

**MODELLING OF VECTOR PREVALENCE VIS-A-VIS ITS
VULNERABILITY TO CLIMATE CHANGE**

SUDESH KUMAR SOHANI



**DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
APRIL 2015**

© Indian Institute of Technology Delhi (IITD), New Delhi, 2015

**MODELLING OF VECTOR PREVALENCE VIS-A-VIS ITS
VULNERABILITY TO CLIMATE CHANGE**

by

SUDESH KUMAR SOHANI

Department of Civil Engineering

Submitted

In fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

NEW DELHI, INDIA

APRIL 2015

CERTIFICATE

This is to certify that the thesis entitled *Modelling of Vector Prevalence vis-a-vis Its Vulnerability to Climate Change* submitted by **Mr. Sudesh Kumar Sohani** to the **Indian Institute of Technology Delhi**, for the award of the degree of **Doctor of Philosophy**, is a record of the original bonafide research work carried out by him under our guidance and supervision. The thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted in part or full to any other University or Institute for the award of any degree or diploma.

Dr. A. K. Nema

Professor

Department of Civil Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016

Dr. B. R. Chahar

Professor

Department of Civil Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016

ACKNOWLEDGEMENTS

I want to take this great opportunity to express my deep gratitude to all those great people who directly or indirectly supported me and contributed during the course of my PhD thesis work.

With high regards, I express my sincere gratitude and appreciation to my supervisors, Prof. B. R. Chahar, Department of Civil Engineering and Prof. A. K. Nema, Department of Civil engineering, Indian Institute of Technology Delhi for their expert guidance, constant encouragement, unconditional trust, support and constructive criticism during the course of my research work. I am deeply indebted to them for their suggestions during the preparation of the manuscript of this thesis and help to achieve my goal.

I express my special regards and thanks to my Student Research Committee members Prof. Mukesh Khare, Prof. S. V. Veeravalli and Dr. D. R. Kaushal for their critical review, valuable suggestions and support. Their suggestions helped me a lot to come through this research work.

I express my sincere thanks to the Authorities of IIT Delhi in general, and the Department of Civil Engineering in particular for providing facilities to carry out the research work.

I gratefully acknowledge the help of the faculty members Prof. A. K. Gosain, Prof. Shashi Mathur, Prof. A. K. Jain, Dr. Anushree Malik, Dr. Dhanya C. T. for their support, teaching and encouragement. I am especially grateful to my fellow research scholar and presently a Project Scientist at IIT Delhi Dr. Sudheer Ch. for his

unconditional cooperation and diligent guidance throughout this journey of my research work.

I am thankful to the Authorities of National Institute of Malaria Research (NIMR, Indian Council of Medical Research), New Delhi for providing the study data. I express my gratitude to Dr. R.C. Dhiman for sharing his expertise and providing specific data for the study. His constant encouragement and guidance helped me to decide appropriate line of action for my study. I wish to thank Dr. V.P. Sharma for his kind initial enlightening discussions and for interacting with me time to time, in turn, helping me to understand the concept of the vector borne diseases. I am thankful to the Authorities of CRDT, IIT Delhi for providing the required facilities for the study, I specially wish to thank Dr. Anushree Malik for the encouragement and for extending all the help.

My special thanks are due to my well-wishers, mentor and sponsorer Er. Ashok Sojatia, Chairman (BOG) and Er. Ashish Sojatia, Group Chairman, Acropolis Group of Institutions for encouraging me to pursue higher studies and providing me necessary support.

This acknowledgement will remain incomplete unless I thank my fellow research scholars: Raktim Haldar, R Maheswaran, Deepak Kumar, Jayshri Patel, Boini Narsimlu, Pratiksha Pandey, Shailendra Jain, Aniket Chakravorty, Jatin Anand, Sridebi Basu, Pawan Kumar Rai, Jyoti, Pankaj Prajapati, Ashwani Soni and the staff members: Mr. Rajbir Agrawal, Mr. N. R. Gehlot, Mr. Vikram Chand, Mr. Rajiv Sharma, Mr. Amit Bundela, Mr. Chakresh, for their utmost cooperation at every stage, and for good ideas and humour to keep the study period pleasant.

I am indeed lucky to live in a joint family of we three brothers, my elder Dr. Susheel Sohani, younger Mr. Sanjeev Sohani, and parents. I do not find words to express my gratitude to every single member of my big family, my mother, my wife Alpana and son Shivam for their all-out support who also accepted my challenging task at this stage. My late Father and the almighty GOD shall remain the ideal to take all the credit.

New Delhi, April 2015

Sudesh Kumar Sohani

ABSTRACT

Malaria is one of the most dangerous vector borne diseases, causing more than one million deaths every year. Many researchers envisioned the control and elimination of malaria by inventing new drugs, developing newer strategies and by developing various models. In the process, many new concepts and complexities have been introduced. Still these inventions and new models are like tip of the iceberg compared to the existing complexities involved in malaria.

In India, malaria is a major public health problem and it contributes significantly to the overall malaria burden in Southeast Asia. To find a solution to the problem, mathematical models do help explain the dynamics of such infectious disease and can facilitate comparisons among competing control strategies that can inform policy decisions. Further to strengthen the prevention and control measures, this study has been carried out in order to develop appropriate modelling approaches of forecasting malaria transmission.

The study specifically has been aimed to develop three possible approaches to model and forecast the propagation of malaria with respect to the climate change variables. The first approach is to predict the variation in vector density with respect to climate variables in the subtropical climatic zones of lower Himalayan regions having increasing altitudes using Genetic Programming (GP). The study is done at three locations namely Raikhalkhatta, Bhorsa and Jyadi. Models have been developed based on the vector density as indicator. Calibration and validation of the simulated vector density indicated very good significance levels. To assess the impact of temperature and rainfall on the

vector density sensitivity analyses also have been done for all the locations. The analyses showed the variation in vector density on increasing and decreasing the temperature.

The second forecasting model has been developed through another novel approach by coupling the Fire Fly Algorithm (FFA) and Support Vector Machines (SVM) to determine the malaria transmission at two locations in the plains of Rajasthan under arid and semi-arid climatic zones. The performance of SVM models depend upon the appropriate choice of SVM parameters. In this study FFA has been employed for determining the parameters of SVM. In addition, three more forecasting approaches are developed by employing Support Vector Machines (SVM), Artificial Neural Network (ANN) and Auto Regressive Moving Average (ARMA). At both stations malarial incidences have been computed using SVM-FFA, SVM, ANN and ARMA models. Simulated and actual malaria incidences values are compared. SVM-FFA model performed best out of four models, and showed very good significance level for the simulated malaria incidences on comparing it with the actual incidences. Box and Whiskers Plot has been used to find out the normalized residual errors to compare all the above algorithms. Sensitivity analyses also have been done which showed that there is variation in incidences on increasing and decreasing the temperature.

The Anopheles mosquito is one of the major vectors of malaria. Anopheles culicifacies species play a major role in spreading malaria in Asian countries, especially in Indian subcontinent. Hence, thirdly a mathematical model has been worked out to understand interactions of entomological and climatic parameters of malaria transmission. MATLAB has been used to develop this deterministic model, in the process three major steps are followed. First of all the endogenous variables, which have significant influence on the

parasite transmission cycle, are identified and their relationship as a function of temperature is determined. Secondly, vector ecology, and entomological parameters are identified in order to determine exogenous parameters, which have direct influence on the malaria transmission. Finally a set of Coupled Ordinary Differential Equations is solved in order to determine the malarial incidences using Fourth Order Runge-Kutta Methods. The developed model is validated using already published data. Results match quite well with the data.

- - -

TABLE OF CONTENTS

CERTIFICATE	i
ACKNOWLEDGEMENTS	iii
ABSTRACT.....	vii
TABLE OF CONTENTS	xi
LIST OF FIGURES	xvii
LIST OF TABLES	xxi
LIST OF SYMBOLS AND NOTATIONS.....	xxiii
CHAPTER 1	1
INTRODUCTION.....	1
1.1 Background	1
1.2 Research Motivation	4
1.3 Objectives and Scope of the work.....	5
1.3.1 Objectives	5
1.3.2 Scope of the work	5

1.4	Organization of the Research Work in the Thesis.....	6
CHAPTER 2		9
LITERATURE REVIEW		9
2.1	Mechanistic Modelling.....	11
2.1.1	Ronald Ross Model.....	12
2.1.2	Mac Donald Model	14
2.1.3	Anderson and May Model	15
2.1.4	Reproductive Number	15
2.2	Statistical Modelling	16
2.3	Effect of climate parameters on malaria transmission	17
2.3.1	Rainfall.....	17
2.3.2	Temperature	18
2.4	Miscellaneous.....	19
CHAPTER 3.....		25
DEVELOPMENT OF A GENETIC PROGRAMMING MODEL FOR		
ESTIMATING THE <i>AN. CULICIFACIES</i> VECTOR DENSITY		25
3.1	Introduction	25
3.2	Data Collected	27
3.3	Genetic Programming	28
3.4	Data Analysis	32
3.5	Determination of Input Variables.....	34
3.6	Model Development, Calibration and Validation	38

3.7	Sensitivity analysis for GP model	44
3.8	Projected Vector Densities	49
CHAPTER 4.....		55
A SUPPORT VECTOR MACHINE-FIRE FLY ALGORITHM BASED		
FORECASTING MODEL TO DETERMINE MALARIA TRANSMISSION		55
4.1	Introduction	55
4.2	Support Vector Machines	57
4.3	Fire Fly Algorithm in selecting the parameters of SVM.....	59
4.4	Results and Discussions	62
4.4.1	Training and testing of SVM models.....	64
4.4.2	Estimation of SVM parameters using FFA.....	65
4.4.3	Model Development.....	68
4.4.3.1	Auto Regressive Moving Average Method (ARMA)	68
4.4.3.2	Artificial Neural Network Model (ANN)	69
4.4.3.3	Support Vector Machine (SVM)	70
4.4.4	Comparison of the proposed/ recommended SVM-FFA model with the traditional models.....	71
4.4.5	Sensitivity analysis for SVM-FFA Model	80
CHAPTER 5.....		83
MODELLING MALARIAL INCIDENCES USING ENTOMOLOGICAL AND		
CLIMATIC PARAMETERS.....		83
5.1	Introduction	83
5.2	Methodology	84
5.2.1	Description of major parameters.....	86
5.2.2	Effect of temperature on entomological parameters.....	89

5.2.2.1	Sporogonic Cycle	90
5.2.2.2	Feeding Rate	91
5.2.2.3	Daily Survival Probability rate	91
5.2.2.4	Average Infectious Period (P_F , years).....	92
5.2.2.5	Natural Mortality Rate of Mosquitoes (μ_M)	93
5.2.2.6	Vectorial Capacity	94
5.2.2.7	Average Immune Period.....	95
5.2.3	Governing Equations	96
5.3	Study area	100
5.4	Results and discussions	101
CHAPTER 6		105
CONCLUSIONS AND RECOMMENDATIONS.....		105
6.1	Conclusions	105
6.2	Justification and comparison of all the three approaches of modelling	110
6.3	Future Scope.....	120
REFERENCES.....		121
APPENDIX 1		141
SUPPORT VECTOR MACHINE		141
APPENDIX 2.....		149
SAMPLE OUTPUT OF SVM-FFA.....		149
APPENDIX 3.....		151

MATLAB Code	151
A3.1 SVM-FFA	151
A3.2 Programme for SVM Jodhpur/Bikaner	164
A3.3 Programme for ANN Bikaner/Jodhpur	201
A3.4 Programme for ARMA Jodhpur/Bikaner	203
APPENDIX 4.....	205
MALARIA INCIDENCE DATA.....	205
A4.1 Monthly malaria incidence data in Bikaner region	205
A 4.2 Monthly malaria incidence data in Jodhpur region.....	209
LIST OF PUBLICATIONS IN INTERNATIONAL JOURNALS	213
BRIEF BIODATA OF THE AUTHOR	215

LIST OF FIGURES

Figure 3-1 (a) Schematic map depicting the study areas. (b) Elevation profile of the study areas	28
Figure 3-2 (a) Examples of Parse trees. (b) An illustration depicting Mutation process. (c) An illustration depicting the process of Cross over in Genetic programming model.....	30
Figure 3-3 (a) Box plot representations of relative vector density. (b) Box plot representations of temperature. (c) Box plot representations of rainfall.	33
Figure 3-4 (a) Cross correlation plots between mean precipitation and <i>An. Culicifacies</i> Density in Bhorsa region. (b) Cross correlation plots between temperature and <i>An. Culicifacies</i> Density in Bhorsa region	35
Figure 3-5 (a) Cross correlation plots between mean precipitation and <i>An. Culicifacies</i> Density in Raikhalkhatta region. (b) Cross correlation plots between temperature and <i>An. Culicifacies</i> Density in Raikhalkhatta region	37
Figure 3-6 (a) Cross correlation plots between mean precipitation and <i>An. Culicifacies</i> Density in Jyadi region. (b) Cross correlation plots between temperature and <i>An. Culicifacies</i> Density in Jyadi region	37
Figure 3-7(a) Calibration and validation of <i>An. Culicifacies</i> predicted density based on GP model for Bhorsa (b) Correlation Coefficient between observed and simulated density of calibrated data for <i>An. Culicifacies</i> for Bhorsa (c) Correlation Coefficient between observed and simulated density of validated data for <i>An. Culicifacies</i> for Bhorsa	40
Figure 3-8 (a) Calibration and validation of <i>An. Culicifacies</i> predicted density based on GP model for Raikhalkhatta (b) Correlation Coefficient between observed and simulated density of calibrated data for <i>An. Culicifacies</i> in Raikhalkhatta (c) Correlation Coefficient	

between observed and simulated density of validated data for <i>An. Culicifacies</i> in Raikhalkhatta	41
Figure 3-9 (a) Calibration and validation of <i>An. Culicifacies</i> predicted density based on GP model for Jyadi (b) Correlation Coefficient between observed and simulated density of calibrated data for <i>An. Culicifacies</i> in Jyadi (c) Correlation Coefficient between observed and simulated density of validated data for <i>An. Culicifacies</i> in Jyadi.....	42
Figure 3-10 Changes in Model Output with reference to changes in Temperature	46
Figure 3-11 Changes in Model Output with reference to changes in Precipitation.....	48
Figure 3-12 Projected vector densities for Raikhalkhatta.....	50
Figure 3-13 Projected vector densities for Jyadi.....	51
Figure 4-1 India map depicting the study areas	62
Figure 4-2 Mean monthly variations of Temperature, Relative Humidity and Rainfall in Bikaner.....	63
Figure 4-3 Mean monthly variations of Temperature, Relative Humidity and Rainfall in Jodhpur.....	64
Figure 4-4 Schematic picture depicting the SVM-FFA methodology.....	66
Figure 4-5 Comparison between simulated and actual malaria incidences values of Bikaner using different model	74
Figure 4-6 Comparison between simulated and actual malaria incidences values of Jodhpur using different models	76
Figure 4-7 (a) Box plot of Normalized residual error of Bikaner (b) Box plot of Normalized residual error of Jodhpur	77

Figure 4-8 Scatter plots between simulated and actual malaria incidences values of Bikaner	78
Figure 4-9 Scatter plots between simulated and actual malaria incidences values of Jodhpur.....	79
Figure 5-1 Flowchart to develop vector transmission models	85
Figure 5-2 Transformation of Sporozoite into Merozite.....	86
Figure 5-3 Schematic diagram showing the entire process of Erythrocytic cycle and Exo-Erythrocytic cycle (Source: http://dpd.cdc.gov/dpdx)	87
Figure 5-4 Schematic sketch depicting the sporogonic cycle (Source: http://dpd.cdc.gov/dpdx)	89
Figure 5-5 Effect of temperature on Sporogonic cycle.....	90
Figure 5-6 Effect of Temperature on Feeding Interval	91
Figure 5-7 Effect of Temperature on Daily Survival Probability rate	92
Figure 5-8 Temperature (Temp, °C) V/s Average Infectious Period (P_F, years)	93
Figure 5-9 Temperature (Temp, °C) V/s Natural Mortality Rate of Mosquitoes (Mu_M).	94
Figure 5-10 Temperature (Temp, °C) V/s Vectorial Capacity (VC, Infected bites per human per night).	95
Figure 5-11 Temperature (Temp, °C) V/s Average Immune Period (P_M, years).....	96
Figure 5-12 Study area - Nuqui region	100
Figure 5-13 The time series plot depicting temperatures and precipitations of Nuqui..	101
Figure 5-14 Comparison between simulated and actual malaria incidences values of Nuqui region.	103

Figure 6-1 Flow chart shows use of Genetic Programming to obtain initial condition for vector population (susceptible) to model vector and host interaction. 108

Figure 6-2 Schematic sketch depicting the procedure to prioritize the selection of appropriate model based on the available data. 115

LIST OF TABLES

Table 3.1 Variations in R^2 values with changes in temperature.....	45
Table 3.2 Variations in R^2 values with changes in Rainfall.....	47
Table 4-1 SVM parameters determined by FFA for Jodhpur and Bikaner station	67
Table 4-2 The NMSE and R^2 statistics of SVM-FFA for Bikaner and Jodhpur for testing and training data.....	68
Table 4-3 Parameters of ANN models.....	69
Table 4-4 Number of training and validation data sets used in training the models	71
Table 4-5 Comparison of R^2 and NMSE values for Bikaner and Jodhpur Stations	72
Table 4-6 Variations in R^2 values with changes in Temperature and Rainfall for Bikaner and Jodhpur Stations	80
Table 5-1 Population Details in Nuqui Region.....	102
Table 6-1 Exogenous variables required for mosquito population.....	116
Table 6-2 Exogenous variables considered for the human population	117
Table 6-3 Exogenous variables considered in the infectious disease models.....	117
Table 6-4 Time series of populations.....	118
Table 6-5 Metrological Data.....	118
Table 6-6 Additional* Data required	119
Table 6-7 Data of the endemic area to model malarial incidences	119
Table 6-8 Data of the endemic area to model vector density	119

LIST OF SYMBOLS AND NOTATIONS

The following listed symbols and notations are used in the Thesis

Symbol	Definition
ANN	Artificial Neural Network
$ARMA$	Auto Regressive Moving Average method
EIR	Entomological Inoculation Rate
FFA	Fire Fly Algorithm
GP	Genetic Programming
HBI	Human Blood Index
HUF	Infectious Human
HUI	Infected Human
HUM	Recovered Human
HUS	Susceptible Human
Ma	Human Biting rate
N	Sporogonic Cycle
P_{hu}	Total Human Population
PSO	Particle Swarm Optimization
R	Rainfall /precipitation
R^2	Coefficient of determination
RH	Relative humidity
R_{mo}	Mortality rate
$RMSE$	Root Mean Square Error

R_{na}	Natality rate
SVM	Support Vector Machine
$Temp$	Temperature
VC	Vectorial Capacity
VF	Infectious Vector
VI	Infected Vector
VS	Susceptible Vector
μ_m	Natural Mortality rate
α_m	Induced Mortality Rate