THE EUROPEAN
PHYSICAL JOURNAL PLUS

Monkeypox: a model-free analysis

V. R. Saiprasad[1], **R. Gopal**[1], **D. V. Senthilkumar**[2,a], **V. K. Chandrasekar**[1,b]

[1] Department of Physics, Centre for Nonlinear Science and Engineering, School of Electrical and Electronics Engineering, SASTRA Deemed University, Thanjavur 613 401, India
[2] School of Physics, Indian Institute of Science Education and Research, Thiruvananthapuram 695016, India

Abstract Monkeypox is a zoonotic disease caused by a virus that is a member of the orthopox genus, which has been causing an outbreak since May 2022 around the globe outside of its country of origin Democratic Republic of the Congo, Africa. Here we systematically analyze the data of cumulative infection per day adapting model-free analysis, in particular, statistically using the power law distribution, and then separately we use reservoir computing-based Echo state network (ESN) to predict and forecast the disease spread. We also use the power law to characterize the country-specific infection rate which will characterize the growth pattern of the disease spread such as whether the disease spread reached a saturation state or not. The results obtained from power law method were then compared with the outbreak of the smallpox virus in 1907 in Tokyo, Japan. The results from the machine learning-based method are also validated by the power law scaling exponent, and the correlation has been reported.

1 Introduction

Monkeypox is a virus [1, 2] that belongs to the Poxviridae family under the genus orthopox virus, which originated in the Democratic Republic of the Congo, Africa, in the 1970s and has been endemic there in African countries. But its worldwide outbreak has resulted in over 50,000 cases, as of 27th September, 2022, across 50+ countries since May 2022 [1, 3]. It is the second largest outbreak of this century followed by COVID-19 [4]. A study [5] has reported that the withdrawal of smallpox vaccine in the late 1980s could be a reason for this outbreak which is efficient in preventing the monkeypox spread as well. This poses a serious threat as a large number of people could be susceptible to the disease. Moreover, several studies indicated that the efficacy of the smallpox vaccine has been waning posing an increasing probability for the disease to spread. Though monkeypox is severe but not fatal in adults and cures on its own, it is severe among children and young adults who have poor health conditions and malnutrition.

Due to the widespread of COVID-19 around the globe [4], we have seen the possibility for the virus to mutate into more virulent ones and can cause damage to people from personal level to societal scale. So proper prediction of the outbreak needs to be done which might help to monitor the disease spread and to implement appropriate control measures by the health care system in bringing the disease spread under control. A well-augmented prediction of the disease spreading is deemed to be necessary for the benefit of the policymakers and the health care systems to strategize to limit the disease spread. A broad spectrum of model-based techniques such as parameterized compartmental SIR models [6, 7], time-series modeling [8], probabilistic master equation-based models [9], iterative maps, fractal-based models [10, 11] and logistic equation [12] are available to predict the disease spread. These methods can be useful only if we have proper information about the epidemiological parameters. The less amount of data and the lack of information on epidemiological parameters would pose serious limitations on the above methods for their prediction. Fortunately, on the other hand, we have model-free data-driven methods such as statistical modeling [13–15] and the machine learning model [16], which were helpful in prediction.

For our current investigation, we consider the realtime cumulative monkeypox infection per day data from its initial spread to its spread on 27th September 2022 and use the power law [15] to estimate the scaling exponent of the disease spread in nine most infected countries. We find that the USA stands first in the list of nine countries with highest scaling exponent elucidating the degree of the spread of the monkeypox infection. In contrast, small values of the scaling exponent for the countries like Brazil and Peru are observed. Being a virus from the orthopox family, we have also calculated the scaling exponent for the smallpox outbreak in Tokyo, Japan, during 1907–1908 and found a high degree of scaling exponent quite close to that of the scaling exponent for the monkeypox spread in the USA providing a vital clue about the current spread of the monkeypox virus. Further, we have also analyzed the number of infections using the scale-free echo state network (ESN), which is actually a system of modified recurrent neural networks (RNN)

[a] e-mail: skumar@iisertvm.ac.in (corresponding author)

[b] e-mail: chandrasekar@eee.sastra.edu (corresponding author)

[16, 17]. We have predicted the cumulative monkeypox infection, which agrees fairly well with the realtime cumulative spread for the last ten days from 18th September 2022 to 27th September 2022 and forecasted the cumulative infection for the next ten days from 28th September 2022, which is then used to estimate the scaling exponent. The results from the scaling exponents and those from the machine learning technique are found to be in excellent agreement with each other. Usually, the predictions from the machine learning approach are more accurate in prediction than any other models, but the downside of it is that the results cannot be quantified which severely affects the interpretability of the prediction. In this paper, we combine the results from the power law method along with the machine learning approach and bring out the two important features of prediction, namely accuracy and its interpretability. When it comes to forecasting in realtime systems, our study puts together the better predictability of machine learning and quantitative results from power law approach provide interpretability to the results that help health care system in making more tangible predictions.

The plan of the paper is as follows: In Sect. 2, we provide the description of the power law approach and the scaling exponent of monkeypox outbreak for the USA, Spain, Brazil, France, Canada, Peru, Netherlands, UK and Germany were calculated and reported. We also compare the smallpox outbreak of 1907 in Japan [18] with the current monkeypox outbreak. We present the description of ESN network that we used for prediction and forecasted the results along with its interpretation through the power law fit in Sect. 3. Finally, we provide a conclusion in Sect. 4.

2 Power law approach

For our analysis, throughout this paper, we took the data from the website (https://ourworldindata.org/monkeypox) accessed on 27th September 2022. Epidemiological models make the assumption that the number of person $n(t)$ infected by a sick person will grow exponentially over time 't' as

$$n(t) = Bt^{\gamma}, \tag{1}$$

where B is a proportionality constant and γ is the scaling exponent which reveals the growth pattern of the disease in the chosen demography or a particular nation. We determine the range of power law behavior by plotting the cumulative infection per day data corresponding to different countries in a log-log plot. The scaling exponent for each country was then calculated. To illustrate this method, we took the cumulative infection data for the United States of America (USA) as an example. The cumulative cases of the monkeypox outbreak in the USA are plotted in Fig. 1. The two vertical lines (see Fig. 1b) indicate the range used to extract the scaling exponent. The data points are overlaid with the best fit as determined by linear regression.

We have also carried out the aforementioned analysis for the nine most infected countries by the monkeypox, listed in the first column of Table 1. We found the scaling exponent during the initial power law growth for all the nine countries and tabulated the results in the second column of Table 1 against the respective countries. We have calculated the scaling exponent for all the other eight countries for the window (days) where there exists a power law growth similar to that calculated for the USA as in Fig. 1. From the table, one can say that the USA tops the tabulated list of countries with the highest scaling exponent of 4.421, whereas in countries like Brazil and Peru still a significant amount of growth of the disease is yet to occur, which can be inferred from rather smaller values of their scaling exponents. Further, we have also calculated the scaling exponent for the last ten days (from 27th Sep. 2022) from which one can infer whether the disease spread is either in the growth state or slowing down phase or saturated. The scaling exponent calculated for the last ten days was given in the third column of Table 1. It is evident from the table that in all the nine countries the disease spread was significantly dropped with Germany having the lowest scaling exponent $\gamma = 0.219$ elucidating that the disease spread is in the state of getting saturated in Germany. Countries like Germany, the UK, the USA and Spain with significantly low values of the scaling exponent unveil the reduced rate of the spread of the disease, whereas in Brazil scaling exponent only dropped a little revealing that there lies an imminent power law growth. The results from these scaling exponents were also helpful in validating the corresponding results predicted from the machine learning approach.

There were studies that showed the withdrawal of the smallpox vaccination in the late 1980s could be the reason for the outbreak of monkeypox as the efficacy of the smallpox vaccine is found to be around 80 percent against the monkeypox virus also. As it has been withdrawn, the waning of immunity due to the vaccine could be a reason for the emergence of the orthopox virus. We have

Fig. 1 a Total number of monkeypox infections in the USA was plotted as a function of time from 3rd June, 2022 to 27th September 2022. **b** Log–Log plot of number of infections of the USA as a function of time, the two vertical lines indicate the spectrum of scale-free behavior. The data points are overlaid with a linear regression's best fit
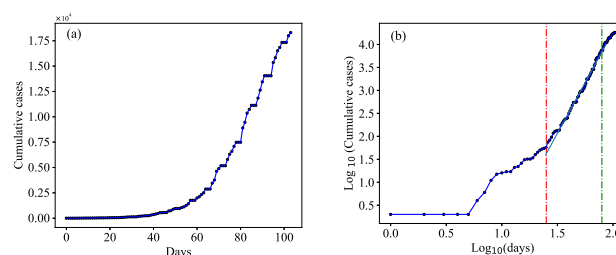
Table 1 Country-wise scaling exponent (γ) for initial power law growth

Country	γ (for initial growth)	γ (for last 10 days)
USA	4.424	1.260
Brazil	3.445	3.091
Peru	2.707	1.492
UK	2.543	0.254
Spain	2.377	0.427
Germany	2.213	0.219
France	2.115	0.308
Netherlands	2.015	0.244
Canada	1.770	0.521

considered the daily disease spread data of smallpox outbreak in Tokyo, Japan, in 1907–1908 (see Fig. 2a) and investigated the similarities among the disease spreading rate of the current monkeypox outbreak with that of smallpox. From the daily infection data, we have calculated the cumulative cases and estimated the power law fit for the time window (days) represented by two vertical dotted lines depicted in Fig. 2b. The scaling exponent for the smallpox outbreak is found to be 4.983 which is quite close to that of the scaling exponent for the monkeypox spread in the USA. This is an important and interesting result, and is also an alarming one as the recent monkeypox spread is similar to the spread of the smallpox virus in Japan in 1907–1908, which accounts for one of the long spanning epidemics causing a significant amount of loss at both individual and societal levels in the human history.

3 Machine learning approach

3.1 Architecture of echo state network (ESN)

In this section, we use a machine learning technique known as reservoir computing, a class of recurrent neural networks called echo state network for prediction [16, 17]. The primary reason for the choice of the ESN over the other schemes is that the echo state network has all the advantages of the recurrent neural network (RNN). Note that ESN also avoids the training issues such as vanishing and exploding gradient problems that are inherent to the traditional RNNs. Furthermore, the training process in ESN is also relatively straightforward since most of the weights are chosen randomly and only once. Only the output weights need to be trained in the ESN, which is not the case in the traditional RNNs. Nevertheless, ESN can imbibe the characteristics of the dynamical systems and capture complicated dynamics throughout time. ESN showed encouraging results of handling multiple different inputs of temporal data in a wide spectrum of works [19, 20]. Its ability to trace the correlation between the sequential data, inspired us to use the power of the ESN to create a method for forecasting the spread of monkeypox outbreak. In this model, for a given input vector there is a reservoir that drives the model to train the weights of the output layer accordingly so as to learn the specific temporal pattern of the input data points. The schematic representation of this ESN model is given in Fig. 3. An ordinary leaky tanh governs the dynamics of the nodes of the ESN, which is represented by the following recursive relation

$$x(t + 1) = (1 - \alpha)x(t) + \alpha. \tanh(W_{\text{res}}x(t) + W_{\text{in}}u(t)). \tag{2}$$

Here, $u(t)$ is a M-dimensional input vector and $x(t)$ is a N_{res}-dimensional vector corresponding to the state of the reservoir nodes at the time instant t. The weights of the internal connections between reservoir nodes are embedded in the matrix W_{res} (dimension: $N_{\text{res}} * N_{\text{res}}$) and the matrix W_{in} (dimension: $N_{\text{res}} * M$) represents the input weight matrix. The leakage constant, denoted by the parameter α, can indeed vary in the range of zero to unity. It is important to remember that the tanh function is the activation function for each node constituting the reservoir. The initial reservoir weight matrix W_{res} is created by uniformly distributed random values in the range $(-1, 1)$.

Fig. 2 **a** Total number of infections for the smallpox outbreak of Tokyo, Japan, in 1907 was plotted as a function of time from 18th December 1907 to 25th July 1908. **b** Log–Log plot of number of infections as a function of time, the two vertical lines indicate the spectrum of scale-free behavior. The data points are overlaid with a linear regression best fit
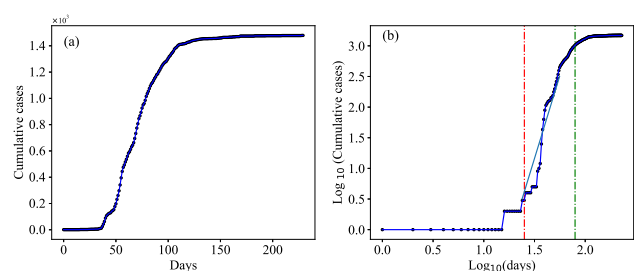
Fig. 3 Schematic depiction of
architecture of Echo state model

Input layer   Dynamical Reservoir   Output layer
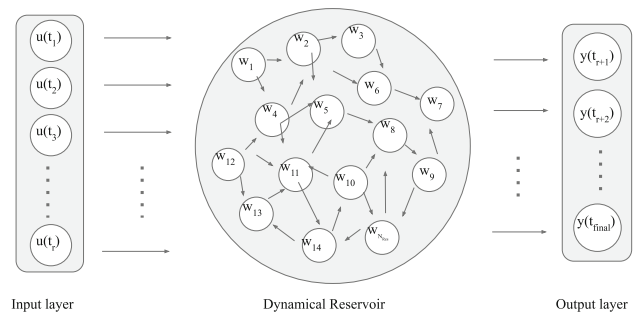
Now, we consider the time-series data consisting of N data points corresponding to the cumulative infection per day for N days. This includes the cumulative infection per day for M days that is to be fed as the input to the reservoir for training and the remaining $(N - M)$ days of the cumulative infection data points are used as a benchmark to test the predicted cumulative infection per day by the model. The initial goal is to determine appropriate input weights of the input layer and connection weights of the reservoir to mimic the specific data pattern embedded in the first M (when $t = 0, 1, \ldots, t_r$) data points of the input corresponding to the cumulative infection per day during the training phase. At each time step, t, the evolution of the input connection weight matrix of the reservoir nodes can be given as

$$\begin{pmatrix} W_{in}(1,1) & \cdots & W_{in}(1,M) \\ W_{in}(2,1) & \cdots & W_{in}(2,M) \\ \vdots & \vdots & \vdots \\ W_{in}(N_{res},1) & \cdots & W_{in}(N_{res},M) \end{pmatrix} \times \begin{pmatrix} I_1(t) \\ I_2(t) \\ \vdots \\ I_M(t) \end{pmatrix},$$

where $I(t)$ is the input data corresponding to the realtime cumulative infection per day for M days. In the training phase, at each instant t, the reservoir state $x(t)$ and input $u(t)$ are accumulated in $V(t) = [1; u(t); x(t)]$. The output vector $Y(t)$ can be written in the form as

$$Y(t) = W_{out}V(t), \tag{3}$$

where $Y(t)$ is a column vector of length $(1 + M + N_{res})$ and $V(t)$ is the column vector of length of $(1 + M + N_{res})$, which is represented as

$$V(t) = \begin{pmatrix} 1 \\ I_1(t) \\ I_2(t) \\ \vdots \\ I_M(t) \\ x(1,t) \\ x(2,t) \\ \vdots \\ x(N_{res},t) \end{pmatrix}. \tag{4}$$

Here t is the time duration that runs from $0, 1, \ldots, t_r$. Ridge regression method is used to determine the output weight matrix W_{out}, represented as

$$W_{out} = YV^T(VV^T + \lambda I)^{-1}, \tag{5}$$

where λ is the regularization factor used to match the forecasted data to the test data. I is the identity matrix of dimension same as VV^T. Using the output weight matrix W_{out}, the output data corresponding to the prediction of cumulative infection per day made by the model is given by the vector $Y(t)$ expressed as [16],

$$Y(t) = \begin{pmatrix} I_{M+1}(t_{r+1}) \\ I_{M+2}(t_{r+2}) \\ \vdots \\ I_N(t_{final}) \end{pmatrix}. \tag{6}$$

Fig. 4 Prediction of monkeypox cumulative infection per day for nine most infected countries listed in the Table 1. Red solid line indicates the realtime cumulative infection data, and open blue circles represent the prediction of the test set by the ESN model and forecasting done by the ESN model is represented by green triangles extended from the predicted set
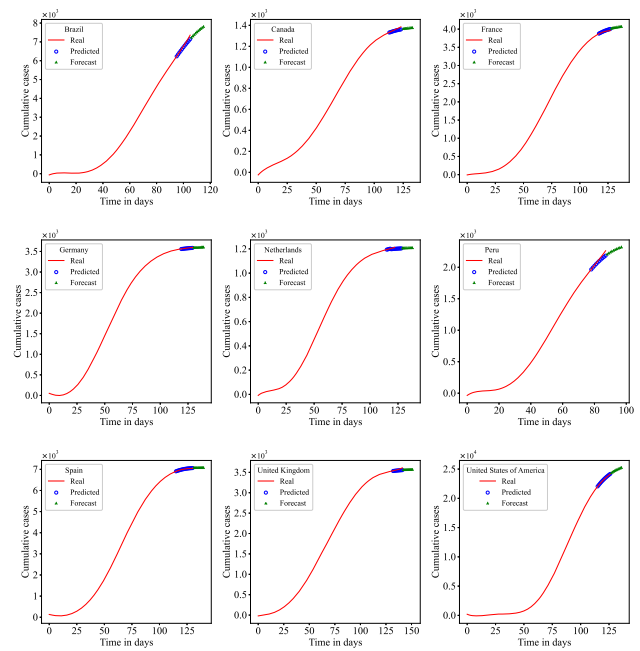
Table 2 Country-wise ReMSE values for different N_{res}

Country	$N_{\text{res}} = 5000$	$N_{\text{res}} = 2000$	$N_{\text{res}} = 1000$
USA	0.0175	0.0197	0.0223
Brazil	0.0835	0.0982	0.0994
Peru	0.0553	0.0687	0.0712
UK	0.0798	0.1024	0.1453
Spain	0.0178	0.0937	0.2474
Germany	0.0505	0.0861	0.0934
France	0.0749	0.0958	0.1427
Netherlands	0.0749	0.0862	0.1284
Canada	0.03281	0.0723	0.1003

3.2 Analysis of realtime cumulative infection

To perform the prediction of realtime cumulative infection per day, we divide the real-world infection data to the proportion of 9:1, in which 9/10th of the data has been given as input to the system and let the machine evolve the input weight and the connection weight to mimic the input data pattern and the rest 1/10th of the data has been used as a metric to validate the system prediction. The data was pre-processed, that is smoothed using Savitzky-Golay filter [21], before being fed into the reservoir. For this prediction, the system has to be fed with proper hyper-parameters based on the nature of the data to be trained. For this optimization process, the hyperopt package [22] in python has been used, which will provide the best parameters that can be used for the prediction. Note that in all our simulations, we use the hyper-parameter ρ to be negative because of non-chaotic nature of the infection data and the magnitude of the spectral radius is scaled to be less than unity. The value of N_{res} is taken to be 5000.

Both the predicted and forecasted data for the nine most infected countries listed in Table 1 are shown in Fig. 4. The predicted data represented by blue open circles is depicted in Fig. 4 and is superimposed along with the realtime cumulative infections (red solid line). It is evident that the ESN model has predicted the realtime cumulative infections fairly well. For this, we took the cumulative infection data from initial infection up to 27th September 2022. The last ten data point from the initial infection data is used as the test set, which is used for the validation of the model. We have forecasted the evolution for the next ten days. From the predicted data, represented by green triangles, it is evident that countries like France, Brazil, Peru and the USA show still a growth pattern despite that the growth pattern is slowed down.

Moreover, one can infer that for countries like Germany, the Netherlands and the UK the infection spread curve looks like a logistic one which shows that in those countries the disease spread is reduced as the value of cumulative cases reaches a constant. It is to be noted that these results were in good agreement with the results obtained from the power law method given in Table 4, where the above-said countries like Germany, Netherlands and the UK have the scaling exponent being dropped below one. This shows that the ESN model is well suited for forecasting the spread for a week or two from the current state. But the drawback of the ESN

Table 3 Country-wise ReMSE values with various training and testing time split ratio

Country	Split ratio 9:1	Split ratio 8:2	Split ratio 7:3	Split ratio 6:4	Split ratio 5:5	Standard Deviation
USA	0.0176	0.0358	0.1006	0.2407	0.3343	0.1226
Brazil	0.0841	0.1172	0.1449	0.1742	0.2821	0.0678
Peru	0.0553	0.1774	0.2313	0.3921	0.4615	0.1331
UK	0.0797	0.1015	0.1793	0.2813	0.4936	0.1508
Spain	0.0178	0.0351	0.0751	0.1298	0.2506	0.0838
Germany	0.0505	0.0913	0.1179	0.1643	0.1987	0.0523
France	0.0749	0.0809	0.1813	0.2273	0.2614	0.0757
Netherlands	0.0738	0.1298	0.2801	0.3391	0.4615	0.1406
Canada	0.0828	0.1183	0.2678	0.3241	0.4226	0.0523

Table 4 Country-wise scaling exponent for forecasted data

Country	γ (for forecatsed data)
USA	0.525
Brazil	0.917
Peru	0.504
UK	0.0362
Spain	0.0175
Germany	0.0418
France	0.181
Netherlands	0.0267
Canada	0.117

model is that it might not be suited for long-term prediction because, during the initial spread, the curve looks like an exponential one, so the model might misinterpret the growth as an exponential one forever, which might not a meaningful result. To overcome this limitation, we need to update the data in realtime and then the prediction needs to be done.

Now, we validate the predictions of our model with the realtime data using the normalized root mean square value error defined by

$$ReMSE = \sqrt{\frac{\frac{1}{T} \int_0^T (I_a(t) - I_m(t))^2 \mathrm{d}t}{\max(I_a)}}, \tag{7}$$

where I_a is the realtime data of the total number of infected individuals, and I_m is the number of infected individuals predicted by the proposed model. The ReMSE values of all the nine most infected countries for the various output weights are calculated and tabulated in Table 2. It is evident from the table that increasing N_{res}, correspondingly increase in the output weights W_{out}, results in better accuracy of the predictions made by the model. Further, the ReMSE values of the nine most infected countries for various data proportions are estimated and tabulated along with their standard deviation in Table 3, which elucidates that the prediction error will be minimized only when a large data set is used for the training. Note that we have used a proportion of 9:1 of the input data for training and testing because of the availability of relatively less realtime data. The model can be provided with a reasonably large data for training than that for the testing process by using this proportion.

3.3 Power law fit for the forecasted data

Using the machine learning technique, we forecasted the disease spread for ten days from the data available (as of 27th September 2022). We have also performed the power law fit for the forecasted cumulative infections, which will provide a quantitative measure for the disease spreading. The calculated scaling exponent corresponding to the forecasted data is tabulated in Table 4. Using the power law fit for the forecasted cumulative infection, indicated by green triangles in Fig. 4, by the ESN, one can infer whether the disease attained the endemic state or not. From that, one can get quantitative information about the infection rate and its trend. These results elucidate that in all nine countries the power law growth has ended and a saturation stage has been attained. For the countries like Netherlands, Spain and UK, the regime of power law has completely ended as the number of new infections is dropped near to zero. Moreover, our results elucidate that even the countries which still show the power law growth in last ten days like the USA, Brazil and Peru have the trend of moving toward saturation stage over time. Using this method has an advantage over usual machine learning method because whenever health care systems make decisions it has been backed up by a quantitative measure of scaling exponent which reveals to us the disease spread rate rather than the results from traditional machine learning methods.

Fig. 5 Total number of monkeypox infections in all the countries that were taken for analysis is plotted as a function of time from initial infection to 4th November 2022
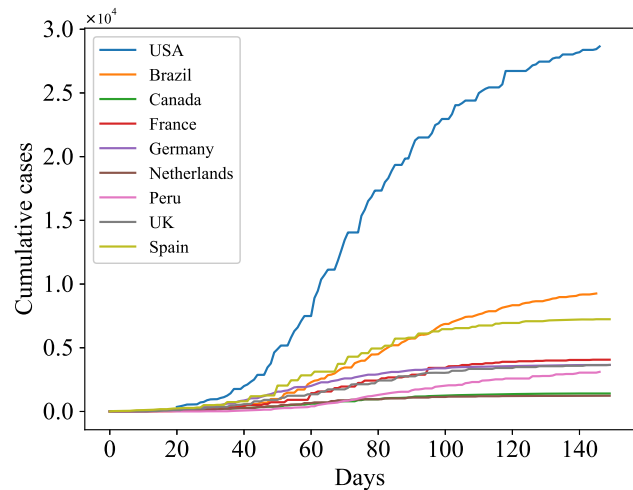
4 Conclusion

In this paper, we have adopted a machine learning-based technique coupled with a power law-based approach for accurate yet quantitatively interpretable prediction of the monkeypox infection. We have accomplished this by choosing the realtime data corresponding to the cumulative infection per day for the nine most infected countries listed in Table 1 and performing a power law fit to the infection data of all these countries. The scaling exponents deduced from the power law fit elucidate the trend of the infection rate. For instance, the initial scaling exponent for the USA is found to be around $\gamma = 4.424$, which reveals the exponential growth of the infection rate.

The scaling exponent for all nine countries is tabulated in the second column of Table 1, which illustrates us the nature of the infection trend in the respective countries. For instance, the infection rate for Brazil and Peru is found to be the least, but above unity, revealing that there lies a possibility of imminent exponential growth of the infection rate in those countries. In order to understand the complete trend of the outbreak, we have performed the power law fit for the last ten days (from 18th September 2022 to 27th September 2022) and tabulated the corresponding scaling exponents in the third column of Table 1. The decimal values of the scaling exponent of the infection rate in several countries for the last ten days elucidate that the infection rate in those countries is in the saturation state, whereas the scaling exponents for Brazil, Peru and USA are still greater than unity revealing the exponential growth pattern of the infection rate is still intact. Further, the scaling exponent for the smallpox outbreak in Japan in 1907–1908 is found to be $\gamma = 4.983$, which is in the order of the initial monkeypox spread in the USA depicting the similarity in the infection pattern of these two viruses signaling that monkeypox could also be potential threat as an epidemic.

We have also used a modified neural network (ESN) to predict and forecast the infection growth pattern. The predicted data is found to match well with the realtime infection data. We have validated the predictions of our model with the realtime data using the normalized root mean square value error and the results are tabulated in Tables 2 and 3. The ReMSE values in these tables elucidates that the accuracy of predictions will be increased for a large output weights and a large input data for the training phase. We have also estimated the scaling exponents for all nine countries from the forecasted data, listed in Table 4, which largely agrees with the scaling exponents for the last ten days tabulated in the third column of Table 1. These observations elucidate that the trend in most countries, other than Brazil, Peru and USA, is almost in saturation phase. It is to be noted that now, at the time of finalizing the manuscript, we have the realtime data until 3rd November 22 as depicted in Fig. 5. It is evident that our results are in excellent agreement with the trend observed in Fig. 5 for the forecasted period.

It is also to be noted that though the power law model is simple to interpret, it cannot take into account the pharmaceutical protocols like quarantine, lockdown, etc. Nevertheless, as of now, there were no such things followed by the infected countries and hence our model provides a valid and interpretable analysis that might help healthcare systems to back their decisions with a quantitative measure. We need to carefully track and control the monkeypox outbreak before it spreads all across the globe because there will be a possibility of this virus getting mutated and intensified. Added to the fact that currently there is no vaccination practice against any of the viruses from the orthopox family, we need more careful analysis of the realtime infection as well as its spread until it becomes completely endemic.

Data Availability Statement This manuscript has associated data in a data repository. [Authors' comment: The data sets on the current study are available from the corresponding author on reasonable request.].

References

1. M. Hraib, S. Jouni, M.M. Albitar, S. Alaidi, Z. Alshehabi, The outbreak of Monkeypox 2022: an overview. Ann. Med. Surg. **79**, 104069 (2022)
2. O.J. Peter, S. Kumar, N. Kumari, F.A. Oguntolu, K. Oshinubi, R. Musa, Transmission dynamics of Monkeypox virus: a mathematical modeling approach. Model. Earth Syst. Environ. **8**, 3423–3434 (2022)
3. R. Grant, L.B.L. Nguyen, R. Breban, Modelling human-to-human transmission of Monkeypox. Bull. World Health Organ. **98**(9), 638 (2020)
4. Y.C. Wu, C.S. Chen, Y.J. Chan, The outbreak of COVID-19: an overview. J. Chin. Med. Assoc. **83**(3), 217 (2020)
5. E.M. Bunge, B. Hoet, L. Chen, F. Lienert, H. Weidenthaler, L.R. Baer, R. Steffen, The changing epidemiology of human Monkeypox-A potential threat? A systematic review. PLoS Negl. Trop. Dis. **16**(2), e0010141 (2022)
6. W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. Ser. A **115**, 700–721 (1927)
7. V.R. Saiprasad, R. Gopal, V.K. Chandrasekar, M. Lakshmanan, Analysis of COVID-19 in India using a vaccine epidemic model incorporating vaccine effectiveness and herd immunity. Eur. Phys. J. Plus. **137**(9), 1–11 (2022)
8. H. Qi, S. Xiao, R. Shi, M.P. Ward, Y. Chen, W. Tu, Z. Zhang, COVID-19 transmission in Mainland China is associated with temperature and humidity: a time-series analysis. Sci. Total Environ. **728**, 138778 (2020)
9. J. Hindes, M. Assaf, I.B. Schwartz, Outbreak size distribution in stochastic epidemic models. Phys. Rev. Lett. **128**(7), 078301 (2022)
10. A. Gowrisankar, T.M.C. Priyanka, S. Banerjee, Omicron: a mysterious variant of concern. Eur. Phys. J. Plus **137**, 100 (2022)
11. D. Easwaramoorthy, A. Gowrisankar, A. Manimaran, S. Nandhini, L. Rondoni, S. Banerjee, An exploration of fractal-based prognostic model and comparative analysis for second wave of COVID-19 diffusion. Nonlinear Dyn. **106**, 1375–1395 (2021)
12. E. Pelinovsky, A. Kurkin, O. Kurkina, M. Kokoulina, A. Epifanova, Logistic equation and COVID-19. Chaos Solit. Fractals **140**, 110241 (2020)
13. A. Gowrisankar, L. Rondoni, S. Banerjee, Can India develop herd immunity against COVID-19? Eur. Phys. J. Plus **135**, 526 (2020)
14. C. Kavitha, A. Gowrisankar, S. Banerjee, The second and third waves in India: when will the pandemic be culminated? EurT. T. Phys. J. Plus **136**, 596 (2021)
15. H.M. Singer, Short-term predictions of country-specific Covid-19 infection rates based on power law scaling exponents. arXiv:2003.11997v1
16. S. Ghosh, A. Senapati, A. Mishra, J. Chattopadhyay, S.K. Dana, C. Hens, D. Ghosh, Reservoir computing on epidemic spreading: a case study on COVID-19 cases. Phys. Rev. E **104**(1), 014308 (2021)
17. M. Lukoševičius, A practical guide to applying echo state networks, in *Neural Networks: Tricks of the Trade* (Springer, Berlin, Heidelberg, 2012), pp. 659–686
18. K. Nakajo, H. Nishiura, Estimation of $R(t)$ based on illness onset data: an analysis of 1907–1908 smallpox epidemic in Tokyo. Epidemics **38**, 100545 (2022)
19. H. Jaeger, H. Haas, Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science **304**(5667), 78–80 (2004)
20. J. Pathak, B. Hunt, M. Girvan, Z. Lu, E. Ott, Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. Phys. Rev. Lett. **120**(2), 024102 (2018)
21. R.W. Schafer, What is a Savitzky-Golay filter? IEEE Signal Process. Mag. **28**(4), 111–117 (2011)
22. N. Trouvain, L. Pedrelli, T.T. Dinh, , X. Hinaut, Reservoirpy: an efficient and user-friendly library to design echo state networks, in *International Conference on Artificial Neural Networks* (Springer, Cham, 2020), pp. 494–505