# Fuzzy Based Model For Predicting Malaria Outbreak In South-West, Nigeria

**WILLIAMS, Kehinde O.**
Department of Computer Science,
Federal Polytechnic, Ile-Iluji Ondo State, Nigeria.
kehindewilliams@gmail.com

**FALOHUN, A. S.**
Department of Computer Science and Engineering,Ladoke Akintola University of Technology (LAUTECH), Ogbomoso.
asfalohun@lautech.edu.ng

**ADEGOKE, Benjamin O.**
Department of Computer Engineering
Faculty of Engineering,
Federal Polytechnic, Ile-Oluji
Ondo State, Nigeria
adegokebo35@gmail.com , benadegoke@fedpolel.edu.ng

***Abstract*— Malaria is a major problem to Nigerians today and this threat is a serious pandemic in many developing nations at large especially those with insufficient health facilities and is unable to cater for their national health needs. This study provided a means for monitoring the risk of malaria among Nigerians and thus a means of mitigating the risk of the deadly scourge of malaria. Malaria risk factors were employed as inputs into the Fuzzy Inference System (FIS) employed in the system and it used one hundred and twenty-seven (127) fuzzy rules. The results from the expert system reflects a good ability of the developed system to predict outbreak of malaria.**

***Keywords—fuzzy logic, malaria, membership functions, risk factors***

## I. Introduction

Malaria is one of the most serious vector-borne diseases (VBDs) and remains a leading cause of morbidity and mortality and a major health burden in many developing countries [1, 2, 3, 4]. According to [4], malaria is a disease of the tropics and subtropics. It is a global problem which is very endemic in over 106 countries with an estimated 3.3 billion people at risk of malaria [5]. In 2010 there were 216 million cases of malaria worldwide, with an estimated 655,000 deaths [6]. In Indonesia around half the population lives in malaria-prone areas [7]. Control of malaria often focuses on control of mosquitoes either by excluding them from contact with people using insecticide treated nets (ITNs) or by indoor residual spraying (IRS) of insecticide [8].

Despite more than 100 years of control efforts [2], epidemics of malaria still occurred in 104 countries in 2012 [9]. In 2011, there was an estimation of about 216 million malaria cases and approximately 665,000 deaths. International funding for malaria control increased sharply from less than US$ 100 million in 2000 to US$ 1.84 billion in 2012 [9]. Hay estimated that although the number of regions where malaria transmission occurs has reduced from 1900 to 2010, with land area ranged from 77 to 39 million km2 and the population exposing at risk has grown from 890 million to 3,410 million worldwide [10]. Malaria is a heavy burden of disease and a threat to global prosperity, economic growth and development [2]. In some countries with heavy malaria burden, the disease account for as much as 40% of public health expenditure, 30 to 50% of inpatient admissions, and up to 50% of outpatient visits [11].

Modeling the likelihood of malaria risk can help physicians identify the variables that are important to its incidence and have potential application in the management of the disease. The importance of modeling in understanding malaria risk is evident in the range of approaches available in modeling literature. These include disease dynamics such as transmission, vector/disease relationships with environment (especially with climate) as well as more detailed clinical studies, vaccine development, entomologic modeling of vectors and genetic research. Mathematical disease models for malaria may consider the dynamics [12] or transmission [13] including entomological parameters [14] or transmission and resistance [15].

Physicians are prone to making decision errors, because of high complexity of medical problems and due to cognitive limitations. Decision making in medicine is complex because, a vast amount of knowledge is required even to solve seemingly simple problems [16]. A physician is required to remember and apply knowledge of a large array of entities like disease presentations, diagnostic parameters, drug combinations and guidelines. However, the physician's cognitive abilities are restricted due to factors like multi-tasking, limited reasoning and memory capacity [17].

Malaria risk factors among others include: pregnancy status in women, education level [18, 19], age/age group, location [20, 21], proximity of residence to potential mosquito breeding sites [22], artificial agricultural sites, socio-economic status [22, 23]. Researches on malaria infections have been conducted in Neast Tanzania [19], Ghana [21], south-west Ethiopea [19]. A number of research had been conducted on malaria outbreak but all were treating its outbreak but not many had tried to predict its outbreak so as to guide against its occurrence. There is a need for a model that can be used in the identification of the likelihood of the risk of malaria using identified non-clinical variables relevant to malaria risk, hence this study. This research employed fuzzy logic, one of the soft-computing tool to predict outbreak of Malaria.

## II. METHODOLOGY

The process of fuzzification required two main stages which include derivation of the membership functions for both the input and the output variables alongside the linguistic representation of their functions. Membership function chosen for this study was the triangular membership functions because of the variation in the data representing the input variables. Input variables used for development of the fuzzy logic model for malaria risk prediction used a number of triangular membership function that was proportional and equal to the number of values of each respective variable considered for risk of malaria. Fuzzification parameters for the triangular membership function were proposed as represented by equations (1 to 9).

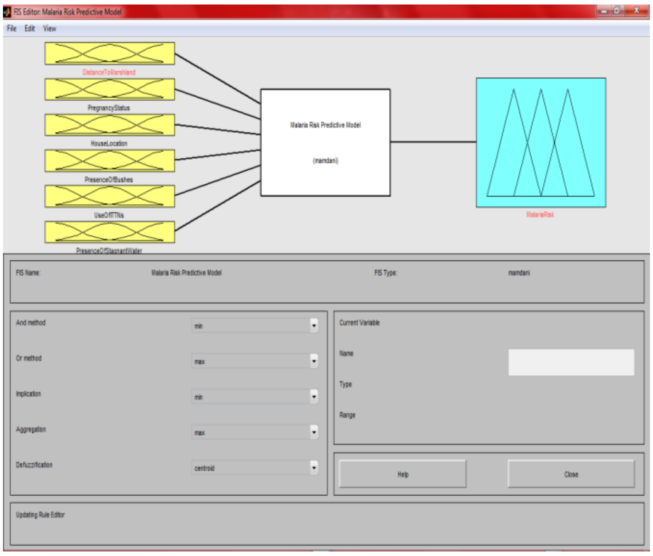

Figure 1: Fuzzy Inference System for Malaria Risk

Diagrammatic representation of the fuzzification are shown in Appendix A. It represents distance to marshlands, pregnancy status, presence of bushes, use of insecticides treated net (ITNs) and the presence of stagnant waters respectively.

Formulation of fuzzifiers for 2, 3, and 4 membership functions (mf) are shown in equations 1 and 2; 3, 4 and 5; and 6, 7, 8, and 9 respectively.

a. Risk factors with two variables, the triangular membership functions used were formulated using the two expressions as follows:

$$f(variable1; 0.00, 0.30, 0.60) = \begin{cases} 0, & x \le 0.00 \\ \frac{x}{0.30}, & 0.00 \le x \le 0.30 \\ \frac{0.60 - x}{0.30}, & 0.30 \le x \le 0.60 \\ 0, & 0.60 \le x \end{cases} \quad (1)$$

$$f(variable2; 0.50, 0.75, 1.00) = \begin{cases} 0, & x \le 0.50 \\ \frac{x - 0.50}{0.25}, & 0.50 \le x \le 0.75 \\ \frac{1.00 - x}{0.25}, & 0.75 \le x \le 1.00 \\ 0, & 1.00 \le x \end{cases} \quad (2)$$

b. Risk factors with three variables, the triangular membership functions used were formulated using the three expressions as follows:

$$f(Low; 0.00, 0.20, 0.40) = \begin{cases} 0, & x \le 0.00 \\ \frac{x}{0.20}, & 0.00 \le x \le 0.20 \\ \frac{0.40 - x}{0.20}, & 0.20 \le x \le 0.40 \\ 0, & 0.40 \le x \end{cases} \quad (3)$$

$$f(Moderate; 0.30, 0.50, 0.70) = \begin{cases} 0, & x \le 0.30 \\ \frac{x - 0.30}{0.20}, & 0.30 \le x \le 0.50 \\ \frac{0.70 - x}{0.20}, & 0.50 \le x \le 0.70 \\ 0, & 0.70 \le x \end{cases} \quad (4)$$

$$f(High; 0.60, 0.80, 1.00) = \begin{cases} 0, & x \le 0.60 \\ \frac{x - 0.60}{0.20}, & 0.60 \le x \le 0.80 \\ \frac{1.00 - x}{0.20}, & 0.80 \le x \le 1.00 \\ 0, & 1.00 \le x \end{cases} \quad (5)$$

$$f(Slum; 0.00, 0.15, 0.30) = \begin{cases} 0, & x \le 0.00 \\ \frac{x}{0.15}, & 0.00 \le x \le 0.15 \\ \frac{0.30 - x}{0.15}, & 0.15 \le x \le 0.30 \\ 0, & 0.30 \le x \end{cases} \quad (6)$$

$f(Low\ Cost; 0.25, 0.35, 0.50)$

$$= \begin{cases} 0, & x \le 0.25 \\ \dfrac{x - 0.30}{0.20}, & 0.25 \le x \le 0.35 \\ \dfrac{0.70 - x}{0.20}, & 0.35 \le x \le 0.50 \\ 0, & 0.50 \le x \end{cases} \quad (7)$$

$f(Medium\ Cost; 0.45, 0.55, 0.70)$

$$= \begin{cases} 0, & x \le 0.45 \\ \dfrac{x - 0.45}{0.10}, & 0.45 \le x \le 0.55 \\ \dfrac{0.70 - x}{0.25}, & 0.55 \le x \le 0.70 \\ 0, & 0.70 \le x \end{cases} \quad (8)$$

$f(High\ Cost; 0.65, 0.80, 1.00)$

$$= \begin{cases} 0, & x \le 0.65 \\ \dfrac{x - 0.65}{0.25}, & 0.65 \le x \le 0.80 \\ \dfrac{1.00 - x}{0.20}, & 0.80 \le x \le 1.00 \\ 0, & 1.00 \le x \end{cases} \quad (9)$$

Each of the variable were observed to be described using two (2) different values and as such were all formulated using the expressions in equations 1and 2 for each of their respective values

Figure 2 shows the schematic diagram of the defuzzification of the output variables – the risk of malaria, where each membership function represents the values of low, moderate and high by equations 3, 4, and 5 respectively. Figure 3 shows the schematic diagram of the fuzzification of the input variable – house location, where each membership function represents the values of slum, low cost, medium cost and high cost represented by equations 6, 7, 8 and 9 respectively.

***Rules formation for the inference engine***

The IF-THEN rules that were formulated were used to define the conditional statements that comprises of the fuzzy values generated from the crisp inputs. These domain specific conditional statements were generated from information provided by human experts. Variation of rules that guide the prediction of the risk of malaria among experts may be as a result of either or all of the following: number of years of experience of experts; number of cases that the expert has been exposed too; and/or instrument of diagnosis used by the experts.

For the purpose of this study, following the identification of the variables, predictive risk factors of malaria and the values of each respective risk factor; possible number of rules that were elucidated by expert were first identified and processed. A total of 128 rules were generated for the system. Rule editor of the fuzzy system is shown in Figure 4 while Appendix B gives a description of the rules that were elicited from the expert based on the combination of the values of the attributes of the risk of malaria. Figure 5 shows a description of the rule editor interface used in inserting all 128 rules elicited from the expert into the inference engine of the fuzzy model for the prediction of malaria risk. The values of each variable were selected from the bottom part of the interface where each attribute were defined and the rules were

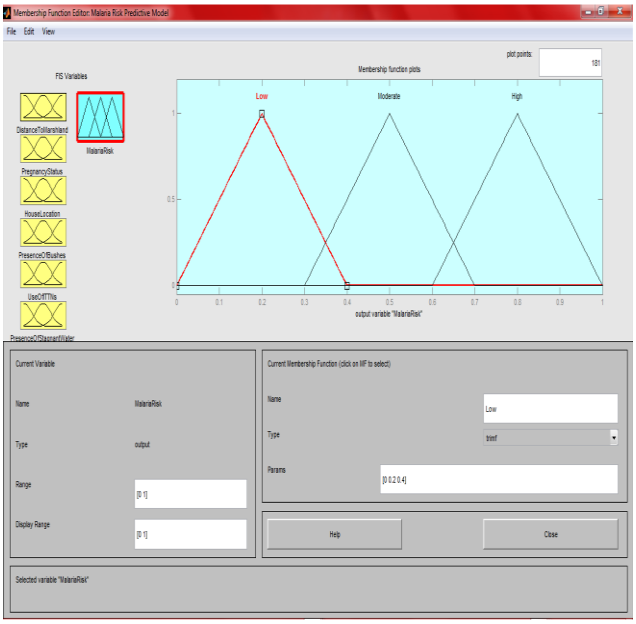

Figure 2: Defuzzification of Malaria Risk

added following the selection of the required values for the rules provided.

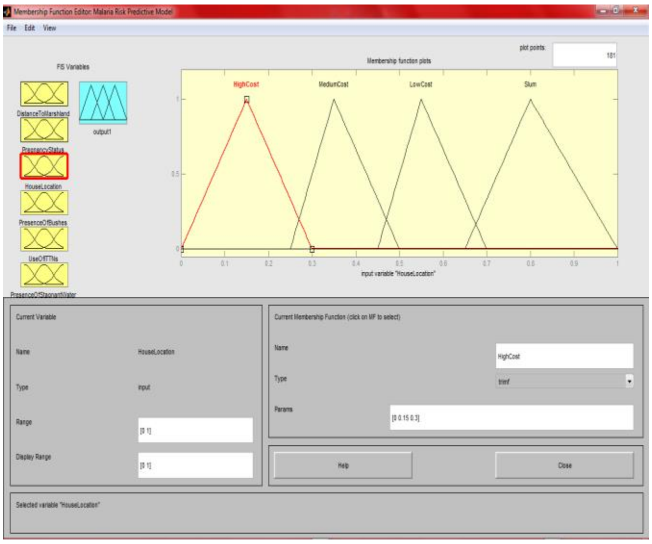

Figure 3: Fuzzification of House Location

***Aggregation of all the outputs***

The decision made by fuzzy logic inference systems were based on the testing of all the 128 rules in the Fuzzy Inference System (FIS), the rules were combined in such a way that decisions could be made. The process by which the fuzzy sets that represent the outputs of each rule were combined into a single fuzzy set is called aggregation. This process only occurs once of each output variable, just prior to the final step of defuzzification. Input of the aggregation process is the list of truncated output functions returned by the implication process for each rule. Hence, the output for the aggregation is a fuzzy set for each output variable. As long as the aggregation method is commutative (which it always should be), then the order in which the rules are executed is unimportant (Figure 4).

***Defuzzification of the output variable to its crisp value.***

The input into the defuzzification process is a fuzzy set (the aggregate output fuzzy set) and the output is a single number. As much as fuzziness helps the rule evaluation process during the intermediate steps, the final desired output for each variable is generally a single number. However, the aggregate of a fuzzy set encompasses a range of output values, and so must be defuzzified in order to resolve a single output value from the set. Centroid was used to return the center of area under the curve which gives the final output of the prediction system as shown in Figure 5.

III RESULTS AND DISCUSSION

Following the formulation and simulation of the fuzzy logic model for predicting the risk of malaria, the results were discussed using plots of the surface diagrams developed from MATLAB® Fuzzy Logic Tool-Box Surface Viewer interface. The surface viewer interface shows a 3-dimensional plot of the relationship between the values of two (2) variables (on the horizontal xy-plane) alongside their respective proportional value of the risk of

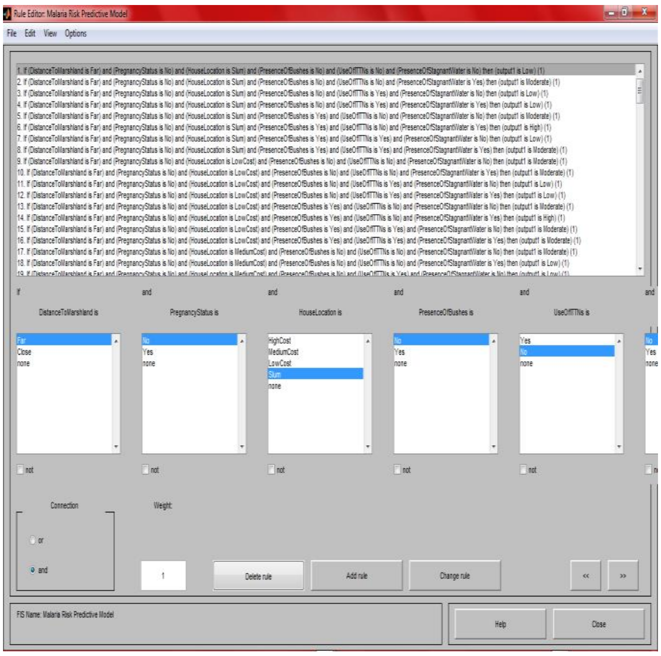

Figure 4: Rule Base consisting of the Inference engine

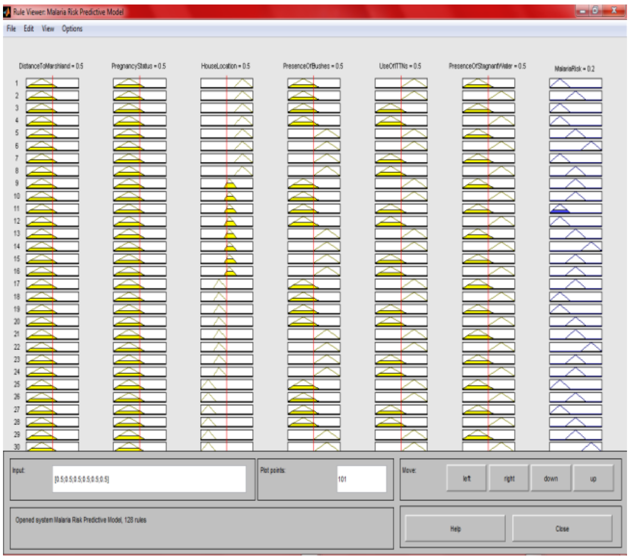

Figure 5: Aggregation and deffuzification of malaria risk

malaria on the z-axis which is the vertical axis perpendicular to the horizontal xy-plane. The surface plot is also affected by the rules presented by the expert which was used to develop the inference engine of the system.

Figure 6 shows the surface plot for the pregnancy status against the distance to marshlands with respect to their respective risk of malaria. The diagram shows that the risk of malaria is very low in most of the part of the diagram which is shown by the empty hollow portions residing within the shape plot of the distribution of the variables. Although, whenever the pregnancy status is Yes and whenever the presence to marshlands is near, the associated risk of malaria is seen to be more likely to rise. Figure 7 shows the surface plots of the presence of bushes against the distance to marshlands with respect to their risk of malaria. The diagram shows that irrespective of the distance to marshlands, whenever there is presence of bushes then the associated risk of malaria is always high – shown by the elevated flat region on the surface plot. Although, whenever the presence of bushes is No then the risk of malaria is always observed to be as low as possible – shown by the depressed region in the surface plots. Figure 8 shows the surface plots of the use of insecticides treated nets (ITNs) against the distance to marshlands with respected to their associated risk of malaria.

The diagram shows that whenever the use of insecticide treated net is No and the distance to the marshlands is Near then the risk of malaria is high – shown the elevated region of the surface plot. Although, whenever the use of ITNs is Yes and the distance to marshlands is Near then the associated risk of malaria is observed to be as low as possible – shown by the depressed region beside the elevated region. Also, it was discovered that whenever the distance to marshlands is Far then the risk of malaria is Low irrespective of whether ITNs were used or not. Figure 9 shows the surface plots of the presence of stagnant water against distance to marshlands with respect to their respective risk of malaria. The diagram shows that irrespective of the distance to marshlands the risk of malaria is lower whenever there is no presence of stagnant water. Although, whenever there is the presence of stagnant water then the risk of malaria is observed to be high – shown by the elevated region on the surface plot.

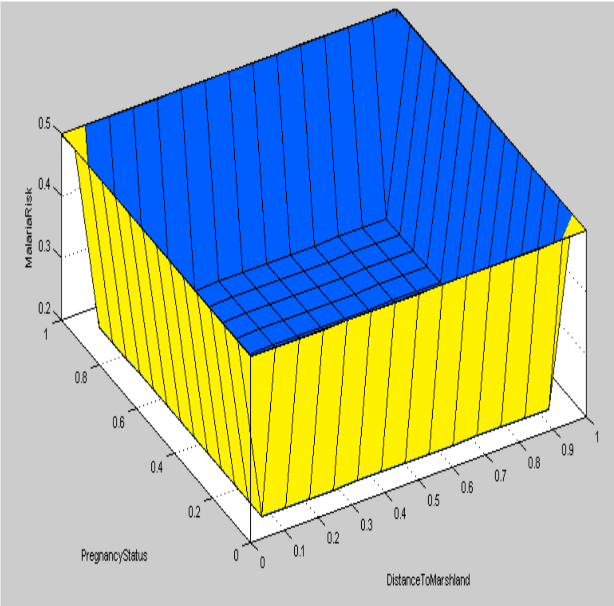


Figure 6: Surface plot of Pregnancy status against distance to marshlands

## IV CONCLUSION AND RECOMMENDATION

The fuzzy-based model for predicting malaria was simulated using the formulated description defined for the triangular membership function for all variables via the MATLAB Fuzzy Logic

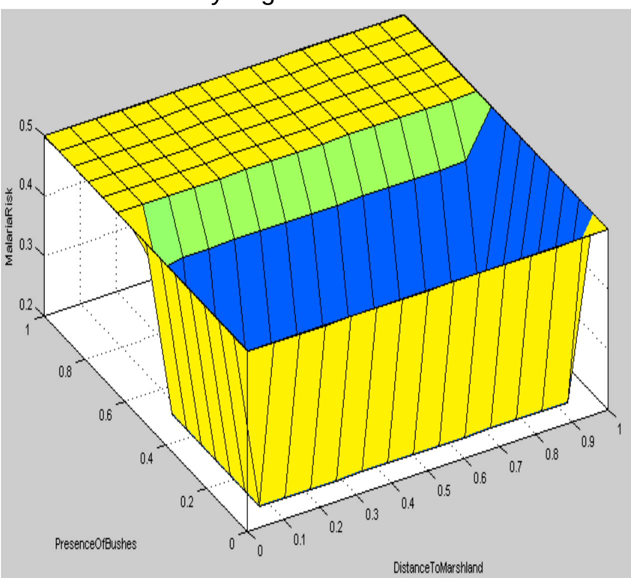


Figure 7: Surface plot of presence of bushes against distance to marshlands

Toolbox environment available in the MATLAB R2013b software used for the study. The results of the system by viewing the surface diagram showed that the system had the capacity to

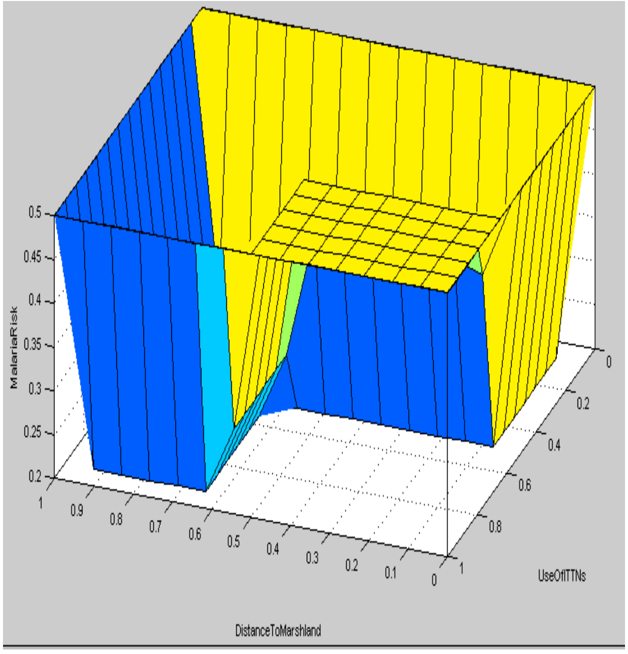


Figure 8: Surface plot of use of ITNs against distance to marshlands

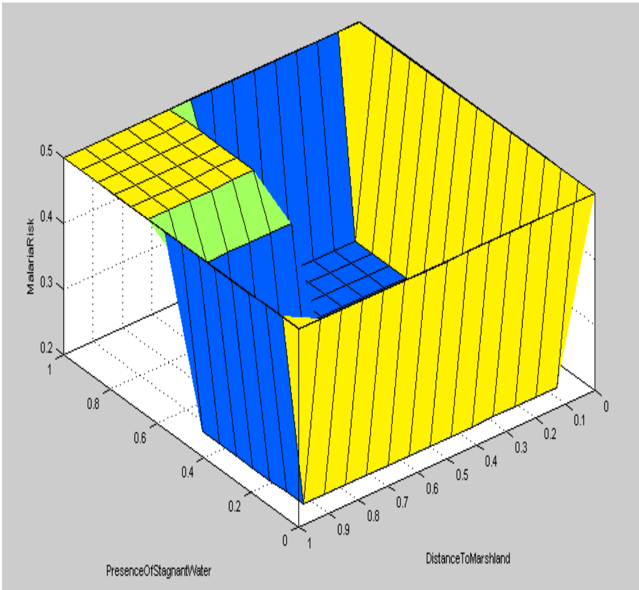


Figure 9: Surface plot of presence of stagnant water against distance to marshlands

properly determine the risk of malaria based on the respective values of the considered input variables. The system can be integrated as a software which can be used by individuals and public health officers alike to monitor the risk of malaria by identifying the values of the underlying variables. It is also believed that the model can help assist people in the identification of the risk of malaria so as to assist in the fight against the spread of malaria and also aid the quick support of malaria eradication in Nigeria. More input variables can be incorporated into the system.

REFERENCES

[1] Sachs, J., and Malaney, P. (2002).The economic and social burden of malaria. *Nature,* 415(6872): 680-685.

[2] Guinovart, C., Navia, M. and Tanner, M. (2006). Malaria: burden of disease. *Current molecular medicine,* 6(2): 137-143.

[3] Alonso, D., Bouma, M. J., and Pascual, M. (2011). Epidemic malaria and warmer temperatures in recent decades in an East African highland. *Proceedings of the Royal Society B: Biological Sciences,* 278(1712): 1661-1669.

[4] Ramathilaga, A., Murugesan, A. G. and Sathesh, P. C., (2012). Bio-larvicidal activity of Peanibacillusmacerans and Bacillus subtilisisolated from the dead larvae against Aedesaegypti- Vector for Chikungunya. *In Proceedings of the International Academy of Ecology and Environmental Sciences*, 2(2): 90-95.

[5] Visser, G., Dale, P., Dowe, D., Ndoen, E., Dale, M. and Sipe, N. (2012). A novel approach for modeling malaria incidence using complex categorical household data: The Minimum Message Length (MML) method applied to Indonesian data. *Computational Ecology and Software*, 2(3): 140 – 159.

[6] World Health Organization (2012). Global Malaria Programme. World malaria report. World Health Organization.

[7] World Health Organization, (2012). World Malaria Day. Accessed from http://www.who.int/mediacentre/news/releases/2012/malaria_20120424/en/on January 31, 2016.

[8] World Health Organization (WHO) (2009). Economic costs of malaria. Accessed from https://apps.who.int/inf–fs/en/InformationSheet10.pdf on February 21, 2016.

[9] World Health Organization (2012). Global Malaria Programme. World malaria report. World Health Organization.

[10] Hay, S., Cox, J. and Rogers, D. (2002).Climate change and the resurgence ofmalaria in the East African highlands. Nature 415(6874): 905-909.

[11] World Health Organization (WHO) (2002). Economic costs of malaria. Accessed from https://apps.who.int/inf–fs/en/InformationSheet10.pdf on February 21, 2016.

[12] Tumwiine, J., Mugisha, J.Y.T. and Luboobi, L.S. (2007).A mathematical model for the dynamics of malaria in a human host and mosquito vector with temporary immunity. Applied Mathematics and Computation, 189(2): 1953- 1965.

[13] Ishikawa, H., Ishii, A. and Nagai, N. (2003).A mathematical model for the transmission of Plasmodium vivax malaria. Parasitology International, 52(1): 81-93.

[15] Smith, M.W., McKlin, M.G. and Thomas, C.J. (2013).Hydrological and geomorphological controls of malaria transmission. Earth-Science Review 116(0): 109 – 127.

[16] Bell D., Alessandra E. F., Hehh, M. C., Boomgard, J. D. and McConnico, C. C., (2016). Development of new malaria diagnostics: matching performance and need. *Malaria Journal,* 15,

[17] Badu K, Brenya R. C., Timmann, C., Garm R. Kruppa, T. F, (2013).Malaria transmission intensity and dynamics of clinical malaria incidence in a mountainous forest region of Ghana, *Malaria World Journal,* 4(14): 1-8.

[18] Adam I., Khamis A. H. and Elbashir M. I., (2005). Prevalence and risk factors for Plasmodium falciparum malaria in pregnant women of eastern Sudan. Malaria Jounal, 4(18): 1-4.

[19] Hassan, A.N., Nogoumy, N.E. and Kassem, H.A. (2013). Characterization of landscape features associated with mosquito breeding in urban Cairo using remote sensing. *The Egyptian Journal of Remote Sensing and Space Science,* 2(1): 1 – 6.

[20] Winskill P., Rowland M, Mtove G., Malima, R. C., Kirby, M. J., (2011). Malaria risk factors in north-east Tanzania. Malaria Journal, 10 (98): 1-7.

[21] Doko-Gyeke D. and Kofie H. M., (2015). Factors Influencing Prevention and Control of Malaria among pregnant Women Resident in Urban . African Journal of Reproductive Health, 19(1): 44-53.

[22] De-Silva P. and Marshall J. M., (2012). Factors Contributing to Urban Malaria Transmission Sub-Sahara Africa: A Systematic Review. Journal of Tropical Medicine, 1-11.

[23] Vargas-De-Leon, C. (2012). Global analysis of a delayed vector-bias model for malaria transmission with incubation period in mosquitoes. *Mathematical Biosciences and Engineering*, 9(1): 165-174.

APPENDIX A

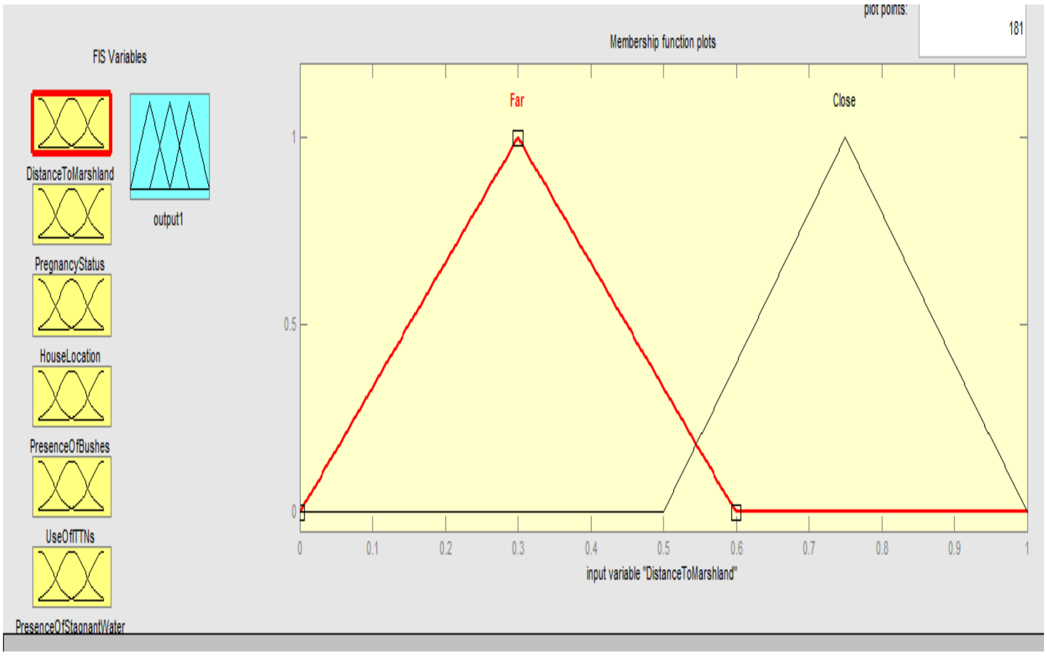


Figure A.1: Fuzzification of Distance to Marshlands

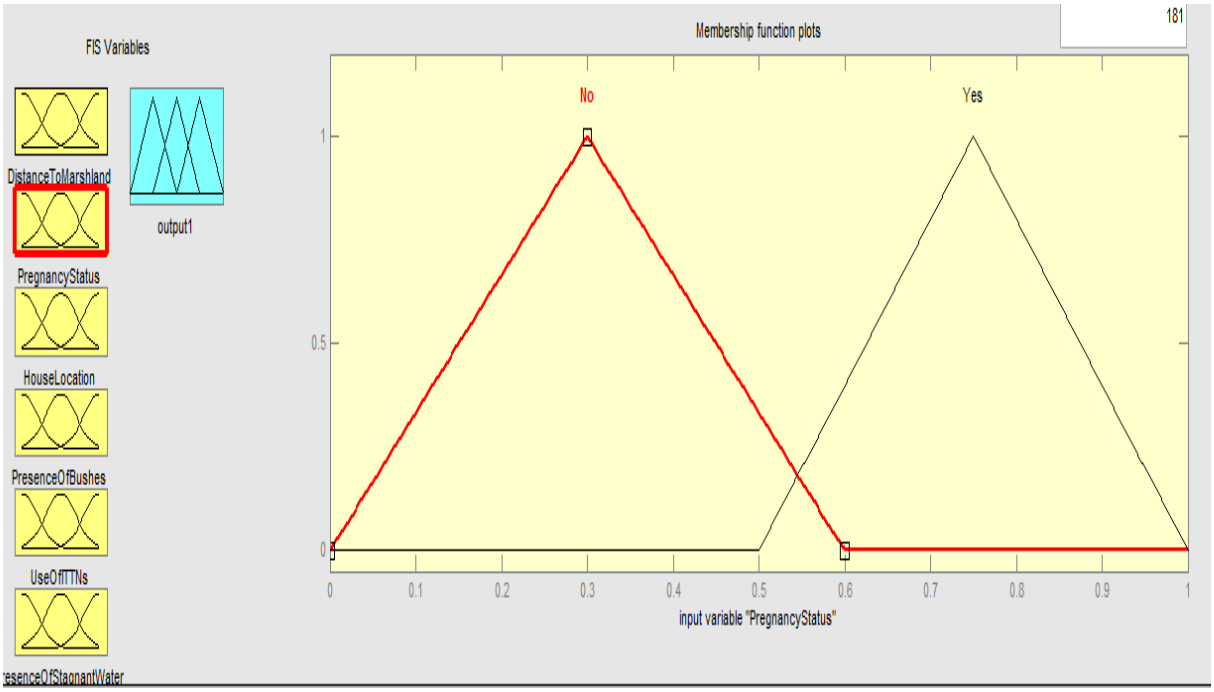


Figure A.2: Fuzzification of Pregnancy Status

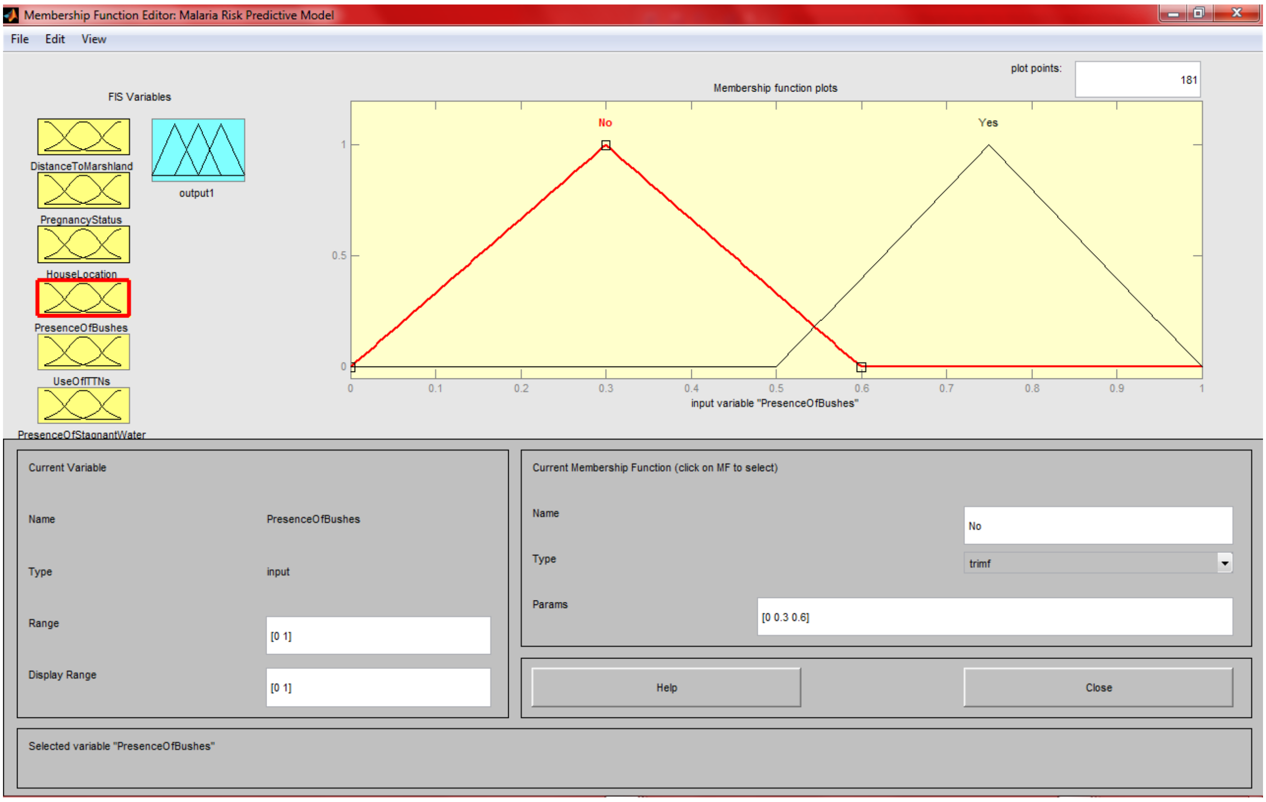


Figure A.3: Fuzzification of Presence of bushes

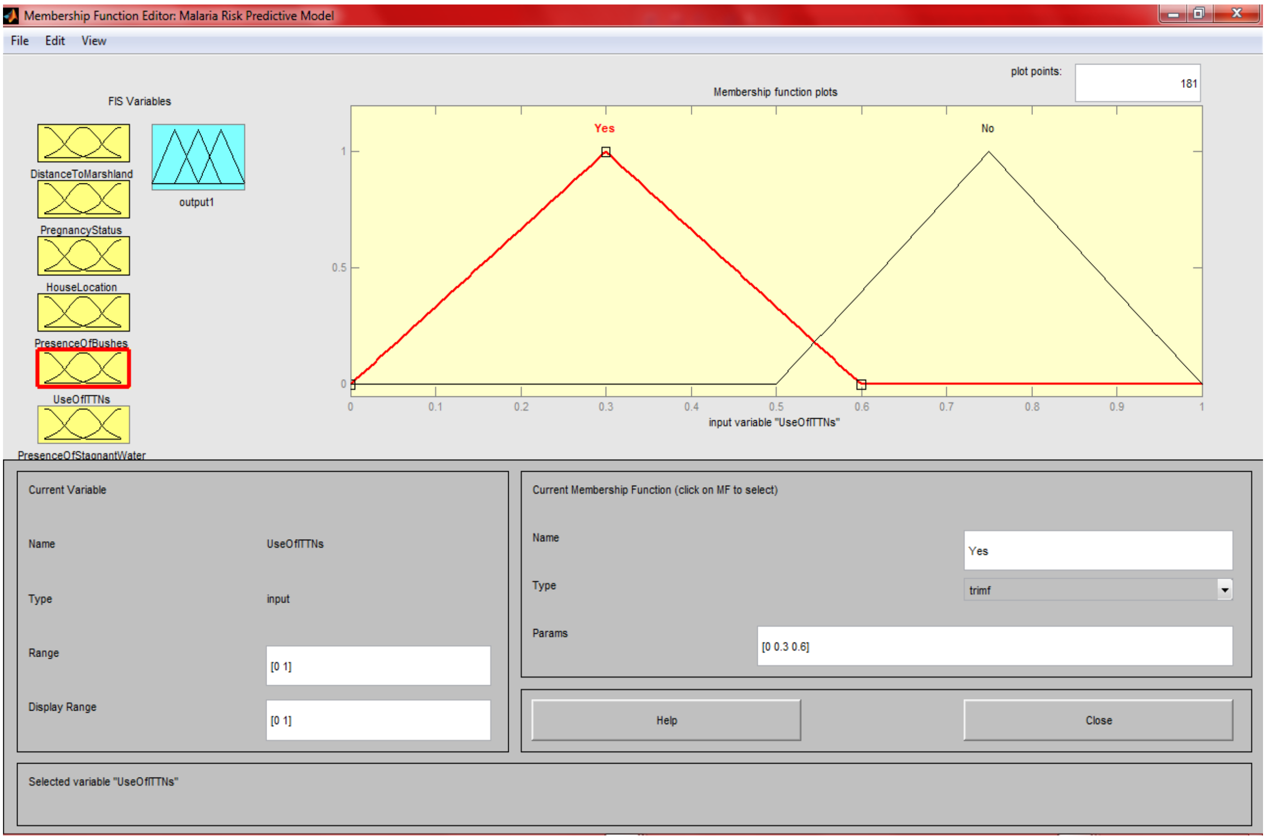


Figure A.4: Fuzzification of Use of Insecticide Treated Nets (ITNs)

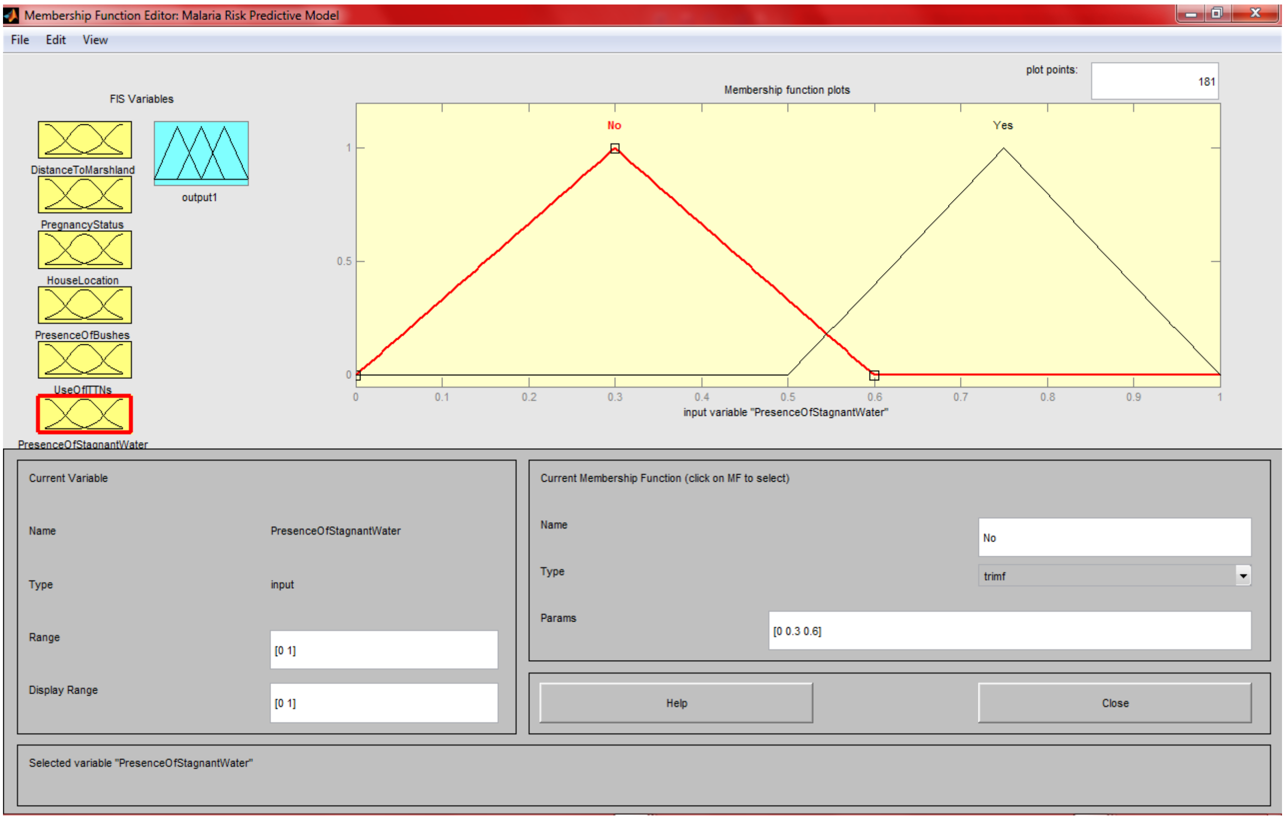


Figure A.5: Fuzzification of Presence of Stagnant Water

## APPENDIX B

Malaria Risk Rules Elicited from Expert Physician

| Rules | Distance to Marshland | Pregnancy Status | House Location | Presence of Bushes | Use of ITNs | Presence of stagnant water | Risk of Malaria |
|---|---|---|---|---|---|---|---|
| **1** | Far | No | Slum | No | No | No | Low |
| **2** | Far | No | Slum | No | No | Yes | Moderate |
| **3** | Far | No | Slum | No | Yes | No | Low |
| **4** | Far | No | Slum | No | Yes | Yes | Low |
| **5** | Far | No | Slum | Yes | No | No | Moderate |
| **6** | Far | No | Slum | Yes | No | Yes | High |
| **7** | Far | No | Slum | Yes | Yes | No | Low |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| **127** | Close | Yes | High cost | Yes | Yes | No | Moderate |
| **128** | Close | Yes | High cost | Yes | Yes | Yes | Moderate |