Original research

DIGITAL HEALTH

DIGITAL HEALTH
Volume 9: 1–17

DOI: 10.1177/20552076231204748
journals.sagepub.com/home/dhj

Sage

A hybrid forecasting technique for infection and death from the mpox virus

Hasnain Iftikhar[1], Muhammad Daniyal[2], Moiz Qureshi[3], Kassim Tawaiah[4,5], Richard Kwame Ansah[4,6] and Jonathan Kwaku Afriyie[5]

Abstract

Objectives: The rising of new cases and death counts from the mpox virus (MPV) is alarming. In order to mitigate the impact of the MPV it is essential to have information of the virus's future position using more precise time series and stochastic models. In this present study, a hybrid forecasting system has been developed for new cases and death counts for MPV infection using the world daily cumulative confirmed and death series.

Methods: The original cumulative series was decomposed into new two subseries, such as a trend component and a stochastic series using the Hodrick–Prescott filter. To assess the efficacy of the proposed models, a comparative analysis with several widely recognized benchmark models, including auto-regressive (AR) model, auto-regressive moving average (ARMA) model, non-parametric auto-regressive (NPAR) model and artificial neural network (ANN), was performed.

Results: The introduction of two novel hybrid models, HPF_1^1 and HPF_3^4, which demonstrated superior performance compared to all other models, as evidenced by their remarkable results in key performance indicators such as root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE), is a significant advancement in disease prediction.

Conclusion: The new models developed can be implemented in forecasting other diseases in the future. To address the current situation effectively, governments and stakeholders must implement significant changes to ensure strict adherence to standard operating procedures (SOPs) by the public. Given the anticipated continuation of increasing trends in the coming days, these measures are essential for mitigating the impact of the outbreak.

Keywords

forecasting, mpox virus, decomposition, Hodrick–Prescott filter, time series, hybrid models

Submission date: 20 May 2023; Acceptance date: 14 September 2023

[1]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
[2]Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur, Pakistan
[3]Department of Statistics, Shaheed Benazir Bhutto University, Shaheed Benazirabad, Pakistan
[4]Department of Mathematics and Statistics, University of Energy and Natural Resources, Sunyani, Ghana
[5]Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[6]Department of Mathematics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Corresponding author:
Kassim Tawaiah, Department of Mathematics and Statistics, University of Energy and Natural Resources, Sunyani, P. O. Box 214, Sunyani, Ghana.
Email: kassim.tawiah@uenr.edu.gh

Introduction

Mathematical modelling has become a useful tool for understanding the spread and transmission of infectious diseases such as mpox.[1] In recent years, there has been a growing body of literature on the mathematical modelling of mpox. The first studies on the mathematical modelling of mpox focused on understanding the basic dynamics of the disease. These models typically used simple compartmental models, such as Susceptible–Exposed–Infectious–Recovered (SEIR) models, to describe the flow of individuals between different states of infection. These models were used to estimate the basic reproductive number of the disease, which represents the average number of secondary cases generated by a single primary case in a fully susceptible population.[1]

More recent studies have expanded on these basic models to incorporate more realistic and complex dynamics. For example, some studies have incorporated demographic factors, such as age structure and population mixing patterns, to better understand the transmission dynamics of mpox. Other studies have considered the impact of interventions, such as vaccination or treatment, on the spread of the disease. One area of particular interest has been the potential for cross-species transmission of mpox from animals to humans.[2]

Mathematical models have been used to estimate the risk of zoonotic transmission and to explore the potential impact of control measures, such as vaccination of wildlife populations, on reducing the risk of transmission. Overall, the literature on the mathematical modelling of mpox has made significant contributions to our understanding of the dynamics of the disease. These models have provided valuable insights into the transmission dynamics and have helped inform the design of control and intervention strategies. However, there is still much to be learnt about the epidemiology of mpox, and on-going research in this field is needed to further refine our understanding of the disease.

Mathematical models are used to simulate the transmission dynamics of the disease and to predict the number of cases in a population. The models take into account various factors such as the incubation period of the virus, the rate of transmission and the susceptibility of individuals in the population. Different types of models can be used to study mpox, including compartmental models, network models and spatial models. Compartmental models divide the population into groups based on their infection status, while network models consider the interactions between individuals in the population. Spatial models take into account the geographical distribution of the disease and the movement of infected individuals. The results of these models can inform public health interventions and strategies to prevent the spread of mpox.[3] For example, they can help identify high-risk populations and areas and guide resource allocation for disease control and prevention efforts.[4]

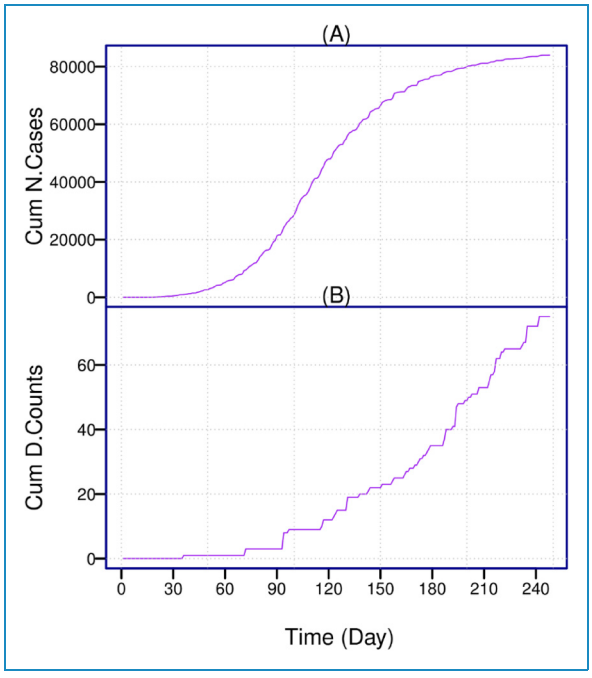

Figure 1. World mpox virus data: cumulative time series new cases (a) and cumulative time series death counts (b) over the period of 1 May 2022 to 3 January 2023.

Table 1. Descriptive statistics of the data.

	Cumulative new cases	Cumulative death cases
Count	18,386	18,386
Mean	587.623	0.592
Median	13.000	0.000
Standard deviation	2522.373	4.035
Kurtosis	83.790	184.306
Skewness	8.481	12.548
Minimum	0.000	0.000
Maximum	29603.00	75.000
First quartile	3.000	0.000
Third quartile	169.000	0.000

It is important to note that while modelling can provide valuable insights into the spread and impact of mpox, it is only one aspect of a comprehensive approach to disease control and prevention.[5] It is essential to complement modelling with other methods, such as surveillance, laboratory testing and vaccine deployment, to effectively control and prevent the spread of the disease.

This cutting-edge study suggests a novel hybrid strategy for improving the daily cumulative confirmed and fatality counts of mpox predicting accuracy. The proposed methodology is based on a filter and different linear and nonlinear time series models. To accomplish this, using the Hodrick–Prescott filter, the initial time series of cumulative confirmed cases and death counts is divided into two new subseries: the nonlinear long-term trend series and a residual series. Four well-known time series models, including the auto-regressive (AR), auto-regressive moving average (ARMA), non-parametric auto-regressive (NPAR) and artificial neural networks (ANN), as well as all of their conceivable model combinations, were taken into consideration to predict the filtered subseries. In order to get the final estimates for mpox that were one day ahead of the total number of confirmed cases and death cases, the different prediction models were immediately combined. The introduction of this new hybrid approach proposed based on filtering and various combinations of linear and nonlinear time series models will improve the accuracy of time series modelling and forecasting in infectious disease modelling and forecasting. This will be an essential contribution to literature. Once more, the suggested approach can be generalized and evaluated for different data sets related to infectious diseases, such as COVID-19 confirmed cases, death counts and recovered cases just as it were implemented in.[6,7] It can also be applied in all areas where time series modelling is implemented.

The rest of the work is organized as follows. Section 2 provides an overview of the literature review and related studies. The general methodology of the suggested hybrid forecasting method is described in Section 3. The suggested modelling methodology is empirically tested in Section 4 utilizing a time series of daily mpox cumulative confirmed cases and deaths. Section 5 provides a discussion of the results and comparison with existing literature. Finally, Section 6 presents conclusions, limitations and future research directions.

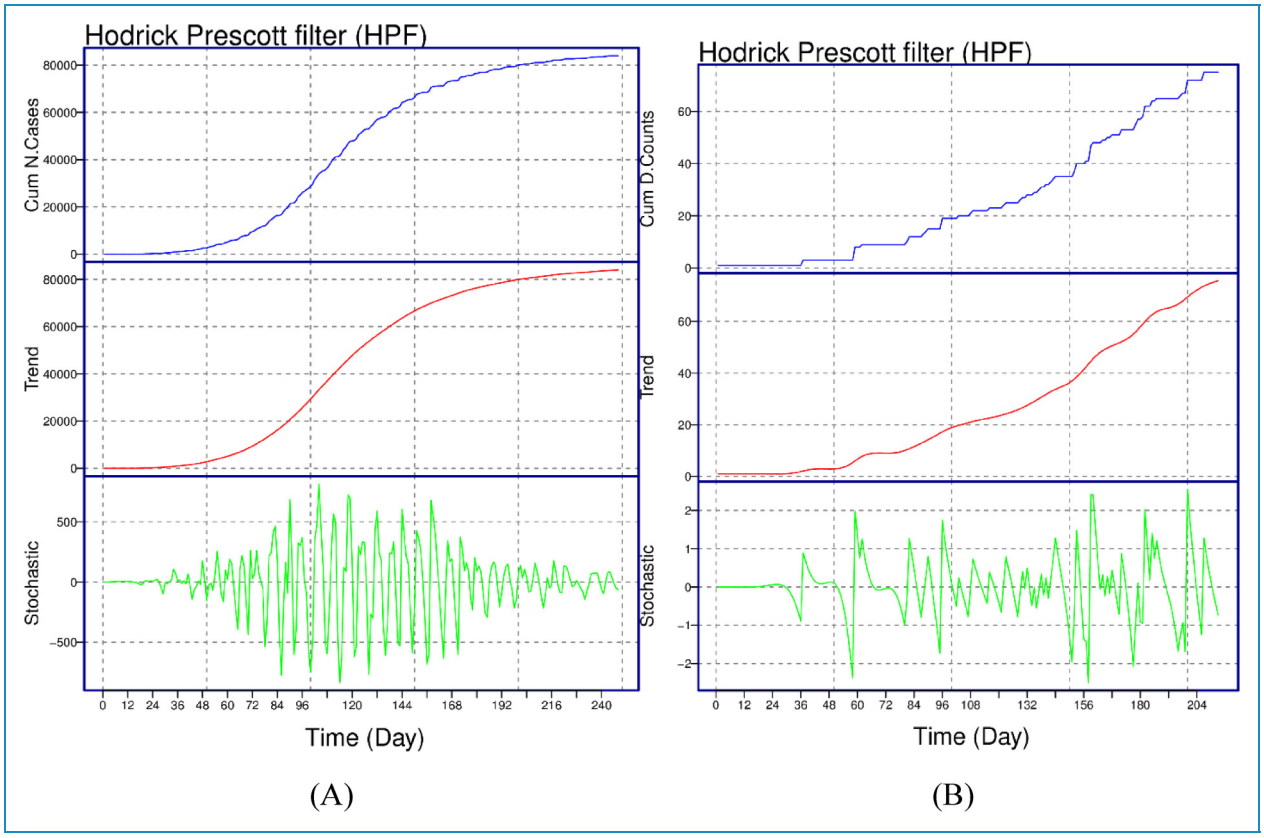

Figure 2. World mpox virus data: The daily cumulative confirmed cases and death counts of the monkeypox virus are filtered by the HPF. Within each subfigure, the top panel shows the daily cumulative confirmed cases and death counts (blue curve), the middle panels show the long-term trend (red curve) and the bottom panels show the stochastic part (green curve).

Literature reviews and related works

According to the World Health Organization (WHO), mpox virus (MPV) presents similar symptoms to smallpox and can spread to all areas.[8,9] In 1970, the first recorded case of human MPV infection was in a 9-month-old boy in the Democratic Republic of the Congo, as reported by the WHO[8] and the Centers for Disease Control and Prevention (CDC).[10] MPV has become one of the leading orthodox viruses that infect humans. Recently, the number of reported MPV cases increased significantly compared to previous decades.[8,10] The CDC reports that as of 2 August 2022, there have been 25,391 confirmed cases of MPV globally, with 25,047 instances in nonendemic regions (countries without a history of recorded cases) and 344 cases in endemic regions (countries with a history of documented cases).[10] Although the main source of the presence of MPV in humans is currently unknown,

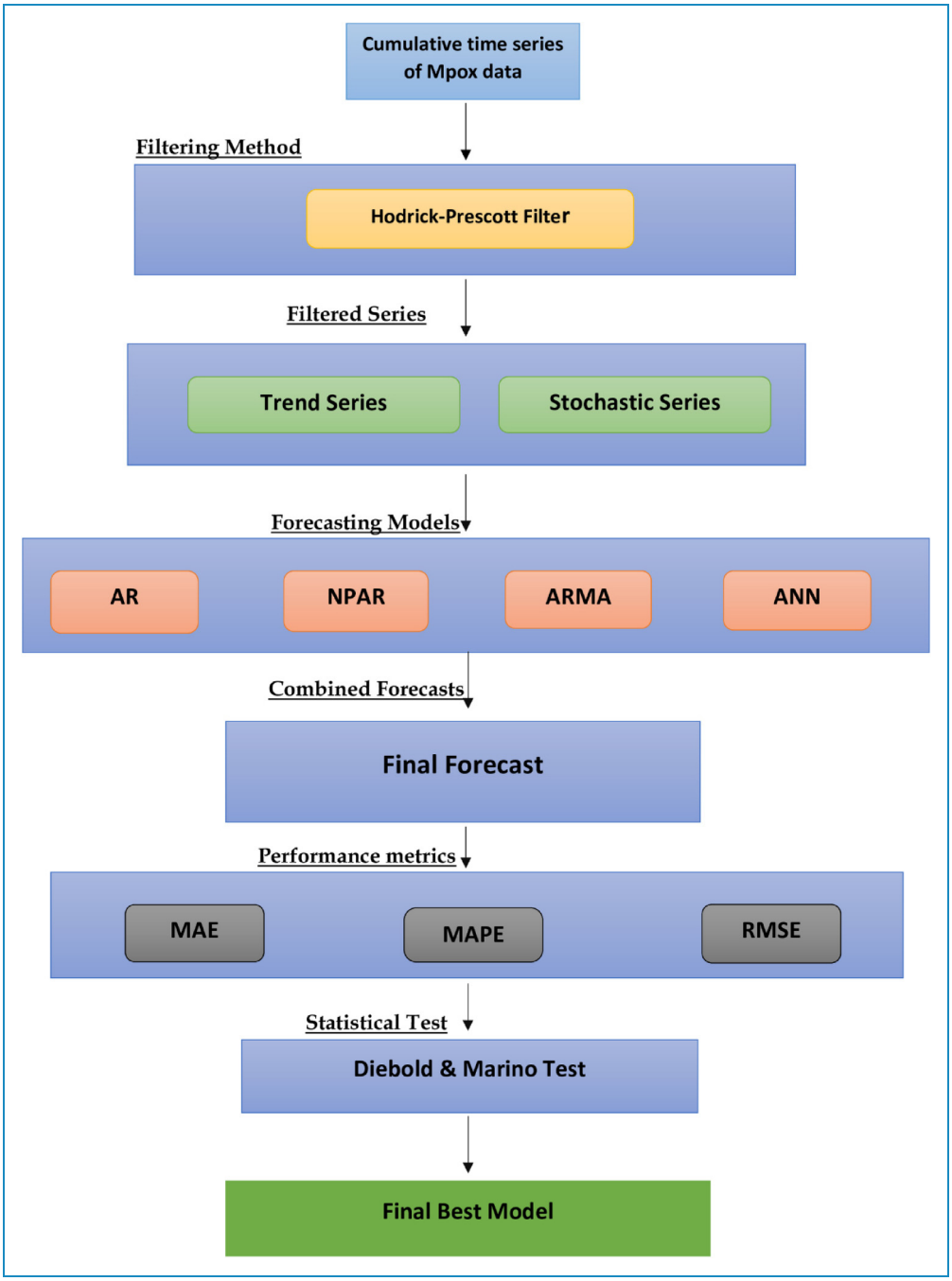

Figure 3. Flow chart of the proposed modelling procedure for global confirmed cumulative new and death cases mpox infection.

studies have proposed that animals such as rodents and non-human primates may be potential reservoirs.[8,10] Globally, there have been a total of 21,099 confirmed cases as of 28 July 2022. Similar to this, 42 nations in the five WHO regions have also recorded one fatality from MPV infection.[8]

Ninety-eight percent of the confirmed cases, or the majority of them, have been recorded after May 2022. The MPV has had a significant impact on people's lives and the global economy, particularly in respect of the management of the infection and the lookout for efficient local or national initiatives. There are several mathematical and statistical models and techniques that have been successfully utilized to study how epidemic illnesses and pandemics behave. In [11] researchers conducted a study to forecast human mpox based on human judgement utilizing probabilistic models. They suggested that forecasting based on human judgement can generate probabilistic predictions for public health importance.

Table 2. World mpox data: out-of-sample cumulative new cases for day-ahead mean forecast error for all combination models.

Models	MAPE	MAE	RMSE
HPF_1^1	0.0547	44.8904	63.3184
HPF_2^1	0.0638	52.2712	70.6252
HPF_3^1	0.0665	54.5975	65.4382
HPF_4^1	0.0597	49.0543	65.3335
HPF_1^2	0.0563	46.1637	64.6279
HPF_2^2	0.0652	53.4423	72.0460
HPF_3^2	0.0674	55.3443	66.3498
HPF_4^2	0.0602	49.4720	65.7891
HPF_1^3	0.0556	45.6072	63.9815
HPF_2^3	0.0646	52.9030	71.3286
HPF_3^3	0.0671	55.1286	66.0678
HPF_4^3	0.0601	49.3553	65.6847
HPF_1^4	0.1201	98.7468	116.1344
HPF_2^4	0.1214	99.8275	120.2476
HPF_3^4	0.1243	102.1602	116.4604
HPF_4^4	0.1043	85.7481	102.0147

In [12] a comparative analysis between machine learning and classical time series models was used to forecast the mpox cumulative cases. This result showed that the machine learning model outperformed the classical time series model in forecasting the cumulative MPV cases. [13] used five different time series and machine learning techniques to forecast the cumulative human mpox cases. The authors found that the neural prophet model achieved a high level of accuracy compared to the other models. Multiple candidate ARIMA models have been used to predict human mpox phenomena utilizing the Akaike information criteria (AIC), root mean square error (RMSE) and mean absolute error (MAE) as goodness of fit.[14] In comparison, ARIMA (5, 2, 3) emerged as the best fit for the daily confirmed cases.[14]

A short forecast of the mpox cases presented in [15] utilized the ensemble n-sub-epidemic modelling based on a 10-week calibration. The results showed a global downward trend in mpox. The work of [16] found an association between mpox and metrological factors. The authors assumed that temperature, relative humidity, precipitation, wind speed, surface pressure and dew point are significant factors in the increasing phenomena of mpox cases. They utilized the auto-regressive integrated moving average with exogenous variable (ARIMAX) model to find the association between the factors and seasonal exponential smoothing (SES), ARIMA and prophet models to forecast the cumulative cases. Based on the revealed results, the authors found that temperature, surface pressure and relative humidity have a significant relation with the increasing phenomena of mpox cases.

The prediction of many chronic diseases has regularly employed both machine learning and traditional statistical methods.[17,18] For example, in [19] the authors analysed and forecasted the mpox outbreak based on some statistical models by shedding light on hair forecasting phenomena through which the prophet model was found effective. [20] used machine learning and power law distribution algorithms to achieve a similar result. The results revealed that the methods based on machine learning may perform well in predicting and forecasting mpox cases. Their work shed light on modelling and predicting the spread of mpox in the United States of America (USA).

For the authors in [21], they used neural network time series models and compared them with long-short-term memory and GRU models by utilizing the augmented dynamic adaptive model (ADAM) optimizer. Their results revealed that the ANN model performed well than the other models in prediction of the spread of mpox in the USA, thereby proposing a machine and stacking learning techniques to forecast the transmission of mpox outbreaks. [22] compared the proposed stacking ensemble learning (SEL) technique with AdaBoost, random forest (RF), GBOOST, least absolute shrinkage and selection operator (LASSO), ridge, and ordinary least squares (OLS) to forecast the transmission of mpox

using key performance indicators (KPI) such as RMSE, MAE and mean square error (MSE). The study found that the proposed SEL technique outperformed the others.[22]

In [23] the ARIMA model was utilized to model and forecast the mpox cases for those countries that have the highest infection rate. The results showed that the ARIMA model pointed to an increasing trend in USA, Spain, Germany, England, and France for the upcoming days. However, [24] used seasonal ARIMA (SARIMA) and ARIMA to model and forecast mpox cumulative cases for a short-term period, with the SARIMA model achieving the highest accuracy than the ARIMA model in modelling and forecasting confirmed and cumulative cases.

A comparative analysis between machine learning and classical time series methods using ARIMA, RF, logistic regression (LR), decision tree (DT), ensemble network (EN), ANN, and convolutional neural network (CNN) have been performed with the CNN model outperforming the other existing models.[25] Countless deep convolutional neural network (DCNN) models have also been used in the detection as well as prediction of mpox with significant accuracy.[26–28] Other models, like transfer learning models, have also been implemented in mpox detection and classification.[29]

Hybrid models have been widely used in disease forecasting.[30,31] It has been implemented in modelling and forecasting dengue fever infection. [32,33] conducted a review study on hybrid methodologies applied to epidemics. They concluded that by using the appropriate methods that are suitable to the data, one can achieve consistent results in modelling and predicting any series of an epidemic. For more epidemic hybrid time series forecasting, see, for example, [34–42].

Materials and methods

Data

This study utilized world daily cumulative new cases and death counts of MPV for all analysis. The data set was extracted from Our World in Data website.[43] The data set

Table 3. World mpox infection data: *p*-value of the DM test using the loss-squared function.

Models	HPF_1^1	HPF_2^1	HPF_3^1	HPF_4^1	HPF_1^2	HPF_2^2	HPF_3^2	HPF_4^2	HPF_1^3	HPF_2^3	HPF_3^3	HPF_4^3	HPF_1^4	HPF_2^4	HPF_3^4	HPF_4^4
HPF_1^1	0	0.81	0.12	0.27	0.98	0.83	0.11	0.28	0.96	0.82	0.11	0.28	0.57	0.78	0.01	0.29
HPF_2^1	0.19	0	0.11	0.12	0.21	0.89	0.1	0.13	0.2	0.82	0.11	0.13	0.19	0.67	0.02	0.07
HPF_3^1	0.88	0.89	0	0.63	0.89	0.89	0.36	0.65	0.89	0.89	0.32	0.64	0.78	0.84	0.26	0.6
HPF_4^1	0.73	0.88	0.37	0	0.75	0.88	0.36	0.95	0.75	0.88	0.36	0.89	0.71	0.84	0.22	0.51
HPF_1^2	0.02	0.79	0.11	0.25	0	0.81	0.1	0.26	0.1	0.8	0.1	0.25	0.53	0.77	0.01	0.26
HPF_2^2	0.17	0.11	0.11	0.12	0.19	0	0.1	0.12	0.19	0.14	0.1	0.12	0.16	0.65	0.02	0.07
HPF_3^2	0.89	0.9	0.64	0.64	0.9	0.9	0	0.66	0.9	0.9	0.44	0.65	0.78	0.85	0.26	0.6
HPF_4^2	0.72	0.87	0.35	0.05	0.74	0.88	0.34	0	0.73	0.87	0.34	0.1	0.7	0.83	0.2	0.48
HPF_1^3	0.04	0.8	0.11	0.25	0.9	0.81	0.1	0.27	0	0.81	0.1	0.26	0.55	0.77	0.01	0.27
HPF_2^3	0.18	0.18	0.11	0.12	0.2	0.86	0.1	0.13	0.19	0	0.1	0.12	0.17	0.66	0.01	0.07
HPF_3^3	0.89	0.89	0.68	0.64	0.9	0.9	0.56	0.66	0.9	0.9	0	0.65	0.78	0.85	0.26	0.6
HPF_4^3	0.72	0.87	0.36	0.11	0.75	0.88	0.35	0.9	0.74	0.88	0.35	0	0.7	0.83	0.21	0.49
HPF_1^4	0.43	0.81	0.22	0.29	0.47	0.84	0.22	0.3	0.45	0.83	0.22	0.3	0	0.86	0.04	0.24
HPF_2^4	0.22	0.33	0.16	0.16	0.23	0.35	0.15	0.17	0.23	0.34	0.15	0.17	0.14	0	0.05	0.09
HPF_3^4	0.99	0.98	0.74	0.78	0.99	0.98	0.74	0.8	0.99	0.99	0.74	0.79	0.96	0.95	0	0.77
HPF_4^4	0.71	0.93	0.4	0.49	0.74	0.93	0.4	0.52	0.73	0.93	0.4	0.51	0.76	0.91	0.23	0

of total confirmed cases spans from 6 May 2022 to 3 January 2023, and the total death counts are over the range from 15 July 2022 to 3 January 2023. The graphical presentation of both series, cumulative new cases and death counts, can be seen in Figure 1. The complete data set of cumulative new cases covers 248 days, which, from 6 May 2022 to 15 November 2022 (199 days), was used to train the model, while from 16 November to 3 January 2023 (49 days) was used for cumulative new case day-ahead post-sample (testing) prediction. On the other hand, the death counts covered 213 days, of which, from 5 July 2022 to 21 November 2022, a total of 170 days were used for model estimation and from 22 November to 3 January 2023 (43 days) for cumulative death counts day-ahead post-sample (testing) prediction. The summary of the data is provided in Table 1.

The following subsections detail our inferences about the outcomes from different perceptions.

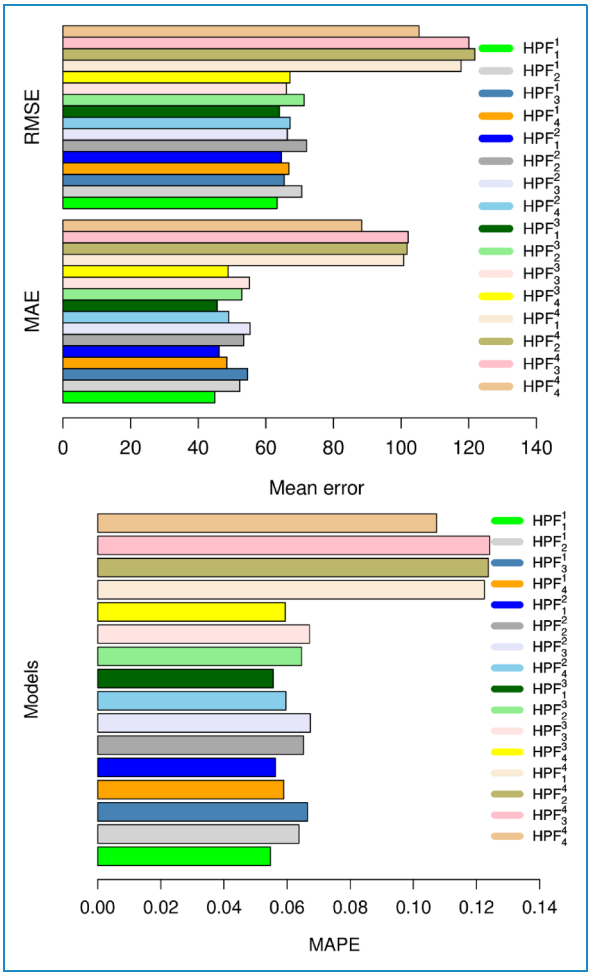

Figure 4. Graphical representation of accuracy measures (MAE, MAPE and RMSE) hybrid models for new cases of global mpox cases. The RMSE and MAE are presented in top box plot, while the MAPE is presented in the bottom box plot.

Methods

Proposed hybrid forecasting system

This section explains the proposed hybrid forecasting methods for cumulative new cases and death counts globally. Generally, the cumulative time series includes an increasing linear or nonlinear long-run trend component. To do this, the cumulative new cases series (C_t) and death count cumulative series (D_t) are decomposed into two components: the long-run trend component (l_t) and a stochastic component (s_t) by utilizing the Hodrick–Prescott filter (HPF).[44] Mathematically, the decomposed components[44] can be written as

$$C_t = l_t + s_t$$
$$D_t = l_t + s_t \quad (1)$$

A detailed description of the HPF is illustrated in the following section.

Hodrick–Prescott filter

A smoothed curve representation of a time series that is more complex to long-term variations than to short-term fluctuations is obtained by using the HPF.[44] The adjustment of the sensitivity of the trend to short-term fluctuations is achieved by modifying a multiplier λ. Let C_t $(t = 1, 2, \ldots, N)$ denote the time series data. The series C_t is made up of a trend component denoted by τ_t and error component denoted by ε_t, thereby yielding an equation defined[44] as

$$C_t = \tau_t + \varepsilon_t \quad (2)$$

Here, a component for the long-term trend, which may be calculated by minimizing the expression, is included as follows:

$$\min_t \left(\sum_{t=1}^{N} (C_t - \tau_t)^2 + \lambda \sum_{t=2}^{N-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right) \quad (3)$$

The loss function is the first term in the equation above, and the second term is a penalty term multiplying the second differencing sum of the squares of the trend components, penalizing variations in the trend component growth rate.

For a visual representation of the HPF's performance, the new subseries decomposed are shown in Figure 2: first, the cumulative series of new cases and second the trend (l_t) as well as the third the stochastic component in both panels. From the figure, it is observed that the HPF decompose (C_t or D_t) captures the dynamics well. The nonlinear long-run trend is adequately captured by both the cumulative new cases and death counts series.

Modelling and comparison of the decomposed components

Once the components are extracted from the cumulative new cases series (C_t) and death counts cumulative series (D_t) using the HPF, the estimate of the extracted components is acquired by four well-known time series models, including two linear and two nonlinear models, such as AR, ARMA, NPAR and ANN.[6,7,44] The models are briefly explained as in the following subsections. The proposed hybrid models are then compared with each other, and the most appropriate among them is compared with the benchmark models for the accuracy of prediction.

Auto-regressive model. A linear and parametric auto-regressive (AR) process describes the short-term dynamics of Q_t and considers a linear combination of the previous time t observations of Q_t[6] is denoted as

$$Q_t = \mu + \gamma_1 Q_{t-1} + \gamma_2 Q_{t-2} + \ldots + \gamma_t Q_{t-n} + \epsilon \quad (4)$$

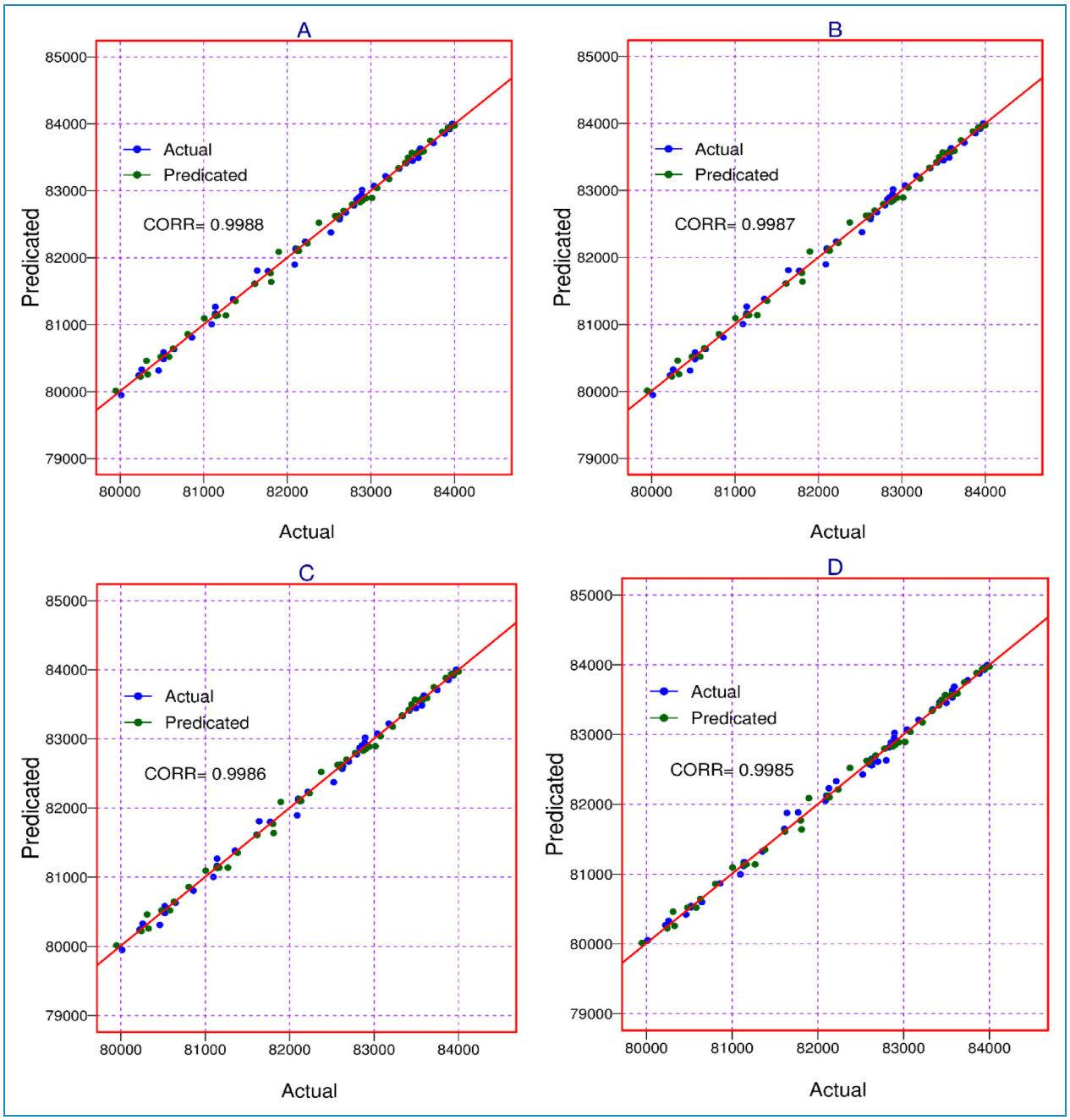

Figure 5. Correlation plots for four other models for cumulative new global mpox cases.

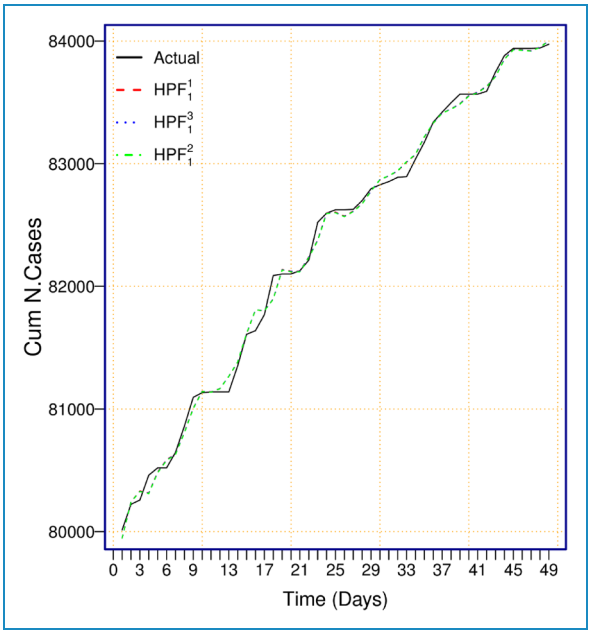

Figure 6. The observed and predicted values for the three best models for cumulative new global mpox cases.

where $Q_t = C_t$ *or* $Q_t = D_t$ (as described in the previous section), μ and $\gamma_i (i = 1, 2, \ldots, t)$ are the average value and slope parameters of the underlying auto-regressive process and ϵ is the error of the model.

Auto-regressive moving average model. The ARMA model not solely incorporates the past values of the target variable but also utilizes important information in the form of moving average(s). In the recent case, the study variable C_t is explained by the previous n terms, the lagged values of residuals as well.[42] Mathematically,

$$Q_t = \mu + \gamma_1 Q_{t-1} + \gamma_2 Q_{t-2} + \ldots + \gamma_r Q_{r-N} + \epsilon_t + \varnothing_1 \epsilon_{t-1} + \varnothing_2 \epsilon_{t-2} + \ldots + \varnothing_m \epsilon_{t-m} \quad (5)$$

μ represents the average value of the model; $\gamma_i (i = 1, 2, \ldots, t)$ and $\varnothing_k (k = 1, 2, \ldots, m)$ denote the parameters of the ARMA processes, respectively; and ϵ_t is normally distributed with white noise with constant mean and variance. Model selection of ARMA is guessed by inspecting the correlograms.

Non-parametric auto-regressive model. The additive non-parametric counterpart of the AR process results in the additive model (NPAR), where the link between Q_t and its preceding terms lacks any particular parametric form,[45] likely allowing for any type of nonlinearity that is expressed as

$$Q_t = g_1 Q_{t-1} + g_2 Q_{t-2} + \ldots + g_t Q_{t-N} + \epsilon_t \quad (6)$$

$g_i (i = 1, 2, \ldots, k)$ denotes the smoothing constant and shows the link between Q_t and its recent past lag value.

Table 4. World mpox death data: out-of-sample cumulative death counts day-ahead mean forecast error for all combination models.

Models	MAPE	MAE	RMSE
HPF$_1^1$	1.0265	0.6525	0.9557
HPF$_2^1$	1.0345	0.6533	1.0218
HPF$_3^1$	0.9991	0.6336	0.8673
HPF$_4^1$	1.0263	0.6507	0.9107
HPF$_1^2$	1.0351	0.6581	0.9623
HPF$_2^2$	1.0429	0.6585	1.0261
HPF$_3^2$	0.9956	0.6312	0.8662
HPF$_4^2$	1.0288	0.6522	0.9142
HPF$_1^3$	1.0349	0.6580	0.9602
HPF$_2^3$	1.0417	0.6575	1.0242
HPF$_3^3$	0.9991	0.6337	0.8660
HPF$_4^3$	1.0292	0.6525	0.9124
HPF$_1^4$	0.9300	0.6019	0.9669
HPF$_2^4$	0.9476	0.6075	1.0507
HPF$_3^4$	0.8538	0.5427	0.8244
HPF$_4^4$	0.8958	0.5803	0.8934

Table 5. Comparison of the proposed best model and the benchmark models: out-of-sample cumulative new counts day-ahead mean forecast error.

Models	MAPE	MAE	RMSE
HPF$_1^1$	0.0547	44.8904	63.3184
AR	0.0929	76.0600	98.4571
NPAR	0.1040	85.5860	101.4200
ARMA	0.0804	65.8699	81.5731
NNR	0.0887	72.6908	101.6149

In our case, the function g_i represents cubic regression splines. Similar to the parametric form, the lag 2 is utilized while estimating NPAR.

Table 6. World mpox virus data: *p*-values of the DM test using the loss-squared function for cumulative new cases.

Models	HPF_1^1	AR	NPAR	ARMA	NNR
HPF_1^1	0	0.08	0.78	0.89	0.02
AR	0.92	0	0.95	0.96	0.01
NPAR	0.22	0.05	0	0.78	0.01
ARMA	0.11	0.04	0.22	0	0.01
NNR	0.98	0.99	0.99	0.99	0

Artificial neural network. To model a wide range of nonlinear issues, ANNs are adaptable computing frameworks. In comparison to other nonlinear models, the main benefit of ANN is its ability to approximate a variety of functions with greater accuracy. Due to the concurrent processing of the data, they are efficient. The process of model building does not relate to acquire any information of the models' form. However, the data characteristic plays a key role in establishing the network model. The multilayer perceptron (MLP), especially with one hidden layer, is one of the most often used ANN types for time series modelling and forecasting.[46] The network that establishes the model is composed of three layers of basic processing units connected by cyclic linkages. The following equation is used to denote the link between output (Q_t) and inputs ($Q_{t-1}, Q_{t-2}, \ldots, Q_{t-n}$),[47] where

$$Q_t = \alpha + \sum_{k=1}^{z} g_k \varphi \left(g_{ok} + \sum_{j=1}^{n} Q_{t-j} \right) \tag{7}$$

The model parameters are denoted as $g_{(j,k)}(j = 1, 2, \ldots, n; k = 1, 2, \ldots z)$ and $g_j(j = 1, 2, \ldots, z)$, where n denotes the length of input nodes, while z denotes the length of hidden nodes in the equation above.

For evaluating the prediction accuracy, three accuracy measure criteria, namely, MAPE, MAE and RMSE, for each combination model are computed[6,7,44] as follows:

$$MAPE = Mean \left(\frac{|Q_t - \hat{Q}_t|}{|Q|} \right) \times 100$$

$$MAE = Mean \left(|Q_t - \hat{Q}_t| \right)$$

$$RMSE = \sqrt{Mean(Q_t - \hat{Q}_t)^2} \tag{8}$$

where Q_t is the original and $\hat{Q}_t$ is the forecasted cumulative new case or death counts values for the t^{th} observation. Using the above accuracy measures, the best model within the proposed hybrid prediction system is obtained. A flowchart of our proposed hybrid forecasting procedure is shown in Figure 3.

Results

To determine the efficient model among the proposed combined models for cumulative new cases, three standard key performance indicators were computed, and the results were presented in Table 2. Table 2 indicates that the HPF_1^1 model outperforms the other models because it has the lowest errors than the others for the cumulative new case prediction. The MAPE, MAE and RMSE values for HPF_1^1 model are 0.0547, 44.8904 and 63.3184, respectively. The HPF_1^3 and HPF_1^2 models follow in the second and third best models. The second and third best models' mean errors are MAPE = 0.0556, MAE = 45.6072, RMSE = 63.9815 and MAPE = 0.0563, MAE = 46.1637 and RSME = 64.6279, respectively. After the accuracy measurements have been computed, the evaluation of these outcomes followed suit. To achieve this, the Diebold and Mariano test (DM) was implemented.[48–51] The hypothesis for the DM is

H_o: *the models on the rows and colums are equally accurate*

H_1: *the models on the colums are more accurate than the models on the rows*

The DM is applied to each combination of the models.[52] The DM test outputs (*p*-value) are demonstrated in Table 3. We rejected the null hypothesis at a 5% level of significance. The output shows that among all possible combinations of the models in Table 2, the HPF_1^1, HPF_1^3 and HPF_1^2 models statistically outperform the others at the 95% level of confidence.

Graphical representation of the accuracy measures (MAE, MAPE and RMSE) for all possible combinations of the models is shown in Figure 4, wherein RMSE and MAE (top box plot), and the MAPE is (bottom box plot). From these plots, we can see that the HPF_1^1 model produces the lowest mean error values (MAPE, MAE and RMSE) in comparison to the other existing models, while the HPF_1^3 and HPF_1^2 are the best, comparatively.

On the other hand, Figure 5 shows the correlation plots for the other four best models. From Figure 5, the model with the highest correlation value is acknowledged as the best model since it has the highest correlation value between the observed and predicted values.[53,54] Additionally, Figure 6 shows the observed and predicted values for the three best models. In light of the descriptive statistics, statistically significant tests and exploratory findings, the suggested hybrid system performs better than all previous models for predicting the novel MPV case from all angles.

Table 7. World mpox death data: output of sample DM test.

Models	HPF_1^1	HPF_2^1	HPF_3^1	HPF_4^1	HPF_1^2	HPF_2^2	HPF_3^2	HPF_4^2	HPF_1^3	HPF_2^3	HPF_3^3	HPF_4^3	HPF_1^4	HPF_2^4	HPF_3^4	HPF_4^4
HPF_1^1	0	0.99	0.6	0.63	1	1	0.65	0.65	1	0.99	0.63	0.65	1	1	1	1
HPF_2^1	0.01	0	0.27	0.35	0.02	1	0.31	0.36	0.01	1	0.3	0.36	1	1	1	1
HPF_3^1	0.4	0.73	0	0.58	0.46	0.77	0.98	0.61	0.43	0.75	0.99	0.6	1	1	1	1
HPF_4^1	0.37	0.65	0.42	0	0.42	0.7	0.47	0.81	0.39	0.68	0.45	0.93	1	1	1	1
HPF_1^2	0	0.98	0.54	0.58	0	0.99	0.58	0.6	0.01	0.99	0.57	0.6	1	1	1	1
HPF_2^2	0	0	0.23	0.3	0.01	0	0.25	0.31	0.01	0	0.25	0.31	1	1	1	1
HPF_3^2	0.35	0.69	0.02	0.53	0.42	0.75	0	0.55	0.38	0.72	0.1	0.55	1	1	1	1
HPF_4^2	0.35	0.64	0.39	0.19	0.4	0.69	0.45	0	0.38	0.67	0.43	0.44	1	1	1	1
HPF_1^3	0	0.99	0.57	0.61	0.99	0.99	0.62	0.62	0	0.99	0.6	0.62	1	1	1	1
HPF_2^3	0.01	0	0.25	0.32	0.01	1	0.28	0.33	0.01	0	0.27	0.33	1	1	1	1
HPF_3^3	0.37	0.7	0.01	0.55	0.43	0.75	0.9	0.57	0.4	0.73	0	0.57	1	1	1	1
HPF_4^3	0.35	0.64	0.4	0.07	0.4	0.69	0.45	0.56	0.38	0.67	0.43	0	1	1	1	1
HPF_1^4	0	0	0	0	0	0	0	0	0	0	0	0	0	0.95	0.6	0.13
HPF_2^4	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0.43	0.08
HPF_3^4	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.57	0	0.01
HPF_4^4	0	0	0	0	0	0	0	0	0	0	0	0	0.87	0.92	0.99	0

To choose which model is efficient among the proposed combined models for cumulative death counts, accuracy measures are calculated and presented in Table 4. The results in Table 4 indicate that the HPF_3^4 model yields the least model error value compared to all other existing models for the cumulative death counts prediction. The MAPE, MAE and RMSE values for the HPF_3^4 model are 0.8538, 0.5427 and 0.8244, respectively. However, results from the models HPF_4^4 and HPF_1^4 were the second- and third-best, respectively. The models HPF_4^4 and HPF_1^4 produced the second- and third-best results, respectively. The second- and third-best models' mean errors are as follows: MAPE = 0.8958, MAE = 0.5803, RMSE = 0.8934 and MAPE = 0.9300, MAE = 0.601, and RMSE = 0.9669. Thus, based on the results obtained, the proposed hybrid system, the HPF_3^4 model is more accurate for prediction than its competitors. Our best hybrid model is compared with our benchmark models in terms of their mean errors in Table 5 with the results of the corresponding DM test presented in Table 6. These results all point to the fact that our hybrid model is more efficient.

To check the efficiency of the best model listed in Table 4, the DM test is applied to each combination of the model.[52] Table 7 displays the results of the DM test. The hypothesis here is the same as that in Table 3. Table 7 shows that among all combination models in Table 4, the HPF_3^4, HPF_4^4 and HPF_1^4 models are statistically efficient at the 95% confidence level. A graphical representation of the accuracy measures is presented in Figure 7. From the figure, the HPF_3^4 model produced the lowest mean error values (MAPE, MAE and RMSE) in comparison to the other models, with the HPF_4^4, HPF_1^4 and HPF_2^4 models being the best competitors. Figure 8 displays the graphical descriptions of the correlation graphs for these three top models. From Figure 8, the optimum model yields the highest correlation value and results in a significant correlation between the observed and predicted values. Figure 9 also shows the observed and predicted values for the three top models. The best model's predictions match the observed cumulative death numbers, as seen in the figure. Our best hybrid model is compared with our benchmark models in terms of their mean errors in Table 8 with

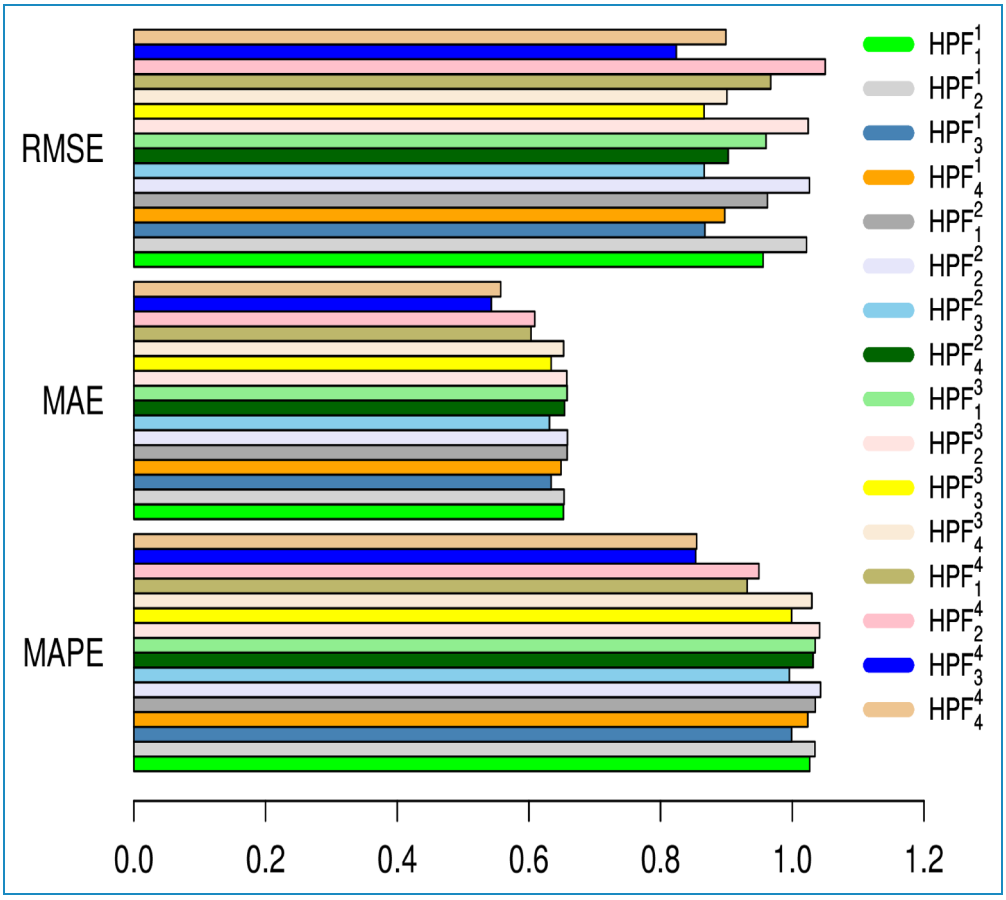

Figure 7. Graphical representation of accuracy measures (RMSE, MAE and MAPE) hybrid models for global death mpox cases.

the results of the corresponding DM test presented in Table 9. These results all point to the fact that our hybrid model is more efficient.

Table 10 presents a short forecast of the confirmed global mpox new cases and death counts. The results point to marginal new cases and significantly low deaths.

Discussion

This study introduced a valuable contribution to the field of disease prediction by presenting two innovative hybrid models, namely, HPF^1_1 and HPF^4_3. These models are designed to forecast the cumulative new cases and cumulative deaths associated with MPV infection, which is a critical global health concern. This hybrid model's way outperformed the benchmark models employed in the study.

Our results proved to be superior to the results in[25] where the same data sets were used. Again, they were of the view that ARIMA is satisfactory for forecasting cases of MPV, which vehemently contract our results. Again, our mean error values are far better than those obtained in[25]. Again, our results contradict the results in[24] where SARIMA and ARIMA were used for forecasting. Just like in[25], our mean error values were far better than those obtained in[24].

More so, our hybrid models again performed way better than the raw machine learning time series model, ANN model, model-free and all other models implemented in the analysis and forecasting of MPV in [19–22] and [12–15], respectively. Again, our forecasted values are more precise and accurate than those presented in these works.

This study plays a pivotal role in advancing the understanding of hybrid forecasting methodologies and, specifically, their application in predicting mpox outbreak trends. By shedding light on their accuracy and efficiency, the study contributes substantially to the field of infectious disease modelling and serves as a stepping stone toward more precise and effective forecasting tools for public health emergencies.

Conclusions

In the disease prediction analysis, the decomposition of the original cumulative series has been done for the very first time by filtering method; then to enhance the forecasting performance, different linear and nonlinear time series models were compared for the best results, and the

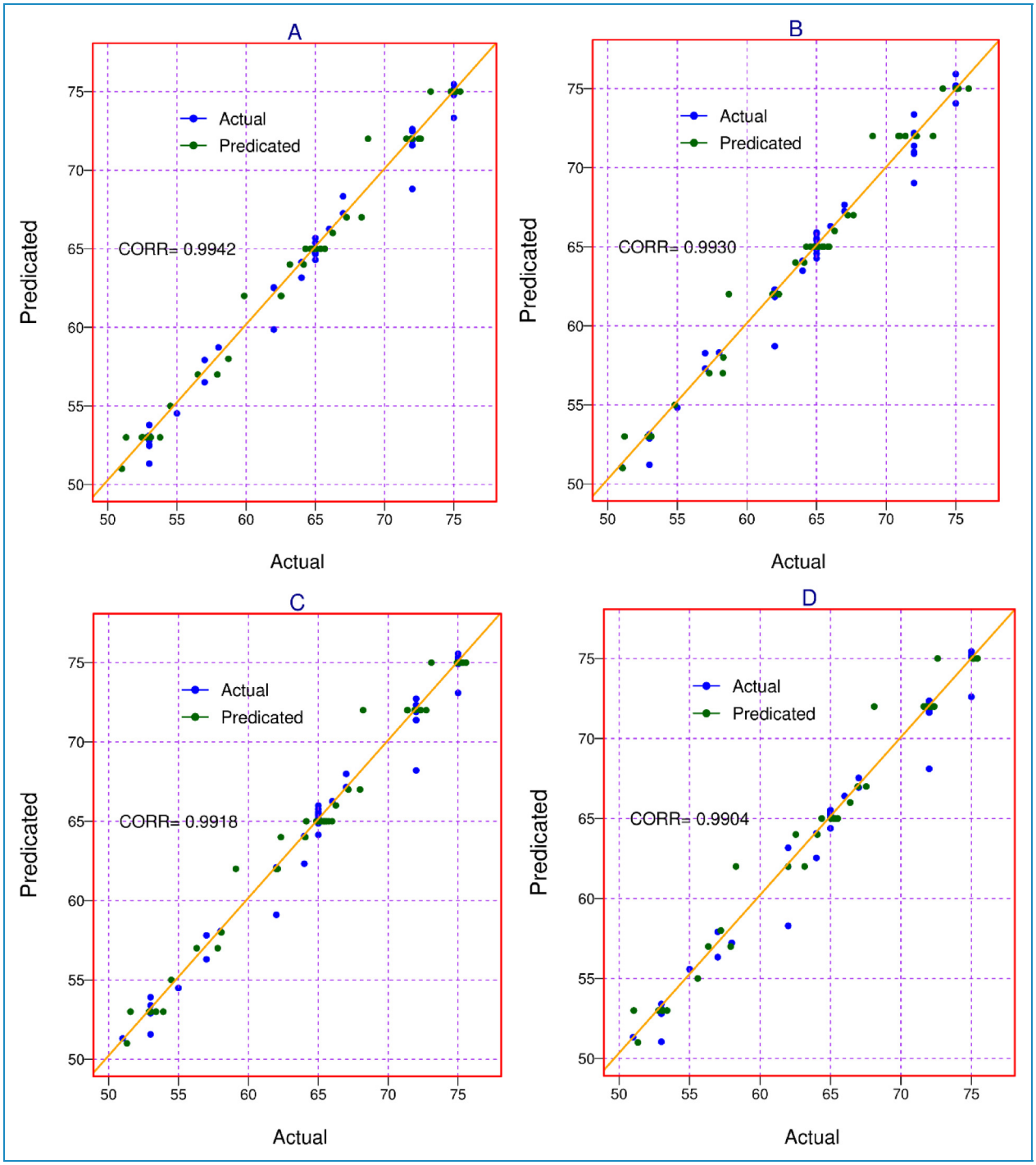

Figure 8. Correlation plots for four other models for cumulative death global mpox cases.

suggested model can be applied for disease deaths and cases in the future. The study proposed a new hybrid forecasting system for predicting MPV new cases and deaths worldwide. The original cumulative series was decomposed into trend and stochastic subseries using the Hodrick–Prescott filter. Two linear and two nonlinear time series models were considered to forecast the decomposed subseries, resulting in a total of 16 models in the hybrid system. Performance evaluation involved three standard mean errors, graphical analysis and the Diebold and Mariano test. The HPF_1^1 and HPF_3^4 hybrid models showed superior accuracy and efficiency among the proposed hybrid forecasting models for the new and death cases, respectively. These hybrid models also outperformed a

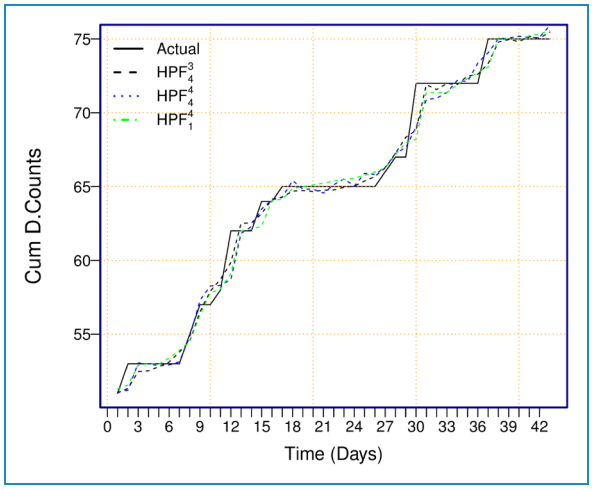

Figure 9. The observed and predicted values for the three best models for cumulative global death mpox cases.

Table 8. Comparison of the proposed best model and the benchmark models: out-of-sample cumulative death counts day-ahead mean forecast error.

Models	MAPE	MAE	RMSE
HPF$_3^4$	0.8538	0.5427	0.8244
AR	1.0560	0.6784	1.2387
NPAR	1.3145	0.8450	1.1621
ARMA	1.3984	0.9039	1.3044
NNR	1.2034	0.7722	1.2991

benchmark model in forecasting outcomes. Governments and stakeholders are advised to ensure adherence to standard operating procedures and prioritize vaccination to combat the expected growth in MPV trends.

Limitations of the study

The study is limited to cumulatively confirmed cases and death counts for mpox data; however it might be broadened to incorporate more elements like daily infection case counts and daily death rates. This would allow for the evaluation of the proposed hybrid time series forecasting system's performance. Additionally, the algorithm might be used to forecast COVID-19 confirmed cases, deaths and recovered cases over the next several days. Despite the fact that the suggested system only utilizes univariate linear and nonlinear time series models, in the future, multivariate time series models and machine learning models such as vector auto-regressive, vector ARIMA, ANN, support vector regression model, random forest and decision tree learning models can be incorporated.

Table 9. World mpox virus data: *p*-values of the DM test using the loss-squared function for cumulative death cases.

Models	HPF$_3^4$	AR	NPAR	ARMA	NNR
HPF$_3^4$	0	0.6	0.01	0.58	0
AR	0.4	0	0.02	0.46	0
NPAR	0.99	0.98	0	0.98	0.01
ARMA	0.42	0.54	0.02	0	0.01
NNR	1	1	0.99	0.99	0

Acknowledgements: The authors are grateful to E. Mathieu, F. Spooner, S. Dattani, H. Ritchie and M. Roser for making the data available for free on Our World in Data website (Uhttps://ourworldindata.org/monkeypox).

Contributorship: All authors contributed to the conceptualization, validation and visualization of the study. HI, MQ, MD and KT were involved in the data curation, formal analysis and methodology. HI, MQ, MD, and KT were responsible for the writing of the original draft. MD, KT, RKA, and JKA were involved in the writing of the review and editing. All listed authors contributed to the manuscript, approved its claims and agreed to be authors. All authors meet the ICMJE criteria.

Declaration of conflicting interest: The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding: The authors received no financial support for the research, authorship and/or publication of this article.

Guarantor: KT.

ORCID iDs: Muhammad Daniyal https://orcid.org/0000-0002-1415-4970
Kassim Tawaiah https://orcid.org/0000-0003-0195-225X

Data availability: The data used in this study is available at https://ourworldindata.org/monkeypox.

References

1. Molla J, Sekkak I, Mundo Ortiz A, et al. Mathematical modeling of MPOX: a scoping review. *One Health* 2023: 100540. https://doi.org/10.1016/j.onehlt.2023.100540

Table 10. World mpox virus data: the forecasted new cumulative cases and death counts using the best proposed model over 2 weeks.

Date	Cumulative new cases	Cumulative death cases	Daily new cases	Daily death cases
4 January 2023	84,063	76	0	0
5 January 2023	84,140	76	77	0
6 January 2023	84,189	76	49	0
7 January 2023	84,225	77	36	1
8 January 2023	84,264	77	39	0
9 January 2023	84,309	77	45	0
10 January 2023	84,356	77	47	0
11 January 2023	84,402	78	46	1
12 January 2023	84,448	78	46	0
13 January 2023	84,492	78	44	0
14 January 2023	84,537	78	45	0
5 January 2023	84,583	79	46	1
16 January 2023	84,628	79	45	0
17 January2023	84,673	79	45	0

2. Bankuru SV, Kossol S, Hou W, et al. A game-theoretic model of monkeypox to assess vaccination strategies. *PeerJ* 2020; 8: e9272.
3. Beeson A, Styczynski A, Hutson CL, et al. MPOX respiratory transmission: the state of the evidence. *The Lancet Microbe* 2023. https://doi.org./10.1016/S2666-5247(23)00034-4
4. Pan D, Nazareth J, Sze S, et al. Transmission of monkeypox/MPOX virus: a narrative review of environmental, viral, host, and population factors in relation to the 2022 international outbreak. *J Med Virol* 2023; 95: e28534.
5. White J, Rivero MJ, Mohamed AI, et al. Male sexual health implications of the 2022 global monkeypox outbreak. *Res Rep Urol* 2022: 415–421. https://doi.org./10.2147/RRu.S381191
6. Lakshmi M and Das R. Classification of monkeypox images using LIME-enabled investigation of deep convolutional neural network. *Diagnostics* 2023; 13: 1639.
7. Manohar B and Das R. Artificial neural networks for prediction of COVID-19 in India by using backpropagation. *Expert Syst* 2023; 40: e13105.
8. Jezek Z, Szczeniowski M, Paluku KM, et al. Human monkeypox: confusion with chickenpox. *Acta Trop* 1988; 45: 297–307.
9. McAndrew T, Majumder MS, Lover AA, et al. Early human judgment forecasts of human monkeypox, May 2022. *The Lancet. Digital Health* 2022; 4: e569–e571.
10. Qureshi M, Khan S, Bantan RA, et al. Modeling and forecasting monkeypox cases using stochastic models. *J Clin Med* 2022; 11: 6555.
11. Long B, Tan F and Newman M. Forecasting the monkeypox outbreak using ARIMA, prophet, NeuralProphet, and LSTM models in the United States. *Forecasting* 2023; 5: 127–137.
12. Khan MI, Qureshi H, Bae SJ, et al. Predicting monkeypox incidence: fear is not over! *J Infect* 2023; 86: 256–308.
13. Bleichrodt A, Dahal S, Maloney K, et al. Real-time forecasting the trajectory of monkeypox outbreaks at the national and global levels, July-October 2022. *BMC Med* 2023; 21: 19.
14. Islam M, Sangkham S, Tiwari A, et al. Association between global monkeypox cases and meteorological factors. *Int J Environ Res Public Health* 2022; 19: 15638.
15. Iftikhar H, Khan M, Khan Z, et al. A comparative analysis of machine learning models: a case study in predicting chronic kidney disease. *Sustainability* 2023; 15: 2754.
16. Iftikhar H, Khan M, Khan MS, et al. Short-term forecasting of monkeypox cases using a novel filtering and combining technique. *Diagnostics* 2023; 13: 1923.
17. Yasmin F, Hassan MM, Zaman S, et al. A forecasting prognosis of the monkeypox outbreak based on a comprehensive statistical and regression analysis. *Computation* 2022; 10: 77.
18. Saiprasad VR, Gopal R, Senthilkumar DV, et al. Monkeypox: a model-free analysis. *European Physical Journal Plus* 2023; 138: 38.
19. Manohar B and Das R. Artificial neural networks for the prediction of monkeypox outbreak. *Tropical Medicine and Infectious Disease* 2022; 7: 24.

20. Dada EG, Oyewola DO, Joseph SB, et al. Ensemble machine learning for monkeypox transmission time series forecasting. *Applied Sciences* 2022; 12: 12128.
21. Spicknall IH, Pollock ED, Clay PA, et al. Modeling the impact of sexual networks in the transmission of monkeypox virus among gay, bisexual, and other men who have sex with men—United States, 2022. *MMWR Morb Mortal Wkly Rep* 2022; 71: 1131–1135.
22. Pramanik A, Sultana S and Rahman MS. Time series analysis and forecasting of monkeypox disease using ARIMA and SARIMA Model. In: 2022 *13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. 2022, October, pp.1–7). IEEE. DOI: 10.1109/ICCCNT54827.2022.9984345.
23. Priyadarshini I, Mohanty P, Kumar R, et al. Monkeypox outbreak analysis: an extensive study using machine learning models and time series analysis. *Computers* 2023; 12: 36.
24. Gülmez B. A novel deep neural network model based Xception and genetic algorithm for detection of COVID-19 from X-ray images. *Ann Oper Res* 2022: 1–25. https://doi.org/10.1007/s10479-022-05151-y
25. Ahamed BS, Usha R and Sreenivasulu G. A deep learning-based methodology for predicting monkey pox from skin sores. In: 2022 *IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, 2022, p.1–6. IEEE. DOI: 10.1109/MysuruCon55714.2022.9972746.
26. Syed SF, Singh A, Rakhra M, et al. Monkey pox detection using deep learning. In 2022 *2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*. 2022, October, p.767–771. IEEE. DOI: 10.1109/ICTACS56270.2022.9987977.
27. AlZu'Bi S, AbuShanab S, AlMi'Ani MM, et al. transfer learning enabled CAD system for monkey pox classification. In: 2022 *9th International Conference on Internet of Things*: *Systems*, *Management* and *Security* (*IOTSMS*). 2022, pp.1–5. IEEE. DOI: 10.1109/IOTSMS58070.2022.10062163.
28. Biggerstaff M, Johansson M, Alper D, et al. Results from the second year of a collaborative effort to forecast influenza seasons in the United States. *Epidemics* 2018; 24: 26–33.
29. Manohar B and Das R. Comparison of hybrid artificial neural networks with GA, PSO, and RSA in predicting COVID-19 cases: a case study of India. In: *Multi-Disciplinary applications of fog computing: responsiveness in real-time*. IGI Global, 2023, pp.207–244. https://doi.org/10.4018/978-1-6684-4466-5.ch011
30. Chakraborty T, Chattopadhyay S and Ghosh I. Forecasting dengue epidemics using a hybrid methodology. *Physica A* 2019; 527: 121266.
31. Petropoulos F, Makridakis S and Stylianou N. COVID-19: forecasting confirmed cases and deaths with a simple time series model. *Int J Forecast* 2022; 38: 439–452.
32. Wang Y, Xu C, Li Y, et al. An advanced data-driven hybrid model of SARIMA-NNNAR for tuberculosis incidence time series forecasting in Qinghai Province, China. *Infect Drug Resist* 2020; 13: 867–880.
33. Rahman MM, Islam MM, Manik MMH, et al. Machine learning approaches for tackling novel coronavirus (COVID-19) pandemic. *SN computer Science* 2021; 2: 84.
34. Li J, Li Y, Ye M, et al. Forecasting the tuberculosis incidence using a novel ensemble empirical mode decomposition-based data-driven hybrid model in Tibet, China. *Infect Drug Resist* 2021; 14: 1941–1955.
35. Jutinico AL, Vergara E, García CEA, et al. Robust Kalman filter for tuberculosis incidence time series forecasting. *IFAC-PapersOnLine* 2021; 54: 424–429.
36. Safi SK, Sanusi OI and Tabash MI. Forecasting the impact of COVID-19 epidemic on China exports using different time series models. *Advances in Decision Sciences* 2022; 26: 1–25.
37. Prajapati S, Swaraj A, Lalwani R, et al. Comparison of traditional and hybrid time series models for forecasting COVID-19 cases. *arXiv preprint arXiv:2105.03266*. 2021.
38. Yan W, Xu Y, Yang X, et al. A hybrid model for short-term bacillary dysentery prediction in Yichang City, China. *Jpn J Infect Dis* 2010; 63: 264–270.
39. Cao S, Wang F, Tam W, et al. A hybrid seasonal prediction model for tuberculosis incidence in China. *BMC Med Inform Decis Mak* 2013; 13: –7.
40. Singh S, Parmar KS, Kumar J, et al. Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos Solitons FrActals* 2020; 135: 109866.
41. Mathieu E, Spooner F, Dattani S, et al. MPOX (monkeypox). *Our World in Data* 2022. https://ourworldindata.org/monkeypox
42. Dritsaki M and Dritsaki C. Comparison of HP filter and the Hamilton's regression. *Mathematics* 2022; 10: 1237.
43. Jurečková J, Arslan O, Güney Y, et al. Nonparametric tests in linear model with autoregressive errors. *Metrika* 2023; 86: 443–453.
44. Yu L, Cao Y, Zhou C, et al. Landslide susceptibility mapping combining information gain ratio and support vector machines: a case study from Wushan segment in the three gorges reservoir area, China. *Applied Sciences* 2019; 9: 4756.
45. Wang J, Wang Y, Li Z, et al. A combined framework based on data preprocessing, neural networks and multi-tracker optimizer for wind speed prediction. *Sustainable Energy Technologies and Assessments* 2020; 40: 100757.
46. Shah I, Iftikhar H and Ali S. Modeling and forecasting electricity demand and prices: a comparison of alternative approaches. *Journal of Mathematics* 2022; 2022. https://doi.org/10.1155/2022/3581037
47. Eya CU, Salau AO, Braide SL, et al. Improved medium term approach for load forecasting of Nigerian electricity network using artificial neuro-fuzzy inference system: a case study of University of Nigeria, Nsukka. *Procedia Comput Sci* 2023; 218: 2585–2593.
48. Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2020; 20: 553–558.
49. Alshanbari HM, Iftikhar H, Khan F, et al. On the implementation of the artificial neural network approach for forecasting different healthcare events. *Diagnostics (Basel, Switzerland)* 2023; 13: 1310.
50. Diebold FX. Comparing predictive accuracy, twenty years later: a personal perspective on the use and abuse of Diebold–Mariano Tests. *J Bus Econ Stat* 2015; 33: 1.
51. Barzegar R, Aalami MT and Adamowski J. Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. *Stoch Environ Res Risk Assess* 2020; 34: 415–433.

52. Mitra A and Wang L. Improving artificial neural network based stock forecasting using Fourier de-noising and Hodrick-Prescott Filter. In: 2015 *10th International Conference on Information, Communications and Signal Processing (ICICS)*. 2015, p.1–5. DOI: 10.1109/ICICS.2015.7459920.
53. Lounis M and Riad A. Monkeypox (MPOX)-related knowledge and vaccination hesitancy in non-endemic countries: concise literature review. *Vaccines (Basel)* 2023; 11: 29.
54. Haripriya KP and Inbarani HH. Performance analysis of machine learning classification approaches for monkey pox disease prediction. In: 2022 *6th International Conference on Electronics, Communication and Aerospace Technology*. 2022, p.1045–1050. IEEE. DOI: 10.1109/ICECA55336.2022.10009407.