



KIRINYAGA UNIVERSITY

AFRICAN JOURNAL OF SCIENCE, TECHNOLOGY AND ENGINEERING (AJSTE)

Volume 4(2) 2024

KIRINYAGA UNIVERSITY, KENYA

ISO 9001: 2015 Certified

AFRICAN JOURNAL OF SCIENCE, TECHNOLOGY AND ENGINEERING (AJSTE)

Editorial Board

Editors

Prof. Charles Omwandho, PhD

Prof. Jane Kerubo, PhD

Dr. Jotham Wasike, PhD

Advisory Board

Prof. Mary Ndung'u, PhD

Prof. Charles Omwandho, PhD

Administration & Logistics

George Ngorobi

Nellyann Kathomi

© Copyright 2023, Kirinyaga University

Copyright Statement

All rights reserved. Seek the University permission to reproduce, distribute, display or make derivative content or modification.

AFRICAN JOURNAL OF SCIENCE, TECHNOLOGY AND ENGINEERING (AJSTE)

Email: journals@kyu.ac.ke

ISSN 1356-6282

**KIRINYAGA UNIVERSITY
P.O Box 10500
NAIROBI, KENYA**

Preamble

African Journal of Science, Technology and Engineering (AJSTE) is an academic peer reviewed publication that hosts original, innovative research and scholarly articles that contribute to growth of knowledge in Science, Technology, Engineering and related disciplines. The journal provides an interactive platform for researchers to showcase the latest innovations and developments in the disciplines. The articles featured in this issue showcase groundbreaking research dedicated to advancing our understanding of the complex scientific phenomena describing presence of heavy metals in petroleum refinery wastewater, a review on application program interface - based android malware detection, malaria vector control: challenges and future strategies, variation in genetic and chemical constituents of *strychnos henningsii* populations in Kenya, cloud data security audit reporting techniques using bat inspired algorithm and application of data analytics in management of SMEs' performance and sustainability in post-COVID-19 Kenya.

Table of Contents

<u>HEAVY METALS IN REFINERY WASTEWATER: ASSESSMENT AND TREATMENT USING DIATOMACEOUS EARTH</u>	8
<u>SYSTEMATIC LITERATURE REVIEW ON APPLICATION PROGRAM INTERFACE -BASED ANDROID MALWARE DETECTION</u>	17
<u>MALARIA VECTOR CONTROL: CHALLENGES AND FUTURE STRATEGIES</u>	32
<u>REVIEW ON VARIATION IN GENETIC AND CHEMICAL CONSTITUENTS OF <i>Strychnos henningsii</i> POPULATIONS IN KENYA</u>	43
<u>CLOUD DATA SECURITY AUDIT REPORT TECHNIQUES USING BAT INSPIRED ALGORITHM: A REVIEW</u>	65
<u>BIG DATA ANALYTICS FOR SMES' PERFORMANCE SUSTAINABILITY IN POST-COVID-19 KENYA</u>	81

HEAVY METALS IN REFINERY WASTEWATER: ASSESSMENT AND TREATMENT USING DIATOMACEOUS EARTH

OYUGI C. A¹., OHOWA B²., SHEE A³.

¹Kenya Petroleum Refineries Ltd, Kenya

²Kenya Marine and Fisheries Research Institute, Kenya

³ Technical University of Mombasa, Kenya

Correspondence: sheali@tum.ac.ke

Abstract

In refining crude oil, freshwater is applied in distillation, hydration, desalination, system cooling, firefighting and other cleaning operations. The Kenya Petroleum Refineries Limited consumes huge quantities of freshwater daily for various applications leading to generation of contaminated effluents. The effluents contain a cocktail of contaminants such as aliphatic hydrocarbons, polycyclic aromatic compounds, dissolved gases such as Hydrogen Sulphide like algae, fungi, and heavy metals. Except for heavy metals, the refinery wastewater treatment plant is able to remove most of the other contaminants from the wastewater through a combination of processes including filtration, coagulation, sedimentation, softening, de-aeration, chlorination, desulfurization, bioremediation, and ion-exchange. However, these techniques are expensive, generate huge amounts of sludge, and are also not effective for removal of trace levels of heavy metals. Using inductively coupled plasma coupled with optical emission spectroscopy (ICP-OES), the effluents were found to contain significant levels of Iron, Cadmium, Chromium, Manganese, Nickel, Lead, Vanadium and Zinc. Although the installed wastewater treatment could reduce heavy metals load to trace, effluents storage in the tank farm prior to disposal leads to accumulation of heavy metals over time. Diatomaceous earth (DE) was applied to treat the heavy metals from the tank farm to within allowable limits as recommended by the World Health Organisation. Hence DE should be applied to effluents exiting the wastewater treatment plant prior to storage in the tank farm.

Keywords: Adsorption, Diatomaceous Earth, Heavy Metals, ICP-OES, Refinery Wastewater

Introduction

Crude oil refining consumes huge amounts of freshwater in various unit operations and unit processes such as desalting, distillation and thermal and catalytic cracking. For every volume of crude oil processed, the amount of wastewater produced is 0.4-1.6 (Ishak et al., 2012) times as much. Untreated petroleum refining wastewater (PRW) contains a cocktail of contaminants including hydrocarbons (benzene, toluene, ethyl benzene, xylenes, and polycyclic aromatic hydrocarbons), phenols, dissolved minerals and heavy metals. Most of these contaminants are toxic and possible carcinogens (Zarooni and Elshorbagy, 2006; El-Naas et al., 2014). Therefore, before discharge into the environment or municipal sewerage system (MSS), PRW needs to be treated to reduce contaminant load to acceptable levels. The Kenya Petroleum Refinery Limited (KPRL) located in Mombasa-Kenya is responsible for processing and storage of petroleum products. It serves Kenya and the wider East African region. Freshwater at KPRL is supplied by the Mombasa Water and Sewerage Company Limited (MWSC). In addition to petroleum processing operations, freshwater at KPRL is also used in firefighting and general cleaning operations. In order to comply with set environmental regulations, the PRW is first treated before discharge into MSS for further treatment at the Kipevu Sewage Treatment Plant (KSTP), Mombasa. The KPRL wastewater treatment plant employs a combination of physical and chemical processes including filtration, coagulation, sedimentation, softening, de-aeration, chlorination, desulfurization, bioremediation, and ion-exchange. These processes are expensive and require specialized skills for operation. Notably, low concentrations of heavy metals are not easily removed by most of these processes.

Adsorption of contaminants onto activated carbons, zeolites, resins, polymers and natural biomass is a cost-effective and easy to operate process. Adsorption process involves transfer of contaminants from the bulk aqueous phase to the adsorbent surface or into the adsorbent pores. Adsorption is essential in water treatment since it transfers a contaminant from a dilute solution to a small concentrated volume which can be subjected to further chemical treatment.

Although activated carbon is a widely used adsorbent in water treatment applications due to its high porosity and surface area ($> 1000 \text{ m}^2/\text{g}$), it is expensive and its regeneration results in significant loss (20-30%). Therefore, cheaper adsorbents derived from locally available materials are preferred alternatives especially in low and middle income countries such as Kenya.

Diatomaceous earth (DE) is a natural adsorbent that has been applied in wastewater treatment applications. Fine DE is amorphous mineral clay with high porosity and permeability, low thermal conductivity and density and moderate surface area ($10\text{-}30 \text{ m}^2/\text{g}$). DE mined from Kariandusi, Kenya contains mainly Silicon dioxide (SiO_2 ($> 70\%$)) and 10% Aluminum Oxide (Al_2O_3 (9-11%)) (Jemutai-Kimosop et al. 2020). DE is cheap and readily available and has its Natural and modified forms been extensively used as adsorbent for water treatment in removal of fluoride (Simiyu et al. 2023). Equally, carbamazepine (Jemutai-Kimosop et al. 2020) and E. coli and rotavirus (Simiyu et al. 2023) have been used. In the current work, DE will be evaluated for heavy metals removal in PRW obtained at the KPRL.

Materials and methods

Sample preparation

The DE was obtained from African Diatom Industries Limited (ADIL) based at Kariandusi, Kenya. It was cleaned, dried at 120°C for 24 h and ball milled to fine powder of $< 300 \mu\text{m}$ mesh size. 200 g of the powder was dispersed in 500 mL distilled water. In order to remove dissolved ions, the pH of the suspension was adjusted to 11 using 0.1 NaOH and agitated on an orbital shaker at 300 rpm for 1 h. The mixture was then centrifuged at 5000 rpm to obtain the solid residue which was dried at 120°C for 24 h. The dry solid was suspended in 500 mL aqueous solution acidified with 0.1 M H_2SO_4 at pH 1. The solid was recovered by centrifugation and washed with distilled water until the supernatant pH was 6.5. The clean DE was dried at 120°C for 24 h and stored in an air-tight container for further use.

Characterization of the adsorbent

To determine pH, 5 g DE in 50 mL of 1 M KCl was agitated on an orbital shaker at 180 rpm for 24 h. 10 mL of the extract was used for pH measurement using a pH meter S400 (Mettler-Toledo). The point-of-zero charge (pH_{pzc}) of DE was determined by alkalimetric method (Hosseinzadeh and Mohammadi, 2015).

Heavy metals analysis from KPRL wastewater

Sampling glass bottles (1 L) were washed with detergent in running tap water, rinsed with distilled to eliminate any traces of contaminants. Further cleaning of the sampling bottles was done using aqua regia (mixture of concentrated HCl and concentrated HNO₃ in 3:1 volume to volume) and rinsed with double distilled water. Composite sampling of 1 L of PRW was obtained from 7 interceptors over a period of three months (June to August, 2014) namely the main interceptor, blending area interceptor, tank farm interceptor, input tank age interceptor, interceptor 2, DAF 2781 interceptor of the wastewater treatment plant and raw water (municipal freshwater supplied by MWSC) interceptor. Immediately after collection, 2 mL of concentrated HNO₃ (65%) was added to preserve the metals in their metallic form and minimize their precipitation.

For sample dissolution and digestion, 100 mL of water sample was mixed with 10 mL of aqua regia, 1 mL of perchloric acid and incubated at 80 °C in a water bath. After acid-digestion, the sample was mixed with 50 mL double distilled water. Heavy metal content was analyzed using inductively-coupled plasma-optical emission spectrometry (ICP-OES).

Continuous adsorption experiments

All experiments were carried out in triplicate under identical conditions. Composite sampling was performed to collect wastewater for filtration by DE. A 1 m column with internal diameter of 3.5 cm was filled with 400 mg DE to a bed height of about 75 cm. The composited PRW samples were filled to the 1 m mark of the column and allowed to flow through the DE bed filter under gravity. Eluates from the column were collected periodically and subjected

to ICP-OES for heavy metal analysis. Sorption capacity at bed breakthrough point was calculated using equation 1:

$$q_e = \frac{(C_0 - C_e)V}{m} \quad ($$

Where q_e refers to sorption capacity at breakthrough (mg/kg), C_0 and C_e are the initial and equilibrium metal concentrations (mg/L), V is the volume of PRW (L) and m is the mass of adsorbent (kg).

Analytical techniques

Multi-element standards of analytical grade were applied for preparation of a series of standard solutions. Using ICP-OES (PerkinElmer Optima 8000), emission readings of heavy metals in standard samples were determined and calibration curves prepared. The readings were determined at emission wavelengths of 226.5 nm, 220.4 nm, 213.9 nm, 267.7 nm, 239.6 nm, 267.6 nm, 231.6 nm and 292.4 nm for cadmium (Cd), lead (Pb), zinc (Zn), chromium (Cr), iron (Fe), manganese (Mn), nickel (Ni), and vanadium (V), respectively. For determination of heavy metals in refinery wastewater, 25 mL of the PRW sample was mixed with 2% HNO_3 in distilled water and 4 mL aliquot applied for ICP-OES measurement. Control samples were prepared by mixing only 25 mL distilled water and 2% HNO_3 .

Results and discussion

Composited samples from seven interceptors (sampling points) at KPRL were collected and analyzed for heavy metals. Eight heavy metals namely Cd, Pb, Zn, Cr, Fe, Mn, Ni and V were detected in the samples. The concentration levels of these metals at the different sampling points are presented in Figure 1.

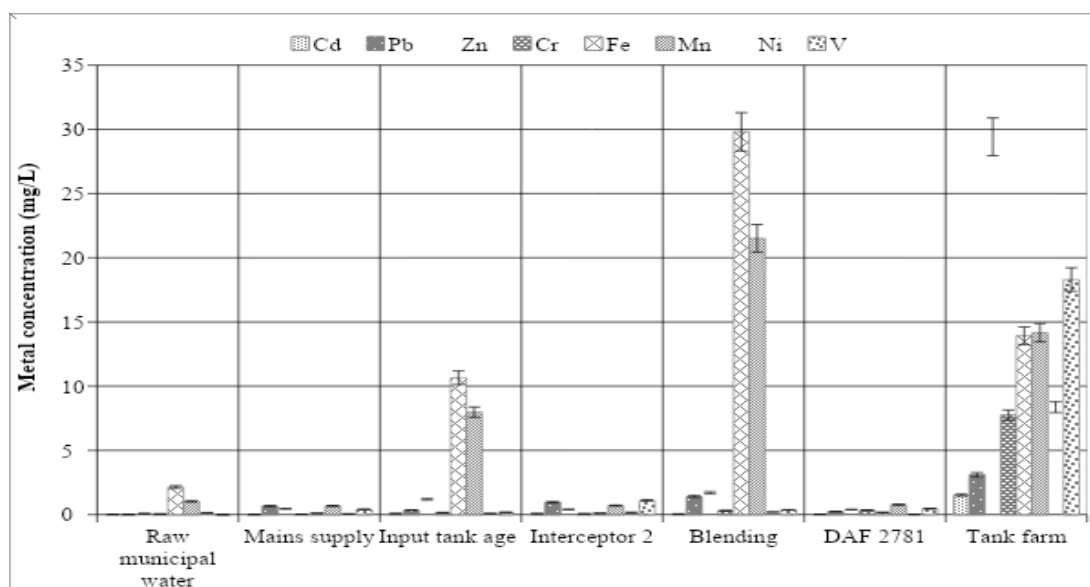


Figure 1: The levels of heavy metals in refinery wastewater at the Kenya Petroleum Refineries Limited (pH = 6.5; volume of composited sample = 1 L; number of replicate experiments = 3)

As can be seen in Figure 1, no significant levels of heavy metals were found in raw municipal water and the storage mains supply. The main source of freshwater at KPRL is supplied by MWSC. The input tank age and interceptor 2 are channels for distributing crude oil from the supplier to the processing zone. No significant levels of heavy metals were found in interceptor 2. The crude storage input tank age had high amounts of Fe and Mn. The blending area showed even a higher amount of Fe and Mn ($\approx 28\text{-}30\text{ mg/L}$) with increased levels of other metals. The blending area is where all crude operations are carried out. This observation clearly indicates that heavy metals are derived from refinery operation rather than from freshwater supplied by MWSC. After processing, the effluents are passed through a wastewater treatment plant. The KPRL wastewater treatment plant uses a combination of treatment processes including filtration, coagulation, sedimentation, softening, de-aeration, chlorination, desulfurization, bioremediation, and ion-exchange. The DAF 2781 was selected as the sampling point for effluents exiting the wastewater treatment plant. As can be seen in Figure 1, nearly all the heavy metals are removed by the treatment plant to below 1 mg/L which is within the world health organization (WHO) limits. The tank farm is where all

treated effluents are stored before discharge to KSTP. The storage takes a couple of days or weeks. Although the wastewater treatment plant is able to reduce heavy metal load to recommended WHO limits, it can be seen in Figure 1 that heavy metals load in tank farms is significantly high. This could be due to trace heavy metals accumulation over time. The DAF 2781 shows trace levels of heavy metals. Since the effluents are not immediately released from the tank farm, trace levels of heavy metals leaving the wastewater treatment plant accumulate over time.

In order to treat heavy metals load in refinery wastewater, composited samples from the tank farm were passed through DE in column studies. The wastewater flow through the DE bed filter in the column was done under gravity flow. The results are shown in Figure 2.

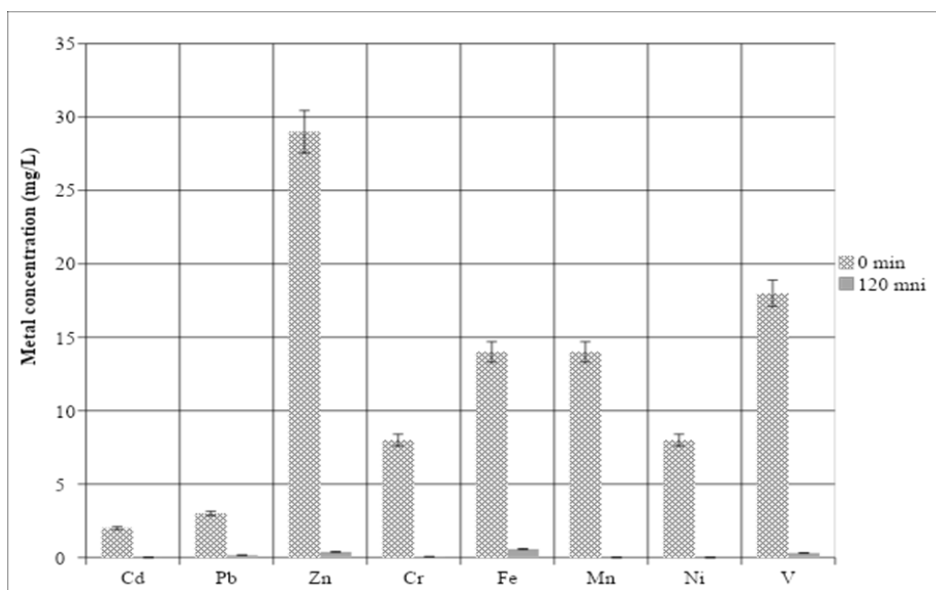


Figure 2: Removal of heavy metals from refinery wastewater by a fixed bed of diatomaceous earth under gravity flow (1 m glass column with 3.5 cm internal diameter; pH = 6.5; DE bed height = 75 cm; DE mass = 400 mg)

Figure 2 shows heavy metal removal by DE bed filter under gravity flow. A contact of time of about 120 min was essential for heavy metals removal by DE under the experimental conditions. This shows that heavy metals could easily be removed by adsorption onto DE. DE is a microporous clay mineral rich in Aluminum and Silicon with moderate surface area

(20-30 m²/g) (Jemutai-Kimosop et al., 2020). Since the point of zero charge of DE was found to be 3.5, it implies that under the experimental conditions (pH = 6.5), negative charged species are responsible for adsorption of heavy metals. The adsorption process is due to electrostatic forces between the positively charged metal ions and negative charged surface groups on DE.

Conclusions

Eight types of heavy metals namely Cadmium (Cd), Lead (Pb), Zinc (Zn), Chromium (Cr), Iron (Fe), Manganese (Mn), Nickel (Ni) and Vanadium (V) were detected and quantified in refinery wastewater at KPRL. Among all the metals, Fe and Mn had the highest concentrations. The heavy metals were found to have resulted from refinery operations and not from freshwater supplied by the MWSC. The installed wastewater treatment plant at the refinery is able to reduce heavy metals load to within WHO recommended limits. However, storage of treated wastewater at the tank farm in the refinery leads to accumulation of trace metals to higher levels with time. Thus, DE bed filters should be applied to effluents exiting the wastewater treatment plant prior to storage in the tank farm.

Acknowledgement

The authors are grateful to KPRL for their permission to carry out this work.

References

- El-Naas M.H., Alhaija M.A., Al-Zuhair, S. (2014). Evaluation of a three-step process for the treatment of petroleum refinery wastewater. *J. Environ. Chem. Eng.* 2, 56-62. <https://doi.org/10.1016/j.jece.2013.11.024>
- Hosseinzadeh, H., Mohammadi, S., (2015). Quince seed mucilage magnetic nanocomposites as novel bioadsorbents for efficient removal of cationic dyes from aqueous solutions. *Carbohydr. Polym.* 10(134): 213-221.
- Ishak, S., Malakahmad, A., Isa, M.H., (2012). Refinery wastewater biological treatment: A short review. *J. Sci. Ind. Res.* 71: 251-256.
- Jemutai-Kimosop, S., Orata, F., Shikuku, V.O., Okello, V.A., Getenga, Z.M., (2020). Insights on adsorption of carbamazepine onto iron oxide modified diatomaceous earth: Kinetics,

isotherms, thermodynamics, and mechanisms. Environ. Res. 180, 108898.
<https://doi.org/10.1016/j.envres.2019.108898>

Simiyu, M.T., Nyongesa, F.W., Aduda, B.O., Birech, Z., Mwebaze, G.A., Njenga, L., Ilin, N., Krauss, T.F., Pitruzzello, G. (2023). Application of molasses in improving water purification efficiency of diatomaceous earth waste ceramic membranes. MRS Adv. 8: 538-544
<https://doi.org/10.1557/s43580-023-00537-x>

Zarooni A.L., Elshorbagy M.W., (2006). Characterization and assessment of Al Ruwais refinery wastewater. J. Hazard. Mater. 136(3):398–405.
<https://doi.org/10.1016/j.jhazmat.2005.09.060>

SYSTEMATIC LITERATURE REVIEW ON APPLICATION PROGRAM INTERFACE - BASED ANDROID MALWARE DETECTION

EHODA E.,¹ ADEBAYO O.S.,² ISMAILA I.,³ OJENIYI J.A.,⁴ OLALERE M⁵

^{1,2,3,4} Cyber Security Science Department,

Federal University of Technology, Niger State Nigeria.

⁵Islamic University, Uganda

Correspondence : eehoda@gmail.com

Abstract.

Over the years, various malware detection approaches have been proposed in a bid to address evolving malware threats landscape in android operating system. Systematic literature reviews to analyze these detection approaches have been carried out, but none have been tailored to identifying challenges with android malware detection based on the use of Android program interface (API) features, hence there is no aggregated information on what work has been done by researchers in this area. This research, therefore, presents a systematic literature review on API feature based android malware detection literatures between 2018 to 2022 collected systematically using PRISMA frameworks. This study seeks to identify the challenges faced in android malware detection over the years, methodologies used to address them and limitations of API based feature detection. These useful insights documented in this research will serve as valuable resources which researchers can leverage on to improve the detection of android malware.

Keywords: Android Platform, Malware Detection, Application Program Interface, PRISMA Framework

Introduction

The number of smartphone users in the world has grown from 4.435 billion to 6.648 billion from 2017 to 2022 and this number constitutes 90.72% of the world's population according to a report by Statista. The large percentage of mobile phone users motivates attackers to target mobile phone platforms, predominantly android operating systems with malwares as more persons are interconnected and exposed to the threats. With more people exposed to the threat and proliferation of various applications in the android platform, comes the increased burden to protect users' devices against malicious applications and attacks. Over the years, a number of malware detection approaches have been proposed in a bid to address evolving malware threats landscape in android operating system. Researchers have employed the use of static features such as permission and strings, however this approach contends with the challenge of code obfuscation and other malware evasion techniques. Systematic literature reviews to analyze these detection approaches have been carried out by some researchers, but none have been tailored to android malware detection based on the use of API feature, hence there is no aggregated information on what work has been done by researchers in this area. This research, in a bid to identify the recent challenges associated with the detection of malicious applications on the android devices, carried out a systematic review of the existing detection strategies using API features. API feature based android malware detection papers between 2018 to 2022 were collected systematically using PRISMA frameworks and challenges faced in android malware detection over the years, methodologies used to address them and limitations API based feature detection is subject to have been identified and documented by this research. Also type of analysis employed and datasets used by the researchers as well as performance reported by the papers were highlighted. The remaining parts of the research are organized as follows: the related work is discussed in the section 2

of this work while the section 3 is used to discuss the methodology adopted to carry out the review. Section 4 discusses the analysis and result and section 5 presents the conclusion.

Related Work

Several works have been done in the area of android malware analysis and detection but more of the works were based on static features. Application Programming Interface is one useful feature that can define the behavior of applications therefore considering the limitation of permission based static features; malware researchers have drifted to approaches that use the behavioral pattern of applications in their quest to improve android malware detection mechanism. There are systematic literature reviews that have been conducted by different researchers on android malware detection but we have not come across any that examined android malware detection based on the use of API features. There is therefore no aggregated information on what work has been done by researchers in the area of API feature based analysis and detection. (Ashawa and Morris, 2019) carried out a systematic review of the malware detection techniques used for android devices. The review highlighted strengths and limitations of various detection techniques but not much was said on techniques using API based features. Similarly, (Ehsan et al. 2022) conducted a systematic literature review on android platform, analyzing articles focused on permission analysis for malware detection. On the other hand, Ya et al (2020) examined static analysis techniques for malware detection. They categorized static analysis into methods that are opcode based, program graph based, symbolic execution based and android characteristics based. They observed that static analysis methods are effective but are however faced with some challenges that needed to be addressed to improve android malware detection. Other SLRs dealt with the subject of android malware detection from a general perspective, hence our motivation to carry out a systematic literature review narrowed down to techniques that employ use of API based features.

Research Methodology

This section shows the methodology used in carrying out the systematic literature review. The steps followed are presented using figure1.

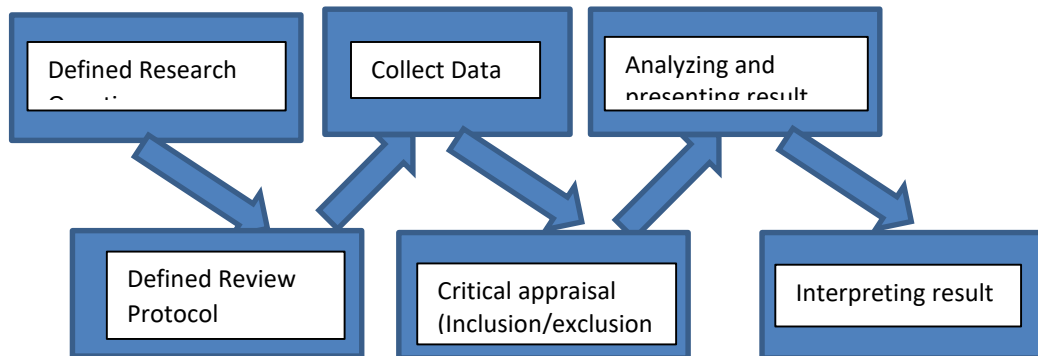


Figure 1. Systematic Literature Review Steps

Review Protocol

PRISMA framework was used to carry out this review. PRISMA stands for Preferred Reporting Items for Systematic Reviews and Meta-Analysis. It is an evidence-based minimum set of items for reporting systematic reviews and meta- analysis.

Data Collection

Articles searched and used for this review were systematically collected using the PRISMA framework.

a) Identification Stage

Material used for review were identified using search keywords and literature identified from IEEE, SCIENCEDIRECT and Google Scholar. The breakdown is as shown below;

A = Articles identified from IEEE - 98

B = Articles identified from Science Direct - 65

ADs = Articles identified with Google Scholar – 50

Table 1: Review Data Collection Parameters

Criteria	
Search Keyword	“Android Malware Detection” AND “API Call” was used to search for the review articles and databases searched are Science Direct and IEEE

Inclusion	A1 = Journals and conference papers A2 = Papers that discussed API as feature A3 = Papers between 2018 - 2022
Exclusion	B1 = Papers later than 2018 B2 = Paper not explicitly related to android malware detection using API features B3 = duplicate copies indexed in other databases

b) Screening Stage

The papers collected were subjected to a screening procedure applying inclusion criteria A1, A2, A3 and exclusion criteria, B3 where 27 duplicates were removed. Furthermore, applying exclusion criteria B2, title and abstracts were reviewed to reduce the articles to the subject area in focus. This reduced the papers collected for review to 54. The breakdown from each database is as shown in table 2. below

Table 2: Papers Collected from Database

DATABASE	Number of Papers
IEEE	25
SCIENCE DIRECT	18
GOOGLE SCHOLAR (Springer, ACM, Research gate)	11
TOTAL	54

c) Eligibility Stage

Quality of papers collected was assessed to ensure they are useful for the research. Duplicates were removed, abstract were thoroughly reviewed in line with the research focus and no paper was removed as a total of 54 papers were retained.

d) Inclusion Stage

Based on the inclusion and exclusion parameter earlier defined, all 54 papers screened were used for the review.

Critical Appraisal (Inclusion/exclusion)

The processes carried out between screening stage and eligibility stage defined in the PRISMA framework constitute the step of critical appraisal in the methodology. The papers returned after searching the database using the above keywords and inclusion/exclusion criteria were critically appraised. Papers within the scope of research bordering on API calls were kept while papers outside the scope which discussed other methods of detection other than use of API were screened. Furthermore, papers not related such as those that discussed cloud-based detection; windows OS or PE based and IoT based detection were screened leaving a total of 54 papers for synthesis in the review. Figure 2 shows the article search flow from the identification stage to articles included for the study

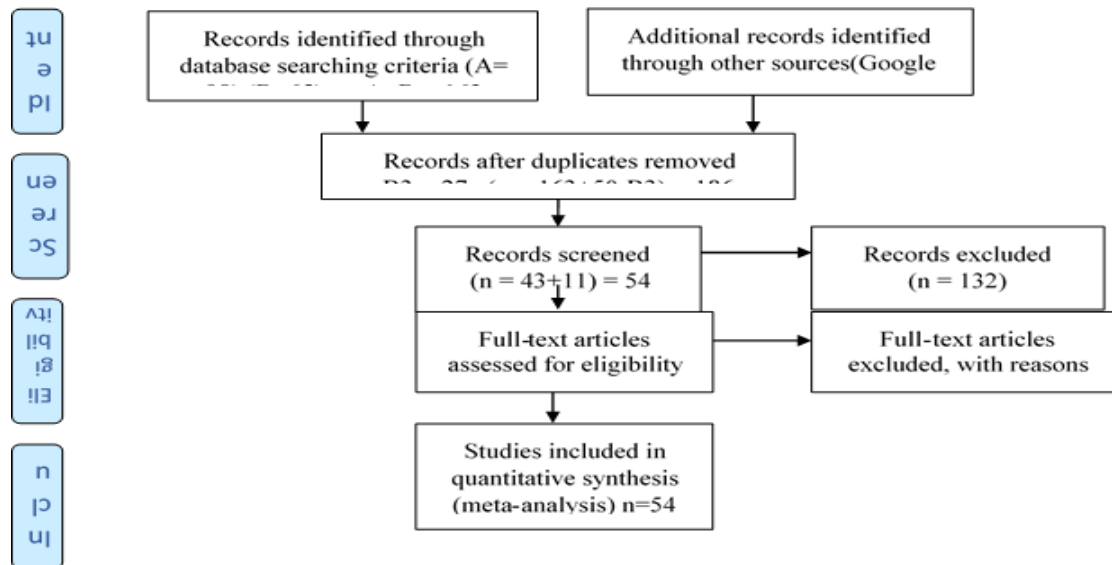


Figure 2: Prisma Flow Diagram

1. Analyses and Result Presentation

In this section, we provide answers to our research questions using papers studied. Figure 3 shows frequency distribution of papers from database sources, figure 4 shows frequency distribution of datasets used.

RQ1 - Are there challenges with Android Malware Detection approaches that prompted consideration for use of API features?

There have been a lot of efforts put into proffering solutions to malware attacks in android platforms by different researchers. However, these solutions have been fraught with myriad challenges which prompted researchers to consider the use of API call features with different models to address.

Wang et al (2020) observed that lack of descriptive distinctive feature of malware behavior and intent poses a challenge for android malware detection and proposed invocation of local sensitive API call using function call graph to address the challenge. Elsewhere, the work by Alzaylaee et al (2020) noted that employing static analysis using extracted features like API calls, commands and intent are prone to obfuscation where malicious code is concealed to prevent detection. Therefore, an approach to detection which considers extracting android permissions (static features) before execution of application and then extracting the API calls and Intents during execution (dynamic feature) was proposed using multilayer perceptron classifier, (MLP. Kumar and Ciza, 202) also alluded to the problem of code obfuscation with static analysis but employed a different approach to solving it. Obfuscation has also been mentioned in other reports (Lu, et al. 2019, Moutaz et al. 2020, Elayan and Mustafa 2021 and Michele et al, 2019). Other reports (Arindaam et al. 2020, Stuart et al. 2021) noted inability to detect zero-day malware as a challenge while (Roopak et al. 2020) cited the problem of multicollinearity and data overfitting in most classifiers used. The challenge of dynamic code loading was also stated by Elayan and Mustafa (2021) while Pang and Bian (2019) averred

that dynamic analysis is not efficient and malware detection can be limited due to execution time and code trigger condition. The research by Feng et al, (2020) also noted poor efficiency of dynamic analysis due to limitation in code coverage while (Kumar et al 2021) indicated that the problem of unbalanced dataset affected precision, recall, and F1 score values of the classifiers used and should be considered for investigation in the future.

RQ2 - Are there challenges with use of API in Android Malware Detection and can the challenges affect the detection?

Literature synthesized suggests that API features can be employed dynamically or statically with appropriate models to solve malware detection issues effectively. However, challenges also exist that can affect android malware detection where API features are used. Roopak et al, (2020) used conditional dependencies among relevant static and dynamic features (API calls, permissions and system calls) which are required for an app to work were used in a Tree Augmented Naive Bayes based hybrid malware detection mechanism and observed that few malicious software can evade the detection model by using adversarial techniques. Consequently, the researchers suggested future work for more powerful Bayesian models to be built for effectively identifying such adversarial malware applications by employing reinforcement learning techniques. Elsewhere, Michele et al, (2019) in their research relied on system API information to distinguish ransomware from other malicious and benign applications. Although inclusion of sample into the training set, meant the approach worked well against string obfuscation and heavy antistatic obfuscation done with class encryption, evasion is possible with this approach using semantically equivalent user implemented packages/classes/methods. Further to this, adversarial attack can also affect the outcome. According to Millar et al (2021), APIs usage as features requires a lot of feature-engineering and domain insight hence not effective in zero-day scenarios. Good accuracy and F1 score was achieved in work by Feng et al (2021) when they used graph neural networks to automatically capture critical information from call graphs rather than manual selection of

API calls, API call sequences and call traces. However, obfuscation techniques like packing, dynamic code loading and bytecode encryption could not be handled. Also, the approximate call graph used cannot capture reflection, implicit callback and implicit control flow and this could be exploited to evade detection.

In another research, Hadiprakoso et al (2020) declared that classical machine learning algorithms are dependent on feature engineering. This brings to the fore issues of expert domain knowledge needed for representation of the features as well as attackers' capability to evade detection once the features. Shen et al (2019) argued that evolution in modern malware has made reliance on simple information flow ineffective because modern malware performs complex computations before, during, and after collecting sensitive information and also, benign applications now use the same information that malicious applications gather. By performing N-gram analysis on sequences of API calls that occur along Complex-Flows' control flow paths to identify unique and common behavioral patterns present in Complex-Flows they developed a new mobile malware detection technique based on information flow.

Call Graphs gives information about API calls and shows relationship between methods in applications. However, Feng et al (2020) revealed that use of precise call graph consumes resources and results in poor efficiency. The work by Yang et al, (2021) also raised the same concern of resource consumption. They noted that thousands of APIs are provided by android platforms, therefore analysis of all function call graphs would consume large resources. Wang et al, (2022) on their part stated that API call sequences are usually too long, therefore a truncated segment of the API call sequences or its statistical features in malware detection was used by some researchers but it suffers from high false alarm because execution order information of the applications are lost. From the foregoing it can be seen that code coverage limitation, trigger conditions, dependence on expert knowledge for

feature selection, high volume of resource consumption, imbalance dataset among others constitute challenges that can affect malware detection where API call features are employed.

RQ3- Are there appropriate methods available to address the challenges of android malware detection using API features?

Various researchers have proposed solutions to address perceived challenges in android malware detection. Alzaylaee et al, (2020) proposed DL-Droid, which used dynamic stateful input generation to enhance code coverage. DL-Droid also focused on dealing with code obfuscation and employed real devices to avoid anti-emulator tendencies of malware. Their approach to detection considered extracting static features of android permissions before program execution, and then extracting the API calls and intents dynamically during execution using Multilayer Perceptron classifier, MLP. Kumar and Ciza, (2021) in their research also attempted to deal with the problem of code obfuscation with static analysis. The authors identified suspicious API classes and methods used by Malware apps and generated MSA (Multiple Sequence Alignment) corresponding to API class sequences present in malware applications to overcome malware evasion techniques using machine learning classifier Profile Hidden Markov Model(PHMM).

Elsewhere, Arindam et al, (2020) proposed a light weight detection framework that operates upon only 50 features. The method adopted for their research analyzed API calls extracted from smali code, maps the API Calls to certain features (permission) and constructed a frequency-based feature vector for each application. The approach bridges the existing gap of high need of resources such as time, space and computational power in existing work. Similarly, a malware detection system, MAPAS which learns behaviors of malwares by using a deep learning algorithm (CNN) and detects malware based on common patterns of API call graphs of malware was proposed to effectively deal with issues of high computing resources (Kim et al, 2022). Stuart et al, (2021) on the other hand addressed issues of expert domain knowledge required for feature engineering and attendant inability to detect zero

day scenarios. They proposed a solution that used Convolutional Neural Network CNN to learn from a limited set of only 210 proprietary Android API packages that have no expert pre-categorization as sensitive or otherwise.

Exploiting the advantages of deep learning to address featuring engineering in classical machine learning, (Hadiprakoso et al 2020) designed a new system that compiles static and dynamic analysis features such as API call sequence, system command, manifest permission, intent and process the data using a deep neural network. (Wang et al, 2022a) introduced an efficient extraction algorithm for API call sequences, which contains two sub-algorithms. The first sub-algorithm simplifies the function call graph from a multigraph to a simple graph, and the second develops a pruning depth-first search. The authors posit that the existing API call sequence extraction methods are laborious and time- consuming, which seriously decreases the efficiency of static analysis, hence the need to adopt the methodology.

Wang et al, (2022b) addressed the challenge of high false alarm caused by use of statistical features or truncated segments of API call sequences with their proposed FGL_Droid. FGL_Droid converts dynamic API call sequence into a function call graph, joins the function call graph feature and extracted permission request feature to carry out malware detection. The function call graph retains most of the application execution order information with significantly reduced sequence size and missed behavior information during conversion is made up for with the advanced features of permission requests extracted.

RQ4 - Are the available detection methods effective?

The effectiveness of detection methods used in various studies was largely measured with the performance evaluation metrics Accuracy, Precision, Recall and F1 Score. The approaches to android malware detection employed had good results with the various metrics hence taken to be effective. However, these performances could be investigated further as the type, size and nature of the sample of dataset used could result in bias in performance.

RQ 5 - Are there datasets being used in the primary studies?

The datasets used by selected primary studies and their frequency is represented in the figure 3.

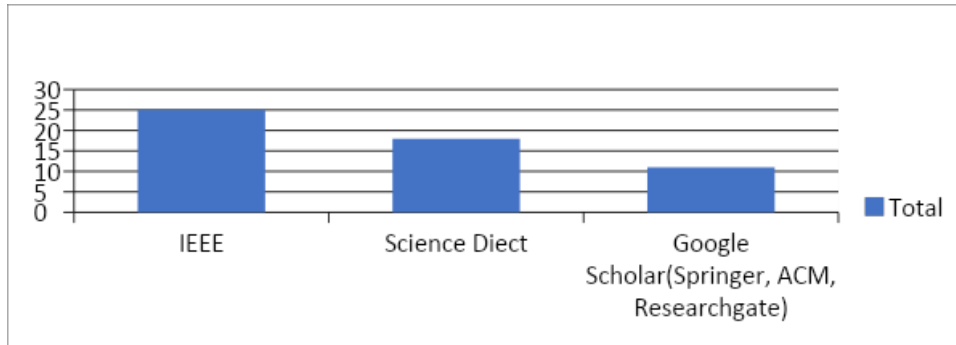


Figure 3: Frequency distribution of papers from Database Sources

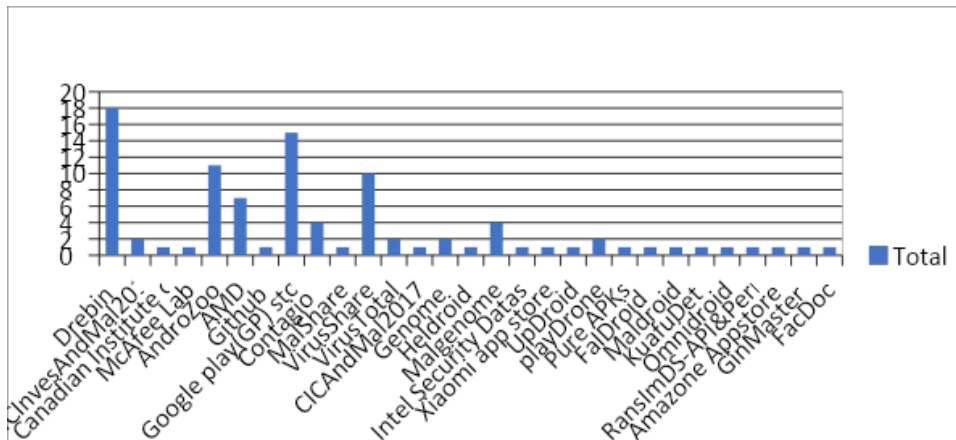


Figure 4: Frequency Distribution of Datasets used

5. Conclusion

This review examined literature on android malware detection with API call features. The study revealed that various researchers have employed the use of static analysis and dynamic analysis using API call features in detecting malicious android applications. The review also outlined different methodologies/algorithms with their performances and datasets used by authors in malware detection. Furthermore, challenges with android

malware detection in the research area which include code obfuscation, dynamic code loading issues with static analysis as well as limited code coverage, high resource consumption, execution time and trigger condition issues with dynamic analysis were identified among others. Various machine learning/deep learning methods and approach employed in the detection and analysis as revealed in the review provide useful insight researchers can leverage to improve android malware detection and is considered as a valuable contribution in this work.

References

- Alzaylaee M. K., Yerima S. Y., Sezer, S. (2020). "DynaLog; An automated dynamic Analysis Framework for characterizing android applications". Journal of Computers and Security. 89 <https://www.sciencedirect.com/science/article/pii/S0167404819300161>
- Arindam R., Singh J. S., Gitanjali J., Kapil S. (2020) "Android Malware Detection based on Vulnerable Feature Aggregation" in Procedia Computer Science. 173, 345-353.
- Ashawa M.A., Morris S. (2019). Analysis of android malware detection techniques systematicreview,dspace.lib.cranfield.ac.uk.8(3):177-187
<https://dspace.lib.cranfield.ac.uk/handle/1826/15057>
- Ehsan A., Catal C. Mishra A. (2022). Detecting Malware by Analyzing App Permissions on Android Platform: A Systematic Literature Review. Sensors, mdpi.com. 22(20): 7928
<https://www.mdpi.com/1424-8220/22/20/7928>
- Elayan N.O., Mustafa M. A. (2021). "Android Malware Detection Using Deep Learning" in Procedia Computer Science. 184:847 – 852,
<https://www.sciencedirect.com/science/article/pii/S1877050921007481>.
- Feng P., Ma J., Li T., Ma X., Xi N., Lu D. (2020). "Android Malware Detection Based on Call Graph via Graph Neural Network".
<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9353764>.
- Feng P., Ma J., Li T., Ma X., Xi N., Lu D. (2021). Android Malware Detection via Graph Representation Learning. Mobile Information Systems, hindawi.com,
<https://www.hindawi.com/journals/misy/2021/5538841/>

Hadi Prakoso R. B., Buana I. K., Pramadi Y. R. (2020). "Android Malware Detection Using Hybrid-Based Analysis and Deep Neural Network" <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9332066>.

Kim J., Ban Y., Ko E., Cho H., Yi J.H. (2022). MAPAS: a practical deep learning-based android malware detection system. *International Journal of Information*, Springer, <https://doi.org/10.1007/s10207-022-00579-6>

Kumar S., Ciza T. (2021). "ProDroid; An Android malware detection framework based on profile hidden Markov model" in *Pervasive and Mobile Computing* 72. <https://www.sciencedirect.com/science/article/pii/S1574119221000109>.

Kumar S., Mishra D., Shukla S. K. (2021). "Android Malware Family Classification: What Works API Calls, Permissions or API Packages" <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9699322>

Lu X., Jiang F., Zhou X., Yi S., Sha J., Pietro L. (2019). "ASSCA: API sequence and statistics features combined architecture for malware detection" in *Computer Networks*, Vol 157, 99-111. <https://www.sciencedirect.com/science/article/pii/S138912861930461X>

Michele S., Davide M., Francesco M, Aaron V. C., Fabio M., Giorgio G. (2019). "On the effectiveness of system API-related information for Android ransomware detection" in *Computers and Security*. 86:168-182, <https://www.sciencedirect.com/science/article/pii/S0167404819301178>

Moutaz A., Mamoun A., Andrii S., Abdel Wadood M., Albara A. (2020). "Intelligent mobile malware detection using permission requests and API calls" in *Future Generation Computer Systems*. 107:509-521, <https://www.sciencedirect.com/science/article/pii/S0167739X19321223>.

Pang J., Bian J. (2019). "Android Malware Detection Based on Naive Bayes", <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9040796>

Roopak S., Tony T., Sabu E. (2020). "A TAN based hybrid model for android malware detection" in *Journal of Information Security and Applications*. 54: 102483 - 102483 <https://www.sciencedirect.com/science/article/pii/S2214212618308263>.

Shen F., Vecchio J. D., Mohaisen A., Ko S Y., Ziarek L. (2019). "Android Malware Detection Using Complex-Flows" in *IEEE Transactions on Mobile Computing*. 18(6) <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8423084>

- Stuart M., Niall M., Jesus M., Paul M. (2021). "Multi-view deep learning for zero-day Android malware detection" in Journal of Information Security and Applications. 58(3) <https://www.sciencedirect.com/science/article/pii/S2214212620308577>
- Wang W., Wei J., Zhang S., Luo X. (2020). Malware Detection Based on Local Sensitive API Invocation Sequences", IEEE Transactions on Reliability.69(1)174-187, <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8776652>
- Wang W., Ren C., Song H., Zhang S., Liu P. (2022a). FGL_Droid: An Efficient Android Malware Detection Method Based on Hybrid Analysis. Security and Communications Networks, <https://www.hindawi.com/journals/scn/2022/8398591/>
- Wang T., Xu Y., Zhao X., Jiang Z., Li R. (2022b). Android malware detection via efficient application programming interface call sequences extraction and machine learning classifiers. IET Software, Wiley Online Library, <https://doi.org/10.1049/sfw2.12083>
- Yang Y., Du X., Yang Z., Liu X. (2021). Android malware detection based on structural features of the function call graph. Electronics, mdpi.com, <https://www.mdpi.com/961828>
- Ya P., Xiuting G., Chunrong F., Yong F. (2020). A Systematic Literature Review of Android Malware Detection Using Static Analysis. 8: 116363-116379 DOI 10.1109/ACCESS.2020.3002842, *IEEE Access*.

MALARIA VECTOR CONTROL: CHALLENGES AND FUTURE STRATEGIES

WANGAI, L. N., KAMAU, K. K., WAIRIMU, B.M., KAMAU, L. N., NJUGUNA, M. N.,
ALWORA, A.

Department of Health Sciences, Kirinyaga University, KENYA

Correspondence: kkamau@kyu.ac.ke

Abstract

The current demand for the eradication of malaria marks a new-fangled chapter in the antiquity of this illness. This has been brought about by the striking decreases in malaria caused by administration of efficient medications and vector control. However, the emergence of pesticide resistance poses a challenge to this approach. Alternative tools must be developed to continue supporting or potentially replace insecticide-based vector control methods. Long-lasting insecticidal nets (LLINs) and indoor residual spraying (IRS) continue to be the mainstays of the majority of National Malaria Control Programs in Africa, despite the large number of promising control tools tested against mosquitoes. These strategies are not enough to successfully control malaria. While these techniques are successful in lowering malaria incidence, their overall effectiveness in lowering malaria prevalence is often limited. Additionally, efficiency of LLINs and IRS is threatened by the rising rates of pesticide resistance in the targeted mosquito populations. Thus, although larvicidal treatments can be beneficial, using them in rural regions is not advised. To enhance mosquito vector control efforts and improve their quality and delivery, it is important to focus on integrated approaches. Successful malaria eradication requires close collaboration between parasitologists and entomologists, along with a comprehensive evaluation of epidemiological impact of innovative mosquito vector control strategies. This review discusses current malaria vector control strategies and highlights challenges, and promising tools that are expected to contribute to malaria eradication.

Keywords: Malaria, Vector Control, Current Challenges and Future Strategies.

Introduction

Approximately 619 000 people still die from malaria each year, and 247 million new cases are reported in 84 different countries (WHO 2022). A 33% decrease in mortality associated with malaria has been recorded in Africa. This has been attributed to preventive measures that have been scaled up in attempt to achieve the Millennium Development Goals (MDGs) of reduced maternal and child mortality. This represents a significant improvement over the past ten years. A definite pointer for MDG 6 is malaria, which also supports MDGs for education, maternal health, child survival, and poverty (Hemingway 2014). The majority of malaria prevention efforts focus on reducing the mosquito vectors to human contact by encouraging people to sleep under treated mosquito nets and to apply indoor residual spraying (IRS).

Roughly 780 million individuals in Africa who are at risk of malaria would need access to long-lasting insecticidal nets (LLINs), and roughly 150 million bed nets would need to be supplied annually. This estimate makes the assumption that the nets are still functional after five years on average. However, only 66 million nets had been distributed by 2015, failing to meet the aim. Additionally, mounting data indicate that LLINs are unlikely to last the full 5-year timeframe (Mejia et al 2013, Wills et al 2013). The LLIN market is very price sensitive, and poorer quality, less durable nets have dominated. Because of this, the anticipated 5-year lifespan of these nets is greatly shortened and they start to fall apart rapidly. In 2011, 11% of the world's population was protected from malaria thanks to the use of indoor residual spraying (IRS), up from 5% in 2005. 77 million of those persons were in Africa. The President's Malaria Initiative (PMI) in 19 target African countries, where 30.3 million individuals were shielded by the IRS in 2012, is largely responsible for this surge.

It is crucial to significantly expand funding for disease control initiatives and operational actions if we are to maintain the gains made in disease control and go closer to regional and

global eradication of malaria. Meeting the global malaria targets will require an estimated \$5.1 billion annually from 2011 to 2020. However, in 2011, only \$2.3 billion was available, less than half of the required amount. Recent years have seen some progress, with more pledges made for 2022, including over \$12 billion that has already been pledged in comparison to the \$15 billion required for 2014–2016 for the Global Fund PMI 2023.

Notably, it won't be enough to merely increase financial support, though. Thus, to preserve the long-term durability of the insecticides and medications used for prevention and treatment of malaria, improved stewardship of these resources is essential. Resistance will certainly reduce our ability to successfully prevent and manage malaria if we ignore this obligation. Because Malaria is an infectious disease spread by mosquitoes that affects millions of people worldwide, malaria vector control is essential in the fight against malaria transmission. There have been several methods used to manage malaria vectors, and research is always being done to create new ones. Here is a summary and review of the problems with and potential solutions for malaria vector control.

Present Vector Control Techniques

Insecticide-Treated Nets (ITNs): ITNs are among the widely used and effective vector control interventions. They provide a physical barrier between humans and mosquitoes while also delivering insecticide that kills or repels mosquitoes. ITNs have shown significant success in reducing malaria transmission, especially when used consistently and universally (Hill et al 2006, Lengler et al 2007).

IRS: Insecticides are sprayed on interior walls of homes to kill mosquitoes that come into contact with the treated surfaces. This strategy has been successful in limiting the spread of malaria, especially in regions with high vector densities and pesticide susceptibility (Protopopoff et al 2015, Okumu et al 2011). However, challenges such as insecticide resistance and logistical requirements limit its widespread implementation. Larval Source Management (LSM), targets mosquito breeding sites, such as stagnant water bodies, to

reduce mosquito populations. It involves interventions like draining, larviciding, or modifying habitats to disrupt mosquito breeding. LSM can be an effective complementary approach to ITNs and IRS, particularly in areas with specific breeding sites or where mosquitoes have developed resistance to insecticides.

Biological Control: This strategy encompasses introduction of natural predators, such as fish or predatory insects, to reduce mosquito populations. While it has shown some promise in certain settings, biological control methods require careful consideration to avoid unintended ecological consequences and disruption of local ecosystems (Beneli et al 2016, Dahmana and Mediannikov 2020). Advances in genetic engineering have opened up possibilities for genetically modifying mosquitoes to reduce their ability to transmit malaria. One notable example is the development of genetically modified mosquitoes that carry genes that suppress the mosquito population or make them resistant to the malaria parasite (Raghavendra et al, 2011). This approach shows promise but raises ethical, regulatory, and safety concerns that need to be addressed. Integrated Vector Management (IVM) sums it all up as it involves combining multiple vector control strategies in a coordinated manner, considering local epidemiological and ecological factors (Benelli and Beier 2017). By integrating various interventions, such as ITNs, IRS, LSM, and others, IVM aims to achieve sustainable and effective vector control (Figure 1).

Challenges and future strategies

Malaria vector control faces several challenges that can hinder its effectiveness in reducing transmission. Understanding and addressing these challenges is crucial for the success of vector control interventions.

Insecticides used in vector control interventions, such as pyrethroids, might cause mosquito resistance which may in turn reduce the effectiveness of indoor residual spraying (IRS) and insecticide-treated nets (ITNs). Resistance develops and spreads as a result of over-reliance on a small number of insecticides and insufficient rotation of other pesticide classes (Takken

and Kohls 2009, Corbel and N’Guessan 2013, Killen and Ranson 2018). Developing and deploying new insecticides, as well as implementing resistance management strategies, are necessary to combat this challenge.

Mosquitoes have consistently changed their behaviour in response to vector control interventions. For example, some mosquitoes may shift their feeding habits from indoors to outdoors or change their biting times to avoid contact with insecticide-treated surfaces. Such behavioural adaptations reduce the impact of interventions that primarily target indoor-biting mosquitoes, making it more challenging to control malaria transmission. Malaria is transmitted by various mosquito species, and their behaviour, ecology, and susceptibility to control interventions can vary. A one-size-fits-all approach may not effectively target all vector species in different regions (Raghavendra et al 2011, Massebo et al 2015). Understanding the local vector species composition and their behaviours is crucial for designing appropriate and targeted vector control strategies.

Achieving high coverage and sustained use of vector control interventions, such as ITNs and IRS, can be challenging in resource-limited settings. Thus, limited access to these interventions, financial constraints, and cultural factors can impede their widespread adoption. Ensuring equitable access to vector control tools and addressing barriers to their implementation is essential to achieving effective coverage Tizifa et al 2018. Vector control interventions require ongoing maintenance, monitoring, and replacement to remain effective. Challenges related to infrastructure, funding, supply chain management, and community engagement can hinder sustainability of interventions. Building robust systems for monitoring and evaluation, as well as strengthening health systems, is crucial to sustaining vector control efforts.

Climate change has greatly influenced the distribution, abundance, and behaviour of mosquito vectors. Rising temperatures, altered rainfall patterns, and changes in land use can impact mosquito populations and their ability to transmit malaria (Glunt et al 2013).

Adapting vector control strategies to account for climate change and considering environmental factors in intervention design is important for long-term effectiveness.

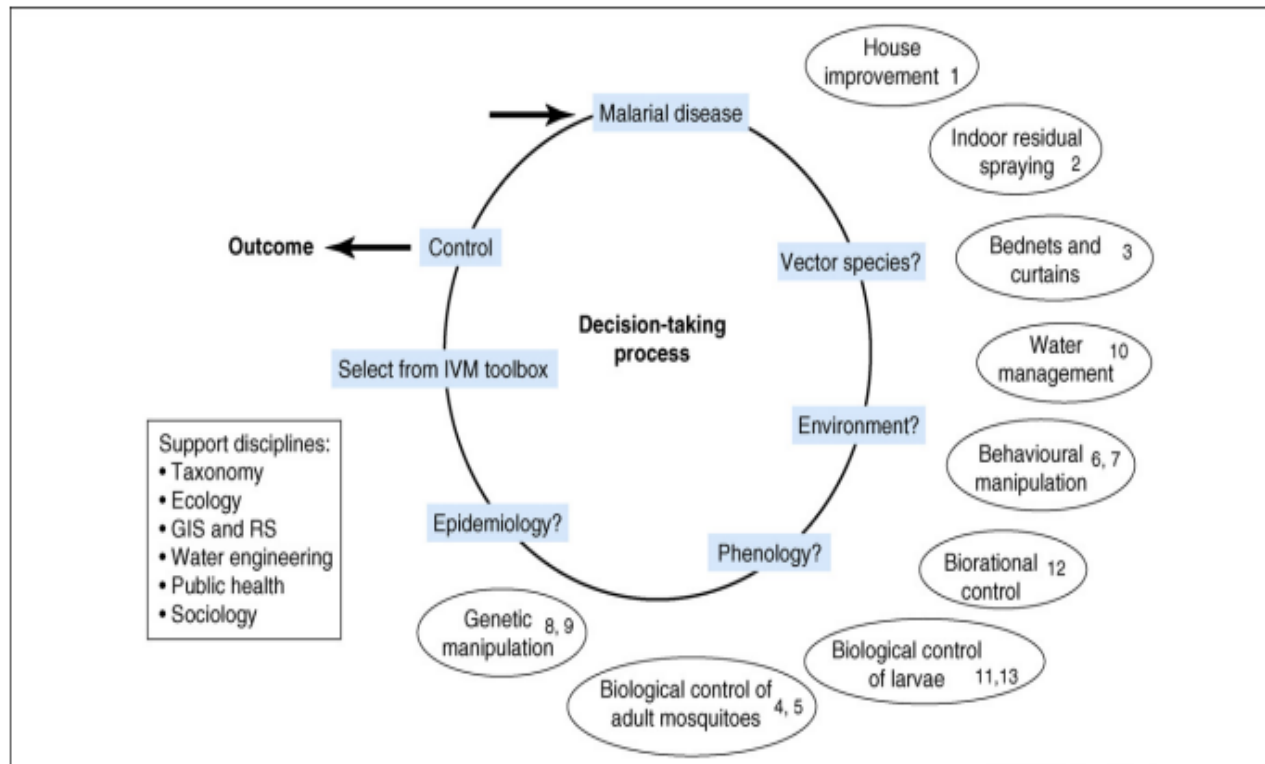


Figure 1. Integrated vector control interventions (Tekken and Knolls, 2009)

Future vector control strategies

Future strategies for malaria vector control are being explored and developed. These may include innovative repellents that deter mosquitoes from entering human living spaces. Approaches such as new insecticides and formulations, developing and deploying alternative insecticides and formulations that overcome insecticide resistance and have improved safety profiles (Figure 2) have been considered. Spatial repellents: Creating long-

lasting spatial conflict and instability. Malaria-endemic regions often face political instability, conflicts, and displacement of populations, which can disrupt or hinder vector control efforts. Inadequate infrastructure, limited resources, and population movements make it challenging to implement and sustain control interventions in these contexts. Coordinated efforts involving humanitarian agencies, governments, and international organizations are necessary to address vector control challenges in conflict-affected areas (Hemingway 2014). Addressing these challenges requires a multi-faceted and integrated approach, including research and innovation, strengthening health systems, community engagement, and collaboration among stakeholders. Continuous monitoring, evaluation, and adaptation of strategies based on local contexts are crucial to improving the effectiveness and sustainability of malaria vector control efforts.

Attractive Toxic Sugar Baits (ATSB): Using sugar baits laced with toxic substances to attract and kill mosquitoes outside human dwellings. ***Targeting outdoor transmission***: Developing interventions that specifically target outdoor biting and transmission, as some malaria vectors exhibit outdoor feeding behaviour. ***Novel vector control tools***: Exploring new technologies, such as trapping systems, genetic control methods, and use of unmanned aerial vehicles (drones) for targeted delivery of interventions (Benelli and Beier 2017, Tizifa et al 2018, Echodu et al, 2020).

It's paramount to note that the success and success of these future strategies may vary contingent on the local context, including vector species, insecticide resistance patterns, and community engagement (Ranson 2017, Nasir et al 2020, Barreaux et al 2017, Sougoufara et al 2017). Continued research, collaboration, and investment in malaria vector control are essential to develop and implement sustainable strategies that can reduce the burden of malaria globally.

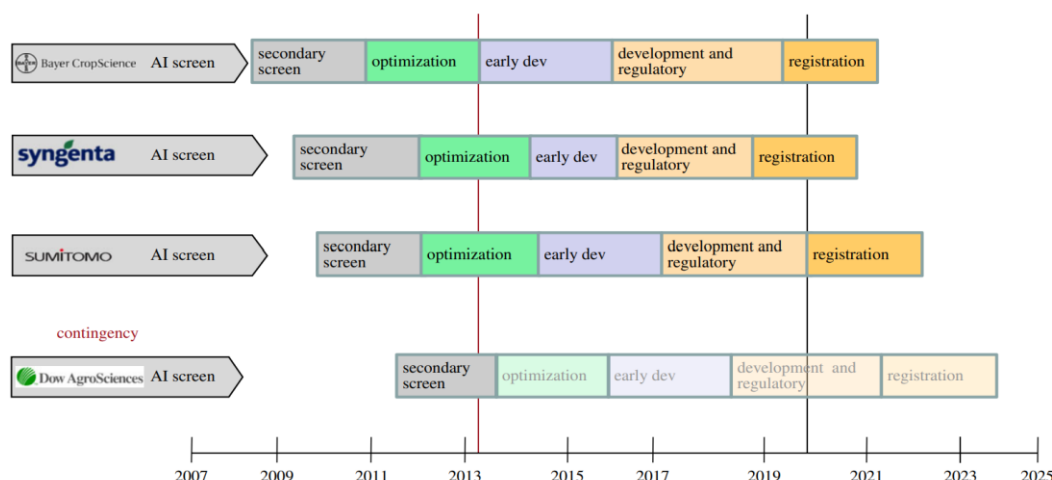


Figure 2. An overview of the market trends for new insecticides in the IVCC product line, together with the timelines for their introduction (Hemingway, 2014).

Conclusion

To enhance mosquito vector control efforts and improve their quality and delivery, it is important to focus on the following areas of research: Process and delivery approaches for vector control measures optimization. employing trusted methods to keep an eye on mosquito populations and biting activity. Creation of efficient and ecologically friendly instruments to lessen or possibly completely eradicate local cases of malaria and other diseases spread by mosquitoes thorough assessment of applicability and efficiency of new mosquito control methods in actual settings to show how they affect disease transmission. monitoring of environmental changes that might affect mosquito populations and spread of malaria. collaboration among numerous academic and research disciplines, including parasitology, tropical medicine, ecology, entomology, and ecotoxicology. This interdisciplinary approach must be kept up if we are to better understand the behavioral ecology of malaria vectors. The effectiveness of vector control strategies is hampered by a number of ecological issues. These include the diversity of vector species, interactions within

the mosquito food web and competition, variations in mosquito behavior, emergence of insecticide resistance, propensity of mosquitoes to avoid specific control methods, knowledge gap regarding mosquito dispersal and mating behavior, and the influence of environmental changes on mosquito traits.

References

- Barreaux P., Barreaux A.M., Sternberg E.D., Suh E., Waite J.L. (2017). Whitehead SA, et al. Priorities for broadening the malaria vector control tool kit. *Trends Parasitol.* 33(10):763–774.
- Benelli G., Jeffries C.L., Walker T. (2016). Biological control of mosquito vectors: past, present, and future. *Insects.* 7(4):52.
- Benelli G., Beier J.C. (2017). Current vector control challenges in the fight against malaria. *Acta Trop.* 174:91–96.
- Corbel V., N’Guessan R. (2013). Distribution, mechanisms, impact and management of insecticide resistance in malaria vectors: a pragmatic review. In: *Anopheles Mosquitoes-New insights into malaria vectors.* IntechOpen. 4(3):31-39.
- Dahmana H., Mediannikov O. (2020). Mosquito-borne diseases emergence/resurgence and how to effectively control it biologically. *Pathogens.* 9(4):310.
- Echodu R., Iga J., Oyet W.S., Mireji P., Anena J., Onanyang D., et al. (2020). High insecticide resistances levels in *Anopheles gambiae* sl in northern Uganda and its relevance for future malaria control. *BMC Res Notes.* 13(1):1–6.
- Glunt K.D., Blandford J.I., Paaijmans K.P. (2013). Chemicals, climate, and control: increasing the effectiveness of malaria vector control tools by considering relevant temperatures. *PLoS Pathog.* 9(10):e1003602.
- Hemingway J. (2014). The role of vector control in stopping the transmission of malaria: threats and opportunities. 365(1645) *Philos Trans R Soc B Biol Sci.* 20130431.
- Hill J., Lines J., Rowland M. (2006). Insecticide-treated nets. *Adv Parasitol.* 61:77–128.
- Initiative PM. PMI - President’s Malaria Initiative [Internet]. PMI. [cited 2023 June 19]. Available from: <https://www.pmi.gov/>

- Killeen G.F., Ranson H. (2018). Insecticide-resistant malaria vectors must be tackled. *The Lancet*.391(10130):1551-2.
- Lengeler C., Grabowsky M., McGuire D., deSavigny D. (2007). Quick wins versus sustainability: options for the upscaling of insecticide-treated nets. *Define Defeating Intolerable Burd Malar III Prog Perspect Suppl* 77 (6): 222-226.
- Massebo F., Balkew M., Gebre-Michael T., Lindtjørn B. (2015). Zoophagic behaviour of anopheline mosquitoes in southwest Ethiopia: opportunity for malaria vector control. *Parasit Vectors*. 8(1):1-9.
- Mejía P., Teklehaimanot H.D., Tesfaye Y., Teklehaimanot A. (2013). Physical condition of Olyset® nets after five years of utilization in rural western Kenya. *Malar J*. 12:1-11.
- Nasir S.M., Amarasekara S., Wickremasinghe R., Fernando D., Udagama P. (2020). Prevention of re-establishment of malaria: historical perspective and future prospects. *Malar J*. 19(1):1-16
- Okumu F.O., Moore S.J. (2011). Combining indoor residual spraying and insecticide-treated nets for malaria control in Africa: a review of possible outcomes and an outline of suggestions for the future. *Malar J*.10(1):1-13.
- Protopopoff N., Wright A., West P.A., Tigererwa R., Mosha F.W., Kisinza W. (2015). Combination of insecticide treated nets and indoor residual spraying in northern Tanzania provides additional reduction in vector population density and malaria transmission rates compared to insecticide treated nets alone: a randomized control trial. *PloS One*. 10(11).
- Raghavendra K., Barik T.K., Reddy B.N., Sharma P., Dash A.P. (2011). Malaria vector control: from past to future. *Parasitol Res*. 108:757-79.
- Ranson H. (2017). Current and future prospects for preventing malaria transmission via the use of insecticides. *Cold Spring Harb Perspect Med*. 7(11):a026823.
- Sougoufara S., Doucouré S., Sembéne P.M.B., Harry M., Sokhna C. (2017). Challenges for malaria vector control in sub-Saharan Africa: resistance and behavioral adaptations in *Anopheles* populations. *J Vector Borne Dis*. 54(1):4.
- Takken W., Kohls B.G. (2009). Malaria vector control: current and future strategies. *Trends Parasitol*. 25(3):101-104.

Tizifa T.A., Kabaghe A.N., McCann R.S., van den Berg H., Van Vugt M., Phiri K.S. (2018). Prevention efforts for malaria. *Curr Trop Med Rep*. 5:41–50.

Wills A.B., Smith S.C., Anshebo G.Y., Graves P.M., Endeshaw T., Shargie E.B.(2013). Physical durability of PermaNet 2.0 long-lasting insecticidal nets over three to 32 months of use in Ethiopia. *Malar J*.12:1–13.

World Health Organization (WHO) (2022). Guidelines for malaria, World Health Organization

REVIEW ON VARIATION IN GENETIC AND CHEMICAL CONSTITUENTS OF *Strychnos henningsii* POPULATIONS IN KENYA

WAHU K.M.

School of Education, Humanities and Social Sciences

Kabarak University, Kenya

Correspondence: mkuria@kabarak.ac.ke

Abstract

Strychnos henningsii is an indigenous medicinal plant species widely used in tropical Africa. Studies have revealed that this plant has been used as a remedy for various ailments including rheumatism, gastrointestinal complications, abdominal pains, syphilis, snakebites, diabetes malaria, and arthritis amongst others. Phytochemical and pharmacological studies have identified various compounds such as alkaloids, anthraquinones, cardiac glycosides, chalcones, flavonoids, phenolics, proanthocyanidins, saponins, steroids, tannins and triterpenes from the crude extracts of *S. henningsii*. These chemical constituents exhibited analgesic, antibacterial, antidiabetic, anti-inflammatory, antioxidant, antiplasmodial, antiprotozoal, antispasmodic as well as cytotoxicity activities. Secondary metabolites are known to aid plants in coping with various environmental stresses. Environmental stress triggers expression of genes for the enzymes involved in biosynthesis of secondary metabolites, many of which have higher medicinal value despite being useful in plant defense mechanisms. This paper is a review on the chemical constituents, pharmacological properties and genetic variation of *S. henningsii* across its geographical range.

Key Words: *Strychnos henningsii*, chemical constituents, genetic, medicinal, variation.

INTRODUCTION

Botanical information of *S. henningsii*.

S. henningsii belongs to the family Strychnine but was earlier included in the family Loganiaceae. The species epithet honors Professor Paul Christopher Henning, a mycologist at the Royal Botanical Gardens, Berlin-Dahlem. The common names are Red bitter berry (English) (Gachathi, 2007), Henning's Strychnos (Maundu and Tengäs, 2005). The local names include Muteta (Kikuyu and Kamba), Maset (Kipsigis), Entuyesi (Maasai), Mutambi (Mbeere), Muchimbi (Meru), Kapkamkam (Pokot), Nchipilikwa (Samburu), Hadesa (Somali), Turukukwa (Tugen) and Yapoliss (Turkana) (Maundu and Tengäs, 2005).

It varies in size from a shrub or small erect tree, much-branched tree of about 2 to 15 m tall with green-reddish stem. The bark is pale grey and smooth in young trees but becomes dark brown and somewhat flaky in specimens. The twigs have pale ashy or straw-colored and waxy skin splitting lengthwise. Lenticles are few and inconspicuous. Leaves are opposite, sub-sessile or ovate, 2.5 to 6.5 cm long and 0.8 to 4.5 cm wide. They have an entire margin and acuminate leaf tips. The leaves are strongly; three to five nerved from base cuneate or rarely sub-cordate at base; a characteristic feature in *Strychnos* species (Van Wyk et al., 1997). Floral cymes are borne on flat clusters in the leaf axils, 2 to 2.5 mm long and 4 mm wide when open, scented, yellowish-green in color turning orange with age. The ovary is globose with a short style. The fruit is up to 1.9 cm long and 6 to 11 cm wide, oblong or roundish with one to two seeds (coffee-like) red, brown or orange when ripe (Figure 1) (Beentje, 1994; Gachathi, 2007; Maundu and Tengäs 2005).

S. henningsii is a semi-deciduous plant commonly occurring in the dry and moist forests, wooded hillsides and thickets, on rocky hills, coastal forests and stream banks. It is native to Angola, Mozambique, South Africa, Swaziland, Tanzania and Uganda. In Kenya, it is widely distributed in Nairobi, Kakamega, and in the Central province. It is often associated with dry

Podocarpus and Olea forests, hillsides, thickets and *Combretum* bushland (Maundu and Tengäs, 2005). It is raised from seedlings or wildings. The species also suckers well. The pulp is removed before sowing the seeds. The seeds exhibit orthodox storage behavior. It is managed through pruning and coppicing (Maundu and Tengäs, 2005).

In the African traditional medicine, it is used for treatment of various ailments including rheumatism, gastrointestinal complications, abdominal pains, syphilis, and possibly of value in dysmenorrhoea (Hutchings, 1989; Watt and Breyer, 1962; Pujol, 1993; Hutchings, 1996; Oyedemi *et al.*, 2009). Root's bark and green fruits of *Strychnos* species are used as a remedy for snakebites (Tits *et al.*, 1991; Van Wyk *et al.*, 1997) and hookworm infections in Tanzania (Oyedemi *et al.*, 2009). The bark decoction is employed as a remedy for rheumatism and arthritis (Palgrave, 1988; Beentje, 1994). A decoction of the plant has been used in traditional Kenyan medicine for the treatment of rheumatism, gynecological complaints, chest pain, internal injuries and malaria (Kareru *et al.*, 2007). The ground bark is a mouth antiseptic and applied on the wounds in cattle and horses to hasten healing (Gachathi, 2007). In South Africa, the decoction or infusions from the stem bark is widely used for the management Diabetes mellitus (Oyedemi *et al.*, 2009). The aqueous bark extract is also used in South Africa for the treatment of stomach, colic, dizziness and as a purgative agent (Oyedemi *et al.*, 2013). Almost all parts of *S. henningsii* are used as a source of medicine, however studies conducted elsewhere (Alfred, 2021 and Kuria *et al.*, 2012) revealed that the roots, stem and the bark are the most commonly used parts for medicinal purposes in different parts of Africa. This plant species is mainly used as an anthelmintic, appetizer, blood cleanser, purgative, and tonic as well as in ethnoveterinary medicine (Alfred, 2021). In traditional medicine, it is mainly used as a remedy for abdominal pain, bilharziasis, colic, diabetes mellitus, gastro-intestinal complications, headache, malaria, menstrual problems, respiratory diseases, rheumatism, snake bites and syphilis (Alfred, 2021 and Kuria *et al.*, 2012).

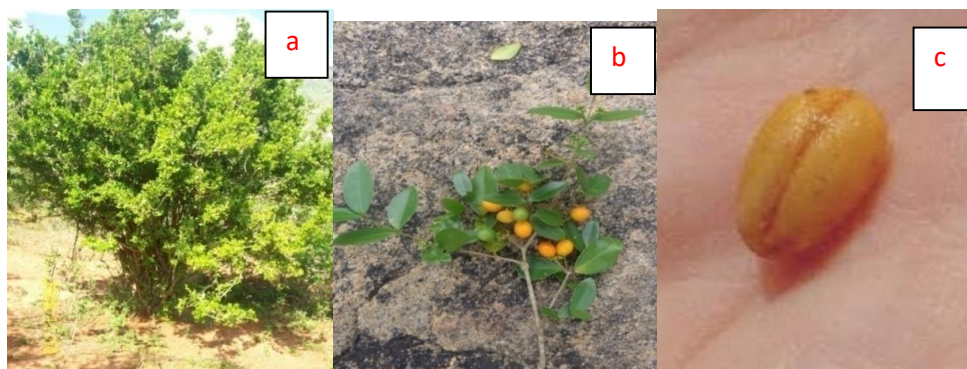


Figure 1 a, b and c: *S. henningsii* Shrub, fruits and seed

Variation in chemical constituents and pharmacological properties of *S. henningsii*

Various studies conducted by different authors have revealed various bioactive compounds isolated from different parts of *S. henningsii*. Such compounds include the indolinic alkaloids, strychnine, brucine, curanine, and bitter glycosides with significant values (Penelle *et al.*, 2000; Oyedemi *et al.*, 2010a). Other compounds including holstine, diaboline, strychnochromine and guianensine have been isolated from the stem and root bark of *S. henningsii* (Angenot and Tits, 1981). A research conducted by (Alfred, 2021) also revealed a wide range of biological compounds such as alkaloids, anthraquinones, cardiac glycosides, chalcones, flavones, flavonoids, flavonols, phenolics, proanthocyanidins, saponins, steroids, sterols, tannins and triterpenes produced by this plants species. These compounds have been isolated from the bark, leaves, roots, root bark, stem bark and twigs of *S. henningsii*. Some of these phytochemical compounds may be responsible for the various pharmacological properties exhibited by *S. henningsii*.

Pharmacological studies have shown that the different phytochemical compounds identified from extracts of *S. henningsii* have various biological activities. They included antibacterial (Tirop *et al.*, 2019; Njire *et al.*, 2010), antidiabetic (Ngugi *et al.*, 2011; Oyedemi *et al.*, 2012; 2013), anti-inflammatory (Tits *et al.*, 1991), antioxidant, (Oyedemi *et al.*, 2010a; 2013)

antiplasmodial (Phillipe *et al.*, 2005; Kirira *et al.*, 2006; Frederich *et al.*, 1999), antiprotozoal (Wright *et al.*, 1994), antispasmodic (Tits *et al.*, 1991), cytotoxicity (Oyedemi *et al.*, 2012 and 2013) and toxicity (Ogeto *et al.*, 1984; Oyedemi *et al.*, 2010a; Tirop *et al.*, 2018). Although there are several reports about phytochemical constituents and pharmacological properties of *S. henningsii*, information about the variation of these chemical constituents of *S. henningsii* based on its geographical location is unavailable.

Plants of the same species growing in different geographical locations are subjected to a wide range of biotic and abiotic environmental factors (Hartmann *et al.*, 2005; Rapinski *et al.*, 2014 and 2015; Baille *et al.*, 2016; Mahmoud *et al.*, 2016). These environmental factors trigger an adaptive response by stimulating gene expression for enzymes responsible for production of a wide array of secondary metabolites in plants which in turn may have medicinal value (Mahmoud *et al.*, 2016). The environmental stressors include temperature, low precipitation, solar radiation as well as edaphic factors. These environmental factors are subject to latitudinal, longitudinal and altitudinal gradients and hence the differences in the chemical constituents of plants (Dixon *et al.*, 2006; Asensio *et al.*, 2020).

There is a general perception among the traditional healers and elders that plants in the Northern latitude and coastal regions are more efficient sources of traditional medicine as well as higher altitude because they accumulate higher concentrations of secondary metabolites (Baille *et al.*, 2016). For instance, higher contents of rutin were reported in the populations of *Casearia sylvestris* (SW) growing in Savannah (poor soils and higher solar radiation) as well as those in higher altitudes (Silva *et al.*, 2006). Plants growing in higher altitudes were reported to contain high levels of flavonoids in *Calluna vulgaris* populations (Monschein *et al.*, 2010), *Arnica montana* (Perry, *et al.*, 2009) and *Quera robur* (Abdala-Roberts *et al.*, 2016). Higher contents of flavonoids and anthocyanins were reported in plants growing in higher latitude due to longer daylight periods and lower night temperature (Lätti *et al.*, 2010). Higher levels of phenolic compounds were also reported in Bearberry plants growing

in the areas of higher radiation and temperature (Asensio *et al.*, 2020). The island populations of *Prunus Africana* have been overexploited for medicinal purposes than the inland populations (Kadu *et al.*, 2012).

S. henningsii is widely distributed in the tropical and subtropical areas in Africa. It occurs in wooded and open forests from sea level up to 2200 m altitude (Ruijter *et al.*, 2008). This may explain the wide range of chemical variation from individuals of this plant species from various geographical locations. Additionally, ethnobotanical studies also revealed that different parts of this plant species are used as sources of medicine in different geographical areas (Kuria *et al.*, 2012 and Alfred 2021). Plants exhibit differences in their chemical components not only according to their locality but their tissue types as well (Fraster *et al.*, 2007). Differences in phenolic compounds and activities observed according to tissue type were supported by traditional healers who used decoctions from specific parts of the plant for different symptoms (Fraster *et al.*, 2007). Inner bark is more preferred medicinally because it tends to show higher degree of antioxidant activity than leaves (McCune and Johns 2007; Fraster *et al.*, 2007; Rapinski *et al.*, 2014 and 2015). Ethnobotanical studies of *S. henningsii* revealed that the roots and the stem were the most widely used plant parts for medicinal purposes in the areas of study (Kuria *et al.*, 2012).

Further genetic variation in plants is partially reflected in the variation of concentrations and types of chemical constituents produced in a plant species growing in different geographical locations (Baille *et al.*, 2016). This is because the genetic makeup of plants provides the ability or inability to synthesis certain compounds (Chaplain, 1975; Baille *et al.*, 2016) and such potential depends on the differences in the environmental conditions prevailing in the geographical location of specific plant population (McCune and Johns 2007; Theis and Lerda, 2003; Dixon and Paiva, 1995; Figueredo *et al.*, 2008; Fraster *et al.*, 2007).

In literature, there is no available information explaining association between genetic variation and the phytochemical constituents of *S. henningsii*. Evaluating the link between

genetic diversity and chemical constituents of this plant will provide useful insight for designing strategies for sustainable utilization and conservation of this important plant species. Additionally, this information will also be useful in developing drugs from populations that show higher potential as a remedy for various ailments.

Genetic variation of *S. henningsii* from different populations in Kenya

Genetic variation accounts for the chemical diversity in plants (Moore *et al.*, 2014). Genetic diversity promotes the adaptation of organisms to environmental conditions (Onda *et al.*, 2016). Environmental factors such as soil nutrients, temperature, water availability and light amongst others influence the genetic and chemical diversity of plant populations (Pacheco-Hernández *et al.*, 2021). These environmental conditions exert strong selective pressures that could influence the evolutionary course of plant populations (Pacheco-Hernández *et al.*, 2021). This natural phenomenon causes plant populations consisting of single species to show varied genetic patterns and chemical variations in different geographical locations (Chen *et al.*, 2015). Expression of genes for enzymes involved in production of secondary metabolites in plants varies and it's higher in plants subjected to areas characterized with stressful environmental conditions (Baille *et al.*, 2016). Plants adapt to new environmental conditions due to their genetic variation that may be associated with specific chemical compounds produced (Via and Conner, 1995; Younsi *et al.*, 2018). Maintenance of genetic diversity of plant species is vital for selecting the best fit (adaptable) individuals and self-sustaining populations (Reed and Frankham, 2003).

The effects of genetic variation in the biosynthesis of secondary metabolites in medicinal plants have been reviewed (Iannicelli *et al.*, 2020). Variation in genetic and chemical constituents has been reported in various plants (Silva *et al.*, 2006; Khan *et al.*, 2017; Asensio *et al.*, 2020). High contents of flavonoids and anthocyanins were reported in plants growing in high latitude due to longer daylight durations and lower night temperatures (Lätti *et al.*, 2010). High contents of phenolic compounds in bearberry plants were also reported in plant

species growing in areas of higher radiation and temperature (Asensio *et al.*, 2020). Altitudinal variations also influence production of secondary metabolites, for instance higher levels of flavonoids were reported in *Calluns vulgaris* populations growing in higher altitudes, *Arnica Montana* (Perry *et al.*, 2009), *Quera robur* (Abdala-Roberts *et al.*, 2016) and in *Casearia sylvestris* (SW). Silvia *et al.*, (2006) reported higher levels of rutin production in plants growing in high altitude and savannah regions.

A study on genetic diversity of *S. henningsii* was conducted by (Kurira *et al.*, (2018) using ISSR markers (Figure 2). Nine populations were selected from areas identified from the following places: Taita-Taveta (Mwache forest), Kilifi (Arabuko Sokoke forest), Narok (Tipilikwani forest in Talek near Maasai Mara game reserve), Baringo (Tugen hills), Kitui (Ndumooni hills), Marsabit (Marsabit forest reserve), Nyeri (Kabiruini forest), Kiambu (Karura forest) and Kajiado (Ngong forest). Each population comprised thirty individuals and therefore a total of two hundred and seventy individuals were randomly selected from the nine populations to conduct a study on genetic diversity of *S. henningsii* in Kenya. Nine markers that gave clear and reproducible bands were selected to help determine the genetic diversity among *S. henningsii* populations.

ISSR markers have been successfully used in other studies to determine the genetic diversity of medicinal plant species such as in *Croton heliotropiifolius* in (Rocha *et al.*, 2016), *Varronia curassavica* (Jacq.) in (Brito *et al.*, 2016), *Rheum* spp, in (Tabin *et al.*, 2016), *Withania Somnifera* in (Khan and Shah 2016) and *Croton tetrandenius* in (Almeida-Pereira *et al.*, 2017) amongst others.

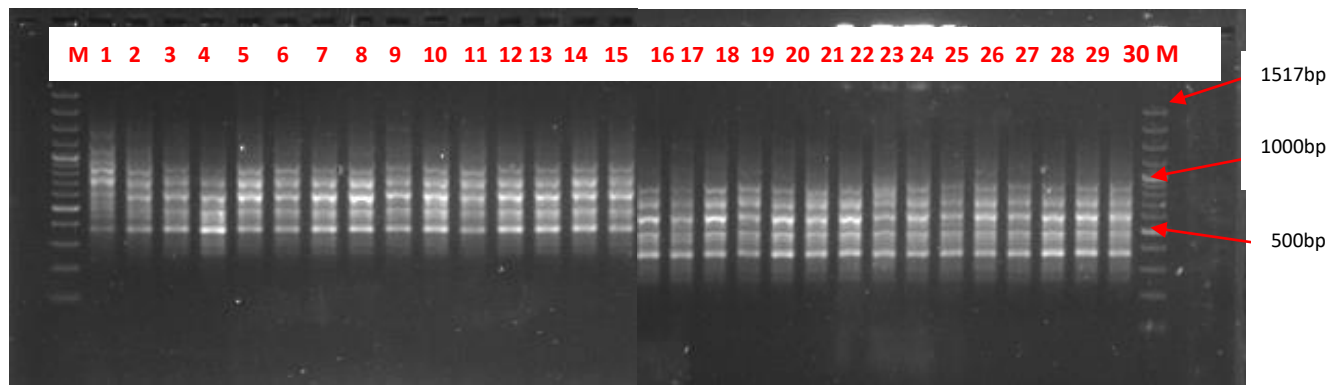


Figure 2: ISSR marker profile of amplified loci of samples from Baringo population using primer 862. Lane (1-30) are samples, M:-Marker DNA 100bp ladder

In this study, ISSR detected and amplified a total of 96 loci among *S. henningsii* genotypes, all of which were polymorphic. The mean percentage of polymorphism detected was 43.40%. The most polymorphic population was Ngong with 51 polymorphic loci (53.12 %) while Baringo was the least polymorphic population with 28 polymorphic loci (29.17%) (Table 1). Similar results were reported in other studies on genetic diversity using ISSR markers. For example, a percentage polymorphism of 42.47% was revealed in *Costus pictus* (Naik *et al.*, 2017) and 59.13% percentage polymorphism in *Peganum harmals* L. (Zebarjadi *et al.*, 2016). High polymorphism (94.8%) was reported in *Croton tetradenius* (Almeida-Pereira *et al.*, 2017), 93.4% in *Ziziphus sphi-christi* l. (Alansi *et al.*, 2016) and 76.1% in *Thuja sutchuenensis* (Liu *et al.*, 2013). Other studies reported low polymorphism using the same markers such as 24.36% in *Bruguiera gymnorrhiza* and 12.73% in *Heritiera fomes* (Dasgupta *et al.*, 2015).

Table 1 Genetic diversity analysis of nine populations of *S. henningsii* as revealed by ISSR markers in GenAlex software.

ISSR Markers								
Population	%P	N	Na	Ne	I	He	UHe	PSL
	43.75	30.00						
Kitui	%	0	0.917	1.307	0.251	0.172	0.175	5.000

	41.67	30.00						
Marsabit	%	0	0.865	1.255	0.219	0.147	0.149	1.000
	29.17	30.00						
Baringo	%	0	0.688	1.159	0.145	0.096	0.097	0.000
	39.58	30.00						
Nyeri	%	0	0.802	1.282	0.232	0.159	0.162	0.000
	42.71	30.00						
Narok	%	0	0.885	1.271	0.230	0.156	0.158	2.000
	51.04	30.00						
Karura	%	0	1.063	1.376	0.299	0.207	0.211	2.000
	53.13	30.00						
Ngong	%	0	1.115	1.315	0.273	0.183	0.186	2.000
	37.50	30.00						
Jilore	%	0	0.781	1.235	0.203	0.137	0.139	0.000
	52.08	30.00						
Taveta	%	0	1.052	1.298	0.267	0.177	0.180	2.000
	43.40	30.00						
Mean	%	0	0.907	1.278	0.236	0.159	0.162	

N= population size, PPL= population polymorphic loci, % P= percentage polymorphism, Na= Number of observed alleles, Ne= number of effective alleles, H= Nei's genetic diversity, I= Shannon information indices, He= expected Heterozygosity, UHe= unbiased expected Heterozygosity, PSL=population specific loci.

According to Nei (1978), percentage polymorphism is not a significant measure of genetic variation despite being the most commonly used indicator of genetic variation in many studies on natural population and that the parameter of genetic diversity (H) is more

appropriate. The values for genetic diversity (H) and Shannon index (I) ranged from (0.0955 – 0.1828 and 0.1448-0.2728 respectively according to a study conducted by Kuria *et al.*, (2018) (Table 2). According to genetic diversity and Shannon index values this study showed that the Ngong population was the most diverse while Baringo was the least diverse. These values indicate a low genetic (allelic) diversity for *S. henningsii* populations. The results obtained could be attributed to the pollination, propagation and seed dispersal mechanisms.

Table 2 Genetic diversity analysis of nine populations of *S. henningsii* has revealed using ISSR markers in PopGene software

ISSR Markers							
Population	N	PPL	%P	Na*	Ne*	H*	I*
Kitui	30	42	43.75	1.4375	1.3067	0.1720	0.2514
Marsabit	30	40	41.67	1.4167	1.2548	0.1469	0.2189
Baringo	30	28	29.17	1.2917	1.1594	0.0955	0.1448
Nyeri	30	38	39.58	1.3958	1.2823	0.1590	0.2317
Narok	30	41	42.71	1.4271	1.2715	0.1558	0.2303
Karura	30	49	51.04	1.5104	1.3764	0.2071	0.2994
Ngong	30	51	53.12	1.5312	1.3148	0.1828	0.2728
Jilore	30	36	37.5	1.3750	1.2346	1.1366	0.2030
Taveta	30	50	52.08	1.5208	1.2977	0.1773	0.2673
Overall	270	96	100	2.0000	1.4683	0.2889	0.4473

Key words:

N= population size, PPL= population polymorphic loci, % P= percentage polymorphism, Na = Number of observed alleles, Ne = number of effective alleles, H= Nei's genetic diversity, I = Shannon information indices.

This plant species has cleistogamous reproduction (self-pollinating) (Bruce and Lewis, 1960). It bears small and brightly colored flowers which indicate a high possibility of entomophilous pollination. Insects transfer pollen for short distance mainly on flowers in a single tree resulting in the production of inbred seeds with poor germination (Bryndum and Hedgegart, 1969; Mathew *et al.*, 1987; Indira and Mohandas, 2002; Tangmitcharoen *et al.*, 2009). *S. henningsii* is also known to have restricted geographical zones within its natural environment. All these factors may have contributed to the low genetic diversity due to the narrow and common gene pools in the populations of this plant species.

Analysis of molecular variance (AMOVA) revealed a higher genetic variation $p < 0.001$ (58 %) among than within (42%) the *S. henningsii* provenances (Figure 3). This was possibly due to the self-pollinating nature of the species. Khan and Shah, (2016) reported a higher genetic variation among populations than within population in *Withania somnifera*, a self-pollinating plant species. Genetic drift may have also contributed to the higher genetic variation due to loss of some alleles through successive generations. The preserved alleles may be responsible for the adaptation of this plant species in its specific but wide geographical distribution range from the sea level up to about 2220 m above sea level. The geographical locations vary in the environmental conditions hence the individuals from the different geographical areas (populations) cope up differently by producing various chemical compounds that aid in the adaptation process (Baille *et al.*, 2016). This may explain the wide range of chemical constituents produced by this plant species.

Habitat fragmentation and reduction in population size in the wild medicinal plants as a result of over exploitation is one of the main causes of increased genetic differentiation among populations and reduced gene flow between populations (Panda *et al.*, 2015). In the study conducted by (Kuria *et al.*, 2018), Kitui, Taita-Taveta and Nyeri populations revealed

higher genetic variation due to the reduction in gene pool within these populations as a result of over exploitation for medicinal purposes.

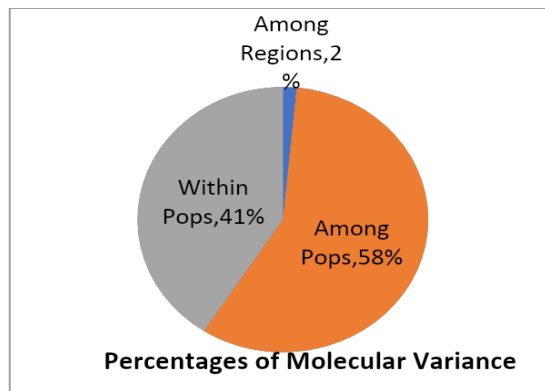


Figure 3 Percentage of Molecular variance of ISSR data

Cluster analysis of ISSR data based on the Nei's, (1978), unbiased genetic distance generated a dendrogram with three major groups. Cluster I consisted of three populations namely Kitui, Ngong and Jilore. Cluster II consisted of five populations (Marsabit, Taveta, Nyeri, Narok and Karura) and cluster III consisted of Baringo population (Kurua *et al.*, 2018) (Figure 4). The Principal Coordinate Analysis confirmed the results of the clustering analysis where there was dispersion in the genetically diverse populations (Figure 5). However, the UPGMA and PCA analyses did not reveal a clear pattern of clustering and the geographical trend among the populations (Figure 4 and 5). Therefore, genetic divergence did not match to the geographical places of collection. The lack of correlation between genetic distance and geographical locations indicate that genetic drift has played a significant role in shaping the genetic structure and variation among populations of *S. henningsii* (Fischer *et al.*, 2000).

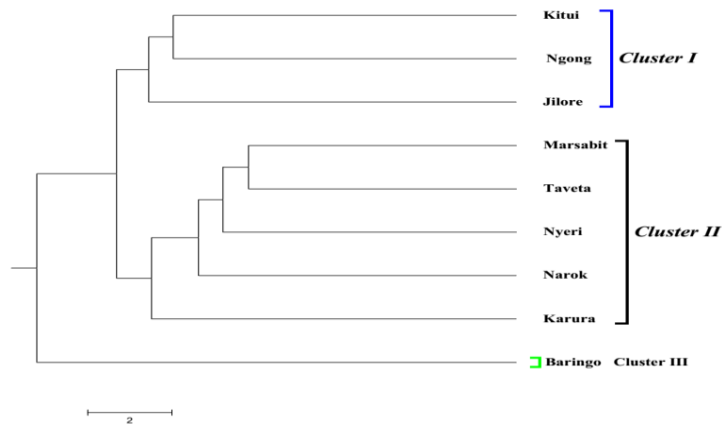


Figure 4 UPGMA clustering analysis of nine *S. henningsii* populations based on Nei's (1978) unbiased genetic distance

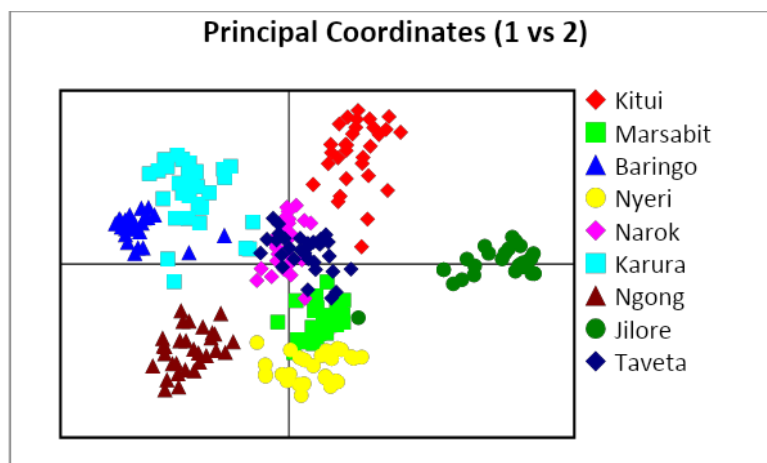


Figure 5 a three dimensional plot of the Principal Coordinate Analysis (PCA) of ISSR data showing the clustering of *S. henningsii* populations.

RECOMMENDATIONS AND CONCLUSIONS

Strychnos henningsii is a traditional medicinal plant species widely used in tropical Africa. Overview of this plant species has revealed that it has been used as a remedy for various ailments including rheumatism, gastrointestinal complications, abdominal pains, syphilis, among others in African traditional medicine. It is a source of several important chemical constituents which could possibly be responsible for its various pharmacological activities

exhibited by this plant species. Genetic diversity study revealed higher genetic variation among the populations than within the population. However, there are no reports describing association between the variation in genetic and chemical constituents of this plant species across its geographical range. There is therefore a need for further studies to provide more insights in the conservation strategies of genotypes that show superiority in their genetic and chemical constituent's variability. These genotypes can serve as sources of raw material for the development of new drugs that could be useful in treatment of chronic diseases in both medical and veterinary institutions.

References

- Abdalla-Roberts L., Rasmann S., Berny-Mier Y., Terán J.C., Covelo F., Glauser G. M. (2016). Biotic and abiotic factors associated with altitudinal variations in plant traits and herbivory in dominant oak species. *American Journal of botany* 103: 2070-2078.
- Alansi S., Tarroum M., Al-Qurainy F., Khan S., Nadeem M. (2016). Use of ISSR markers to assess the genetic diversity in wild medicinal *Ziziphus spina-christi* (L.) Wild collected from different regions of Saudi Arabia, *Biotechnology and Biotechnological Equipment*. 30(5): 942-947.
- Alfred Maroyi (2021). Evaluation of medicinal uses, phytochemistry and pharmacological properties of *Strychnos henningsii* Gilg. (Strychnine). *International Journal of Scientific and Technological Research* 10: 10-18.
- Almeida M.C., Pina E.S., Hernandes, C. *et al.* (2018). Genetic diversity and chemical variability of *Lippia* spp. (Verbenaceae). *BMC Res Notes* **11**: 725 <https://doi.org/10.1186/s13104-018-3839-y>.
- Angenot L., Tits M. (1981). Isolation of a new alkaloid (O-Acetyl Retuline) and a triterpenoid (Friedelin) from *Strychnos henningsii* of Zaire. *Planta Medica* 41(3): 240-243.
- Asensio E., Vitales D., Pérez I., Peralba L., Viruel J., Montaner C., Vallès J., Garnatje T., Sales E. (2020). Phenolic compounds contents and genetic diversity at population level across the

natural distribution range of bearberry (*Arctostaphylos uva-ursi*, Ericaceae) in the Iberian Peninsula. 9(9):1250. doi: 10.3390/plants 90911250.

Bailie A., Renaut S., Uhalijoro, E., Guerrero-Anajco J.A., Saleem A., Haddad P., Cuerrier A. (2016). Phytogeography and genetic variation in *Sorbus* a traditional antidiabetic medicine - adaptation in action in both a plant and a discipline. Peer Journal 4e2645. Doi 10.7717/peerj.2645.4(6)

Beentje H.J. (1994). Kenya Trees, Shrub and Lianas. Nairobi: National Museums of Kenya. Pp 325.

Brito, F.A., Nizio, D.A.C., Silva, A.V C., Diniz, LEC. (2016). Genetic diversity analysis of *Varronia curassavica* (Jacq.) accessions using ISSR markers. *Genetic Molecular Research* **15**: 1-10. <http://dx.doi.org/10.4238/gmr.15038681>.

Bruce, E.A., Lewis, J. (1960). *Loganiaceae* in: Hubbard C.C., Milne- Redlead E (Eds) London: Flora of tropical East Africa. Pp 225.

Bryndum, K., Hedgegart, T. (1969). Pollination of teak (*Tectona grandis*). *Silvae Genetica* **18**: 77-80.

Chaplain J.F. (1975). Genetic influence on chemical constituents of tobacco leaf and smoke. Tobacco research Laboratory. Agriculture Research service. N. Carolina U.S.A.

Chen, J., Xu, Y., Wei, G., Liao, S., Zhang, Y., Huang, W., Yuan, L., Wang, Y. (2015). Chemotypic and genetic diversity in *Epimedium sagittatum* from different geographical regions of China. *Phytochemistry* **116**: 180-187.

Dasgupta, N., Nandy, P., Sengupta, C., Da, S. (2015). RAPD and ISSR marker mediated genetic polymorphism of two mangroves *Bruguiera gymnorhiza* and *Heritiera* from Indian Sundarbans in relation to their sustainability. *Physiology and Molecular Biology of Plants* **21** (3), 375-384. doi: 10.1007/s12298-015-0308-0.

De-Ruijters A. (2008). *Strychnos henningsii* (Gill) in Schmelzer G.H. and Gurib-Fakim A (Eds), *PROTA* (Plant Research of Tropical Africa) **11**: Medicinal Plant 1. Backhuys Publishers, Leiden, the Netherlands, 570-571.

Dixon R.A., Gang D.R., Charlton A. J., Fiehn O., Kuiper H.A., Reynolds T.L., *et al.* (2006). Application of metabolomics in agriculture. *Journal of Agricultural Food Chemistry*. **54**:8984–94.

Dixon R.A., Paiva N.L. (1995). Stress-induced phenylpropanoid metabolism. *The Plant Cell* **7**:1085-1097 DOI 10.1105/tpc.7.7.1085.

Figueiredo A.C., José G.B., Luis G.P., Johannes J.C.S (2008). Factors affecting secondary metabolite production in plants: volatile components and essential oils. *Flavour and Fragrance Journal* **23**:213-226 DOI 10.1002/ffj.1875.

Fischer, M., Husi, R., Prati, D., Peintinger, M., Van Kleunen, M., Schmid, B. (2000). RAPD variation among and within small and large populations of the rare clonal plant *Ranunculus reptans* (Ranunculaceae). *American Journal of Botany* **87**: 1128–1137.

Fraser, M.H., Cuerrier A., Haddad, P.S., Arnason, J.T., Owen, P.L., Johns, T. (2007). Medicinal metabolite production in plants: Volatile components and essential oils. *Flavour and Fragrance Journal* **23**: 213-226.

Gachathi, M. (2007). Kikuyu Botanical dictionary. A guide to plant names, uses and culture values (2nd Ed.), Tropical Botany **143**: 212.

Hartmann U., Sagasser M., Mehrrens F., Stracke R., Weisshaar B. (2005). Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue specific activation of phenylpropanoid biosynthesis genes. *Plant Molecular Biology* **57**(2):155-171 DOI 10.1007/s11103-004-6910-0.

Hutchings, A. (1989). A survey and analysis of traditional medicinal plants as used by the Zulu, Xhosa and Sotho. *Bothalia* **19**: 111–123.

Hutchings, A. (1996). Zulu medicinal plants. Pietermaritzburg's Natural University Press. Pp 205.

Iannicelli, J., Guariniello, J., Tossib, V.E., Regalado, J. J., Di Ciaccioa, L., van Barene, C.M., Pitta Álvarez, S.I., Escandón, A.S. (2020). The “polyploid effect” in the breeding of aromatic and medicinal species. *Scientific Horticulture*. **26**: 108854.

Indira, E.P., and Mohandas, K., (2002). Intrinsic and extrinsic factors affecting pollination and fruit productivity in teak (*Tectona grandis* Linn. f.). *Indian Journal of Genetics* **63**: 208-214.

Kadu C.A.C., Parich A., Schuelar S., Kanrad H., Muluvi G.M., Dog-mating O., Muchugi A., Williams V. I., Ramamonjisoa C., Hafashimana D., et al., (2012). Bioactive constituents in *Prunus Africana* geographic variation throughout Africa and association with environmental and genetic parameters. *Phytochemistry* 83: 70-78.

Kareru, P.G. Kenji, G.M., Cachanja, AN., Keriko, J.M., Mungai, G. (2007). Traditional medicines among the Embu and Mbeere people of Kenya. *Journal of Traditional, Complementary and Alternative medicines (AJTCAM)* 4(1): 75-86.

Khan, S., Shah, R.A. (2016). Assessment of genetic diversity among Indian Ginseng *Withania somnifera* (L.) Durnal using RAPD and ISSR markers. *Research in Biotechnology* 7: 1-10.

Khan M., Feroza H.W., Israr Ahmed., Ikram Mohamad., Sahib Gul-Afridi., Muhamad Hamid S.W., Mustafa Kamal S.S. (2017). Assessment of genetic diversity and phytochemical analysis of *Nigella sativa* genotypes from Pakistan. *Asian Journal Biological Sciences* 10: 56-63.

Kuria, M.W., Njenga, P.K., Ngumi, V.W. (2012). Ethnobotanical studies of *Strychnos henningsii* in five (Gilg.) natural habitats in Kenya. *International Journal of medicinal Plant research* 1(16): 063-074.

Kuria, M. W., Ngumi, V.W., Njenga, P.K., Wangai L.N., Magiri E. (2018). Assessing genetic diversity of an endangered medicinal plant *Strychnos henningsii* (Gilg.) in nine populations in Kenyans Counties as revealed by ISSR Markers. *International Journal of Innovative Research and Knowledge* 3(12): 81-94.

Lätti, A.K., Riihinen K.R., Kainulainen P.S. (2010). Analysis of anthocyanin variation in wild populations of berberry (*Vaccinium myrtillus* L.) *Journal of Agriculture and Food Chemistry* 58: 427-433.

Liu, J., Shi, S., Chang, E., Yang, W., Jiang, Z. (2013). Genetic Diversity of the critically endangered *Thuja sutchuenensis* as revealed by ISSR Markers and the Implications for Conservation. *International Journal of Molecular Sciences* 14: 14860-14871; doi: 10.3390/ijms140714860

- Mathew, G., Mathew, P.K., Mohandas, K. (1987). Preliminary studies on Insects' visitors to teak (*Tectona grandis* L.f.) inflorescence in Kerala, India. *Indian Forester* **113**: 61-64.
- Maundu, P., Tengäs, T. (2005). Useful Trees and Shrubs of Kenya. Technical handbook No. 35. Nairobi, Kenya: World Agroforestry Center-Eastern and Central Africa regional Program (ICRAF-ECA). Pp 400.
- McCune, L.M., Johns, T., (2007). Antioxidant activity relates to plant part, life form and growing condition in some diabetes remedies. *Journal of Ethnopharmacology* **112**:461- 469 DOI 10.1016/j.jep.2007.04.006.
- Moore, B.D., Andrew, R.L., Külheim, C., Foley, W.J. (2014). Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytology* **201**: 733–750.
- Monschein, M.; Iglesias, J.; Kunert, O.; Bucar, F (2010). Phytochemistry of heather (*Calluna vulgaris* (L.) Hull) and its altitudinal alteration. *Phytochem. Rev.*, **9**, 205–215.
- Naik, A., Prajapat, P., Krishnamurthy, R., Pathak, J.M. (2017). Assessment of genetic diversity in *Costus pictus* accessions based on RAPD and ISSR markers. *Biotechnology* **7**(1): 70-75.
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a smaller number of individuals. *Genetics* **89**: 583-590.
- Ngugi M.P., Murugi N.J., Kibiti M.C., Ngeranwa J.J., Njue M.W., Maina D., Gathumbi K.P., Njagi N.E. (2011). Hypoglycemic activity of some Kenyan plants traditionally used to manage diabetes mellitus in Eastern province. *Diabetes and Metabolism* **2**: 8.
- Njire M.M., Bundambula N, M., Kiiru J.N. (2010). Antimicrobial effects of selected herbal extractson multidrug resistance gram negative bacteria strains. *Proceedings of 2010 JKUAT Scientific, Technological and Industrialization Conference* Pp 305.
- Ogeto J.O., Juma F.D., Muriuki G. (1984). Practical therapeutics: Some investigations of the toxic effects of the alkaloids extracted from *Strychnos henningsii* (Gilg) 'muteta. *East African Medical Journal* **61**: 427-432.
- Onda, Y., Mochida, K. (2016) Exploring genetic diversity in plants using high-throughput sequencing techniques. *Current Genome* **17**: 358–367.

Oyedemi, S.O., Bradley, G., and Afolayan, A.J., (2009). Ethnobotanical survey of medicinal plants used for the management of Diabetes mellitus in the Nkonkobe municipality of South Africa. *Journal of medicinal plants research* 3(12): 1040- 1044.

Oyedemi, S.O., Bradley, G., Afolayan, A.J. (2010a). In vitro and in vivo antioxidant activities of aqueous extract of *Strychnos henningsii* Gilg. *African Journal of Pharmacology* 4: 70-78.

Oyedemi, S.O., Koekemoer, T., Bradley, G., Van de Venter, M., Afolayan, A. (2013). In vitro anti-hyperglycemia properties of the aqueous stem bark extract from *Strychnos henningsii* (Gilg.). *International Journal Diabetes in developing countries* 33(2): 120-127.

Pacheco-Hernández, Y., Villa-Ruano, N., Lozoya-Gloria, E., Barrales-Cortés, C.A., Jiménez-Montejo, F.E., Cruz-López, M.D.C. (2021). Influence of Environmental Factors on the Genetic and Chemical Diversity of *Brickellia veronicaefolia* Populations Growing in Fragmented Shrub Lands from Mexico. *PLANTS* 10: 325. <https://doi.org/10.3390/plants10020325>.

Palgrave, K.C. (1988). *Trees of South Africa* (5th Ed.) Cape Town: Struik publishers.

Panda, S., Naik, D., Kamble, A. (2015). Population structure and genetic diversity of perennial medicinal shrubs from *Plumbago*. *AOB PLANTS* 7 plv048; doi:10.1093/aobpla/plv048.

Penelle, J., Tits, M., Christen, P., Molgo, J., Brandt, V., Frederich, M., Angenot, L. (2000). Quaternary indole alkaloids from the stem bark of *Strychnos guianensis*. *Phytochemistry* 53: 1057-1066.

Perry N.B., Burgess E.J., Rodriguez G., Romero M.A., Franco R., López Mosquera E., Smallfield B. M., Joyce N.T., Littlejohn R. p. (2009). Sesquiterpene Lactones in *Arnica montana* Helenalin and dihydrohelenalin chemotypes in Spain. *Plant Medicine* 75: 660-666.

Pujol, J. (1993). *Naturafrica-the herbalists' handbook* 4th ed. Jean Pujol Natural Healer Foundation, Durban Pergamon press, Pp 270.

Rapinski M., Liu R., Saleem A., Arnason J.T., Cuerrier A. (2014). Environmental trends in the variation of biologically active phenolic compounds in Labrador Tea, *Rhododendron groenlandicum*, from Northern Quebec, Canada. *Botany* 92:783-794 DOI 10.1139/cjb-2013-0308.

- Rapinski M., Musallam L., Arnason J.T., Haddad P., Cuerrier A. (2015). Adipogenic activity of wild populations of *Rhododendron groenlandicum*, a medicinal shrub from the James Bay Cree traditional pharmacopeia. Evidence-Based Complementary and Alternative Medicine 2015: Article 492458 DOI 10.1155/2015/492458.
- Reed, D.H., Frankham, R. (2003). Correlation between Fitness and Genetic Diversity. Conservation Biology **17**: 230–237.
- Rocha T.O., Freitas J.S., Santos E.S.L., and Scaldaferrri M.M. (2016). Estimate of genetic diversity in Casting (*Croton heliotropiifolius*) based on molecular markers. African Journal of Biotechnology 15: 518-523
- Silvia M.A. S. Da., Ming L.C., Pereira A.M.S., Bertoni B.W., Bartistini A.P., Pereira P. S. (2006). Phytochemical and genetic variability of *Casearia sylvestris* SW. from Sao Paulo State Atlantic forest and Cerrado populations. Rev. Bras. Plant Medicine Botucatu **8**: 159-166.
- Tabin, S., Kamili A.N., Ganie S.A., and Zargar O. (2016). Genetic diversity and population structure of *Rheum* species in Kashmir Himalaya based on ISSR markers. Flora 223: 121-128.
- Tangmitcharoen, S., Tasen, W., Owen, J.N., and Bhodthipuks J. (2009). Fruit set as affected by pollinators of teak (*Tectona grandis* L.f.) at two tree spacing in a seed orchard. Songklanakarin Journal of Science and Technology **31**: 255-259.
- Theis N., Lerchau M. (2003). The evolution of function in plant secondary metabolites. International Journal of Plant Sciences **164**: S93-S102 DOI 10.1086/374190.
- Tirop E. K., Maina N.N., Njenga P.K., Magiri E., Ngumi V.W. (2018). Evaluation of toxicity of *Strychnos henningsii* (Gilg) Loganiaceae leaves and roots aqueous extracts in mice. European Journal of medicinal plants **25**: 1-11.
- Tirop E. K., Ngumi V.W., Njenga P.K., Magiri E., Maina N.N. (2019). Antimicrobial activity of *Strychnos henningsii* (Gilg) Loganiaceae. Jomo Kenyatta university of Agriculture and Technology **19**: 79-83
- Tits, M., Damas, J., Qurtin-Leclercq, J., Angenot, L. (1991). From ethnobotanical uses of *Strychnos henningsii* to anti inflammatories, analgesics and antispasmodics. Journal of ethnopharmacology 34 (2-3): 261-267.

Van Wyk, B.E., Bosch Van, O., Nigel, G. (1997). Medicinal plants of South Africa. (2nd Ed) Pretoria, South Africa. Pp 244-245.

Via, S., Conner, J. (1995). Evolution in heterogeneous environments: Genetic variability within and across different grains in *Tribolium castaneum*. *Heredity* **74**: 80-90.

Watt, M.P., and Breyer, N.G. (1962). Medicinal and poisonous plants of Southern and Eastern Africa. Livingstone: Edinburgh. Pp 225.

Younsi, F., Rahali, N., Mehdi, S., Boussaid, M., Messaoud, C. (2018). Relationship between chemotypic and genetic diversity of natural populations of *Artemisia herba-alba* Asso growing wild in Tunisia. *Phytochemistry* **148**: 48-56.

Zebarjadi A., Ahmadvandi H.R., Kahrizi D., Chenghamirza K. (2016). Assessments of genetic diversity by application of inter-simple sequence repeat primers on Iranian Harmal (*Peganum harmala* L.). Germplasm as an important medicinal plant. *Journal of Applied Biotechnology Reports* **3 (3)**: 441-445

CLOUD DATA SECURITY AUDIT REPORT TECHNIQUES USING BAT INSPIRED ALGORITHM: A REVIEW

ONUJA, A. M.¹, ADEBAYO, O. S², OLALERE, M.³, ALIYU, H. O.⁴

^{1,2,4}Federal University of Technology, Minna, Nigeria.

³Islamic University, Uganda

Correspondence: onuja14@gmail.com

Abstract

Cloud computing has become an interesting attraction for Information and Communication Technology (ICT) practitioners. It is helping to reduce the cost of procuring computing devices and services. To this end, security attacks have been rising with cloud service providers always having to under-report the prevalence in order not to scare away potential cloud service subscribers. In this review, eighty (80) papers were acquired from reputable academic publications using Google Scholar as the search engine. The eighteen (18) carefully selected from the pool of 80, for review, have a total of seventy-four (74) versions on the internet. Furthermore, the authors agree with the research publications, that there is need to bring on board the client into security breach mitigation strategies, being developed to strengthen cloud data security. Findings from this review, reveals that research interest in cloud data security audit has increased from four (4) papers between 2011 and 2016, to thirteen (13) papers between 2017 and 2022. These existing frameworks and techniques are still being tested to improve performance. This research, therefore, proposed a cloud data security audit report technique using Bat inspired algorithm (CDSART-BA) to improve cloud data security.

Keywords: Audit Report, Bat Algorithm, Cloud computing, Data, Security, Smart Environment.

Introduction and Background

Smart environments are gradually gaining research attention and trends with respect to adoption and usability in many human endeavours (Adebesin et al., 2021). Current developments in technology have paved the way for increased safety, security and at the same time increased user experience (Ikuomola and Bashir, 2021) and privacy threats. The major threat to the electronic banking system and other services that support smart nations are the concerns of security and privacy of information (Umoren et al 2021). Cloud computing services is a modern trend in information and communication technology (Ngumbi and Wasike, 2022) that is evolving. It is powered with high-speed internet connection. Commonly used resources like hardware, system and user applications can be sourced as service, over the internet (Shankarwar and Pawar, 2015). Figure 1 one below, presents a diagram of cloud computing.

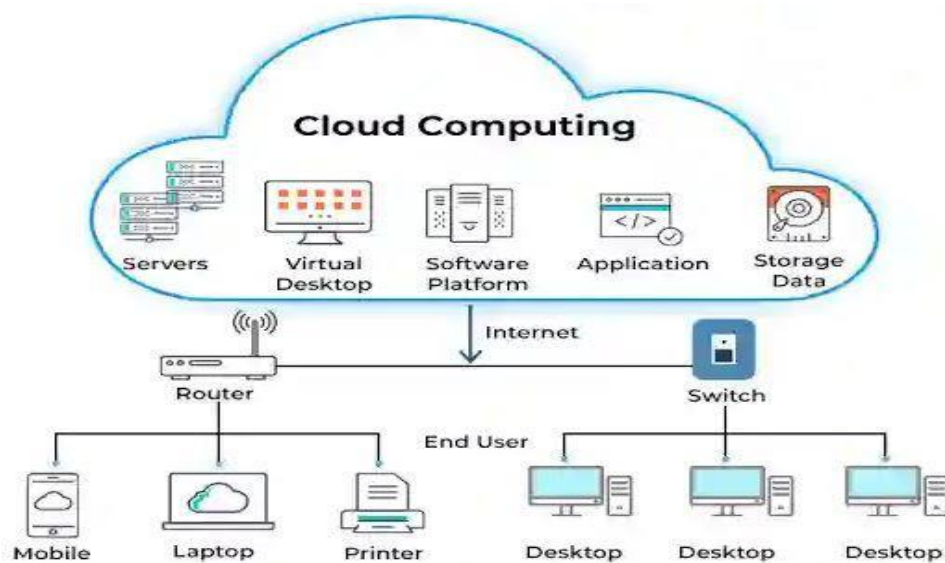


Figure 1: Cloud computing (Fahad, Ahmed and Kahar, 2018). [6].

Cloud computing services have markets that are growing fast, but mainly, security reasons restrict broad wide industrial acceptance (Doelitzscher, 2012). The high number of research with financial support, is a reflection of the expanding cloud security concerns. An example is the European Union 7th Framework Projects, whose objective 1.2 is centered on internet of services (Doelitzscher, 2012). Germany included a trusted cloud, as part of the government's economic development plan (Doelitzscher, 2012). In this concept, Cloud users must adopt their security, backup and business continuity plans to avoid losing valuable data in the cloud (Doelitzscher, 2012).

Enron and WorldCom financial scandals raised concerns by the government about accounting errors and fraudulent practices created within organizations (Fahad et al. 2018). The Sarbanes–Oxley Act (SOX) of 2002 was legislated (Fahad et al. 2018) and Governments policy for Information Security is recommending security audits and certificates as the preferred method of proof to clients of cloud services (Doelitzscher, 2012). IT auditing thus becomes mainstream practice (Fahad, et al. 2018). Figure 2 below is a diagram identifying the role of auditing in cloud data security and accountability life cycle.

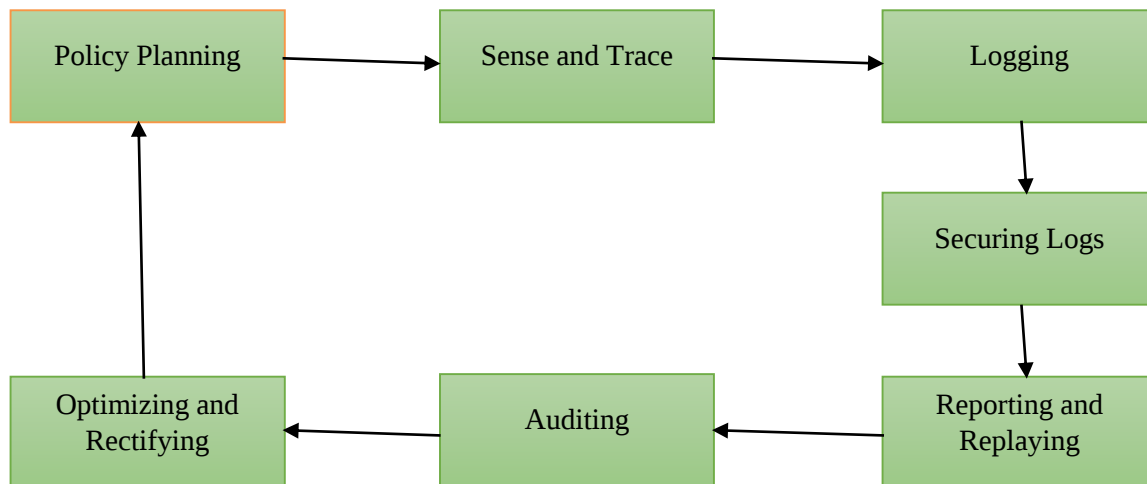


Figure 2: Auditing in Cloud data security accountability life cycle (Ko et al., 2018).

Related Work

Many researchers have actually worked on cloud data security auditing with different techniques and trade off. This serves a basis for this research work aimed to improve on the findings of the papers surveyed.

A. Cloud data security

In supporting the aim of this research, Ko et al., (2018) agrees that preventive controls for security and privacy measures are actively being researched and that there is still little focus on detective controls related to cloud accountability and auditability. The paper presents the TrustCloud framework as a solution but the overhead cost was not highlighted. Many researches have been conducted with the help of Third-Party Auditors (TPA), to check stored data in the cloud to assure the integrity of the environment (El-Booz, Attiya, and El-Fishawy, 2016). Since the cloud service provider (CSP) has the monopoly of interaction with TPA, the CSP most times, is able to influence the concealment of data loss, from cloud subscribers, to prevent defamation (El-Booz et al 2016). The proposed system enhances the authentication level of security by using two authentication techniques; time-based one-time password (TOTP), for cloud users' verification and automatic blocker protocol (ABP), to fully protect the system from unauthorized third-party auditors. However, the mechanism by which auditing reports are generated was not discussed. The proposal in (Ismail and Islam, 2020), is a framework for cloud security transparency, and audit in addition to developing a tool, through which users can collect and analyze evidences from cloud service providers, for determining conformity to requirements, as well as for the specification of remedial actions. The concept of cloud data auditing system in (Mohan and Gladston, 2020) integrates MerkleTree based Cloud audit together with blockchain based audit recording system, thereby recording each verification result into a blockchain as a transaction with a timestamp. This paper suggests the use of the load balancing ability of Bat algorithm to generate audit reports with lower overhead cost.

B. Application of bat algorithm

Bat algorithm (Valarmathi and Sheela, 2019) is used to improve optimization for scheduling of tasks in a cloud environment. The proposed model in (Sreeram and Vuppala, 2019) is the bat algorithm. It is applied as a technique to achieve prompt and fast detection (Mutheu and Wasike, 2020) of distributed denial of service, in the application layer, through HTTP flooding. This justifies the view of this research that bio-inspired bat algorithms can be used as a technique for cloud data security audit reports. Similarly, (Raja Sree, and Mary Saira Bhanu, 2020) present fuzzy bat clustering techniques that read activity logs in cloud virtual machines, and investigates HTTP flooding attacks by grouping identical input patterns. It then determines the anomalous behaviour using deviated anomalous scores.

Sagnika et al (2018) presents an application of the Bat Algorithm (BA) for workflow schedules. The algorithm is successful in laboratory implementation, gives an optimal processing cost, convergence, fair load balancing versus Cat Swarm Optimization (CSO) and Particle Swarm Optimization (PSO).

The discussion in (Ullah et al 2020) is concerned with virtual machines as a tool for resource distribution. This is because when users send data into virtual machines, it can be overloaded sometimes, thus leading to breakdown, thereby affecting data security goals. The researchers used the bat algorithm as a load balancing technique to improve and enhance even the resource allocation system for cloud virtual machines (VM). Bat algorithm also has echolocation potentials that enables it to send auditing reports to users as proposed by this research.

Punitha and Indumathi, (2019) submitted that data integrity protection becomes difficult, as users' subscriptions to cloud computing services increase. There is therefore a need for clients to become part of the security control process with increasing security threats. They suggested accountability with a key generation framework, to maintain users' data in the cloud. This would encircle the logging mechanism together with users' data and policies.

However, bat key generation will require more computing overhead cost hence this paper proposes that the bat algorithm should rather just send an audit report to client to take necessary security measures.

Proposed Method

A. Survey of Research Problems

The research applies the method used in (Yakubu et al., 2019) to carry out the survey for the study. Eighty (80) papers were downloaded between January and July, 2022. The source of the 17 papers cited or reviewed for this work are Nigeria Computer Society library archive, Springer, IEEE, EURASIP, ACM, Elsevier, Research Gate. Notably, Google Scholar served as the gateway to all these research databases. The remaining 18 were sorted from a total of 74 versions. The time range or year interval for the publication used in the review is between 2011 and 2021 as grouped with a class interval of 3 years in table 1 below.

Table 1: Publication Year Interval

2011-2013	[5], [7].	2
2014-2016	[4], [8].	2
2017-2019	[6], [11], [12], [14], [16], [17].	6
2020-2022	[1], [3], [9], [10], [13], [15], [18], [19].	9
Total		18

B. Security Audit Report Model

This research paper is proposing a methodology of using bat algorithms as a technique for cloud data security reports for clients using cloud services. Hitherto, cloud data security audit has been the sole privilege of the cloud service provider (CSP). This was to avoid dissent in patronage. However, current research in cloud computing is bearing the fact that cloud computing clients have to be active partners in the security of data in the cloud. So, thus, there is security requirement for cloud data security audit reports to include clients. Existing research models are shown in figure 3, in which cloud service organization's administrators grant access to cloud users to use their platform as the cloud service provider. There is synchronization between the CSP, administrative and operational third-party auditor, while the client is isolated from the auditor. The model in figure 3 looks at security auditing from the perspective of authentication with CSP deciding what eventual security breach report it relays to the client.

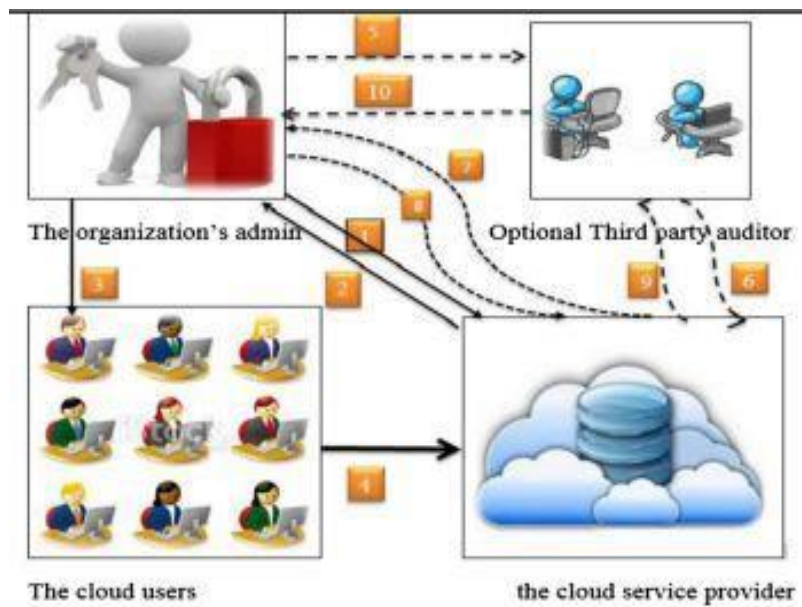


Figure 3: Existing security audit model (El-Booz, Attiya, and El-Fishawy, 2016) [8].

C. CDSART-BA Workflow

Given that cloud service clients have signed into the model cloud environment proposed in this paper, the embedded bat algorithm keeps track of all data transiting from the client to the cloud server. The BA algorithm protects against data loss and flooding of traffic into the communication channel. If the size of source packets of data in bits from the client is equal to the size of packets arriving in the cloud server, the CDSART-BA model sends a secured audit report to the client through email. Also, if an addition or subtraction is noticed in the packets of data in bits from client side to cloud server end, an audit report signaling security threat is sent, enabling the client to take necessary security actions in collaboration with the cloud service provider (CSP). Figure 4 presents an improved technique on the existing model, that gives direct security report access from the cloud service provider to the cloud service client (CSC)

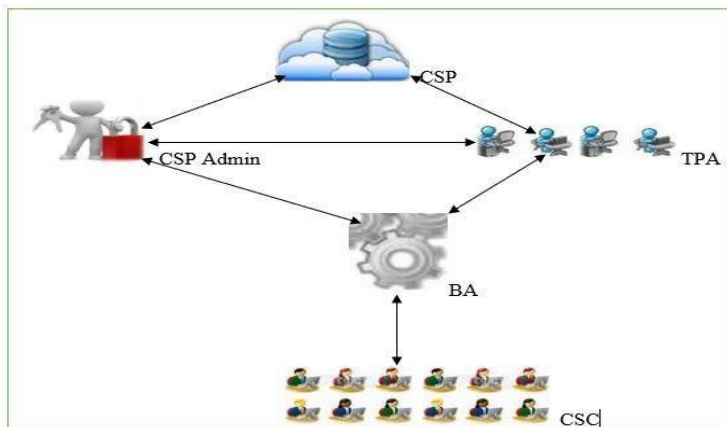


Figure 4: CDSART-BA Model

D. CDSART-BA Mathematical Model

Mathematically, let the initial client data cd_0 , the initial stored cloud data sd_0 , and the subsequent client data cd_n and subsequently stored cloud data sd_n be variables for auditing security and transaction between client or cloud service user and cloud service provider. And given that, the number of packets transmitted is greater or equal to zero.

Security is achieved when;

$$\sum_{i=0}^n cd = \sum_{i=0}^n sd \dots \dots \dots \text{equation 1}$$

Security is breached when;

$$\sum_{i=0}^n cd \leq \sum_{i=0}^n sd \dots \dots \dots \text{equation 2}$$

Security is also breached when;

$$\sum_{i=0}^n cd \geq \sum_{i=0}^n sd \dots \dots \dots \text{equation 3}$$

The equation above shows three scenarios that trigger a security audit report to the client. The equation one (1) describes a good security report. Equation two (2) describes data loss scenario and equation three (3) explains data intrusion and unauthorized addition while on transit.

Discussion of Survey Outcomes and Results

This review employs use of google scholar as a gateway into academic research databases, for the survey of resourceful research publications.

Existing techniques include; time-based one-time password (TB-OTP) and automatic blocker protocol (ABP) (El-Booz et al 2016). The method used in (El-Booz et al 2016) secures client data against the access of CSPs and TPAs without authorization using TB-OTP and ABP. So, an audit report cannot be generated without the permission of the client. Sending permission requests to individual cloud users improves security of client data but increases overhead.

Merkle Tree and Blockchain based cloud data (Mohan and Gladston, 2020) effects timespan and enables cloud users to verify that TPAs performed their auditing services as at when due. Centralized cloud information accountability with bat key generation algorithm (CCIA-BKGA) framework in cloud computing environment (Punitha and Indumathi, (2019). This

approach applies use of the bat algorithm to generate encryption keys to authenticate and account for access into cloud user data thereby adding to processing requirements of computing devices.

In contrast, the proposed CDART-BA uses a bat algorithm to trigger audit reports to cloud users, CSPs and TPAs, for an all-inclusive security mitigation strategy. The summary of survey outcomes and results is presented on table 2.

Table 2: Survey outcomes and Results

1.	Delitzscher et al., (2012)	IEEE	Validating cloud infrastructure changes by cloud audits.	The model is for Service provider-side audit reports.
2.	Ko et al., (2011)	IEEE	TrustCloud: A framework for accountability and trust in cloud computing.	Service provider oriented.
3.	Shankarwar and Pawar, (2015).	Springer	Security and privacy in cloud computing: A survey.	Research findings are yet to be implemented.
4.	El-Booz, Attiya, and El-Fishawy (2016).	EURASIP	Automatic blocker protocol and one-time password for	Did not affect the audit report.

			secure cloud storage.	
5.	Fahad, Ahmed and Kahar (2018).	Springer	Combines artificial neural network and whale swarm to detect intrusion cloud computing.	The focus of the research is detection and not audit reports for improved security strategies.
6.	Valarmathi, and Sheela, (2019).	Springer	Applies bat algorithm and particle swarm optimization for task scheduling in cloud environments.	Basically about the application of the bat algorithm and did not utilize the algorithms for the audit report.
7.	Sreeram and Vuppala, (2019).	Elsevier	HTTP flood attack detection in application layer using machine learning metrics and bat inspired algorithm for the	The research did not solve the audit report requirement.

			detection of HTTP flood attacks.	
8.	Sagnika, Bilgaiyan and Mishra, (2018).	Springer	Scheduling of data workflow in cloud computing using bat algorithm.	The aim of the paper is about load balancing and not audit report.
9.	Punitha and Indumathi, (2019).	Springer	The paper centralized the authentication of information with keys generated using the bat algorithm.	The implementation of the framework.
10.	Yakubu et al., (2019).	Springer	Security challenges in fog-computing environments.	Developing a model that improves the security in fog-computing.
11.	Adebesin et al., (2021).	Nigeria Computer Society	A framework for mitigating phishing attacks.	The implementation of the framework in a cloud environment.

12.	Umoren, Inyang, and Gilean (2021).	Nigeria Computer Society	Security threats prediction using Bayesian network algorithm in banks digital channels.	Developing the system for cloud environments.
13.	Ismail and Islam, (2020).	Nigeria Computer Society.	Framework for cloud security and audit.	Authors do not specify the security algorithm proposed in the framework.
14.	Mohan and Gladston, (2020).	ACM	Merkle tree and Blockchain-based cloud data auditing.	The use of blockchain technology has associated high computational power requirements.
15.	Raja Sree, and Mary Saira Bhanu, (2020).	Springer	Detection of HTTP flooding attacks in the cloud using fuzzy bat clustering.	Affecting audit report in the model for improved security.

16.	Ullah, Nawî and Khan, (2020).	INDERSCIEN CE	Load balancing using bat algorithm.	Development of the system model.
17.	Ngumbi and Wasike, (2022).	Kirinyaga University	ICT use in teaching and learning.	Improving the security of cloud environments for teaching and learning.
18.	Mutheu and Wasike, (2020).	Kirinyaga University	Decision Support (IDS) in Software Risk Management Based on Data Mining, Rough Sets and Decision Theory.	Applying decision theory in cloud data audit reports.

Conclusion

In this review, the authors were able to contribute to research on cloud data security with respect to auditing reports that involve cloud service clients as important stakeholders in the quest for the confidentiality, integrity and availability of cloud data and services. The review is able to achieve system design and mathematical models from the literature review of related work. The authors look forward to implementing the prototype of cloud data security report techniques using bat algorithm (CDSART-BA), to send timely audit reports to cloud users about the security situation of their data online. This is hoped to be achieved in Google Firebase, a google cloud computing lab environment that supports the use of backend codes.

References

1. Adebessin A.A, Sodiya A.S, Orunsolu A.A, Lawal O.A, Kareem S.O. (2021). A Proposed Distributed Anti-Phishing Framework for mitigating Cyber-attacks in Smart Environments. International Journal of Information Security, Privacy on Digital Forensics. 5(2). Nigeria Computer Society
2. Doelitzscher F., Fischer C., Moskal D., Reich C., Knahl M., Clarke N. (2012). Validating cloud infrastructure changes by cloud audits. In 2012 IEEE Eighth World Congress on Services. 377-384. IEEE.
3. El-Booz S. A., Attiya G., El-Fishawy N. (2016). A secure cloud storage system combining time-based one-time password and automatic blocker protocol. EURASIP Journal on Information Security. (1) 1-13.
4. Fahad A. M., Ahmed A. A., Kahar M. N. M. (2018). Network intrusion detection framework based on whale swarm algorithm and artificial neural network in cloud computing. In International Conference on Intelligent Computing and Optimization 56-65. Springer, Cham.
5. Ismail U. M., Islam S. (2020). A unified framework for cloud security transparency and audit. Journal of Information Security and Applications, 54, 102594. Nigeria Computer Society.
6. Mohan A. P., Gladstone A. (2020). Merkle tree and Blockchain-based cloud data auditing. International Journal of Cloud Applications and Computing (IJCAC). 10(3): 54-66. ACM.
7. Mutheu R., Wasike J. (2020). Intelligent Decision Support (IDS) in Software Risk Management Based on Data Mining, Rough Sets and Decision Theory. African Journal of Science, Technology and Engineering (AJSTE), 1(1):82-96. Kirinyaga University.
8. Ngumbi E., Wasike J. (2022). Use of ICT in the Classroom: The Teacher and Learner. African Journal of Science, Technology and Engineering (AJSTE). 3(1):34-48. Kirinyaga University.
9. Punitha A., Indumathi G. (2019). Centralized cloud information accountability with bat key generation algorithm (CCIA-BKGA) framework in cloud computing environment. Cluster Computing, 22(2):3153-3164. Springer.

10. Sagnika, S., Bilgaiyan, S., Mishra, B. S. P. (2018). Workflow scheduling in a cloud computing environment using the bat algorithm. In Proceedings of first international conference on smart system, innovations and computing Springer, Singapore. (149-163).
11. Shankarwar M. U., Pawar A. V. (2015). Security and privacy in cloud computing: A survey. In Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) (1-11). Springer, Cham.
12. Sree T.R., Bhan, S.M. (2020). Detection of HTTP flooding attacks in the cloud using fuzzy bat clustering. *Neural Computing and Applications*, 32(13):9603-9619. Springer.
13. Sreeram I., Vuppala V. P. K. (2019). HTTP flood attack detection in application layer using machine learning metrics and bio inspired bat algorithm. *Applied computing and informatics*, 15(1): 59-66. Elsevier
14. Ullah, A., Nawli, N. M., and Khan, M. H. (2020). BAT algorithm used for load balancing purpose in cloud computing: an overview. *International Journal of High Performance Computing and Networking*, 16(1):43-54. INDERSCIENCE
15. Umoren I., Inyang S., Gilean O. (2021). Bayesian Network Algorithm for Predictive Modelling of Cyber Security for Efficient Bank Channels Digitalization. *International Journal of Information Security, Privacy on Digital Forensics* 5(2).
16. Valarmathi R., Sheela T. (2019). Ranging and tuning based particle swarm optimization with a bat algorithm for task scheduling in cloud computing. *Cluster Computing*, 22(5): 11975-11988.
17. Yakubu, J., Abdulhamid, S. I. M., Christopher, H. A., Chiroma, H., and Abdullahi, M. (2019). Security challenges in fog-computing environment: a systematic appraisal of current developments. *Journal of Reliable Intelligent Environments*, 5(4): 20

BIG DATA ANALYTICS FOR SMES' PERFORMANCE SUSTAINABILITY IN POST-COVID-19 KENYA

Musyoka C., Wanjohi. P.

Department of Pure and Applied Sciences

Kirinyaga University, Kenya

Correspondences: cmusyoka@kyu.ac.ke

Abstract

Small and medium-sized businesses (SMEs) play a critical role in a nation's economy, contributing significantly to its wealth and fostering innovation. Globally, they account for up to half of all jobs and 90% of all businesses. However, SMEs often grapple with limited access to credit from suppliers, compounded by liquidity challenges, decreased sales, and defaults, as supply chains struggle to secure credit. Factors like advance payments, penalties for delayed credit payments, and recurring expenses further exacerbate the vulnerability of Smashing the context of Kenya's post-COVID-19 landscape. This study aimed to explore the potential of Big Data Analytics and Data Science in sustaining SMEs' performance. Drawing from theories such as Complex Adaptive System and Strategic Choice Theory, a descriptive survey design was employed, on a target population of 287 managers of SMEs in each sub County, in Nairobi. Employing a stratified sampling method, a total of 260 respondents were interviewed. Data was analyzed using descriptive statistics, including frequencies, percentages, mean, and standard deviation, while inferential statistics like multiple regression and Pearson correlation were used to examine relationships between variables. The study revealed that Business Intelligence, with a mean score of 3.9 (std. dv = 0.851), and Machine Learning, with a mean of 3.7 (std. dv = 0.928), both had a positive impact on SMEs' sustainability, with an overall average mean of 3.8 (std. dv = 0.8895). Similarly, Data Analytics, comprising Predictive Analytics (mean = 3.73, std. dv = 0.850) and Prescriptive Analytics (mean = 3.85, std. dv = 0.684), positively influenced SMEs' performance, with an average mean of 3.79 (std. dv = 0.767). These findings underscore the potential of Data Science drivers like Business Intelligence and Machine Learning in helping SMEs tackle unforeseen challenges in competitiveness. The study further highlights the importance of implementing a robust legal framework to safeguard data in the context of Data Analytics, particularly in predictive and prescriptive analysis, as a means to enhancing SMEs' performance, survival, and growth in the post-COVID-19 era.

Keywords: Big Data Analytics, Data Analytics, Business Intelligence, Machine Learning, Predictive Analytics

1.0 Introduction

Small and medium-sized businesses (SMEs) contribute to approximately half of all jobs and 90% of businesses worldwide but were significantly affected globally due to COVID 19 pandemic because of their small size and limited resources (Kumar *et al.*, 2021). They continue to face limited credit availability from creditors, who also face liquidity problems, lower sales, and higher rates subjecting the sector to severe financial constraints. The sector continues to face severe financial constraints as a result of supply chains not receiving credit from suppliers. A decrease in household income of approximately 36% was experienced by 95% of the entrepreneurs. Household income continues to be impacted by lower sales and family members losing jobs or seeing their wages reduced.

According to KBA MSMEs Survey Report, (2021) 35% of them missed or delayed loan repayments. Thus more small and medium-sized enterprises (SMEs) required financing to continue operating yet business owners struggle to obtain the necessary liquidity to revive and flourish. The study indicated that the majority of merchants accepted digital payments before the pandemic. However, a significant increase in the use of digital financial services resulted from exclusion of transactions exceeding 1,000 Ksh. (10 USD). Thus between February and October 2020, the Central Bank of Kenya reported that the monthly volume of P2P transactions increased by 87 percent. The volume of transactions below Ksh. 1,000 (USD10) increased by 114 percent during this time, and 2.8 million more customers began using mobile money, resulting in significant growth in digital business transactions. A slight improvement of 8% was observed among SMEs that use partnerships or an e-commerce model to mitigate the negative effects of COVID-19. More SMEs in rural areas – 89 percent compared to 62 percent in urban areas have yet to adopt any digital channel to grow their businesses and adapt to the shifting business environment (KBA MSMEs Survey Report, 2021).

Majority of SMEs are microfinance institution customers who have drained their reserve funds and offered resources to relieve the extreme monetary effect of the pandemic but still ran the risk of having their credit limits reduced as their savings decrease, making them more vulnerable and less able to handle unexpected events in the future. In order to keep their businesses afloat, SMEs have been learning how to diversify their sources of income (KBA MSMEs Survey Report, 2021). Thus, have been employing a variety of

strategies to stay afloat. These strategies include seeking credit and using loan proceeds to infuse funds into their businesses while also prioritizing business continuity and seeking external support to recover. Vitari, and Raguseo (2020) stated that Data Analytics will help SMEs to survive in today's dynamic business environment based on Predictive Analytics and Prescriptive Data Analytics to enable them make more data-driven decisions based on historical and present data and to predict future business dynamics and respond as appropriate.

Statement of the problem

In the wake of the COVID-19 pandemic, Kenya's small and medium-sized enterprises (SMEs) faced formidable obstacles in their quest for sustained performance and long-term viability. The pandemic has brought havoc on traditional business models and consumer behavior, necessitating a rapid adaptation of SMEs to the ever-changing business landscape. However, a significant impediment lies in the fact that many SMEs in Kenya lack the essential tools and strategies required to harness the potential of big data analytics effectively. This deficiency hampers their ability to make data-driven decisions that could bolster their performance and ensure sustainability in a post-pandemic world (Chumba *et.al.*, 2020).

Small and Medium Enterprises (SMEs) face a critical challenge in that their ability to navigate unforeseen situations is hindered by limited innovative technological capacities and underdeveloped research and development capabilities (Nath and Agrawal, 2020). These challenges become particularly pronounced during and post the COVID-19 pandemic. Survivability for SMEs therefore hinges on their capacity to address these problems by fortifying technological capabilities, especially in customer relations management (CRM). Additionally, there is a pressing need for SMEs to establish a flexible work system, allowing employees to work seamlessly from any location at any time. Failure to address these technological shortcomings poses a significant obstacle to SMEs' adaptability and sustainability in the face of ongoing and unforeseen challenges.

These challenges are compounded by various factors, including financial constraints, limited access to advanced technology, and lack of awareness and expertise in the realm of big data analytics. Additionally, SMEs in Kenya are operating in an intensely competitive environment, where adaptability and responsiveness to shifting market

dynamics are critical for survival. Without the capacity to leverage big data analytics, SMEs may struggle to identify emerging trends, understand customer preferences, and optimize their operational processes, potentially leaving them disadvantaged

Objective of the study

To establish the extent to which Big Data Analytics influence sustainability and performance improvement of small and medium-sized enterprises (SMEs) in Kenya after COVID-19 pandemic.

Research Question

The research was guided by the following research question:

To what extent does applications of Big Data Analytics influence the sustainability and performance improvement of small and medium-sized enterprises (SMEs) in Kenya the Post-COVID-19 pandemic?

Empirical Review

Maroufkhani et al. (2020) provides compelling insights into the transformative potential of BDA on SME sustainability and performance. The study effectively highlights how leveraging extensive and diverse datasets empowers SMEs to gain valuable insights into customer behaviors, identify market trends, and optimize operational efficiency. The emphasis on data-driven decision-making and strategic resource allocation underscores BDA's role as a catalyst for overall business improvement. The recognition of real-time analysis as a tool for proactive problem-solving and uncovering growth opportunities is commendable. Given the challenges posed by the post-COVID-19 era, the study rightly emphasizes BDA's contribution to equipping SMEs with the resilience and agility required to navigate dynamic market conditions. However, a critical examination of the study should consider the specifics of the proposed adoption model, potential limitations, and the generalizability of findings across different SME contexts to ensure the robustness and applicability of the proposed insights. Additionally, a deeper exploration of the challenges and ethical considerations associated with BDA adoption in SMEs would contribute to a more comprehensive understanding of its implications.

According to Nasrollahi *et al.*, (2021) study investigates the impact of big data adoption (BDA) on the performance of small and medium enterprises

(SMEs), presenting a comprehensive model. The research unveils BDA's multifaceted influence on SMEs, positively affecting both operational (efficiency, productivity) and economic (profitability, revenue) performance. Notably, no direct association was found between BDA and social performance. The study highlighted the mediating role of operational performance, where improvements in efficiency due to BDA contributed to enhanced economic performance. The model is conceptualized as a pyramid, with BDA components forming the base and arrows indicating causal relationships between layers. The findings emphasize BDA's potential for SMEs to streamline operations, make data-driven decisions, and gain a competitive edge. While acknowledging limitations, such as a small sample size and focus on Iranian SMEs.

Utilization of big data for informed decision-making across various sectors, as highlighted by Jin *et al.*, (2020), presents a significant contemporary challenge. While the literature consistently emphasizes the substantial value and competitive advantages organizations can derive from accurate information, there is a potential downside to this increasing dependence on big data. Businesses are actively seeking to leverage big data, aiming to enhance decision-making processes and introduce new technologies for innovative data utilization and knowledge discovery (Storey and Song, 2021). However, the problem lies in the potential overreliance on big data, with organizations possibly neglecting the importance of human judgment and qualitative insights. Relying solely on data-driven decisions may overlook contextual nuances and intangible factors crucial for comprehensive decision-making. Additionally, the challenge extends to ethical considerations, including issues related to data privacy, security, and potential biases within the datasets. The rush to adopt big data solutions without addressing these concerns could lead to unintended consequences and undermine the intended benefits of enhanced decision-making and business value. Therefore, a critical examination of the ethical and human-centric aspects of big data utilization is imperative to ensure responsible and effective implementation across businesses of all sizes.

Sangpetch and Ueasangkomsate's (2023) studied the intricate interplay of big data analytics (BDA), circular economy (CE), and sustainable performance for small and medium-sized enterprises (SMEs). The research uncovers BDA's role as a potent enabler of CE practices within SMEs, utilizing data analysis on material flows, resource consumption, and product lifecycles to inform strategies for

product redesign, waste reduction, and closed-loop systems. The study underscores how embracing CE principles directly enhances SMEs' sustainable performance, reducing environmental impact, improving resource efficiency, and enhancing brand reputation.

The existing literature on Big Data Analytics (BDA) for SMEs' Performance Sustainability in the post-COVID-19 era underscores its transformative potential, emphasizing operational efficiency, market prediction, and long-term sustainability. However, a critical analysis reveals a notable gap in empirical research, particularly in understanding the specific challenges and opportunities faced by small and medium-sized enterprises (SMEs) in the aftermath of the pandemic. The literature lacks comprehensive insights into the practical implementation barriers, adoption challenges, and contextual factors influencing the successful integration of BDA tools within the distinctive operational structures of SMEs navigating the post-COVID-19 business landscape. Additionally, while broad benefits are acknowledged, there is a need for more focused exploration of sector-specific applications and varying impacts across different industries within the SME ecosystem. Closing these gaps would provide valuable insights for SMEs, policymakers, and researchers seeking to foster resilience and sustainability in the evolving post-pandemic business environment

Theoretical Framework

It highlights literature that corresponds with the study's general and specific objectives, fundamental theories correlated to the research and research gaps.

Complex Adaptive System theory provides a perspective for comprehending complex and dynamic systems characterized by numerous interconnected elements, whose behavior can often be unpredictable, as exemplified in various contexts, including businesses. It asserts that these systems are in a continual state of adaptation to their surroundings and are far from static entities. In the context of the study on Data Analytics for SMEs' Performance Sustainability in Post-COVID-19 Kenya, CAS theory gains relevance by recognizing that SMEs function within intricate and swiftly shifting environments. According to Tammissalo (2020), applying CAS theory to SMEs, one gains insights into how data analytics can be harnessed to navigate uncertainties, respond effectively to evolving market dynamics, and uphold sustained performance. This perspective underscores the study's focus on fostering adaptability, agility, and flexibility

in SMEs' adoption of data analytics to bolster their sustainability amid the post-pandemic challenges

The Strategic Choice Theory (SCT), pioneered by Richard Cyert and James G. March postulates that organizational decisions are shaped by a limited pool of information and are heavily influenced by the goals, values, and perceptions of key decision-makers within the organization. SCT underscores the concept of bounded rationality, highlighting that organizations make choices based on cognitive limitations and available information. Hazen, et.al (2020) in their study stated that SCT holds relevance as it sheds light on the decision-making processes of SMEs, particularly concerning adoption of data analytics. It suggests that SME leaders' cognitive processes and perceptions regarding the advantages and risks associated with data analytics will play a pivotal role in determining its adoption. SCT thus informs the study by accentuating the significance of leadership and organizational decision-making in the context of data analytics adoption and its potential influence on SME performance sustainability

Conceptual Framework

This study investigates to what extent do applications of Big Data Analytics, with a focus on Post-COVID-19 circumstances, influence sustainability and performance improvement of small and medium-sized enterprises (SMEs) in Kenya

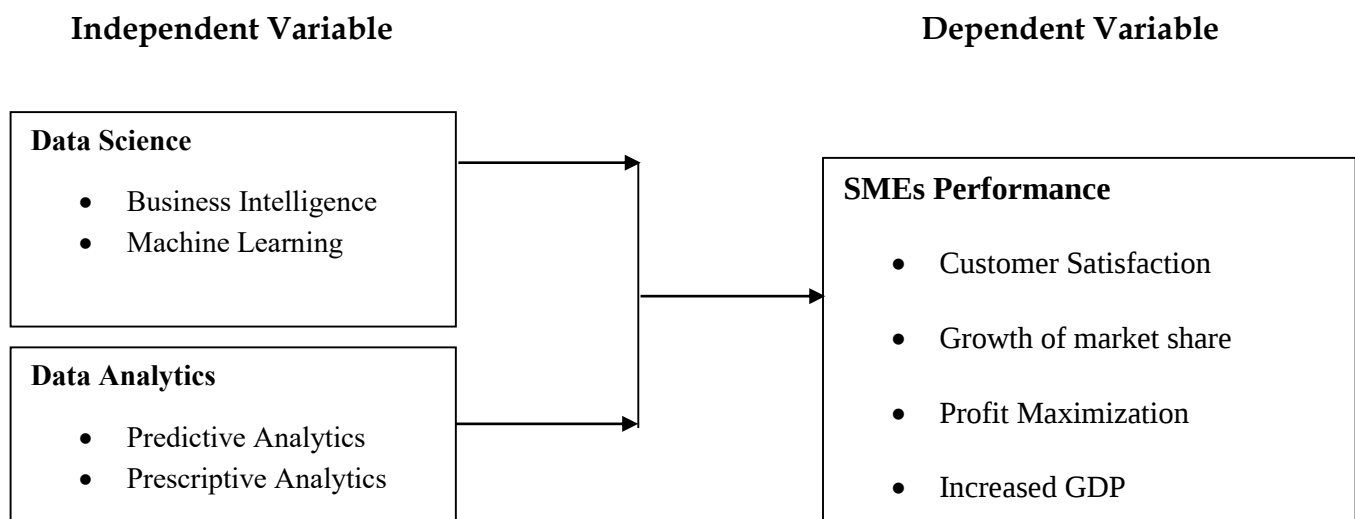


Figure 1: Conceptual Framework

Research Methodology

A descriptive survey design was used in this study. The target population was 287 respondents comprising managers of SMEs in Nairobi. The sample for the study was 260 respondents selected using the stratified sampling technique. Questionnaires were used to collect data as shown in Table 3.1.

Table1 Sampling size

Sub-County	Target Population	Sample Size	Percentage (%)
Eastlands	16	14	5.57
Dagoretti North	14	10	4.88
Lang'ata	20	19	6.97
Kibra	17	14	5.92
Roysambu	20	19	6.97
Kasarani	11	10	3.84
Ruaraka	20	19	6.97
Embakasi South	14	14	4.88
Embakasi North		14	6.27
Embakasi Central	18	13	4.53
Embakasi East	13	14	5.92
Embakasi West	17	19	6.62
Makadara	19	10	4.18
Kamuknji	12	19	6.97
Mathare	20	14	5.57
Starehe	16	19	6.97
Dagoretti North	20	19	6.97
Total	287	260	100

Source: Author

Research Findings and Discussions

Table 2 presents data including two key variables, Business Intelligence and Machine Learning, along with their respective mean and standard deviation values. The mean value for Business Intelligence was 3.9 (std. dv = 0.851) indicating that on average, the SMEs in the study exhibited a relatively high level of Business Intelligence utilization in their operations. A mean value above 3 suggests that the majority of SMEs in the sample actively employed Business Intelligence practices to gather insights and make informed decisions. The mean value for Machine Learning is 3.7 (std. dv = 0.851) indicating that, on average, SMEs in the study exhibited a reasonably high level of Machine Learning adoption. While the mean is slightly lower than that of Business Intelligence, it still reflects a notable utilization of Machine Learning techniques.

Statement	Mean	Std Dev.
Business Intelligence	3.9	0.851
Machine Learning	3.7	0.928
Average	3.8	0.8895

Table 3 below provides descriptive results that focus on the impact of Data Analytics on sustainability of SMEs' performance in Kenya following the COVID-19 pandemic. The table includes two crucial variables, Predictive Analytics and Prescriptive Analytics, along with their respective means and standard deviation values, offering insights into the extent of their utilization among the SMEs in the study. The mean score of 3.73 signifies that the SMEs have a relatively strong adoption of Predictive Analytics. This suggests that a considerable number of SMEs actively used Predictive Analytics to forecast future trends to inform their decision-making and strategic planning. The standard deviation of 0.850 implies moderate variability in the responses regarding Predictive Analytics. While many SMEs employ this practice, there may be some diversity in the extent of its application across the sampled SMEs.

The mean score of 3.85 indicates that the SMEs robustly adopted Prescriptive Analytics to provide specific recommendations and actions based on data analysis. The standard deviation of 0.684 suggests relatively low variability in the responses concerning

Prescriptive Analytics implying that the majority of SMEs exhibited a consistent pattern of adopting Prescriptive Analytics.

Statement	Mean	Std Dev.
Predictive analytics	3.73	0.850
Prescriptive Analytics	3.85	0.684
Average	3.79	0.767

Table 3: Descriptive Results Data Analytics and Sustainability of SMEs' Performance In Kenya after COVID-19

5.0 Conclusions and Recommendations.

Descriptive results presented on Tables 2 and 3 shed light on the utilization of Data Science and Data Analytics, respectively, in enhancing the sustainability of SMEs' performance in Kenya post-COVID-19. The findings indicate that the SMEs demonstrated notable adoption of Business Intelligence, with a mean score of 3.9, suggesting active utilization to gather insights and make informed decisions. Similarly, the study reveals a reasonably high level of Machine Learning adoption, with a mean score of 3.7. Moving to Data Analytics, the results highlight a strong average adoption of Predictive Analytics (mean of 3.73), indicating widespread use for forecasting future trends, albeit with some variability. These SMEs showcase a robust adoption of Prescriptive Analytics, with an average mean score of 3.85, suggesting consistent utilization for providing specific recommendations and actions based on data analysis. These findings collectively underscore the significance of advanced data-driven techniques in SMEs, portraying a positive landscape for leveraging Data Science and Analytics to enhance sustainability in the post-COVID-19 business environment.

References

- Cumba L. T., Huang X., Kholaf M. M. N. H. K. (2020). The impact of digital transformation on big data analytics and firm's sustainability performance in a post-pandemic era. *Human Systems Management*, (Preprint), 1-22.
- Hazen B. T., Skipper J. B., Ezell J. D., Boone C. A. (2020). Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers and Industrial Engineering*, 101:592-598.
- Jin J., Liu Y., Ji P., Liu H. (2020). Understanding big consumer opinion data for market-driven product design. *International Journal of Production Research*, 54(10):3019-3041.
- KBA Micro, Small and Medium Enterprises (MSMEs) Survey Report, 2021. <https://www.centralbank.go.ke/2023/06/13/2022-survey-report-on-msme-access-to-bank-credit/>.
- Kumar A., Sharma K., Singh H., Naugriya, S. G., Gill, S. S., Buyya, R. (2021). A drone-based networked system and methods for combating coronavirus disease (COVID-19) pandemic. *Future Generation Computer Systems*, 115, 1-19.
- Maroufkhani P., Wan Ismail, W. K., Ghobakhloo M. (2020). Big data analytics adoption model for small and medium enterprises. *Journal of Science and Technology Policy Management*. 11(4): 483-513.
- Nasrollahi, M., Ramezani, J., Sadraei, M. (2021). The impact of big data adoption on SMEs' performance. *Big Data and Cognitive Computing*, 5(4):68.
- Nath, V., Agrawal, R. (2020). Agility and lean practices as antecedents of supply chain social sustainability. *International Journal of Operations and Production Management*, 40(10):1589-1611.
- Sangpetch, P., Ueasangkomsate, P. (2023). The Influence of the Big Data Analytics and Circular Economy on the Sustainable Performance of SMEs. *Thammasat Review*, 26(1): 114-139.
- Storey V. C., Song I.-Y. (2017). Big data technologies and management: What conceptual modeling can do. *Data and Knowledge Engineering*, 108:50-67.
- Tammisalo T. (2020). Harnessing big data for business purposes in Finnish SMEs: adaptive marketing capabilities perspective.
- Vitari C., Raguse, E. (2020). Big data analytics business value and firm performance: linking with environmental context. *International Journal of Production Research*, 58(18): 5456-5476.



Kirinyaga University

**AFRICAN JOURNAL OF SCIENCE, TECHNOLOGY AND ENGINEERING
(AJSTE)**

Published by:

P.O BOX 143 - 10300, KERUGOYA, KENYA

MOBILE +254709742000/+254729499650

Email: info@KyU.ac.ke

Website: www.KyU.ac.ke