 patients was sourced from kaggle.com, <https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset>. The Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were used to verify that the dataset satisfied the assumption for Factor. Descriptive statistics were generated for the seven (7) variables using the SPSS software.

2.2 Factor Analysis Model

Factor analysis which is basically a one-sample procedure for possible applications to data with groups was used in the analysis. Assume a random sample y_1, y_2, \dots, y_p from a homogeneous population with mean vector μ and covariance matrix Σ . The factor analysis model expresses each variable as a linear combination of underlying factors f_1, f_2, \dots, f_m with an accompanying error term to account for that part of the variable that is unique (not in common with the other variables). For y_1, y_2, \dots, y_p in any observation vector y , the model is presented in equation 1.

$$\begin{aligned} y_1 - \mu_1 &= \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1m}f_m + \varepsilon_1 \\ y_2 - \mu_2 &= \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2m}f_m + \varepsilon_2 \\ &\vdots \\ y_p - \mu_p &= \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pm}f_m + \varepsilon_p \dots \dots \dots (1) \end{aligned}$$

Where:

y 's are the random variables

μ 's are called mean vectors

f 's are called latent factors

λ_{ij} 's are called loadings

ε 's are the residuals

Model (1) can be written in matrix notation as:

$$y - \mu = \Lambda f + \varepsilon \quad \dots\dots\dots (2)$$

Where:

$$y = (y_1, y_2, \dots, y_p)'$$

$$\mu = (\mu_1, \mu_2, \dots, \mu_p)'$$

$$f = (f_1, f_2, \dots, f_m)'$$

$$\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)'$$

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pm} \end{pmatrix} \dots\dots\dots (3)$$

The aim is to find f 's from y 's such that m is substantially smaller than p ($m < p$). Before fitting the model, Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were employed to justify the factor analysis model assumptions after which the factor loading (λ_{ij}) were estimated using the principal component method which were used to describe each extracted factor component. The varimax rotation, a widely used technique in factor analysis which serves in enhancing the interpretability of latent factors was used to streamline the understanding of how specific tumor attributes align with underlying factors.

2. Results and Discussion

3.1 Data Presentation

The data for this work is secondary data consisting of breast cancer tumor characteristics from kaggle.com <https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset>. The dataset comprises 569 subjects each with seven (7) distinct breast cancer tumor characteristics. The tumor characteristics include: radius mean (average size of the tumor's radius), texture mean (mean value of the grayscale levels in the tumor's texture), perimeter mean (mean perimeter of the tumor), area mean (mean area covered by the tumor), smoothness mean (mean smoothness of the tumor's boundaries), compactness mean (mean compactness of the tumor, $\text{perimeter}^2 / \text{area} - 1.0$), and concavity mean (mean level of concavity within the tumor). The data is presented in Appendix 1

3.2 Justification of Factor Analysis Assumption.

The result of the KMO and Bartlett spherical test is presented in Table 1. The KMO value of 0.719 which lies within the range of 0.7 to 0.8 indicating that the degree of information among the variables overlaps greatly. Hence factor analysis can be conducted on the dataset.

This is further justified by Bartlett's test of sphericity with a highly significant p-value (sig. = 0.00), indicating a significant relationship between the variables in the dataset. This is a positive outcome because factor analysis assumes that variables are related.

TABLE 1
KMO and Bartlett's Test of Sphericity for Breast Cancer Tumor

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.719
Bartlett's Test of Sphericity	Approx. Chi-Square	8086.053
	Df	21
	Sig.	<.001

3.3 Descriptive statistics

The result of the descriptive statistics of the variables is presented in Table 2. From the result, it can be observed that the columns display the number of data points (N), range, minimum, maximum, mean, standard deviation, and variance. The radius mean contains 569 data points with a range of 21.1290, varying from a minimum of 6.9810 to a maximum of 28.1100. The mean radius is approximately 14.13, with a standard deviation of 3.52, indicating the spread of data around the mean. The variance, at 12.42, quantifies the degree of dispersion within the dataset.

The statistics for other features are similarly presented, offering insights into the characteristics of the dataset. The area mean has a much larger variance compared to the other features, suggesting greater variability in the data points. Additionally, some features like smoothness mean exhibit very small standard deviations, indicating that the data points are clustered closely around the mean.

TABLE 2
Descriptive Statistics of the Breast Cancer Tumor Characteristics

Descriptive Statistics							
	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Radius mean	569	21.1290	6.9810	28.1100	14.127	3.524	12.419
Texture mean	569	29.5700	9.7100	39.2800	19.290	4.301	18.499
Perimeter mean	569	144.7100	43.7900	188.5000	91.969	24.2989810	590.440

Area mean	569	2357.500 0	143.5000	2501.0000	654.889	351.914	123843.554
Smoothness mean	569	.1108	.0526	.1634	.096360	.0140641	.000
Compactness mean	569	.3260	.0194	.3454	.104341	.0528128	.003
Concavity mean	569	.4268	.0000	.4268	.088799	.0797198	.006
Valid N (listwise)	569						

3.4 Factor Analysis Using Correlation Matrix

The correlation matrix of the dataset is presented in Table 3. The correlation matrix shows how variables in the analysis correlate with each other. Rows and columns correspond to different variables, while the values reflect the strength and direction of their linear relationship. For "radius mean" and "texture mean," the correlation is 0.324, suggesting a weak positive relationship between these variables. Similarly, high positive correlations are observed between "radius mean" and "perimeter mean" (0.998), "radius mean" and "area mean" (0.987), "perimeter mean" and "area mean" (0.987), "compactness" and "concavity mean" (0.883), indicating strong positive relationships between these variables. The "Sig. (1-tailed)" section provides p-values indicating the significance of the correlations. In this case, most of the correlations have extremely low p-values (typically < 0.001), suggesting that the observed correlations are statistically significant.

Table 4 shows the communalities of the variables. The initial communality of radius mean is 1.000 indicating that the variability in the variable is initially attributed to the variable itself. However, after conducting the factor analysis (using Principal Component Analysis), the extraction communality is 0.952. This suggests that the extracted factors explain 95.2% of the variance in radius mean, and the remaining variance (4.8%) is not captured by the factors. For texture mean, the extraction communality is 0.262 indicating that the extracted factors explain 26.2% of the variance in the variable while the remaining variance (73.8%) is not captured by the factors.

The total variance explained by the components is presented in Table 5. The result reveals that two (2) factors were extracted with a cumulative percentage of 81.748. The remaining five factors have lower cumulative loadings and minimal variation. Therefore, their contribution to the total variance is relatively small to be considered less important in this analysis, this result is further collaborated by the Scree plot presented in Figure 1.

TABLE 3
Correlation Matrix of Tumor Characteristics

Correlation Matrix a								
		radius mean	texture mean	perimeter mean	area mean	smoothness mean	compactness mean	concavity mean
Correlation	radius	1.000	0.324	0.998	0.987	0.171	0.506	0.677

	mean							
	texture mean	0.324	1.000	0.330	0.321	-0.023	0.237	0.302
	perimeter mean	0.998	0.330	1.000	0.987	0.207	0.557	0.716
	area mean	0.987	0.321	0.987	1.000	0.177	0.499	0.686
	smoothness mean	0.171	-0.023	0.207	0.177	1.000	0.659	0.522
	compactness mean	0.506	0.237	0.557	0.499	0.659	1.000	0.883
	concavity mean	0.677	0.302	0.716	0.686	0.522	0.883	1.000
Sig. (1-tailed)	radius mean		<.001	<.001	<.001	<.001	<.001	<.001
	texture mean	0.000		0.000	0.000	0.289	0.000	0.000
	perimeter mean	0.000	0.000		0.000	0.000	0.000	0.000
	area mean	0.000	0.000	0.000		0.000	0.000	0.000
	smoothness mean	0.000	0.289	0.000	0.000		0.000	0.000
	compactness mean	0.000	0.000	0.000	0.000	0.000		0.000
	concavity mean	0.000	0.000	0.000	0.000	0.000	0.000	
a. Determinant = 6.063E-7								

TABLE 4
Communalities of the Breast Cancer Tumor Characteristics

Communalities		
	Initial	Extraction
radius mean	1.000	0.952
texture mean	1.000	0.262
perimeter mean	1.000	0.963
area mean	1.000	0.946

smoothness mean	1.000	0.842
compactness mean	1.000	0.883
concavity mean	1.000	0.875
Extraction Method: Principal Component Analysis.		

TABLE 5
Total Variance Explained by the Components

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.328	61.822	61.822	4.328	61.822	61.822	3.371	48.152	48.152
2	1.395	19.926	81.748	1.395	19.926	81.748	2.352	33.596	81.748
3	0.850	12.136	93.885						
4	0.327	4.672	98.557						
5	0.086	1.234	99.791						
6	0.014	0.205	99.996						
7	0.000	0.004	100.000						
Extraction Method: Principal Component Analysis.									

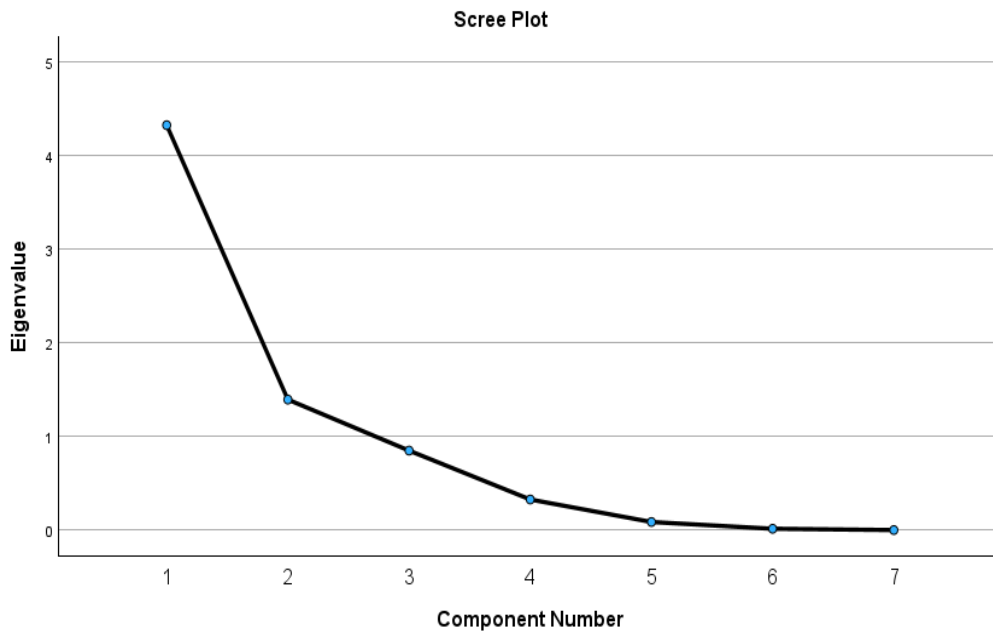


FIG1: Scree Plot of Breast Cancer Tumor Characteristics

3.8 Factor Loading

The factor loading is presented in Table 6, from where it can clearly be seen that tumor perimeter mean has a strong positive loading (0.937) on the first component and a negative loading (-0.292) on the second component. Similarly, tumor radius mean has a high positive loading (0.915) on the first component and a negative loading (-0.338) on the second component and so on. From the analysis with $p = 7$ and $m = 2$, the resulting factor model is given as:

$$\begin{aligned} y_1 - \mu_1 &= 0.945f_1 + 0.245f_2 + \varepsilon_1 \\ y_2 - \mu_2 &= 0.940f_1 + 0.249f_2 + \varepsilon_2 \\ y_3 - \mu_3 &= 0.936f_1 + 0.296f_2 + \varepsilon_3 \\ y_4 - \mu_4 &= 0.511f_1 + \varepsilon_4 \\ y_5 - \mu_5 &= 0.913f_2 + \varepsilon_5 \\ y_6 - \mu_6 &= 0.352f_1 + 0.872f_2 + \varepsilon_6 \\ y_7 - \mu_7 &= 0.571f_1 + 0.741f_2 + \varepsilon_7 \end{aligned}$$

where, $y_1, y_2, y_3, \dots, y_7$ are the radius, texture, perimeter, area, smoothness, compactness, and concavity means respectively and f_1 and f_2 are respectively the tumor size tumor shape.

The rotated component matrix provided in Table 6 applies varimax rotation method to simplify the loadings and enhance interpretability.

Factor 1

Variables such as radius mean, perimeter mean, and area mean, have strong positive loadings (around 0.9), indicating that they are highly positively associated with Component 1. This

suggests that these variables tend to vary together, and their values contribute significantly to the variability explained by Component 1. Variables such as compactness mean and smoothness mean on the other hand have very low to no loading on Component 1, which suggests a weak relationship. Furthermore, tumor texture mean and concavity mean have moderate positive loadings on Component 1, indicating that even though they contribute to Component 1, the extent of the contribution is low compared to the variables with high positive loadings.

Factor 2

The Varimax rotation also indicate that tumor smoothness mean, compactness mean and concavity mean have strong positive loadings on Component 2 indicating their high positive association with Component 2. Tumor radius mean, area mean, and perimeter mean on the other hand have relatively weak positive loadings on Component 2.

In general, these results provide insights into the underlying structure of the data. Component 1 seems to capture a set of variables related to tumor size (radius, perimeter, and area), while Component 2 appears to be associated with characteristics related to tumor shape (smoothness mean, compactness mean, and concavity mean).

TABLE 6
Component Matrix and Rotated Component Matrix for Tumor Characteristics

Component Matrix a			Rotated Component Matrix a		
	Component			Component	
	1	2		1	2
perimeter mean	0.937	-0.292	radius mean	0.945	0.245
radius mean	0.915	-0.338	area mean	0.940	0.249
area mean	0.914	-0.333	perimeter mean	0.936	0.296
concavity mean	0.892	0.282	texture mean	0.511	
compactness mean	0.787	0.514	smoothness mean		0.913
texture mean	0.404	-0.315	compactness mean	0.352	0.872
smoothness mean	0.447	0.801	concavity mean	0.571	0.741
Extraction Method: Principal Component Analysis.			Extraction Method: Principal Component Analysis.		
a. 2 components extracted.			Rotation Method: Varimax with Kaiser Normalization.		
			a. Rotation converged in 3 iterations.		

3.9 Component Transformation Matrix

The component transformation matrix is presented in Table 7. The value 0.821 in the first row, first column (1, 1) indicates that the first initial component has a strong positive relationship with the first rotated component. The value 0.571 in the first row, second column (1, 2) shows that the first initial component also has a moderate positive relationship with the second rotated component. Similarly, the value -.571 in the second row, first column (2, 1) demonstrates that the second initial component has a moderate negative relationship with the first rotated component. The value .821 in the second row, second column (2, 2) suggests that the second initial component has a strong positive relationship with the second rotated component.

Varimax rotation with Kaiser normalization has been applied to enhance the clarity of the loadings structure, making it easier to comprehend the relationships between components and variables.

TABLE 7
Component Transformation Matrix for Tumor Characteristics

Component Transformation Matrix		
Component	1	2
1	0.821	0.571
2	-0.571	0.821
Extraction Method: Principal Component Analysis.		
Rotation Method: Varimax with Kaiser Normalization.		

3. Conclusion

In exploring patterns of breast cancer tumor characteristics namely radius mean, texture mean, perimeter mean, area mean, smoothness mean, compactness mean, and concavity mean using factor analysis, two major underlying factors that are involved in describing tumor characteristics based. The first component captured a set of variables related to tumor size namely radius mean, perimeter mean, and area mean while the second component is associated with characteristics related to tumor shape namely smoothness mean, compactness mean, and concavity mean.

References

- [1] Luo, Y., Wu, J., Lu, J., Xu, X., Long, W., Yan, G., Tang, M., Zou, L., Xu, D., Zhuo, P., Si, Q., & Zheng, X. (2020). Investigation of COVID-19-related symptoms based on factor analysis. *Annals of Palliative Medicine*, 9(4), 1851-1858.doi:10.21037/apm-20-1113 <https://apm.amegroups.org/article/view/45369/html>
- [2] Tavakol, M., & Wetzel, A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. *International Journal of Medical Education*, 11, 245-247.
- [3] World Health Organization Annual Report (2015). World Health Organization, Ethiopia. <https://www.afro.who.int/news/global-effort-raise-awareness-breast-cancer-october-has-been-designated-pink-month-pink-month>

- [4] World Cancer Research Fund International (2022). *Breast cancer statistics*. World Health Organization/International Agency for Research on Cancer. <https://www.wcrf.org/cancer-trends/breast-cancer-statistics/>
- [5] Wu, A. H., Tseng, C., Vigen, C., Yu, Y., Cozen, W., Garcia, A. A., & Spicer, D. (2020). Gut microbiome associations with breast cancer risk factors and tumor characteristics: a pilot study. *Breast Cancer Research and Treatment*, 182(2), 451–463. <https://doi.org/10.1007/s10549-020-05702-6>
- [6] Yanai H., & Ichikawa M. (2006). Factor Analysis. *Handbook of Statistics*, 26, 257-296. [https://doi.org/10.1016/S0169-7161\(06\)26009-7](https://doi.org/10.1016/S0169-7161(06)26009-7)

NCPS-2024