GIS-ENHANCED BAYESIAN REINFORCEMENT LEARNING FOR PREDICTING SPREAD OF VECTOR-BORNE INFECTIOUS DISEASES: A CASE STUDY OF MALARIA IN KENYA

KIPNGETICH GIDEON

A THESIS SUBMITTED IN PARTIAL FULFILLMENT FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN STATISTICS IN THE SCHOOL OF PURE AND APPLIED SCIENCES, KIRINYAGA UNIVERSITY

AUGUST, 2026

DECLARATION

This thesis is my original work and has not been presented for a degree or any other award in any other Institution.

Signed..... Date.....

Kipngetich Gideon

PA301/S/17846/2022

This thesis has been submitted with our approval as University Supervisors.

Signed..... Date.....

Dr. Victor Musau

Department of Pure and Applied Sciences,

Kirinyaga University, Kenya

Signed..... Date.....

Dr. Margaret Kinyua

Department of Mathematics, Statistics and Actuarial Sciences,

Karatina University, Kenya

COPYRIGHT

© 2026 Kipngetich Gideon

All rights reserved

No part of this thesis may be reproduced, stored in any retrieval system or transmitted in any form or by any means; electronic, mechanical, photocopying, recording or otherwise without prior written permission of the author or Kirinyaga University, except for citations and quotations used for academic purposes with proper acknowledgement

DEDICATION

To my lovely mum Florence Kirui and dad Joseph Kirui.

ACKNOWLEDGMENT

I thank God for his care and protection during my study period and for placing people in my life who have benefited me. I want to express my boundless gratitude to my supervisors Dr. Victor Musau and Dr. Margaret Kinyua for their support, guidance, and encouragement throughout this research. They directed me and provided valuable information and suggestions on the best ways to improve this research. I also thank Kirinyaga University for giving me an opprtunity to pursue my degree of doctor of philosophy in statistics.

ABSTRACT

Vector-borne infectious diseases have continued to pose a significant challenge to global public health, and accurate and timely spread prediction is crucial for effective intervention and control. The development of accurate and efficient predictive models for vector-borne infectious diseases has been a current trend in disease modeling. Modeling the spread, however, has assumed a discrete transmission space, which is not always ideal in the real world, and has paid little attention to overestimation biases in predictions. The effectiveness of a predictive model is also evaluated by its ability to capture a significant number of data characteristics, thereby enabling robust and accurate case prediction. The study developed a GIS-enhanced Bayesian Reinforcement learning model for predicting the spread of vector-borne infectious diseases, and assessed its performance. Bayesian quantifies the uncertainty in model parameter estimation, and min-max ensemble Q-value estimation minimizes overestimation bias in the model. A simulation study was employed to evaluate model performance, success rate, and infection rate, case prediction error, and cross-location adaptability, and the Kenya Malaria Indicator Survey 2015 and 2020 from the endemic western counties of Kenya were used to validate the model performance. Results showed that humans have a success rate of up to 97.4% and an infection rate of 2.6% when location-varying intervention rates of spray repellent and treated mosquito nets are adopted. Temperature, rainfall, age, wealth index (middle), wealth index (richer), wealth index (richest), and education significantly influence the risk of contracting malaria in the region. The model demonstrates better performance in describing the spread of malaria than the GIS-MAB model, with an average case prediction error of 0.0211 and cross-location adaptability greater than 1 across all human interventions. The model can be used to determine appropriate intervention rates to reduce human-vector interactions.

ABBREVIATIONS AND ACRONYMS

ACT	Artemisinin-Based Combination Therapy
ARI	Acute Respiratory Infection
AIDS	Acquired Immunodeficiency Syndrome
Ada-Boosting	Adaptive Boosting
AGLSTM	Adaptive Gradient Long-Short Term Memory
ARIMA	Auto-regressive Integrated Moving Average
DL	Deep Learning
DNN	Deep Neural Network
DQN	Deep Q-Network
FACMAC	Factored Multi-Agent Centralized Strategy Gradient
GLEaM	Global Epidemic and Mobility
GIS	Geographic Information System
HIV	Human Immunodeficiency Virus
KNN	K-Nearest Neighbors
LASSO	Least Absolute Shrinkage and Selection Operator
LGBM	Light Gradient Boosting Machine
LSTM	Long-Short Term Memory
LSTMSeq2Seq	Long-Short Term Memory Sequence-to-Sequence
MIC	Malaria Incidence Classification
MAB	Multi-Armed Bandit

MADDPG	Multi-Agent Deep Deterministic Policy Gradient
MARL	Multi-Agent Reinforced Learning
MAS	Multi-Agent System
MCA	Markov Chain Approach
MFRL	Mean-Field Reinforced Learning
MFVDN	Mean-Field Value Decomposition Network
MTMFQ	Multi-Type Mean Field Q-Learning
ODE	Ordinary Differential Equation
QMIX	Q-Mixing Network
QPLEX	Q-function with Parallelized Experience Replay and Loss Extrapolation
QTRAN	Quantum Temporal Recurrent Attention Networks
RF	Random Forest
RL	Reinforced Learning
RNN	Recurrent Neural Network
SARIMA	seasonal autoregressive integrated moving average
SARS	Severe Acute Respiratory Syndrome
SIR	Susceptible Infected Recovered
SEIR	Susceptible Exposed Infected Recovered
SEQIJR	Susceptible-Exposed-Quarantine-Infected-Isolation Recovered
SG	Stochastic Game

SVM	Support Vector Machine
TD3	Twin Delayed Deep Deterministic Policy Gradient
VDN	Value Decomposition Network
WHO	World Health Organization

LIST OF SYMBOLS AND NOTATIONS

y_{it}	Observed case indicator at location i, wave t
$\mathbf{S}_{it}$	Covariate vector at location i, wave t
u_i	Spatial random effect, location i
γ_t	Temporal fixed effect, wave t
ϵ_{it}	Spatiotemporal unstructured noise, location i, wave t
π_{it}	Fitted malaria risk probability, location i, wave t
ℓ	Location conditioning variable in the RL state/policy
β	Coefficients of covariates $\mathbf{S}_{it}$
ψ	Auxillary variable
σ	Variance
α	Acceptance probability
β	Coefficients of covariates $\mathbf{S}_{it}$
$\propto$	Proportionality symbol
$\forall$	For all
θ	Vector of parameters β
$Q(\cdot)$	Q-value (action-value function)

TABLE OF CONTENTS

DECLARATION i

COPYRIGHT ii

DEDICATION iii

ACKNOWLEDGMENT iv

ABSTRACT v

ABBREVIATIONS AND ACRONYMS vi

LIST OF SYMBOLS AND NOTATIONS ix

LIST OF TABLES xiii

LIST OF FIGURES xiv

LIST OF APPENDICES xv

CHAPTER ONE

1 INTRODUCTION 1

1.1 Background of the Study 1
1.2 Statement of the Problem 6
1.3 Objectives of the Study 7
1.3.1 General Objective 7
1.3.2 Specific objectives 8
1.4 Research Questions 8
1.5 Significance of the Study 8
1.6 Scope of the Study 9
1.7 Limitations of the Study 10
1.8 Thesis Outline 10

CHAPTER TWO

2 LITERATURE REVIEW **11**

2.1 Introduction . . . 11
2.2 Models for Infectious Disease Prediction . . . 11
2.2.1 Traditional Models . . . 11
2.2.2 Machine Learning Models . . . 14
2.2.3 Reinforcement Learning . . . 18
2.2.4 Integration of GIS and Reinforcement Learning . . . 20
2.3 Summary and Critique of Literature . . . 21
2.4 Research Gap . . . 23

CHAPTER THREE

3 METHODOLOGY **25**

3.1 Introduction . . . 25
3.2 GIS Enhanced Bayesian Reinforced Learning Model . . . 25
3.3 Estimation of Model Parameters . . . 31
3.3.1 Algorithm for Estimating σ_i^2 Parameters . . . 33
3.3.2 Algorithm for Estimating β_j Parameters . . . 35
3.4 Convergence of the Model Parameters . . . 36
3.4.1 Schwartz's theorem for posterior consistency . . . 36
3.4.2 Bernstein-von Mises Theorem . . . 38
3.4.2.1 RMSE Convergence . . . 38
3.4.2.2 Credible Interval Coverage . . . 38
3.4.2.3 Bias Properties . . . 39
3.4.3 Gelman-Rubin Statistic ($\hat{G}$) . . . 39
3.5 Model Operation . . . 39
3.5.1 Computational Cost and Scalability . . . 39
3.5.2 Identifiability . . . 40
3.5.3 High-Dimensional Inference . . . 40
3.6 Implementation of the Model . . . 40
3.6.1 State Space . . . 40
3.6.2 Action Space . . . 41
3.6.3 Reward Function . . . 42
3.7 Evaluation of Model Performance . . . 43
3.7.1 Simulation Study . . . 43
3.7.2 Evaluation Metrics . . . 46
3.7.2.1 Root Mean Squared Error . . . 46
3.7.2.2 Success Rate . . . 46

3.7.2.3 Infection Rate 47
3.7.2.4 Case Prediction Error 48
3.7.2.5 Cross-Location Adaptivity 48
3.8 Determination of Optimal Strategies 49
3.9 Data Description 50
3.10 Statistical Software 51

CHAPTER FOUR

4 RESULTS AND DISCUSSIONS 52
4.1 Introduction 52
4.2 GIS-Enhanced Bayesian Reinforcement Learning Model 52
4.3 Estimation of Model Parameters 54
4.3.1 Algorithm for Estimating σ_i^2 Parameters 56
4.3.2 Algorithm for Estimating β_j Parameters 58
4.3.3 Monte Carlo Simulation 59
4.4 Model Performance under Simulation 62
4.5 Malaria Data 66
4.5.1 Data Preprocessing 66
4.5.2 Parameter Estimation 67
4.5.3 Prescriptive Optimal Interventions and Prediction of Malaria Cases 68
4.5.4 Intervention Effect Gradients 70
4.6 Model Performance 71
4.7 Discussion 72

CHAPTER FIVE

5 CONCLUSIONS AND RECOMMENDATIONS 75
5.1 Introduction 75
5.2 Conclusion 75
5.3 Recommendations 76

References 78

APPENDICES 91

LIST OF TABLES

4.1 Parameter Estimation and its Evaluation Measures 61

4.2 Comparison of Model Performance 63

4.3 County predicted cases with corresponding optimal action taken by the population (GIS-RL Model) 66

4.4 Parameter Estimates and Credible Intervals for the Estimated parameters 68

4.5 Optimal Strategies and Predicted Malaria Cases 69

4.6 Model Performance 72

LIST OF FIGURES

3.1 Interaction between Vector and Human in the environment 26

3.2 Framework for Twin Deep Deterministic Policy Gradient Algorithm: Source (Li et al., 2024) 28

3.3 Spatial Distribution of Malaria Cases in Eight Counties 51

4.1 Traceplots for Model parameters 62

4.2 Success and Interaction Rates under RL-MAB and GIS-Enhanced reinforcement learning models 64

4.3 Reward values 65

4.4 Spatial Distribution of Predicted Malaria Cases in the Region . . 70

4.5 Intervention Gradients 71

LIST OF APPENDICES

Appendix I Publications

Appendix II Research License

Appendix III Python Algorithm (GitHub Link)

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Vector-borne transmitted infectious diseases are human illnesses caused by bacteria, parasites, and viruses; these diseases include Chikungunya, Dengue and severe dengue, Yellow fever, Zika virus, Malaria, Japanese encephalitis, Lymphatic filariasis, Leishmaniasis, Crimean-Congo hemorrhagic fever, Chagas disease (American Trypanosomiasis), Onchocerciasis, Trypanosomiasis, human African (sleeping sickness), Plague, and Schistosomiasis. Infectious diseases can be transmitted from person to person, animal to person, or person to animal after parasites and microorganisms infects animals or human beings (Bin et al., 2019). The disease has three key characteristics: infectiousness, epidemic, and uncertainty. World Health Organization (WHO) reported that infectious diseases, malaria, HIV/AIDS, lower respiratory infections, and tuberculosis were among the top ten leading causes of death worldwide in 2019; however, the number of deaths has been falling significantly where the most significant decrease is for HIV/AIDS by 59% fewer deaths in 2019 compared to 2000 (WHO, 2020a). WHO 2020 also reports that every year, more than 700000 deaths are attributed to diseases such as yellow fever, malaria, schistosomiasis, human African Trypanosomiasis, Chagas disease, leishmaniasis, dengue, onchocerciasis, and Japanese encephalitis (WHO, 2020b), with 2020 world malaria report indicating that more than 400000 deaths were due to malaria (WHO, 2020c). Infectious diseases have remained a

formidable challenge to public health globally (Lipsitch and Viboud, 2009); the key strategy for successful infectious disease prevention, control, and eradication is detailed research of the spread causes, spreading paths, spreading process, and epidemic laws. Therefore, accurate spread prediction is paramount for effective mitigation and intervention strategies (Corley et al., 2014). Active disease management benefits from accurate and timely predictions of where, when, and how badly contagious diseases will spread. To address the prediction challenge, innovative and interdisciplinary efforts are needed, where research studies have made significant valuable insights to this effort.

Researchers have developed mathematical studies of infectious diseases that primarily rely on infectious disease dynamics theory and methods, providing valuable tools for understanding epidemiological patterns and developing quantitative evidence for global health decision-making (Heesterbeek et al., 2015). Mathematical approaches to infectious disease prediction and control help to address public health challenges and optimize epidemic control strategies (Dimitrov and Meyers, 2010). Due to changing disease dynamics ranging from disease biology, environmental factors, and socioeconomic, disease modeling has been advancing since the development of the first compartmental model in 1927 by Kermack and McKendrick (1927), where they proposed Susceptible-Infected-Recovered (SIR) compartmental model for studying epidemic law of blacks deaths in Europe, the "threshold theory" also distinguished the spread of the disease. Since its proposal, data has proven Susceptible-Infected-Recovered model validity regarding large-scale infectious diseases; therefore, a deterministic model with differential equations has widely been adopted (Sun and Bin, 2017). With deepening research, factors that are considered in establishing mathematical models are evolving and increasing, and models dimensions are also increasing. In 1984 Susceptible-Exposed-Infective-Recovered (SEIR) model was proposed, the model is improvement of SIR mathematical model (Sun and Bin, 2017). The SEIR model takes into account the latency of infectious diseases, therefore making the model more realistic. Safi and Gumel proposed Susceptible-Exposed-Quarantine-

Infective-Isolation-Recovered (SEQIJR) for modeling Severe Acute Respiratory Syndrome (SARS) (Safi and Gumel, 2011), these proposed models assume disease modeling in discrete space.

Reversible Jump Markov Chain Monte Carlo algorithm which combines spatial and contact network model constructed under Bayesian context to provide a real-time risk prediction for emerging infectious diseases during an epidemic, the model showed greater capability than the traditional models, which did not consider the spatial aspect of the disease and uncertainty in model parameter estimation, however the proposed model have a high potential of bias and overfitting which leads to inaccurate predictions (Jewell et al., 2009). To address this challenge of overfitting and inaccuarate prediction of infectious diseases machine learning algorithms have been proposed for control of infectious diseases through spatial-temporal prediction of evolution and spread of infectious diseases (WHO, 2021). Machine learning algorithms are adept at analyze a large, intricate datasets and identifying trends and patterns that are challenging to be detected and predict. Machine learning techniques have shown promising results in predicting infectious diseases, considering demographics for human an vectors, environmental conditions, and individual behaviors. However, the availability of high-quality and comprehensive data is a key challenge. Infectious disease surveillance systems often collect data attributed to various factors, but this data can be incomplete, biased, or noisy, affecting performance of machine learning models. Additionally, some infectious diseases have long incubation period, making data about past outbreaks not accurately reflect current disease conditions. Researchers have recently employed various machine learning algorithms, such as reinforced learning, decision trees, deep-learning networks, support vector machines, and random forests to overcome these challenges. These algorithms have been applied to datasets, such as epidemiological data, social media posts and genomic data. Chae et al. (2018) proposed Deep Learning and Big Data for predicting infectious disease with deep neural network (DNN) outperforming long-short term memory (LSTM) model when infectious disease is spreading. Sharma used Deep Ensem-

ble Learning Methods to analyse and predict COVID-19 multivariate data which help to increase accuracy and reduce training and prediction time (Sharma et al., 2023).

Machine-learning algorithms have demonstrated their capability of accurately predicting the onset and spread of infectious diseases, showing more excellent performance compared to traditional statistical models (Palaniappan et al., 2022). Algorithms have also been utilized in the identification of potential outbreak sources based on pathogen genetic makeup and infection patterns (Ghannam and Techtmann, 2021; Roy et al., 2021); additionally, machine learning algorithms have been used in predicting individual's risk of contracting infectious disease due to personal characteristics and behaviours (Atkinson et al., 2023). Machine learning is, therefore, an exciting area of research for predicting infectious diseases, used in epidemiology, clinical studies, and public health. ML learning however faces some challenges which includes lack of high-quality data, complexity of the underlying phenomena, biases, and overfitting (Roth et al., 2018). These lead to less accurate predictions as a consequence of incomplete records, missing values, and varying data formats. The complexity of transmission of infectious diseases also makes it challenging to develop accurate models (Roth et al., 2018). Geographic information systems (GIS) and remote sensing technology are increasingly being applied to analyze disease geography, focusing on relationships between pathological and environmental factors. However, full GIS potential will only be realized when environmental and disease surveillance systems are developed, which distribute data on environmental conditions, disease agent, and health outcome over time based on user-defined queries for selected geographical areas (Cromley, 2003). Studies have demonstrated that use of GIS and dynamic modelling algorithms can offer a well-timed solution in understanding and tracking disease outbreaks, vulnerabilities to population health, spread and adaption (Saran et al., 2020). Due to GIS importance in disease distribution and pattern prediction it is significant to integrate GIS in developing new machine learning algorithms for prediction of infectious disease to improve on accuracy and repre-

sentative prediction results.

Bayesian statistics is method of data analysis and estimation of parameter using Bayes theorem (van de Schoot et al., 2021). The Bayesian approach has been adopted due to its ability to use prior knowledge, therefore providing a framework for handling uncertainty and sequential learning; it also ensures that there is an updating of beliefs as new data becomes available. The Bayesian framework is useful for parameter estimation and regularization which prevent overfitting with limited data (Bharadiya, 2023). The Bayesian technique is useful in reinforcement learning for logically updating uncertainty and estimating parameters. It entails maintaining a probability distribution over an agent's beliefs about the environment and potential behaviors. The distribution is consistently updated when the agent interact with environment, absorbing new observation and developing grasp of the underlying dynamics (Asheralieva and Niyato, 2020).

Recently, researchers have developed new algorithmic models integrating GIS to predict disease patterns and transmission. Khiem et al. (2022) proposed GIS - machine learning model for predicting occurrence of infectious disease in Vietnam, and the study determined that the neural networks perform better in predicting the infection. Karras et al. (2023) developed integrated GIS-based reinforced learning with Multi-Armed Bandit (MAB) algorithm for efficient prediction of disease transmission. The algorithm shows more excellent performance compared to Random Forest, SVM, Neural Network, Markov process, and RL (Q-Learning); however, the superiority of Multi-Armed Bandit was context-specific with possible high overestimation bias. Also, the lack of current and high-quality data affects the accuracy of the algorithm. Therefore, from the literature, there has yet to be a defined and accurate algorithm for the prediction of the spread and transmission of vector-borne infectious diseases, and there is a need for further research for classical models and machine learning algorithms to improve prediction accuracy. Additionally, application of integrated GIS-Bayesian reinforced learning is still in its development stages.

Within the unique context of Kenya, with diverse environmental and socio-

economic factors that plays a critical role in disease dynamics, there is need for advanced predictive model. Therefore, to enhance accurate and adaptive prediction of vector-borne infectious disease while reducing overestimation biases, the study developed a GIS-Enhanced Bayesian reinforced learning model.
The motivation of enhancing the model was:

a) To develop a model with more excellent learning capabilities for accurate prediction of spread of vector-borne infectious diseases.

b) To determine optimal strategies that reduce the spread of vector-borne infectious diseases.

The research is a proactive step towards strengthening global efforts to combat vector-borne infectious diseases. By addressing the gaps in current predictive models, the study aims to empower public health authorities and policymakers with timely, accurate, and actionable information.

1.2 Statement of the Problem

Vector-borne infectious diseases present a significant threat to public health in Kenya and worldwide; therefore, making accurate and timely spread prediction is critical for effective intervention strategies (Corley et al., 2014). Prediction of infectious disease has relied on traditional models that struggle to capture the dynamic interplay of socioeconomic and environmental factors, and the development of accurate and efficient predictive models for vector-transmitted infectious diseases has been the latest trend in disease modeling. Compartmental models have shown extreme flexibility and adaptability in predicting infectious disease in a discrete space (Zhang et al., 2022), and machine learning models such as neural networks, deep neural networks, deep learning, and Q-learning have also been used mainly in discrete space (Chae et al., 2018; Sharma et al., 2023) to address the problem of the assumption of homogenous mixing and fixed parameters associated with compartmental models. Despite the excellent performance of the

ML model in the prediction of disease cases, there has been no framework that simultaneously captures interaction-level (host-vector interaction) transmission dynamics in a continuous state-action space while learning optimal and location-specific intervention rates for vector-borne infectious disease control in Kenya. This is because the existing vector-borne infectious disease prediction models represent transmission as an averaged, population-level process that relies on discretized states and fixed parameters that overlook host-vector interaction dynamics. Recent machine learning frameworks integrate diverse environmental data but stop short of modeling host-vector contacts as one of the determinants of case occurrence, also failing to learn intervention strategies that are spatially distinct to different locations characterized by varying environmental drivers. These limitations associated with the existing models therefore introduce a systematic estimation bias in predicted values, failing to reflect a fine-grained mechanism that ideally generates infectious disease cases, and cannot adapt intervention decisions to the environmental heterogeneity present across different locations, and this limits their utility for guiding a targeted location-specific disease control rates of interventions. This research addressed the gap by developing a GIS-integrated reinforcement learning framework for predicting spread of disease cases which considers host-vector interactions and records interactions as cases and no interaction as no-cases, using 5-D3 with Bayesian and min-max ensemble estimation to learn adaptive intervention rates across the environmentally heterogeneous vector-borne infectious disease burden locations in Kenya.

1.3 Objectives of the Study

1.3.1 General Objective

The general objective of the study was to develop a GIS-Enhanced Bayesian reinforcement learning model for predicting spread of vector-borne infectious diseases: a case of malaria in Kenya.

1.3.2 Specific objectives

The specific objectives of the study were;

i. To develop a GIS-Enhanced Bayesian reinforcement learning model.

ii. To derive estimators of parameters for the developed model.

iii. To assess performance of the developed model using simulation and real world dataset.

iv. To determine optimal strategies that reduce the spread of vector-borne infectious disease.

1.4 Research Questions

i. How accurately can GIS-Enhanced Bayesian reinforced learning model predict the spread of vector-borne infectious disease?

ii. What are the estimators of the model parameters?

iii. How is the model performing under simulated and real world dataset?

iv. What are the optimal strategies for reducing spread of vector-borne infectious disease?

1.5 Significance of the Study

Epidemiological modeling has progressed from classical compartmental frameworks such as SIR and SEIR-type models. These models represent disease spread as averaged, population-level flows between discretized states that are governed by fixed transmission parameters to offer analytical tractability. This, however, hides underlying host-vector contact that generates disease cases. This development was followed by statistical and machine learning models that incorporate environmental and climatic factors such as temperature, rainfall, vector density

and more to improve prediction accuracy and capture non-linear relationships between the environment and disease risk. This innovation moves away from fixed-parameter assumptions. These developments have largely been predictive and population-averaged, and have treated environmental disease drivers as predictors of disease incidence instead of modeling the interaction-level transmission that generates disease, and stopping short of learning adaptive and location-specific intervention policies within a continuous state-action space. Integrating Geographic Information System (GIS) into reinforcement learning offers significant means to address some of the areas that the ML model has not covered well, ranging from consideration of host-vector interaction during prediction to reducing systematic biases. Bayesian estimation for reinforcement learning was employed to address overfitting and parameter estimation, and min-max ensemble estimation to reduce overestimation bias in reward estimation. This development ensured that the model estimated accurate prediction values and learned adaptive intervention rates across the environmentally heterogeneous vector-borne infectious disease burden locations in Kenya. The study makes a novel contribution to the public health novel contribution to public health (SDG 3 Good health and well-being) and academic literature in bridging the gap between traditional epidemiological models and cutting-edge machine learning techniques for predicting the spread of vector-borne infectious diseases.

1.6 Scope of the Study

The study developed a GIS-enhanced Bayesian Reinforcement learning model for predicting the spread of vector-borne infectious diseases such as dengue fever, chikungunya, zika virus, and malaria. The model provides the number of disease cases and identify high-risk areas by integrating spatial, environmental and socioeconomic data. Geographical focus was across eight regions (Kisumu, Siaya, Homa Bay, Migori, Kakamega, Bungoma, Busia and Vihiga) in Kenya that has many cases of malaria, considering environmental conditions and socioeconomic

factors. The model was tested using simulated data and validated using historical real world dataset from DHS 2015 - 2020, serving as a decision-support for public health policymakers.

1.7 Limitations of the Study

The study successfully developed a model for predicting the spread of vector-borne infections diseases and estimating optimal strategies that humans can adopt to prevent the spread of infectious diseases. The study, however, had the following limitations

i. Prediction of disease cases was performed with consideration of two interventions (spray repellent and treated mosquito nets) based on the available malaria dataset. With the availability of data related to other scientifically proven interventions (such as vaccination, larva management) the model should be replicated.

ii. The model is computational intensive and requires longer training time for accurate estimation.

1.8 Thesis Outline

This thesis comprises five chapters, including this chapter. Chapter one presents the introduction of the study, chapter two presents the literature review of models for predicting the spread of infectious disease and integration of GIS to machine learning models, chapter three presents methodology and important concepts used throughout the thesis, chapter four presents GIS-Enhanced Bayesian Reinforcement Learning Model and the model results, and chapter five presents the conclusions and recommendations drawn from the study.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The chapter discusses models that have been developed for prediction of infectious diseases and identifies the research gap to be addressed by the study.

2.2 Models for Infectious Disease Prediction

Vector-borne infectious diseases are human illnesses caused by bacteria, parasites, and viruses. The diseases can be transmitted from person to person, animal to person, or person to animal after parasites and microorganisms infects animals or human beings (Bin et al., 2019). The infectious diseases includes Chikungunya, Dengue and severe dengue, Yellow fever, Zika virus, Malaria, Japanese encephalitis, Lymphatic filariasis, Leishmaniasis, Crimean-Congo hemorrhagic fever, Chagas disease (American Trypanosomiasis), Onchocerciasis, Trypanosomiasis, human African (sleeping sickness), Plague, and Schistosomiasis.

2.2.1 Traditional Models

Development of mathematical models for modeling infectious diseases has recently gained much attention from researchers (White, 2017). A model with a system of differential equations (ODE) was one of the earlier models developed to model disease transmission Kermack and McKendrick (1927). However, since

its development in 1927, there has been renewed interest in developing new models to improve the capabilities of early models, including Susceptible Infectious Recovery (SIR) models. The approaches that have been used to improve the performance of traditional differential equations (ODE) and compartmental models are the addition of parameters for disease compartments to capture different disease dynamics ranging from environmental factors to disease structure.

Zhang et al. (2022) proposed the compartmental model to predict infectious disease due to its flexible framework and high adaptivity that can allow additional approaches such as Bayesian inference and mixed-effects modeling; the study results show that compartmental model has highly adaptive powerful tool making it flexible to be customized. Osman and Adu (2017) developed the Susceptible Exposed Infectious Recovery - Susceptible Exposed Infectious (SEIR-SEI) model, which is an extension of SEIR to model malaria transmission; they demonstrated that the model is locally and asymptotically stable when reproduction numbers are less than one and unstable when the numbers are greater than one. The model has been used to describe 2016 outbreak of the Zika virus in Brazil Dantas et al. (2018); the study found that the calibrated results presented realistic parameters with reasonable descriptions. However, the need for more data for initial conditions made calibration more complex. From the review of compartmental models, the study observes that the models assume constant parameters and homogeneous mixing, which is unrealistic with vector-borne diseases where interaction between vectors and hosts is largely heterogeneous. Additionally, the models lack spatial heterogeneity and do not adequately incorporate environmental and behavioral factors.

To address these shortcomings associated with compartmental models, researchers have proposed using non-compartmental models such as linear models, time-series analysis, and stochastic models Paolella (2018) for predicting the spread of infectious diseases. Anokye et al. (2018) forecast monthly malaria incidence in Kumasi Metropolis using the Autoregressive integrated moving average (ARIMA), the model forecasts future incidences at time points. Although the model effec-

tively predicted malaria cases, it fails to draw a comparison between different locations. Mopuri et al. (2020) applied seasonal autoregressive integrated moving average (SARIMA) and multi-step polynomial regression and models to forecast malaria cases. The study found that polynomial model had high predictive power in predicting malaria cases, and population and environmental factors such as mean temperature, rainfall, and normalized difference vegetation index build play significant role in malaria transmission. Application of SARIMA in predicting malaria cases using monthly environmental variables demonstrated that total rainfall is the most significant predictor of malaria; this indicates that incorporation of environmental factors can significantly improve the performance of predictive model Adeola et al. (2019).

Arenas et al. (2020) proposed the use of a Microscopic Markov Chain Approach (MMCA) metapopulation mobility model, which is a stochastic model to predict the incidence of epidemics in a spatial population through time; the model was modified to capture human behaviour characteristics such as daily recurrent mobility flows, age strata, urban demography, and age-structured contact patterns, the study improved estimates of epidemiological parameters. From the demonstration of the great performance of stochastic models, studies have continued to develop models that utilize stochastic, metapopulation, multiscale, agent-based, and approaches to simulate the global spread of infectious disease, Balcan et al. (2010) developed a Global Epidemic and Mobility (GLEaM) computational model for predicting the spread of infectious diseases. The model integrated socio-demographic and population mobility; the approach simulated spread of epidemics at worldwide scale. It also suggested a flexible structure that can include different disease structures and local intervention policies, therefore making Global Epidemic and Mobility suitable for computational modelling and anticipation of global epidemic spreading spatio-temporal patterns, understanding of historical epidemics and assessing the role of human mobility with respect to global epidemics.

Bayesian methods have also been used in predicting spread of infectious diseases.

Stojanovic proposed a Bayesian Monte Carlo approach to derive a spatiotemporal kernel for the prediction of the spread of infectious diseases (Stojanovic et al., 2019); the study showed that the spatiotemporal kernel model performed at par with the hhh4 model however, the employed Bayesian Monte Carlo framework is easily extensible and allows for incorporating prior domain knowledge, making it suitable for use on limited, yet complex data sets in epidemiology. Funk et al. applied the Bayesian semi-mechanistic model to Ebola forecasting (Funk et al., 2018). A significant number of existing studies often focus on classical statistical methods in estimation, leaving room for incorporation Bayesian frameworks in disease modeling. Bayesian methods have great power in parameter estimation and uncertainty quantification (Gelman et al., 2014).

Although mathematical and compartmental models are sufficient in modelling infectious diseases, the models are less suited to train Reinforced Learning (RL) agents as they lack randomness. Mathematical models does not incorporate super-spreaders and also lack randomness (Galvani and May, 2005). Additionally, RL environments need to be dynamic to explain the dynamics associated with disease and population dynamics. ODE models are static, and to transform it to dynamic state require addition of parameters Giordano et al. (2020) that often become complex. Therefore, it is prudent to avoid implementation of ODE models and instead adopt implementation of virtual environment that will easily mimics disease spread. The use of equation-based model has been used extensively to model diseases; however, due to non-linearity and complex patterns in infectious disease outbreaks and advancement in technology, machine learning and agent-based models have recently been adopted to improve prediction capability and accuracy (Hunter et al., 2020; Santangelo et al., 2023).

2.2.2 Machine Learning Models

To improve the performance of compartmental models, Shi et al. (2016) integrate machine learning and Least Absolute Shrinkage and Selection Operator (LASSO)

models to predict dengue incidences in Singapore; the study demonstrated that the model effectively predicts dengue disease. However, the extension of the forecast beyond 12 weeks was not possible because local weather conditions have a short-term but significant effect on the dengue disease, which requires the integration of the prediction model with the weather forecasting model. Chen et al. (2018) demonstrated that LASSO model was effective in predicting trends of dengue diseases in Taiwan, Japan, Singapore, and Thailand; although the model was better in short-term prediction, the model was less effective in longer-term predictions, therefore requiring this prediction approach to be implemented repeatedly. Although LASSO model has demonstrated good performance, the comparison of ARIMA, LASSO, Random Forest (RF), and Random Forest with Uncertainty Forest Analysis (RF-UFA) models on the prediction of weekly dengue disease cases showed that the LASSO model was less effective, however for near-term prediction, Random Forest (RF) had more prediction accuracy than Poisson regression and autoregressive integrated moving average models, autoregressive integrated moving average performed better for long-term predictions, and Random Forest with Uncertainty Forest Analysis outperformed RF and logistic regression in predicting dengue outbreaks (Benedum et al., 2020).

Kolesnikov et al. (2019) compared long-short term memory (LSTM), seasonal autoregressive integrated moving average (SARIMA), Random Forest (RF), gradient boosting (XGB), LightGBM, K-Nearest Neighbors (KNN), Keras, and CatBoost models in predicting the incidences of dengue diseases; the study demonstrated that the effective predictions were attributed to a mathematical model that combines spatial analysis techniques and neural networks based on Long-Short term memory (LSTM) architecture. Random forest (RF) algorithms and gradient boosting (XGB) are considered universal for space-time prediction. Other researchers who have demonstrated the effectiveness of the Long-short term memory (LSTM) model in predicting incidences of dengue include the work of Xu et al. (2020), Nguyen et al. (2022) compared SARIMA, Long-short term memory (LSTM), Long-short term memory with Attention (LSTM-ATT), gradient

boosting (XGB), Poisson regression, SVR-L, Support Vector Regression (SVR), and Convolutional Neural Network (CNN) and Long-short term memory with Attention (LSTM-ATT) demonstrated higher performance. Some research, however, failed to identify the efficient model for predicting dengue incidences, Patil and Pandya (2021) compared performance of multiple linear regression, polynomial regression, Seasonal Autoregressive Integrated Moving Average (SARIMA), Autoregressive Integrated Moving Average (ARIMA), Support Vector Regression (SVR), PROPHET, Random Forest (RF), ElasticNet regression, and decision tree regression, they demonstrated that despite high accuracy no model prevailed over the other, but their performance differs depending on the city. However, a study by Roster et al. (2022) demonstrated that Random Forest (RF) performs well in forecasting dengue incidences, but the model did not capture temporal dependencies in the data.

A study to compared the performance of deep neural network (DNN), long-short-term memory (LSTM), and autoregressive integrated moving average (ARIMA) models in predicting the spread of infectious disease found that deep neural network (DNN) and long-short term memory (LSTM) performed better than autoregressive integrated moving average (ARIMA) (Chae et al., 2018). Although the study showed that deep neural network is stable and long-short term memory (LSTM) model is accurate when infectious disease is spreading, the study considered a narrow range of parameters in deep learning model with less subdivision of spatial ranges; this makes it difficult to conclude that a most effective model was created. Santosh et al. (2020) used long short-term memory (LSTM) on a big data platform to model malaria transmission; the study found that clinical and environmental variables are vital in enabling the presence and transmission of malaria. The effectiveness of long short-term memory (LSTM) in prediction of infectious diseases have also been demonstrated in other research works by Chimmula and Zhang (2020), Guo et al. (2020), Masum et al. (2020), Adiga et al. (2021), Goo et al. (2021), Kumar et al. (2021), Shen et al. (2022).

Although long short-term memory (LSTM) has excelled at capturing the temporal

dependencies of disease, the model has struggled to capture dynamic and complex nature of infectious diseases. Kamana et al. (2022) compared performance of long short-term memory (LSTM) and LSTM Sequence-to-Sequence (LSTM-Seq2Seq) deep neural network model for accessing the impact of climate and prediction of malaria cases; they demonstrated that with consideration of climatic factors, LSTMSeq2Seq improved prediction of malaria re-emergence. Therefore, LSTMSeq2Seq can effectively be used in predicting re-emergence of malaria. The model, however, require incorporating an attention mechanism that will improve the model's ability to capture temporal dependencies and variations in the data more effectively.

Studies have also proposed disease-specific models, Shaier et al. (2021) developed a Disease Informed Neural Networks (DINNs) as a reliable candidate model for effective learning of dynamics of spread, progression forecast for infectious diseases to future and finding disease unique parameters such as death rate. Applying Disease Informed Neural Networks (DINNs) to zika and dengue infectious diseases, the study found that Disease Informed Neural Networks tend to perform highly with initial parameter guesses. From the study findings, to improve the effectiveness and reliability of the forecast, high-resolution and real-time data must be incorporated to enhance model interpretability. Nkiruka et al. (2021) proposed a machine learning-based model called Malaria Incidence Classification (MIC) which adopts k-means clustering process for outlier detection, and XGBoost algorithm by Chen and Guestrin (2016) for classification of malaria incidence using climate variability. The study demonstrated that the model outperformed Naïve Bayes, Logistic regression models, and Support Vector Machine in classification and prediction. Results also suggested that incorporating temporal dynamics and geospatial data utilization allows the model to capture complex interactions between temporal trends, climatic factors, and spatial distributions. The modification will improve the prediction capabilities of the model.

Other research on predicting vector-borne infectious diseases, Brock et al. (2019) study demonstrated that the Boosted regression tree model effectively predicted

malaria outbreak in Malaysia. However, the model must incorporate temporal dependence crucial for accurate malaria prediction. Ajith et al. (2020) demonstrated that the Random Forest (RF) method is one of the best models for accurately and effectively predicting West Nile virus cases. The model did not include predicting variations of West Nile virus cases over different locations and times. Therefore, including spatial-temporal features could lead to more accurate and effective predictions. Roster et al. (2022) who demonstrated that Random Forest (RF) performs better than Neural Network models in predicting Zika virus, though the model depends on the quantity of data available. The models have demonstrated more excellent performance in predicting infectious diseases with less capability in controlling dynamic and uncertain environments.

Keshavamurthy et al. (2022) found that machine learning and algorithms significantly predict infectious diseases and improve capability in handling diverse, large data sets and uncovering connections in the data. However, machine learning algorithms must develop interconnections between vector, human, and environmental factors that drives disease presence. Because of this, Keshavamurthy et al. (2022) has suggested building robust artificial intelligence and data science capacities in resource-scare settings across all regions in the world and diseases to capitalize on Machine Learning and Deep Learning techniques potential prediction of Infectious Disease.

2.2.3 Reinforcement Learning

Reinforcement learning techniques have demonstrated better capabilities compared to other machine learning techniques. Reinforced Learning offers unique advantages for tasks involving dynamic environments, sequential decision-making, sparse rewards, and long-term planning. However, it comes with challenges of exploration instability, sample inefficiency, and need for tuning and careful reward design. Reinforced Learning (RL) has been applied in disease diagnostic, progression, and disease treatment Li (2017) where Single-Agent Reinforced Learning has

been extensively applied (Zhou et al., 2023). Mnih et al. (2013) proposed Deep Q-Network (DQN), which combines deep neural networks and Q-Learning. Some of these studies have adopted the Multi-Armed Bandit (MAB) technique (Karras et al., 2023); however, this technique suffers from overestimation bias in deep Q-learning. Deep reinforced Learning has been applied for malaria likelihood prediction based on household questions (Rajpurkar et al., 2017); the study demonstrated that learning DQN generates good results. Due to increased complexity in modeling real-life problems, studies have demonstrated that Single-Agent Reinforced learning provides inadequate results in complex environments where agent cooperation is needed (Bucsoniu et al., 2010; Zhou et al., 2023). Multiple studies have developed Multiple-Agent Reinforced Learning (MARL) to address this complexity (Zhou et al., 2023).

Multiple agents have been applied in the field of transportation such as traffic light control Li et al. (2016); Zhu et al. (2022), Public Health and Medical Diagnosis Khalilpourazari and Doulabi (2021); Khalilpourazari and Hashemi Doulabi (2021), manufacturing such as industrial robots Agrawal et al. (2021), Portfolio management in financial trade Pham et al. (2021), and in smart education. Agents in Multi-Agent Systems (MAS) employ trial-and-error contact with the environment to solve sequential decision problems. However, due to complexity of Multiple-Agent Learning, a stochastic game (SG) was proposed to model multi-agent sequential decision problems of non-Markovian nature (Hu et al., 1998; Bowling and Veloso, 2000). Tampuu et al. (2017) extended Deep Q-Network to multi-agent scenario and equipping an independent Deep Q-Network for each agent, where it consider agent's interaction with the environment, demonstrating that the training algorithm produce good results for simple MAS but faces difficult in convergence for complex tasks and with credit assignment issue. (Fujimoto et al., 2018) developed a Twin Delayed Deep Deterministic Policy Gradient (TD3) technique that proves highly effective in adaptive learning environments with significantly small overestimation bias in deep Q-learning. Despite its success in different applications, a noticeable dearth of literature applying the TD3

Policy Gradient to infectious disease prediction. The identified gap present an opportunity to innovate and extend application of reinforcement learning in public health contexts.

2.2.4 Integration of GIS and Reinforcement Learning

Geographic Information System (GIS) is a technology that creates, manages, analyses, and maps data. It connects data to map, integrating location data with all types of descriptive information, providing the foundation for mapping and analyzing used in science. Geographic Information System helps understand patterns, relationships, and geographic context, improving communication, management, efficiency, and decision-making (Esri, 2024). GIS and remote sensing technology are increasingly applied in analyzing disease geography, focusing on relationships between pathological and environmental factors. A full GIS potential can be realized when environmental and disease surveillance systems are developed, distributing data on environmental conditions, disease agents, and health outcomes over time based on user-defined queries for selected geographical areas (Cromley, 2003). Utilizing Machine Learning (ML) and Deep Learning (DL) for infectious disease prediction fully, the models must include the entire disease ecology in a One-Health context and underrepresented hotspots Keshavamurthy et al. (2022). Studies have demonstrated that using GIS and dynamic modeling algorithms can offer a well-timed solution in understanding and tracking disease outbreaks, vulnerability to population health, spread, and adaption (Saran et al., 2020). Studies have also shown that spatial information is significant in predicting spread and understanding the disease pattern Chowell and Rothenberg (2018). Due to the demonstrated importance of GIS in disease distribution and pattern prediction, it is significant to integrate GIS in developing new machine learning algorithms for predicting infectious diseases to improve accuracy and representative prediction results.

GIS integration in reinforced learning has been applied in different fields. A GIS-

based reinforcement learning model was derived for optimal planning of a rooftop PV system; the model demonstrated good results where the model outperformed existing models in volatile scenarios with lower solar radiation (Jung et al., 2021). A study on modeling of forest cover change demonstrated that integrating GIS and reinforced learning captures emerging forest cover patterns and the ability to evaluate trade-offs between different harvesting objectives (Bone, 2009). Khiem et al. (2022) proposed a GIS-machine learning model for predicting the occurrence of infectious disease in Vietnam, and the study determined that the neural networks perform better in predicting the infection. Karras et al. (2023) developed integrated GIS-based reinforced learning with Multi-Armed Bandit (MAB) algorithm for efficient prediction of disease transmission. The algorithm shows more excellent performance compared to Random Forest, SVM, Neural Network, Markov process, and RL (Q-Learning); however, the superiority of Multi-Armed Bandit was context-specific with possible high overestimation bias, technique struggle with exploration-exploitation trade-off in complex environments. Therefore, from the literature, there has yet to be a defined and accurate algorithm for the prediction of the spread and transmission of vector-borne infectious diseases such as malaria, and there is a need for further research for classical models and machine learning algorithms to improve prediction accuracy. Additionally, the integration of Geographic Information System (GIS) in reinforcement learning and the application of integrated GIS-Bayesian reinforced learning model in the prediction of malaria is still lagging.

2.3 Summary and Critique of Literature

In recent years, the development of predictive models for infectious diseases has received significant attention. Models developed after 1927 including discrete deterministic models have been found to be less flexible in continuous real world environment and problem-specific with restrictive generalization, this is due to models assumption of a "well-mixed" population, where individual have equal

probability of contacting each other making it untrue in reality given that in real life there exist host and vector heterogeneity, it also ignore randomness of transmission while treating it as fixed, treatment of variable parameters as fixed values also create uncertainty and systematic bias, the models also assumes existence of a stable, disease-free equilibrium state yet in reality population dynamics are non-stationary. Continuous compartmental models faces a number of problems that include: reliance on the averaged rates, therefore failing to capture parasite replication dynamics (systematic bias), assumes that key transmission parameters are constant, therefore ignoring seasonal changes, intervention rollouts and this limits long-term forecasting, failure to capture spatial heterogeneity that is individual-level differences in exposure and behavior, assumes that parameters are fixed, therefore overlooking parameter uncertainty, struggles in the integration of diverse and high-dimensional data such as climate. Researchers have significantly developed accurate models, providing better prediction results compared to traditional compartmental models. Adding parameters for disease compartments and considering environmental effects in predictive models has also significantly improved model performance. From the literature, research has demonstrated that compartmental models struggle in capturing intricate characteristics and non-linear dynamics associated with the disease transmission due to model's rigid, pre-defined structures; classical statistical models lack the flexibility required to model complex, non-linear interactions between the host and vector of the disease. To address the shortcoming associated with compartmental and classical statistical models, machine learning models have been proposed to significantly improved the prediction of infectious diseases, where ML has improve analysis of complex patterns in disease data with great error reduction, as demonstrated by deep learning models, ML models also integrate diverse data such as climate, land use, and vector density for accurate advanced prediction of disease outbreaks, capturing non-linear dependencies for timely and actionable recommendations for public health responses. Literature also shows that datasets used to train these ML have been discrete in nature, making the ML models applied in a continu-

ous reality to discretize the space for prediction problem solvable with classical parameter estimation and overestimation biases.

2.4 Research Gap

With application of the machine learning and reinforcement learning in prediction of infectious disease cases, models that have been developed are still limited in their prediction capability in two areas identified from the literature (1) Literature shows that reserchers have continued to model disease transmission as an averaged, population-level rate process instead of considering it as the outcome of discrete host-vector interaction events, this was demonstrated by the compartmental models assuming homogeneous mixing, with advanced machine learning capbable of integrating climate, land use, and vector density data predicts outbreak as aggregate, without considering varying locations and time where host|vector contact occured. This hides underlying behavioral and spatial heterogeneity, which drives transmission dynamics in the real world. This is due to the fact that not all host-vector proximity/contact results in disease transmission, and assuming that all exposures are uniformly risky introduces systematic biases to the predicted values. (2) Most of the prediction methods discretize the continuous disease and environmental states to make the estimation solvable by using classical parameter estimation methods. Discretization, treatment of transmission and intervention parameters as fixed and not adaptive, introduces overestimation bias, therefore limiting the model's ability to be generalized across spatially heterogeneous environments or build accurate location-specific public health interventions and make exact decisions relating to disease. Consequently, no existing framework that represents host-vector interactions as a unit of case occurrence and learns, in a continuous state-action space, the optimal intervention rates required for spatially different locations whose environmental factors are associated with transmission varies. The research addressed this gap by developing a GIS-integrated reinforcement learning model that is capable of recording

host-vector interactions and labeling them as cases and no cases if the interaction has not happened, using and modifying the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm to learn optimal, location-specific intervention rates across different region spatial polygons, each characterized by distinct environmental conditions. The Bayesian approach and min-max ensemble method to reduce overestimation bias in the estimation of future reward estimates. The proposed approach addressed discretization and homogeneity of data to ensure that there are adaptive, spatially differentiated intervention strategies grounded in continuous, interaction-level transmission dynamics and predict the resulting number of cases folowing optimal interventions selection. The study was built from the work of Zhang et al. (2021); Karras et al. (2023).

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter presents the concepts used to achieve the study's objectives. The developed model for predicting the spread of vector-borne infectious disease, algorithms, and Bayesian estimation are discussed.

3.2 GIS Enhanced Bayesian Reinforced Learning Model

The environment where hosts and disease vectors exist is dynamic and continuous due to population migration, population density, vector behavior, host-vector interactions, spatial dynamics factors, temperature, humidity, and terrain, which vary continuously across the space. Due to these dynamic factors, there is need for development of a dynamic model that can capture the dynamic movement of factors and the spatial nature of disease spread to improve disease prediction and identify optimal interventions/actions to reduce the spread of disease. A dynamic reinforced learning environment described by a virtual environment mimicking disease transmission was implemented. Virtual environment in Figure 3.1, is designed as a 2-dimensional grid where human can move randomly and interact. The virtual environment generates states and the results based on actions taken

by human-agent.

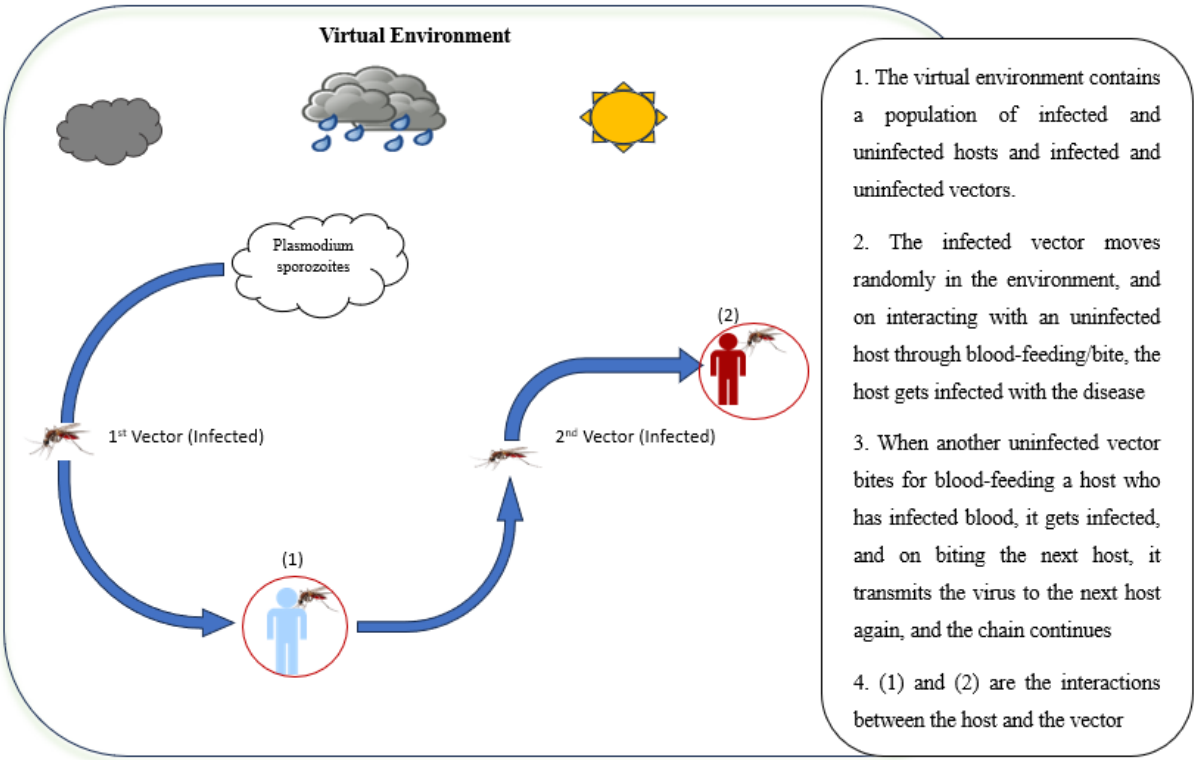


Figure 3.1: Interaction between Vector and Human in the environment

The model has the following assumptions

i) In the environment there exist an infected vector.

ii) Vectors and human moves randomly in the environment.

iii) The interaction between vector and human is involuntary.

iv) On interaction between vector and human, the host get infected with parasite with a probability (with probability ≈ 0.3 (Churcher et al., 2017)).

Deriving the study model, the study build on the methodology developed by (Karras et al., 2023), where Karras et al. (2023) proposed multi-Armed Bandit (MAB) approach for predicting disease transmission in aquatic, which requires agents to select among various strategies to optimize reward over time where reward is influenced by spatial effect of the selected strategy. Since the study is considering the prediction of disease in a dynamic environment, an approach

designed for continuous action space is needed. Therefore, a modified TD3 algorithmic model proposed by Fujimoto et al. (2018) that is utilized in a continuous control domain was used in the study. Twin Deep Deterministic Policy Gradient is an actor-critic, off-policy model-free reinforcement learning described as

$$\begin{aligned} \Delta\theta_i &= \nabla\theta_i(y_t - Q_{\theta_i}(s_t, a_t))^2 \\ y_t &= r_t + \gamma \min_{i=1,2} Q_{\theta_i'}(s_{t+1}, a_t^{'}) \\ a_t^{'} &= P_{\phi'}(s_{t+1}) + u \end{aligned} \tag{3.1}$$

where the error u is a clipped error that obeys normal distribution given by $u \sim clip(N(0, \sigma^2), -b, b)$

r_t agent immediate reward from taking action a_t in state s_t, $\gamma \in (0, 1)$ is discount factor which determines how future rewards are discounted relative to immediate rewards, θ_i's and ϕ are random parameters for updating target networks, current and next state s_t, s_{t+1}, current and next action a_t, $a_t^{'}$, truncated value b with boundary $-b$ and b (small constants greater than zero), P is a target policy where $P_{\phi'}$ is target actor network policy.

Action a_t is calculated at state s_t based on actor network P_ϕ, the objective action a' is computed through actor target network $P_{\phi'}$ at state s_{t+1}. $Q_{\theta_i}(s_{t+1}, a_t)$ and the target $Q_{\theta_1}(s_{t+1}, a_t')$ Q-values determines action-quality in s_t, and actor network policy update. y_t is obtained through selection of minimum Q-value of ($Q_{\theta_1'}$ and $Q_{\theta_2'}$) estimated by critic target network and apply it to $\Delta\theta_i = \nabla\theta_i(y_t - Q_{\theta_i}(s_t, a_t))^2$. The algorithm updates θ_i based on $\Delta\theta_i$ of the critic network parameters, that is obtained through the gradient increase of the loss equation $(y_t - Q_{\theta_i}(s_t, a_t))^2$ for $i = 1, 2$.

The information obtained from the update step, actions, current states, rewards, and next states are stored in tuple format in experience replay buffer. The data needed for updating actor and critic network parameters are then randomly sampled from experience replay buffer.

The soft update on actor and target networks are described as

$$\begin{aligned} \Delta\phi &= \nabla_{\phi} Q_{\theta_i}\left(s_t, P_{\phi}\left(s_t\right)\right) \qquad i = 1, 2 \\ \theta_i' &\leftarrow \delta\theta_i + (1-\delta)\theta_i' \\ \phi' &\leftarrow \delta\phi + (1-\delta)\phi' \end{aligned} \tag{3.2}$$

The framework for Twin Deep Deterministic Policy Gradient is described as shown in the figure 3.2

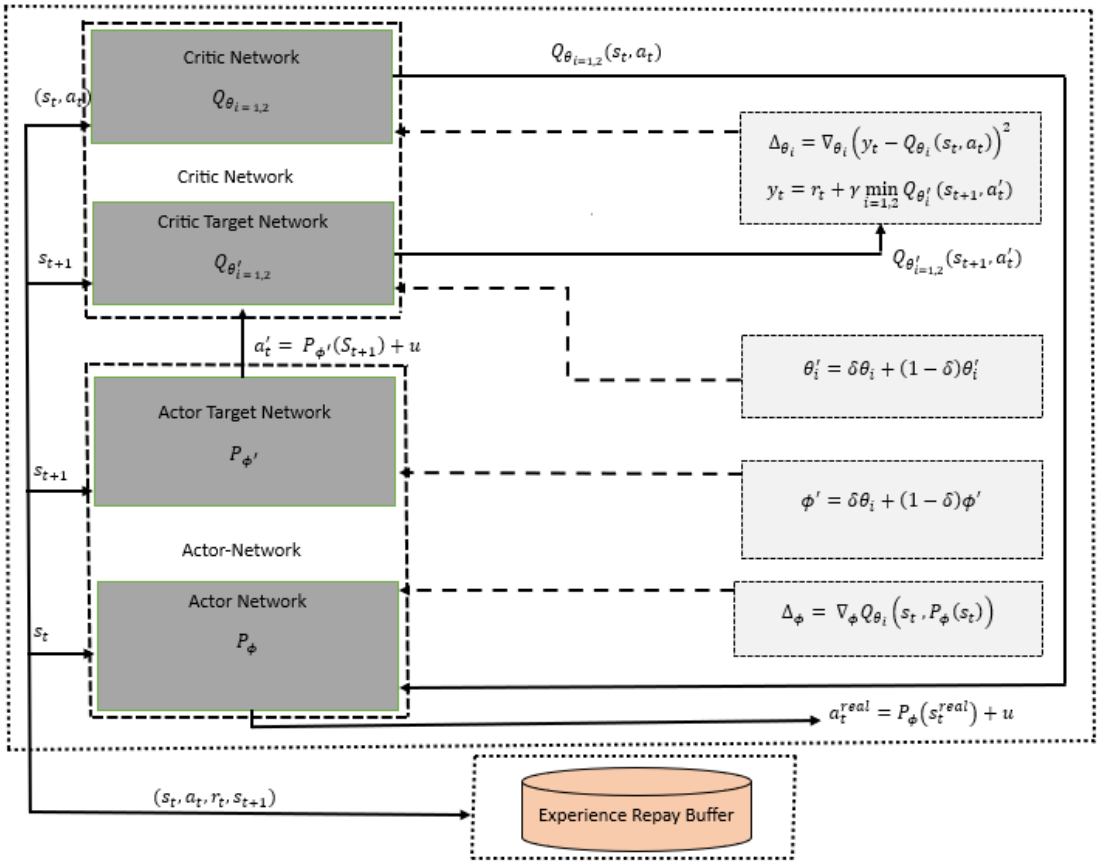


Figure 3.2: Framework for Twin Deep Deterministic Policy Gradient Algorithm: Source (Li et al., 2024)

The first step for TD3 in equation (3.1) modification is by introducing spatial influence and economic factors, which encompass temperature, rainfall, vector density, host density ,wealth, and education to have influence on the reward $r_t \in (-1, 1)$ given by

$$R_t = r_t \times I_{(s_t)}|_{(\text{Max,Min})} \tag{3.3}$$

where $I_{(s_t)}|_{(\text{Max,Min})} = \hat{\pi}_{it}$ embodies spatial and economic influence at state s_t. The fitted risk surface $\hat{\pi}_{it}$ was treated as a fixed, location and time-varying input that scales the reward r_t in equation (3.1) during RL training. This mechanism ensures that the environmental heterogeneity across the eight counties enters the reinforcement learning loop. $\hat{\pi}_{it}$ differs across locations because it is fitted separately for each location; this location-varying quantity captures the heterogeneity the policy is adapting to. The reward R_t is define as a modified reward the agent gain after taking action a_t at state s_t, $I_{(s_t)}|_{(\text{Max,Min})} = \hat{\pi}_{it}$ is given by

$$I_{(s_t)}|_{(\text{Max,Min})} = \hat{\pi}_{it}|_{(\text{Max,Min})} = \left(1 + e^{(-(\mathbf{S}_{it}\boldsymbol{\beta} + u_i + \gamma_t + \epsilon_{it}))}\right)^{-1} \tag{3.4}$$

$\pi_{it}|_{(\text{Max,Min})}, \pi_{it} \in [0, 1]$ is the malaria risk probability at spatial location m and time $t \in (1, \ldots, T)$, $S_i = (S_1, S_2, \ldots, S_n)', i \in (1, \ldots, n)$ are vector of predictors measured at location ℓ and time t, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)$ are regression coefficients, $\mathbf{u} = (u_1, u_2, \ldots, u_n)'$ is a spatial random effect for each spatial location m_i (random effect is added to introduce spatial dependence through dependence structure of $\mathbf{u}$), γ_t temporal fixed effects for time t (to capture year-specific baseline risks) $\gamma_t \sim N(0, \sigma_\gamma^2)$, and $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ is the spatiotemporal unstructured random effect (model uncorrelated noise) in the model.

Equation (3.1) presented as

$$\begin{aligned} \Delta\theta_i &= \nabla\theta_i (y_t - Q_{\theta_i}(s_t, a_t))^2 \\ y_t &= R_t + \gamma \min_{i=1,2} Q_{\theta_i'}(s_{t+1}, a_t') \\ a_t' &= P_{\phi'}(s_t) + u \end{aligned} \tag{3.5}$$

R_t correlates with disease occurrence, actions that prevent host-vector interaction get a higher reward.

The second step of model modification is on determination of y_t where the study proposes concurrent learning of five Q-values $Q_{\theta_1'}, Q_{\theta_2'}, \ldots, Q_{\theta_5'}$ by adopting Min-Max ensemble estimation method to reduce the overestimation and underestimation bias conservatively to minimize bias during learning Wang et al. (2021);

Chen et al. (2021). The Min-Max ensemble estimation is described as

$$\hat{Q}(s_{t+1}, a_t^{'}) = \max\left(\min\left(Q_{\theta_i'}(s_{t+1}, a_t^{'}) \mid i = 2, \ldots, N-1\right), Q_{\theta_N'}(s_{t+1}, a_t^{'})\right), \\ Q_{\theta_1'}(s_{t+1}, a_t^{'}) \tag{3.6}$$

Which can be generally represented as

$$\hat{Q}(s_{t+1}, a_t^{'}) = \max_{a_t' \in A} \min_{i \in N} Q_{\theta_i'}(s_{t+1}, a_t') \tag{3.7}$$

$Q_{\theta_1^{\mathrm{j}}}(s_{t+1}, a_t^{'})$ ensure that estimation is not overly conservative. Equation (3.7) is simplifies as

$$\hat{Q}(s_{t+1}, a_t') = \max(\min(Q_2, Q_3, Q_4),\ Q_5,\ Q_1) \tag{3.8}$$

where $Q_i = Q_{\theta_i'}(x)\ i = 1, \ldots, 5$ and $x = (s_{t+1}, a_t')$, where equation (3.8) retains the benefits of conservative estimation while effectively mitigating both overestimation and underestimation biases while maintaining stability and performance. In the last step, the actor network $a_t^{'} = P_{\phi'}(s_t) + u$ in eqn (3.5) for human learn learns updated deterministic policy

$$a_t^{'} = P_{\phi}(s_t, \ell) + u \tag{3.9}$$

which map the current interaction state s_t and location indicator $\ell \in \{1, 2, \ldots, 8\}$ to an intervention rate.

$$a_t^{'} = (a_{\text{spraying}},\ a_{\text{treated nets}}), \qquad a_{\text{spraying}}, a_{\text{treated nets}} \in (0, 1) \tag{3.10}$$

Conditioning the policy on ℓ allows a single trained framework to produce different, location-appropriate intervention rates for counties with varying environmental and socio-economic covariate.

The model equation is therefore derived by combining equation (3.5), equation (3.8) and equation (3.9).

$$\begin{aligned} \Delta\theta_i &= \nabla\theta_i(y_t - Q_{\theta_i}(s_t, a_t))^2 \qquad i = 1, \ldots, 5 \\ y_t &= R_t + \gamma \max_{a' \in A} \min_{i \in N} Q_{\theta_i'}(s_{t+1}, a_t') \\ a_t^{'} &= P_\phi(s_t, \ell) + u \end{aligned} \tag{3.11}$$

The soft updates on actor network and target networks is given by

$$\begin{aligned} \Delta\phi &= \nabla_\phi Q_{\theta_i}\left(s_t, P_\phi\left(s_t, \ell\right)\right) \quad i = 1, \ldots, 5 \\ \theta_i' &\leftarrow \delta\theta_i + (1-\delta)\theta_i' \\ \phi' &\leftarrow \delta\phi + (1-\delta)\phi' \end{aligned} \tag{3.12}$$

Where δ is a small positive soft update attenuation coefficient ($\delta \leq 1$), the update ensures iterative target network updating, utilizing linear combination of network parameters ϕ, θ_i and target network parameters ϕ', θ_i'. The soft update ensures gradual and smooth changes of target network parameters, and the target values also change smoothly. Smaller value of soft update δ coefficient of attenuation leads more stable algorithm and slower variation target network, with more sluggish algorithm convergence rate.

3.3 Estimation of Model Parameters

Estimating parameters for Bayesian model entails updating prior belief of model parameter with the observed data to derive posterior distribution by multiplying prior density of the parameter by with the likelihood function. The model likelihood function is derived from the observed data to quantify how probable the data is given the parameters (Florens et al., 2019). Samples for parameter estimates are drawn from the posterior distribution using the computation method of non-reversible metropolis hastings (Bierkens, 2016), the method lead to faster convergence, more effective exploration compared to the traditional reversible MCMC method, quantification of uncertainty and making probabilistic inferences (Bierkens, 2016).

The likelihood equation for $I_{(s_t)}$ described in equation (3.4), and used in deriving the reward $R_t = r_t \times I_{(s_t)}$ in model equation (3.8) is given by

$$\begin{aligned}
\mathcal{L}(y_{it}, S_{it}|\boldsymbol{\beta}) &= \prod_{i=1}^{n} p(y_{it}, S_{it}|\boldsymbol{\beta}) \\
&= \prod_{i=1}^{n} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-y_{it}} \left(1 - \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-1}\right)^{(1-y_{it})} \\
&= \prod_{i=1}^{n} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-y_{it}} \left(e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-1}\right)^{(1-y_{it})} \\
&= \prod_{i=1}^{n} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_i)}\right)^{-1} \left(e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{1-y_{it}} \\
&= \left(\left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-1} \left(e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{1-y_{it}}\right)^{n}
\end{aligned} \tag{3.13}$$

where $y_{it} \in [0, 1]$ is the observed outcome for i^{th} observation at time t. The prior distribution for parameters β_j for $j = 1 \ldots, k$ is assumed to be normal prior distribution given by

$$\beta_j \sim N(\mu_i, \sigma_i^2 I) \tag{3.14}$$

where

$$p(\beta_j) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\left(\frac{-(\beta_j-\mu_i)^2}{2\sigma_i^2}\right)} \tag{3.15}$$

The distribution of spatial random effect $\boldsymbol{u}$ is represented as

$$u_i|u_{-i} \sim N(\bar{\mu_{\eta_i}}, \frac{\sigma_u^2}{n_{\eta_i}}) \tag{3.16}$$

Where $\bar{\mu_{\eta_i}} = \frac{1}{n_{\eta_i}} \sum_{j \in \eta_i} u_j$ is the mean of u_j (random effect of the neighboring region j) over the neighboring region $j \in \eta_i$, n_{η_i} is the # of neighbors of region i and σ_u^2 is variance of spatial random effect.

The hyper-parameters in equation (3.14) have non-informative prior $\mu_i \sim N(0, \tau_i^2)$ where τ_i^2 is large enough to cover wide range of possible values of μ_i, the prior provide regularization without being too restrictive therefore allowing data to influence the posterior distribution (Grzenda, 2016). σ_i^2 has inverse gamma prior

$\sigma_i^2 \sim Inverse - Gamma(\alpha_i, \lambda_i)$ where α_i and λ_i are small values shape and scale parameters chosen to be non-informative. The prior for σ_i^2 is conjugate prior for normal distribution variance, simplifying the computation of the posterior distribution (Grzenda, 2016).

The prior density function for σ_i^2 is given by

$$p(\sigma_i^2) = \frac{\lambda_i^{\alpha_i}}{\Gamma(\alpha_i)}(\sigma_i^2)^{-\alpha_i - 1} e^{\left(\frac{-\lambda_i}{\sigma_i^2}\right)} \tag{3.17}$$

To avoid fixing the values of σ_i^2 on estimating for β_j, the study randomly estimate σ_i^2's values by sampling from it's posterior distribution.

$$p(\sigma_i^2|\beta_j) \propto p(\beta_j) \cdot p(\sigma_i^2) \tag{3.18}$$

where

$$\begin{aligned} p(\sigma_i^2|\beta_j) &\propto \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\left(\frac{-(\boldsymbol{\beta_j} - \mu_i)^2}{2\sigma_i^2}\right)} \cdot \frac{\lambda_i^{\alpha_i}}{\Gamma(\alpha_i)}(\sigma_i^2)^{-\alpha_i - 1} e^{\left(\frac{-\lambda_i}{\sigma_i^2}\right)} \\ &\propto (\sigma_i^2)^{-\alpha_i - \frac{3}{2}} exp\left(\frac{-1}{\sigma_i^2}\left(\lambda_i + \frac{(\beta_j - \mu_i)^2}{2}\right)\right) \end{aligned} \tag{3.19}$$

Equation (3.18) is a kernel of an Inverse-gamma distribution, therefore

$$\sigma_i^2|\beta_j \sim Inverse - Gamma\left(\alpha_i + \frac{1}{2}, \lambda_i + \frac{(\beta_j - \mu_i)^2}{2}\right) \tag{3.20}$$

Given the derived posterior distribution (3.20) the samples are drawn for parameter estimates σ_i^2's using the following steps

3.3.1 Algorithm for Estimating σ_i^2 Parameters

1. **Initialization:** Choose initial values for parameter $\alpha_i, \lambda_i, \sigma_i^2$ and β_j and auxillary variable ϕ which takes the values in $[-1, 1]$, this variable determines the direction of the proposal in parameter space

2. Define proposal mechanism $\sigma_i^{2'} = \sigma_i^2 + \phi\delta$, δ is a stepwise parameter adjusted

according to the acceptance rate

3. Compute acceptance probability for new parameter $\sigma_i^{2'}$

$$\alpha(\sigma_i^{2'}|\sigma_i^2) = Min\left(1, \frac{p(\sigma_i^{2'}|\beta_j)}{p(\sigma_i^2|\beta_j)}\right)$$

Draw a uniform random variable $u \in [0,1]$ if $u < \alpha$ accept the new parameters and set $\sigma_i^{2(i+1)} = \sigma_i^{2'}$ otherwise $\sigma_i^{2(i)} = \sigma_i^{2(i-1)}$

4. **Update auxiliary variable:** After updating parameters, potentially flip the auxiliary variable ϕ with some probability q which can be fixed or depend on the state of the chain

$$\phi^{i+1} = \begin{cases} -\phi^{(i)} & \text{with probability} \quad q \\ \phi^{(i)} & \text{with probability} \quad 1-q \end{cases}$$

5. Repeat 2, 3, and 4 for some iterations

The estimated values σ_i^2's are therefore used in estimation of β_j's. The posterior density of β_j is derived as

$$p(\beta_j|y_{it}, S_{it}) \propto \mathcal{L}(y_{it}, S_{it}|\beta_j) \cdot p(\beta_j) \tag{3.21}$$

and estimating $p(\beta_j|y_{it}, S_{it})$ as

$$
\begin{aligned}
p(\beta_j|y_{it}, S_{it}) \propto & \prod_{i=1}^{N} \frac{\left(e^{-(S_{it}\beta_j+u_i+\gamma_t+\epsilon_{it})}\right)^{1-y_{it}}}{1+e^{-(S_{it}\beta_j+u_i+\gamma_t+\epsilon_{it})}} \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\left(\frac{-(\beta_j-\mu_i)^2}{2\sigma_i^2}\right)} \\
\propto & \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_i} \frac{\left(e^{-(S_{it}\beta_j+u_i+\gamma_t+\epsilon_{it})}\right)^{1-y_{it}}}{1+e^{-(S_{it}\beta_j+u_i+\gamma_t+\epsilon_{it})}} \cdot e^{\left(\frac{-(\beta_j-\mu_i)^2}{2\sigma_i^2}\right)} \\
= & 2^{-\frac{n}{2}} \pi^{-\frac{n}{2}} \left(\frac{\mathrm{e}^{\frac{\left(-2S_{it}\beta_j-2u_i-2\gamma_t-2\varepsilon_{it}\right)\sigma_i^2-\left(-\beta_j+\mu_i\right)^2}{2\sigma_i^2}} \left(\mathrm{e}^{-S_{it}\beta_j-u_i-\gamma_t-\varepsilon_{it}}\right)^{-y_{it}}}{\sigma_i\left(1+\mathrm{e}^{-S_{it}\beta_j-u_i-\gamma_t-\varepsilon_{it}}\right)} \right)^n \\
= & \left(\sqrt{2\pi}\right)^{-n} \left(\frac{\mathrm{e}^{\frac{\left(-2S_{it}\beta_j-2u_i-2\gamma_t-2\varepsilon_{it}\right)\sigma_i^2-\left(-\beta_j+\mu_i\right)^2}{2\sigma_i^2}}}{\left(\mathrm{e}^{-S_{it}\beta_j-u_i-\gamma_t-\varepsilon_{it}}\right)^{y_{it}}} \left(\sigma_i\left(1+\mathrm{e}^{-S_{it}\beta_j-u_i-\gamma_t-\varepsilon_{it}}\right)\right)^{-1} \right)^n
\end{aligned}
\tag{3.22}
$$

Posterior equation (3.22) was used to draw samples for parameter estimates of β_j in the model.

3.3.2 Algorithm for Estimating β_j Parameters

1. **Initialization:** Choose initial values for parameter β_j and auxillary variable ϕ which takes the values in $[-1, 1]$, this variable determines the direction of the proposal in parameter space. The non-reversibility is due to following modified detailed-balance (skew-detailed-balance) condition

$$
\pi(\theta, \phi)\, p((\theta, \phi) \to (\theta', \phi')) = \pi(\theta', -\phi')\, p((\theta', -\phi') \to (\theta, -\phi)) \tag{3.23}
$$

   where $\pi(\theta, \phi),\ \beta_j \in \theta,\ j = 1, \ldots, k$ is the joint stationary distribution of the parameter and the auxiliary direction variable

2. Define proposal mechanism $\beta_j^{'} = \beta_j + \phi\delta$, δ is a step-wise parameter that can be adjusted according to the acceptance rate

3. Compute acceptance probability for new parameter $\beta_j^{'}$

$$
\alpha(\beta_j^{'}|\beta_j) = Min\left(1, \frac{p(\beta_j^{'}|y, S_{it}}{p(\beta_j|y, S_{it})}\right)
$$

Draw a uniform random variable $u \in [0,1]$ if $u < \alpha$ accept the new parameters and set $\beta_j^{(j+1)} = \beta_j'$ otherwise $\beta_j^{(j)} = \beta_j^{(j-1)}$

4. **Update auxiliary variable:** After updating parameters, potentially flip the auxiliary variable ϕ with some probability q which can be fixed or depend on the state of the chain

$$\phi^{j+1} = \begin{cases} -\phi^{(j)} & \text{with probability} \quad q \\ \phi^{(j)} & \text{with probability} \quad 1-q \end{cases}$$

5. Repeat 2, 3, and 4 for some iterations

The parameters convergence diagnostic checks were performed using trace plots.

3.4 Convergence of the Model Parameters

The convergence of parameters was supported by the theorems; Schwartz's theorem for posterior consistency and Bernstein-von Mises Theorem

3.4.1 Schwartz's theorem for posterior consistency

The theorem proposed by Schwartz (1965) states that as the number of data observations increase, the posterior distribution tend to concentrates on neighborhoods of the true data generating distribution which implies that the posterior mean estimates and the true posterior values have very negligible difference providing more reliable parameter estimates and their asymptotic measures.

Theorem 3.4.1 (Schwartz's Posterior Consistency (1965))**.**

Consider a Bayesian model with iid observations $X_1, X_2, \ldots \sim F_0$, where F_0 has density f_0 with respect to a σ-finite measure μ. Let F denote the model class of densities and π be a prior distribution on F. The posterior distribution given

$X_1, \ldots, X_n$ is defined as

$$\pi(S \mid X_1, \ldots, X_n) = \frac{\int_S \prod_{i=1}^n f(X_i)\, d\pi(f)}{\int_F \prod_{i=1}^n f(X_i)\, d\pi(f)}, \quad \text{for measurable } S \subseteq F \tag{3.24}$$

Assumes that the conditions 1 and 2 hold

1. **KL support condition** states that the prior π assign positive mass to all Kullback-Leibler (KL) neighborhoods of f_0 where $\pi\left(\{f \in F : K(f_0, f) < \epsilon\}\right) > 0, \quad \forall \epsilon > 0$ where $K(f_0, f) = \int f_0 \log \frac{f_0}{f}\, d\mu$ is the KL divergence.

2. **Existence of uniformly consistent tests** states that for every $\epsilon > 0$, there exists a sequence of tests $\phi_n : X^n \to [0, 1]$ such that

$$\begin{aligned} &\lim_{n\to\infty} F_0^{(n)} \phi_n = 0 \quad \text{(Type I error)}, \\ &\sup_{f \in U_\epsilon^c} F^{(n)}(1 - \phi_n) \leq e^{-Cn} \quad \text{(Type II error)}, \end{aligned} \tag{3.25}$$

   where $U_\epsilon = \{f \in F : d_H(f_0, f) < \epsilon\}$ is a Hellinger neighborhood of f_0, $d_H(f_0, f) = \sqrt{\int (\sqrt{f_0} - \sqrt{f})^2 d\mu}$, and $C > 0$ is a constant.

Therefore, the posterior is consistent at f_0 for every $\epsilon > 0$,

$$\pi\left(U_\epsilon \mid X_1, \ldots, X_n\right) \xrightarrow[n\to\infty]{F_0^\infty (\text{a.s})} 1 \tag{3.26}$$

The theorem applies to the convergence of posterior mean estimates to the true parameter values, where let θ_0 be the true posterior estimate and $\hat{\theta}_n$ be the mean posterior estimate, the theorem states that if the posterior is consistent then for any metric d which is equivalent to euclidean L^2, then

$$\forall_{\epsilon>0}, \pi(d(\hat{\theta}_n, \theta_0) > \epsilon | x_{1:n}) \xrightarrow[n\to\infty]{F_0^{(\infty)} (\text{a.s})} 1 \tag{3.27}$$

This implies that the posterior distribution concentrate around θ_0 and posterior

estimates $\hat{\theta}_n$ converge almost surely (a.s) to θ_0, that is

$$\hat{\theta}_n \xrightarrow[n\to\infty]{F_0^\infty(\text{a.s})} \theta_0 \tag{3.28}$$

3.4.2 Bernstein-von Mises Theorem

For a posterior distribution $\pi_n(\theta|Y_{1:n})$ for parameters $\theta \in \Theta \subseteq \mathbb{R}^d$ with true value θ_0, the Bernstein-von Mises Theorem states that for $n \to \infty$

$$\text{TV}\left(\pi_n(\theta|Y_{1:n}), \mathcal{N}\left(\hat{\theta}_n, (nF^*)^{-1}\right)\right) \leq \delta_3^* \sqrt{\frac{d^2}{n}} + \varepsilon_n \tag{3.29}$$

where $\hat{\theta}_n$ is the MLE, $F^* = \nabla^2 \ell^*(\theta_0)$ is the Fisher information matrix, ε_n is an exponentially small term, TV is the total variation distance, δ_3^* is model dependent coefficient related to third derivative of population log likelihood, and d is the dimension of the parameter space (Katsevich, 2023). The theorem support the explanation of the RMSE convergence, credible interval coverage, and Bias properties of the posterior mean estimates.

3.4.2.1 RMSE Convergence

For posterior mean $\hat{\theta}_n = \mathbb{E}[\theta|Y_{1:n}]$

$$\mathbb{E}_{\theta_0}\left[\|\hat{\theta}_n - \theta_0\|^2\right]^{1/2} \lesssim \sqrt{\frac{\text{tr}(F^{*-1})}{n}} \lesssim \sqrt{\frac{d}{n}} \tag{3.30}$$

holds under the improved dimension condition $n \gg d^2$ (Katsevich, 2023).

3.4.2.2 Credible Interval Coverage

For $1-\alpha$ Bayesian credible sets $B_{1-\alpha}$

$$|\mathbb{P}_{\theta_0}(\theta_0 \in B_{1-\alpha}) - (1-\alpha)| \leq C\delta_3^* \sqrt{\frac{d^2}{n}} \tag{3.31}$$

with $C > 0$ absolute, justifying frequentist calibration of Bayesian uncertainty quantification.

3.4.2.3 Bias Properties

The posterior mean satisfies

$$\|\mathbb{E}[\theta|Y_{1:n}] - \theta_0\| \leq \underbrace{\|\hat{\theta}_n - \theta_0\|}_{\text{MLE error}} + \underbrace{C\frac{\delta_0^*}{\sqrt{n}}}_{\text{Prior bias}} + O_P(1) \tag{3.32}$$

where δ_0^* quantifies prior concentration around θ_0.

3.4.3 Gelman-Rubin Statistic ($\hat{G}$)

The gelman $\hat{R}$ is computed across M parallel chains to evaluate convergence of the estimated parameters. The statistics is given by

$$\hat{G} = \sqrt{\frac{\hat{V}}{W}} \tag{3.33}$$

where $\hat{V}$ is the pooled between/within-chain variance estimate and W is the within-chain variance. The statistics $\hat{G} \to 1$ indicates convergence of parameter.

3.5 Model Operation

3.5.1 Computational Cost and Scalability

The model trained a five-critic ensemble (3.8) evaluated across eight locations, where a single policy and critic ensemble was trained with the location indicator ℓ included as part of the state in equation (3.9), that is, the parameter sharing across locations with location-conditioning. This helped in keeping the network cost fixed at five critics and one actor regardless of the number of locations, and was the design consistent with the objective of a single model producing comparable, location-adaptive prescriptive adoption rates.

3.5.2 Identifiability

The spatial random effect u_i and the location conditioning ℓ in the RL state absorb the location-specific signal. Therefore, identifiability in the model was preserved because u_i is estimated once, offline, and enters only as a fixed multiplier via $\hat{\pi}_{it}$ in the reward equation (3.3). RL policy therefore only learns the intervention rate to the fixed signal and does not re-estimate or absorb u_i itself.

3.5.3 High-Dimensional Inference

With the eight locations and two waves, the 16 (eight locations $\times$ two waves) location-wave cells were smaller relative to the typical hierarchical model requirements. Partial pooling (hierarchical shrinkage) was used so that β_j and u_i estimates for counties with fewer observations borrow strength from the pooled distribution across all eight counties via the shared prior in equation (3.14).

3.6 Implementation of the Model

This subsection describes the implementation of the model equation (3.11) in modeling the spread of disease in the environment. To implement the model, dynamic environment was created where a human agent try to avoid any interaction with a vector agent. The vector move randomly while interacting with the human agent. Rewards were based on successfully avoiding interaction between human and vector agents.

3.6.1 State Space

The human-vector distance was fixed as d_t randomly sampled at time t. A boolean variable "Interaction" at time t was defined to describe the state variable s_t, where interaction is True when host-vector interaction happens and False otherwise, as described in the equation.

$$\text{Interaction}_t = \begin{cases} \text{True}, & d_t = 0 \\ \text{False}, & d_t > 0 \end{cases} \tag{3.34}$$

When $d_t = 0$, this implies that there is a human-vector interaction, and therefore interaction occurs, and the vector transmits disease parasites to humans. From the model assumption, the human is recorded as infected if $d_t > 0$, there is no human-vector interaction; no infection is recorded. The distance d_t is given

$$d_t = \sqrt{(x_h - x_v)^2 + (y_h - y_v)^2} \tag{3.35}$$

where (x_h, y_h) is the point coordinate of human and (x_v, y_v) is the point coordinate of vector in the environment.

The state is therefore defined as

$$s_t = (\text{Interaction}_t,\ \ell) \tag{3.36}$$

where Interaction_t is the Boolean variable of equation (3.34) and $\ell \in \{1, \ldots, 8\}$ identifies the county. ℓ in the state is required for the objective of learning intervention rates that adapt across the eight environmentally heterogeneous locations for the policy to distinguish one county's environment from another.

3.6.2 Action Space

The actions encompass spraying rate and rate of using treated nets. Human and vector agents moves randomly in the environment in (x, y) direction where we set the maximum linear speed of the human agent to be 0.5 m/sec and maximum linear speed of vector to be 0.2 m/sec. The vector in environment moves faster covering more ground therefore having a higher chance of interacting with human agent. Human agent is expected to take strategic actions (spraying, use treated nets) to minimize interaction with the vector. The actions are selected to maximize expected reward, and the optimal action is selected during the learning

process.

The human agent actions representation at each time step is a vector of two continuous intervention rates

$$\text{action} = \{a_{\text{spraying}}, a_{\text{treated nets}}\} \tag{3.37}$$

where $a_{\text{spraying}} \in (0,1), a_{\text{treated nets}} \in (0,1)$. The agent can select values for each of the three actions or get one optimal action that maximize the expected return (a_{spraying} and $a_{\text{treated nets}}$ represent the proportion of the location's population accessed by the intervention (a coverage rate), which is consistent with how spraying and net-use coverage are recorded in the KDHS data, making the learned rates prescriptive, comparable to real intervention coverage statistics reported for each county).

3.6.3 Reward Function

In the dynamic environment of disease spread, an appropriate reward function enables human agents to reach the target (avoiding interaction for 30 seconds). Human agents pursue expected higher reward during training to find an optimal action to avoid interaction (being infected with the disease). Therefore, in this study the reward function design consists of two representations corresponding to the two states; interaction with a vector and reaching target (avoiding interaction with any vector for 30 seconds). The reward values are shown in equation.

$$\text{reward } (R_t) = \begin{cases} r_t \times I_{(s_t)}|_{(\text{Max})}, & d_t = 0 \\ r_t \times I_{(s_t)}|_{(\text{Min})}, & d_t > 0 \end{cases} \tag{3.38}$$

d_t is the human-vector distance. When $d_0 = 0$, human-vector agents interaction occurs and gets minimum reward value given by $r_t \times I_{(s_t)}|_{(\text{Max})}$, and when $d_t > 0$, there is no human-vector interaction, and human agent gets maximum reward value given by $r_t \times I_{(s_t)}|_{(\text{Min})}$. Human agent aim is to maximize the number of

instances where $d_t > 0$ (no interaction with vector).

The number of interactions were be recorded. If the human agent avoids interaction for 30 seconds, the human agent receives a maximum reward, and the instant is recorded as "no case. " When the agent interacts with the vector agent, the human agent receives a minimum reward, and the instant is recorded as "a case.".

3.7 Evaluation of Model Performance

Simulation study and malaria data were used to evaluate the model's performance.

3.7.1 Simulation Study

In this sub-section, the simulation study is provided to evaluate the model performance. The models in (3.4) and (3.11) were used in simulation. The algorithm that was used is as follows

Step 1: Simulate data variables for each of the eight counties $\ell = 1, \ldots, 8$

1) Malaria cases$_\ell \sim$ np.random.binomial$(1, p = 0.3, n = 11587)$

2) Use of mosquito nets$_\ell \sim$ np.random.binomial$(1, p = 0.5, n = 11587)$

3) Use of mosquito repellant spray$_\ell \sim$ np.random.binomial$(1, p = 0.2, n = 11587)$

4) Presence of mosquito species falciparum (Pf)$_\ell \sim$ np.random.binomial$(1, p = 0.2, n = 11587)$

5) Resident$_\ell \sim$ np.random.binomial$(1, p = 0.94, n = 11587)$

6) Age$_\ell \sim$ np.random.uniform$(min = 5, max = 110, n = 11587)$

7) Wealth index$_\ell \sim$ np.random.choice$(levels = [1, 2, 3, 4], size = 11587, replace = True)$

8) Education level$_\ell \sim$ np.random.choice($levels = [1, 2, 3, 4], size = 11587, replace = True$)

9) Temperature$_\ell \sim$ np.random.uniform($min = 18.2, max = 31.9, n = 11587$)

10) Rainfall$_\ell \sim$ np.random.uniform($min = 0.19, max = 10, n = 11587$)

To simulate the dynamic environment in step 2, the following addition assuptions were made

i) One human agent and one vector agent per location represent that location's average host-vector interaction pressure, not full population-level vector dynamics

ii) The simulated vector represents *Anopheles gambiae*/*Anopheles funestus*, the principal *Plasmodium falciparum* vectors in the study area

iii) Environmental covariates are held constant within a location's simulation window; heterogeneity is modelled between locations only

iv) No within-host disease progression (SEIR-type compartments) is modelled; infection is recorded as a binary case at the moment of interaction, consistent with the study's case/no-case definition

v) Intervention rates are agent-controlled and, once the location-conditioned policy is used, are location-adaptive rather than uniform

Step 2: Create dynamic environment

1) Create a continuous 2D space with boundaries $X = [0, 100], \quad Y = [0, 100]$

2) Create human and vector (infected with malaria parasites) agents

   a) Define state space - interaction status (`True/False`) between host and vector

   b) Define the action space: rate of spraying $\in (0, 1)$, and Rate of using mosquito nets $\in (0, 1)$.

c) Initialize parameters, agents, Q-values, and data recording structures

i) Transmission rate between the vector and human $\rho = 0.1$

ii) Initial infection status - human not infected and vector is infected.

iii) Search radius for an infected vector $r = 0.01$ meters

iv) Number of initially infected vectors = 1

v) Human agent initial position $(\text{human_x}, \text{human_y}) = (30.5, 45.2)$, $\text{human_speed} = 0.5$, actions: $\text{spray_action} = 0.2, \text{net_use_action} = 0.5$

vi) Vector agent initial position $(\text{vector_x}, \text{mosquito_y}) = (70.3, 25.1)$, $\text{mosquito_speed} = 0.2$

vii) Set interaction threshold = 0.01

viii) Soft update(τ) = 0.005, learning rate (α) = 0.001, discount factor(γ) = 0.99, exploration noise = 0.1, and batch size = 64.

d) Interaction and reward calculation

i) In each time step the human agent select an action based on its current state using the Actor network. The vector moves randomly and its infection status is updated.

ii) Interaction, the distance between human and mosquito $d_t = d(\text{Human position}, \text{Mosquito position})$ is calculated where interaction occurs when $d_t < 0.01$ metres (interaction threshold)

iii) Reward - Human get maximum expected reward when it avoid any interaction $(r_t \times I_{(s_t)}|_{(Min)})$ and minimum reward when interaction occurs $(r_t \times I_{(s_t)}|_{(Max)})$

e) Record actions and rewards

i) The action log records human actions (spray action, using mosquito nets action) to replay buffer.

ii) The reward log records the human rewards to replay buffer.

f) Repeat the simulation for multiple time steps (100000 time steps with evaluation after every 300 times-steps), updating positions, actions, rewards, and Q-values at each step to allow the host to learn optimal strategies for reducing malaria transmission.

3.7.2 Evaluation Metrics

This section presents the evaluation metrics: Root Mean Squared Error, success, and infection rates to evaluate the model's performance.

3.7.2.1 Root Mean Squared Error

Root mean squared error metric measure the accuracy of estimated model parameters, given by

$$\text{RMSE}(\hat{\beta}) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(\text{E}[\beta_j \mid data] - \beta_i^{\text{true}}\right)^2} \tag{3.39}$$

Where k is the number of parameters, the measure is in the parameter's unit and a smaller root mean squared error indicates better parameter accuracy.

3.7.2.2 Success Rate

This metric measures the proportion of human agents taking a significantly long time without interaction with the vector over n number of runs per test to minimize infection. It is estimated by

$$\text{Success Rate} = \frac{\text{Number of successful episodes without interaction}}{\text{Number of episodes}} \times 100\% \tag{3.40}$$

Where; Successful episode - is an episode where the host agent avoids interaction with the vector. Number of episodes - Total simulations/steps conducted during the learning process.

A higher success rate indicates agent's learned actions are more effective in reducing malaria transmission, while a lower success rate suggests that the agent's

policy is less effective. The success rate comprehensively evaluates the model's effectiveness in reducing malaria spread.

3.7.2.3 Infection Rate

The metric measures the rate of vector-human interaction in the dynamic environment. It is the ratio of successive interactions leading to infection between humans and the disease vector to the total number of executions over n test episodes. The measure is given by

$$\text{Infection Rate} = \frac{\text{Number of Interactions with the vector}}{\text{Number of episodes}} \times 100\% \tag{3.41}$$

A lower infection rate indicate better policy performance, while a decrease in infection rate over time due to the model's learning actions (rate of spraying, and rate of mosquito net use) demonstrates the effectiveness of the intervention. Infection rates during and after training were examined for model performance assessment across different policies.

Remark 3.1 (Prescriptive use of the Policy). Once the policy is trained, it is used to explore for reward maximization and queried prescriptively. Given location ℓ and a learned optimal rate a_t, the trained environment dynamics are used to determine how many host-vector interactions result. That is, the trained actor-environment pair is used as a conditional simulator, given by

$$\hat{C}_\ell(a) = f_{\text{sim}}\left(a_t \mid \ell, \hat{\pi}_{it}\right) \tag{3.42}$$

where $f_{\text{sim}}(\cdot)$ denotes rolling the trained environment forward under a fixed rate a_t at location ℓ and counting the resulting interaction events, and $\hat{\pi}_{it}$ is the fitted, location and time-varying Bayesian risk term from equation (3.4).

Remark 3.2 (Prediction Process). Predicted cases at location ℓ under rate a_t are defined as the count of time steps at which an interaction is recorded when the trained environment is simulated forward under the rate over an evaluation

horizon T_{eval}

$$\hat{C}_\ell(a_t) = \sum_{t=1}^{T_{\text{eval}}} I\left[d_t = 0 \mid \ell, a_t\right] \tag{3.43}$$

where $I[\cdot]$ is the indicator function and d_t is the human-vector distance defined in equation (3.35). This equation makes the framework a case-prediction tool rather than only an intervention optimizer. The model predicts malaria cases indirectly, as the number of interaction events occurring under a learned intervention rate, rather than directly regressing case counts on covariates.

3.7.2.4 Case Prediction Error

This metric evaluate prediction error of cases by comparing predicted and observed case counts, per county

$$\text{Case Prediction Error}_\ell = \left|\hat{C}_\ell(a_t) - C_\ell^{\text{observed}}\right| \tag{3.44}$$

where $\hat{C}_\ell(a_t)$ is the simulated interaction count at county ℓ under intervention rate a_t and C_ℓ^{observed} is the corresponding observed case count for county ℓ from the KDHS data, normalized as

$$\text{Case Prediction Error } (\%)_\ell = \frac{\left|\hat{C}_\ell(a_t) - C_\ell^{\text{observed}}\right|}{C_\ell^{\text{observed}}} \times 100\% \tag{3.45}$$

This metric validates the study objective of prescriptive intervention rate at a given location where the predicted case should be close to the observed county-level case count after human intervention.

3.7.2.5 Cross-Location Adaptivity

To check whether the policy has genuinely learned to adapt to heterogeneity across the eight counties instead of just converging to a single rate applied uniformly, the variance of the learned optimal rate across locations is calculated as

$$\mathrm{Var}_\ell\left(\pi^*(\cdot,\ell)\right) = \frac{1}{8}\sum_{\ell=1}^{8}\left(a_\ell^* - \bar{a}^*\right)^2 \tag{3.46}$$

where a_ℓ^* is the learned optimal intervention rate at location ℓ and $\bar{a}^*$ is its mean across the eight locations. A variance near zero indicates the policy has not learned location-adaptive behaviour, providing a direct checkable diagnostic of the central claim of the study.

3.8 Determination of Optimal Strategies

To determine optimal strategies for reducing malaria spread, where the host agent tries to minimize interactions with the disease vector. The state space is described by boolean interaction variable (TRUE/FALSE), defined by $s_t = (\text{Interaction}_t, \ell)$. The action space is the rate of spraying $\in [0, 1]$ and rate of net use $\in [0, 1]$. The algorithm optimizes the agent's policy by maximizing the expected return from the agent actions, adjusting based on the random movements of the vector and the host's strategic responses. The action combinations that maximize the reward (minimizing infection rates) was recorded throughout the learning process, and each instance of action taken was logged along with the corresponding reward. The recorded action combinations were then retrieved for further analysis, and the optimal set of actions were estimated by comparing action combinations across various scenarios, showing which intervention strategies effectively reduce malaria transmission.

The standard reinforcement learning quantities are defined, where each is explicitly conditioned on location ℓ to reflect that optimal for each of the eight heterogeneous locations, not a single rate applied uniformly across all locations.

$$\pi(a \mid s, \ell): \text{ policy-probability of taking action } a \text{ in state } s \text{ at location } \ell \quad (3.47)$$

$$V(s, \ell) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} \,\middle|\, s_t = s, \ell \right] \quad (3.48)$$

$$Q(s, a, \ell) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} \,\middle|\, s_t = s, a_t = a, \ell \right] \quad (3.49)$$

$$\pi^*(s, \ell) = \arg\max_a Q(s, a, \ell) \quad (3.50)$$

The objective maximized during training is the expected discounted return conditional on location.

$$J(\phi) = \mathbb{E} \left[\sum_{t=0}^{T} \gamma^t R_t \,\middle|\, \ell \right] \quad (3.51)$$

Optimization is performed via the actor network gradient ascent step of the model deterministic policy gradient update equation (3.11) with the target Q-value supplied by the five-critic min-max ensemble of equation (3.8).

3.9 Data Description

The data was derived from the Demographic and Health Survey (DHS) 2015 and 2020, `https://dhsprogram.com/data/dataset/Kenya_MIS_2015.cfm?flag=1`, `https://dhsprogram.com/data/dataset/Kenya_MIS_2020.cfm?flag=1`. The data is Malaria Indicator Survey data, which covers a wide range of variable data; however, for the scope of this study, some variables selected for analysis, including presence of malaria cases, use of mosquito nets, use of mosquito repellant spray, presence of mosquito species falciparum (Pf), use of ACT for malaria treatment, resident, wealth index, education level, county, malaria endemic zones, temperature, and rainfall. The study analyzed data from eight counties (Kisumu, Siaya, Homa Bay, Migori, Kakamega, Bungoma, Busia and Vihiga) that have high cases of malaria.

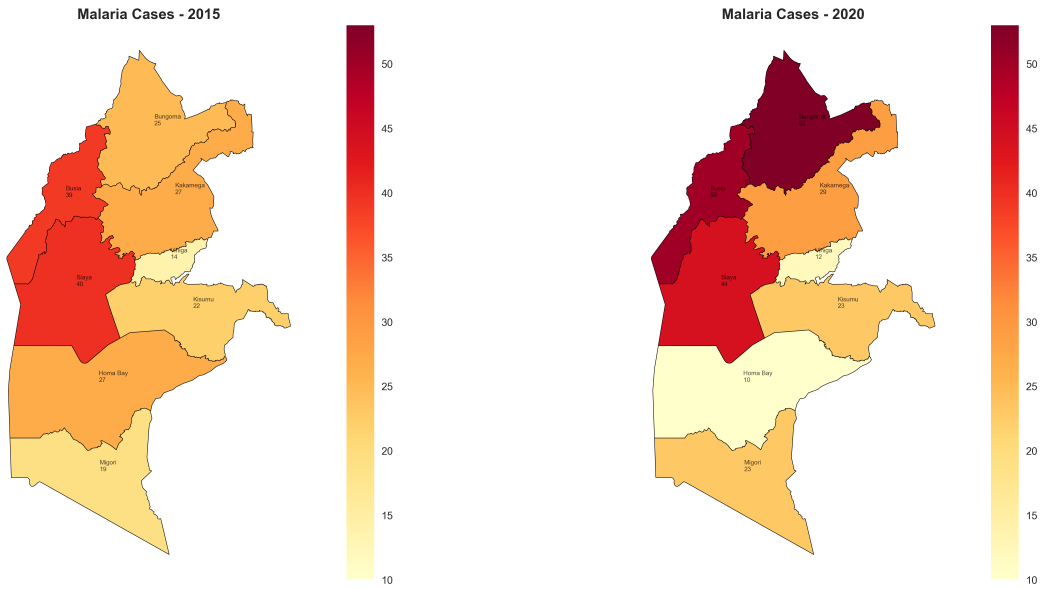


Figure 3.3: Spatial Distribution of Malaria Cases in Eight Counties

3.10 Statistical Software

Model implementation, data presentation, analysis, and spatial visualization was done using Python programming language McKinney (2012); Shreemathi et al. (2024).

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter presents the results of GIS-Enhanced Bayesian Reinforcement Learning model for predicting spread of vector-borne infectious disease (malaria).

4.2 GIS-Enhanced Bayesian Reinforcement Learning Model

The GIS-Enhanced Bayesian Reinforcement Learning Model is given by the equation

$$\begin{aligned} \Delta\theta_i &= \nabla\theta_i(y_t - Q_{\theta_i}(s_t, a_t))^2 \qquad i = 1, \ldots, 5 \\ y_t &= R_t + \gamma \max_{a' \in A} \min_{i \in N} Q_{\theta_i'}(s_{t+1}, a_t') \\ a_t^{'} &= P_\phi(s_t, \ell) + u \end{aligned} \tag{4.1}$$

where the error $u \sim clip(N(0, \sigma^2), -b, b)$, R_t is the agent immediate reward from taking action a_t in state s_t, $\gamma \in (0, 1)$ is discount factor which determines how future rewards are discounted relative to immediate rewards, θ_i's and ϕ are random parameters for updating target networks, current and next state s_t, s_{t+1}, current and next action a_t, $a_t^{'}$, truncated value b with boundary $-b$ and b (small constants greater than zero), P is a target policy where $P_{\phi'}$ is target actor network policy.

Action a_t is calculated at state s_t based on actor network P_ϕ, the objective action a' is computed through actor target network $P_{\phi'}$ at state s_{t+1}. $Q_{\theta_i}(s_{t+1}, a_t)$ and the target $Q_{\theta_1}(s_{t+1}, a'_t)$ Q-values determines action-quality in s_t, and actor network policy update. y_t is obtained through selection of minimum Q-value of ($Q_{\theta'_1}$ and $Q_{\theta'_2}$) estimated by critic target network and apply it to $\Delta\theta_i = \nabla\theta_i (y_t - Q_{\theta_i}(s_t, a_t))^2$. The algorithm updates θ_i based on $\Delta\theta_i$ of the critic network parameters, that is obtained through the gradient increase of the loss equation $(y_t - Q_{\theta_i}(s_t, a_t))^2$ for $i = 1, \ldots, 5$.

The information obtained from the update step, actions, current states, rewards, and next states are stored in tuple format in experience replay buffer. The data needed for updating actor and critic network parameters are then randomly sampled from experience replay buffer.

And the soft updates on actor network and target networks are given by

$$\begin{aligned} \Delta\phi &= \nabla_\phi Q_{\theta_i}(s_t, P_\phi(s_t)) \quad i = 1, \ldots, 5 \\ \theta'_i &\leftarrow \delta\theta_i + (1-\delta)\theta'_i \\ \phi' &\leftarrow \delta\phi + (1-\delta)\phi' \end{aligned} \tag{4.2}$$

Where δ is a small positive soft update attenuation coefficient ($\delta \leq 1$), the update ensures iterative target network updating, utilizing linear combination of network parameters ϕ, θ_i and target network parameters ϕ', θ'_i. The soft update ensures gradual and smooth changes of target network parameters, and the target values also change smoothly. Smaller value of soft update δ coefficient of attenuation leads more stable algorithm and slower variation target network, with more sluggish algorithm convergence rate.

The reward R_t in equation (4.1) is given by

$$R_t = r_t \times I_{(s_t)}|_{(\text{Max,Min})} \tag{4.3}$$

where $I_{(s_t)}|_{(\text{Max,Min})} = \hat{\pi}_{it}$ embodies spatial and economic influence at state s_t, The

fitted risk surface $\hat{\pi}_{it}$ was treated as a fixed, location and time-varying input that scales the reward r_t in equation (4.3) during RL training. This mechanism ensures that the environmental heterogeneity across the eight counties enters the reinforcement learning loop. $\hat{\pi}_{it}$ differs across locations because it is fitted separately for each location; this location-varying quantity captures the heterogeneity the policy is adapting to. The reward R_t is define as a modified reward the agent gain after taking action a_t at state s_t, $I_{(s_t)}|_{(\text{Max,Min})} = \hat{\pi}_{it}$ is given by

$$I_{(s_t)}|_{(\text{Max,Min})} = \hat{\pi}_{it}|_{(\text{Max,Min})} = \left(1 + e^{(-(\mathbf{S}_{it}\boldsymbol{\beta} + u_i + \gamma_t + \epsilon_{it}))}\right)^{-1} \tag{4.4}$$

$\pi_{it}|_{(\text{Max,Min})}, \pi_{it} \in [0, 1]$ is the malaria risk probability at spatial location m and time $t \in (1, \ldots, T)$, $S_i = (S_1, S_2, \ldots, S_n)', i \in (1, \ldots, n)$ are vector of predictors measured at location ℓ and time t, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)$ are regression coefficients, $\mathbf{u} = (u_1, u_2, \ldots, u_n)'$ is a spatial random effect for each spatial location m_i (random effect is added to introduce spatial dependence through dependence structure of $\mathbf{u}$), γ_t temporal fixed effects for time t (to capture year-specific baseline risks) $\gamma_t \sim N(0, \sigma_\gamma^2)$, and $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ is the spatiotemporal unstructured random effect (model uncorrelated noise) in the model.

4.3 Estimation of Model Parameters

In this section, the estimators for the unknown parameters of the model were derived using Bayesian estimation. Let $Y_1, Y_2, \ldots, Y_m$ be a random sample, the likelihood equation $\mathcal{L} = \mathcal{L}(y_{it}, S_{it}|\boldsymbol{\beta})$ for $I_{(s_t)}$ in equation (4.4) is given by

$$
\begin{aligned}
\mathcal{L}(y_{it}, S_{it}|\boldsymbol{\beta}) &= \prod_{i=1}^{n} p(y_{it}, S_{it}|\boldsymbol{\beta}) \\
&= \prod_{i=1}^{n} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-y_{it}} \left(1 - \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-1}\right)^{(1-y_{it})} \\
&= \prod_{i=1}^{n} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-y_{it}} \left(e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-1}\right)^{(1-y_{it})} \\
&= \prod_{i=1}^{n} \left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{i})}\right)^{-1} \left(e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{1-y_{it}} \\
&= \left(\left(1 + e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{-1} \left(e^{-(S_{it}\boldsymbol{\beta}+u_i+\gamma_t+\epsilon_{it})}\right)^{1-y_{it}}\right)^{n}
\end{aligned}
\tag{4.5}
$$

where $y_{it} \in [0, 1]$ is the observed outcome for i^{th} observation at time t. The prior distribution for parameters β_j for $j = 1 \ldots, k$ is assumed to be normal prior distribution given by

$$
\beta_j \sim N(\mu_i, \sigma_i^2 I) \text{ where } p(\beta_j) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\left(\frac{-(\beta_j-\mu_i)^2}{2\sigma_i^2}\right)} \tag{4.6}
$$

The distribution of spatial random effect $\boldsymbol{u}$ is represented as

$$
u_i|u_{-i} \sim N(\bar{\mu_{\eta_i}}, \frac{\sigma_u^2}{n_{\eta_i}}) \tag{4.7}
$$

Where $\bar{\mu_{\eta_i}} = \frac{1}{n_{\eta_i}} \sum_{j \in \eta_i} u_j$ is the mean of u_j (random effect of the neighboring region j) over the neighboring region $j \in \eta_i$, n_{η_i} is the # of neighbors of region i and σ_u^2 is variance of spatial random effect.

The hyper-parameters in equation (4.6) have non-informative prior $\mu_i \sim N(0, \tau_i^2)$ where τ_i^2 is large enough to cover wide range of possible values of μ_i, the prior provide regularization without being too restrictive therefore allowing data to influence the posterior distribution (Grzenda, 2016). σ_i^2 has inverse gamma prior $\sigma_i^2 \sim Inverse - Gamma(\alpha_i, \lambda_i)$ where α_i and λ_i are small values shape and scale parameters chosen to be non-informative. The prior for σ_i^2 is conjugate prior for normal distribution variance, simplifying the computation of the posterior distribution (Grzenda, 2016).

The prior density function for σ_i^2 is given by

$$p(\sigma_i^2) = \frac{\lambda_i^{\alpha_i}}{\Gamma(\alpha_i)}(\sigma_i^2)^{-\alpha_i-1}e^{\left(\frac{-\lambda_i}{\sigma_i^2}\right)} \tag{4.8}$$

To avoid fixing the values of σ_i^2 on estimating for β_j, the study randomly estimate σ_i^2's values by sampling from it's posterior distribution.

$$p(\sigma_i^2|\beta_j) \propto p(\beta_j) \cdot p(\sigma_i^2) \tag{4.9}$$

where

$$\begin{aligned} p(\sigma_i^2|\beta_j) \propto & \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\left(\frac{-(\boldsymbol{\beta_j}-\mu_i)^2}{2\sigma_i^2}\right)} \cdot \frac{\lambda_i^{\alpha_i}}{\Gamma(\alpha_i)}(\sigma_i^2)^{-\alpha_i-1}e^{\left(\frac{-\lambda_i}{\sigma_i^2}\right)} \\ \propto & (\sigma_i^2)^{-\alpha_i-\frac{3}{2}} exp\left(\frac{-1}{\sigma_i^2}\left(\lambda_i + \frac{(\beta_j-\mu_i)^2}{2}\right)\right) \end{aligned} \tag{4.10}$$

Equation (4.9) is a kernel of an Inverse-gamma distribution, therefore

$$\sigma_i^2|\beta_j \sim Inverse - Gamma\left(\alpha_i + \frac{1}{2}, \lambda_i + \frac{(\beta_j-\mu_i)^2}{2}\right) \tag{4.11}$$

Given the derived posterior distribution (4.11) the samples are drawn for parameter estimates σ_i^2's using the following steps

4.3.1 Algorithm for Estimating σ_i^2 Parameters

1. **Initialization:** Choose initial values for parameter $\alpha_i, \lambda_i, \sigma_i^2$ and β_j and auxillary variable ϕ which takes the values in $[-1, 1]$, this variable determines the direction of the proposal in parameter space

2. Define proposal mechanism $\sigma_i^{2'} = \sigma_i^2 + \phi\delta$, δ is a stepwise parameter adjusted according to the acceptance rate

3. Compute acceptance probability for new parameter $\sigma_i^{2'}$

$$\alpha(\sigma_i^{2'}|\sigma_i^2) = Min\left(1, \frac{p(\sigma_i^{2'}|\beta_j)}{p(\sigma_i^2|\beta_j)}\right)$$

Draw a uniform random variable $u \in [0,1]$ if $u < \alpha$ accept the new parameters and set $\sigma_i^{2(i+1)} = \sigma_i^{2'}$ otherwise $\sigma_i^{2(i)} = \sigma_i^{2(i-1)}$

4. **Update auxiliary variable:** After updating parameters, potentially flip the auxiliary variable ϕ with some probability q which can be fixed or depend on the state of the chain

$$\phi^{i+1} = \begin{cases} -\phi^{(i)} & \text{with probability} \quad q \\ \phi^{(i)} & \text{with probability} \quad 1-q \end{cases}$$

5. Repeat 2, 3, and 4 for some iterations

The estimated values σ_i^2's are therefore used in estimation of β_j's. The posterior density of β_j is derived as

$$p(\beta_j|y_{it}, S_{it}) \propto \mathcal{L}(y_{it}, S_{it}|\beta_j) \cdot p(\beta_j) \tag{4.12}$$

and estimating $p(\beta_j|y_{it}, S_{it})$ as

$$\begin{aligned} p(\beta_j|y_{it}, S_{it}) \propto & \prod_{i=1}^{N} \frac{\left(e^{-(S_{it}\beta_j + u_i + \gamma_t + \epsilon_{it})}\right)^{1-y_{it}}}{1 + e^{-(S_{it}\beta_j + u_i + \gamma_t + \epsilon_{it})}} \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\left(\frac{-(\beta_j - \mu_i)^2}{2\sigma_i^2}\right)} \\ \propto & \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_i} \frac{\left(e^{-(S_{it}\beta_j + u_i + \gamma_t + \epsilon_{it})}\right)^{1-y_{it}}}{1 + e^{-(S_{it}\beta_j + u_i + \gamma_t + \epsilon_{it})}} \cdot e^{\left(\frac{-(\beta_j - \mu_i)^2}{2\sigma_i^2}\right)} \\ = & 2^{-\frac{n}{2}} \pi^{-\frac{n}{2}} \left(\frac{\mathrm{e}^{\frac{\left(-2S_{it}\beta_j - 2u_i - 2\gamma_t - 2\varepsilon_{it}\right)\sigma_i^2 - \left(-\beta_j + \mu_i\right)^2}{2\sigma_i^2}} \left(\mathrm{e}^{-S_{it}\beta_j - u_i - \gamma_t - \varepsilon_{it}}\right)^{-y_{it}}}{\sigma_i \left(1 + \mathrm{e}^{-S_{it}\beta_j - u_i - \gamma_t - \varepsilon_{it}}\right)} \right)^n \\ = & \left(\sqrt{2\pi}\right)^{-n} \left(\frac{\mathrm{e}^{\frac{\left(-2S_{it}\beta_j - 2u_i - 2\gamma_t - 2\varepsilon_{it}\right)\sigma_i^2 - \left(-\beta_j + \mu_i\right)^2}{2\sigma_i^2}}}{\left(\mathrm{e}^{-S_{it}\beta_j - u_i - \gamma_t - \varepsilon_{it}}\right)^{y_{it}}} \left(\sigma_i \left(1 + \mathrm{e}^{-S_{it}\beta_j - u_i - \gamma_t - \varepsilon_{it}}\right)\right)^{-1} \right)^n \end{aligned} \tag{4.13}$$

Posterior equation (4.13) was used to draw samples for parameter estimates of β_j in the model.

4.3.2 Algorithm for Estimating β_j Parameters

1. **Initialization:** Choose initial values for parameter β_j and auxillary variable ϕ which takes the values in $[-1, 1]$, this variable determines the direction of the proposal in parameter space. The non-reversibility is due to following modified detailed-balance (skew-detailed-balance) condition

$$\pi(\theta, \phi)\, p((\theta, \phi) \to (\theta', \phi')) = \pi(\theta', -\phi')\, p((\theta', -\phi') \to (\theta, -\phi)) \qquad (4.14)$$

   where $\pi(\theta, \phi)$, $\beta_j \in \theta$, $j = 1, \ldots, k$ is the joint stationary distribution of the parameter and the auxiliary direction variable

2. Define proposal mechanism $\beta_j^{'} = \beta_j + \phi\delta$, δ is a step-wise parameter that can be adjusted according to the acceptance rate

3. Compute acceptance probability for new parameter $\beta_j^{'}$

$$\alpha(\beta_j^{'}|\beta_j) = Min\left(1, \frac{p(\beta_j^{'}|y, S_{it}}{p(\beta_j|y, S_{it})}\right)$$

   Draw a uniform random variable $u \in [0, 1]$ if $u < \alpha$ accept the new parameters and set $\beta_j^{(j+1)} = \beta_j^{'}$ otherwise $\beta_j^{(j)} = \beta_j^{(j-1)}$

4. **Update auxiliary variable:** After updating parameters, potentially flip the auxiliary variable ϕ with some probability q which can be fixed or depend on the state of the chain

$$\phi^{j+1} = \begin{cases} -\phi^{(j)} & \text{with probability} \quad q \\ \phi^{(j)} & \text{with probability} \quad 1 - q \end{cases}$$

5. Repeat 2, 3, and 4 for some iterations

4.3.3 Monte Carlo Simulation

In this sub-section, the simulation study evaluate the root mean squared errors (RMSE) (3.39), credible intervals, bias, and convergence of the model parameter estimates. The posterior distribution given in equation (4.13) was used to sample parameter estimates using the algorithm.

1. A random sample of size 1000 was generated from the posterior distribution distribution using selected initial values for parameters $\alpha_i = 0.1, \lambda_i = 0.1, \sigma_i^2 = 0.001$ and $\beta_i = 0$ and auxillary variable ϕ which takes the values in $[-1, 1]$.

2. For the sample size, parameter estimate is obtained

3. Step 1 to 2 is repeated 10000 times (iterations) and parameter estimate and root mean squared errors recorded for every iteration.

The simulation study was conducted to estimate the model's parameters and evaluate their convergence. Parameter estimates, credible intervals, and root mean square errors are presented in Table 4.1. The rainfall parameter -0.0298 indicates a negative association with malaria risk, where a unit increase in rainfall decreases the log-odds of malaria risk by 0.0298. CI $[-0.1078, -0.0052]$ shows that there is a 95% chance that the true parameter is within the interval with 0.0120 bias; the root mean square error 0.0003 implies a high precision of the parameter estimate. The temperature parameter 0.0779 indicates a positive association with malaria risk, where a unit increase in temperature increases the log-odds of malaria risk by 0.0779. CI $[-0.0499, 0.6268]$ shows that there is a 95% chance that the true parameter is within the interval with 0.0028 bias; the root mean square error 0.0003 implies a high precision of the parameter estimate. The age parameter 0.2350 indicate positive association with malaria risk where a unit increase in age increases the log-odds of malaria by 0.2350, indicating that older ages are associated with significantly higher risk of contracting the disease, the

CI $[0.2119, 0.3313]$ shows that there is 95% chances that the true parameter is within the interval with 0.0067 bias; the root mean square error 0.0003 implies a high precision of the parameter estimate.

Primary level of education parameter -2.0294 indicate negative association with malaria risk, where a positive change in education (primary) level reduces the risk log-odds of malaria by 2.0294 compared to individual with no education, the CI $[-2.3300, -0.1384]$ shows that there is 95% chances that the true parameter is within the interval with 0.0113 bias; the root mean square error 0.0029 implies a high precision of the parameter estimate. Secondary level of education parameter -2.1181 indicate negative association with malaria risk, where a positive change in education (secondary) level reduces the risk log-odds of malaria by 2.1181 compared to individual with no education, the CI $[-2.6160, -0.1279]$ shows that there is 95% chances that the true parameter is within the interval with 0.0134 bias; the root mean square error 0.0022 implies a high precision of the parameter estimate. Tertitairy level of education parameter -2.0154 indicate negative association with malaria risk, where a positive change in education (tertiary) level reduces the risk log-odds of malaria by 2.0154 compared to individual with no education, the CI $[-2.1023, -0.7413]$ shows that there is 95% chances that the true parameter is within the interval with 0.0023 bias; the root mean square error 0.0032 implies a high precision of the parameter estimate.

Middle level of wealth parameter -1.0921 indicate negative association with malaria risk, where a positive change in wealth (middle) level reduces the risk log-odds of malaria by 1.0921 compared to individual in poor wealth bracket, the CI $[-1.2015, -0.3668]$ shows that there is 95% chances that the true parameter is within the interval with 0.0043 bias; the root mean square error 0.0026 implies a high precision of the parameter estimate. Richer level of wealth parameter -2.0013 indicate negative association with malaria risk, where a positive change in wealth (richer) level reduces the risk log-odds of malaria by 2.0013 compared to to individual in poor wealth bracket, the CI $[-3.0127, -1.0822]$ shows that there is 95% chances that the true parameter is within the interval with 0.0346

bias; the root mean square error 0.0023 implies a high precision of the parameter estimate. Richest level of wealth parameter -2.1117 indicate negative association with malaria risk, where a positive change in wealth (richest) level reduces the risk log-odds of malaria by 2.1117 compared to individual in poor wealth bracket level, the CI $[-2.7014, -0.5057]$ shows that there is 95% chances that the true parameter is within the interval with 0.0457 bias; the root mean square error 0.0026 implies a high precision of the parameter estimate. The parameter estimates for treated mosquito nets (-2.1267) and spray repellent (-1.8249) have a significantly lower log-odds of contracting malaria.

Table 4.1: Parameter Estimation and its Evaluation Measures

Variable	Estimate	Credible Interval	RMSE	Bias	Gelman-Rubin
Rainfall	-0.0298	[-0.1078,-0.0052]	0.0003	0.0120	1.0004
Temperature	0.0779	[0.0499,0.6268]	0.0003	0.0028	1.0003
Age	0.2350	[0.2119,0.3313]	0.0003	0.0067	1.0016
$\text{Education}_{\textbf{Primary}}$	-2.0294	[-2.3300,-0.1384]	0.0029	0.0113	1.0016
$\text{Education}_{\textbf{Secondary}}$	-2.1181	[-2.6160,-0.1279]	0.0022	0.0134	1.0002
$\text{Education}_{\textbf{Tertiary}}$	-2.0154	[-2.1023,-0.7413]	0.0032	0.0023	1.0032
$\text{Wealth Index}_{\textbf{Middle}}$	-1.0921	[-1.2015,-0.3668]	0.0026	0.0043	1.0004
$\text{Wealth Index}_{\textbf{Richer}}$	-2.0013	[-3.0127,-1.0822]	0.0023	0.0346	1.0007
$\text{Wealth Index}_{\textbf{Richest}}$	-2.1117	[-2.7014,-0.5057]	0.0026	0.0457	1.0005
Treated mosquito Net	-2.1267	[-3.1483,-2.0100]	0.0027	0.0042	1.0031
Spray repellent	-1.8249	[-2.3167,-1.0113]	0.0029	0.0234	1.0003

The trace plots in figure 4.1 provide strong evidence of convergence for the parameters associated with temperature, rainfall, age, wealth, education, use of treated net, and use of spray repellent. This is demonstrated by the trace plot, which shows clear mixing of the chains. The uniform spread and stable trajectories indicate that the chains have efficiently explored the posterior distributions and achieved stationarity. The overlapping traces suggest excellent chain mixing, thereby providing a reliable estimate of the parameter. Stability also enhances confidence in the accuracy of the posterior distribution and demonstrates that the MCMC process successfully captures the underlying relationships in the data. Convergence of the parameters is also demonstrated by Gelman-Rubin statistics being approximately 1 for all parameters (Table (4.1)).

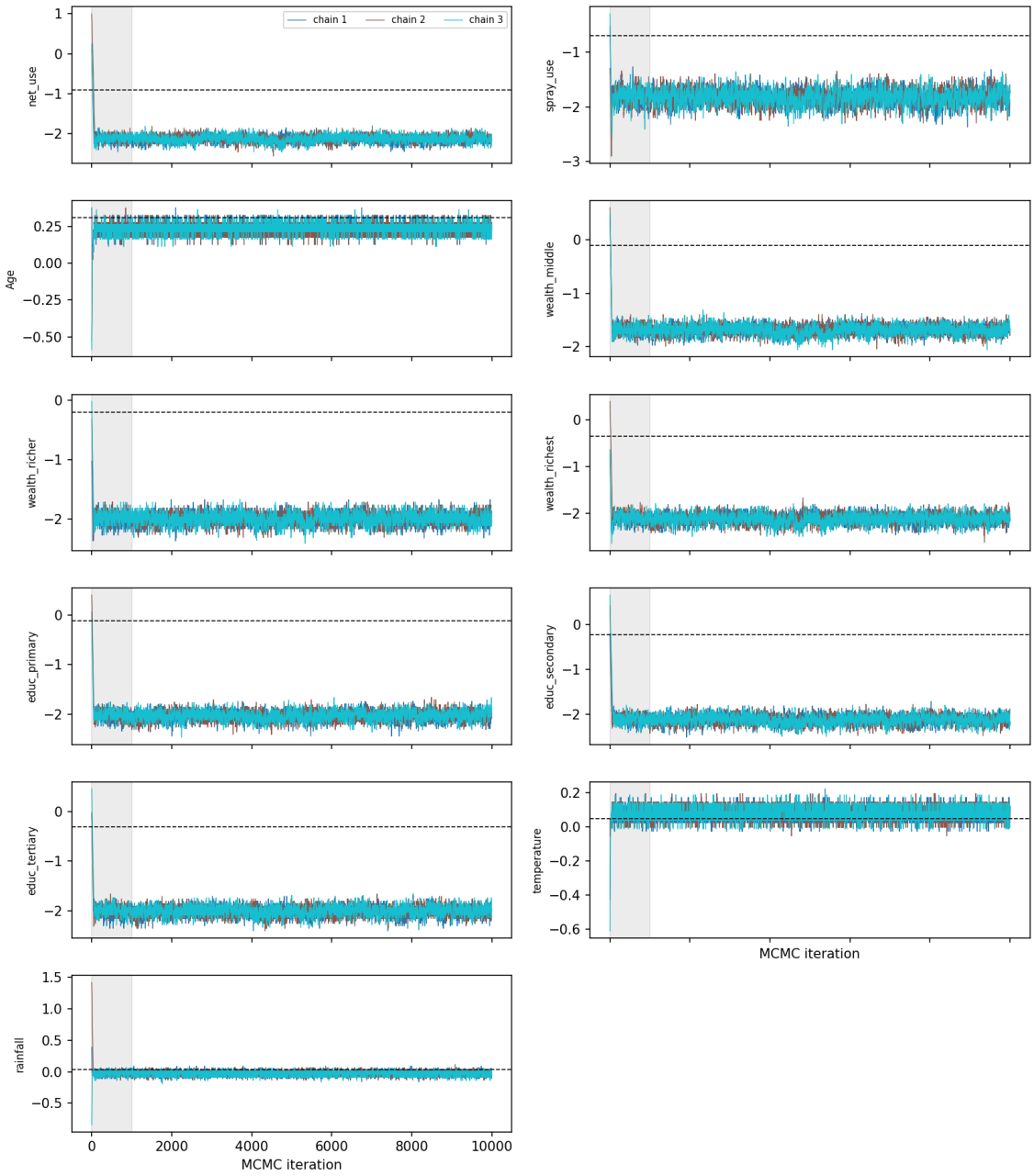


Figure 4.1: Traceplots for Model parameters

4.4 Model Performance under Simulation

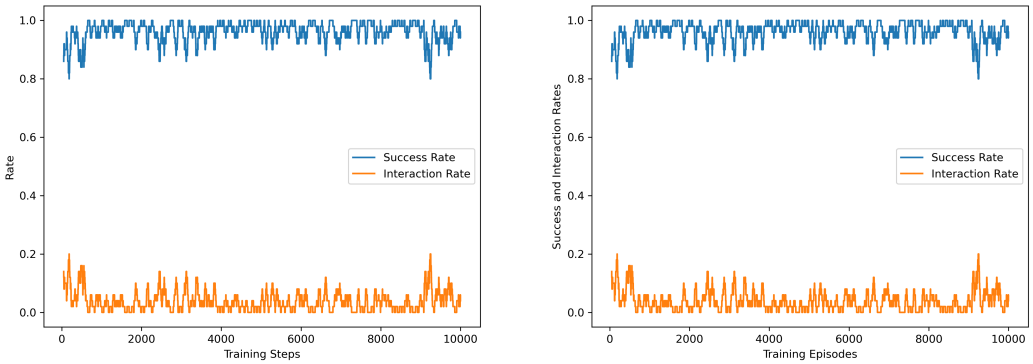
The section presents performance results of the model in predicting the spread of malaria based on simulated data. Success and infection rates of the mosquito and human, performance comparison of the proposed model with the multi-arm bandit model, are presented, model training duration, case prediction error (3.45) and cross-location adaptivity (3.46). The RL-multi-arm bandit model achieved

95.9% success rate and an infection rate of 4.1%, this implies that the model requires more trial and error interactions to achieve success due to less optimised policies and greater exploration. The GIS-enhanced reinforcement learning model achieved a success rate of 96.6% and a lower infection rate of 3.4% compared to the RL-multi-arm bandit model. These results demonstrate that the model is able to efficiently perform prediction with minimal environmental engagement. This effectively solves the issue of weak efficiency and low success rate attributed to overestimation and sample errors; the infection rate also shows that the majority of unsuccessful movements of the human agent are due to interaction with the mosquito Table 4.2). GIS-Enhanced Bayesian reinforcement learning model excels in precision and autonomy, making it suitable for stable and goal-oriented environments. At the same time, the RL-multi-arm bandit Model's higher interaction rate reflects adaptability in dynamic and uncertain scenarios. The proposed model has a significantly higher training duration of 0.659 hours compared to the RL-multi-arm bandit model with a training duration of 0.316 hours. These varying interactions and training durations underscore the importance of aligning model choice based on operational priorities to maximise success with minimal intervention, as demonstrated by the GIS-Enhanced Bayesian reinforcement learning model, and the balancing of exploration and exploitation demonstrated by the RL-multi-arm bandit (MAB) model.

Table 4.2: Comparison of Model Performance

	Performance Metrics		Training Duration (Hours)				
Model	Success rate	Interaction rate	100	200	500	800	1000
RL-MAB Model	95.9%	4.1%	0.291	0.294	0.300	0.306	0.316
GIS-BRL Model	96.6%	3.4%	0.639	0.644	0.647	0.651	0.659

The figure 4.2 shows success and interaction rates of the human agent with the disease vector, the figures demonstrates that there is a significantly small but significant difference in learning and interaction rates of human under the two models.



(a) RL-MAB Model

(b) GIS-Enhanced reinforcement learning model

Figure 4.2: Success and Interaction Rates under RL-MAB and GIS-Enhanced reinforcement learning models

The figure 4.3 shows tracking of cumulative rewards earned by human agent per training episode. The initial episodes show possible low and negative rewards as the agent explores actions randomly therefore incurring suboptimal outcomes. As the training progresses, the rewards gradually rise, this indicate improvement in decision decision-making aligning with the environmental rewards. The fluctuations signals exploration phase in refining policies. The figure also shows that there are some sustained upward trends, this indicates successful learning with few stagnation which suggest that there was also convergence to local optimum. Therefore, these rewards trajectories explain the human agent evolving ability to balance exploration and exploitation to maximize its long-term success.

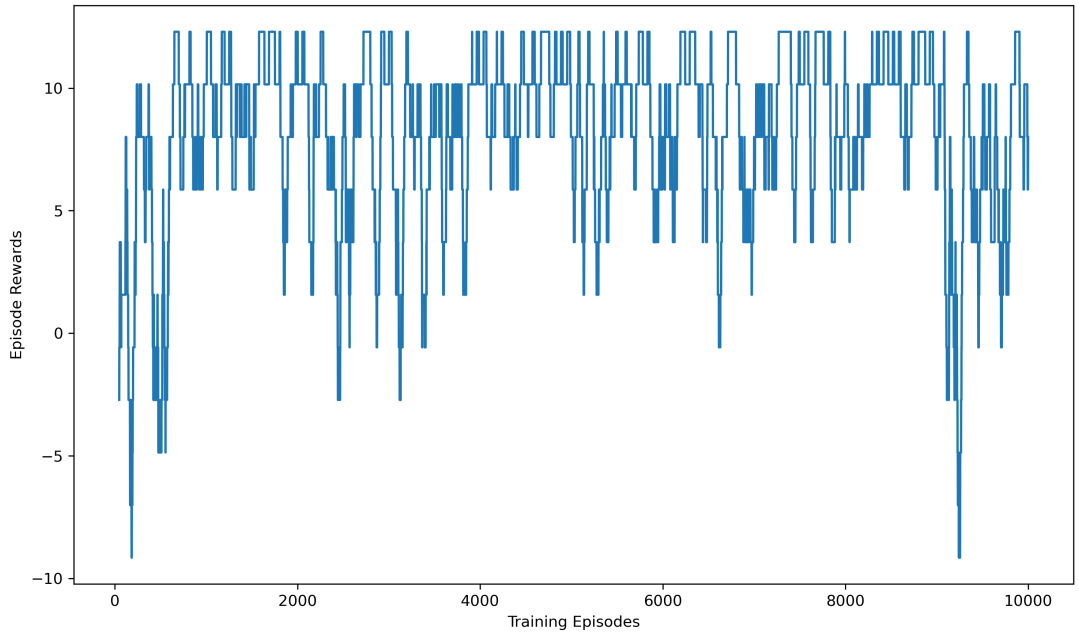


Figure 4.3: Reward values

Table 4.3 shows the number of predicted malaria cases based on sampled cases and the optimal actions that the human took to prevent the spread of malaria in eight counties of Kenya. Use of mosquito nets requires a higher adoption rate between 74.3% and 99.9% to significantly reduce the number of cases, which indicates the importance of the action in the prevention of malaria. Using spray mosquito repellent requires an adoption rate between 10.6% and 70%, which is lower than that of treated mosquito nets. The results on optimal actions are consistent with WHO malaria control frameworks that emphasize control and preventive measures such as the use of insecticide-treated nets (ITNs) and localized spraying/indoor residual spraying (IRS) to reduce mosquito populations WHO (2023). These results reflect a balanced and evidence-based approach specific to local needs. The cross-location adaptivity of spray repellent is 0.6184 and 0.9349 for treated nets; these are greater than zero, indicating the policy learned location-adaptive behaviour. Homa Bay had a small (0.0088) case prediction error compared to other counties, Bungoma(0.0664), Kakamega(0.0167), Busia(0.0195), Siaya(0.0858), Vihiga(0.0329), Kisumu(0.0151), and Migori has the highest case prediction error of 0.0250.

Table 4.3: County predicted cases with corresponding optimal action taken by the population (GIS-RL Model)

		Optimal Intervention (%)			
County	Cases	Spray rep	T-nets	Pred Cases	Pred Error
Bungoma	233	65	99.4	143	0.0664
Kakamega	153	15	74.3	80	0.0167
Busia	103	20	99.9	61	0.0195
Siaya	140	11.8	99.4	103	0.0858
Vihiga	164	46.0	99.3	73	0.0329
Kisumu	127	29	98.5	62	0.0151
Homa Bay	136	10.6	87.2	41	0.0088
Migori	131	70	95.7	46	0.0250
Cross-location adaptivity		0.6184	0.9349		

4.5 Malaria Data

In this subsection, the applications of 4.4 was demonstrated using malaria data sets. Parameter estimates, RMSE and indicator $I_{(s_t)}|_{(\text{Max,Min})}$ were determined. The data set represents malaria cases collected in Kenya for 2015 and 2020 malaria indicator DHS survey. The data was obtained from `https://dhsprogram.com/data/dataset/Kenya_MIS_2020.cfm?flag=1`.

4.5.1 Data Preprocessing

After retrieving the data from the KDHS, the data was read into the Python workspace, and data inspection was done, where the following processes were performed

1. Missing data were removed using dropna() calls before analysis, and replacing infinite values (np.inf/-np.inf) with NaN before dropping
2. Normalization of continuous variables and inverse-distance weights for spatial autocorrelation (Moran's I) were row-normalized
3. Shapefiles were loaded via geopandas, CRS checked/reprojected (to_crs(epsg=3857)), and county statistics were merged onto the geometry; centroids were computed for labeling/weighting for spatial mapping of the data

4. Data was split into training (80%) and testing (20%) for the reinforment learning model

4.5.2 Parameter Estimation

The temperature parameter 0.1917 indicates a positive association with malaria risk, where a unit increase in temperature increases the log-odds of malaria risk by 0.1917. CI $[0.0496, 0.9529]$ shows that there is a 95% chance that the true parameter is within the interval. Rainfall parameter -0.2311 indicates a negative association with malaria risk, where a unit increase in rainfall decreases the log-odds of malaria risk by 0.2311. CI $[-0.4812, -0.0016]$ shows that there is a 95% chance that the true parameter is within the interval. Age parameter 0.4666 indicates a positive association with malaria risk, where a unit increase in age increases the log-odds of malaria by 0.4666, indicating that older ages are associated with significantly higher risk of contracting the disease, the CI $[0.0128, 0.9675]$ shows that there is 95% chances that the true parameter is within the interval.

The primary level of education parameter (-0.1279) indicates a negative association with malaria risk, where a positive change in primary education level reduces the log-odds of malaria by 0.1279 compared to an individual with no education. The CI $[-1.0612, -0.0173]$ shows that there is 95% chances that the true parameter is within the interval. Secondary level of education parameter -0.1060 indicates a negative association with malaria risk, where a positive change in education (secondary) level reduces the risk log-odds of malaria by 0.1060 compared to an individual with no education, the CI $[-1.4397, -0.0168]$ shows that there is 95% chances that the true parameter is within the interval. Tertitairy level of education parameter -1.0108 indicates a negative association with malaria risk, where a positive change in education (tertiary) level reduces the risk log-odds of malaria by 1.0108 compared to an individual with no education, the CI $[-1.8976, -0.1329]$ shows that there is 95% chances that the true parameter is within the interval.

Middle level of wealth parameter -0.0428 indicates a negative association with malaria risk, where a positive change in wealth (middle) level reduces the risk log-odds of malaria by 0.0428 compared to an individual in wealth (poor) level, the CI $[-0.6107, -0.0215]$ shows that there is 95% chances that the true parameter is within the interval. Richer level of wealth parameter -0.5404 indicates a negative association with malaria risk, where a positive change in wealth (richer) level reduces the risk log-odds of malaria by 0.5404 compared to an individual in wealth (poor) level, the CI $[-1.1169, -0.0042]$ shows that there is 95% chances that the true parameter is within the interval. The richest level of wealth parameter -0.1607 indicates a negative association with malaria risk, where a positive change in wealth (richest) level reduces the risk log-odds of malaria by 0.1607 compared to an individual in wealth (poor) level, the CI $[-0.6976, -0.1197]$ shows that there is 95% chances that the true parameter is within the interval. Parameter estimates for treated mosquito nets (-0.5422) and spray repellent (-0.8839) have a significantly lower log-odds of contracting malaria.

Table 4.4: Parameter Estimates and Credible Intervals for the Estimated parameters

Variable	Estimate	Credible Interval
Temperature	0.1917	[0.0496, 0.9529]
Rainfall	-0.2311	[-0.4812,-0.0016]
Age	0.4666	[0.0128, 0.9675]
$\text{Education}_{\textbf{Primary}}$	-0.1279	[-1.0612,-0.0173]
$\text{Education}_{\textbf{Secondary}}$	-0.1060	[-1.4397,-0.0168]
$\text{Education}_{\textbf{Tertiary}}$	-1.0108	[-1.8976,-0.1329]
$\text{Wealth index}_{\textbf{Middle}}$	-0.0428	[-0.6107,-0.0215]
$\text{Wealth index}_{\textbf{Richer}}$	-0.5404	[-1.1169,-0.0042]
$\text{Wealth index}_{\textbf{Richest}}$	-0.1607	[-0.6976,-0.1197]
Treated mosquito Net	-0.5422	[-1.3840,-0.2450]
Spray repellent	-0.8839	[-2.4890,-0.2570]

4.5.3 Prescriptive Optimal Interventions and Prediction of Malaria Cases

Fitting the model to the malaria dataset, results showed that malaria cases significantly reduced due to varying human interventions in different regions. The

prescriptive optimal interventions in Bungoma are mosquito spray repellent applied at (54%) and treated mosquito nets at (78.4%) reducing malaria cases to 29 from 53 cases, in Kakamega mosquito application rate of spray repellent at (25.1%) and treated mosquito nets at (90%) reduce malaria cases to 15 from 29 cases, in Busia mosquito application rate of spray repellent at (40.1%) and treated mosquito nets at (94.5%) reduce malaria cases to 23 from 40 cases, in Siaya application rate of mosquito spray repellent at (68%) and treated mosquito nets at (75.1%) reduce malaria cases to 21 from 44 cases, in Vihiga mosquito application rate of spray repellent at (35.6%) and treated mosquito nets at (99.2%) reduce malaria cases to 5 from 12 cases, in Kisumu application rate of mosquito spray repellent at (50.3%) and treated mosquito nets at (87.3%) reduce malaria cases to 9 from 23 cases, in Homa bay application rate of mosquito spray repellent at (29.5%) and treated mosquito nets at (96.1%) reduce malaria cases to 3 from 10 cases, and in Migori application rate of mosquito spray repellent at (71.2%) and treated mosquito nets at (90.6%) reduce malaria cases to 12 from 23 cases (Table 4.5). The model demonstrated a significant effect of human combined interventions in reducing the number of malaria cases in the region. The cross-location adaptivity of spray repellent is 0.568 and 0.6305 for treated nets; these are greater than zero, indicating the policy learned location-adaptive behaviour. Kisumu had a small (0.0065) case prediction error compared to other counties, Bungoma(0.0622), Kakamega(0.065), Busia(0.0879), Migori(0.02), Vihiga(0.0230), Homa Bay(0.0195), and Siaya has the highest case prediction error of 0.1008.

Table 4.5: Optimal Strategies and Predicted Malaria Cases

	Cases		Prescriptive Optimal Interventions (%)			
County	2015	2020	Spray repel	T-nets	Pred Cases	Pred Error
Bungoma	25	53	54.0	87.4	29	0.0622
Kakamega	27	29	25.1	90.0	15	0.065
Busia	39	40	40.1	94.5	23	0.0879
Siaya	40	44	68.0	75.1	21	0.1008
Vihiga	14	12	35.6	99.2	5	0.0230
Kisumu	22	23	50.3	87.3	9	0.0065
Homa Bay	27	10	29.5	96.1	3	0.0195
Migori	19	23	71.2	90.6	12	0.0200
Cross-location adaptivity			0.568	0.6305		

Figure 4.4 shows a spatial distribution of malaria cases in 8 of the endemic western Kenya counties. The figure presents the initial number of cases (based on the 2020 survey cases) before humans learned interventions using spray repellent and treated mosquito nets, and the number of cases in each county after humans learned simultaneous adoption of spray repellent and treated mosquito nets to prevent the spread of malaria. The figure visualizes the cases in Table 4.5.

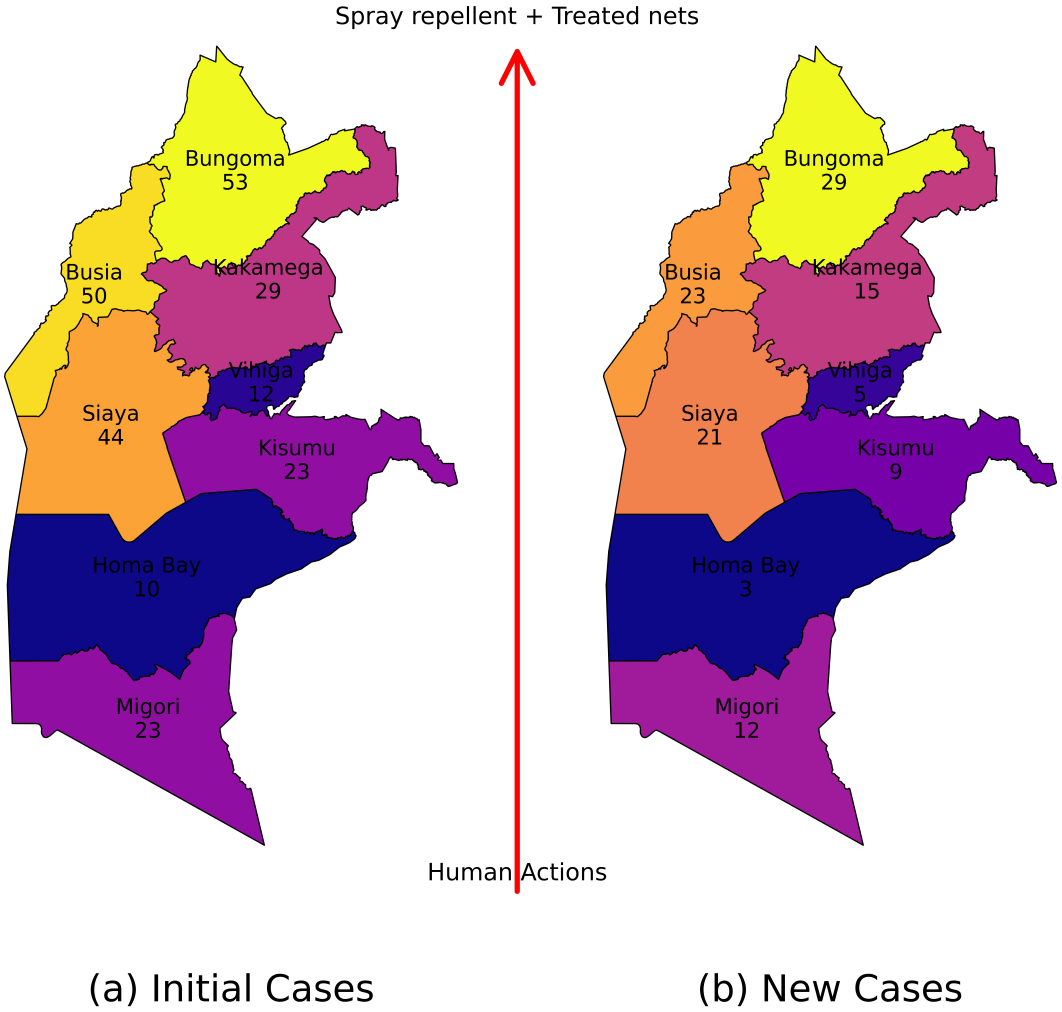


Figure 4.4: Spatial Distribution of Predicted Malaria Cases in the Region

4.5.4 Intervention Effect Gradients

The intervention effect gradient describes the way in which the effectiveness of the intervention changes across counties, affecting the number of malaria cases. Homa Bay, Kisumu, and Vihiga counties demonstrated a significant reduction in malaria

cases due to a positive change in the use of treated mosquito net interventions. Kakamega, Migori, and Bungoma experienced a moderate reduction in malaria cases due to a positive change in the use of treated mosquito nets. The upward use of spray repellent showed a reduction in the number of cases. Homa Bay, Kisumu, and Vihiga counties observed a positive reduction in malaria cases. The adoption of spray repellents and treated mosquito nets demonstrates a significant positive effect in reducing malaria cases in Homa Bay, Kisumu, and Vihiga (Figure 4.5). The varying effect of the interventions underscores that a one-size-fits-all approach is ineffective, and there is a need for county-specific implementation of the intervention rather than blanket implementation.

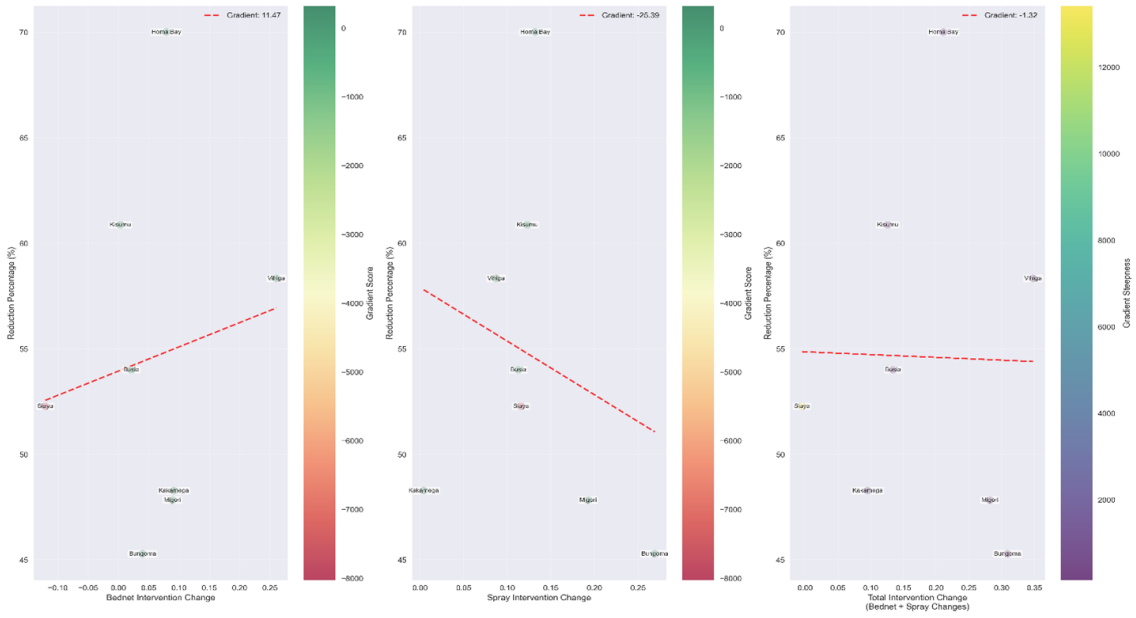


Figure 4.5: Intervention Gradients

4.6 Model Performance

The GIS-enhanced Bayesian reinforcement learning model demonstrated better performance in predicting the number of malaria cases given simultaneous adoption of interventions by humans and estimating prescriptive optimal interventions with a success rate of 97.4 per cent and an infection rate of 2.6% compared to the GIS RL Multi-armed Bandit algorithm that has a success rate of 89.1 per cent and an infection rate of 10.9%. The model, however, has a longer training time

of 17104.98 sec compared to the GIS reinforcement learning multi-armed bandit algorithm (441.28 Sec) (Table (4.6)).

Table 4.6: Model Performance

Model	Success rate	Interaction rate	Training Duration (Sec)
GIS-BRL Model Model	97.4%	2.6%	17104.98
GIS-RL MAB Model	89.1%	10.9%	441.28

4.7 Discussion

The effects of environmental, socioeconomic factors, and spatial effects on the spread of malaria were incorporated into the model, rewarding humans for avoiding interaction with mosquitoes. Descriptive analysis of the dataset demonstrated a decline in malaria cases across the counties from 2015 to 2020 which can be attributed to significant implementation of preventive and control interventions such as use of insecticide-treated nets (ITNs), Indoor residual spraying (IRS) and seasonal malaria chemoprevention (SMC) based on the approach to malaria prevention Kenyan Ministry of Health (2025) and Kenya malaria strategy 2019 – 2023 National Malaria Control Programme (2019). The strategy aimed to reduce malaria incidence and deaths by 75% of 2016 levels by 2023, based on the knowledge that adequate control requires strong collaboration with the public and private sectors. The spatial variation of malaria cases across counties is consistent with the findings of Kioko and Blanford (2025); Elnour et al. (2023), which demonstrate significant variation in malaria cases across different regions due to environmental, economic, and geographical factors. The trace plots demonstrate convergence of parameters associated with temperature, rainfall, age, wealth index (middle), wealth index (richer), wealth index (richest) and education of an individual, indicating that the chains have efficiently explored the posterior distributions and achieved stationarity. This shows reliable inference given that the model includes spatially correlated random effects. From data analysis, rainfall consistently exhibits a negative association with malaria risk, which is consistent

with studies that demonstrate that excessive rainfall leads to flushing of mosquito breeding sites, therefore altering local hydrology Mordecai et al. (2013). Temperature shows a positive effect, which aligns with the temperature-dependent development of Plasmodium parasites and Anopheles vectors Parham and Michael (2009). Temperature, humidity, and rainfall have also been shown to have a strong association with malaria incidence in the region (Mordecai et al., 2019; Kioko and Blanford, 2025). Education and wealth gradients reveal that only secondary or higher education and the richest wealth quintiles contribute a significant protection, which is consistent with the findings that household assets and health literacy are threshold dependent Tusting et al. (2013). Gender (female) indicates lower risk, highlighting a differential exposure and care-seeking behavior by females compared to males Ngutu et al. (2025); Diiro et al. (2016). Model evaluation results show that the model is highly reliable for predicting future malaria cases, given human interventions and the estimation of optimal interventions that humans can take to reduce the spread of vector-borne infectious diseases such as malaria. The study presented a performance comparison of the GIS-enhanced Bayesian reinforcement learning model against GIS RL and GIS reinforcement learning using the MAB algorithm, with the GIS-enhanced Bayesian reinforcement learning model showing better performance in predicting the number of malaria cases and identifying optimal interventions. The study, however, had limitations associated with the data collected in 2020, where even though the data was adequate for testing the model capability, data was collected during the COVID-19 pandemic, where movement restrictions were in place, hindering the collection of data; the data also had many missing values and missing data on the GPS location of participants. The study is original and the first of its kind to establish a GIS-enhanced Bayesian reinforcement learning model for estimating prescriptive optimal intervention intensities to prevent the spread of malaria and for predicting malaria cases in the lake-endemic region and other regions across sub-Saharan Africa. ML and statistical methods have been significantly used in the prediction of malaria cases. However, a GIS-enhanced Bayesian

reinforcement learning model has not been explored for estimating prescriptive optimal interventions to prevent the spread of malaria and for predicting malaria cases. The model demonstrated adequacy due to its consideration of human-mosquito interactions and interventions by humans (Shakya et al., 2023). AI and ML models have extensively demonstrated greater success in other medical fields and forecasting vector-borne infectious diseases (Egbuna et al., 2025).

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the conclusion drawn from the study's findings, in the order of the study's objectives. Recommendations from the study are also presented in this chapter, highlighting suggestions for adopting the study's findings and the future direction of the developed model.

5.2 Conclusion

The research developed an advanced predictive system that models the spread of vector-borne infectious diseases within the population in different counties in the presence of control and preventive interventions by humans. The model was developed by modifying the TD3 algorithm for continuous spaces, integrating a GIS system, and performing Bayesian estimation of model parameters. The developed model provides flexibility, enabling users to simulate disease spread across different regions and underscoring its practicality. Applying the model facilitated the selection of interventions that humans should adopt to reduce the transmission of vector-borne diseases and the prediction of the number of cases based on their adoption.

Bayesian estimation of parameters shows that the model parameters incorporate uncertainty and vary spatially across regions, capturing real-world scenarios, as regions have varying characteristics that influence disease spread and human adoption of interventions, thereby facilitating region-specific decision-making to prevent and control vector-borne infectious diseases.

The study results show that a higher adoption rate of mosquito nets and spray repellent combined significantly reduces the number of malaria cases towards a single-digit number in the endemic western counties of Kenya. Temperature, rainfall, age, wealth index (middle), wealth index (richer), wealth index (richest), and education significantly influence the risk of contracting malaria in the region. The study also underscored the importance of using GIS-Enhanced Bayesian reinforcement learning to predict possible future malaria cases, given the adoption of combined malaria control interventions. Additionally, the study described the space distribution of malaria incidences to identify counties that need urgent consideration for interventions to reduce the number of cases futher; these counties include Bungoma, Kakamega, and Busia. Further, the study is the first of its kind to apply GIS-Enhanced Bayesian reinforcement learning in predicting malaria cases in the area and comparative performance with the MAB algorithm serving as a baseline, and estimating combined optimal interventions for controlling vector-borne infectious diseases such as malaria. Although the model shows superior performance metrics compared to other methodologies, it is crucial to acknowledge that each algorithm has its own strengths and weaknesses, and that the proposed model's superiority was context-specific. While the study primary focus was vector-borne infectious disease in the population, the underlying principles of our tool possess potential applicability to other sectors.

5.3 Recommendations

The findings of this research have broader implications for public health, providing a reference for the national malaria control program department in determining

the intensity of preventive and control measures to reduce vector-borne infectious disease cases. The model was, however, context-specific and based on the available data, with incomplete or unavailable data, which is a real-life scenario, especially in public health research.

The study recommends future research to apply the model to large real-world datasets, when available, to further validate its performance, given that the developed model performed well with large simulated data. Future research could also replicate the results and explore additional interventions that humans can adopt to prevent and control malaria transmission. The model can also be applied in other regions within Kenya and in other sub-Saharan African countries affected by malaria and other vector-borne infectious diseases to further generalize the developed model.

REFERENCES

Adeola, A. M., Botai, J. O., Olwoch, J. M., Rautenbach, H. C. J. d. W., Adisa, O. M., De Jager, C., Botai, C. M., and Aaron, M. (2019). Predicting malaria cases using remotely sensed environmental variables in nkomazi, south africa. *Geospatial health*, 14(1).

Adiga, A., Wang, L., Hurt, B., Peddireddy, A., Porebski, P., Venkatramanan, S., Lewis, B. L., and Marathe, M. (2021). All models are useful: Bayesian ensembling for robust high resolution covid-19 forecasting. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 2505–2513.

Agrawal, A., Won, S. J., Sharma, T., Deshpande, M., and McComb, C. (2021). A multi-agent reinforcement learning framework for intelligent manufacturing with autonomous mobile robots. *Proceedings of the Design Society*, 1:161–170.

Ajith, A., Manoj, K., Kiran, H., Pillai, P. J., and Nair, J. J. (2020). A study on prediction and spreading of epidemic diseases. In *2020 International Conference on Communication and Signal Processing (ICCSP)*, pages 1265–1268. IEEE.

Anokye, R., Acheampong, E., Owusu, I., and Isaac Obeng, E. (2018). Time series analysis of malaria in kumasi: Using arima models to forecast future incidence. *Cogent social sciences*, 4(1):1461544.

Arenas, A., Cota, W., Gomez-Gardenes, J., Gomez, S., Granell, C., Matamalas, J. T., Soriano, D., and Steinegger, B. (2020). A mathematical model for the spatiotemporal epidemic spreading of covid19. *MedRxiv*, pages 2020–03.

Asheralieva, A. and Niyato, D. (2020). Bayesian reinforcement learning and bayesian deep learning for blockchains with mobile edge computing. *IEEE Transactions on Cognitive Communications and Networking*, 7(1):319–335.

Atkinson, A., Ellenberger, B., Piezzi, V., Kaspar, T., Salazar-Vizcaya, L., Endrich, O., Leichtle, A. B., and Marschall, J. (2023). Extending outbreak investigation with machine learning and graph theory: Benefits of new tools with application to a nosocomial outbreak of a multidrug-resistant organism. *Infection Control & Hospital Epidemiology*, 44(2):246–252.

Balcan, D., Gonccalves, B., Hu, H., Ramasco, J. J., Colizza, V., and Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model. *Journal of computational science*, 1(3):132–145.

Benedum, C. M., Shea, K. M., Jenkins, H. E., Kim, L. Y., and Markuzon, N. (2020). Weekly dengue forecasts in iquitos, peru; san juan, puerto rico; and singapore. *Plos Neglected tropical diseases*, 14(10):e0008710.

Bharadiya, J. P. (2023). A review of bayesian machine learning principles, methods, and applications. *International Journal of Innovative Science and Research Technology*, 8(5):2033–2038.

Bierkens, J. (2016). Non-reversible metropolis-hastings. *Statistics and Computing*, 26(6):1213–1228.

Bin, S., Sun, G., and Chen, C.-C. (2019). Spread of infectious disease modeling and analysis of different factors on spread of infectious disease based on cellular automata. *International journal of environmental research and public health*, 16(23):4683.

Bone, C. (2009). Integrating complexity science and artificial intelligence: Gis, agents and reinforcement learning for modeling forest cover change.

Bowling, M. and Veloso, M. (2000). *An analysis of stochastic game theory for multiagent reinforcement learning*. Citeseer.

Brock, P. M., Fornace, K. M., Grigg, M. J., Anstey, N. M., William, T., Cox, J., Drakeley, C. J., Ferguson, H. M., and Kao, R. R. (2019). Predictive analysis

across spatial scales links zoonotic malaria to deforestation. *Proceedings of the Royal Society B*, 286(1894):20182351.

Bucsoniu, L., Babuvska, R., and De Schutter, B. (2010). Multi-agent reinforcement learning: An overview. *Innovations in multi-agent systems and applications-1*, pages 183–221.

Chae, S., Kwon, S., and Lee, D. (2018). Predicting infectious disease using deep learning and big data. *International journal of environmental research and public health*, 15(8):1596.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Chen, X., Wang, C., Zhou, Z., and Ross, K. (2021). Randomized ensembled double q-learning: Learning fast without a model. *arXiv preprint arXiv:2101.05982*.

Chen, Y., Chu, C. W., Chen, M. I. C., and Cook, A. R. (2018). The utility of lasso-based models for real time forecasts of endemic infectious diseases: A cross country comparison. *Journal of biomedical informatics*, 81:16–30.

Chimmula, V. K. R. and Zhang, L. (2020). Time series forecasting of covid-19 transmission in canada using lstm networks. *Chaos, solitons & fractals*, 135:109864.

Chowell, G. and Rothenberg, R. (2018). Spatial infectious disease epidemiology: on the cusp.

Churcher, T. S., Sinden, R. E., Edwards, N. J., Poulton, I. D., Rampling, T. W., Brock, P. M., Griffin, J. T., Upton, L. M., Zakutansky, S. E., Sala, K. A., et al. (2017). Probability of transmission of malaria from mosquito to human is regulated by mosquito parasite density in naive and vaccinated hosts. *PLoS pathogens*, 13(1):e1006108.

Corley, C. D., Pullum, L. L., Hartley, D. M., Benedum, C., Noonan, C., Rabinowitz, P. M., and Lancaster, M. J. (2014). Disease prediction models and operational readiness. *PloS one*, 9(3):e91989.

Cromley, E. K. (2003). Gis and disease. *Annual review of public health*, 24(1):7–24.

Dantas, E., Tosin, M., and Cunha Jr, A. (2018). Calibration of a seir–sei epidemic model to describe the zika virus outbreak in brazil. *Applied Mathematics and Computation*, 338:249–259.

Diiro, G. M., Affognon, H. D., Muriithi, B. W., Wanja, S. K., Mbogo, C., and Mutero, C. (2016). The role of gender on malaria preventive behaviour among rural households in kenya. *Malaria Journal*, 15(1):14.

Dimitrov, N. B. and Meyers, L. A. (2010). Mathematical approaches to infectious disease prediction and control. In *Risk and optimization in an uncertain world*, pages 1–25. INFORMS.

Egbuna, I. K., Orisakahunsi, B. A., Agboro, H., Uduh, I. T., Ojo, S. S., Mustapha, F. A., and Olanrewaju, A. O. (2025). Artificial intelligence for forecasting climate-driven vector-borne disease outbreaks.

Elnour, Z., Grethe, H., Siddig, K., and Munga, S. (2023). Malaria control and elimination in kenya: economy-wide benefits and regional disparities. *Malaria journal*, 22(1):117.

Esri (2024). What is gis? | geographic information system (gis).

Florens, J. P., Mouchart, M., and Rolin, J.-M. (2019). *Elements of Bayesian statistics*. CRC Press.

Fujimoto, S., Hoof, H., and Meger, D. (2018). Addressing function approximation error in actor-critic methods. In *International conference on machine learning*, pages 1587–1596. PMLR.

Funk, S., Camacho, A., Kucharski, A. J., Eggo, R. M., and Edmunds, W. J. (2018). Real-time forecasting of infectious disease dynamics with a stochastic semi-mechanistic model. *Epidemics*, 22:56–61.

Galvani, A. P. and May, R. M. (2005). Dimensions of superspreading. *Nature*, 438(7066):293–295.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian data analysis (Vol. 2)*. Taylor & Francis Boca Raton.

Ghannam, R. B. and Techtmann, S. M. (2021). Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring. *Computational and Structural Biotechnology Journal*, 19:1092–1107.

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., and Colaneri, M. (2020). Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. *Nature medicine*, 26(6):855–860.

Goo, T., Apio, C., Heo, G., Lee, D., Lee, J. H., Lim, J., Han, K., and Park, T. (2021). Forecasting of the covid-19 pandemic situation of korea. *Genomics & Informatics*, 19(1).

Grzenda, W. (2016). Informative versus non-informative prior distributions and their impact on the accuracy of bayesian inference. *Statistics in Transition. New Series*, 17(4):763–780.

Guo, Y., Feng, Y., Qu, F., Zhang, L., Yan, B., and Lv, J. (2020). Prediction of hepatitis e using machine learning models. *Plos one*, 15(9):e0237750.

Heesterbeek, H., Anderson, R. M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., Eames, K. T. D., Edmunds, W. J., Frost, S. D. W., Funk, S., et al. (2015). Modeling infectious disease dynamics in the complex landscape of global health. *Science*, 347(6227):aaa4339.

Hu, J., Wellman, M. P., et al. (1998). Multiagent reinforcement learning: theoretical framework and an algorithm. In *ICML*, volume 98, pages 242–250.

Hunter, E., Mac Namee, B., and Kelleher, J. D. (2020). A model for the spread of infectious diseases in a region. *International journal of environmental research and public health*, 17(9):3119.

Jewell, C. P., Kypraios, T., Christley, R. M., and Roberts, G. O. (2009). A novel approach to real-time risk prediction for emerging infectious diseases: a case study in avian influenza h5n1. *Preventive veterinary medicine*, 91(1):19–28.

Jung, S., Jeoung, J., Kang, H., and Hong, T. (2021). Optimal planning of a rooftop pv system using gis-based reinforcement learning. *Applied Energy*, 298:117239.

Kamana, E., Zhao, J., and Bai, D. (2022). Predicting the impact of climate change on the re-emergence of malaria cases in china using lstmseq2seq deep learning model: a modelling and prediction analysis study. *BMJ open*, 12(3):e053922.

Karras, A., Karras, C., Sioutas, S., Makris, C., Katselis, G., Hatzilygeroudis, I., Theodorou, J. A., and Tsolis, D. (2023). An integrated gis-based reinforcement learning approach for efficient prediction of disease transmission in aquaculture. *Information*, 14(11):583.

Katsevich, A. (2023). Improved dimension dependence in the bernstein von mises theorem via a new laplace approximation bound. *arXiv preprint arXiv:2308.06899*.

Kenyan Ministry of Health (2025). Kenya's comprehensive approach to malaria prevention. https://www.health.go.ke/kenyas-comprehensive-approach-malaria-prevention.

Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.

Keshavamurthy, R., Dixon, S., Pazdernik, K. T., and Charles, L. E. (2022). Predicting infectious disease for biopreparedness and response: A systematic review of machine learning and deep learning approaches. *One Health*, page 100439.

Khalilpourazari, S. and Doulabi, H. H. (2021). Using reinforcement learning to forecast the spread of covid-19 in france. In *2021 IEEE international conference on autonomous systems (ICAS)*, pages 1–8. IEEE.

Khalilpourazari, S. and Hashemi Doulabi, H. (2021). Designing a hybrid reinforcement learning based algorithm with application in prediction of the covid-19 pandemic in quebec. *Annals of Operations Research*, pages 1–45.

Khiem, N. M., Takahashi, Y., Yasuma, H., Oanh, D. T. H., Hai, T. N., Ut, V. N., and Kimura, N. (2022). Use of gis and machine learning to predict disease in shrimp farmed on the east coast of the mekong delta, vietnam. *Fisheries Science*, pages 1–13.

Kioko, C. and Blanford, J. (2025). Malaria survey data and geospatial suitability mapping for understanding spatial and temporal variations of risk across kenya. *Parasite Epidemiology and Control*, 28:e00399.

Kolesnikov, A. A., Kikin, P. M., and Portnov, A. M. (2019). Diseases spread prediction in tropical areas by machine learning methods ensembling and spatial analysis techniques. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:221–226.

Kumar, S. L. et al. (2021). Predictive analytics of covid-19 pandemic: statistical modelling perspective. *Walailak Journal of Science and Technology (WJST)*, 18(16):15583–14.

Li, L., Lv, Y., and Wang, F.-Y. (2016). Traffic signal timing via deep reinforcement learning. *IEEE/CAA Journal of Automatica Sinica*, 3(3):247–254.

Li, P., Chen, D., Wang, Y., Zhang, L., and Zhao, S. (2024). Path planning of mobile robot based on improved td3 algorithm in dynamic environment. *Heliyon*.

Li, Y. (2017). Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*.

Lipsitch, M. and Viboud, C. (2009). Influenza seasonality: lifting the fog. *Proceedings of the National Academy of Sciences*, 106(10):3645–3646.

Masum, A. K. M., Khushbu, S. A., Keya, M., Abujar, S., and Hossain, S. A. (2020). Covid-19 in bangladesh: a deeper outlook into the forecast with prediction of upcoming per day cases using time series. *Procedia Computer Science*, 178:291–300.

McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.".

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Mopuri, R., Kakarla, S. G., Mutheneni, S. R., Kadiri, M. R., and Kumaraswamy, S. (2020). Climate based malaria forecasting system for andhra pradesh, india. *Journal of Parasitic Diseases*, 44:497–510.

Mordecai, E. A., Caldwell, J. M., Grossman, M. K., Lippi, C. A., Johnson, L. R., Neira, M., Rohr, J. R., Ryan, S. J., Savage, V., Shocket, M. S., et al. (2019). Thermal biology of mosquito-borne disease. *Ecology letters*, 22(10):1690–1708.

Mordecai, E. A., Paaijmans, K. P., Johnson, L. R., Balzer, C., Ben-Horin, T., De Moor, E., McNally, A., Pawar, S., Ryan, S. J., Smith, T. C., et al. (2013). Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecology letters*, 16(1):22–30.

National Malaria Control Programme (2019). Kenya malaria strategy 2019-2023: Towards a malaria-free kenya.

Ngutu, M., Omia, D. O., Ngage, T. O., Oduor, C. A., Ouko, N. O., Oingo, B., Oluoch, I., Kariuki, S., Chikovore, J., Onyango-Ouma, W., et al. (2025). Gender-related factors affecting community malaria-related perceptions and practices in migori county, kenya. *Malaria journal*, 24(1):196.

Nguyen, V.-H., Tuyet-Hanh, T. T., Mulhall, J., Minh, H. V., Duong, T. Q., Chien, N. V., Nhung, N. T. T., Lan, V. H., Minh, H. B., Cuong, D., et al. (2022). Deep learning models for forecasting dengue fever based on climate data in vietnam. *PLoS Neglected Tropical Diseases*, 16(6):e0010509.

Nkiruka, O., Prasad, R., and Clement, O. (2021). Prediction of malaria incidence using climate variability and machine learning. *Informatics in Medicine Unlocked*, 22:100508.

Osman, M. and Adu, I. (2017). Simple mathematical model for malaria transmission. *Journal of Advances in Mathematics and Computer Science*, 25(6):1–24.

Palaniappan, S., Ragavi, V., and David, B. (2022). Prediction of epidemic disease dynamics on the infection risk using machine learning algorithms. *SN computer science*, 3(1):47.

Paolella, M. S. (2018). *Linear models and time-series analysis: regression, ANOVA, ARMA and GARCH*. John Wiley & Sons.

Parham, P. E. and Michael, E. (2009). Modeling the effects of weather and climate change on malaria transmission. *Environmental health perspectives*, 118(5):620.

Patil, S. and Pandya, S. (2021). Forecasting dengue hotspots associated with variation in meteorological parameters using regression and time series models. *Frontiers in public health*, 9:798034.

Pham, U., Luu, Q., and Tran, H. (2021). Multi-agent reinforcement learning approach for hedging portfolio problem. *Soft Computing*, 25:7877–7885.

Rajpurkar, P., Polamreddi, V., and Balakrishnan, A. (2017). Malaria likelihood prediction by effectively surveying households using deep reinforcement learning. *arXiv preprint arXiv:1711.09223*.

Roster, K., Connaughton, C., and Rodrigues, F. A. (2022). Forecasting new diseases in low-data settings using transfer learning. *Chaos, Solitons & Fractals*, 161:112306.

Roth, J. A., Battegay, M., Juchler, F., Vogt, J. E., and Widmer, A. F. (2018). Introduction to machine learning in digital healthcare epidemiology. *Infection Control & Hospital Epidemiology*, 39(12):1457–1462.

Roy, S., Biswas, P., and Ghosh, P. (2021). Spatiotemporal tracing of pandemic spread from infection data. *Scientific reports*, 11(1):17689.

Safi, M. A. and Gumel, A. B. (2011). Mathematical analysis of a disease transmission model with quarantine, isolation and an imperfect vaccine. *Computers and Mathematics with Applications*, 61(10):3044–3070.

Santangelo, O. E., Gentile, V., Pizzo, S., Giordano, D., and Cedrone, F. (2023). Machine learning and prediction of infectious diseases: a systematic review. *Machine Learning and Knowledge Extraction*, 5(1):175–198.

Santosh, T., Ramesh, D., and Reddy, D. (2020). Lstm based prediction of malaria abundances using big data. *Computers in Biology and Medicine*, 124:103859.

Saran, S., Singh, P., Kumar, V., and Chauhan, P. (2020). Review of geospatial technology for infectious disease surveillance: use case on covid-19. *Journal of the Indian Society of Remote Sensing*, 48:1121–1138.

Schwartz, L. (1965). On bayes procedures. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 4(1):10–26.

Shaier, S., Raissi, M., and Seshaiyer, P. (2021). Data-driven approaches for predicting spread of infectious diseases through dinns: Disease informed neural networks. *arXiv preprint arXiv:2110.05445*.

Shakya, A. K., Pillai, G., and Chakrabarty, S. (2023). Reinforcement learning algorithms: A brief survey. *Expert Systems with Applications*, 231:120495.

Sharma, S., Gupta, Y. K., and Mishra, A. K. (2023). Analysis and prediction of covid-19 multivariate data using deep ensemble learning methods. *International Journal of Environmental Research and Public Health*, 20(11):5943.

Shen, L., Jiang, C., Sun, M., Qiu, X., Qian, J., Song, S., Hu, Q., Yelixiati, H., and Liu, K. (2022). Predicting the spatial-temporal distribution of human brucellosis in europe based on convolutional long short-term memory network. *The Canadian Journal of Infectious Diseases & Medical Microbiology= Journal Canadien des Maladies Infectieuses et de la Microbiologie Médicale*, 2022.

Shi, Y., Liu, X., Kok, S.-Y., Rajarethinam, J., Liang, S., Yap, G., Chong, C.-S., Lee, K.-S., Tan, S. S. Y., Chin, C. K. Y., et al. (2016). Three-month real-time dengue forecast models: an early warning system for outbreak alerts and policy decision support in singapore. *Environmental health perspectives*, 124(9):1369–1375.

Shreemathi, M., Senthilkumar, B., Sujithra, S. M., Praisoodi, A., and Rithika, S. (2024). Mastering geospatial analysis with python: Understanding geopandas, gdal, fiona, matplotlib, data integration, and gis tools. In *Ethics, Machine Learning, and Python in Geospatial Analysis*, pages 120–149. IGI Global.

Stojanovic, O., Leugering, J., Pipa, G., Ghozzi, S., and Ullrich, A. (2019). A bayesian monte carlo approach for predicting the spread of infectious diseases. *PloS one*, 14(12):e0225838.

Sun, G. and Bin, S. (2017). Router-level internet topology evolution model based

on multi-subnet composited complex network model. *Journal of Internet Technology*, 18(6):1275–1283.

Tampuu, A., Matiisen, T., Kodelja, D., Kuzovkin, I., Korjus, K., Aru, J., Aru, J., and Vicente, R. (2017). Multiagent cooperation and competition with deep reinforcement learning. *PloS one*, 12(4):e0172395.

Tusting, L. S., Willey, B., Lucas, H., Thompson, J., Kafy, H. T., Smith, R., and Lindsay, S. W. (2013). Socioeconomic development as an intervention against malaria: a systematic review and meta-analysis. *The Lancet*, 382(9896):963–972.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Martens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., et al. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1):1.

Wang, H., Lin, S., and Zhang, J. (2021). Adaptive ensemble q-learning: Minimizing estimation bias via error feedback. *Advances in neural information processing systems*, 34:24778–24790.

White, P. J. (2017). Mathematical models in infectious disease epidemiology. In *Infectious Diseases (Fourth Edition)*, pages 49–53.e1. Elsevier.

WHO (2020a). The top 10 causes of death. `https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death`.

WHO (2020b). Vector-borne diseases. `https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases`.

WHO (2020c). World malaria report 2020. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2020.

WHO (2021). Ethics and governance of artificial intelligence for health: Who guidance.

WHO (2023). Malaria vector control. https://www.who.int/teams/global-malaria-programme/prevention/vector-control.

Xu, J., Xu, K., Li, Z., Meng, F., Tu, T., Xu, L., and Liu, Q. (2020). Forecast of dengue cases in 20 chinese cities based on the deep learning method. *International journal of environmental research and public health*, 17(2):453.

Zhang, P., Feng, K., Gong, Y., Lee, J., Lomonaco, S., and Zhao, L. (2022). Usage of compartmental models in predicting covid-19 outbreaks. *The AAPS Journal*, 24(5):98.

Zhang, T., Ye, Q., Bian, J., Xie, G., and Liu, T.-Y. (2021). Mfvfd: A multi-agent q-learning approach to cooperative and non-cooperative tasks. In *IJCAI*, pages 500–506.

Zhou, Z., Liu, G., and Tang, Y. (2023). Multi-agent reinforcement learning: Methods, applications, visionary prospects, and challenges. arxiv 2023. *arXiv preprint arXiv:2305.10091*.

Zhu, Y., He, Z., and Li, G. (2022). A bi-hierarchical game-theoretic approach for network-wide traffic signal control using trip-based data. *IEEE Transactions on Intelligent Transportation Systems*, 23(9):15408–15419.

APPENDICES

APPENDIX I: PUBLICATIONS

Published

- Kipngetich Gideon, V. Musau, Margaret Kinyua. GIS-Enhanced Bayesian Reinforcement Learning for Vector-Borne Infectious Disease Transmission. International Journal of Data Science and Analysis, 10.11648/j.ijdsa.20261202.12

- Kipngetich Gideon, V. Musau, Margaret Kinyua. Bayesian Geospatial Calibration of Reinforcement Learning for Malaria Transmission Control: Parameter Estimation. International Journal of Data Science and Analysis, 10.11648/j.ijdsa.20261202.11

Under Review for Publication

- Kipngetich Gideon, V. Musau, Margaret Kinyua. Optimal Interventions and Prediction of Malaria Cases in Lake-Endemic Region of Western Kenya using GIS-Enhanced Bayesian Reinforcement Learning Model. Parasite Epidemiology and Control, PAREPI-D-26-00019

APPENDIX II: RESEARCH LICENSE

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Ref No: **475722** Date of Issue: **15/January/2025**

RESEARCH LICENSE



This is to Certify that Mr.. Gideon Kipngetich of Kirinyaga University, has been licensed to conduct research as per the provision of the Science, Technology and Innovation Act, 2013 (Rev.2014) in Baringo, Bomet, Bungoma, Busia, Elgeyo-Marakwet, Homabay, Kakamega, Kericho, Kilifi, Kisumu, Kwale, Migori, Mombasa, Nandi, Siaya, Transnzoia, Uasin-Gishu, Vihiga, Westpokot on the topic: GIS-ENHANCED BAYESIAN REINFORCED LEARNING FOR PREDICTING SPREAD OF VECTOR-BORNE INFECTIOUS DISEASES. A CASE STUDY OF MALARIA IN KENYA for the period ending : 15/January/2026.

License No: **NACOSTI/P/25/415058**

475722

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

See overleaf for conditions

APPENDIX III: Python Algorithm

GitHub Link

https://github.com/gideonlangat/GIS-BRL-Model/blob/main/GIS-BRL-Sim.ipynb