International Journal of Data Science and Analysis
2026; 12(2): 25-36
http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/j.ijdsa.20261202.12

GIS-Enhanced Bayesian Reinforcement Learning for Vector-Borne Infectious Disease Transmission

Kipngetich Gideon[1, *], Victor Muthama Musau[1], Margaret Wambui Kinyua[2]

[1]Pure and Applied Sciences, Kirinyaga University, Kerugoya, Kenya
[2]Mathematics, Statistics and Actuarial Sciences, Karatina University, Karatina, Kenya

Email address:

gideonlangat90@gmail.com (Kipngetich Gideon), vmusau@kyu.ac.ke (Victor Muthama Musau),
mwambui@karu.ac.ke (Margaret Wambui Kinyua)
*Corresponding author

To cite this article:

Kipngetich Gideon, Victor Muthama Musau, Margaret Wambui Kinyua. (2026). GIS-Enhanced Bayesian Reinforcement Learning for Vector-Borne Infectious Disease Transmission. *International Journal of Data Science and Analysis, 12*(2), 25-36.
https://doi.org/10.11648/j.ijdsa.20261202.12

Received: 12 March 2026 ; **Accepted:** 26 March 2026 ; **Published:** 10 June 2026

Abstract: Vector-borne infectious diseases has continue to pose a significant challenge to global public health, and accurate and timely transmission prediction are crucial for effective intervention and control. The development of accurate and efficient vector-borne infectious disease predictive models has been the current trend in disease modeling. Modeling transmission, however, has assumed a discrete transmission space, which is not always ideal in the real world, and little attention has been paid to overestimation biases in predictions. The effectiveness of a predictive model is also determined by its capability to capture a significant number of data characteristics to enhance robust and accurate prediction of cases. The study proposed the application of a GIS-enhanced Bayesian Reinforcement learning model for the transmission of vector-borne infectious diseases, and model performance assessment was determined. The Bayesian quantifies the uncertainty in the parameters of models, and min-max ensemble Q-value estimation minimizes overestimation bias in the model. Simulation study was used to evaluate model performance, success rate, and interaction rate. The findings show that vector and human can avoid interaction with a success rate of 96.2% when human select combined intervention actions spray repellent, insecticide treated nets, larval management, and vaccination. Other variables such as education, wealth index, community participation, and gender empowerment significantly influence the transmission of the disease in the area. The model demonstrates a better performance in describing the transmission of the disease, therefore setting the stage for future research in predictive modeling within sub-Saharan disease-prone regions. The model can be used to determine the appropriate actions that the human should adopt to reduce human-vector interaction.

Keywords: Geographic Information System, Bayesian, Reinforcement Learning, Infectious Disease, Transmission

1. Introduction

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 [1]. 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. Models for Infectious Disease Prediction

Development of mathematical models for modeling infectious diseases has recently gained much attention from researchers [2]. Zhang et al. 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 [3]. Osman et al. developed the Susceptible Exposed Infectious Recovery - Susceptible Exposed Infectious (SEIR-SEI) model, which is an extension of SEIR to model malaria transmission [4], the model has also been used to describe 2016 outbreak of the Zika virus in Brazil [5]. 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.

Due to compartmental models shortcomings, researchers have applied non-compartmental models such as linear models, time-series analysis, and stochastic models [6] to predict the spread of infectious diseases. Forecasting monthly malaria incidence in Kumasi Metropolis applied Autoregressive integrated moving average (ARIMA) [7], seasonal autoregressive integrated moving average (SARIMA) and multi-step polynomial regression and models to forecast malaria cases [8]. 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 [9]. 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.

Stojanovic proposed a Bayesian Monte Carlo approach to derive a spatiotemporal kernel for the prediction of the spread of infectious diseases [10], 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 [11]. 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 [12].

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 that is an important requirement so that RL agents do not overfit for a certain number of initial parameters and generate more uncertainty, these models also does not incorporate super-spreaders and also lack randomness [13].

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 [14] 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 [15, 16].

Integration of machine learning and Least Absolute Shrinkage and Selection Operator (LASSO) models has also been proposed to predict dengue incidences in Singapore [17], LASSO has also demonstrated its effectiveness in predicting trends of dengue diseases in Taiwan, Japan, Singapore, and Thailand, despite models significant performance in short-term prediction, the model was less effective in longer-term predictions [18]. 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 [19].

Comparison of long-short term memory (LSTM), Long-short term memory with Attention (LSTM-ATT), seasonal autoregressive integrated moving average (SARIMA), Random Forest (RF), gradient boosting (XGB), LightGBM, K-Nearest Neighbors (KNN), Keras, Poisson regression, SVR-L, Support Vector Regression (SVR), Convolutional Neural Network (CNN) and CatBoost models in predicting the incidences of dengue diseases 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 to achieve better prediction performance [20, 21].

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) [22]. Application of Long short-term memory (LSTM) on a big data platform to model malaria transmission demonstrated that clinical and environmental variables are vital in enabling the presence and transmission of malaria [23]. The effectiveness of long short-term memory (LSTM) in prediction of infectious diseases have also been demonstrated in other research works by [24–30].

Studies have also proposed disease-specific models, Disease Informed Neural Networks (DINNs) was proposed for effective learning of dynamics of spread, progression forecast for infectious diseases to future and finding disease unique parameters such as death rate [31]. 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.

Boosted regression tree model effectively predicted malaria outbreak in Malaysia, however, the model must incorporate temporal dependence crucial for accurate malaria prediction [32]. Random Forest (RF) method demonstrated that the is one of the best models for accurately and effectively predicting West Nile virus cases, however, the model did not include prediction of variations of West Nile virus cases over different locations and times [33], therefore, including spatial-temporal features could lead to more accurate and effective predictions. Random Forest (RF) has also been found to perform better than Neural Network models in predicting Zika virus, though the model depends on the quantity of data available[34]. The proposed models have demonstrated more excellent performance in predicting infectious diseases with less capability in controlling dynamic and uncertain environments.

Machine learning and algorithms significantly predict infectious diseases and improve capability in handling diverse, large data sets and uncovering connections in the data [35]. However, machine learning algorithms must develop interconnections between vector, human, and environmental factors that drives disease presence. Therefore, suggestion for building a 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 [35].

2.1. Reinforced Learning Technique

Reinforced 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, where Single-Agent Reinforced Learning has been extensively applied [36–38]. Multi-Armed Bandit (MAB) technique was proposed to predict transmission of disease in aquatic [39], 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 [40]; 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 [37, 41]. Multiple studies have developed Multiple-Agent Reinforced Learning (MARL) to address this complexity [37].

Multiple agents have been applied in the field of transportation such as traffic light control [42, 43], Public Health and Medical Diagnosis [44, 45], manufacturing such as industrial robots [46], Portfolio management in financial trade [47], 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 [48, 49]. [50] 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. Twin Delayed Deep Deterministic Policy Gradient (TD3) technique proves highly effective in adaptive learning environments with significantly small overestimation bias in deep Q-learning [51]. 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. 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 [52]. 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 [53]. 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 [35]. 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 [54]. Studies have also shown that spatial information is significant in predicting spread and understanding the disease pattern [59]. 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 [60]. 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 [61]. GIS-machine learning model proposed to predict occurrence of infectious disease in Vietnam showed that neural networks perform better in predicting the infection [62]. An integrated GIS-based reinforced learning with Multi-Armed Bandit (MAB) algorithm has also been developed for efficient prediction of disease transmission [39]. 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, 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.

3. Methods

This section presents the basic concepts used to achieve the study objective. The proposed model for transmission of vector-borne infectious disease and algorithms discussed.

3.1. 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. Therefore, developing a dynamic model that capture dynamic movement of factors and the spatial nature of disease spread significantly improve transmission prediction of vector borne infectious disease and actions that reduce the spread of disease. A dynamic reinforced learning environment described by a virtual environment mimicking disease transmission was implemented. Virtual environment in Figure 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.

The model makes the following assumptions about the environment

1. Exist an infected vector in the environment.
2. Vectors and human move randomly in the environment.
3. The interaction between vector and human is involuntary.
4. On vector-human interaction, the human is infected with parasite with probability ≈ 0.3 [55].

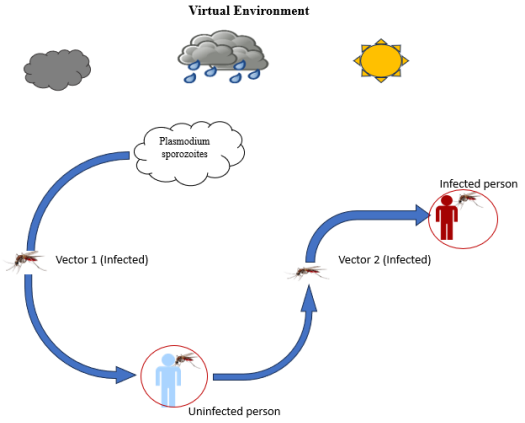

Figure 1. *Interaction between Vector and Human in the environment.*

Deriving the study model, the study build on the methodology developed by [39], where 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 form of TD3 algorithm that is utilized in a continuous control domain was used in the study [51]. Twin Deep Deterministic Policy Gradient is an actor-critic, off-policy model-free reinforcement learning described as

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

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 framework for Twin Deep Deterministic Policy Gradient is described as shown in the figure 2.

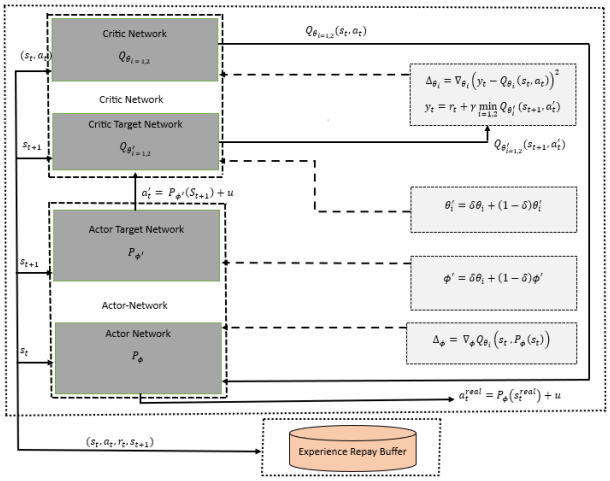

Figure 2. *Framework for Twin Deep Deterministic Policy Gradient Algorithm: Source [56].*

Spatial effect, economic factors (wealth, and education), temperature, rainfall that influence the reward (r_t) are introduced into the model equation (1)

$$R_t = r_t \times I_{(s_t)}|_{(\text{Max,Min})} \tag{2}$$

where $I_{(s_t)}|_{(\text{Max,Min})}$ is the malaria risk probability that embodies spatial and economic influence at state s_t, the reward R_t is define as a modified reward agent obtain after taking action a_t at state s_t and $I_{(s_t)}|_{(\text{Max,Min})}$ is given by

$$I_{(s_t)}|_{(\text{Max,Min})} = \left(1 + e^{(-(\mathbf{S}_{it}\boldsymbol{\beta} + u_i + \gamma_t + \epsilon_{it}))}\right)^{-1} \tag{3}$$

$S_i = (S_1, S_2, \ldots, S_n)'$ are $m \times 1$ vector of predictors measured at location i 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 introduce spatial dependence through dependence structure of **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 spatial-temporal unstructured random effect (model uncorrelated noise) in the model. Equation (1) is modified to

$$\begin{aligned} \Delta\theta_i &= \nabla\theta_i(y_t - Q_{\theta_i}(s_t, a_t))^2 \\ y_t &= R_t + \gamma \min_{i=1,2} Q_{\theta_i'}(s_{t+1}, a_t^{'}) \\ a_t^{'} &= P_{\phi'}(s_{t+1}) + u \end{aligned} \tag{4}$$

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

y_t is derived by concurrent learning of five Q-values $Q_{\theta_1'}, Q_{\theta_2'}, \ldots, Q_{\theta_5'}$ Min-Max ensemble estimation $\hat{Q}(s_{t+1}, a_t^{'}) = \max_{a_t' \in A} \min_{i \in N} Q_{\theta_i'}(s_{t+1}, a_t')$ to reduce the overestimation and underestimation bias conservatively to minimize bias during learning [57, 58].

$Q_{\theta_1^{\text{j}}}(s_{t+1}, a_t^{'})$ ensure that estimation is not overly conservative. The estimate retains the benefits of conservative estimation while effectively mitigating both overestimation and underestimation biases while maintaining stability and performance. The model equation is therefore given by

$$\hat{Q}(s_{t+1}, a_t') = \max\big(\min(Q_2, Q_3, Q_4),\ Q_5,\ Q_1\big) \tag{5}$$

where $Q_i = Q_{\theta_i'}(x)\ i = 1, \ldots, 5$ and $x = (s_{t+1}, a_t')$

$$\begin{aligned} \Delta\theta_i &= \nabla\theta_i(y_t - Q_{\theta_i}(s_t, a_t))^2 \qquad i = 1, \ldots, 5 \\ y_t &= R_t + \gamma \max_{a' \in A} \min_{i \in N} Q_{\theta_i'}(s_{t+1}, a_t') \\ a_t^{'} &= P_{\phi'}(s_{t+1}) + u \end{aligned} \tag{6}$$

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

$$\begin{aligned} \Delta\phi &= \nabla_\phi Q_{\theta_i}(s_t, P_\phi(s_t)) \quad i = 1, \ldots, 5 \\ \theta_i' &\leftarrow \delta\theta_i + (1-\delta)\theta_i' \\ \phi' &\leftarrow \delta\phi + (1-\delta)\phi' \end{aligned} \tag{7}$$

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.

3.2. Implementation of the Model

To implement the model (6), we created a dynamic environment where a human agent tries to avoid interaction with a vector agent. The vector agent moves randomly while interacting with the human agent. Rewards were based on successfully avoiding interaction between human and vector agents.

3.2.1. State Space

The human-vector distance was fixed as d_t randomly sampled at time t. A boolean variable "Interaction" was used as state variable s_t, where interaction is True when host-vector interaction happens and False otherwise (8).

$$\text{Interaction} = \begin{cases} \text{True}, & d_t = 0 \\ \text{False}, & d_t > 0 \end{cases} \tag{8}$$

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

$$d_t = \sqrt{(x_h - x_v)^2 + (y_h - y_v)^2} \tag{9}$$

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.

3.2.2. Action Space

The actions encompass treatment rate, spraying rate, 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, IT-nets, larval source management, and vaccination) to minimize interaction with the vector. The actions are selected to maximize expected reward, and the optimal action is selected during the learning process. Human agent actions representation at each time step is given by a vector of four continuous strategies

$$\text{action} = \{a_{\text{spray}}, a_{\text{IT-nets}}, a_{\text{Larval-Mgnt}}, a_{\text{vaccination}}\} \tag{10}$$

where $a_{\text{spray}} \in (0, 1), a_{\text{IT-nets}} \in (0, 1), a_{\text{larval-Mgnt}} \in (0, 1)$ and $a_{\text{vaccination}} \in (0, 1)$. The agent selects values for each of the four actions or get one optimal action that maximize the expected return.

3.2.3. 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. Therefore, the reward function (11) design consists of two representations corresponding to the two states; interaction with a vector and reaching target (avoiding interaction with the vector for 30 seconds).

$$\text{reward } (R_t) = \begin{cases} r_t \times I_{(s_t)}|_{(\text{Max})}, & d_t = 0 \\ r_t \times I_{(s_t)}|_{(\text{Min})}, & d_t > 0 \end{cases} \tag{11}$$

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 with no vector-interaction $(d_t > 0)$

3.3. Evaluation of Model Performance

3.3.1. 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

$$\text{Success Rate} = \frac{\text{No. of successful episode}}{\text{Number of episodes}} \times 100\% \tag{12}$$

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.

3.3.2. Infection Rate

The metric measures the rate of vector-human interaction in the 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

$$\text{Infection Rate} = \frac{\text{Number of Interactions}}{\text{Number of episodes}} \times 100\% \tag{13}$$

Lower infection rate indicate better policy performance, while a decrease in infection rate over time due to the model's learning actions demonstrates the effectiveness of the intervention.

4. Results and Discussion

4.1. Results

The model was applied to the scenario, and, as part of the evaluation, the following results from the proposed Bayesian RL algorithm are presented in Table 1. The table shows the performance metrics for each model, their success rate, and interaction rate. Various models we assessed for this evaluation include GIS RL (TD3), Multi-armed Bandit, and the proposed GIS Bayesian RL (5D3).

The statistical values for each algorithm's success rate in avoiding interaction for 30 seconds (equivalent to 300 episodes) can be used to calculate each algorithm's success ratio. The statistical values for interaction with vectors for each algorithm can be used to derive the infection ratio between each algorithm and the vectors. Therefore, the proposed GIS Bayesian RL (5D3) has a 96.2% success rate compared to the Multi-Armed Bandit algorithm in the same dynamic simulation environment, and the total navigation duration is 4.184 hours. Compared to the GIS RL(TD3) algorithm, the model improved the success rate by 3.4%; however, the GIS RL(TD3) algorithm has a shorter training navigation

time of 1.281 hours. The experimental outcome validates the improved performance of the proposed GIS Bayesian RL (5D3) method, which effectively solves the issues of weak efficiency and low success rate caused by overestimation and sample error. Based on the interaction rate, it can be deduced that the majority of unsuccessful episodes are due to interaction with vectors. Only a small number of episodes are deemed unsuccessful because the exploration step length exceeded the maximum value.

Table 1. *Performance metrics of various algorithmic choices for vector-borne disease transmission.*

Algorithm	Success Rate	Infection Rate	Model Training(Hrs)
Multi-Armed Bandit	0.926	0.072	1.281
GIS RL (TD3)	0.928	0.074	24.184
GIS Bayesian RL(5D3)	0.962	0.048	0.013

The results in Table 1 are also illustrated in figure 3 where all algorithms are accessed and their evaluation metrics presented.

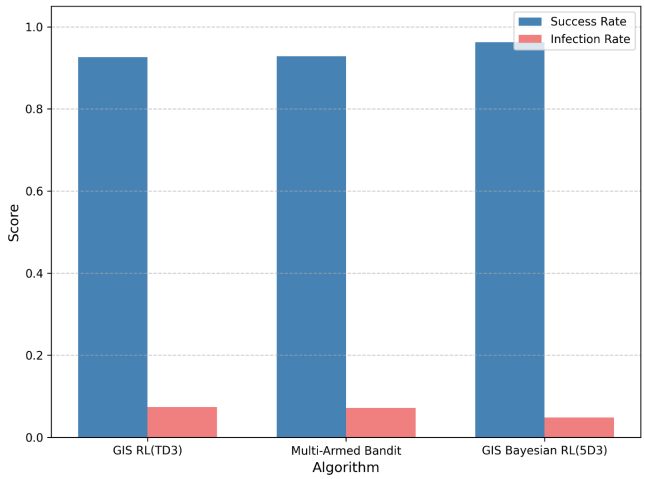

Figure 3. *Graphical representation of performance results for the algorithms.*

Figure 4 show the variation of reward values with variation of rounds when training human in the dynamic environments of each algorithm. The convergence of GIS RL(TD3) and Multi-Arm Bandit training is non-ideal, the algorithm has also erratic stability. Furthermore, the human have a relatively high interaction rate with the vector the subsequent stage of training through observing the training of the human. The algorithms has a slow convergence trend but it is not obvious, and the random exploration process of the algorithms is long, demonstrating the ability of human to finish 30 seconds under the algorithm grows slowly, which ultimately affects the convergence speed. The proposed GIS Bayesian RL (5D3) show an improved success rate and good convergence degree, and the convergence speed in the early stage is also faster, the reward value converges to 12 in the final.

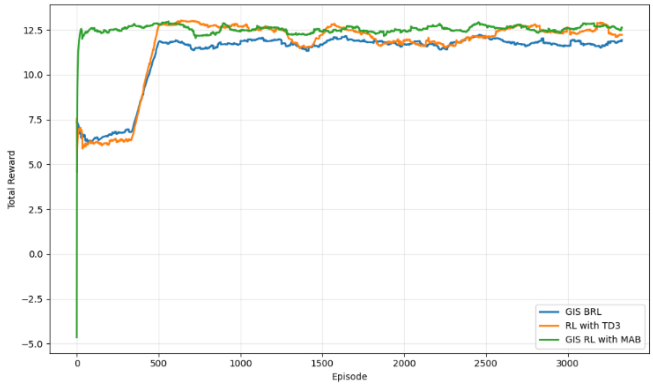

Figure 4. *Reward values of the algorithms.*

The figure 5 shows selection of actions/interventions by human to prevent the interaction with vector over 3600 episodes using the proposed model. The first 350 episodes illustrate action exploration by human, and as the training progresses between episode 350 to 900, spray repellent, insecticide treated nets, larval management, and vaccination stabilizes at an average of 90%. From episode 900 to 3600, human maintains average action rate of 90% for the actions spray repellent, insecticide treated nets, and larval management. Vaccination however fluctuates to an average of 15% between episode 900 to 3600 suggesting policy adjustment, this shows that human favors adoption of vaccination at 15% while maintaining high usage of insecticide treated nets, and larval management at an average rate of 90% with slightly fluctuating spray repellent.

The figure 6 shows selection of actions/interventions by human to prevent the interaction with vector over 3600 episodes using the TD3 model. The first 350 episodes illustrate action exploration by human, and as the training progresses between episode 350 to 1200, spray repellent, insecticide treated nets, larval management, and vaccination stabilizes at an average of 98%. From episode 1200 to 3600, human maintains average action rate of 98% for the actions spray repellent and insecticide treated nets, vaccination and larval management however fluctuates between 98% and 5% suggesting policy adjustment, this shows that human favors adoption of vaccination and larval management at 45% while maintaining high usage of insecticide treated nets, and spray repellent at an average rate of 98% with fluctuating rates of vaccination and larval management.

The figure 7 shows selection of actions/interventions by human to prevent the interaction with vector over 3600 episodes using Multi-armed Bandit model. From the first episode and as the training progresses between 0 to 3600 episodes spray repellent, insecticide treated nets, larval management stabilizes at approximately of 99%. Vaccination however fluctuates between 0-20%, this shows that human favors adoption of vaccination at approximately 10% while maintaining high usage of spray repellent, insecticide treated nets, and larval management at an average rate of 99%.

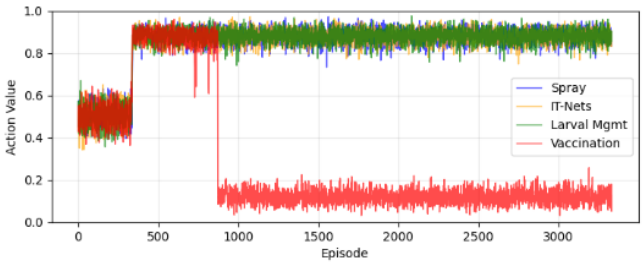

Figure 5. *Action selection by 5D3 algorithms.*

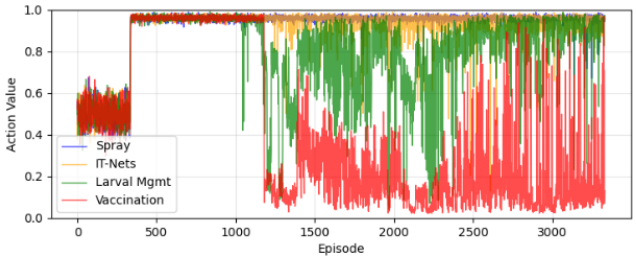

Figure 6. *Action selection by TD3 algorithms.*

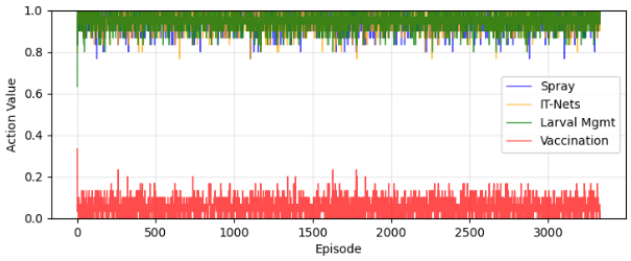

Figure 7. *Action selection by Multi-armed Bandit algorithms.*

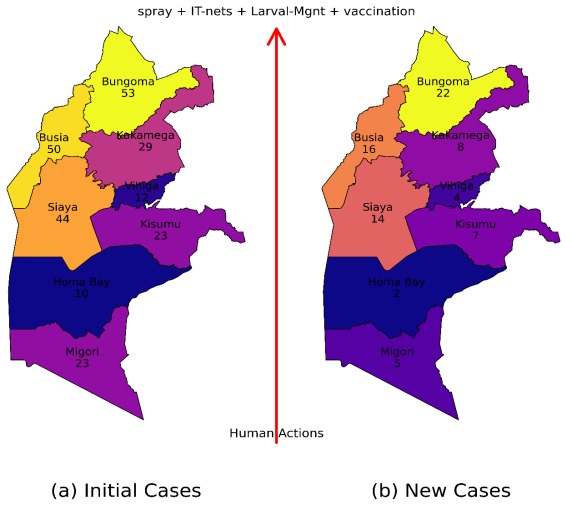

Figure 8. *Case Transmission in 8 Kenyan Regions after Human Actions.*

4.2. Spatial Spread of Cases Using the Proposed Model

Figure 8 shows spatial distribution of the cases for malaria cases in the western region of Kenya when the proposed model is applied to the data to estimate the number of cases when human agents select rates of actions to reduce the number of interactions with mosquitoes. From the figure it was observed that the number of cases reduces significantly with an average rate of 30% across all the regions when human adopts and average action rates of 90% for spray repellent, insecticide treated nets, and larval management, and 15% for vaccination.

4.3. Discussion

The empirical assessment of our proposed model provides a comprehensive understanding of disease transmission. The algorithmic assessment in Table 1 suggests that TD3 model outperforms the Multi-armed Bandit approach algorithm across all metrics, including success and infection rate. Based on performance, our proposed Bayesian RL 5D3 approach achieves a 96.2% success rate. These results reinforce the efficacy of reinforcement learning in modelling and predicting complex processes such as vector-borne disease transmission. In particular, its adaptive nature and its ability to make decisions based on iterative feedback may contribute to its superior performance. However, it is noteworthy that while the success rate was high across all models, relying solely on the success rate in disease modelling can be misleading. The consequences of no interaction between the agents in this context, erroneously predicting an area with low disease transmission, can be critical. Thus, precision, recall, and F1-score, which provide a more rounded view of performance, should be taken into account. For instance, the high recall value for the RL model suggests its ability to detect most interactions, a crucial aspect in epidemic modelling. Furthermore, the optimal action printed during program execution implies a crucial actionable insight. Determining whether the human agent should adopt preventive measures, such as spray repellents, insecticide-treated nets, larval management, and vaccination, is an immediate action that can be extracted from the model. Such outputs not only enhance the utility of the model for real-world application but also underscore its role in decision-making processes during vector borne outbreaks. The observation shown in Figure 8 is the spatial spread of disease cases. Low transmission across regions indicates potential controlled spread under stringent interventions or the presence of effective preventive measures. The findings offer insights into a situation that might be seen in a real-world context with interventions. The reduction in transmission presents an exciting scenario, with the disease affecting a smaller geographical area. Such an understanding is essential for policymakers and health authorities to anticipate and prepare for average vector-borne disease outbreaks and the average rate of interventions required to control their spread.

5. Conclusion

The research developed an advanced predictive system that models. The spread of vector-borne infectious diseases within the population. The flexibility of our model enables users to simulate disease spread across different regions, underscoring its practicality. The model facilitates selecting the actions the human should adopt to reduce the transmission of vector-borne disease and predicting the number of cases based on the human's adoption of the combined actions. By processing human action selection, our tool offers a holistic perspective on potential pathways for vector-borne disease transmission. This predictive prowess facilitates the selection of optimal intervention rates, thereby safeguarding the population's health and vitality. A key achievement of our study is the demonstrated efficacy of the GIS Bayesian 5D3 technique. Although the data indicate superior performance metrics compared to other methodologies, it is crucial to acknowledge that each algorithm has its own strengths and weaknesses, and that GIS Bayesian 5D3's superiority was context-specific. While our primary focus was vector-borne infectious disease in the population, the underlying principles of our tool possess potential applicability to other sectors. This adaptability might lead to innovative predictive models that can transform risk management strategies across different industries.

ORCID

0000-0003-2874-4266 (Kipngetich Gideon)

Abbreviations

RL	Reinforcement Learning
GIS	Geographic Information System
5D3	5-Delayed Deep Deterministic Policy Gradient
MAB	Multi-Armed Bandit
ODE	Ordinary Differential Equation

Author Contributions

Kipngetich Gideon: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing-original draft, Writing-review & editing

Victor Muthama Musau: Supervision, Writing - review & editing

Margaret Wambui Kinyua: Supervision, Writing - review & editing

Acknowledgments

The author would like to thank the School of Pure and Applied Sciences at Kirinyaga University for providing conducive environment for this research. Victor Muthama Musau and Margaret Wambui Kinyua for their constructive suggestions and comments.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] S. Bin, G. Sun, and C.-C. Chen, "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*, vol. 16, no. 23, p. 4683, 2019.

[2] P. J. White, "Mathematical models in infectious disease epidemiology," in *Infectious Diseases (Fourth Edition)*, pp. 49–53.e1, Elsevier, 2017.

[3] P. Zhang, K. Feng, Y. Gong, J. Lee, S. Lomonaco, and L. Zhao, "Usage of compartmental models in predicting covid-19 outbreaks," *The AAPS Journal*, vol. 24, no. 5, p. 98, 2022.

[4] M. Osman and I. Adu, "Simple mathematical model for malaria transmission," *Journal of Advances in Mathematics and Computer Science*, vol. 25, no. 6, pp. 1–24, 2017.

[5] E. Dantas, M. Tosin, and A. Cunha Jr, "Calibration of a seir–sei epidemic model to describe the zika virus outbreak in brazil," *Applied Mathematics and Computation*, vol. 338, pp. 249–259, 2018.

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

[7] R. Anokye, E. Acheampong, I. Owusu, and E. Isaac Obeng, "Time series analysis of malaria in kumasi: Using arima models to forecast future incidence," *Cogent social sciences*, vol. 4, no. 1, p. 1461544, 2018.

[8] R. Mopuri, S. G. Kakarla, S. R. Mutheneni, M. R. Kadiri, and S. Kumaraswamy, "Climate based malaria forecasting system for andhra pradesh, india," *Journal of Parasitic Diseases*, vol. 44, pp. 497–510, 2020.

[9] A. M. Adeola, J. O. Botai, J. M. Olwoch, H. C. J. d. W. Rautenbach, O. M. Adisa, C. De Jager, C. M. Botai, and M. Aaron, "Predicting malaria cases using remotely sensed environmental variables in nkomazi, south africa," *Geospatial health*, vol. 14, no. 1, 2019.

[10] O. Stojanovic, J. Leugering, G. Pipa, S. Ghozzi, and A. Ullrich, "A bayesian monte carlo approach for predicting the spread of infectious diseases," *PloS one*, vol. 14, no. 12, p. e0225838, 2019.

[11] S. Funk, A. Camacho, A. J. Kucharski, R. M. Eggo, and W. J. Edmunds, "Real-time forecasting of infectious disease dynamics with a stochastic semi-mechanistic model," *Epidemics*, vol. 22, pp. 56–61, 2018.

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

[13] A. P. Galvani and R. M. May, "Dimensions of superspreading," *Nature*, vol. 438, no. 7066, pp. 293–295, 2005.

[14] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri, "Modelling the covid-19 epidemic and implementation of population-wide interventions in italy," *Nature medicine*, vol. 26, no. 6, pp. 855–860, 2020.

[15] E. Hunter, B. Mac Namee, and J. D. Kelleher, "A model for the spread of infectious diseases in a region," *International journal of environmental research and public health*, vol. 17, no. 9, p. 3119, 2020.

[16] O. E. Santangelo, V. Gentile, S. Pizzo, D. Giordano, and F. Cedrone, "Machine learning and prediction of infectious diseases: a systematic review," *Machine Learning and Knowledge Extraction*, vol. 5, no. 1, pp. 175–198, 2023.

[17] Y. Shi, X. Liu, S.-Y. Kok, J. Rajarethinam, S. Liang, G. Yap, C.-S. Chong, K.-S. Lee, S. S. Y. Tan, C. K. Y. Chin, *et al.*, "Three-month real-time dengue forecast models: an early warning system for outbreak alerts and policy decision support in singapore," *Environmental health perspectives*, vol. 124, no. 9, pp. 1369–1375, 2016.

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

[19] C. M. Benedum, K. M. Shea, H. E. Jenkins, L. Y. Kim, and N. Markuzon, "Weekly dengue forecasts in iquitos, peru; san juan, puerto rico; and singapore," *Plos Neglected tropical diseases*, vol. 14, no. 10, p. e0008710, 2020.

[20] A. A. Kolesnikov, P. M. Kikin, and A. M. Portnov, "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*, vol. 42, pp. 221–226, 2019.

[21] V.-H. Nguyen, T. T. Tuyet-Hanh, J. Mulhall, H. V. Minh, T. Q. Duong, N. V. Chien, N. T. T. Nhung, V. H. Lan, H. B. Minh, D. Cuong, *et al.*, "Deep learning models for forecasting dengue fever based on climate data in vietnam," *PLoS Neglected Tropical Diseases*, vol. 16, no. 6, p. e0010509, 2022.

[22] S. Chae, S. Kwon, and D. Lee, "Predicting infectious disease using deep learning and big data," *International journal of environmental research and public health*, vol. 15, no. 8, p. 1596, 2018.

[23] T. Santosh, D. Ramesh, and D. Reddy, "Lstm based prediction of malaria abundances using big data," *Computers in Biology and Medicine*, vol. 124, p. 103859, 2020.

[24] V. K. R. Chimmula and L. Zhang, "Time series forecasting of covid-19 transmission in canada using lstm networks," *Chaos, solitons & fractals*, vol. 135, p. 109864, 2020.

[25] Y. Guo, Y. Feng, F. Qu, L. Zhang, B. Yan, and J. Lv, "Prediction of hepatitis e using machine learning models," *Plos one*, vol. 15, no. 9, p. e0237750, 2020.

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

[27] A. Adiga, L. Wang, B. Hurt, A. Peddireddy, P. Porebski, S. Venkatramanan, B. L. Lewis, and M. Marathe, "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*, pp. 2505–2513, 2021.

[28] T. Goo, C. Apio, G. Heo, D. Lee, J. H. Lee, J. Lim, K. Han, and T. Park, "Forecasting of the covid-19 pandemic situation of korea," *Genomics & Informatics*, vol. 19, no. 1, 2021.

[29] S. L. Kumar *et al.*, "Predictive analytics of covid-19 pandemic: statistical modelling perspective," *Walailak Journal of Science and Technology (WJST)*, vol. 18, no. 16, pp. 15583–14, 2021.

[30] L. Shen, C. Jiang, M. Sun, X. Qiu, J. Qian, S. Song, Q. Hu, H. Yelixiati, and K. Liu, "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*, vol. 2022, 2022.

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

[32] P. M. Brock, K. M. Fornace, M. J. Grigg, N. M. Anstey, T. William, J. Cox, C. J. Drakeley, H. M. Ferguson, and R. R. Kao, "Predictive analysis across spatial scales links zoonotic malaria to deforestation," *Proceedings of the Royal Society B*, vol. 286, no. 1894, p. 20182351, 2019.

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

[34] K. Roster, C. Connaughton, and F. A. Rodrigues, "Forecasting new diseases in low-data settings using transfer learning," *Chaos, Solitons & Fractals*, vol. 161, p. 112306, 2022.

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

[36] Y. Li, "Deep reinforcement learning: An overview," *arXiv preprint arXiv: 1701.07274*, 2017.

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

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

[39] A. Karras, C. Karras, S. Sioutas, C. Makris, G. Katselis, I. Hatzilygeroudis, J. A. Theodorou, and D. Tsolis, "An integrated gis-based reinforcement learning approach for efficient prediction of disease transmission in aquaculture," *Information*, vol. 14, no. 11, p. 583, 2023.

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

[41] L. Bucsoniu, R. Babuvska, and B. De Schutter, "Multi-agent reinforcement learning: An overview," *Innovations in multi-agent systems and applications-1*, pp. 183–221, 2010.

[42] L. Li, Y. Lv, and F.-Y. Wang, "Traffic signal timing via deep reinforcement learning," *IEEE/CAA Journal of Automatica Sinica*, vol. 3, no. 3, pp. 247–254, 2016.

[43] Y. Zhu, Z. He, and G. Li, "A bi-hierarchical game-theoretic approach for network-wide traffic signal control using trip-based data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15408–15419, 2022.

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

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

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

[47] U. Pham, Q. Luu, and H. Tran, "Multi-agent reinforcement learning approach for hedging portfolio problem," *Soft Computing*, vol. 25, pp. 7877–7885, 2021.

[48] J. Hu, M. P. Wellman, *et al.*, "Multiagent reinforcement learning: theoretical framework and an algorithm.," in *ICML*, vol. 98, pp. 242–250, 1998.

[49] M. Bowling and M. Veloso, *An analysis of stochastic game theory for multiagent reinforcement learning*. Citeseer, 2000.

[50] A. Tampuu, T. Matiisen, D. Kodelja, I. Kuzovkin, K. Korjus, J. Aru, J. Aru, and R. Vicente, "Multiagent cooperation and competition with deep reinforcement learning," *PloS one*, vol. 12, no. 4, p. e0172395, 2017.

[51] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *International conference on machine learning*, pp. 1587–1596, PMLR, 2018.

[52] Esri, "What is gis? — geographic information system (gis)," 2024.

[53] E. K. Cromley, "Gis and disease," *Annual review of public health*, vol. 24, no. 1, pp. 7–24, 2003.

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

[55] T. S. Churcher, R. E. Sinden, N. J. Edwards, I. D. Poulton, T. W. Rampling, P. M. Brock, J. T. Griffin, L. M. Upton, S. E. Zakutansky, K. A. Sala, *et al.*, "Probability of transmission of malaria from mosquito to human is regulated by mosquito parasite density in naive and vaccinated hosts," *PLoS pathogens*, vol. 13, no. 1, p. e1006108, 2017.

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

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

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

[59] G. Chowell and R. Rothenberg, "Spatial infectious disease epidemiology: on the cusp," 2018.

[60] S. Jung, J. Jeoung, H. Kang, and T. Hong, "Optimal planning of a rooftop pv system using gis-based reinforcement learning," *Applied Energy*, vol. 298, p. 117239, 2021.

[61] C. Bone, "Integrating complexity science and artificial intelligence: Gis, agents and reinforcement learning for modeling forest cover change," 2009.

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