International Journal of Dynamics and Control (2025) 13:319
https://doi.org/10.1007/s40435-025-01837-w

A mpox deep learning prediction framework incorporating a two-strain SIRD epidemic model focusing on the cases of the Democratic Republic of the Congo

Sayani Adak[1] · **Sankar K. Pal**[1]

Received: 21 June 2025 / Revised: 6 August 2025 / Accepted: 13 August 2025

Abstract
Historically endemic to Africa, mpox (formerly known as monkeypox) experienced its greatest outbreak in 2022, spreading to numerous parts of the world and posing a concern to public health. The WHO Director-General proclaimed on August 14, 2024, that the rise of mpox cases in the Democratic Republic of the Congo (DRC) and its spread to nearby nations constitute a Public Health Emergency of International Concern (PHEIC) as per the International Health Regulations (2005). Our present research focuses on the clade Ia and clade Ib outbreaks occurring in the Democratic Republic of the Congo. We formulate a two-strain susceptible–infected–recovered–deceased (TSSIRD) model using a system of ordinary differential equations. We predict the parameters derived from the TSSIRD model incorporating deep learning (DL) architectures such as LSTM, BiLSTM, Stacked LSTM, ConvLSTM, and ConvBiLSTM. The resulting TSSIRD-DL model is used for single-day prediction of the number of infected for both clades. The experimental results show that the TSSIRD-DL models predict significantly well in comparison with the pure deep learning models where the data are taken from January 1, 2023, to December 29, 2024. This approach can be fruitful for the single-day prediction of the infected population when only small historical data sets are available. Additionally, we have performed a global sensitivity analysis on the model parameters to understand their effects on basic reproduction numbers and infected individuals (for both clades). As a result, the significant parameters for the spread of the disease are found to be the disease transmission rate and the rate of recovery.

Keyword mpox modeling; mpox variants; mpox data of the DRC; Deep learning architecture; Global sensitivity analysis

1 Introduction

Numerous epidemics caused by diseases such as COVID-19, rabies, malaria, Nipah, and Ebola demonstrate that zoonotic diseases continue to be a major global public health concern. A case of mpox (formerly known as monkeypox) that was not related to travel to endemic countries was first reported in the UK in mid-2022, sparking the global outbreak of mpox, another zoonotic disease [1].

Historically endemic to Africa, mpox experienced its greatest outbreak in 2022, spreading to numerous parts of the world and posing a concern to public health. Humans can contract this disease through physical contact, body fluids, respiratory droplets, and sexual contact (especially among men who are sexually related to men) [2]. Swollen lymph nodes, fever, and rash are typical signs of mpox. If infected, it takes two to four weeks to recover from mpox without treatment. However, children, women who are pregnant and people with weak immune systems, may be severely affected by the disease. 121 member nations from all six WHO (World Health Organization) regions have reported mpox cases to the WHO since January 1, 2022. As of July 31, 2024, WHO has received reports of 1,02,997 laboratory-confirmed cases [3].

Population immunity to mpox was backed by the smallpox vaccine during the global vaccination efforts against smallpox between 1967 and 1979, as the vaccine of smallpox provided some protection. Nevertheless, vaccination stopped after smallpox was eradicated in 1979. Moreover, the protection of vaccinated individuals against mpox is currently falling due to the combination of the epidemiological evolution of the virus and the loss of vaccine protection. This can be the cause for concern in the ongoing global mpox outbreak.

✉ Sayani Adak
sayaniadak1994@gmail.com

Sankar K. Pal
sankarpal@yahoo.com

[1] Indian Statistical Institute, Kolkata, India

Published online: 02 September 2025

The WHO Director-General proclaimed on August 14, 2024, that the increase in mpox cases in the Democratic Republic of the Congo (DRC) and its spread to nearby nations constitute a Public Health Emergency of International Concern (PHEIC) according to the International Health Regulations (2005) [4]. This spread compels a systemized international response since it poses a risk to the public health of other states. In 2022, the WHO reported multiple international mpox outbreaks in 20 non-endemic European countries, as well as the USA, Canada, Mexico, and much of South America [5].

Mpox is endemic to Central and West Africa. Countries like Burundi, Uganda, Kenya, and Rwanda have seen mpox outbreaks. The Democratic Republic of the Congo has been the center of discussion since the country has seen a rise in mpox more than any other country in Africa [6]. The first recorded case of human mpox in DRC was in 1970. Recently, a new variant of the mpox virus is causing outbreaks in parts of the DRC. Based on genetic and geographic distinctions, two variants (viral genomes that can contain one or more mutations) of mpox virus (MPXV) were identified: the "Central African Clade" (clade I) and the "West African Clade" (clade II). MPXV is currently classified into two distinct clades (I and II) and four distinct subclades: Ia, Ib, IIa, and IIb, each with unique characteristics and epidemiological patterns [6]. The spread of MPXV clade Ia in the Equator causes one outbreak, and the other, i.e., clade Ib is causing outbreaks mostly in the provinces of North Kivu, South Kivu, and Kinshasa. The global mpox outbreak in 2022–2023 was caused by the clade IIb strain, which is still spreading throughout Asia, the Americas, and Europe. Based on current data, the Democratic Republic of the Congo is facing two outbreaks. Therefore, our present study focuses on the cases occurring in the Democratic Republic of the Congo.

Compartmental models are a natural approach for studying infectious disease dynamics as this approach gives qualitative insights into disease behavior. Kermack-Mckendric first proposed a compartmental (susceptible–infected–recovered) model using a system of ordinary differential equations (ODEs) [7]. After that, mathematical modeling in epidemiology has come a long way. Compartmental models have also been utilized to understand multi-strain disease dynamics [8–13]. In these works, the authors have focused on disease-free equilibrium, endemic equilibrium, and whether the competitive exclusion principle holds.

The mpox disease modeling has also been performed using compartmental models. Usman et al. [14] developed an SVEIR model (including a vaccinated component) using ordinary differential equations that accounts for a varied incubation period and individual vaccination status. Yuan et al. [15] assessed transmission risks and control strategies for mpox in a metropolitan area. In the works of Peter et al. [16] and Bhunu et al. [17], they have considered ODEs in their work to study the dynamics of mpox. Emeka et al. [18] used a deterministic model. Many studies have been conducted using fractional-order differential equations, the game theoretic approach, and stochasticity [19–23, 46–50]. In these works, the authors have also studied the utility of one or a combination of several optimal control strategies, such as treatment, vaccination, and isolation.

Although compartmental models are widely used in this field, these models have some important drawbacks. One of the drawbacks is the data sources that are used to simulate the compartmental models. They either assume the parameter values or refer to the values from the available literature. However, in several cases, the precise reasoning behind the derivation of the values was unclear, and crucial parameters such as units to evaluate their validity were also lacking [20, 22]. Due to the uncertainty in estimating the parameters, it is possible that the actual magnitude of the parameters differs from the stated values. Therefore, empirical investigations of mpox epidemiology in human populations are essential for estimating more precise parameters. Some research have been conducted where the parameters are obtained from real data. In the works of Massard et al. and Gao et al. [8, 9], they have performed parameter estimation using real-time data. Massard et al. [8] used the particle swarm optimization algorithm to estimate the model parameters using the data of COVID-19. Gao et al. [9] have used the nonlinear least-squares approach to minimize the root-mean-square error (RMSE) between the actual data and the simulated results for the calibration of the model.

The main objective of parameter estimation is to reduce the error between predicted and actual values. However, with time, the parameters related to the disease change. Using pure numerical fitting cannot correctly capture the trend of epidemics in the spreading process. The consideration of this change remains unanalyzed in the estimation process. Time series analysis of the model parameters can be instrumental in analyzing this change. In this regard, deep learning models can be very fruitful.

Deep learning (DL) techniques have been applied for a wide variety of problems such as object detection, weather prediction, motion modeling, image classification, natural language processing, speech recognition, and synthesis [25–28]. There exist several applications for mpox detection as well. Manohar et al. [29] used artificial neural networks to predict the mpox outbreak. Thieme et al. [30], Jaradat et al. [31] used a deep learning algorithm to classify skin lesions from mpox virus infection. Nayak et al. [32] employed deep learning-based architectures to detect the mpox virus using images of skin lesions. Several other applications of artificial intelligence techniques on mpox detection can be found in [33].

Deep learning methods, especially recurrent neural networks (RNN), and long short-term memory neural networks

(LSTM), are used for time series prediction. These methods excel in training with large datasets. Unfortunately, obtaining a large dataset is not feasible at the beginning of an outbreak. So, a way to predict the number of infected individuals using small historical datasets is very much necessary.

Considering that mathematical epidemic models have background knowledge of infectious diseases and deep learning technology has strong predictive capabilities, combining both technologies can be useful for the prediction of infected individuals while having small datasets. Moreover, it may improve the interpretability of deep learning methods and generate more robust predictions.

This paper proposes a two-strain susceptible–infected–recovered–deceased (TSSIRD) model that shows the transmission dynamics of mpox. Then, utilizing the model, we have fed the model parameters in deep learning models, namely Long short-term memory (LSTM) neural network, Bidirectional LSTM (BiLSTM), Stacked LSTM, and two hybrid models Convolution LSTM and Convolution-Bidirectional LSTM for the prediction of the number of infected population. The main contributions of this work are as follows.

- Formulation of a two-strain SIRD epidemic model, which is suitable to model the current scenario of the mpox outbreak in the DRC.
- Utilize the parameters of the TSSIRD model in the deep learning models.
- Obtain a good hybrid model that combines TSSIRD and deep learning architectures and that has performed significantly well on a small dataset compared to purely deep learning models.
- A global sensitivity analysis has shown the most significant parameters for the infected population.

The article continues as follows: A detailed description of the mpox model is given in Sect. 2. In Sect. 3, equilibria are identified and the basic reproduction number is calculated. The prediction process and description of the deep learning models used are provided in Sect. 4. The findings and discussion are provided in Sect. 5. The global sensitivity analysis is performed in Sect. 6. Finally, in the last section, we have provided a brief conclusion and a future scope of this work.

2 Model formulation

For the present research, we have assumed that the total population is of size N. Time-dependent classes are created by dividing the entire population into susceptible $S(t)$, infected $I(t)$, recovered $R(t)$, and deceased $D(t)$. As the focus of our study is DRC where the new outbreak occurred due to another variant of mpox virus, we have divided the infected class into two subclasses, I_a (infected class due to clade Ia) and I_b (infected class due to clade Ib). We have assumed that the birth rate or recruitment rate of the population is Λ. If the susceptible population gets infected by clade Ia, it goes to class I_a. The disease transmission rate of clade Ia is β_a. In a similar manner, the person infected with clade Ib goes to class I_b. The disease transmission rate in this case is considered to be β_b. Due to several disease intervention processes like treatment, vaccination, etc., or natural immunity, some of the infected population go to the recovered class, and some die due to the disease. Let α be the recovery rates from both clade Ia and clade Ib, and γ_a and γ_b be the death rates due to clade Ia and clade Ib, respectively. As there are cases of reinfection of mpox disease in Africa [34], we have assumed that the population is not completely immune to the disease, and the recovered population becomes susceptible. Parameters σ_a and σ_b are the loss of recovery or immunity due to clade Ia and clade Ib, respectively. The parameter μ is considered the natural death rate. Based on the above assumptions, we propose a two-strain SIRD model in (2.1) by means of a system of ordinary differential equations. The schematic representation of the model is provided in Fig. 1. The details of the parameters are given in Table 1.

$$\begin{aligned}
\frac{dS(t)}{dt} &= \Lambda - \frac{\beta_a I_a(t) S(t)}{N} - \frac{\beta_b I_b(t) S(t)}{N} \\
&\quad + (\sigma_a + \sigma_b) R(t) - \mu S(t), \\
\frac{dI_a(t)}{dt} &= \frac{\beta_a I_a(t) S(t)}{N} - \alpha I_a(t) - \gamma_a I_a(t) - \mu I_a(t), \\
\frac{dI_b(t)}{dt} &= \frac{\beta_b I_b(t) S(t)}{N} - \alpha I_b(t) - \gamma_b I_b(t) - \mu I_b(t), \\
\frac{dR(t)}{dt} &= \alpha I_a(t) + \alpha I_b(t) - (\sigma_a + \sigma_b) R(t) - \mu R(t), \\
\frac{dD(t)}{dt} &= \gamma_a I_a(t) + \gamma_b I_b(t),
\end{aligned} \tag{2.1}$$

with the initial conditions, $S(0) \geq 0$, $I_a(0) \geq 0$, $I_b(0) \geq 0$, $R(0) \geq 0$, $D(0) \geq 0$. Moreover,

$$N = S + I_a + I_b + R + D. \tag{2.2}$$

where N is assumed to be asymptotically constant.

3 Basic reproduction number of the TSSIRD model

In this section, we examine the existence of several equilibrium points of the proposed model and compute the basic reproduction number.

Theorem 1 *The system* (2.1) *has a unique mpox-free equilibrium* $E_0 = (\frac{\Lambda}{\mu}, 0, 0, 0, 0)$. *There is a dominance equilibrium corresponding to clade Ia,* $E_a = (\frac{N(\alpha+\mu+\gamma_a)}{\beta_a}, \frac{(\beta_a\Lambda - N\mu(\alpha+\gamma_a+\mu))(\sigma_a+\mu)}{\beta_a(\mu(\alpha+\gamma_a+\mu))+\sigma_a(\mu+\gamma_a)}, 0, \frac{\alpha(\beta_a\Lambda - N\mu(\alpha+\gamma_a+\mu))}{\beta_a(\mu(\alpha+\gamma_a+\mu)+\sigma_a(\gamma_a+\mu)}, D^*)$ *where* D^* *can be obtained using* (2.2), *if and only if* $\mathcal{R}_a > 1$ *where* $\mathcal{R}_a = \frac{\beta_a\Lambda}{\mu N(\alpha+\mu+\gamma_a)}$ *is the basic reproduction number due to clade Ia. Similarly, there is a*

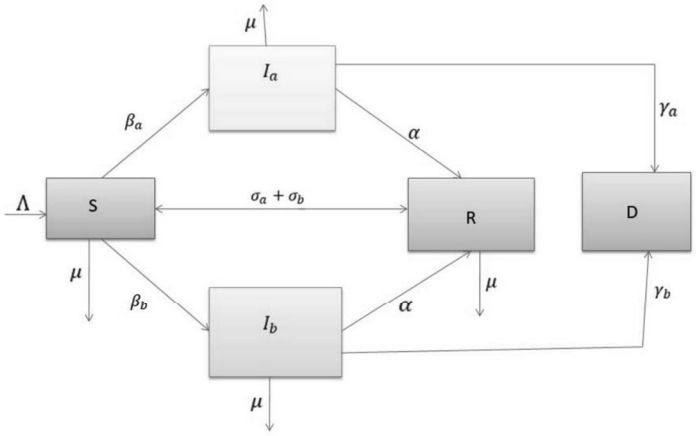


Fig. 1 Schematic diagram of TSSIRD model

Table 1 Parameter summary of the system (2.1)

Used symbol	Interpretation	Unit	Value
Λ	Birth rate/Recruitment rate of human population	week^{-1}	To be estimated from data
β_a	Transmission rate of clade Ia	size$^{-1}\times$ week^{-1}	To be predicted
β_b	Transmission rate of clade Ia	size$^{-1}\times$week^{-1}	To be predicted
α_a	Recovery rate from clade Ia	week^{-1}	To be estimated from data
α_b	Recovery rate from clade Ib	week^{-1}	To be estimated from data
σ_a	Reinfection rate after getting infected by clade Ia	week^{-1}	To be predicted
σ_b	Reinfection rate after getting infected by clade Ib	week^{-1}	To be predicted
γ_a	Death rate due to infection from clade Ia	week^{-1}	To be predicted
γ_b	Death rate due to infection from clade Ib	week^{-1}	To be predicted
μ	Natural death rate	week^{-1}	To be estimated from data

dominance equilibrium corresponding to clade Ib, $E_b = (\frac{N(\alpha+\mu+\gamma_b)}{\beta_b}, 0, \frac{(\beta_b\Lambda - N\mu(\alpha+\gamma_b+\mu))(\sigma_b+\mu)}{\beta_b(\mu(\alpha+\gamma_b+\mu))+\sigma_b(\mu+\gamma_b)}, \frac{\alpha(\beta_b\Lambda - N\mu(\alpha+\gamma_b+\mu))}{\beta_b(\mu(\alpha+\gamma_b+\mu)+\sigma_b(\gamma_b+\mu)}, D^{**})$ *where* D^{**} *can be obtained using* (2.2)*, if and only if* $\mathcal{R}_b > 1$ *where* $\mathcal{R}_b = \frac{\beta_b\Lambda}{\mu N(\alpha+\mu+\gamma_b)}$. *The coexistence equilibrium* $(S^{***}, I_a^{***}, I_b^{***}, D^{***})$ *exists if and only if* $\mathcal{R}_a = \mathcal{R}_b$*, where the basic reproduction number* $\mathcal{R}_0 = max\{\mathcal{R}_a, \mathcal{R}_b\}$.

3.1 Basic reproduction number

Basic reproduction numbers have a critical role in understanding disease dynamics. It is the estimated number of secondary infected cases in a population of susceptible. It is denoted by $\mathcal{R}_0$. A typical scenario is when $\mathcal{R}_0 < 1$, disease vanishes in a population but remains among people when $\mathcal{R}_0 > 1$. There are several approaches in which $\mathcal{R}_0$ can be determined. Here, we intend to determine $\mathcal{R}_0$ using the next-generation matrix method (Driessche et al. [35]). Here, the infected compartment in our system (2.1) is 'I_a' and 'I_b'. Thus, following the approach, the equation of the infected compartment can be written as:

$$\frac{d\zeta}{dt} = \Phi(\zeta) - \Psi(\zeta),$$

where

$$\zeta = (I_a \quad I_b)^T.$$

Here, Φ contains the incidence function $(\frac{\beta_a SI_a}{N} \quad \frac{\beta_b SI_b}{N})^T$, and $\Phi(\zeta) = ((\alpha+\gamma_a+\mu)I_a \quad (\alpha+\gamma_a+\mu)I_b))^T$ The system (2.1) has a disease-free equilibrium (DFE), $(\frac{\Lambda}{\mu}, 0, 0, 0, 0)$. Now, we define F and V as,

$$F = \frac{\partial\Phi}{\partial I_i}, \; V = \frac{\partial\Psi}{\partial I_i}, \; i = 1, \; 2.$$

The basic reproduction number of the system (2.1) is the spectral radius of the matrix FV^{-1} ([35]). In the case of our present model, the basic reproduction number $\mathcal{R}_0$ is given by

$$\mathcal{R}_0 = max\{\frac{\beta_a\Lambda}{N\mu(\alpha+\gamma_a+\mu)}, \frac{\beta_b\Lambda}{N\mu(\alpha+\gamma_b+\mu)}\}.$$

4 Prediction of parameters using deep learning algorithms

In this section, we obtain the parameters from the system of equations (2.1) that will be used in the deep learning models.

4.1 Derivation of the parameters

The parameters of the model are β_a, β_b, γ_a, γ_b, σ_a, σ_b, and α. To consider the changing pattern of the model parameters, we assume the parameters as a function of time. The system (2.1) becomes:

$$\begin{aligned}
\frac{dS(t)}{dt} &= \Lambda - \frac{\beta_a(t)I_a(t)S(t)}{N} - \frac{\beta_b(t)I_b(t)S(t)}{N} \\
&\quad +(\sigma_a(t)+\sigma_b(t))R(t) - \mu S(t), \\
\frac{dI_a(t)}{dt} &= \frac{\beta_a(t)I_a(t)S(t)}{N} - \alpha(t)I_a(t) - \gamma_a(t)I_a(t) - \mu I_a(t), \\
\frac{dI_b(t)}{dt} &= \frac{\beta_b(t)I_b(t)S(t)}{N} - \alpha(t)I_b(t) - \gamma_b(t)I_b(t) - \mu I_b(t), \\
\frac{dR(t)}{dt} &= \alpha(t)I_a(t) + \alpha(t)I_b(t) \\
&\quad -(\sigma_a(t)+\sigma_b(t))R(t) - \mu R(t), \\
\frac{dD(t)}{dt} &= \gamma_a(t)I_a(t) + \gamma_b(t)I_b(t).
\end{aligned} \tag{4.1}$$

To implement the change in the parameters with time in the deep learning model, we reformulate our model into a set of difference equations.

$$\begin{aligned}
S(t+1) - S(t) &= \Lambda - \frac{\beta_a(t)I_a(t)S(t)}{N} - \frac{\beta_b(t)I_b(t)S(t)}{N} \\
&\quad +(\sigma_a(t)+\sigma_b(t))R(t) - \mu S(t), \\
I_a(t+1) - I_a(t) &= \frac{\beta_a(t)I_a(t)S(t)}{N} - \alpha I_a(t) \\
&\quad -\gamma_a(t)I_a(t) - \mu I_a(t), \\
I_b(t+1) - I_b(t) &= \frac{\beta_b(t)I_b(t)S(t)}{N} - \alpha I_b(t) \\
&\quad -\gamma_b(t)I_b(t) - \mu I_b(t), \\
R(t+1) - R(t) &= \alpha I_a(t) + \alpha I_b(t) - (\sigma_a(t) \\
&\quad +\sigma_b(t))R(t) - \mu R(t), \\
D(t+1) - D(t) &= \gamma_a(t)I_a(t) + \gamma_b(t)I_b(t).
\end{aligned} \tag{4.2}$$

Whenever there is no clade Ib in the population, the system (4.2) becomes

$$\begin{aligned}
S(t+1) - S(t) &= \Lambda - \frac{\beta_a(t)I_a(t)S(t)}{N} \\
&\quad +\sigma_a(t)R(t) - \mu S(t), \\
I_a(t+1) - I_a(t) &= \frac{\beta_a(t)I_a(t)S(t)}{N} - \alpha I_a(t) \\
&\quad -\gamma_a(t)I_a(t) - \mu I_a(t), \\
R(t+1) - R(t) &= \alpha I_a(t) - \sigma_a(t)R(t) - \mu R(t), \\
D(t+1) - D(t) &= \gamma_a(t)I_a(t).
\end{aligned} \tag{4.3}$$

Whenever there is no clade Ia in the population, (4.2) becomes

$$\begin{aligned}
S(t+1) - S(t) &= \Lambda - \frac{\beta_b(t)I_b(t)S(t)}{N} + \sigma_b(t)R(t) \\
&\quad -\mu S(t), \\
I_b(t+1) - I_b(t) &= \frac{\beta_b(t)I_b(t)S(t)}{N} - \alpha I_b(t) \\
&\quad -\gamma_b(t)I_b(t) - \mu I_b(t), \\
R(t+1) - R(t) &= \alpha I_b(t) - \sigma_b(t)R(t) - \mu R(t), \\
D(t+1) - D(t) &= \gamma_b(t)I_b(t).
\end{aligned} \tag{4.4}$$

We approximate the parameters $\beta_a(t)$, $\sigma_a(t)$, $\gamma_a(t)$ from the system (4.3) and approximate $\beta_b(t)$, $\sigma_b(t)$, $\gamma_b(t)$ from the system (4.4) and obtain

$$\gamma_a(t) = \frac{D(t+1) - D(t)}{I_a(t)}, \tag{4.5}$$

$$\beta_a(t) = \frac{N}{S(t)I_a(t)}[I_a(t+1) - I_a(t) + (\alpha + \mu + \frac{D(t+1) - D(t)}{I_a(t)})I_a(t)], \tag{4.6}$$

$$\sigma_a(t) = -\mu + \frac{\alpha I_a(t) - R(t+1) + R(t)}{R(t)}, \tag{4.7}$$

$$\gamma_b(t) = \frac{D(t+1) - D(t)}{I_b(t)}, \tag{4.8}$$

$$\beta_b(t) = \frac{N}{S(t)I_b(t)}[I_b(t+1) - I_b(t) + (\alpha + \mu + \frac{D(t+1) - D(t)}{I_b(t)})I_b(t)], \tag{4.9}$$

$$\sigma_b(t) = -\mu + \frac{\alpha I_b(t) - R(t+1) + R(t)}{R(t)}. \tag{4.10}$$

When dominance equilibrium exists, according to Theorem 1, $\mathcal{R}_a = \mathcal{R}_b$, i.e., when $\beta_b = \beta_a(\frac{\alpha+\gamma_b+\mu}{\alpha+\gamma_a+\mu})$. So, when both the clades are present, we approximate β_b as $\beta_b(t) = \beta_a(t)(\frac{\alpha+\gamma_b(t)+\mu}{\alpha+\gamma_a(t)+\mu})$ where β_a is given in equation (4.6).

4.2 Data collection

Democratic Republic of the Congo, a country in Central Africa, consists of 26 provinces. Among them, clade Ib has been detected in Sud-Kivu, Nord-Kivu, Haut Katanga, Tanganyika. Clade Ia has been detected in Bas Uele, Equateur, Haut Uele, Kasai Central, Kongo Central, Kwango, Kwilu, Lualba, Maniema, Mongala, Nord Ubangi, Sankuru, Sud-Ubangi, Tshuapa, Kasai Oriental. Both clade Ia and clade Ib have been detected in Kasai, Kinshasa, Tshopo, Maindombe, and no reports were addressed from Haut Lomami, Ituri, Lomami [6]. The epidemic data were collected from January 1, 2023, to December 29, 2024, in a weekly basis. Since the first outbreak, initially, only one clade had affected several regions of Africa. After 2023, another clade was detected and caused a severe outbreak. Generally, two types of cases were

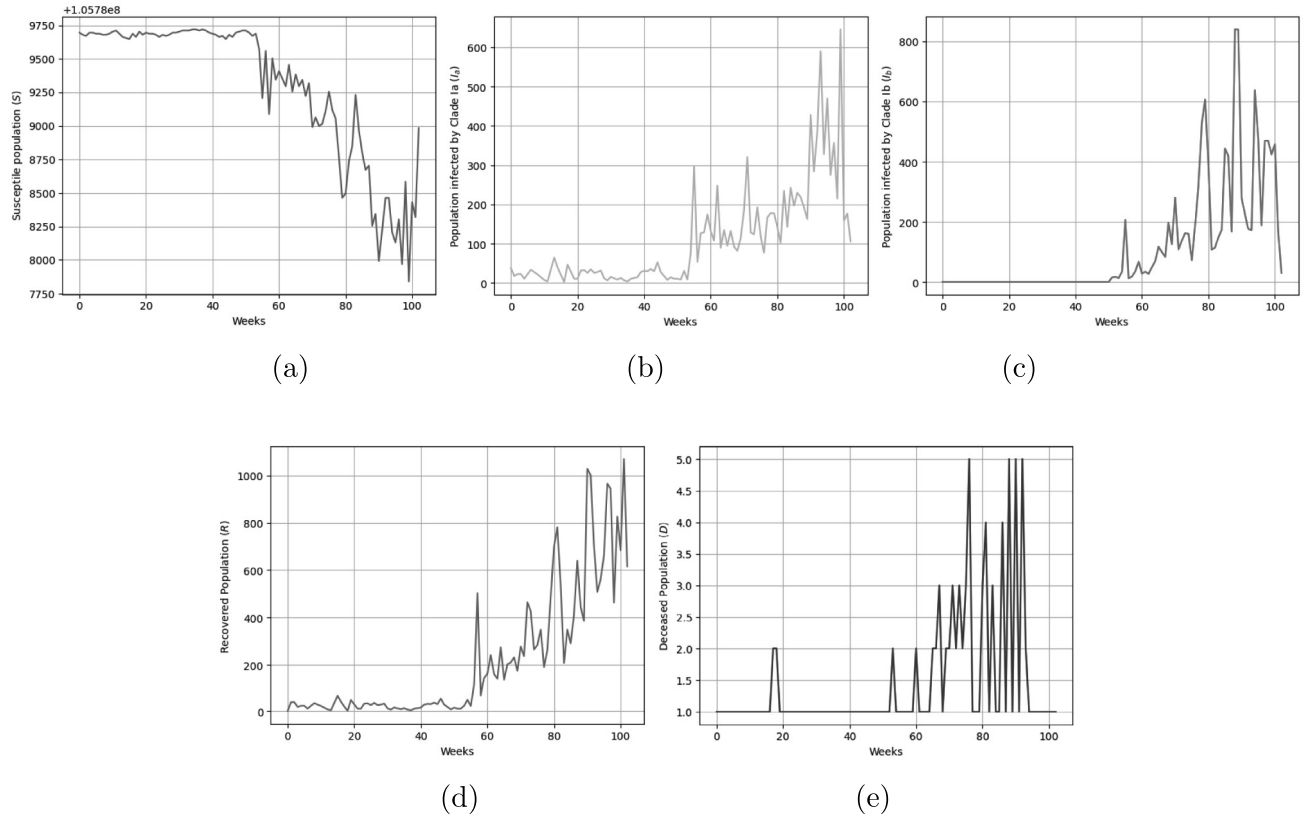


Fig. 2 The data distribution of **a** Susceptible **b** Infected (Clade Ia) **c** Infected (Clade Ib) **d** Recovered **e** Deceased class

reported (i) Suspected case and ii) Confirmed case. The definition of suspected cases and confirmed cases can be found in [6].

In this study, we have considered confirmed cases only as the reporting of suspected cases was closed on November 17, 2024. Any case that has been clinically and epidemiologically diagnosed with mpox and confirmed in the laboratory.

For the numerical simulations performed throughout the remainder of the paper, we consider the following initial population sizes:

I) According to the World Health Organization (WHO), as of 2023 in the Democratic Republic of the Congo, the current population is 105,789,731 [36]. So, we consider $N = 105789731$.

II) The values of I_a and I_b and D are obtained from [6].

III) To obtain the values of R, we have assumed that the infected population recovers within 14 days [34]. The value of R is obtained using the relation $R = I_a + I_b - D$.

IV) The value of S is obtained using the relation $S = N - (I_a + I_b + R + D)$.

V) According to the World Bank data, the crude birth rate of the DRC is 41.74 per 1,000 people in 2022 [37]. So, for the present model, the birth rate of DRC is obtained as $\Lambda = 84,916.2$.

VI) As there is no reported distinction in the recovery time between the clades, we assume they are equal and the recovery day is 14 days. So, $\alpha = \alpha_a = \alpha_b = \frac{1}{14}$.

VII) Whenever there is no disease in the population, we can write $N = \frac{\Lambda}{\mu}$ from Theorem 1. So, μ can be determined as $\mu = \frac{\Lambda}{N} = \frac{84916}{105789731} = 0.00008$.

The several compartments, namely susceptible, infected (by clade Ia), infected (by clade Ib), recovered, and deceased, are shown in Fig. 2. The figures show the change in the susceptible, infected, recovered, and deceased over the stipulated time. The parameters obtained from equations (4.5)-(4.10) are also plotted. The change in the parameters $\beta_a, \beta_b, \gamma_a, \gamma_b, \sigma_a, \sigma_b$ with time is shown in Fig. 3.

4.3 Data augmentation

Data augmentation is the process of utilizing preexisting data to create modified copies of a dataset in order to artificially expand the training dataset. Small training sets are not enough to learn the parameters of deep learning algorithms and are known to cause overfitting. With data augmentation, it is possible to create variants of the training sample without altering the underlying data's meaning. Data augmentation increases the diversity of training data, enhancing the model's generalizing capabilities. There exist several traditional data

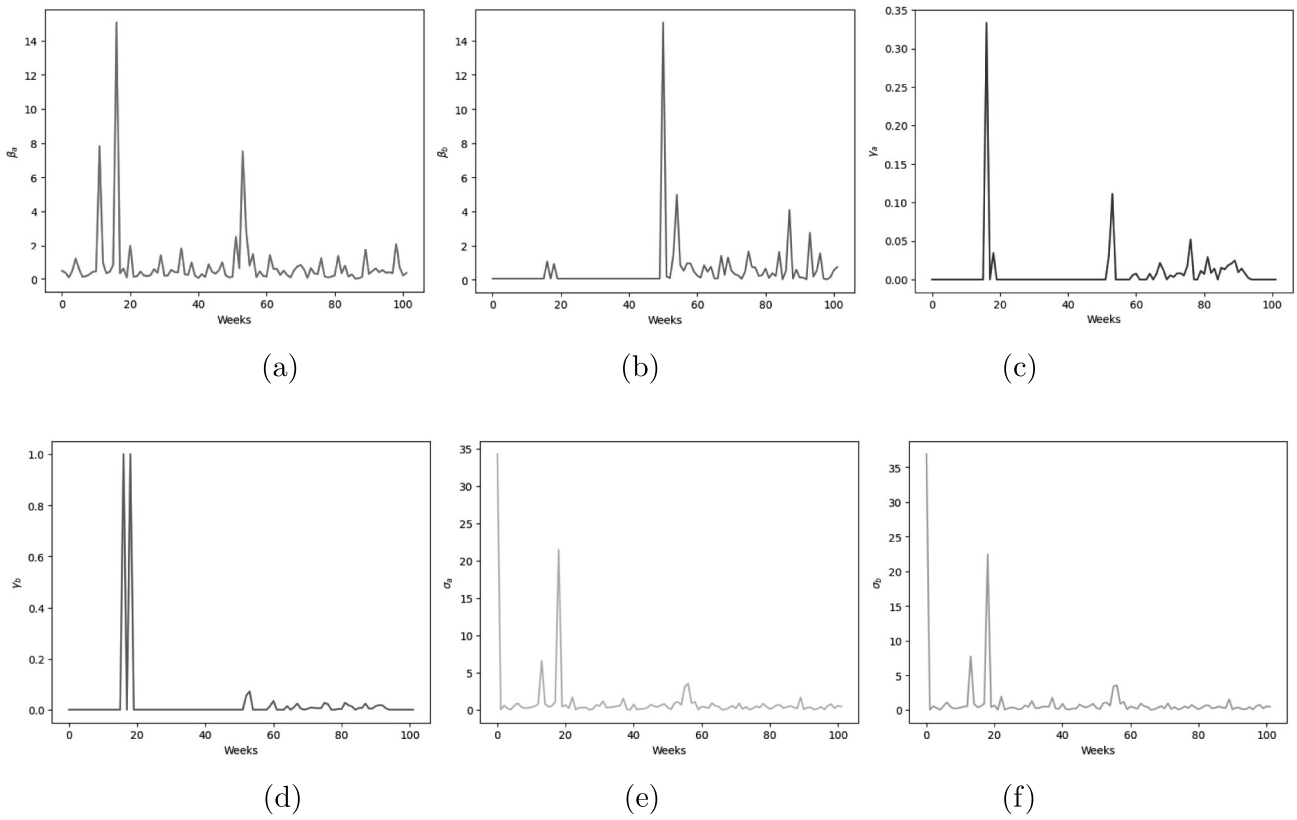


Fig. 3 The data distribution of **a** β_a **b** β_b **c** γ_a **d** γ_b **e** σ_a **f** σ_b

augmentation techniques such as jittering, scaling, rotation, and many more [38].

For this study, the data are available on a weekly basis from January 1, 2023, to December 29, 2024. So, due to the limited amount of data, we have augmented some data in the present work by injecting noise into the original data. Data augmentation through noise addition involves injecting controlled noise into the time series data. Given an original time series $X = [x_1, x_2, ..., x_T]$, where x_T represents the value of x at time T, and a noise signal $N = [n_1, n_2, ..., n_T]$, where n_T is sampled from a predefined noise distribution, the augmented time series is $X^{'} = X + N$. We have introduced noise from a Gaussian distribution with a standard deviation of 0.05 and a mean of 0 to the original data. The mean and standard deviation define the magnitude and shape of the deformation. Figures of S, I_a, I_b, R, D after data augmentation are shown in Fig. 4. From the figures, we see that the pattern of the data is the same. Only slight variations are incorporated into the data points.

4.4 Deep learning prediction model

In this section, we use the time-dependent TSSIRD model to understand the behavior of the underlying parameters. The parameters can be arranged in a time series using the relations discussed in Sect. 4.1. Using several deep learning algorithms, we form TSSIRD-DL models that will predict the parameters and, eventually, the number of infected individuals. Liao et al. [39] used this approach for the prediction of the number of infections during the COVID-19 pandemic. They have used deep learning architectures such as Vanilla LSTM, Bidirectional LSTM, Stacked LSTM, and GRU in their work. In the present work, we have used LSTM and its variations and two hybrid LSTM networks to obtain a single-day prediction of infected individuals. The schematic diagram of the proposed approach is provided in Fig. 5. In the next section, we give a brief discussion about the deep learning techniques that will be used in our analysis.

4.5 Some deep learning techniques

Recurrent neural networks (RNNs) are artificial neural networks that deal with sequential data. RNN retains the recollection of previous inputs, in contrast to feedforward neural networks. However, training the data with long-term dependency is challenging. Long short-term memory (LSTM) [40] is a variant of RNN. Also, a modification of RNN overcomes the problem of long dependencies in sequences, along with the problem of fading and exploding gradients.

LSTM is composed of three logistic gates: a forget gate, an input gate, and an output gate. The forget gate helps to discard unwanted information from the cell state. Through

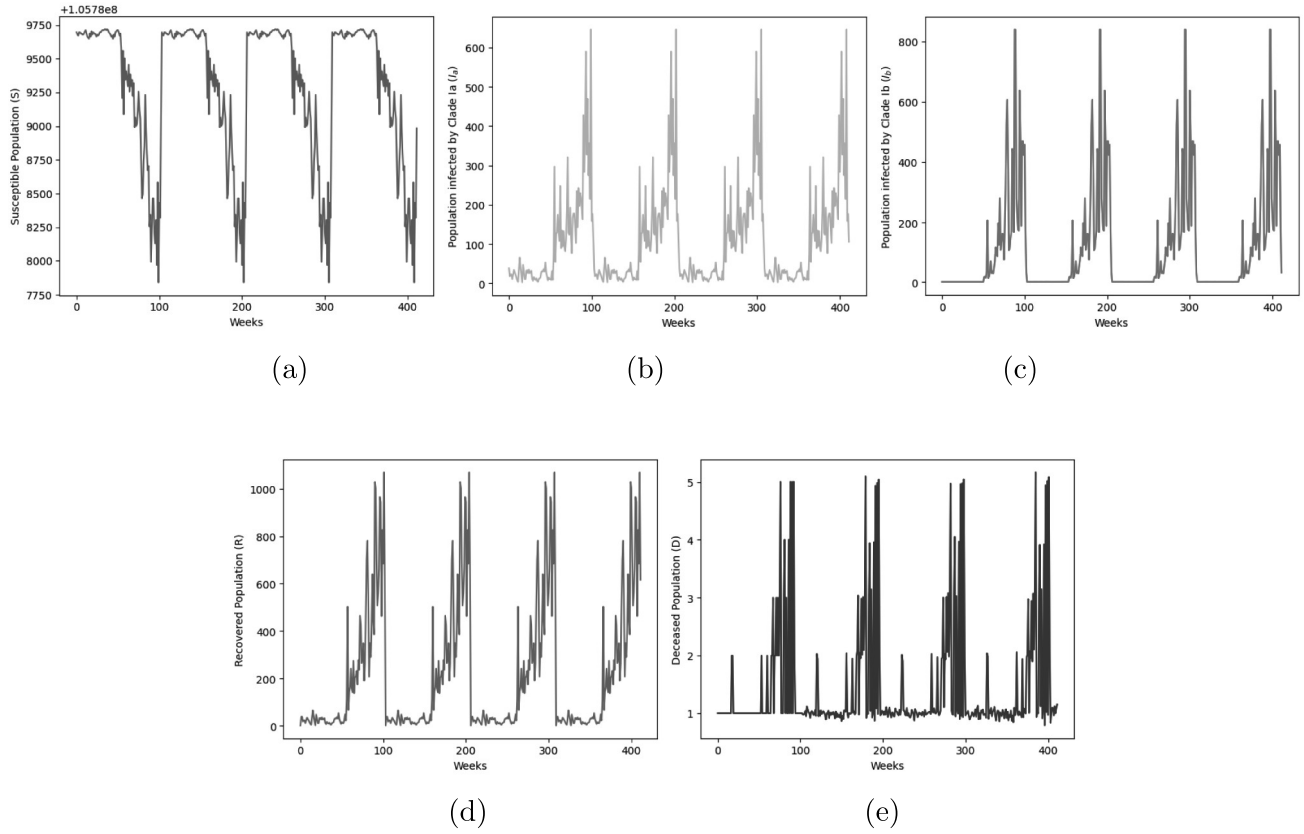


Fig. 4 The data distribution of **a** Susceptible **b** Infected (Clade Ia) **c** Infected (Clade Ib) **d** Recovered **e** Deceased class after performing data augmentation

the input gate, new information is stored in the cell state. Finally, the output gate decides which information to be the output from the cell state.

The Bidirectional LSTM (BiLSTM) [41] is a further modification of the LSTM. The primary distinction between the two is that, while BiLSTM enables additional training by navigating the input data in two directions (forward and backward), improving prediction accuracy in some situations, and LSTM analyzes the data in a unidirectional manner (usually from left to right). The basic layout of the LSTM can be found in [40].

Furthermore, stacked LSTM is used to investigate more deeply and produce more accurate forecasts. We can capture complex temporal connections in the data by stacking deep learning models. It is used to obtain more accurate projections for problems of a similar nature. Although the original LSTM model only has one hidden LSTM layer, the stacked LSTM expands on this architecture by adding many hidden LSTM layers. So, stacking LSTM aids in delving deeper into the architecture and is frequently credited with the method's performance in a variety of difficult prediction tasks.

A convolutional neural network (CNN), on the other hand, is a type of deep learning architecture that is commonly used to classify and recognize particular sections of an image, as well as for object classification and recognition [42]. CNN is made up of multiple layers, each performing a particular task. It is primarily a structure that is mostly composed of convolutional layers that distinguish and identify distinct features, pooling layers that reduce the size of convolved data, and fully connected (FC) layers that link neurons from various layers. However, it may also contain a number of convolutional, pooling, and dropout layers. Depending on the size of the input, the convolution layer can be one-dimensional (1D), two-dimensional (2D), three-dimensional (3D), or even higher dimensional. Pooling layers come in a variety of forms, including global, average, and maximum pooling. However, the most popular of them is max-pooling, which determines the highest value in each feature patch of the map. Furthermore, it should be mentioned that the main advantage of using CNN is its ability to categorize data sets and identify important features without the need for human oversight.

Given that CNN can recognize important features and that LSTM and BiLSTM are efficient at processing time series data, we also use a hybrid deep learning architecture that blends these two deep learning models.

In hybrid models, the input data are first convolved by convolutional layers to create the elementary abstract of the feature map, and the information is subsequently

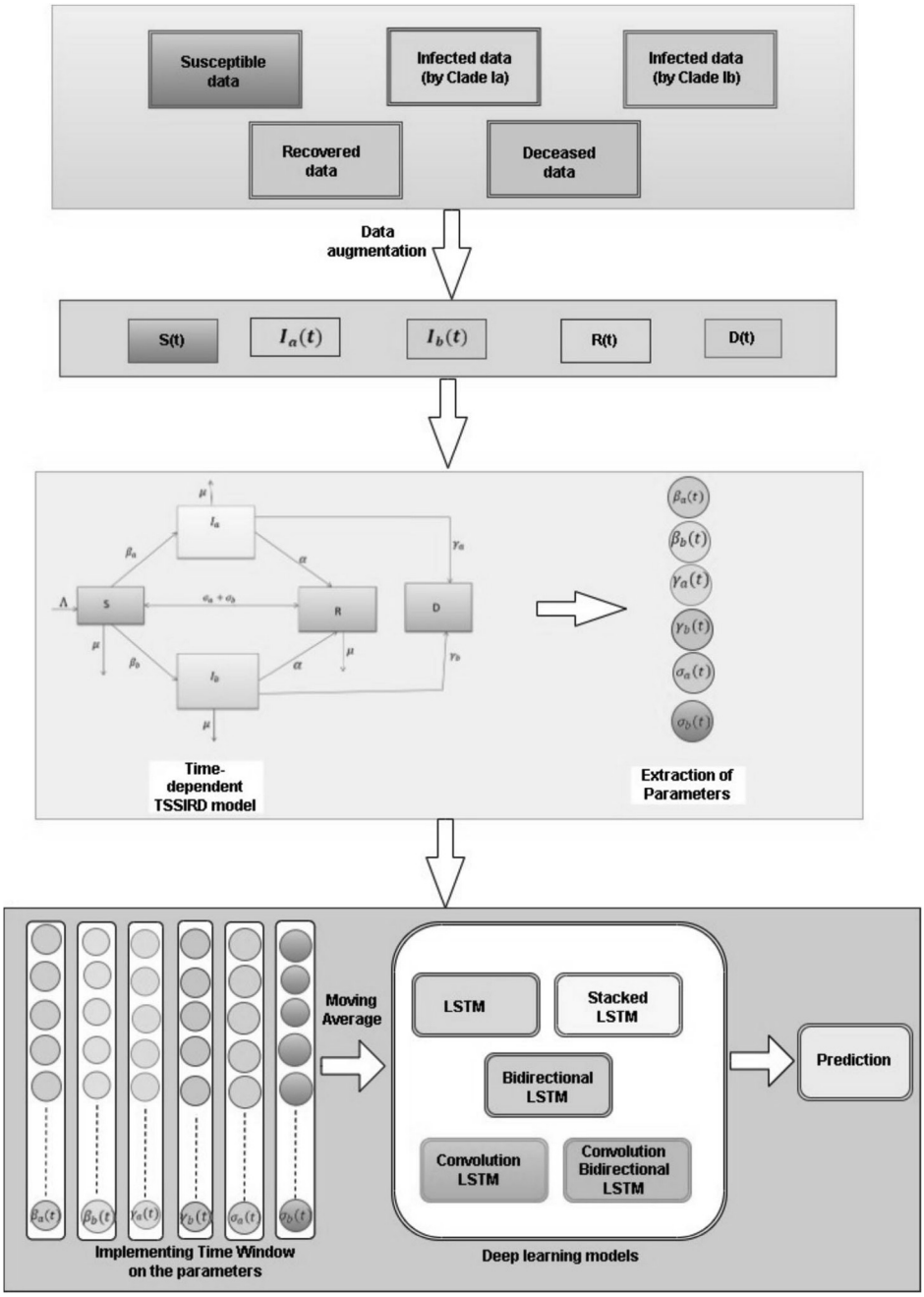


Fig. 5 The overall workflow of the TSSIRD-DL model

down-sampled using the max polling layer to minimize the dimension before being fed into the LSTM and BiLSTM. The processed signal then passes to the FC layer [43], which categorizes the result. Convolutional Long Short-term Memory (ConvLSTM) is an LSTM layer that uses convolutional operations in place of internal matrix multiplication.

In this paper, LSTM, BiLSTM, Stacked LSTM, Convolution LSTM, Convolution BiLSTM are implemented using Keras packages.

The deep learning model's architecture is defined by its hyperparameters. Achieving high-quality models requires appropriate hyperparameter values. To find the ideal model configuration, we tried a variety of hyperparameter combinations. These combinations were produced by varying a number of parameters, including batch size, optimizer selection, and epochs. Each configuration's specifics, including the values applied to these hyperparameters, are listed in Table 2.

By carefully modifying these hyperparameters, we aimed to thoroughly evaluate the impact of each configuration on the model's performance and identify the configuration that most effectively achieved our prediction objectives. To understand

Table 2 Hyperparameters of the models used for training and testing

Hyperparameter	Possible values
Epochs	{50, 80, 100}
Batch size	{10, 15, 20}
Optimizer	{Adam, Adagrad }

the performance, the RMSE values for each scenario are considered. In this experiment, the final combinations used are listed in Table 3.

4.6 Process of prediction

We have the data of $S(t)$, $I_a(t)$, $I_b(t)$, $R(t)$, $D(t)$ for $0 \leq t \leq T-1$, where t is the time variable and T is the total time range. The input of the model $S(t)$, $I_a(t)$, $I_b(t)$, $R(t)$, $D(t)$ is obtained using the process discussed in section (4.1). The model parameters $\{\beta_a(t), \beta_b(t), \gamma_a(t), \gamma_b(t), \sigma_a(t), \sigma_b(t)\}$ for $0 \leq t \leq T-2$ are computed using equations (4.5–4.10).

We use the model parameters $\{\beta_a(t), \beta_b(t), \gamma_a(t), \gamma_b(t), \sigma_a(t), \sigma_b(t)\}$ as the targets of the deep learning model.

Since the parameters continue to change during the propagation of a disease, we have constructed a time window of size W as the dimension of the input time series data for the deep learning models.

To remove the effect of noise in the data, we have used the moving average with a small sliding window to the parameters of the model.

The optimal hyperparameters are obtained by training and testing the model. The prediction algorithm is shown below.

Algorithm of prediction

Input: $\{S(t), I_a(t), I_b(t), R(t), D(t)\}$
1. Preprocessing of data.
2. Derive the input for the TSSIRD model: $\{S(t), I_a(t), I_b(t), R(t), D(t)\}$ using the process discussed in Section 4.1.
3. Augment the data by adding noise from the Gaussian distribution.
4. Calculate the model parameters $\beta_a(t), \beta_b(t), \sigma_a(t), \sigma_b(t), \gamma_a(t), \gamma_b(t)$ using equations (4.5)-(4.10).
5. Implement moving average with window size W = 4 using $x_i^* = \frac{1}{W}\Sigma_{i=1}^{W} x_{t-i+1}$.
6. Normalize the parameters using the robust scaling normalization technique given by $x_{new} = \frac{x - x_{median}}{IQR}$, where IQR is the interquartile range.
7. Split the train and test data set into a 75:25 ratio.
8. Train and evaluate the model using deep learning models.
9. Obtain the predicted parameters $\beta_a(t), \beta_b(t), \sigma_a(t), \sigma_b(t), \gamma_a(t), \gamma_b(t)$
10. Compute the predicted infected values (I_a & I_b) using the second and third equations of (4.2), respectively.

Output: The predicted values of I_a & I_b.

4.7 Model evaluation metrics

This study uses three evaluation metrics to make fair and effective comparisons. They are, respectively, the root-mean-square error (RMSE), mean absolute error (MAE), and determination coefficient (R2). The mathematical formulations are in the following.

$$RMSE = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(\tilde{y}_i - y_i)^2}, \tag{4.11}$$

$$MAE = \frac{1}{N}\Sigma_{i=1}^{N}|y_i - \tilde{y}_i|, \tag{4.12}$$

$$R2 = 1 - \frac{\Sigma_{i=1}^{N}(\tilde{y}_i - y_i)^2}{\Sigma_{i=1}^{N}(\bar{y}_i - y_i)^2}. \tag{4.13}$$

where y_i are the actual observation, $\tilde{y}_i$ are the predicted observations, $\bar{y}$ is the mean of the actual observation, and N is the total observations.

5 Results and discussion

The experiments are conducted using open-source libraries such as NumPy, Pandas, TensorFlow, and Keras. We have used Python programming language for the learning of the deep neural networks. The hyperparameters used for the present task are given in Table 3.

In this section, we show the results derived from the TSSIRD-DL models. We have applied five deep learning models, namely LSTM, BiLSTM, Stacked LSTM, Convolution LTSM, and Convolution BiLSTM. We have shown the training and test results of these models for a single-day prediction in Tables 4 and 5, respectively. In Table 4, single-day prediction results of the infected population due to clade Ia are provided, and in Table 5, prediction results of the number of infected population due to clade Ib are presented.

From Tables 4 and 5, we see the RMSE, MAE, and R2 values of the five models. In the case of the prediction result of the number of infected due to clade Ia, the RMSE and MAE values (for the test set) of the TSSIRD-Convolution BiLSTM are 35.62 and 23.184, which is the lowest from the rest of the four models. In this case, the R2 value is 0.928, which means that this model can capture 92.8 % of the variance of the target set, which is the highest among the rest of the models. From Table 5, we can see that the RMSE and MAE values for the TSSIRD-Convolution LSTM model are 33.948 and 17.28 which is the lowest among others. In addition, the R2 score for this model is 97.2 %, which is the highest among the other models. Hence, the TSSIRD-Convolution LSTM model and the TSSIRD-Convolution BiLSTM model perform better than the other three models for both cases. So, in this scenario, CNN-LSTM hybrid models are cap-

Table 3 Configuration of the used deep learning models

Models	Configuration	Total number of Parameters	Epochs	Batch size	Optimizer
LSTM	2 LSTM layers, 1 dropout layer, 1 dense layer	5,825	80	10	Adam
BiLSTM	1 Bidirectional layer, 1 dropout layer, two dense layers, and activation function ReLU	13,089	80	15	Adam
Stacked LSTM	2 LSTM layers, 1 dropout layer, 1 dense layer	13,089	80	15	Adam
Convolution LSTM	1 1D Convolution layer, 1 Max Pooling layer, 3 Dropout layers, 1 Flatten layer, 1 Repeat vector, 2 LSTM layers, 2 Dense layers with activation function ReLU	5,697	80	15	Adam
Convolution BiLSTM	Two 1D Convolution layers, 1 Max Pooling layer, 2 dropout layers, 1 Flatten layer, 1 bidirectional layer, 2 dense layers with activation function ReLU	19,969	80	15	Adam

Table 4 Single-day prediction result for the clade Ia infected population

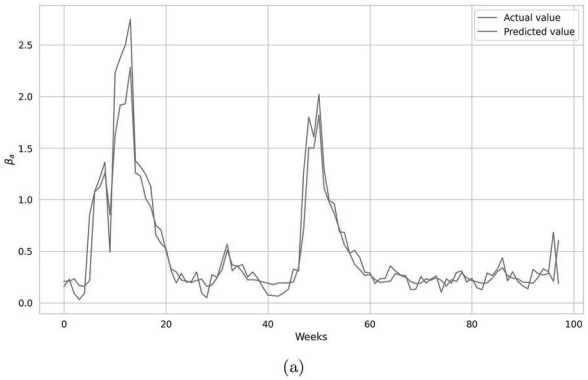
Model	Evaluation metric					
	Train			Test		
	RMSE	MAE	R2	RMSE	MAE	R2
TSSIRD-LSTM	37.036	23.161	0.9313	38.402	24.213	0.9167
TSSIRD-BiLSTM	38.611	24.047	0.9264	40.427	25.379	0.9091
TSSIRD-Stacked LSTM	37.967	24.213	0.9285	41.971	25.854	0.9045
TSSIRD-Convolution LSTM	35.223	21.511	0.9365	38.956	23.045	0.9139
TSSIRD-Convolution BiLSTM	35.95	22.551	0.9371	35.62	23.184	0.928

Table 5 Single-day prediction result for the clade Ib infected population

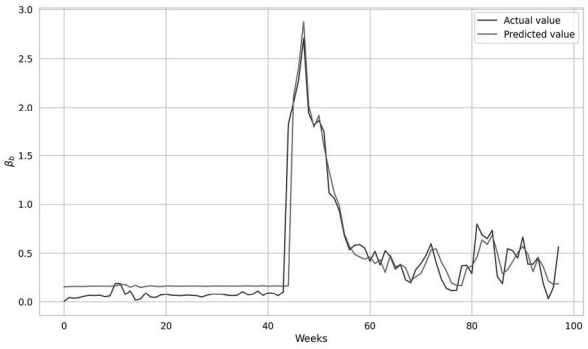
Model	Evaluation metric					
	Train			Test		
	RMSE	MAE	R2	RMSE	MAE	R2
TSSIRD-LSTM	44.626	19.09	0.9557	49.71	21.49	0.9449
TSSIRD-BiLSTM	46.185	19.85	0.9536	49.44	21.95	0.946
TSSIRD-Stacked LSTM	43.86	20.82	0.9594	46.106	21.29	0.9538
TSSIRD-Convolution LSTM	36.136	18.124	0.9708	33.948	17.28	0.972
TSSIRD-Convolution BiLSTM	37.98	18.64	0.968	39.98	19.17	0.9636

turing more insights than non-hybrid deep learning models. Among the two hybrid models, we see that the TSSIRD-Convolution LSTM model is giving better results than the TSSIRD-BiLSTM model in predicting infected cases due to clade Ib. However, TSSIRD-BiLSTM performs better in predicting the infected population due to clade Ia. In Fig. 6, we present the test results of the single-day prediction of the parameters of the TSSIRD model β_a, β_b, γ_a, γ_b, σ_a, and σ_b obtained from the TSSIRD-Convolution LSTM model. In Fig. 7, we show the test result of the single-day prediction model of the infected population due to clade Ia and clade Ib obtained from the TSSIRD-Convolution LSTM model. In the figures, we can observe the actual and predicted values of the TSSIRD model parameters.

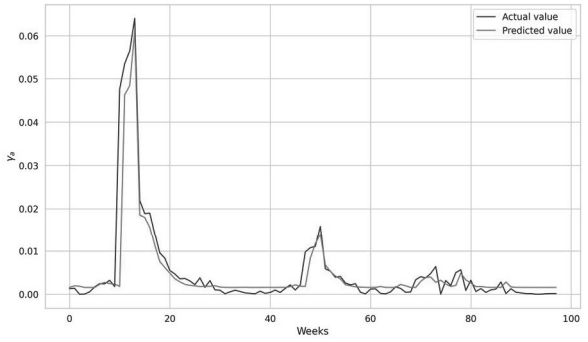
In addition, we compare the single-day prediction result obtained from the purely deep learning models with the proposed TSSIRD-DL models. For the purely deep learning models, we have augmented the same number of data with the original data using a similar technique discussed in Section 4.5. In Tables 6 and 7, we present the test results of the deep learning models with the TSSIRD-DL models, where the configuration of the pure deep learning models is kept the same as in Table 3. From Table 6, we can observe that the lowest RMSE value is 35.626 which occurs in the test result of the



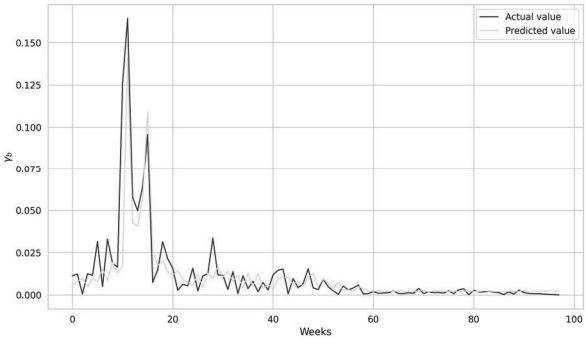
(a)



(b)

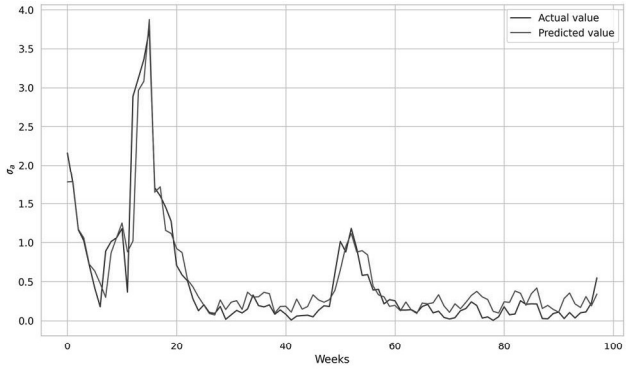


(c)

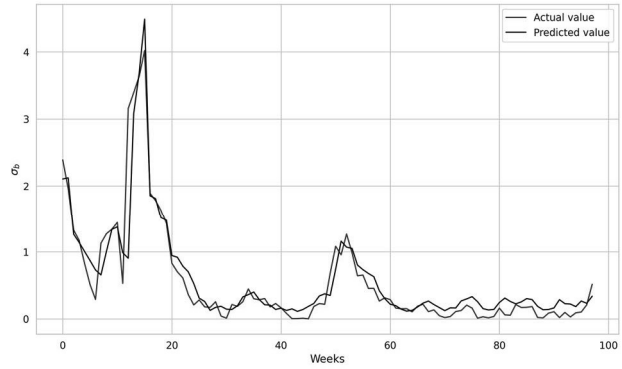


(d)

Fig. 6 The predicted and actual values of the parameters **a** β_a **b** β_b **c** γ_a **d** γ_b **e** σ_a and **f** σ_b obtained from TSSIRD-ConvLSTM model



(e)



(f)

Fig. 6 continued

Table 6 Single-day prediction result for the clade Ia infected population

Model	Evaluation metric		
	Test		
	RMSE	MAE	R2
LSTM	62.982	38.404	0.7633
BiLSTM	56.801	35.707	0.80516
Stacked LSTM	58.228	36.409	0.79521
Convolution LSTM	49.475	31.787	0.85219
Convolution BiLSTM	38.188	21.884	0.9119
TSSIRD-LSTM	38.402	24.213	0.9167
TSSIRD-BiLSTM	40.427	25.379	0.9091
TSSIRD-Stacked LSTM	41.973	25.85	0.9045
TSSIRD-Convolution LSTM	38.956	23.045	0.9139
TSSIRD-Convolution BiLSTM	35.626	23.184	0.928

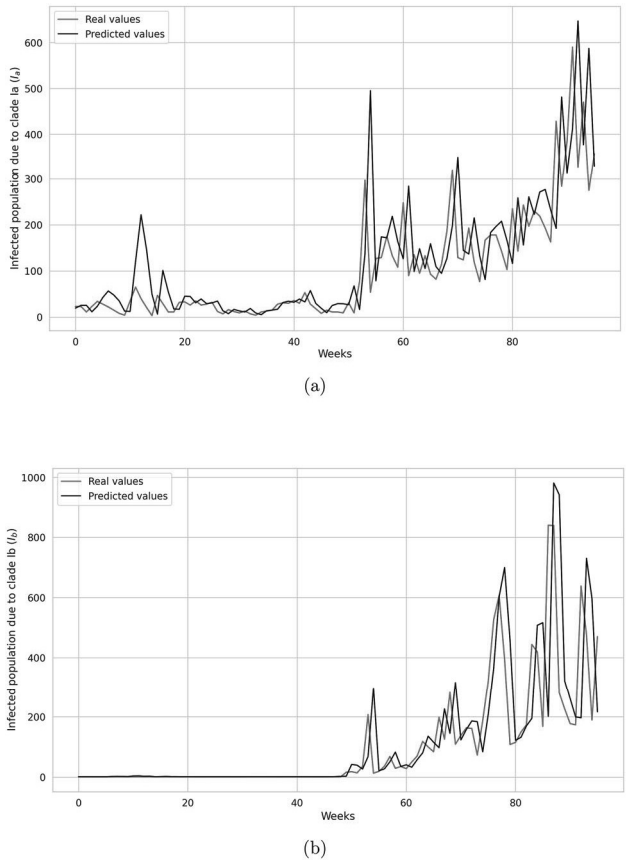


Fig. 7 Single-day prediction results of **a** Infected (Clade Ia) **b** Infected (Clade Ib) using SIRVD-ConvLSTM model

TSSIRD-Convolution BiLSTM model. This model also gives the highest R2 score, which is 0.928. It should be observed that among the purely deep learning models, the hybrid deep learning model Convolution BiLSTM gives an R2 score of 0.9119, an RMSE of 38.188, and an MAE of 21.884. Similarly, from Table 7, we can observe that the lowest values of RMSE and MAE are achieved by the TSSIRD-Convolution LSTM model. In addition, the TSSIRD-Convolution LSTM model captures 97.2% variation of the target set. Among the purely deep learning models, the lowest RMSE and MAE are achieved by the Convolution BiLSTM model. However, the TSSIRD-LSTM model shows better performance than the Convolution BiLSTM model. Hence, we can conclude that all the TSSIRD-DL models perform better than their counterparts in the pure deep learning model.

6 Global sensitivity

Global sensitivity analysis is the process of allocating the uncertainty in each input element throughout its whole range of interest to the uncertainty in outputs. A global sensitivity analysis is one in which the sensitivity is assessed across the whole range of each input element and all the input factors are changed at the same time.

In this section, we perform a global sensitivity analysis of the two different infected compartments I_a, I_b, and two different basic reproduction numbers $\mathcal{R}_a$ and $\mathcal{R}_b$, where we vary all parameters at the same time. To do this, we consider a classical approach that combines the Latin hypercube sampling (LHS) with the partial rank correlation coefficient (PRCC) [44]. The partial rank correlation coefficient (PRCC) is a robust sensitivity measure, ideal to apprehend the nonlinear relationships between the inputs and output. By employing PRCC, we can identify and prioritize specific parameters that exert a significant influence on the reproduction numbers $\mathcal{R}_a$ and $\mathcal{R}_b$ and infected populations I_a and I_b. This method enables a more sophisticated comprehension of the interrelationships among variables. The comprehensive LHS-PRCC method involves several procedural steps:

1. Generate several parameter values in the parameter space within the given range of the parameters using Latin Hypercube Sampling (LHS).
2. Using the parameter values, run the model.
3. Calculate the Partial Rank Correlation Coefficient(PRCC) for the output and each input parameter.
4. Analyze the PRCC values to understand the most significant parameters.

Here, we have considered the parameters of the model as random variables following the uniform probability distribution. The parameters showing a PRCC value that tends to be +1 or -1 are considered the most significant. We also computed the p values of the parameters. The parameter showing a small p value (<0.05) is considered the most significant. For our analysis, the initial parameter values are taken from the prediction result obtained from the TSSIRD-ConvLSTM model, which are $\Lambda = 84916$, $\beta_a = 0.675$, $\beta_b = 0.114$, $\sigma_a = 0.507$, $\sigma_b = 0.4165$, $\gamma_a = 0.00088$, $\gamma_b = 0.0023$, $\mu = 0.00008$, $\alpha = 1/14$. We have assumed that the parameters follow a uniform distribution in the proposed range of the parameters. $\Lambda \in [84900, 85000]$, $\beta_a \in [0.387, 2]$, $\beta_b \in [0.1, 0.2]$, $\sigma_a \in [.3, 0.7]$, $\sigma_b \in [0.15, 0.45]$, $\gamma_a \in [0.0004, 0.001]$, $\gamma_b \in [0.0015, 0.003]$, $\mu \in [0.00007, 0.00009]$, $\alpha \in [1/28, 1/12]$.

In Table 8, we present the PRCC and p values of I_a, I_b and $\mathcal{R}_a$, and $\mathcal{R}_b$. In Fig. 8(c) and 8(d), PRCC and p values of I_a and I_b are given. From the PRCC and p values of I_a and I_b, we can observe that α is taking the highest value. However, the PRCC value of α for I_a is 0.08, which is itself a small number. A similar scenario can also be observed for I_b where α is taking the value 0.27. However, if we observe the PRCC and p values of the parameters for $\mathcal{R}_a$ and $\mathcal{R}_b$ (Fig. 8(a) and 8(b)), a clear relationship of β_a and α with I_a is visible. We see that β_a and α are the most significant with $\mathcal{R}_a$. β_a

Table 7 Single-day prediction result for the clade Ib infected population

Model	Evaluation metric		
	Test		
	RMSE	MAE	R2
LSTM	86.293	46.121	0.789
BiLSTM	80.326	37.066	0.817
Stacked LSTM	84.598	40.711	0.797
Convolution LSTM	54.833	31.898	0.915
Convolution BiLSTM	44.098	25.167	0.945
TSSIRD-LSTM	49.721	21.49	0.945
TSSIRD-BiLSTM	49.448	21.95	0.946
TSSIRD-Stacked LSTM	46.106	21.29	0.953
TSSIRD-Convolution LSTM	33.948	17.2883	0.972
TSSIRD-Convolution BiLSTM	39.98	19.17	0.964

Table 8 PRCC and p values of I_a, I_b, $\mathcal{R}_a$ and $\mathcal{R}_b$

Parameters	PRCC (I_a)	p value (I_a)	PRCC ($\mathcal{R}_a$)	p value ($\mathcal{R}_a$)	PRCC (I_b)	p value (I_b)	PRCC ($\mathcal{R}_b$)	p value ($\mathcal{R}_b$)
Λ	0.04106	0.3639	0.0649	0.15098	−0.00672	0.88179	0.0649	0.15908
β_a	0.0112	0.8035	0.9576	1.7535E-266	−0.0335	0.4588	0.0426	0.3461
β_b	−0.0173	0.70126	−0.0786	0.0817	0.0671	0.1376	0.9579	2.9519E-267
γ_a	0.0189	0.6757	0.0224	0.6191	0.0335	0.4577	−0.0682	0.13108
γ_b	0.0407	0.3671	−0.0352	0.4352	0.0796	0.0779	−0.0632	0.161455
σ_a	−0.0173	0.70126	0.0126	0.7792	−0.03267	0.47001	0.00144	0.9745
σ_b	0.048	0.2782	0.0232	0.6067	0.0299	0.50825	−0.05204	0.2497
α	0.0405	0.3697	−0.96464	2.447E-285	0.3971	5.326E-20	−0.9802	0
μ	0.055	0.2234	0.01308	0.7723	−0.041522	0.358	−0.0481	0.2867

is positively correlated with $\mathcal{R}_a$, and α is negatively related to $\mathcal{R}_a$. A similar relation is also observed for $\mathcal{R}_b$. The most significant parameters for $\mathcal{R}_b$ are β_b and α. For the infected population I_b, α is the most significant parameter.

7 Conclusion

mpox, which is historically endemic to Africa, experienced its greatest outbreak in 2022, spreading to several parts of the world and posing a concern for public health. In this work, we have investigated the mpox cases of the Democratic Republic of the Congo. We formulate a two-strain susceptible–infected–recovered–deceased (TSSIRD) model using a system of ordinary differential equations. We predict the parameters derived from the TSSIRD model that incorporate deep learning models such as LSTM, BiLSTM, Stacked LSTM, ConvLSTM, ConvBiLSTM. To the best of our knowledge, the present work is the first work in which mpox disease modeling is done considering two infected classes and a deceased class. This work can be helpful for a single-day prediction of the disease when we have small historical data sets. This work has used real-time data that shows the relevance of this method in the real world. In addition, we performed a global sensitivity analysis of model parameters using PRCC and p values. We have found that the disease transmission rate plays an important role in the disease dynamics for both clades in the DRC, which is very natural. Most importantly, the parameter α, the rate of recovery, shows a strong correlation with both the basic reproduction numbers. From the values of the predicted parameters, we can say that in the DRC, the disease transmission rate for clade Ia is greater than the disease transmission rate for clade Ib. So, even if the latest mpox outbreak is caused by clade Ib, the capability of the clade Ia variant should not be neglected. On the other hand, the mpox outbreak due to clade Ib is caused by sexual contact within densely populated areas [45]. The comparatively low transmission rate of the clade Ib variant obtained from our study implies that the outbreak due to clade Ib can be controlled with vigilance. Hence, government officials and medical practitioners must focus on intervening in both clades in the DRC. In addition, awareness drives

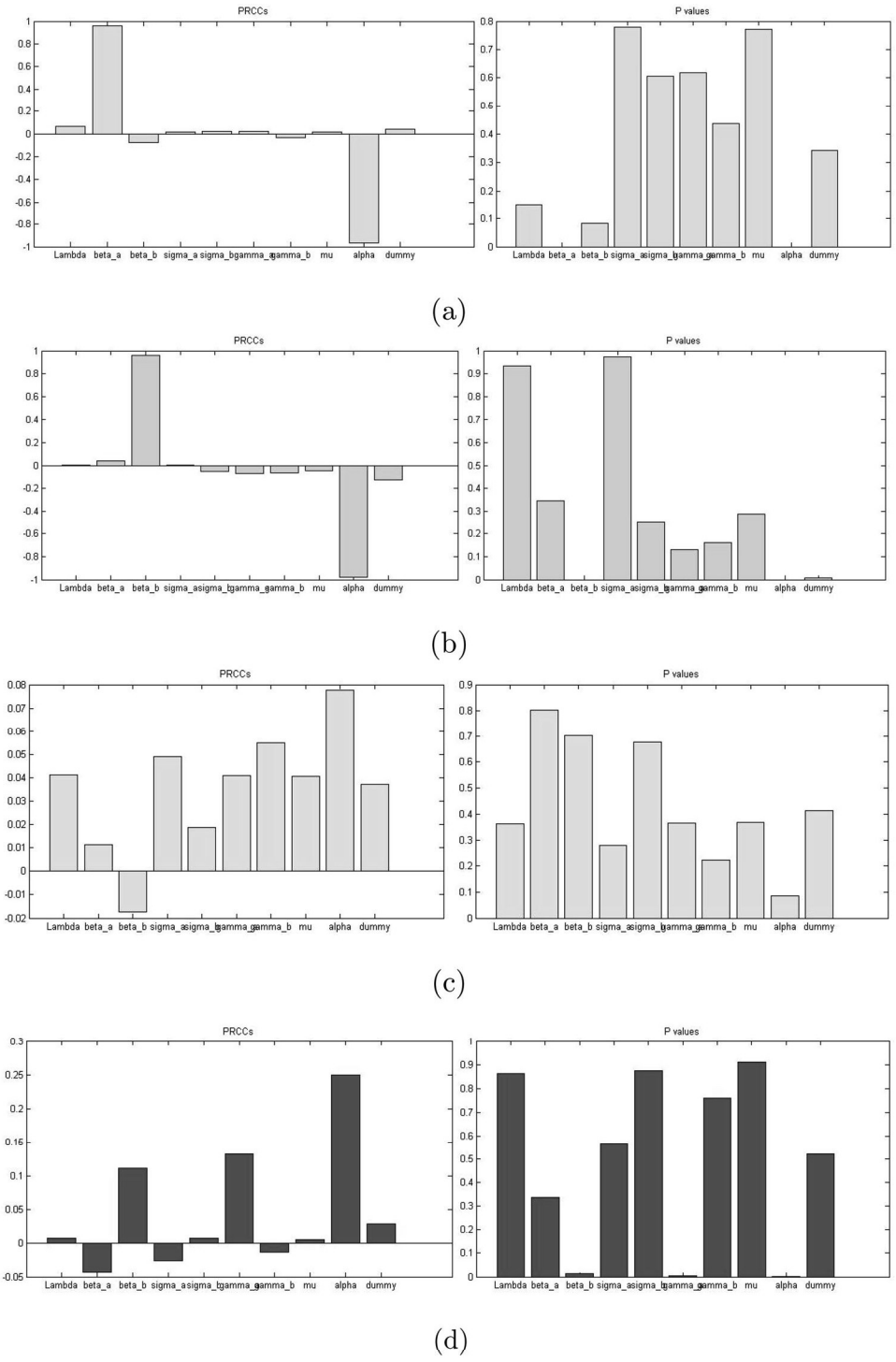


Fig. 8 Bar plot of PRCC and p values of **a** $\mathcal{R}_a$, **b** $\mathcal{R}_b$, **c** I_a and **d** I_b

among people should be conducted so that mpox infection can be controlled. The authorities should also focus on the fact that reinfection is possible. So, mass vaccination should be imposed so that the disease can be eradicated.

In this research, we have obtained a hybrid model that combines TSSIRD and deep learning architectures and that has performed significantly well in a small dataset compared to purely deep learning models. It is known that obtaining and managing large datasets requires time and funding. This methodology can solve the issue of data scarcity and time complexity. This methodology can be applied to the prediction task of several other diseases, such as Nipah, Ebola, and many infectious diseases, for which researchers have difficulty obtaining real-time data. In our future work, we

can generalize the proposed TSSIRD model, which will be helpful in describing multi-strain epidemic models. From the sensitivity analysis, we have seen that α is a crucial parameter for the disease dynamics. Hence, it requires detailed work on the effect of this parameter on the propagation of mpox disease. That is another goal that we have for the future.

Funding The work is supported by the National Science Chair, Science and Engineering Research Board-Department of Science and Technology (SERB-DST), Government of India, (National Science Chair Mode-1) under Grant SB/NSC/2019/01.

Data Availability The data used in this work can be accessed from [6]

Code Availability https://github.com/sayani9931/A-mpox-deep-learning-prediction-frame-work-incorporating-a-two-strain-SIRD-epidemic-model-.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval This research is carried out with integrity and honesty. All procedures adhered to relevant ethical guidelines and regulations.

References

1. Kluge H, Ammon A (2022) Monkeypox in Europe and beyond - tackling a neglected disease together. Euro Surveill 27(24):2200482. https://doi.org/10.2807/1560-7917.ES.2022.27.24.2200482
2. https://www.who.int/news/item/28-11-2022-who-recommends-new-name-for-monkeypox-disease
3. https://ncdc.mohfw.gov.in/wp-content/uploads/2024/08/CD-Alert-mpox-August-2024.pdf
4. Wenham C, Eccleston-Turner M (2022) Monkeypox as a PHEIC: implications for global health governance. Lancet 400(10369): 2169–2171. https://doi.org/10.1016/S0140-6736(22)01437-4
5. Our World in Data: https://ourworldindata.org/explorers/monkeypox?
6. https://worldhealthorg.shinyapps.io/mpx_global/#section-fns
7. Kermack WO, McKendrick AG (1932) Contributions to the mathematical theory of epidemics II-The problem of endemicity. Proce Royal Soc London Ser Containing Papers Math Phys Character 138(834):55–83
8. Massard M, Eftimie R, Perasso A, Saussereau B (2022) A multi-strain epidemic model for COVID-19 with infected and asymptomatic cases: application to French data. J Theor Biol 545:111117
9. Gao S, Shen M, Wang X, Wang J, Martcheva M, Rong L (2023) A multi-strain model with asymptomatic transmission: application to COVID-19 in the US. J Theor Biol 565:111468
10. Li X, Bhattacharya S, Gao S (2013) A two-strain epidemic model with differential susceptibility and mutation. J Biol Syst 21(04):1340009
11. Fudolig M, Howard R (2020) The local stability of a modified multi-strain SIR model for emerging viral strains. PLoS ONE. https://doi.org/10.1371/journal.pone.0243408
12. Dang YX, Li XZ, Martcheva M (2016) Competitive exclusion in a multi-strain immuno-epidemiological influenza model with environmental. J Biol Dyn 10(1):416–456
13. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character. (1927) 115(772) 700–721. https://doi.org/10.1098/rspa.1927.0118
14. Usman S, Adamu II (2017) Modeling the transmission dynamics of the monkeypox virus infection with treatment and vaccination interventions. J Appl Math Phys 5:2335–53
15. Yuan P, Tan Y, Yang L, Aruffo E, Ogden NH, Bélair J, Heffernan JM, Arino J, Watmough J, Carabin H, Zhu H (2023) Assessing transmission risks and control strategy for monkeypox as an emerging zoonosis in a metropolitan area. J Med Virol 95(1):e28137
16. Peter OJ, Kumar S, Kumari N, Oguntolu FA, Oshinubi K, Musa R (2022) Transmission dynamics of monkeypox virus: a mathematical modelling approach. Model Earth Syst Environ 8:3423–3434. https://doi.org/10.1007/s40808-021-01313-2
17. Bhunu CP, Mushayabasa S (2011) Modelling the transmission dynamics of pox-like infections. IAENG Int J Appl Math 41:141–9
18. Emeka PC, Ounorah MO, Eguda FY, Babangida BG (2018) Mathematical model for monkeypox virus transmission dynamics. Epidemiology. https://doi.org/10.4172/2161-1165.1000348
19. Majee S, Jana S, Barman S, Kar TK (2023) Transmission dynamics of monkeypox virus with treatment and vaccination controls: a fractional order mathematical approach. Phys Scr. https://doi.org/10.1088/1402-4896/acae64
20. Bankuru SV, Kossol S, Hou W, Mahmoudi P, Rychtá J, Taylor D (2020) A game-theoretic model of monkeypox to assess vaccination strategies. Peer J 8:e9272
21. Peter OJ, Oguntolu FA, Ojo MM, Oyeniyi AO, Jan R, Khan I (2022) Fractional order mathematical model of monkeypox transmission dynamics. Phys Scr 97:084005. https://doi.org/10.1088/1402-4896/ac7ebc
22. Barman S, Jana S, Majee S, Khatua A, Kar TK (2024) Complex dynamics of a fractional-order monkeypox transmission system with saturated recovery function. Eur Phys J Special Topics. https://doi.org/10.1140/epjs/s11734-024-01283-3
23. Khan A, Sabbar Y, Din A (2022) Stochastic modeling of the monkeypox epidemic with cross-infection hypothesis in a highly disturbed environment. Math Biosci Eng 19(12):13560–13581. https://doi.org/10.3934/mbe.2022633
24. Majee S, Jana S, Barman S, Kar TK (2023) Transmission dynamics of Monkeypox virus with treatment and vaccination controls: a fractional order mathematical approach. Phys Scr. https://doi.org/10.1088/1402-4896/acae64
25. Pal SK, Biswas S, Dutta D (2024) Granulated mask RCNN and eye detection index (EDI) for detection and localization of eye of tropical cyclone from satellite imagery Journal of Data. Inf Manage 6(3):255–275
26. Pal SK, Pramanik A, Maiti J, Mitra P (2021) Deep learning in multi-object detection and tracking: state of the art. Appl Intell 51(9):6400–6429. https://doi.org/10.1007/s10489-021-02293-7
27. Pal SK, Kumar DA (2023) Adaptive granulation-based convolutional neural networks with single pass learning for remote sensing image classification. IEEE J Selected Topics Appl Earth Observ Remote Sens 16:57–70. https://doi.org/10.1109/JSTARS.2022.3223180
28. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Assoc Comput Mach. https://doi.org/10.1145/3065386
29. Manohar B, Das R (2022) Artificial neural networks for the prediction of monkeypox outbreak. Tropical Med Infect Disease 7(12):424
30. Thieme AH, Zheng Y, Machiraju G, Sadee C, Mittermaier M, Gertler M et al (2023) A deep-learning algorithm to classify skin lesions from mpox virus infection. Nat Med 29(3):738–747
31. Jaradat AS, Al Mamlook RE, Almakayeel N, Alharbe N, Almuflih AS, Nasayreh A, Gharaibeh H, Gharaibeh M, Gharaibeh A, Bzizi H (2023) Automated monkeypox skin lesion detection using deep

learning and transfer learning techniques. Int J Environ Res Public Health 20(5):4422
32. Nayak T, Chadaga K, Sampathila N, Mayrose H, Gokulkrishnan N, G. M. B, Prabhu S., Swathi K. S., Umakanth S. (2023) Deep learning based detection of monkeypox virus using skin lesion images. Med Novel Technol Devices 18:100243. https://doi.org/10.1016/j.medntd.2023.100243
33. Chadaga K, Prabhu S, Sampathila N, Nireshwalya S, Katta SS, Tan RS, Acharya UR (2023) Application of artificial intelligence techniques for Monkeypox: a systematic review. Diagnostics 13(5):824
34. Hazra A, Zucker J, Bell E, Flores J, Gordon L, Mitjà O et al (2024) mpox in people with past infection or a complete vaccination course: a global case series. Lancet Infect Dis 24(1):1473–3099. https://doi.org/10.1016/S1473-3099(23)00492-9
35. van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180:29–48
36. https://data.who.int/countries/180
37. https://liveprod.worldbank.org/en/indicator/sp-dyn-cbrt-in
38. Iglesias G, Talavera E, González-Prieto Á et al (2023) Data Augmentation techniques in time series domain: a survey and taxonomy. Neural Comput Appl 35:10123–10145. https://doi.org/10.1007/s00521-023-08459-3
39. Liao Z, Lan P, Fan X, Kelly B, Innes A, Liao Z (2021) SIRVD-DL: A COVID-19 deep learning prediction model based on time-dependent SIRVD. Comput Biol Med 138:104868. https://doi.org/10.1016/j.compbiomed.2021.104868
40. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
41. Chen A, Wang F, Liu W, Chang S, Wang H, He J, Huang Q (2020) Multi-information fusion neural networks for arrhythmia automatic detection. Comput Methods Programs Biomed 193:105479. https://doi.org/10.1016/j.cmpb.2020.105479
42. Yan T, Wong PK, Ren H, Wang H, Wang J, Li Y (2020) Automatic distinction between COVID-19 and common pneumonia using multi-scale convolutional neural network on chest CT scans. Chaos, Solitons Fractals 140:110153. https://doi.org/10.1016/j.chaos.2020.110153
43. Chang P, Grinband J, Weinberg BD, Bardis M, Khy M, Cadena G, Su M-Y, Cha S, Filippi CG, Bota D, Baldi P, Poisson LM, Jain R, Chow D (2018) Deep learning convolutional neural networks accurately classify genetic mutations in gliomas. Am J Neuroradiol 39:1201–7. https://doi.org/10.3174/ajnr.A5667
44. Marino S, Hogue IB, Ray CJ, Kirschner DE (2008) A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol 254(1):178–196. https://doi.org/10.1016/j.jtbi.2008.04.011
45. Masirika LM, Udahemuka JC, Schuele L et al (2025) Epidemiological and genomic evolution of the ongoing outbreak of clade Ib mpox virus in the eastern Democratic Republic of the Congo. Nat Med. https://doi.org/10.1038/s41591-025-03582-1
46. Mondal J, Samui P, Chatterjee AN (2022) Modelling of contact tracing in determining critical community size for infectious diseases. Chaos, Solitons Fractals 159:112141. https://doi.org/10.1016/j.chaos.2022.112141
47. Mondal J, Mondal S, Samui P (2023) Simulation and forecasting CTLs response effectualness against global dynamics of SARS-CoV-2/HIV coinfection. Res Control Optim 13:100323. https://doi.org/10.1016/j.rico.2023.100323
48. Mondal J, Mondal S, Samui P (2025) Impact of vaccination and awareness campaign in reducing mpox transmission: a fractional mathematical study. J Appl Nonlinear Dyn. https://doi.org/10.5890/JAND.2025.12.007
49. Khatua A, Kar TK, Nandi SK, Jana S, Kang Y (2020) Impact of human mobility on the transmission dynamics of infectious diseases. Energy Ecol Environ 5(5):389–406. https://doi.org/10.1007/s40974-020-00164-4
50. Majee S, Jana S, Kar TK, Bhunia B (2024) Complex dynamics of a fractional-order delayed epidemic model incorporating waning immunity and optimal control. Eur Phys J Special Topics. https://doi.org/10.1140/epjs/s11734-024-01221-3