proceedings

Proceediings

An assessment on the hidden ecological factors of the incidence of malaria[†]

Babagana Modu[1], A. Taufiq Asyhari[2] and Yonghong Peng[3]

[1]School of Electrical Engineering and Computer Science, University of Bradford, UK; modubabagana70@yahoo.com

[2]Centre for Electronic Warfare, Information and Cyber, Cranfield University, Defense Academy, Shrivenham, SN6 8LA, UK; taufiq-a@ieee.org

[3]Faculty of Computer Sciences, University of Sunderland, UK; Yonghong.Peng@sunderland.ac.uk

* Correspondence: modubaabgana70@yahoo.com; Tel.: +447435382027

† Presented at the IS4SI 2017 Summit DIGITALISATION FOR A SUSTAINABLE SOCIETY, Gothenburg, Sweden, 12-16 June 2017.

Published: date: 9 June 2017

Abstract: Confounding effects of climatic factors poses a greater influence in the incidence of malaria and its considerable changes over space and time. In this study, we explore a new framework for assessment and identification of hidden ecological factors to the incidence of malaria with cognizance to its ecosystem. Partial least squares path modeling (PLS-PM) and exploratory factor analysis (EFA) techniques are employed for the identification of hidden ecological factors from climate data and the reported cases of malaria in Ejisu-Juaben, Ghana. Three hidden factors are identified: Factor I is related to minimum temperature and relative humidity, Factor II is related to maximum temperature and solar radiation and Factor III is related to precipitation and wind speed, respectively. Factor I is identified as the most influential hidden ecological factor of malaria incidence in the study area, as evaluated by communality and Dillon-Goldstein's indices. This result is also validated using cross-correlation, where Factor I indicate positive association with malaria incidence at Lag1 (a month in advance). We conclude that minimum temperature and relative humidity are accountable for high malaria incidence and should be given consideration in strategizing policy for prevention and control in the study area.

Keywords: Malaria incidence; Climatic factors; Structural equation modeling; Partial least square model and Hidden factors.

1. Introduction

Malaria is vectored by mosquito [1], and it is considered as one of the most serious infectious disease causing public health problem in the world. It is estimated that about two-third of the world population are at risk, with resultant deaths close to a million annually [2]. The prevalence of malaria can be attributed to climate factors, which is, in some cases, worsened by human factors through: poor sanitation, overwhelmed sewage and deforestation. All these identifiable factors are found contributing to the prevalence of malaria, which, in some ways, significantly imposes a greater challenge to human life.

Climate factors are the drivers of malaria transmission, but the ecological relationship that exists among them is still paucity, particularly in the proposed study area (Ejisu-Juaben, Ghana), which is characterized by bi-modal rainfall pattern. Recent works by [3]-[5] used meteorological variables together with malaria incidence data and established some models that predict the malaria incidence. Similarly, reference [6] used regression and correlation analysis between malaria incidence and meteorological variables and determined the trend of malaria incidence. Another study by [7] used time-series modeling and lagged-regression of climate data and the reported malaria incidence cases. Their result showed that malaria incidence in the study area has significant correlation with relative humidity and whereas temperature and precipitation have negligible influence. This finding might particularly reveal that malaria incidence can be strongly influenced by relative humidity alone. However, the methodology fails to capture the existing dependency among the climate factors. In this study, we propose the use of partial least squares path modeling (PLS-PM) to analyze the causal relationship between meteorological variables such as: minimum average temperature, maximum average temperature, relative humidity, altitude, wind speed, precipitation and solar radiation, and explore their impacts on the occurrence of malaria incidence. By doing so, we aim to develop an integrated model that provides us with information to identify the hidden ecological factors that lead to the high malaria incidence in the study area. Furthermore, based on our findings, we aim to inform policy makers and public health workers with the best strategy for prevention and control of malaria transmission from ecological perspective.

The subsequent sections of this paper are presented as follows. In Section 2, we discuss the concept of ecosystem and malaria transmission and its health implications to the change in biodiversity. Section 3 presents the materials and methods which consist of study site and population, data collection and source and factor analysis technique used for construction of hypothetical structural equation models. Section 4 presents results and discussion. Section 5 concludes the paper by providing the summary of our findings.

2. Ecosystem impact to malaria transmission

The malaria ecosystem comprises four main components namely: human host, mosquito's vector, parasites and environmental condition (see Figure 1). These components are very dynamics in nature due to the inherent characteristics of ecology and the anticipatory change to biodiversity because of global warming. The works by [8]-[10] reported that ecological changes would adversely affect human health in some ways that are both obvious and obscure. However, the growing evidence also suggests that due to the rise in temperature because of the anticipated global warming, some previously unexposed regions of malaria transmission would have 50% chance of experiencing its due to the link between malaria incidence and ecological factors [11]. The relationship between environmental changes and human health cannot be overemphasized because of the inherent variability and complexity of human nature. In many circumstances, grasslands and forest are converted for agriculture to reduce communicable disease, including wetland drainage for prevention and control of malaria [10].

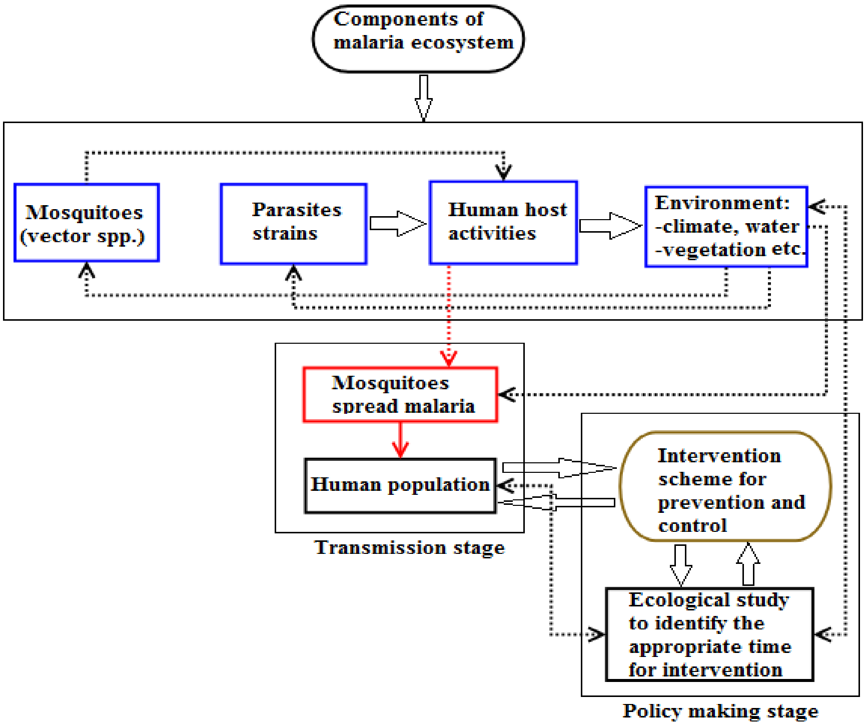

Figure1. Architecture of malaria transmission ecosystem for the prevention and control of its scourge.

These activities can either lead to an unintended negative health effects or succeeding in the designed purpose. Also, transforming forest to augment food production, which in the long run may leads to creating a suitable environment for disease-causing agents' e.g. mosquito for malaria transmission [12].

3. Materials and Methods

In this section, we present the general methodology used in data collection, technique for data presentation and the reference statistical tools for analysis.

3.1 Study site and population

Ejisu-Juaben Municipal has a population of 143,762 [13], lies within latitudes $1^o15^{|}$N and $1^o45^{|}$N also with longitudes $6^o15^{|}$W and $7^o00^{|}$W and occupies a land area of $582.5\,km^2$ [14]. The vegetation of the municipal is a typical semi-deciduous forest (see Figure 2a), with undulating topography and low altitude of about 240m–300m above the sea level [14]. Also, the rainfall pattern of the area is bi-modal (i.e., two distinct seasons in a year), characterized by major and minor rainfall. The major rainfall begins from March to July with average annual rainfall between 1200mm–1500mm, while the minor rainfall begins in September and tapers off in November with annual average of 900mm–1120mm. Usually, December through February is hot, dry and dusty with mean annual temperature 25^oC-32^oC, and the relative humidity is moderately high during the rainy seasons [14]. Figure 2b presents the map of Ejisu-Juaben Municipal, which is indicated by the large dark-shaded portion.

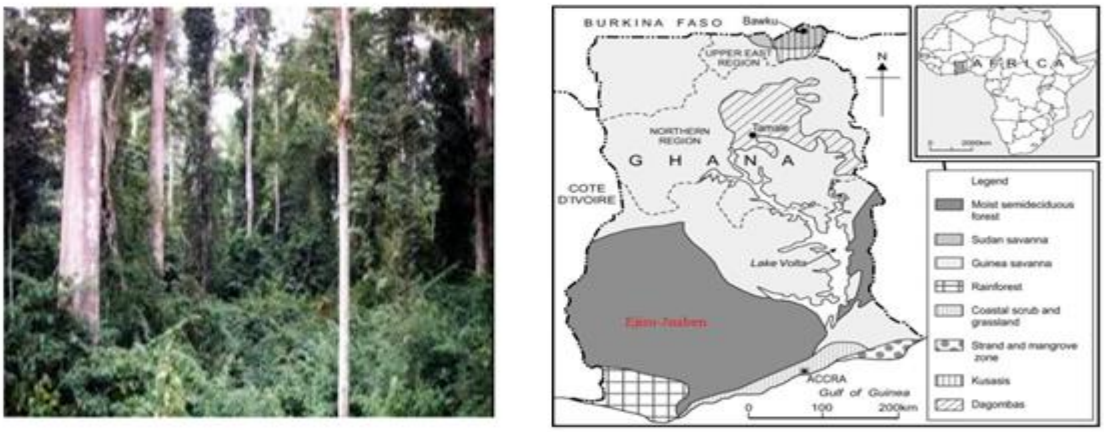

Figure 2. Representation of physical features of the area under the study describing: (a) The vegetation of study area is a typical semi-deciduous forest, and (b) The map of Ejisu-Juaben Municipal, the Ashanti Region of Ghana, retrieved in [14].

3.2 Data collection and source

A total of 85,627 confirmed diagnosed cases of malaria incidence over the period of five years (2009 to 2013) was retrieved from the literature in [15]. The distribution of malaria cases as reported in [15] of the study area shows an indication of high malaria incidence. We also use meteorological data obtained from satellite through the link http://globalweather.tamu.edu/. The boundary metrics of latitude 6.7989^{o}N to 6.6823^{o}S and longitude -1.5656^{o}W to -1.4186^{o}E are used to determine the location of the study area on the satellite globe map. Within the demarcation, one weather station was found, and used to generate data on the meteorological variables of interest.

3.3 Factor analysis

Exploratory factor analysis (EFA) is one of the techniques for factor analysis (FA). It is primarily used in statistics to describe variance among observed correlated variables in terms of potentially smaller number of unobserved variables, usually referred to as factors [17]. In this work, EFA was employed to search for confounding ecological factors that are latent [8], [17] from the set of observed meteorological variables.

3.4 Structural equation modeling

Structural equation model (SEM) is a very popular technique that has multidisciplinary applications. The SEM combines together both the measurement model and the structural model [18]-[20]. In Figure 3a, we present the complex hypothetical SEM showing causal relationship between malaria incidence and latent ecological factors together with their observed variables. We used ellipse shapes to represents latent factors, while the observed variables are represented by rectangular shapes. In the following we demonstrate the SEM technique using simple mathematical sketches in which the observed variables can modeled as a linear combination of the potential factors plus error terms. Mathematical representations for Figure 3b are given as follows:

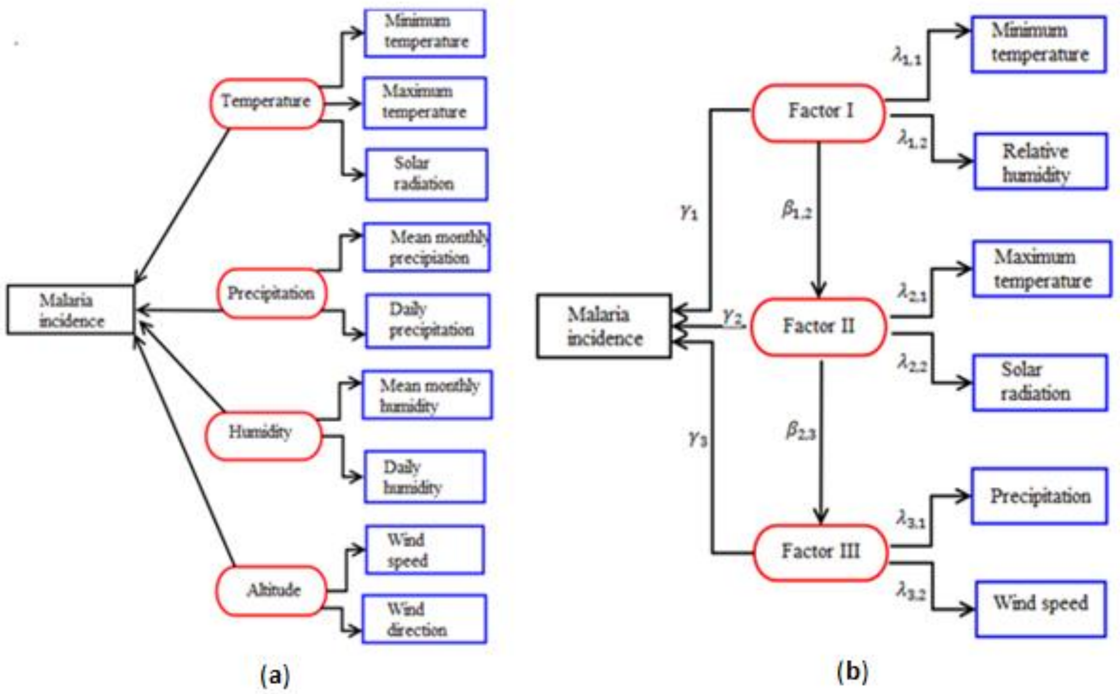

Figure 3. Representation of structural model showing: (a) The structural model that depicts the hypothetical relationship between malaria incidence and meteorological variables, and (b) The structural model that outlines the relationship among malaria incidence, hidden ecological factors and observed meteorological variables.

$$\text{Factor I} = \lambda_{1,1}(\text{minimum temperature}) + \lambda_{1,2}(\text{relative humidity}) + \beta_{1,2}(\text{Factor II}) + \gamma_1(\text{malaria incidence}) + e_1 \quad (1)$$

$$\text{Factor II} = \lambda_{2,1}(\text{maximum temperature}) + \lambda_{2,2}(\text{solar radiation}) + \beta_{2,3}(\text{Factor III}) + \gamma_2(\text{malaria incidence}) + e_2 \quad (2)$$

$$\text{Factor III} = \lambda_{3,1}(\text{precipitation}) + \lambda_{3,2}(\text{wind speed}) + \gamma_3(\text{malaria incidence}) + e_3 \quad (3)$$

3.5 Construction of PLS-PM

The technique called partial least squares path modeling (PLS-PM) or partial least squares structural equation modeling (PLS-SEM) was chosen due to its characteristics in terms of: small sample, non-normality, multi-dimension, and multicollinearity [21]-[23]. We extracted three hidden factors using EFA, and subsequently invoked PLS-PM and constructed SEM (see Figure 3b). The PLS-PM is a component-based estimation technique that uses iteration algorithm, separately analyzes the blocks of the measurement model and estimates the path coefficients in the structural model [21]. For the analysis, we used R software and performed the PLS-PM, in which the path weighting schemes were implemented for the inner estimate of the standardized latent variable. In the PLS technique, the default sample number for bootstrapping analysis was set to 500 selection. Also, the PLS-PM latent variable scores were expressed as a linear combination of their observed variables and treated as an error-free substitute for the observed variables [21].

4. Results and Discussion

The conceptual SEM presented in Figure 3 shows the hypothetical causal relationship between latent (hidden) variables and observed meteorological (manifest) variables to the occurrence of malaria incidence. For the extraction of confounding hidden variables, we performed factor analysis using exploratory factor analysis (EFA) [17]. From the results, three hidden factors were identified namely: Factor I (related to minimum temperature and relative humidity), Factor II (related to maximum temperature and solar radiation) and Factor III (related to precipitation and wind speed), respectively. The identified factors are accounted for 64% of the total variance, and at α = 5% level of significance, χ^2= 13.91, df = 8, p−value = 0.0841., This provides sufficient evidence to explain malaria incidence in the study area. We also used Kaiser Criterion and Cattel scree plots, which reconfirmed three hidden ecological factors to the incidence of malaria.

In Table 1, we present Pearson's cross-correlation between meteorological variables and occurrence of malaria incidence at various lag effects from 0 to 3 months. The result showed that maximum temperature, minimum temperature, and relative humidity were related to the malaria incidence at lagged effects of 1 month in advance except precipitation which has negative association in the study area. This result is attributed to the bi-modal rainfall which washes out egg and larval stages of mosquitos' life cycle. We found that the preceding result is consistent with other relevant studies on the influence of meteorological variables to the malaria incidence [24]. The 1 month time lag in the study area is sufficient to capture the pattern of malaria transmission for various strains of plasmodium parasites with definite lengths of extrinsic incubation period (EIP). The period usually takes about 10-15 days [25] and temporally varies over location, parasites species and climatic resolution. The other important summary statistics are also presented in Table 1, which shows the distributional pattern of the climate indicator of malaria incidence and variance inflation factor (VIF). This VIF use to describe how much multicollinearity (correlation between predictors) exists in a set of predictor variables. From Table 1, the minimum temperature and relative humidity are having VIF of 8.7919 and 9.0065 that gives high degree of multicollinearity factor. Other statistics like kurtosis and standard error describe the meteorological variables distribution.

Table 1. Cross-correlation between meteorological variables and malaria incidence together with

some other statistical indices.

Variables	Lag 0	Lag 1	Lag 2	VIF	Kurtosis	Standard error
Maximum temperature	0.284	0.321**	0.092	2.4096	5.48	0.38
Minimum temperature	-0.122	0.215**	-0.237	8.7919	2.07	0.33
Precipitation	-0.214	-0.292*	-0.155	1.4194	20.73	0.27
Relative humidity	-0.134	0.254**	-0.198	9.0065	1.42	0.02
Solar radiation	-	-	-	1.9000	6.73	0.50
Wind speed	-	-	-	1.3452	-0.58	0.04

* the negative association between meteorological variables and malaria incidence at lag 1.

** the positive association between meteorological variables and malaria incidence at lag 1.

In Table 2, we show the results of factors score estimates for path coefficients used in PLS-PM and evaluated using three different weighting schemes. We observed that Centroid (A) converges faster after 12 iterations, while factorial (B) and path weighting (C) converge after 15 iterations. This shows that Centroid is more robust than factorial and path-weighting schemes.

Table 2. Factor scores for path coefficients in the PLS-PM using three weighting schemes.

Measurement/ Structural model	Parameter	Estimate	Centroid (A)	Factorial (B)	Path-weighting (C)
Min. temp.← Factor I	$\lambda_{1,1}$	0.9479	0.9479	0.9495	0.9495
Rel. hum. ← Factor I	$\lambda_{1,2}$	0.9910	0.9910	0.9903	0.9903
Max. temp.←Factor II	$\lambda_{2,1}$	0.8816	0.8816	0.8675	0.8675
Solar rad. ← Factor II	$\lambda_{2,2}$	0.8735	0.8735	0.8873	0.8873
Precip. ← Factor II	$\lambda_{3,1}$	0.9849	0.9849	0.9852	0.9852
Wind sp. ← Factor IFactor I → Factor II	$\lambda_{3,2}$	0.0017	0.0017	0. 0031	0.0031
	$\beta_{1,2}$	-0.3248	-0.3248	-0.3302	-0.3302
Factor II → Factor III	$\beta_{2,3}$	-0.2774	-0.2774	-0.2690	-0.2690
Factor I → Mal. Inci.	γ_1	0.9700	-	-	-
Factor II → Mal. Inci.	γ_2	0.7700	-	-	-
Factor III → Mal. Inci.	γ_3	0.4900	-	-	-

Table 3 presents the results of bootstrapping sampling for outer loading of the observed variables and path coefficient of the latent variables in the PLS-PM. The results show that all outer loadings and path coefficients are significant at (α = 5%), except for the solar radiation, precipitation and wind speed which contain zero-point in the bootstrap confidence interval. Furthermore, the interaction effects of the Factors between (I and II, II and III) were also investigated and the results show that none of the combination is significant to the incidence of malaria in the study area.

Table 3. Bootstrapping test of outer loadings and path coefficients in the PLS-PM.

Measurement/Structural model	Parameter	Estimate	Bias	Standard error	95% Lower CI	95% Upper CI
Min. temp. ← Factor I	$\lambda_{1,1}$	0.9495	-0.0107	0.0442	0.8410	0.9940
Rel. humid ← Factor I	$\lambda_{1,2}$	0.9903	-0.0031	0.0132	0.9570	1.0000
Max.temp. ← Factor II	$\lambda_{2,1}$	0.8675	-0.0042	0.1037	0.6010	0.9880
Solar rad. ← Factor II	$\lambda_{2,2}$	0.8873	-0.0492	0.1790	0.3860	0.9780
Precip. ← Factor II	$\lambda_{3,1}$	0.9852	-0.1709	0.3975	-0.5500	1.0000
Wind sp. ← Factor III	$\lambda_{3,2}$	0.0031	0.1360	0.4018	-0.6000	0.9350
Factor I → Factor II	$\beta_{1,2}$	-0.3302	-0.0346	0.1675	-0.5680	0.1690
Factor II → Factor III	$\beta_{2,3}$	-0.2690	-0.0373	0.2269	-0.6260	0.3140

The decision for selecting the most influential hidden ecological factor to the incidence of malaria is based on the communality and Dillon-Goldstein's indices. Furthermore, Table 4 summarizes the results indicating some indices for selecting the hidden ecological factors to the high incidence of malaria in the study area. Among the three factors identified by EFA, we find that Factor I, indicated by minimum temperature and relative humidity, influences malaria transmission with communality index (0.94) and Dillon-Goldstein's ρ (0.97). This result is also consistent with the finding in [26], where positive association exists between temperature and occurrence of dengue. Factor II and Factor III appear to have less influence to the malaria incidence.

Table 4. The indices for selection of the ecological hidden factor of high malaria incidence in the study.

Factor	Reflective variables	Communality	Dillon-Goldstein's ρ
I	2	0.94^{***}	0.97^{***}
II	2	0.77	0.87

II	2	0.49	0.49

*** the most significant hidden factor indicator to the high incidence of malaria in the study area

5. Conclusion

This study has provided an overview on ecosystem for malaria modeling and proposed a new framework for the study of malaria transmission ecosystem for prevention and control of its scourge. Also, the framework would enable us assessing and identifying a hidden factor that leads to high malaria incidence in the study area. Our data analysis results have suggested that the minimum temperature and relative humidity, which are related Factor I, have positive association with the incidence of malaria in the study area. The other observed variables like maximum temperature, solar radiation, precipitation and wind speed, which are related to hidden Factor II and Factor III, appear to have mildly influenced malaria incidence.

Acknowledgments: The authors would like to thank National Centre for Environmental Prediction (NCEP), for their support in providing this work with relevant meteorological data. Also, the author B. Modu would like to thank Tertiary Education Trust Fund (tetFund), Nigeria for sponsoring his Ph.D. programme in University of Bradford, UK. No grant is dedicated for this research work.

Author Contributions: A.T.A. and Y.H.P. designed the experiments; B.M. performed the experiments; B.M. analyzed the data; B.M, A.T.A., and Y.H.P. contributed reagents/materials/analysis tools; B.M. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Lindsay, S.W., and M.H. Birley." Climate change and malaria transmission." *Annals of tropical medicine and parasitology* 90.6 (1996): 573- 588.

[2] World Health Organization. "Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 6 (2014-2015)." (2015).

[3] Zacarias, O. P., Bostrm, H. (2013, December). Comparing support vector regression and random forests for predicting malaria incidence in Mozambique. In Advances in ICT for Emerging Regions (ICTer), 2013 *International Conference on* (pp. 217-221). IEEE.

[4] Kumar, V., Mangal, A., Panesar, S., Yadav, G., Talwar, R., Raut, D., and Singh, S. (2014). Forecasting malaria cases using climatic factors in Delhi, India: a time series analysis. Malaria research and treatment, 2014.

[5] Ngarakana-Gwasira, E.T., Bhunu, C.P., Masocha, M., and Mashonjowa, E. (2016). Assessing the Role of Climate Change in Malaria Transmission in Africa. *Malaria research and treatment, 2016.*

[6] Nath, D. C., Mwchahary, D. D. (2013). Association between Climatic Variables and Malaria Incidence: A Study in Kokrajhar District of Assam, India: Climatic Variables and Malaria Incidence in Kokrajhar District. *Global journal of health science*, 5(1), 90.

[7] Modu B., Asyhari A.T., Peng Y.H., (2016). Data analytics of climatic factor influence on the impact of the malaria incidence. *IEEE Symposium series on Computational Intelligence*, Athens, Greece, Dec. 2016.

[8] Myers, S. S., Patz, J. A. (2009). Emerging threats to human health from global environmental change. *Annual Review of Environment and Resources*, 34, 223-252.

[9] Myers, S. S., Gaffikin, L., Golden, C. D., Ostfeld, R. S., Redford, K. H., Ricketts, T. H., Osofsky, S. A. (2013). Human health impacts of ecosystem alteration. *Proceedings of the National Academy of Sciences*, 110(47), 18753-18760.

[10] Bayles, B. R., Brauman, K. A., Adkins, J. N., Allan, B. F., Ellis, A. M., Goldberg, T. L., Ricketts, T. H. (2016). Ecosystem Services Connect Environmental Change to Human Health Outcomes. *EcoHealth*, 13(3), 443-449.

[11] The Potsdam Institute for Climate Impact Research and Climate Analytics, Turn-down the Heat Why a 4-degree Warmer World Must be Avoided, International Bank for Reconstruction and Development and World Bank, Washington, DC, 2012.

[12] de Castro, M. C., Monte-Mr, R. L., Sawyer, D. O., Singer, B. H. (2006). Malaria risk on the Amazon frontier. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2452- 2457.

[13] Nyarko, P. (2014). Population and Housing Census, District Analytical Report, Ejisu-Juaben Municipal. [online] https://www.citypopulation.de/php/ghana-admin.php?adm2id=0117. Available at: http://Population of Ejisu-Juaben district [Accessed 12 Jan. 2017].

[14] Addai, G. and Anyatewon Kwesi, D. (2014). 2010 population and housing census: district analytical report. 1st ed. Accra: *Ghana Statistical Service*.

[15] Appiah, S., T., H., O. and I.B., N. (2015). Times Series Analysis of Malaria Cases in Ejisu-Juaben Municipality. [online] http://www.ijstr.org/final-print/june2015/Times-Series-Analysis- Of-Malaria-Cases-In-Ejisu-juaben-Municipality.pdf. Available at: http://Time Series Analysis of Malaria Cases in Ejisu-Juaben Municipally [Accessed 12 Jan. 2017].

[16] Rummel, R.J. (1967). Understanding factor analysis. *Journal of Conflict Resolution*, 11(4), 444-480.

[17] Ruscio, J., Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological assessment*, 24(2), 282.

[18] Kline, R. B. (2015). Principles and practice of structural equation modelling. *Guilford publications*.

[19] Kelloway, E. K., Santor, D. A. (1999). [Using LISREL for Structural Equation Modeling: *A Researcher's Guide]*. *Canadian Psychology*, 40(4), 381.

[20] Byrne, B. M. (2013). Structural equation modelling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. Psychology Press.

[21] Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., Lauro, C. (2005). PLS path modeling. *Computational statistics data analysis*, 48(1), 159-205.

[22] Bagozzi, R.P., Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the academy of marketing* science, 40(1), 8-34.

[23] Li, X. X., Wang, L. X., Zhang, J., Liu, Y. X., Zhang, H., Jiang, S. W., Zhou, X. N. (2014). Exploration of ecological factors related to the spatial heterogeneity of tuberculosis prevalence in PR China. *Global health action*, 7.

[24] Srinivasulu, N., Gujju.Gandhi, B., Naik, R. and Daravath, S. (2013). Influence of Climate Change on Malaria Incidence in Mahaboobnagar District of Andhra Pradesh, India. [online] https://www.ijcmas.com/Archives/vol-2-5/N.

[25] Paaijmans, Krijn P., Andrew F. Read, and Matthew B. Thomas. "Under- standing the link between malaria risk and climate." *Proceedings of the National Academy of Sciences* 106.33 (2009): 13844-13849.

[26] Xu, L., Stige, L. C., Chan, K. S., Zhou, J., Yang, J., Sang, S., Lu, L. (2016). Climate variation drives dengue dynamics. *Proceedings of the National Academy of Sciences*, 201618558.