



## OPEN Analysis of a mathematical model for malaria using data-driven approach

Adithya Rajnarayanan<sup>1</sup>, Manoj Kumar<sup>1</sup> & Abdessamad Tridane<sup>2✉</sup>

Malaria remains one of the leading causes of global morbidity and mortality, with millions of cases and fatalities annually. Effective intervention strategies by public health authorities and medical practitioners necessitate a robust understanding of disease transmission dynamics. This study presents a novel framework for modeling malaria transmission dynamics by integrating temperature and altitude-dependent transmission functions into a compartmental SIR-SI model. A key innovation lies in the introduction of a new transmission function that explicitly captures environmental dependencies, enhancing realism in the modeling of disease spread. We conduct steady-state analysis of the system, establishing the stability criteria for both disease-free and endemic equilibria through linearization techniques. We used a novel transmission function to model the dependence on temperature and altitude. To address the challenge of accurate parameter estimation, we develop a comparative learning framework using ANNs, RNNs, and PINNs, with PINNs standing out by embedding epidemiological dynamics into the training process. This enables physics-constrained parameter inference, significantly enhancing predictive performance over purely data-driven approaches. Additionally, we implement Dynamic Mode Decomposition (DMD) to derive a data-driven transmission risk index from infection trajectory data, providing a novel and interpretable metric for real-time risk assessment.

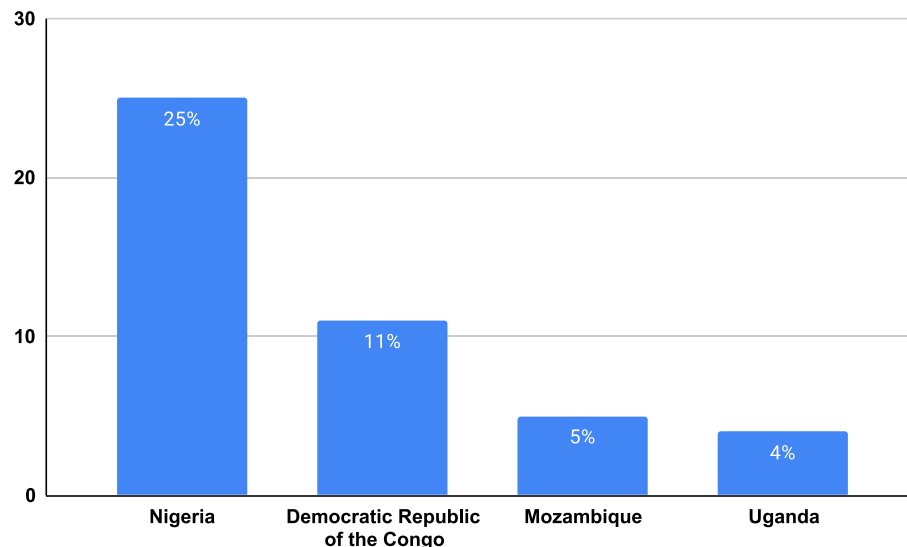
**Keywords** Malaria model, Compartmental model, Data-driven methods, Neural network, Dynamic mode decomposition

According to the 2022 report by the World Health Organization (WHO), Africa remains the region most severely affected by malaria globally. Notably, 94% of the world's 233 million malaria cases and 95% of global malaria-related fatalities (approximately 580,000 deaths) occur on the African continent (Fig. 1). These figures underscore malaria's disproportionate mortality burden in Africa, emphasizing the critical need for advanced modeling frameworks to enhance healthcare professionals' understanding of disease transmission dynamics.

In epidemiological research, compartmental models based on ordinary differential equations (ODEs) are widely used to analyze disease spread. These models partition populations into distinct, mutually exclusive compartments, enabling the mathematical study of transitions between health states (e.g., infection, recovery, vaccination, or death). Commonly employed frameworks include the SIR (Susceptible-Infected-Recovered), SIRD (Susceptible-Infected-Recovered-Deceased), and SIRDV (Susceptible-Infected-Recovered-Deceased-Vaccinated) models, which simulate population flows through ODE-derived rates of change. Early mathematical models of malaria laid the foundation for understanding transmission dynamics. Ross introduced a two-compartment model capturing basic interactions between infected humans and mosquitoes but omitted the parasite's latent period (see<sup>1–4</sup>). MacDonald<sup>5</sup> later incorporated the mosquito latent phase, while Anderson refined the framework by modeling both human and mosquito susceptibility and infection explicitly. For detailed developments and comparisons of these models, see<sup>6</sup>.

In the study by Ogueta et al.<sup>7</sup>, a variant of the physics-informed neural network (PINN), termed the disease-informed neural network (DINN), was employed as a deep learning model. The SIRD compartmental model was utilized to analyze disease dynamics, incorporating the movement of individuals between cities. The primary objective of this work was to predict various parameters, including the rates of transmission, mortality, and recovery for the selected cities, as well as the rate of movement of individuals between them. Several studies have combined compartmental epidemic models with physics-informed and deep learning approaches to model COVID-19 dynamics. These include applications of PINNs to SIR-type models for estimating transmission and

<sup>1</sup>School of Engineering and Science, Indian Institute of Technology Madras Zanzibar, PO Box 394, Bweleo, Zanzibar, Urban West 71215, Tanzania. <sup>2</sup>Department of Mathematical Sciences, United Arab Emirates University, P.O. Box 15551, Al Ain, UAE. ✉email: a-tridane@uae.ac.ae



**Fig. 1.** In 2017, four countries from Africa accounted for 45% of all malaria cases worldwide.

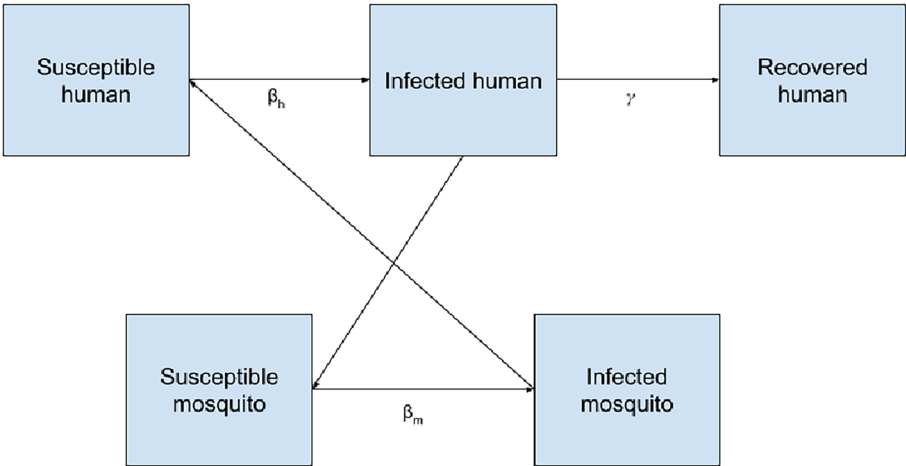
recovery rates<sup>8–10</sup>, and the integration of LSTM and DNN with extended compartmental frameworks to forecast short- and medium-term trajectories across different regions<sup>11–14</sup>. For a broader overview of related statistical and deep learning models, see<sup>15–25</sup>. Several studies have also investigated stochastic models for infectious diseases; for further details, we refer to<sup>26,27</sup>.

In the study by Bhujji et al.<sup>28</sup>, the temperature dependence of the transmission rate was analyzed using the SEIR model for humans and the LSEI model for mosquitoes. The authors conducted various mathematical analyses, including the stability of the disease-free equilibrium and the existence of the endemic equilibrium. Numerical simulations across different temperature scenarios revealed that temperature significantly influences the transmission rate. Keno et al.<sup>29</sup> examined the temperature dependency of the transmission parameter using the SIR model for humans and the SI model for mosquitoes. Their analysis includes assessments of both local and global stability of equilibrium points. The study demonstrated that when the basic reproduction number is less than one, the disease-free equilibrium is both locally and globally asymptotically stable. Additionally, the impact of temperature on transmission dynamics was investigated, reinforcing the conclusion that temperature plays a critical role in disease transmission.

In the study by Proctor et al.<sup>30</sup>, Dynamic Mode Decomposition (DMD) was utilized to incorporate control effects and extract low-order models from high-dimensional, complex systems. Alla and Kutz<sup>31</sup> implemented DMD to reduce the order of a nonlinear dynamical system. Similarly, Andreuzzi et al.<sup>32</sup> extended DMD for forecasting future states of parametric dynamical systems. Watson et al.<sup>33</sup> employed a Bayesian time series model in conjunction with random forests to predict the number of cases and deaths using the SIRD compartmental model, conducting a 21-day forecast for three cities: New York, Colorado, and West Virginia. Additional research on mathematical models of malaria and dengue can be found in<sup>34–43</sup>. Most existing studies on malaria transmission dynamics primarily rely on mathematical modeling. In contrast, this work leverages deep learning methods to analyze the dynamics of malaria transmission. One key advantage of using the neural network approach is that these models are designed to emulate the human brain, allowing them to capture complex patterns in data. This capability makes deep learning particularly well-suited for modeling malaria dynamics.

This study employs a feedforward Artificial Neural Network (ANN) to predict trajectories across all five compartments of the epidemiological model. To infer malaria transmission parameters, we implement a comparative analysis using three machine learning architectures: ANNs, Recurrent Neural Networks (RNNs), and Physics-Informed Neural Networks (PINNs). RNNs are uniquely suited for this task due to their capacity to capture temporal dependencies in sequential data, enabling robust forecasting by leveraging historical trends. PINNs further enhance parameter estimation by integrating domain-specific physical laws—derived from the governing equations of the SIR-SI system—directly into the neural network's loss function, ensuring biologically plausible outputs. For risk quantification, Dynamic Mode Decomposition (DMD) is applied to infected population trajectories; unlike conventional deep learning methods, DMD operates as a data-driven modal decomposition technique, extracting dominant spatial-temporal patterns to characterize transmission risks without requiring a priori mechanistic assumptions. Central to this analysis is the multivariate environmental dependence of the transmission rate, which is rigorously evaluated through the concurrent effects of temperature and altitude—a critical advancement over prior univariate approaches.

The remainder of this work is structured as follows: Section [Model formulation](#) introduces the mathematical formulation of the compartmental epidemiological model, incorporating environmental dependencies such as temperature and altitude into the transmission dynamics. Section [Data-Driven Methods](#) details the analytical and computational framework, beginning with the problem statement, followed by methodology encompassing steady-state stability analysis, parameter estimation using artificial neural networks (ANNs), recurrent neural networks (RNNs), and physics-informed neural networks (PINNs), and concluding with results evaluating



**Fig. 2.** Schematic diagram of SIR-SI system.

| Compartment Symbol | Explanation                                                                                  |
|--------------------|----------------------------------------------------------------------------------------------|
| $S_h$              | It represents the section of the human population who are susceptible to malaria             |
| $I_h$              | It represents the section of humans who are infected with malaria                            |
| $R_h$              | It represents the section of humans who have recovered from malaria                          |
| $S_m$              | It represents the section of mosquitoes that are susceptible to the malaria-causing parasite |
| $I_m$              | It represents the section of mosquitoes that are infected with the malaria-causing parasite  |

**Table 1.** Interpretation of compartments.

trajectory predictions and risk quantification via dynamic mode decomposition (DMD). Finally, Section [Concluding remarks](#) synthesizes the findings, discusses their implications for malaria mitigation strategies, and proposes future research directions to enhance model generalizability and real-world applicability.

**Model formulation**

This section presents a compartmental epidemiological model to capture the interdependent dynamics of malaria transmission between human and mosquito populations, a necessity given the disease’s reliance on cross-species interaction. The human population is divided into three compartments: susceptible ( $S_h$ ), infected ( $I_h$ ), and recovered ( $R_h$ ). In parallel, the mosquito population is categorized into susceptible ( $S_m$ ) and infected ( $I_m$ ) compartments. The model accounts for bidirectional transmission mechanisms: infected mosquitoes ( $I_m$ ) transmit the parasite to susceptible humans ( $S_h$ ), increasing the  $I_h$  population, while infected humans ( $I_h$ ) subsequently infect susceptible mosquitoes ( $S_m$ ), driving the rise of  $I_m$ . These interactions are governed by density-dependent transmission rates, reflecting real-world contact patterns. The model assumes constant birth and death rates, homogeneous mixing of the population, and no demographic or spatial heterogeneity. Figure 2 schematically represents the coupled transmission pathways, emphasizing the feedback loop central to malaria’s persistence.

For the human population, we consider the standard SIR model, and for the mosquito population, we consider the SI model. The following is the system of differential equations.

$$\begin{aligned} \frac{dS_h}{dt} &= \Gamma_h N_h - \frac{\beta_h S_h I_m}{N_h} - \mu_h S_h \\ \frac{dI_h}{dt} &= \frac{\beta_h S_h I_m}{N_h} - (\gamma_h + \mu_h) I_h \\ \frac{dR_h}{dt} &= \gamma_h I_h - \mu_h R_h \\ \frac{dS_m}{dt} &= \Gamma_m N_m - \frac{\beta_m S_m I_h}{N_m} - \mu_m S_m \\ \frac{dI_m}{dt} &= \frac{\beta_m S_m I_h}{N_m} - \mu_m I_m. \end{aligned} \tag{1}$$

The model used in this work is the SIR-SI model. The SIR model is used for the human population, and the SI model is used for the mosquito population. Table 1 and 2 explain the compartments and the parameters used.

| Compartment Symbol | Explanation                             |
|--------------------|-----------------------------------------|
| $\Gamma_h$         | Birth rate of the humans                |
| $N_h$              | Total human population                  |
| $\beta_h$          | Transmission rate of malaria in humans  |
| $\mu_h$            | Mortality rate of humans due to malaria |
| $\gamma_h$         | Recovery rate of humans                 |
| $\Gamma_m$         | Birth rate of mosquitoes                |
| $\beta_m$          | Transmission of malaria in mosquitoes   |
| $\mu_m$            | Mortality rate of mosquitoes            |
| $N_m$              | Total population of mosquitoes          |

**Table 2.** Parameters description.

### Disease-free steady state analysis

In this section, we perform the steady-state analysis. Steady-state solutions play an important role when the analytical solution is not known, and we want to study the qualitative properties of solutions.

We define the basic reproduction number for the model (1) as

$$R_0 = \frac{\beta_h \beta_m \Gamma_h \Gamma_m}{\mu_h \mu_m^2 (\mu_h + \gamma_h)}.$$

Observe that the basic reproduction number depends on recruitment rates, infection rates, recovery rates, and mortality rates.

**Theorem 1** *If  $R_0 < 1$ , the disease-free steady state is locally stable.*

*Proof* For the analysis of the disease-free steady state, we need to equate the infected and recovered populations of both species to zero. Therefore, we obtain  $S_h = \frac{\Gamma_h}{\mu_h}$ ,  $I_h = 0$ ,  $R_h = 0$ ,  $S_m = \frac{\Gamma_m}{\mu_m}$ ,  $I_m = 0$ .

Dividing the first three equations  $N_h$  and the last two equations with  $N_m$  in (1), we get

$$\begin{aligned} \frac{dS_h}{dt} &= \Gamma_h - m\beta_h S_h I_m - \mu_h S_h \\ \frac{dI_h}{dt} &= m\beta_h S_h I_m - (\gamma_h + \mu_h) I_h \\ \frac{dR_h}{dt} &= \gamma_h I_h - \mu_h R_h \\ \frac{dS_m}{dt} &= \Gamma_m - \frac{\beta_m S_m I_h}{m} - \mu_m S_m \\ \frac{dI_m}{dt} &= \frac{\beta_m S_m I_h}{m} - \mu_m I_m. \end{aligned} \quad (2)$$

Now our task is to linearize the system (2) around the disease-free steady state. After linearization, we obtain the following system

$$\begin{aligned} \frac{dS_h}{dt} &= \Gamma_h - m\beta_h \frac{\Gamma_h}{\mu_h} I_m - \mu_h S_h \\ \frac{dI_h}{dt} &= m\beta_h \frac{\Gamma_h}{\mu_h} I_m - (\gamma_h + \mu_h) I_h \\ \frac{dR_h}{dt} &= \gamma_h I_h - \mu_h R_h \\ \frac{dS_m}{dt} &= \Gamma_m - \beta_m \frac{\Gamma_m}{m\mu_m} I_h - \mu_m S_m \\ \frac{dI_m}{dt} &= \frac{\Gamma_m \beta_m I_h}{m\mu_m} - \mu_m I_m. \end{aligned}$$

Constructing the Jacobian matrix, we get

$$J = \begin{bmatrix} -\mu_h & 0 & 0 & 0 & -\frac{m\beta_h\Gamma_h}{\mu_h} \\ 0 & -(\gamma_h + \mu_h) & 0 & 0 & \frac{m\beta_h\Gamma_h}{\mu_h} \\ 0 & \gamma_h & -\mu_h & 0 & 0 \\ 0 & -\frac{\beta_m\Gamma_m}{m\mu_m} & 0 & -\mu_m & 0 \\ 0 & \frac{\beta_m\Gamma_m}{m\mu_m} & 0 & 0 & -\mu_m \end{bmatrix}.$$

The characteristic polynomial of the above matrix is

$$f(\lambda) = -(\lambda + \mu_h)(\lambda + \mu_h)(\lambda + \mu_m) \left[ (\lambda + \mu_m)(\lambda + \mu_h + \gamma_h) - \frac{\beta_h\beta_m\Gamma_h\Gamma_m}{\mu_m\mu_h} \right].$$

For the system to be stable, all the eigenvalues must have negative real parts. It is clear that three of the eigenvalues are negative; we need to check the roots of the quadratic polynomial for the remaining two eigenvalues. By solving the quadratic equation, we get the roots as

$$\begin{aligned} & \frac{-(\gamma + \mu_h + \mu_m) + \sqrt{(\mu_m + \mu_h + \gamma_h)^2 - 4(\mu_m(\mu_h + \gamma_h) - \frac{\beta_h\beta_m\Gamma_h\Gamma_m}{\mu_h\mu_m})}}{2} \\ & \frac{-(\gamma + \mu_h + \mu_m) - \sqrt{(\mu_m + \mu_h + \gamma_h)^2 - 4(\mu_m(\mu_h + \gamma_h) - \frac{\beta_h\beta_m\Gamma_h\Gamma_m}{\mu_h\mu_m})}}{2}. \end{aligned}$$

If we observe, the second root is always negative, and thus we need to find out the condition for which the first root is negative, and that is

$$\mu_m(\mu_h + \gamma_h) - \frac{\beta_h\beta_m\Gamma_h\Gamma_m}{\mu_h\mu_m} > 0.$$

Rearranging the above inequality, we get the expression

$$\frac{\beta_h\beta_m\Gamma_h\Gamma_m}{\mu_h\mu_m^2(\mu_h + \gamma_h)} < 1.$$

So, if  $R_0 < 1$ , it will ensure us that the disease-free steady state is stable.  $\square$

### Endemic steady state analysis

In this subsection, we aim to study the stability of the endemic steady-state.

The given system of equations is

$$\begin{aligned} \frac{dS_h}{dt} &= \Gamma_h - m\beta_h \frac{\Gamma_h}{\mu_h} I_m - \mu_h S_h \\ \frac{dI_h}{dt} &= m\beta_h \frac{\Gamma_h}{\mu_h} I_m - (\gamma_h + \mu_h) I_h \\ \frac{dR_h}{dt} &= \gamma_h I_h - \mu_h R_h \\ \frac{dS_m}{dt} &= \Gamma_m - \beta_m \frac{\Gamma_m}{m\mu_m} I_h - \mu_m S_m \\ \frac{dI_m}{dt} &= \frac{\Gamma_m\beta_m I_h}{m\mu_m} - \mu_m I_m. \end{aligned}$$

Before proceeding with the next theorem, it is essential to define the following quantities

$$\begin{aligned} b &= \mu_h + \frac{\beta_h I_m^*}{N_h} + \mu_h + \gamma + \frac{\beta_h I_h^*}{N_m} + \mu_m \\ c &= (\mu_h + \frac{\beta_h I_m^*}{N_h})(\mu_h + \gamma) + (\mu_h + \gamma)(\frac{\beta_h I_h^*}{N_m} + \mu_m) + (\frac{\beta_h I_h^*}{N_m} + \mu_m)(\mu_h + \frac{\beta_h I_m^*}{N_h}) - \frac{\beta_m\beta_h S_m^* S_h^*}{N_m N_h} \\ d &= \left( \mu_h + \frac{\beta_h I_m^*}{N_h} \right) \left[ (\mu_h + \gamma)(\frac{\beta_h I_h^*}{N_m} + \mu_m) - \frac{\beta_m\beta_h S_m^* S_h^*}{N_m N_h} + \frac{\beta_m\beta_h^2 S_m^* S_h^* I_m^*}{N_m N_h^2} \right], \end{aligned}$$

where  $S_m^*, S_h^*, I_m^*, I_h^*$  are the non trivial equilibrium solutions.

**Theorem 2** If  $R_0 > 1$  and  $b, c, d, bc - d > 0$ , the endemic steady state is locally stable.

*Proof* The equilibrium points will be obtained by equating all of the above time derivatives to zero, and solving them for non-trivial solutions, we obtain the solutions as

$$\begin{aligned}
 S_h^* &= \frac{N_h(\gamma + \mu_h)(\mu_m N_m + \frac{\Gamma_h N_h \beta_m}{\gamma + \mu_h})}{\beta_m(\mu_h N_h + \frac{\Gamma_m N_m \beta_h}{\mu_m})} \\
 I_h^* &= \frac{N_h N_m \mu_m \mu_h (R_0 - 1)}{\beta_m(\mu_h N_h + \frac{\Gamma_m N_m \beta_h}{\mu_m})} \\
 R_h^* &= \frac{N_h N_m \mu_m \gamma_h (R_0 - 1)}{\beta_m(\mu_h N_h + \frac{\Gamma_m N_m \beta_h}{\mu_m})} \\
 S_m^* &= \frac{N_m \mu_m (\mu_h N_h + \frac{\Gamma_m N_m \beta_h}{\mu_m})}{\beta_h(\mu_m N_m + \frac{\Gamma_h N_h \beta_m}{\gamma + \mu_h})} \\
 I_m^* &= \frac{N_h N_m \mu_m \mu_h (R_0 - 1)}{\beta_h(\mu_m N_m + \frac{\Gamma_h N_h \beta_m}{\gamma + \mu_h})}.
 \end{aligned}$$

Linearizing the model (2) around the endemic equilibrium point, we get the following system

$$\begin{aligned}
 \frac{dS_h}{dt} &= \Gamma_h N_h - \frac{\beta_h}{N_h} (S_h^* I_m + I_m^* S_h - S_h^* I_m^*) - \mu_h S_h \\
 \frac{dI_h}{dt} &= \frac{\beta_h}{N_h} (S_h^* I_m + I_m^* S_h - S_h^* I_m^*) - \gamma I_h - \mu_h I_h \\
 \frac{dR_h}{dt} &= \gamma I_h - \mu_h R_h \\
 \frac{dS_m}{dt} &= \gamma_m N_m - \frac{\beta_m}{N_m} (S_m^* I_h + S_m I_h^*) - \mu_m S_m \\
 \frac{dI_m}{dt} &= \frac{\beta_m}{N_m} (S_m^* I_h + S_m I_h^*) - \mu_m I_m.
 \end{aligned}$$

Writing the Jacobian matrix, we get

$$J = \begin{bmatrix} -\frac{\beta_h I_m^*}{N_h} - \mu_h & 0 & 0 & 0 & -\frac{\beta_h S_h^*}{N_h} \\ \frac{\beta_h I_m^*}{N_h} & -(\gamma_h + \mu_h) & 0 & 0 & \frac{\beta_h S_h^*}{N_h} \\ 0 & \gamma_h & -\mu_h & 0 & 0 \\ 0 & -\frac{\beta_h S_m^*}{N_m} & 0 & -\frac{\beta_h S_m^*}{N_m} - \mu_m & 0 \\ 0 & \frac{\beta_h S_m^*}{N_m} & 0 & \frac{\beta_h S_m^*}{N_m} & -\mu_m \end{bmatrix}$$

The characteristic polynomial of the above matrix is

$$f(\lambda) = -(\lambda + \mu_h)(\lambda + \mu_m) \left( \left[ \lambda + \mu_h + \frac{\beta_h I_m^*}{N_h} \right] \left[ (\lambda + \mu_h + \gamma) \left( \lambda + \frac{\beta_m I_h^*}{N_m} + \mu_m \right) - \frac{\beta_m \beta_h S_m^* S_h^*}{N_h N_m} \right] + \frac{\beta_h^2 \beta_m S_h^* I_m^* S_m^*}{N_h^2 N_m} \right).$$

From the above equation, we can see that we have two linear factors and a cubic factor. To analyze the cubic factor, let us state the following lemma:

**Lemma 1** Let  $f(x) = ax^3 + bx^2 + cx + d$  be a cubic polynomial. For  $f(x)$  to have all negative roots or complex roots with negative real parts, the following conditions are necessary:

$$a > 0, b > 0, c > 0, d > 0, bc - ad > 0.$$

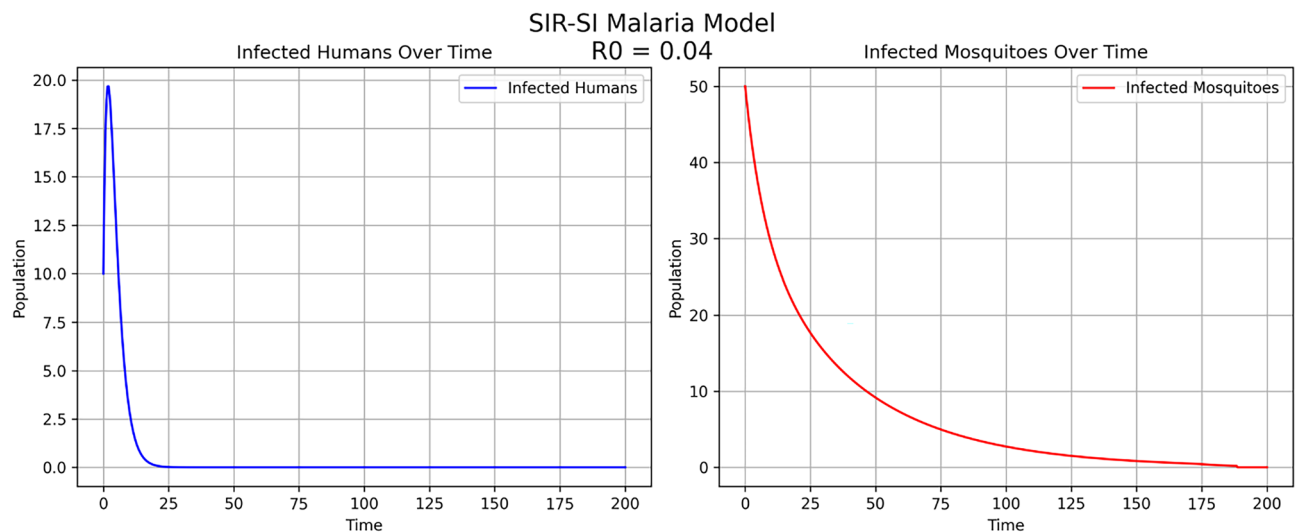
Now, by using the above-stated lemma, we can arrive at our required result.  $\square$

**Remark 1** Equilibrium points are crucial for evaluating malaria control efforts. A disease-free equilibrium indicates that transmission can be halted through vector control and drug administration, while an endemic equilibrium suggests sustained transmission, requiring ongoing interventions. Stability analysis helps identify critical thresholds, such as intervention coverage or mosquito density, to guide effective control strategies.

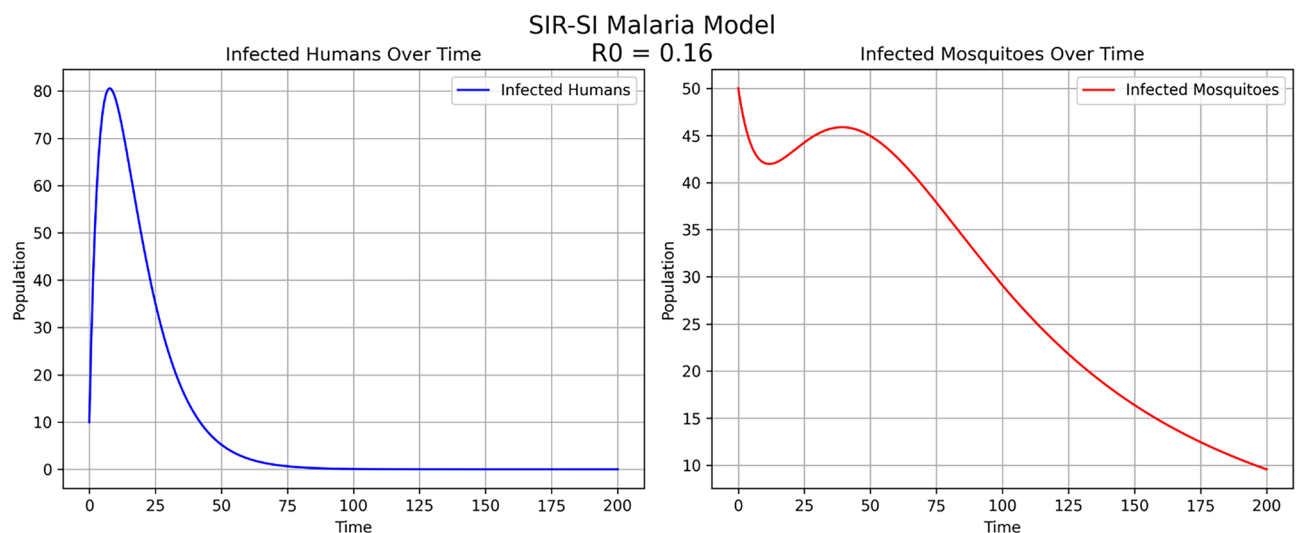
### Numerical validation of the stability theorems

In the previous section, we established that the disease-free steady state is attained when  $R_0 < 1$ . Accordingly, we selected  $R_0 = 0.04, 0.16$ , and  $0.51$ , ensuring  $R_0 < 1$ . As shown in Figs. 3, 4, and 5, the infected populations of humans and mosquitoes decline over time.

From Figs. 3, 4, and 5, we observe that the proposed model for malaria spread exhibits stable dynamics when the reproduction number remains below 1. Furthermore, as the reproduction number increases, the peak infection level intensifies, indicating a heightened disease burden within the population.



**Fig. 3.** For very small value of  $R_0 = 0.04$ , infected population converges to zero.



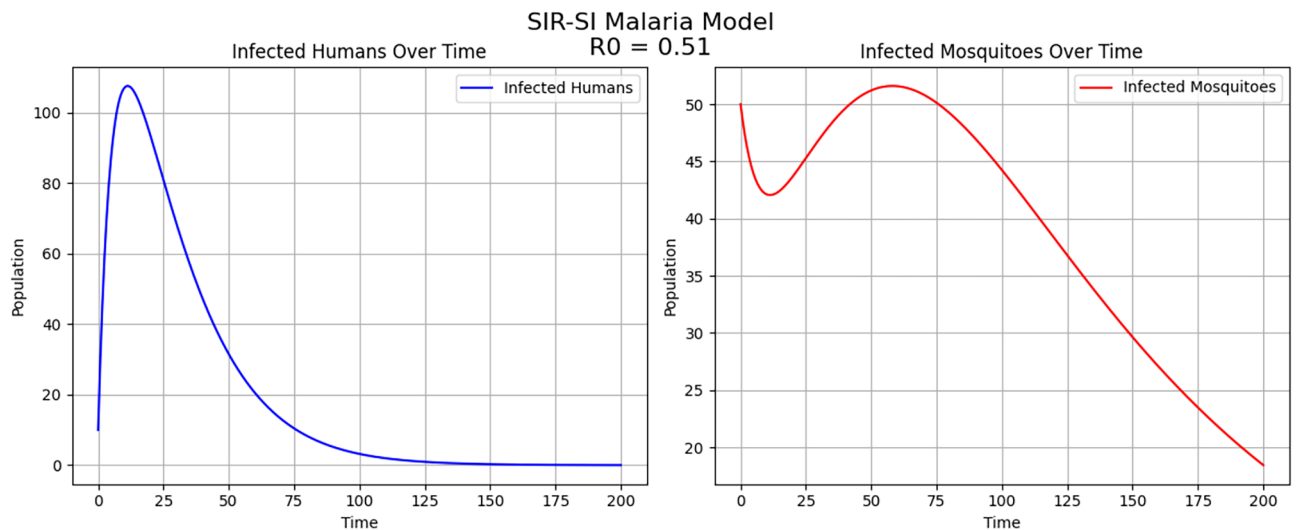
**Fig. 4.** For  $R_0 = 0.16$ , firstly infected population increases and then diminishes to zero.

### Importance of steady-state analysis and equilibrium points

Steady-state analysis in malaria modeling provides insights into the long-term behavior of transmission dynamics, focusing on equilibrium points where the rate of new infections and recoveries stabilizes. These equilibrium points are critical in understanding malaria persistence and eradication potential in specific regions. Identifying whether the system reaches a disease-free equilibrium or a sustained endemic state allows for targeted intervention strategies.

For malaria control, equilibrium points are essential for assessing the effectiveness of interventions. A disease-free equilibrium indicates that malaria transmission can be halted, typically through a combination of vector control measures such as insecticide-treated nets, indoor spraying, and mass drug administration. Conversely, an endemic equilibrium suggests that transmission is sustained within the population despite interventions, indicating that continuous, long-term strategies such as routine treatment, surveillance, and seasonal control programs are needed to keep malaria prevalence low. Stability analysis of these equilibrium points helps to identify critical thresholds, such as coverage levels for interventions or mosquito density, that determine whether malaria transmission will be suppressed or continue to persist, guiding the implementation of more effective control measures.

$$\frac{\beta_h \beta_m \Gamma_h \Gamma_m}{\mu_h \mu_m^2 (\mu_h + \gamma_h)} < 1.$$



**Fig. 5.** Behaviour of infected population for  $R_0 = 0.51$ . It reveals that the number of infected mosquitoes initially increases for a brief period before diminishing and eventually converging to zero.

This inequality indicates that if the basic reproduction number  $R_0$  is less than 1, it ensures that the disease-free steady state is stable, further supporting the potential for malaria eradication through effective control strategies.

### Temperature and altitude dependence of the transmission rate

Several factors influence malaria transmission; however, Patz et al.<sup>44</sup> identify temperature and altitude as among the most significant. Therefore, we model transmission as a function of temperature and altitude as follows:

$$\beta(T, h) = \beta_0 e^{-\frac{(T-25)^2}{\eta^2}} e^{-\frac{h^2}{\xi^2}} (1 - e^{-\frac{h^2}{\xi^2}}), \quad (3)$$

where  $\beta_0$  is the transmission constant of the region,  $T$  is the temperature of the region,  $h$  is the altitude,  $\eta$  and  $\xi$  are the constants associated with the region's temperature and height, respectively. The Gaussian function is centered at  $25^\circ\text{C}$ , as this temperature is biologically optimal for mosquito survival, leading to maximum transmission. This aligns with the findings of Shapiro et al.<sup>45</sup>, which also identify a similar optimal temperature for malaria mosquito breeding.

Here  $e^{-\frac{(T-25)^2}{\eta^2}}$  is used to model the temperature and  $e^{-\frac{h^2}{\xi^2}} (1 - e^{-\frac{h^2}{\xi^2}})$  is used to model the height. Malaria transmission will be very minimal when the temperature is either extremely high or it is extremely low and thus to model this variation, the Gaussian function is used and the reason for shifting it by 25 is because the optimum temperature for mosquito's existence and malaria transmission is  $25^\circ\text{C}$ . Also, malaria transmission is completely zero when the altitude is zero since there won't be any mosquitoes in the sea and in the same way when the altitude is extremely high again the transmission is completely zero since there are no mosquitoes in the space and thus to model both of these conditions the negative exponential function is used in this manner.

Assuming the temperature is from  $T_1$  to  $T_2$  and the height is from  $h_1$  to  $h_2$ , we can write

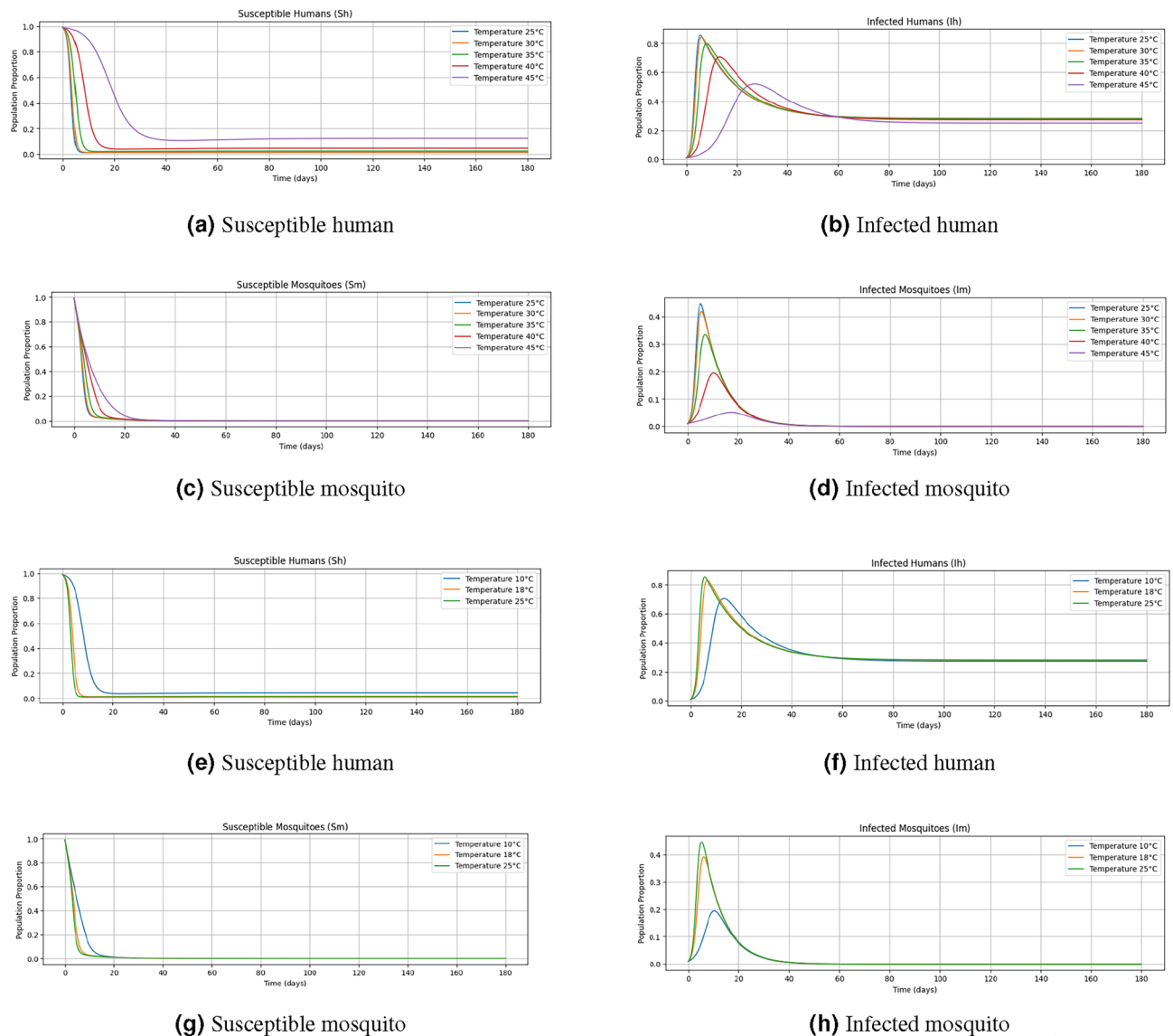
$$\beta_{avg} = \int_{T_1}^{T_2} \int_{h_1}^{h_2} e^{-\frac{(T-25)^2}{\eta^2}} e^{-\frac{h^2}{\xi^2}} (1 - e^{-\frac{h^2}{\xi^2}}) dh dT. \quad (4)$$

Here, the effect of transmission rate is studied (Figs. 6–8) by changing the temperature values for a fixed height. The parameters considered are the following:  $\beta_0 = 10$ , and for human we have taken  $\eta = 200$ ,  $\xi = 20000$  and for mosquitoes,  $\eta = 400$ ,  $\xi = 40000$ . The different temperature values which we used are  $25^\circ$ ,  $30^\circ$ ,  $35^\circ$ ,  $40^\circ$ ,  $45^\circ$ .

From Figs. 9 to 11, we observe that the transmission rate is highest when the altitude is 150. This represents the maximum value among the altitudes considered. Therefore, we conclude that  $h = 150$  may serve as a threshold value for altitude in the model. The function  $f(h) = e^{-\frac{h^2}{\xi^2}} (1 - e^{-\frac{h^2}{\xi^2}})$  achieves its highest value when  $h = \xi \sqrt{\ln(2)}$ , and thus the transmission rate increases as the height value approaches  $h = \xi \sqrt{\ln(2)}$ .

By substituting both values of  $\xi$  and taking the average, we get approximately 142.1; thus, the optimal value for altitude can be concluded to be in the range 140–150, and since it can be observed that the trajectory for all of these values is nearly the same.





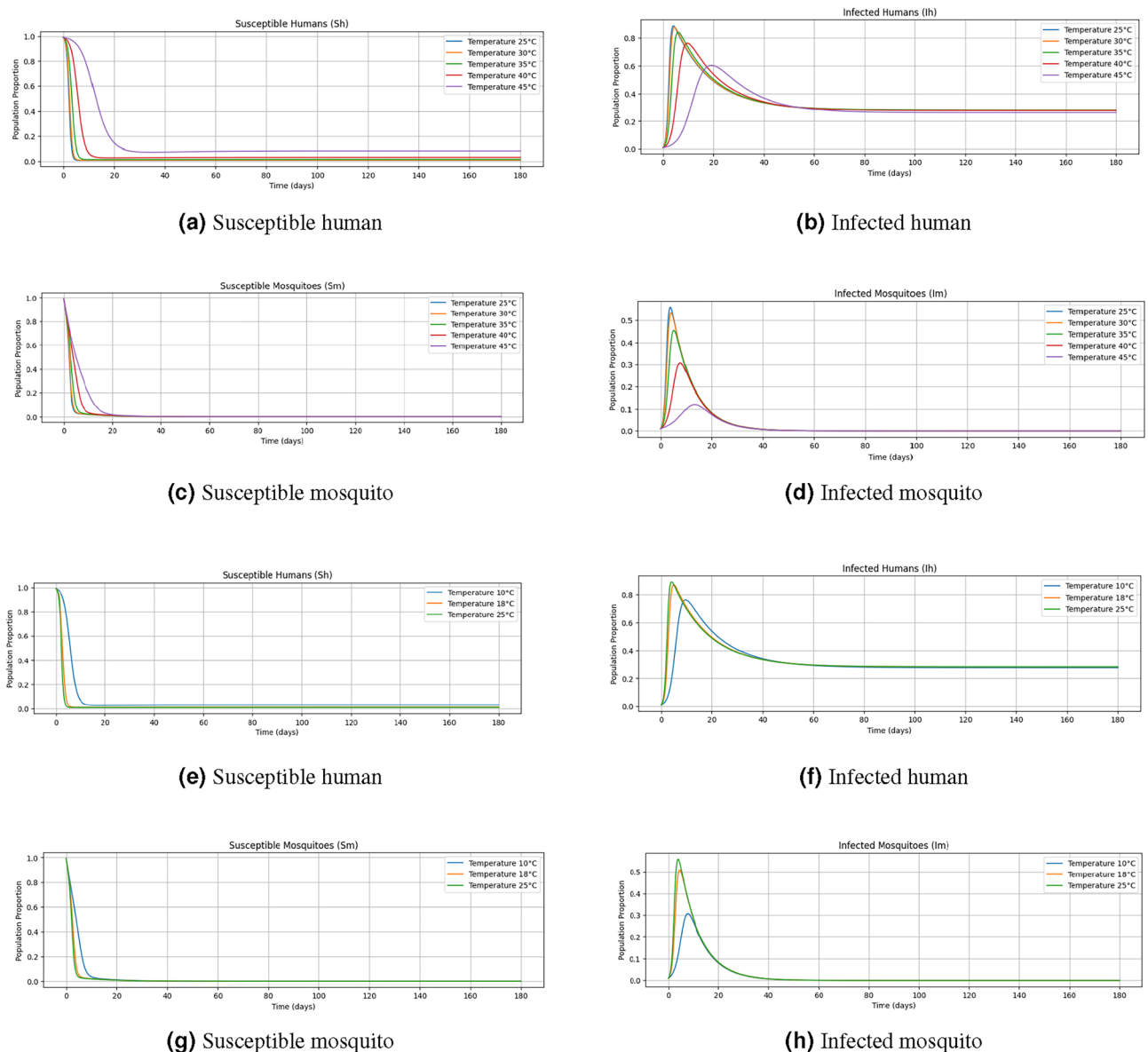
**Fig. 6.** At a constant altitude of 75 meters, temperature variations above and below 25°C were rigorously assessed to determine their influence on transmission dynamics. Our analysis reveals that 25°C serves as the optimal temperature for maximizing transmission rates. Analysis demonstrates a consistent decline in both infected mosquito and human populations as temperatures rise beyond this threshold. Furthermore, population trends for infected mosquitoes and humans align with a Gaussian distribution, reflecting a symmetrical, bell-shaped relationship with temperature. These findings underscore the critical role of temperature regulation in transmission mitigation strategies.

## Data-driven methods

Unlike conventional epidemic modeling approaches grounded in deterministic differential equations, this work adopts a data-driven paradigm to analyze malaria transmission dynamics. Our methodology integrates machine learning techniques, specifically artificial neural networks (ANNs), recurrent neural networks (RNNs), and physics-informed neural networks (PINNs), to infer system parameters and predict trajectories directly from observational data. This shift eliminates the need for explicit implementation of complex mathematical formulations; instead, neural architectures autonomously learn latent patterns and adapt their computations through iterative training. We systematically evaluate these architectures for parameter estimation, with PINNs further constrained by epidemiological principles to ensure biological fidelity. The derived parameters are then used to forecast the infection trajectories, demonstrating how hybrid data-driven and mechanistic approaches can advance predictive modeling of malaria transmission under heterogeneous environmental conditions.

## Parameter estimation

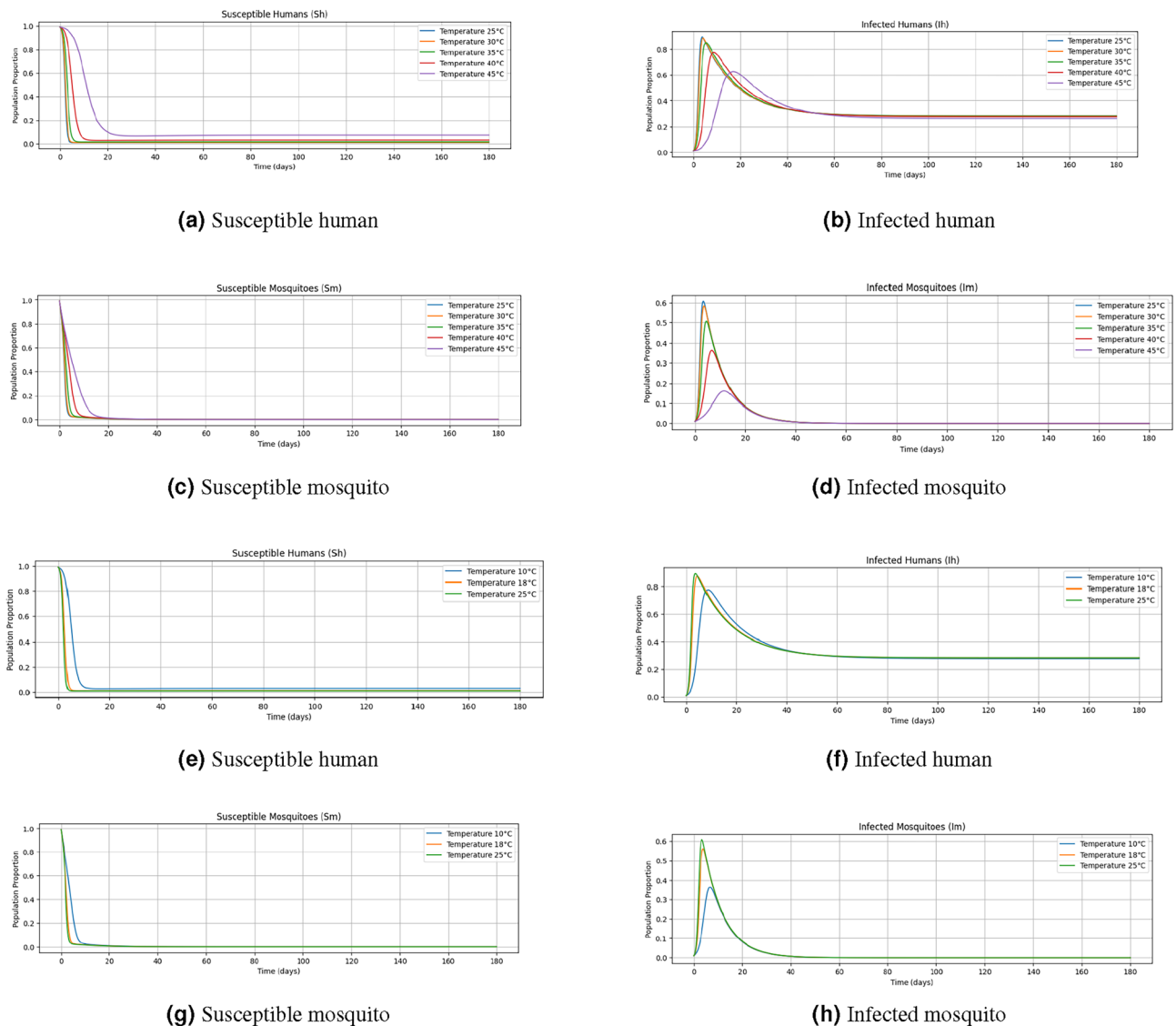
An important factor influencing disease transmission is the parameters associated with transmission dynamics. Understanding these parameters is crucial for medical professionals, as it allows them to assess the severity of the



**Fig. 7.** When the altitude was maintained at 100 m, temperature variations were analyzed to assess their impact on transmission dynamics. The findings confirm that 25°C is the optimal temperature for maximizing the transmission rate. Additionally, the population density of infected mosquitoes was higher compared to an altitude of 75 m, though the variation was marginal.

disease and implement appropriate interventions. This section presents three distinct neural network models. Artificial neural networks (ANNs), recurrent neural networks (RNNs), and physics-informed neural networks (PINNs) to estimate these parameters. Using the predicted parameters, we can forecast the trajectories of the various compartments involved in the disease dynamics. We used synthetic data generated from the system (1) to ensure consistency with the governing equations (critical for PINNs), enable controlled training of ANN and RNN models, and provide a clean benchmark for comparison using realistic parameters that isolate core transmission dynamics.

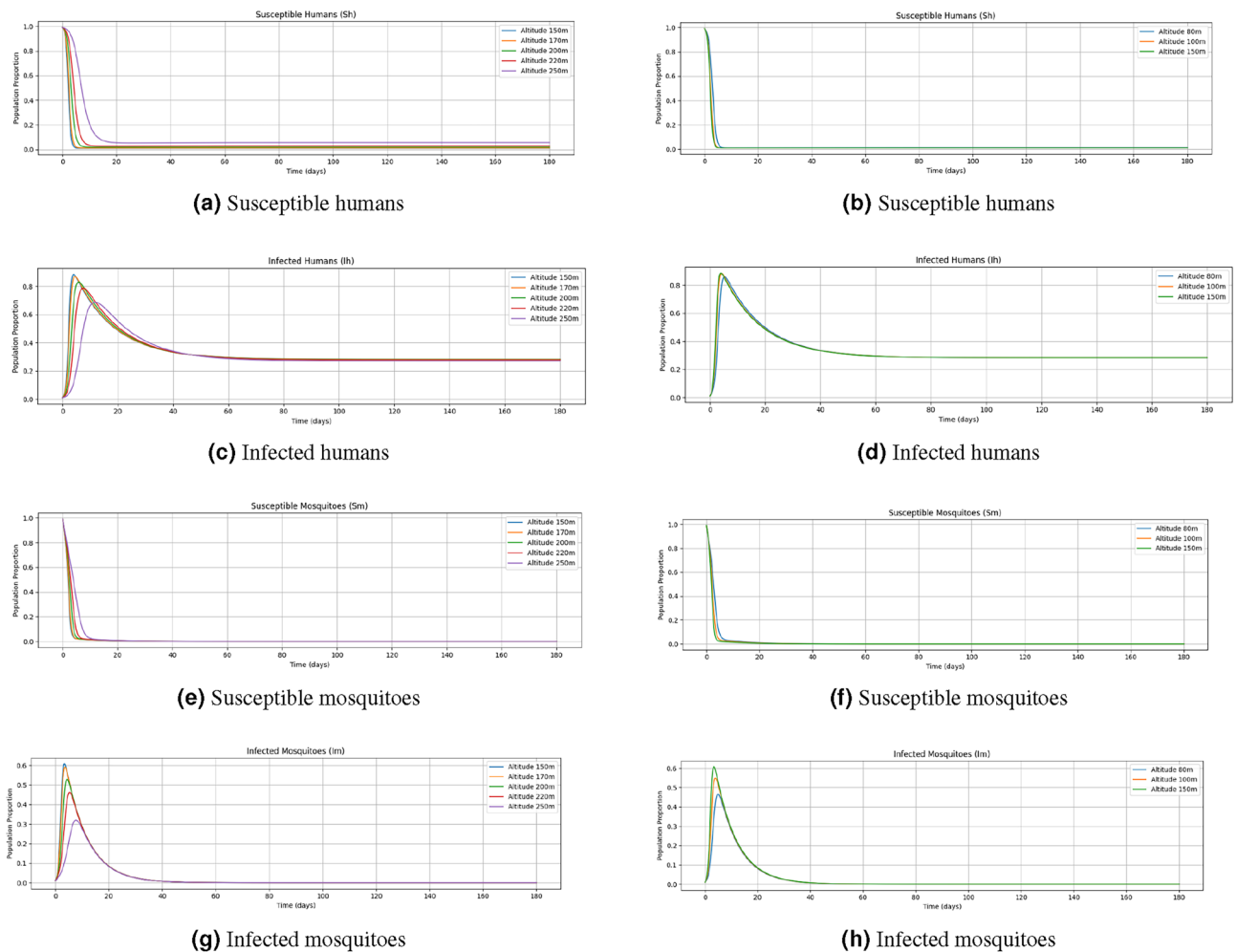
The architecture of Artificial Neural Networks (ANNs) employed in this study consisted of five layers, including three hidden layers. Each hidden layer comprised 15 dense units utilizing the sigmoid activation function, while the output layer contained seven dense units without an activation function. The Recurrent Neural Networks (RNNs) framework implemented in this work featured three layers: an input layer, a dropout layer, and an output layer. The input layer consisted of 50 Long Short-Term Memory (LSTM) units with the ReLU activation function, followed by a dropout layer with a 20% dropout rate to mitigate overfitting, and an output layer with seven dense units without an activation function. The Physics-Informed Neural Networks (PINNs) architecture closely followed the ANN structure, with the primary distinction being the number of nodes in the input and output layers, which were set to one and five, respectively. This modification enabled the



**Fig. 8.** At a fixed altitude of 125 m, the impact of temperature variations on transmission dynamics was analyzed. The findings consistently indicate that 25°C remains the optimal temperature for maximizing the transmission rate. A similar trend was observed, with a slight variation in the population of infected mosquitoes, reaching its peak at a higher point.

PINNs to incorporate physical constraints into the learning process, enhancing predictive accuracy and model generalizability.

Since three distinct neural network architectures are utilized in this study, the methodological approach differs for ANN, RNN, and PINN. For ANN and RNN, both models were trained on a dataset comprising 1000 data points, where the input consisted of the first 10 points of the trajectories for all compartments, and the output corresponded to the estimated parameters. The estimated parameters from ANN and RNN are shown in Table 3 and 4. The ANN and RNN models showed poor parameter estimation performance (errors > 200% for some parameters), highlighting their sensitivity to hyperparameters and training. While tuning might improve results, it is time-consuming and uncertain. In contrast, the methodology for PINN diverges significantly from these conventional models, as it does not rely on a predefined training dataset. Instead, an initial set of assumed parameters is iteratively refined by minimizing the loss function, thereby ensuring convergence to the actual parameter values through a comparison between the predicted and true trajectories. Here, the input variable is time, while the output consists of the trajectories of all five compartments. The selection of parameter values in this preliminary study was deliberately generic to evaluate the PINN framework's capability in recovering parameters under idealized conditions. While these values are not empirically derived, this approach allowed us to assess the algorithm's performance independently of confounding factors such as data noise or parameter interdependencies. Moving forward, parameter refinement will be guided by region-specific epidemiological literature—such as African mosquito mortality rates sourced from the Malaria Atlas Project, and sensitivity analyses to identify the most influential parameters. The ability of PINNs to effectively address inverse problems



**Fig. 9.** At a constant temperature of 28°C, altitude variations were studied. Population density trends for infected mosquitoes and humans exhibit a Gaussian distribution, reflecting a symmetrical, bell-shaped relationship with altitude. These findings highlight the significant impact of altitude on disease transmission and emphasize the importance of considering elevation in vector control strategies.

will further facilitate the integration of sparse field data, such as monthly case reports, enabling a more robust and data-driven parameter estimation process in future research.

For a novel comparison between the models, each model is simulated for 20000 epochs. Details of actual and predicted parameters are given in the tables below.

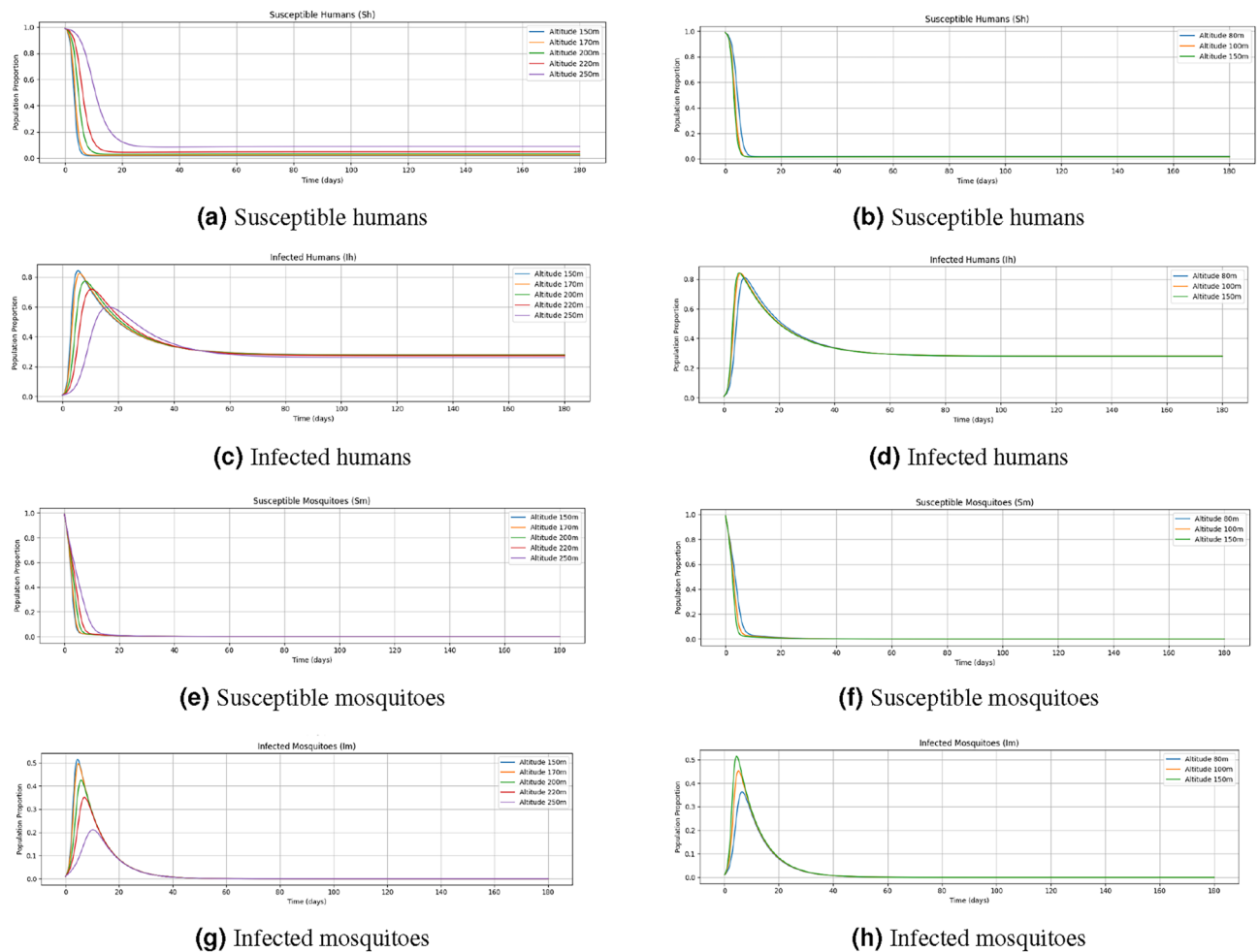
From Fig. 12 and Fig. 13, it can be observed that the predictions made by PINN closely align with the actual values, demonstrating its effectiveness in parameter estimation for the SIR-SI compartment model with trajectory prediction. Additionally, the trend analysis reveals that PINN effectively captures the underlying dynamics of the system, accurately reflecting the growth and decline patterns of infections over time. The model consistently follows the expected trajectory, reinforcing its reliability in forecasting epidemic progression.

PINNs outperform traditional ANN and RNN models by integrating epidemiological equations directly into the learning process. Unlike purely data-driven approaches, PINNs enforce physical consistency through a dual loss function—combining data loss with physics-informed loss—leading to better generalization, especially under sparse data. This makes them particularly effective for parameter estimation and predictive modeling in disease dynamics, where data may be limited but underlying processes are well-understood. Loss function in PINNs involves both data loss and physics loss, which gives better estimates for parameters. Mathematically,

$$\mathcal{L}_{physics} = \sum_i |\mathcal{F}(t_i)|^2 \quad (5)$$

$$\mathcal{L}_{data} = \sum_i |u_\theta(t_i) - u_i^{true}|^2 \quad (6)$$

$$\mathcal{L} = \lambda_{physics} \mathcal{L}_{physics} + \lambda_{data} \mathcal{L}_{data} \quad (7)$$

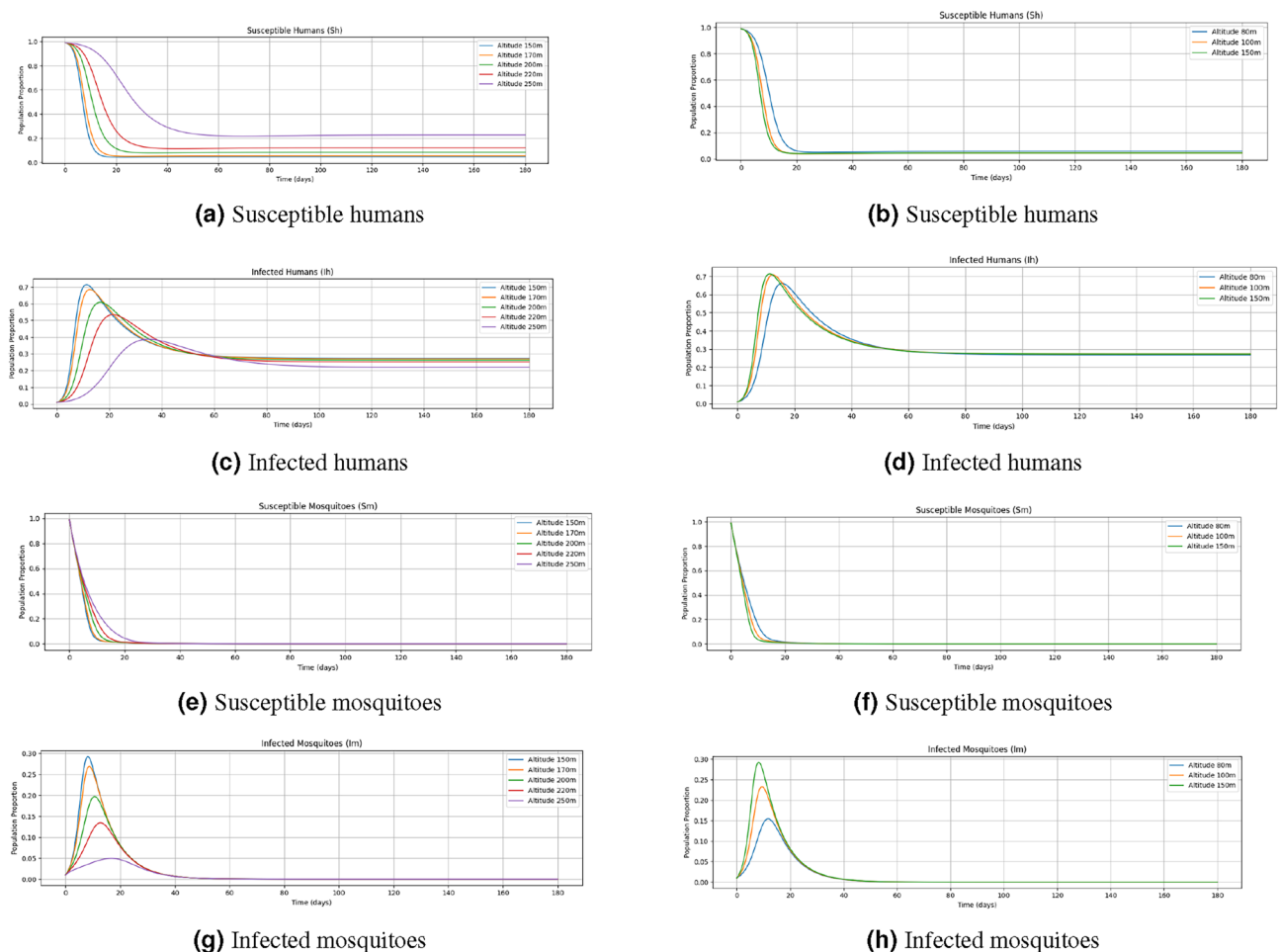


**Fig. 10.** With the temperature maintained at 35°C, the impact of altitude fluctuations on transmission dynamics was further examined. A similar trend was observed; however, at 35°C, the peak population of infected mosquitoes occurred at lower altitudes.

where:

- $\mathcal{L}_{physics}$  - Physics-informed loss, ensuring that the learned function satisfies the governing physical laws.
- $\mathcal{L}_{data}$  - Data loss, ensuring that the learned function closely matches observed data.
- $\mathcal{L}$  - Total loss function combining physics and data losses.
- $\mathcal{F}(t_i)$  - Residual of the governing physical equation at time  $t_i$ .
- $u_\theta(t_i)$  - Neural network approximation of the true solution at time  $t_i$ .
- $u_i^{true}$  - True observed data value at time  $t_i$ .
- $\lambda_{physics}$  - Weighting factor for the physics-informed loss.
- $\lambda_{data}$  - Weighting factor for the data loss.

The parameters predicted by PINNs for all five compartments are presented in Table 5. The temporal evolution of the human population is depicted in Fig. 12, while the mosquito population dynamics are shown in Fig. 13. Previous studies have predominantly relied on traditional time-series or statistical methods, such as LSTM, for epidemic forecasting. For example, Wang et al.<sup>46</sup> applied LSTM to predict COVID-19 trends, Chandra et al.<sup>47</sup> explored ARIMA for dengue incidence modeling, Yadav et al.<sup>48</sup> used regression frameworks for malaria risk assessment, and Elshafee et al.<sup>49</sup> employed Bayesian statistical approaches. In contrast to these data-driven methods, our work leverages PINNs to integrate mechanistic epidemiological principles (e.g., transmission dynamics and compartmental interactions) directly into the parameter estimation process. This physics-informed approach achieves minimal error across nearly all compartments (Table 5), demonstrating its ability to reconcile observed data with domain knowledge. These results highlight the advantages of PINNs for epidemic modeling, as they inherently encode the biophysical processes governing disease spread, enabling robust parameter inference even with sparse or noisy datasets. While Physics-Informed Neural Networks (PINNs) showed promising results in our study, it is important to acknowledge their limitations. The superior performance of PINNs is largely attributed to the integration of known physical laws into the learning process. However, in scenarios where the governing dynamics are poorly understood, highly stochastic, or where the



**Fig. 11.** The impact of altitude variations on transmission dynamics was further analyzed at a fixed temperature of  $42^{\circ}\text{C}$ . In this case, a more significant decline in the population of infected mosquitoes was observed. These findings highlight the critical role of altitude in shaping transmission patterns and optimizing intervention strategies.

| Parameter                       | Predicted Value | Actual Value | Relative Error (%) |
|---------------------------------|-----------------|--------------|--------------------|
| Birth rate of humans            | 0.38273         | 0.40000      | 4.32               |
| Transmission rate of humans     | 0.28249         | 0.30000      | 5.84               |
| Death rate of humans            | 0.09973         | 0.10000      | 0.27               |
| Recovery rate of humans         | 0.04640         | 0.01000      | 364.01             |
| Birth rate of mosquitoes        | 0.07613         | 0.05000      | 52.25              |
| Transmission rate of mosquitoes | 0.01045         | 0.02000      | 47.73              |
| Death rate of mosquitoes        | 0.05757         | 0.04000      | 43.92              |

**Table 3.** Comparison of predicted and actual values with relative errors for ANN.

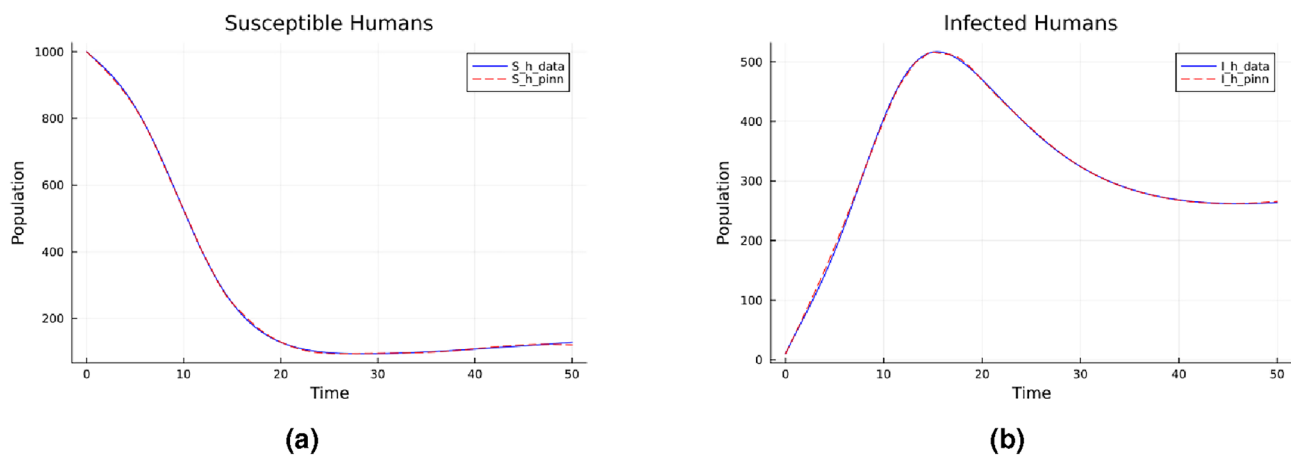
data-generating process deviates significantly from the assumed model structure, as may be the case with real, noisy malaria data, PINNs may underperform compared to conventional data-driven approaches like ANN or RNN. Thus, their applicability is inherently constrained by the availability and accuracy of the underlying physical model.

### Finding the risk of a disease

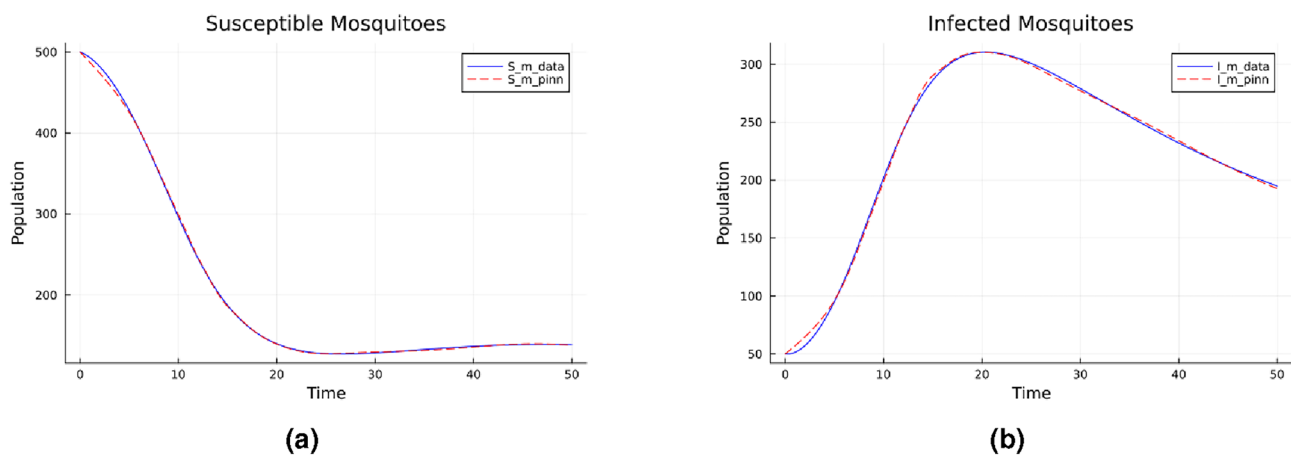
The most important aspect of a disease is the determination of risk, which we define as the number of infected people in a particular region. Whenever there is a disease outbreak in a country, there are some regions where there is more risk compared to the other regions; thus, it is essential to calculate the risk of every region. This problem statement is addressed using the method of DMD (dynamic mode decomposition), and the main

| Parameter                       | Predicted Value | Actual Value | Relative Error (%) |
|---------------------------------|-----------------|--------------|--------------------|
| Birth rate of humans            | 0.19740         | 0.40000      | 50.65              |
| Transmission rate of humans     | 0.18619         | 0.30000      | 37.94              |
| Death rate of humans            | 0.34348         | 0.10000      | 243.48             |
| Recovery rate of humans         | 0.21954         | 0.01000      | 2095.38            |
| Birth rate of mosquitoes        | 0.18898         | 0.05000      | 277.96             |
| Transmission rate of mosquitoes | 0.24995         | 0.02000      | 1149.77            |
| Death rate of mosquitoes        | 0.38605         | 0.04000      | 865.13             |

**Table 4.** Comparison of predicted and actual values with relative errors for RNN.



**Fig. 12.** Prediction made by PINNs for the human population.



**Fig. 13.** Prediction made by PINNs for the mosquito population.

reason for using this method is that DMD can make exact predictions from raw data, unlike other deep learning methods. The complete methodology of the problem statement can be seen in Fig. 14.

Dynamic Mode Decomposition (DMD) extracts spatiotemporal coherent structures from high-dimensional dynamical systems<sup>30–32</sup>. Given a sequence of  $m + 1$  state vectors  $x_k \in \mathbb{R}^n$  sampled at intervals  $\Delta t$ , we construct snapshot matrices:

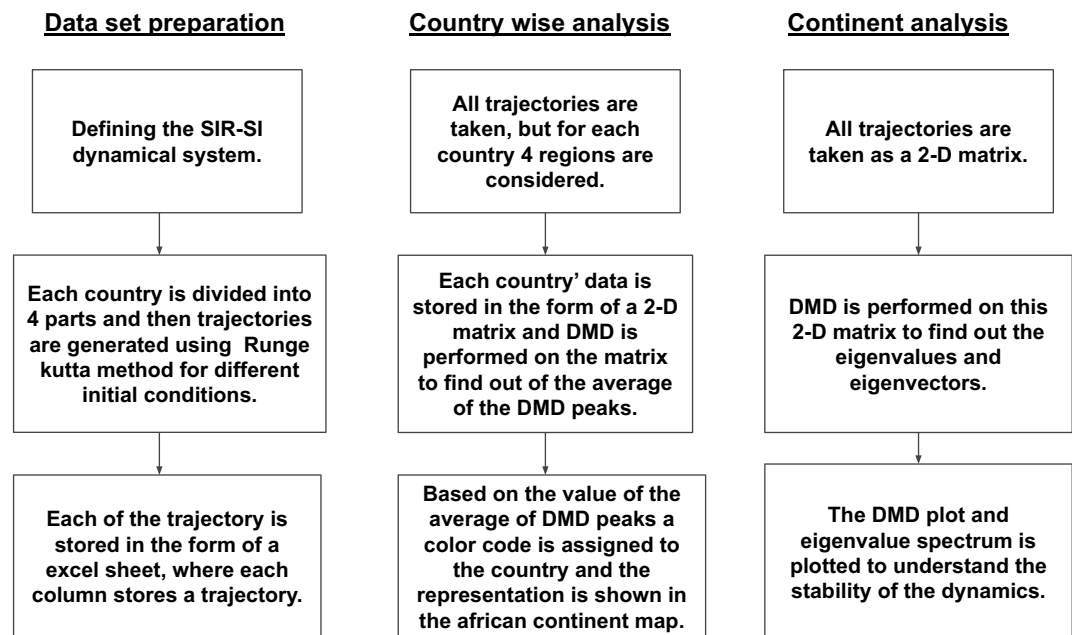
$$X = [x_1, x_2, \dots, x_m], \quad X' = [x_2, x_3, \dots, x_{m+1}] \quad (8)$$

DMD seeks a best-fit linear operator  $A \in \mathbb{R}^{n \times n}$  satisfying:  $X' \approx AX$ . The solution proceeds via truncated singular value decomposition (SVD):



| Parameter                       | Predicted Value | Actual Value | Relative Error (%) |
|---------------------------------|-----------------|--------------|--------------------|
| Birth rate of humans            | 0.39200         | 0.40000      | 2.00               |
| Transmission rate of humans     | 0.28186         | 0.30000      | 6.05               |
| Death rate of humans            | 0.09793         | 0.10000      | 2.07               |
| Recovery rate of humans         | 0.00949         | 0.01000      | 5.07               |
| Birth rate of mosquitoes        | 0.04593         | 0.05000      | 8.14               |
| Transmission rate of mosquitoes | 0.01969         | 0.02000      | 1.56               |
| Death rate of mosquitoes        | 0.03600         | 0.04000      | 10.01              |

**Table 5.** Comparison of predicted and actual values with relative errors for PINN.



**Fig. 14.** Flow chart of methodology.

$$X \approx U_r \Sigma_r V_r^* \quad (\text{rank-}r \text{ approximation})$$

$$\tilde{A} = U_r^* X' V_r \Sigma_r^{-1} \quad (\text{reduced operator})$$

Solving the eigenvalue problem for  $\tilde{A}$ :  $\tilde{A}W = W\Lambda$ ,  $\Lambda = \text{diag}(\lambda_i)$  yields full-state DMD modes:

$$\Phi = X' V_r \Sigma_r^{-1} W \quad (9)$$

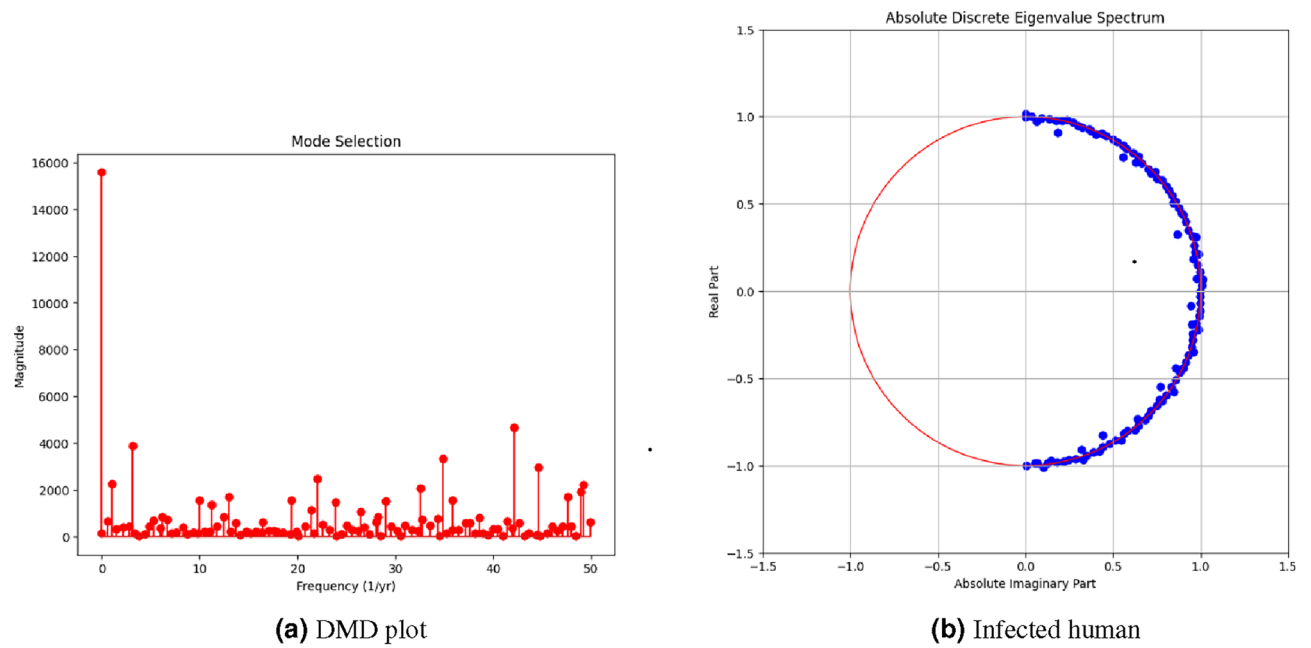
The continuous-time eigenvalues  $\omega_i = \ln(\lambda_i)/\Delta t$  determine mode dynamics (growth/decay rates and frequencies). The reconstructed solution is:  $x(t) \approx \sum_{i=1}^r \phi_i \exp(\omega_i t) b_i$ ,  $b = \Phi^\dagger x_1$ , where  $\phi_i$  are columns of  $\Phi$  and  $b$  contains mode amplitudes.

In this work, Dynamic Mode Decomposition (DMD) is used to calculate the disease risk in a particular region. DMD provides early detection by identifying patterns in time-series data, facilitating timely intervention. Additionally, it performs dimensionality reduction by extracting dominant modes from large datasets, preserving essential dynamical features while reducing computational complexity. Its forecasting capability allows for accurate predictions of disease progression, aiding in proactive decision-making. Moreover, while standard DMD operates linearly, extended DMD (eDMD) can approximate nonlinear disease dynamics, making it adaptable to complex epidemiological models. DMD also captures the oscillations of the dynamics, and by analyzing its peak values, we obtain a measure of risk. The DMD plot and the eigenvalue spectrum can be found in Fig. 15.

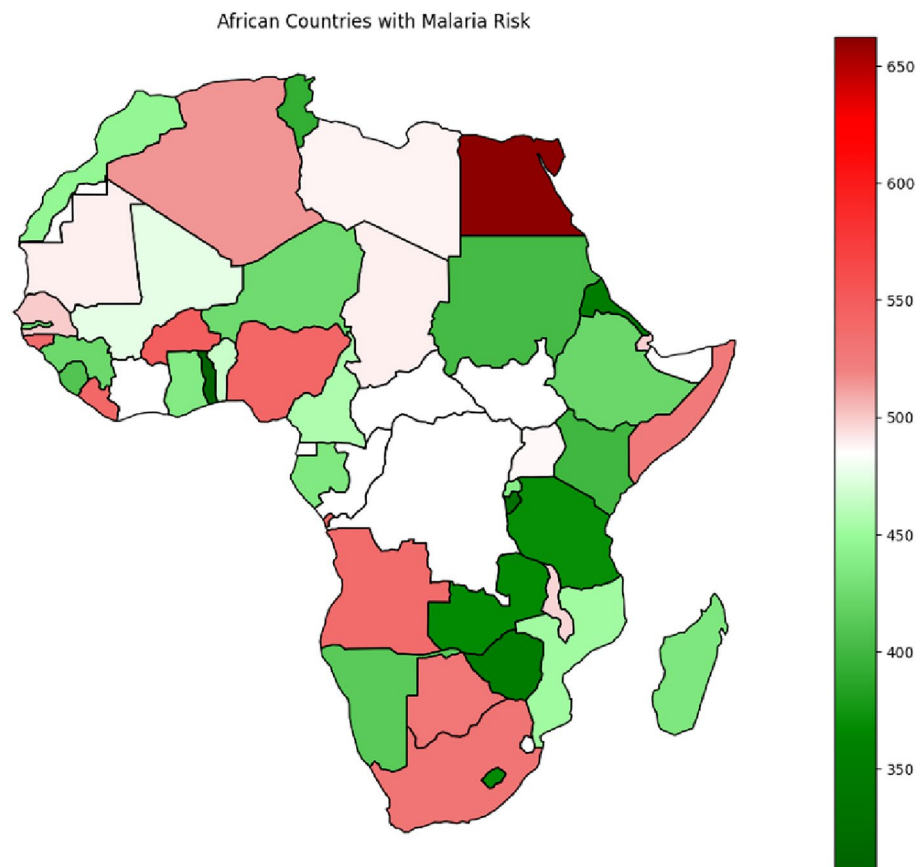
From the eigenvalue spectrum, we can observe that all the points are either on or within the unit circle. This shows that the transmission of malaria in Africa is stable. African map with the corresponding color coding based on the risk can be found in Fig. 16.

From Sub Fig. 15b we can observe that the disease spread is not severe since all of the infections are within the unit circle. DMD extracts governing dynamics directly from observational data, making it advantageous





**Fig. 15.** DMD plot and the eigenvalue spectrum.



**Fig. 16.** A map of Africa with color coding representing risk levels, where green indicates the lowest risk and red denotes the regions with the highest risk.

for rapidly evolving outbreaks where mechanistic understanding is incomplete<sup>50</sup>. Unlike machine learning approaches (e.g., LSTM), which excel at long term prediction but acts as 'black boxes', DMD provides interpretable spatial-temporal modes, albeit with trade-offs in modeling strongly nonlinear interactions. This balance positions DMD as a complementary tool for real-time risk assesment alongside established methods.

## Concluding remarks

In this work, an attempt is made to understand the dynamics of malaria transmission using various mathematical and machine-learning techniques. The proposed theorems related to steady states were validated using numerical simulations. We also examined the influence of temperature and altitude on malaria transmission. Subsequently, parameter estimation was performed using three neural network architectures—ANNs, RNNs, and PINNs—where the best-performing model's predicted parameters were utilized to forecast the trajectories of all compartments. Finally, the study assessed the measure of risk using Dynamic Mode Decomposition (DMD), providing valuable insights into malaria dynamics and prediction.

While these results are preliminary and based solely on synthetic data, they demonstrate the potential of the proposed methodology as a step toward developing actionable tools for policymakers. By accurately quantifying transmission dynamics in controlled settings, this approach lays the groundwork for future integration with real-world data. In particular, incorporating PINNs with real-time surveillance systems could, in principle, support the identification of transmission hotspots and inform scenario-based planning under climatic uncertainties. Such extensions, once validated with empirical data, may ultimately assist in optimizing resource allocation and enhancing outbreak preparedness in climate-sensitive regions.

In the future, we aim to enhance malaria transmission modeling by incorporating advanced physics-informed machine learning techniques. Sparse Identification of Nonlinear Dynamics (SINDy) can help discover parsimonious governing equations from data, potentially uncovering key drivers of transmission. Neural ODEs provide a flexible framework to model continuous-time disease dynamics from irregular time series data. Variants such as VPINNs (Variational PINNs) and Recurrent PINNs can improve model accuracy and scalability, particularly in capturing spatiotemporal variability and memory effects.

## Data availability

This study uses only synthetic data, which was generated for the purpose of this research. The synthetic data are not based on real-world observations and can be made available upon request to the corresponding author at : a-tridane@uaeu.ac.ae

Received: 6 January 2025; Accepted: 15 July 2025

Published online: 26 July 2025

## References

- Ross, R. *The prevention of malaria* (John Murray, 1911).
- Ross, R. Some a priori pathometric equations. *Br. Med. J.* **1**, 546 (1915).
- Ross, R. An application of the theory of probabilities to the study of a priori pathometry. part i. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character* **92**, 204–230 (1916).
- Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry. part iii. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character* **93**, 225–240 (1917).
- Macdonald, G. *The epidemiology and control of malaria* (Oxford University Press, 1957).
- Mandal, S., Sarkar, R. R. & Sinha, S. Mathematical models of malaria—a review. *Malar. J.* **10**, 1–19 (2011).
- Ogueda, A., Martinez, E., Arunachalam, V. & Seshaiyer, P. Machine learning for predicting the dynamics of infectious diseases during travel through physics informed neural networks. *Journal of Machine Learning for Modeling and Computing* (2023).
- Schiassi, E., De Florio, M., D'Ambrosio, A., Mortari, D. & Furfaro, R. Physics-informed neural networks and functional interpolation for data-driven parameters discovery of epidemiological compartmental models. *Mathematics* **9**, 2069 (2021).
- Ning, X., Guan, J., Li, X.-A., Wei, Y. & Chen, F. Physics-informed neural networks integrating compartmental model for analyzing covid-19 transmission dynamics. *Viruses* **15**, 1749 (2023).
- Ning, X., Li, X.-A., Wei, Y. & Chen, F. Euler iteration augmented physics-informed neural networks for time-varying parameter estimation of the epidemic compartmental model. *Front. Phys.* **10**, 1062554 (2022).
- Deng, Q. Dynamics and development of the covid-19 epidemics in the us—a compartmental model with deep learning enhancement. *medRxiv* 2020–05 (2020).
- Bousquet, A., Conrad, W. H., Sadat, S. O., Vardanyan, N. & Hong, Y. Deep learning forecasting using time-varying parameters of the sird model for covid-19. *Sci. Rep.* **12**, 3030 (2022).
- Deng, Q. Modeling the omicron dynamics and development in china: with a deep learning enhanced compartmental model. *medRxiv* 2022–06 (2022).
- Deng, Q. & Wang, G. A deep learning-enhanced compartmental model and its application in modeling omicron in china. *Bioengineering* **11**, 906 (2024).
- Baccegga, D., Castagno, P., Fernández Anta, A. & Sereno, M. Enhancing covid-19 forecasting precision through the integration of compartmental models, machine learning and variants. *Sci. Rep.* **14**, 19220 (2024).
- Chen, W., Luo, H., Li, J. & Chi, J. Long-term trend prediction of pandemic combining the compartmental and deep learning models. *Sci. Rep.* **14**, 21068 (2024).
- Ma, J. *et al.* Rnn enhanced compartmental model for infectious disease prediction. In *2024 IEEE International Conference on Digital Health (ICDH)*, 225–236 (IEEE, 2024).
- Millevoi, C., Pasetto, D. & Ferronato, M. A physics-informed neural network approach for compartmental epidemiological models. *PLoS Comput. Biol.* **20**, e1012387 (2024).
- Islam, M. S., Shahrear, P., Saha, G., Ataulha, M. & Rahman, M. S. Mathematical analysis and prediction of future outbreak of dengue on time-varying contact rate using machine learning approach. *Comput. Biol. Med.* **178**, 108707 (2024).
- Xue, D., Wang, M., Liu, F. & Buss, M. Time series modeling and forecasting of epidemic spreading processes using deep transfer learning. *Chaos, Solitons & Fractals* **185**, 115092 (2024).
- Shukla, S. S. P., Jain, V. K., Yadav, A. K. & Pandey, S. K. Fourth wave covid19 analyzing using mathematical seirs epidemic model & deep neural network. *Multimed. Tools Appl.* **83**, 27507–27526 (2024).

22. Juneja, M., Saini, S. K., Kaur, H. & Jindal, P. Statistical machine and deep learning methods for forecasting of covid-19. *Wireless Personal Communications* **138**, 497–524 (2024).
23. Cumbane, S. P. & Gidófalvi, G. Deep learning-based approach for covid-19 spread prediction. *International Journal of Data Science and Analytics* 1–17 (2024).
24. Hu, H., Kennedy, C. M., Kevrekidis, P. G. & Zhang, H.-K. A modified pinn approach for identifiable compartmental models in epidemiology with application to covid-19. *Viruses* **14**, 2464 (2022).
25. Anwar, N. et al. Stochastic supervised neuro-architecture design for analyzing vector-borne plant virus epidemics with latency and incubation effects. *Eur. Phys. J. Plus* **139**, 1–34 (2024).
26. Anwar, N. et al. Intelligent bayesian neural networks for stochastic svls epidemic dynamics: Vaccination strategies and prevalence fractions with wiener process. *Fluctuation and Noise Letters* **24**, 2550020–198 (2025).
27. Anwar, N. et al. Dynamical analysis of hepatitis b virus through the stochastic and the deterministic model. *Comput. Methods Biomech. Biomed. Engin.* 1–17 (2025).
28. Bhujji, G., Phaijoo, G. & Gurung, D. Mathematical study on impact of temperature in malaria disease transmission dynamics. *Advances in Computer Sciences* **1**, 1–8 (2018).
29. Keno, T. D., Makinde, O. D. & Obsu, L. L. Impact of temperature variability on sirs malaria model. *Journal of Biological Systems* **29**, 773–798 (2021).
30. Proctor, J. L., Brunton, S. L. & Kutz, J. N. Dynamic mode decomposition with control. *SIAM J. Appl. Dyn. Syst.* **15**, 142–161 (2016).
31. Alla, A. & Kutz, J. N. Nonlinear model order reduction via dynamic mode decomposition. *SIAM J. Sci. Comput.* **39**, B778–B796 (2017).
32. Andreuzzi, F., Demo, N. & Rozza, G. A dynamic mode decomposition extension for the forecasting of parametric dynamical systems. *SIAM J. Appl. Dyn. Syst.* **22**, 2432–2458 (2023).
33. Watson, G. L. et al. Pandemic velocity: Forecasting covid-19 in the us with a machine learning & bayesian time series compartmental model. *PLoS Comput. Biol.* **17**, e1008837 (2021).
34. Traoré, B., Koutou, O. & Sangaré, B. A global mathematical model of malaria transmission dynamics with structured mosquito population and temperature variations. *Nonlinear Analysis: Real World Applications* **53**, 103081 (2020).
35. Koella, J. C. On the use of mathematical models of malaria transmission. *Acta tropica* **49**, 1–25 (1991).
36. Chitnis, N., Hyman, J. M. & Cushing, J. M. Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model. *Bull. Math. Biol.* **70**, 1272–1296 (2008).
37. Chitnis, N., Cushing, J. M. & Hyman, J. Bifurcation analysis of a mathematical model for malaria transmission. *SIAM J. Appl. Math.* **67**, 24–45 (2006).
38. Osman, M. & Adu, I. Simple mathematical model for malaria transmission. *Journal of Advances in Mathematics and Computer Science* **25**, 1–24 (2017).
39. Dudley, H. J., Goenka, A., Orellana, C. J. & Martonosi, S. E. Multi-year optimization of malaria intervention: a mathematical model. *Malar. J.* **15**, 1–23 (2016).
40. Gebremeskel, A. A. & Krogstad, H. E. Mathematical modelling of endemic malaria transmission. *American Journal of Applied Mathematics* **3**, 36–46 (2015).
41. Eikenberry, S. E. & Gumel, A. B. Mathematical modeling of climate change and malaria transmission dynamics: a historical review. *J. Math. Biol.* **77**, 857–933 (2018).
42. Yacheur, S., Moussaoui, A. & Tridane, A. Modeling the imported malaria to north africa and the absorption effect of the immigrants. *Math. Biosci. Eng.* **16**, 967–989 (2019).
43. Raza, A., Arif, M. S. & Rafiq, M. A reliable numerical analysis for stochastic dengue epidemic model with incubation period of virus. *Adv. Differ. Equ.* **2019**, 1–19 (2019).
44. Patz, J. A. & Olson, S. H. Malaria risk and temperature: influences from global climate change and local land use practices. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5635–5636. <https://doi.org/10.1073/pnas.0601493103> (2006).
45. Shapiro, L. L., Whitehead, S. A. & Thomas, M. B. Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria. *PLoS Biology* **15**, e2003489 (2017).
46. Wang, P., Zheng, X., Ai, G., Liu, D. & Zhu, B. Time series prediction for the epidemic trends of covid-19 using the improved lstm deep learning method: Case studies in russia, peru and iran. *Chaos, Solitons & Fractals* **140**, 110214. <https://doi.org/10.1016/j.chaos.2020.110214> (2020).
47. Chandra, R., Jain, A. & Singh Chauhan, D. Deep learning via lstm models for covid-19 infection forecasting in india. *PLoS One* **17**, e0262708. <https://doi.org/10.1371/journal.pone.0262708> (2022).
48. Yadav, S. K. & Akhter, Y. Statistical modeling for the prediction of infectious disease dissemination with special reference to covid-19 spread. *Front. Public Health* **9**, 645405 (2021).
49. ElShafee, A., El-Shafai, W., Algarni, A. D., Soliman, N. F. & Aly, M. H. Statistical time series forecasting models for pandemic prediction. *Comput. Syst. Sci. Eng.* **47** (2023).
50. Proctor, J. L. & Eckhoff, P. A. Discovering dynamic patterns from infectious disease data using dynamic mode decomposition. *Int. Health* **7**, 139–145 (2015).

## Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable comments.

## Author contributions

All the authors contribute equally conception and design of the work; modeling, analysis: A.R. and M.K.; visualization and interpretation of the results: A.R., M.K., A.T.; funding and supervising: A.T.; wrote the paper: A.R., M.K.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to A.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025