



Article

Short-Term Forecasting of Monkeypox Cases Using a Novel Filtering and Combining Technique

Hasnain Iftikhar ^{1,2,*}, Murad Khan ³, Mohammed Saad Khan ⁴ and Mehak Khan ^{5,*}

- ¹ Department of Mathematics, City University of Science and Information Technology, Peshawar 25000, Pakistan
- ² Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan
- ³ Department of Statistics, Abdul Wali Khan University Mardan, Mardan 23200, Pakistan
- ⁴ Faculty of Computer Sciences and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Swabi 23640, Pakistan
- ⁵ Department of Computer Science, AI Lab, Oslo Metropolitan University, P.O. Box 4 St. Olavs Plass, 0130 Oslo, Norway
- * Correspondence: hasnain@stat.qau.edu.pk (H.I.); mehakkha@oslomet.no or mehakkhan3@hotmail.com (M.K.)

Abstract: In the modern world, new technologies such as artificial intelligence, machine learning, and big data are essential to support healthcare surveillance systems, especially for monitoring confirmed cases of monkeypox. The statistics of infected and uninfected people worldwide contribute to the growing number of publicly available datasets that can be used to predict early-stage confirmed cases of monkeypox through machine-learning models. Thus, this paper proposes a novel filtering and combination technique for accurate short-term forecasts of infected monkeypox cases. To this end, we first filter the original time series of the cumulative confirmed cases into two new subseries: the long-term trend series and residual series, using the two proposed and one benchmark filter. Then, we predict the filtered subseries using five standard machine learning models and all their possible combination models. Hence, we combine individual forecasting models directly to obtain a final forecast for newly infected cases one day ahead. Four mean errors and a statistical test are performed to verify the proposed methodology's performance. The experimental results show the efficiency and accuracy of the proposed forecasting methodology. To prove the superiority of the proposed approach, four different time series and five different machine learning models were included as benchmarks. The results of this comparison confirmed the dominance of the proposed method. Finally, based on the best combination model, we achieved a forecast of fourteen days (two weeks). This can help to understand the spread and lead to an understanding of the risk, which can be utilized to prevent further spread and enable timely and effective treatment.

Keywords: monkeypox virus; short-term forecasting; machine learning models; time series models; filtering and combining technique



Citation: Iftikhar, H.; Khan, M.; Saad Khan, M.; Khan, M. Short-Term Forecasting of Monkeypox Cases Using a Novel Filtering and Combining Technique. *Diagnostics* **2023**, *13*, 1923. <https://doi.org/10.3390/diagnostics13111923>

Academic Editor: Hiroshi Ikegaya and Shaker El-Sappagh

Received: 12 May 2023
Revised: 27 May 2023
Accepted: 29 May 2023
Published: 31 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In today's world, humans have made incredible advances in the fields of science, technology, and artificial intelligence (AI) by inventing remote-controlled drones, automated responsive robots, and self-driving cars. However, they continue to face numerous natural disasters challenges, such as floods, earthquakes, and droughts, as well as novel viral diseases, including the coronavirus infectious disease 2019 (COVID-19) and the monkeypox virus (MV). The COVID-19 pandemic was one of the worst disasters in human history and caused 6.8 million deaths around the world as of 17 February 2023 [1]. Moreover, it is on the decline and is in the endemic phase, while the world is about to face another crisis in the form of a new viral disease outbreak of MV. MV is transmitted from one human to another through physical contact with the one who is infected, with contaminated matter, or with

infected animals. The symptoms of MV include a rash on the skin, a mucosal lesion, muscle aches, headaches, back pain, fever, low energy, swollen lymph nodes, and a fever that can last for around 2 to 4 weeks. It can be difficult to identify and distinguish MV due to its similarity to other infectious diseases. It is important to differentiate MV from chickenpox, bacterial skin infections, capsules, herpes, scabies, other sexually transmitted infections, and medicine-related allergies. Therefore, diagnostic tests are necessary to distinguish it from other similar diseases, to obtain treatment as soon as possible, and to stop the spread. The most preferred laboratory diagnostic test for MV is the detection of viral DNA by polymerase chain reaction. MV is an infectious disease caused by the MV, a species of the Orthopoxvirus *genus* and a causative agent of smallpox. There are two different clades of MV: Clade I and Clade II. MV was first detected in Copenhagen, Denmark, in 1958 [2,3].

Meanwhile, the first formally documented case was observed in a nine-year-old child in the Democratic Republic of the Congo (DRC) in 1970 [4,5]. Initially, the MV outbreak was limited to the continent of Africa but gradually spread to North America and Europe. Furthermore, confirmed MV cases were recorded in African countries, including one case in Cote d'Ivoire, one in Cameroon, three cases in Nigeria, four in Liberia, and thirty-eight in Congo, between 1970 and 1979. In the year 1986, the total confirmed cases reached 400, with a 10% mortality rate. Moreover, between 1991 and 1999, some small outbreaks were also observed in West Africa and the equatorial central region, and Congo alone observed 500 cases. After that, the continent of Africa witnessed a decline and an endemic stage of the MV outbreak [6,7]. Since 2005, thousands of suspected MV infectious cases have been recorded every year in the DRC. In 2017, MV re-emerged and spread among the people of Nigeria through travelers across the country. However, in May 2022, the sudden increase and rapid spread of the MV outbreak were seen across Europe, America, and all six regions of the WHO, with the reporting of around 87 thousand confirmed cases and 112 deaths in 110 countries. This variant of MV was due to Clade I of MV, and it was detected in a refugee camp in the Republic of Sudan. Later, many countries on different continents, including Asia, Australia, Africa, South America, North America, and Europe, witnessed the same outbreak of MV.

The MV outbreak was declared a “public health emergency of international concern” (PHEIC) by the World Health Organization (WHO) after reporting 82,000 confirmed cases on 23 July 2022 [8]. Recently, the multi-country MV outbreak has attracted global attention, as a cumulative total of 86,017 confirmed cases and 97 deaths have been reported as of 17 February 2023, worldwide in over 111 countries [9,10]. On 11 May 2023, the WHO declared that the emergence of the MV outbreak of international concern in multi-countries is over [11]. The WHO also published a response plan and guidance documents on strategic preparedness to curb the spread of MV. The proper diagnostics, surveillance, risk management, and engagement of the community are most important to stop the spread and eliminate the human-to-human spread of MV. MV poses a major threat to global health and the economy due to the non-availability of a dedicated vaccine, cure, and on-time detection; therefore, it needs community awareness, the education of health workers, and timely forecasting to prevent its spread [12]. For this purpose, time series and machine learning models are needed for modeling and forecasting the daily cumulative deaths and confirmed cases of MV to assist governments and stakeholders in controlling the spread.

In the past, several statistical and machine learning (ML) models have been widely used by researchers for the prediction of various diseases and pandemics [13–19]. There has been tremendous progress in the use of AI tools to detect, diagnose, and classify diseases [20,21]. Hence, it is possible to implement AI models for the disease classification, monitoring, and forecasting of MV outbreaks [22,23]. The researchers in [24] comparatively evaluated the performance of machine-learning-based multilayer perception models (MLP) and (autoregressive integrated moving average) ARIMA models for the forecasting of the cumulative confirmed cases of MV and found that MLP with a sigmoid function outperformed the traditional ARIMA model in forecasting the actual cases of MV. Further, the human historical monkeypox cases were used for the prediction of the MV

transmission rate by the application of stack ensemble learning (SEL) and ML techniques including random forest, adaptive boosting regression (Adaboost), and gradient boosting (Gboost). The experimental results revealed that the SEL outperformed the other used algorithms [25]. For seven-day forecasting of MV cases in the USA, the five ML and time series models, including ARIMA, long short-term memory (LSTM), prophet, neural prophet, and a stacking model, were used, and 95% accuracy was achieved in the neural prophet output [26]. The researchers in [27] proposed an optimized hybrid deep learning approach based on the Al-Biruni earth-radius-based (BER) optimization technique and an optimized LSTM for forecasting confirmed cases of MV outbreaks with high accuracy and low error. Furthermore, researchers have analyzed the spread of MV across multiple countries by using various ML models, including linear regression (LR), random forest (RF), elastic net regression (EN), convolutional neural network (CNN), and artificial neural network (ANN), and found the CNN outperformed the other methods used in forecasting MV outbreaks. They have also used time series models, ARIMA, and seasonal autoregressive integrated moving averages (SARIMA) to measure the events' occurrence over time [28]. In further studies, the ARIMA and feedforward neural network (FFNN) were used for the daily confirmed and death cases of MV outbreaks for the next two months; their forecasted estimates were around 87,276 confirmed and 94 deaths up to 31 January 2023, with a 95% confidence interval [29]. The authors in [30] performed the sentiment analysis of tweets related to the illness of MV by analyzing the three possible responses of positive, negative, and neutral with 94% accuracy via the proposed hybrid deep learning approach based on CNN-LSTM. The researchers in [31] conducted a study for a short-term forecast for 10 weeks of calibration by applying n-sub-epidemic models and found an accurate prediction of the declining trend in MPV cases globally and country-specific cases in the last subsequent period of ten days. Their findings also revealed that the population with behavioral modifications and increased immunity were at higher risk of being affected by viruses. Further, the authors in [32] implemented a lag-correlation analysis, a gradient causality test, and a vector autoregression model to forecast the monkeypox epidemic in 20 countries and found a robust association between 13 days of priority and daily confirmed cases globally.

In this paper, we propose a novel filtering and combination technique for the accurate and efficient short-term forecasting of cumulative confirmed cases of MV. The proposed methodology is based on various filters and machine learning models. The steps of the proposed methodology are as follows: First, we decompose the original time series of the cumulative confirmed cases into new subseries, that is, the long-term trend series and residual series, using the two proposed and one benchmark filter, including the regression spline filter (RSF), smoothing spline filter (SSF), and the Hodrick–Prescott filter (HPF). Then, to predict the decomposed subseries, we consider five well-known machine learning models, including artificial neural network (ANN), support vector machine with two kernels (linear (SVM1) and spline (SVM2)), random forest (RF), decision tree learning (DT), and all their possible combination models. Hence, we combine the individual predictive models directly to obtain the final predictions one day ahead of the cumulatively infected cases of MV. The contributions of this work can be summarized as follows: A novel filtering and combination approach is proposed based on different filters and various combinations of ML models to improve the accuracy of the MV-confirmed-case day-ahead forecasts. First, we verify the performance of the proposed filters compared to the standard benchmark filter method using accuracy measures and a statistical test. Second, within the proposed methodology, we compare the performance of the different combinations of the considered machine learning models using two proposed filters and a benchmark filter. Third, the proposed final combination model is compared with the standard time series and the ML models, and the comparative results are noted. The noted results showed that the proposed final model is highly accurate and efficient for forecasting the daily confirmed cases of MV as compared to the benchmark models. Finally, the proposed methodology can be generalized and tested for other datasets.

The rest of the paper is organized as follows: Section 2 describes the general procedure of the proposed filtering and combination forecasting methodology. Section 3 provides an empirical application of the proposed modeling framework using the daily cumulative time series of the MV. Section 4 comprises a discussion of the proposed best combination model versus some of the best time series and machine learning models. Finally, Section 5 addresses the concluding remarks and future research directions.

2. The Proposed Filtering and Combination Technique

This section explains the proposed filtering and combination forecasting technique for the short-term cumulative confirmed cases forecast. To achieve this, the time series of the cumulative confirmed cases (N_t) is decomposed into two subseries: the long-term nonlinear trend (L_t) and a residual subseries (R_t) using two new proposed filters and a considered benchmark filter, including the regression spline filter, the smoothing spline filter, and the Hodrick–Prescott Filter. The mathematical representation of the decomposed subsequence is formulated as

$$N_t = L_t + R_t. \quad (1)$$

Therefore, for the modeling and forecasting purposes, the long-run nonlinear trend L_t is a function of time t and the residual subseries, which describes the short-run dependence of the cumulative series and is obtained by $R_t = (N_t - L_t)$. Therefore, the new proposed filters and the benchmark filter are described in the following subsection.

- Regression Spline Filter

A regression spline is a general nonparametric approximation of N_t by a piecewise m^{th} degree polynomial, estimating a subinterval bounded by a series of m points (called knots). Any spline function $U(N)$ of order q can be defined as a linear combination of functions $U_i(N)$ called basis functions, whose formula is given by increase.

$$U(N) = \sum_{i=1}^{t+m+1} \alpha_i U_i(N). \quad (2)$$

The unknown parameter is α_i , estimated by the ordinary least squares method. The most important choices are the number of nodes and their positions that define the smoothness of the approximation. In this work, we used cross validation to estimate these quantities.

- Smoothing Splines Filter

To meet the requirements for resolving the knot regions, spline features can be predicted using a least-squares penalty environment to limit the sum of squares. Hence, the equation can be written as

$$\sum_{j=1}^N (N_t - U(N))^2 + \lambda \int (U''(N))^2 dt, \quad (3)$$

where $(U''(N))$ is the second derivative of $U(N)$. The first term describes the goodness of fit, and the second term penalizes the coarseness of the function by the smoothing parameter λ . Moreover, the selection of smoothing parameters is a difficult task and is performed by cross-validation methods in this work.

- Hodrick–Prescott Filter

To assess the performance of the two proposed filters, they are compared to a standard benchmark filter, the Hodrick–Prescott Filter (HPF). The HPF is used to obtain a smoothed-curve representation of a time series that is more complex for long-term than short-term fluctuations. The adjustment of the sensitivity of the trend to short-term fluctuations is achieved by modifying a multiplier λ . Let N_t ($t = 1, 2, \dots, N$) be denoted the time series data.

The series N_t is made up of a trend component denoted by τ_t , and an error component, denoted by ε ; therefore, the equation is

$$N_t = \tau_t + \varepsilon_t. \quad (4)$$

Here, we insert this into in a long-term trend component that can be estimated by minimizing the following expression,

$$\min_t \left(\sum_{t=1}^M (N_t - \tau_t)^2 + \lambda \sum_{t=2}^{M-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right). \quad (5)$$

In the above equation, the first term is the loss function, and the second term is a plenty term λ multiplied by the sum of the square of the trend component second difference, which penalizes the variation in the growth rate of the trend component.

To graphically demonstrate the performance of the proposed filters and the considered benchmark filter (HPF) presented above, the new subseries are shown in Figure 1. In each subfigure, at the first position is a nonlinear increasing trend (L_t) subseries, and in the second position is a residual (R_t) subseries. From the figure, we observe that the proposed and the benchmark filters decomposed (N_t) and captured the long-term nonlinear trend well.

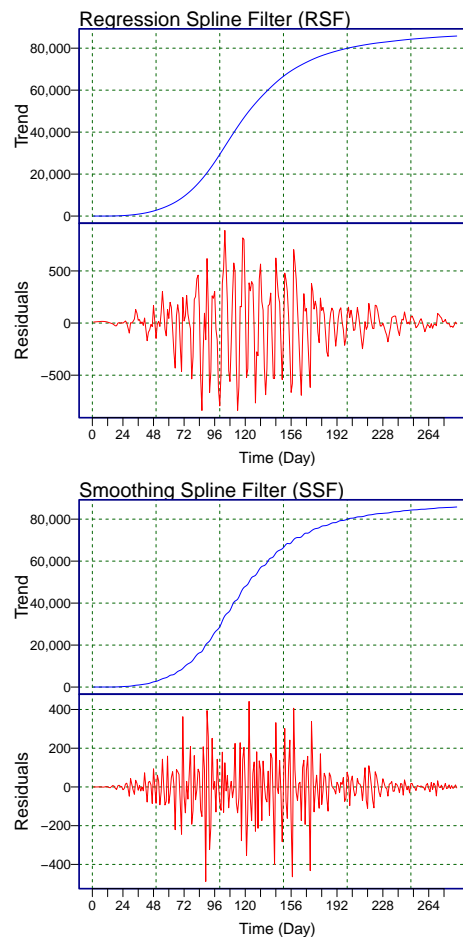


Figure 1. Cont.

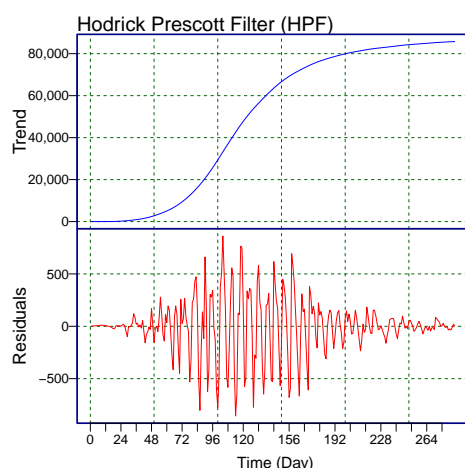


Figure 1. World Monkeypox Virus Data: the daily confirmed cases of the monkeypox virus are filtered by the two proposed filters: (top) RSF, (middle) SSF, and the benchmark filter HPF (bottom). Within each subfigure, the top panel shows the long-term trend (blue curve- L_t), and the bottom panel shows the residual part (red curve- R_t).

2.1. Modeling to Filtered Series

Once the filtered subseries are extracted from the time series of the daily cumulative confirmed cases using the two new proposed filters and a benchmark filter, we estimate the extracted subseries using five standard machine learning models, including artificial neural network, support vector machine with two kernels (linear and spline), random forest, and decision tree learning. Therefore, all the considered models are explained as follows:

- Nonlinear autoregressive neural network

A nonlinear autoregressive neural network (ANN) is a type of machine learning model that is useful for forecasting the future values of an input variable. The ANN network predicts the future values of a time series based on its history using a re-feeding mechanism, in which an expected value can be used as an input for new predictions at later points in time [33]. The network is built and trained in an open loop, using actual target values as feedback to ensure greater training accuracy. The network is converted into a closed loop after training, and the predicted values are used to supply new feedback inputs to the network.

Mathematically, the model predicts the future values of a time series N_t based on its historical values $N_{t-1}, N_{t-2}, \dots, N_{t-d}$, where d is the time delay parameter. The backpropagation algorithm is used to train the network, and the steepest descent method is used to minimize the square error between the actual and predicted values.

- Support Vector Machine

The statistical learning theory and the concept of structural risk reduction, which were first introduced by Cortes and Vapnik in 1995, serve as the foundation for the machine learning algorithm known as support vector machine (SVM). One of the most widely used methods for supervised learning, the SVM is used to solve classification and regression issues [34]. The SVM is quick, easy to operate, consistent, and generates accurate results. On the other hand, SVM models employ a variety of fundamental kernel operations. The functions can be categorized as a polynomial, linear, sigmoid, and exponential radial basis function and a radial basis function. Complex nonlinear decision boundaries can be modeled by the SVM technique. In addition, the SVM model is effective in time series forecasting because it can resolve issues with nonlinear regression estimates. The formula listed below determines the SVM model for a given dataset.

Consider an n set of data $(x_1, y_1), \dots, (x_n, y_n)$, where x_i is the i th input vector, and y_i is the corresponding desired output. Because $i = 1, 2, \dots, n$, where n is the size of the sample, the estimating function assumes the following form:

$$\delta_n = \omega \cdot \phi(\delta_n) + \mu, \tag{6}$$

where ω is the weight vector, μ is the bias, $\phi(\delta_n)$ is the high-dimensional feature space nonlinearly mapped from the input space, and (\cdot) represents the inner product. This work considers two kernels, the linear (SVM1) and spline (SVM2) kernels.

- Random Forest Model

A well-known machine learning algorithm that belongs to the supervised learning class is called the random forest (RF). Random forest often referred to as random decision forest, is an ensemble learning technique for classification and regression that operates by putting several decision trees through training. The RF algorithm determines the outcome based on the decision trees' predictions [35]. The output of different trees is averaged to produce forecasts. As the number of trees increases, the accuracy of the result increases. The more trees in the forest, the more accurate it is, and the issue of overfitting is avoided. To train tree learners, the random forest training algorithm employs the widely used method of bootstrap aggregation or bagging [36]. Bagging repeatedly (M times) takes a random sample with replacement of the training set and fits trees to these samples given a training set $Y = y_1, \dots, y_n$ with responses $X = x_1, \dots, x_n$.

For $m = 1, \dots, M$:

1. m training instances from Y, X are sampled with replacement; they are referred to as Y_m, X_m .

2. On Y_m, X_m , we train a classification or regression tree f_m . After training, summing the predictions from all the various regression trees on y' can be used to make predictions for the unseen samples y' :

$$\hat{f} = \frac{1}{M} \sum_{m=1}^M f_m(y'). \tag{7}$$

This bootstrap technique enhances the model performance by reducing the model variance without increasing the bias. The majority vote is used in classification tree cases.

- Decision Tree Learning

Data mining frequently uses the decision tree learning technique [37]. The goal is to develop a model that predicts the value of a target parameter given a set of input parameters. A tree can be trained to learn by subdividing the source dataset depending on an attribute value test [38]. With the use of a recursive partitioning process, a decision tree algorithm divides a training dataset, $Y = \{Y_1, Y_2, \dots, Y_N\} \rightarrow X = \{X_1, X_2, \dots, X_N\}$, to create a model. Note that each observation $Y_i \rightarrow X_i$ could include M_f features, $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{iM_f})$. If the data at node m are Q with M_m observations, then for each candidate partition $\pi = (f, t_m)$ consisting of a feature f and threshold t_m , the data Q are split into $Q_{left}(\pi)$ and $Q_{right}(\pi)$ subsets, such that

$$Q_{left}(\pi) = (Y, X) | Y_f < t_m,$$

and

$$Q_{right}(\pi) = Q \setminus Q_{left}(\pi).$$

The impurity at node n after each partition is defined as:

$$I(Q, \pi) = \frac{M_{left}}{M_m} G(Q_{left}(\pi)) + \frac{M_{right}}{M_m} G(Q_{right}(\pi)).$$

$G(\dots)$ is the impurity function in this situation.

$$G(Q) = \frac{1}{M_m} \sum_{i \in M_n} (X_i - \bar{X}_m)^2,$$

$$\bar{X}_m = \frac{1}{M_m} \sum_{i \in M_n} X_i.$$

By reducing the impurity $I(Q, \pi)$, the decision tree model's parameters are chosen.

$$\pi^* = \operatorname{argmin}_{\pi} I(Q, \pi).$$

Recursively, the identical partitioning operation is carried out for the subsets $Q_{left}(\pi^*)$ and $Q_{right}(\pi^*)$ up until the maximum permitted depth is reached, or there is only one observation left ($M_m = 1$). The model may then be used to forecast the value of a target variable (X_{test}) based on the independent variables (Y_{test}) after the decision tree has been trained with the training dataset, Y and X .

In the current study, we denote each combined model with each filter method by $RSF_{R_t}^{L_t}$, where the L_t at the top right is associated with a nonlinear trend subseries, and the R_t at bottom right is associated to the residual subseries. In the forecasting models, we assign a code to each model: "0" for the ANN, "1" for the SVM1, "2" for the SVM2, "3" for the RF, and "4" for the DT. For example, RSF_1^0 represents the estimate of the long-term trend (L_t) with the ANN, and the residual series (R_t) estimated using the SVM1. The individual forecast models are summed to obtain the final one-day-ahead cumulative confirmed cases to forecast.

$$\hat{N}_{t+1} = (\hat{L}_{t+1} + \hat{R}_{t+1}). \tag{8}$$

2.2. Accuracy Measures

In the literature, many researchers have used various accuracy measures and statistical tests to check the performance of predictive models [39]. However, in this work, we initially use four accuracy mean errors for the proposed evaluation of all sixteen combination models, such as the root mean square percentage error (RMSPE), the root mean square error (RMSE), the mean absolute error (MAE), and the mean absolute error percentage (MAPE). The mathematical equations for these mean errors are the following:

$$RMSPE = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(\frac{|N_t - \hat{N}_t|}{|N_t|} \right)^2} \times 100, \tag{9}$$

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left(\frac{|N_t - \hat{N}_t|}{|N_t|} \right) \times 100, \tag{10}$$

$$MAE = \frac{1}{T} \sum_{t=1}^T (|N_t - \hat{N}_t|), \tag{11}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (N_t - \hat{N}_t)^2}, \tag{12}$$

where N is the number of observations in the dataset, (N_t) and (\hat{N}_t) are the T^{th} estimated and observed data points, respectively.

The layout of the proposed filtering and combination technique is shown in Figure 2.

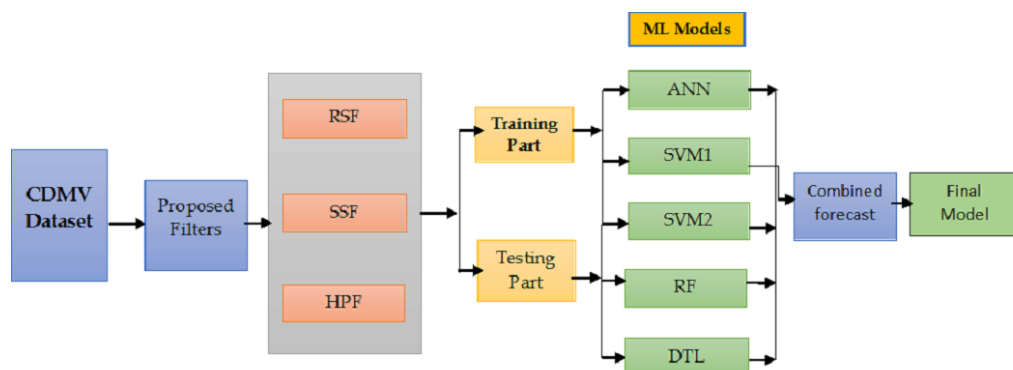


Figure 2. A flowchart of the proposed filtering and combination technique.

3. Data Description and Case Study Results

The main aim of this work is to provide a short-term forecast of the cumulatively infected cases of MV using a dataset from the entire world. The dataset (daily cumulative confirmed cases) of MV was taken from the official website of “Our World in Data” from 7 May 2022 to 10 February 2023. The graphical presentation and the descriptive statistics of daily and cumulative confirmed cases can be seen in Figure 3 and Table 1. Figure 3 (left) shows the daily cumulative infected cases of MV, and Figure 3 (right) shows the daily new infected cases of MV. Further, Figure 3 (left) shows an increasing nonlinear curve, while Figure 3 (right) shows low infected cases at the start and end of the considered series and high cases in the middle of the considered series. It was confirmed that at the start, MV had a low number of infections, and with time, it gradually increased, but after that, it again showed a gradual decline in the infected cases. On the other hand, the descriptive statistics are tabulated in Table 1. This table shows that the minimum number of confirmed cases of MV was zero, and the maximum number was 1814. Moreover, the average number of infected cases was 300 throughout the world, while the variation among the infected cases was observed by the standard deviation with a value of 406. Therefore, the complete dataset of the daily cumulative confirmed cases covering 286 days was split; 7 May to 11 November 2022 (214 days) was used for model training, and 12 November 2022 to 10 February 2023 (72 days) was used for the one-day-ahead cumulative confirmed cases post-sample (testing) forecasts.

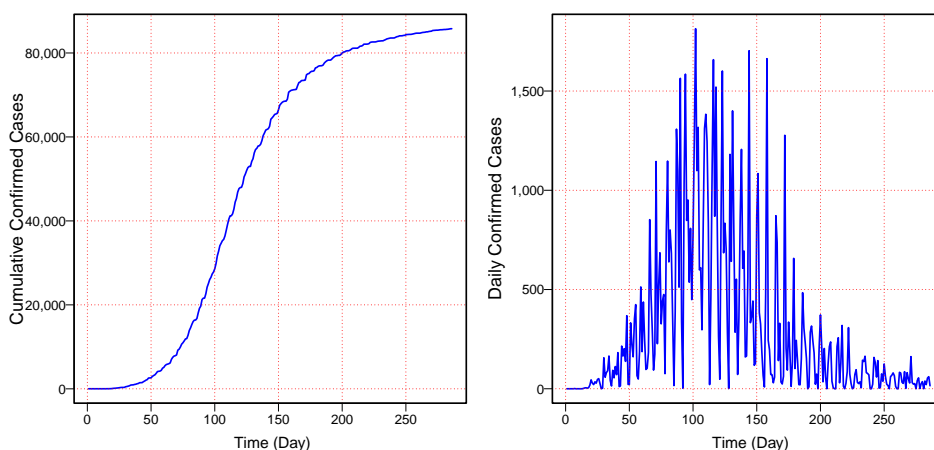


Figure 3. World Monkeypox Virus Data: the cumulative confirmed cases of monkeypox virus (left) and the daily confirmed cases of monkeypox virus (right) from 7 May 2022 to 10 February 2023.

Table 1. Descriptive statistics for the daily confirmed cases of the monkeypox virus dataset.

S.No	Statistic	Value
1	Minimum	0
2	First Quartile	27.00
3	Median	125.50
4	Mean	299.90
5	Variance	164,481.70
6	Standard Deviation	405.56
7	Skewness	1.76
8	Kurtosis	2.39
9	Third Quartile	381.00
10	Maximum	1814.00

To obtain the forecast for the daily confirmed MV cases a day ahead, using the filtering and combination forecasting technique described in Section 2, the following steps were followed: first, the new two proposed filters and a benchmark filter were used to obtain a long-term nonlinear trend (L_t), and residual (R_t) time subseries. Second, the previously described five well-known machine learning models were applied to each subseries. Thereby, the models were estimated, and a day-ahead forecast for 72 days was obtained using the rolling window method. The final daily confirmed MV cases day-ahead forecasts were obtained using Equation (8). The accuracy measures RMSPE, RMSE, MPAE, and MAE were then used to evaluate and compare the performance of the models.

The actual time series of the daily cumulative confirmed MV cases (N_t) was divided into a long-term nonlinear trend (L_t) and a residual subseries (R_t), and two proposed filters and a benchmark filter were used in this work. Forecasts for these subseries were obtained using five machine learning models. To this end, combining the model and subseries forecasts, there were ($5^{L_t} \times 5^{R_t} = 25$) different combinations for each proposed filter. Thus, there were two proposed filters (RSF and SSF) and one benchmark filter (HPF), for a total of 75 (3×25) models. For these 75 models, the out-of-sample forecast accuracy measures for one day ahead (RMSPE, RMSE, MPAE, and MAE) are tabulated in Tables 2–4. The results of the performance measures showed that the RSF_1^2 model produced a better prediction than all the other models using the RS filter. The best forecasting model was RSF_1^2 , which produced 0.1452, 1.2083, 1010.7360, and 1207.9950 for the RMSPE, MAPE, MAE, and RMSE, respectively. However, the RSF_1^1 and RSF_1^3 models produced the second and third-best results. On the other hand, using the SS filter, the lowest forecast errors were found by the SSF_1^2 model with the values of 0.1415, 1.2162, 1020.2140, and 1182.9890 for the RMSPE, MAPE, MAE, and RMSE, respectively. Notwithstanding, the second- and third-best results were achieved by the SSF_1^1 , and SSF_1^0 models, respectively. In contrast, the benchmark filter (HPF) was outperformed by the proposed filters. Therefore, it was confirmed from Tables 2–4, within the proposed filtered and benchmark filters, the RS filter produced the lowest mean errors. Moreover, within all the possible combination models ($5^{L_t} \times 5^{R_t} = 25 = 75$), the RSF_1^2 was declared the best model.

Table 2. World Monkeypox Virus Data: out-of-sample one-day-ahead mean forecast error for all combination models using the RS filter.

S.No	Models	RMSPE	MAPE	MAE	RMSE
1	RSF_0^0	0.2442	1.9733	1642.6080	2022.8460
2	RSF_0^1	0.2467	2.0019	1666.6090	2043.4370
3	RSF_0^2	0.3258	2.9147	2434.1050	2706.2880
4	RSF_0^3	0.2459	1.9922	1658.4570	2036.4830
5	RSF_0^4	0.2445	1.9769	1645.6030	2025.4190

Table 2. Cont.

S.No	Models	RMSPE	MAPE	MAE	RMSE
6	RSF ₁ ⁰	0.1540	1.3494	1139.0140	1303.5960
7	RSF ₁ ¹	0.1529	1.3401	1130.9040	1293.1590
8	RSF ₁ ²	0.1452	1.2083	1010.7360	1207.9950
9	RSF ₁ ³	0.1532	1.3428	1133.2750	1296.6320
10	RSF ₁ ⁴	0.1539	1.3481	1137.8680	1302.2670
11	RSF ₂ ⁰	0.4953	4.7261	3957.2110	4130.2120
12	RSF ₂ ¹	0.4981	4.7562	3982.5270	4154.4740
13	RSF ₂ ²	0.5863	5.6693	4750.3480	4895.3990
14	RSF ₂ ³	0.4972	4.7460	3973.9930	4146.2930
15	RSF ₂ ⁴	0.4956	4.7298	3960.3820	4133.2500
16	RSF ₃ ⁰	0.2988	2.6629	2255.5920	2539.5080
17	RSF ₃ ¹	0.2962	2.6345	2231.6820	2517.1330
18	RSF ₃ ²	0.2203	1.8700	1585.7510	1874.4260
19	RSF ₃ ³	0.2971	2.6441	2239.7430	2524.6700
20	RSF ₃ ⁴	0.2985	2.6594	2252.5970	2536.7020
21	RSF ₄ ⁰	0.2392	1.9147	1593.3890	1980.4850
22	RSF ₄ ¹	0.2416	1.9428	1616.9220	2000.8470
23	RSF ₄ ²	0.3201	2.8522	2381.5920	2659.1560
24	RSF ₄ ³	0.2408	1.9332	1608.8620	1993.9700
25	RSF ₄ ⁴	0.2395	1.9182	1596.2950	1983.0290

Note: Regression spline filter (RSF). In the forecasting models, we assign a code to each model: “0” for the ANN, “1” for the SVM1, “2” for the SVM2, “3” for the RF, and “4” for the DT. For example, RSF₁⁰ represents the estimate of the long-term trend (L_t) with the ANN and the residual series (R_t) estimated using the SVM1.

Table 3. World Monkeypox Virus Data: out-of-sample one-day-ahead mean forecast error for all combination models using the SS filter.

S.No	Models	RMSPE	MAPE	MAE	RMSE
1	SSF ₀ ⁰	0.2453	1.9858	1653.0550	2031.8290
2	SSF ₀ ¹	0.2457	1.9904	1656.9690	2035.2010
3	SSF ₀ ²	0.2732	2.3178	1932.1830	2265.4930
4	SSF ₀ ³	0.2440	1.9712	1640.8210	2021.3110
5	SSF ₀ ⁴	0.2449	1.9812	1649.2080	2028.5190
6	SSF ₁ ⁰	0.1439	1.2690	1067.9890	1211.4960
7	SSF ₁ ¹	0.1438	1.2680	1067.0680	1210.6070
8	SSF ₁ ²	0.1415	1.2162	1020.2140	1182.9890
9	SSF ₁ ³	0.1442	1.2722	1070.8680	1214.3620
10	SSF ₁ ⁴	0.1440	1.2700	1068.8940	1212.3830
11	SSF ₂ ⁰	0.4928	4.7009	3936.0090	4109.9020
12	SSF ₂ ¹	0.4933	4.7058	3940.1540	4113.8720
13	SSF ₂ ²	0.5248	5.0340	4216.1620	4378.9440
14	SSF ₂ ³	0.4914	4.6855	3923.0550	4097.4980
15	SSF ₂ ⁴	0.4924	4.6960	3931.9360	4106.0020
16	SSF ₃ ⁰	0.2833	2.4950	2114.1720	2408.0140
17	SSF ₃ ¹	0.2828	2.4904	2110.2570	2404.4040
18	SSF ₃ ²	0.2549	2.2077	1871.5680	2168.3610
19	SSF ₃ ³	0.2846	2.5095	2126.4060	2419.3060
20	SSF ₃ ⁴	0.2837	2.4996	2118.0190	2411.5630
21	SSF ₄ ⁰	0.2392	1.9151	1593.7240	1980.7790
22	SSF ₄ ¹	0.2396	1.9197	1597.5240	1984.1050
23	SSF ₄ ²	0.2668	2.2426	1868.9750	2211.8310
24	SSF ₄ ³	0.2380	1.9010	1581.8500	1970.4030
25	SSF ₄ ⁴	0.2388	1.9107	1589.9910	1977.5130

Note: Smoothing spline filter (SSF). In the forecasting models, we assign a code to each model: “0” for the ANN, “1” for the SVM1, “2” for the SVM2, “3” for the RF, and “4” for the DT. For example, SSF₁⁰ represents the estimate of the long-term trend (L_t) with the ANN and the residual series (R_t) estimated using the SVM1.

Table 4. World Monkeypox Virus Data: out-of-sample one-day-ahead mean forecast error for all combination models using the HP filter.

S.No	Models	RMSPE	MAPE	MAE	RMSE
1	HPF ⁰ ₀	0.2476	2.0125	1675.5110	2050.9090
2	HPF ¹ ₀	0.2468	2.0035	1667.9020	2044.5200
3	HPF ² ₀	0.3663	3.3588	2807.5720	3046.5700
4	HPF ³ ₀	0.2474	2.0105	1673.7750	2049.4510
5	HPF ⁴ ₀	0.2478	2.0152	1677.7820	2052.8180
6	HPF ⁰ ₁	0.1530	1.3411	1131.7320	1294.3660
7	HPF ¹ ₁	0.1533	1.3435	1133.9060	1297.5640
8	HPF ² ₁	0.1605	1.2661	1055.2190	1329.2900
9	HPF ³ ₁	0.1531	1.3416	1132.2280	1295.0920
10	HPF ⁴ ₁	0.1529	1.3403	1131.0830	1293.4200
11	HPF ⁰ ₂	0.4966	4.7406	3969.4040	4141.8950
12	HPF ¹ ₂	0.4958	4.7313	3961.5770	4134.3950
13	HPF ² ₂	0.6269	6.0870	5101.5360	5236.8690
14	HPF ³ ₂	0.4964	4.7384	3967.6180	4140.1840
15	HPF ⁴ ₂	0.4969	4.7434	3971.7400	4144.1340
16	HPF ⁰ ₃	0.2952	2.6241	2222.9010	2508.9300
17	HPF ¹ ₃	0.2960	2.6329	2230.2930	2515.8350
18	HPF ² ₃	0.1884	1.6163	1369.1300	1602.0430
19	HPF ³ ₃	0.2954	2.6261	2224.5880	2510.5050
20	HPF ⁴ ₃	0.2950	2.6215	2220.6950	2506.8700
21	HPF ⁰ ₄	0.2423	1.9511	1623.9360	2006.8410
22	HPF ¹ ₄	0.2416	1.9423	1616.5440	2000.5240
23	HPF ² ₄	0.3603	3.2943	2753.3300	2996.6570
24	HPF ³ ₄	0.2421	1.9491	1622.2500	2005.3990
25	HPF ⁴ ₄	0.2425	1.9537	1626.1420	2008.7290

Note: Hodrick–Prescott Filter (RSF). In the forecasting models, we assign a code to each model: “0” for the ANN, “1” for the SVM1, “2” for the SVM2, “3” for the RF, and “4” for the DT. For example, HPF¹₁ represents the estimate of the long-term trend (L_t) with the ANN and the residual series (R_t) estimated using the SVM1.

On the other hand, from the proposed filters (RSF and SSF) and the benchmark filter (HPF), one best combination model from each filter was selected and compared. The mean of the accuracy measures numerically are listed in Table 5 and graphically presented in Figure 4. From both presentations, it was confirmed that the RSF²₁ produced the lowest values (RMSPE = 0.1452, MAPE = 1.2083, MAE = 1010.7360, and RMSE = 1207.9950) among the best combination models. Finally, it was concluded that the proposed filter (RSF) resulted in more accurate forecasts than the other proposed filter and a benchmark filter. On the other hand, the RSF²₁ was confirmed as the best model.

Once the accuracy measures were calculated, the next step was to evaluate the dominance of these results. For this purpose, many researchers in the literature have performed the Diebold and Mariano test (DM) [40–43]. In this work, to verify the superiority of the proposed filtering and combination forecasting system results (accuracy measurements) presented in Tables 2–4, we performed tests by Diebold and Mariano (DM) on each pair of models [44]. The DM test results (*p*-values) are shown in Table 6. In contrast to the alternative that each entry in the table is *p* and that the column/row predictor accuracies are more accurate than the column/row predictor values of the hypothesis system, the null hypothesis was that there was no predictor. This table shows that among all the combination models, in Tables 3–5, the RSF²₁, SSF²₁, and SSF²₁ models were statistically superior to the others at the 5% significance level. In the same way, from the proposed filters (RSF and SSF) and the benchmark filter (HPF), one best model from each combination of filters was selected, and their accuracy error is listed in Table 5. Moreover, to verify the superiority of the best-selected models, the DM test results (*p*-values) are shown in Table 6; it was confirmed that the RSF²₁ was statistically superior to the others at the 5% significance level.

Table 5. World Monkeypox Virus Data: the best final combination models results of the out-of-sample cumulative confirmed day-ahead mean forecast error.

S.No	Models	RMSPE	MAPE	MAE	RMSE
1	RSF ₁ ²	0.1452	1.2083	1010.7360	1207.9950
2	SSF ₁ ²	0.1455	1.2162	1020.2140	1222.9890
3	HPF ₁ ²	0.1605	1.2661	1055.2190	1329.2900

Table 6. World Monkeypox Virus Data: results (*p*-value) of the DM test for the null hypothesis that the two models on the rows and columns are equally accurate and the alternative hypothesis that the model on the columns is more accurate than the model on the rows (using the loss square function).

Models	RSF ₁ ²	SSF ₁ ²	HPF ₁ ²
RSF ₁ ²	-	0.18	1.00
SSF ₁ ²	0.82	-	0.99
HPF ₁ ²	0.00	0.01	-

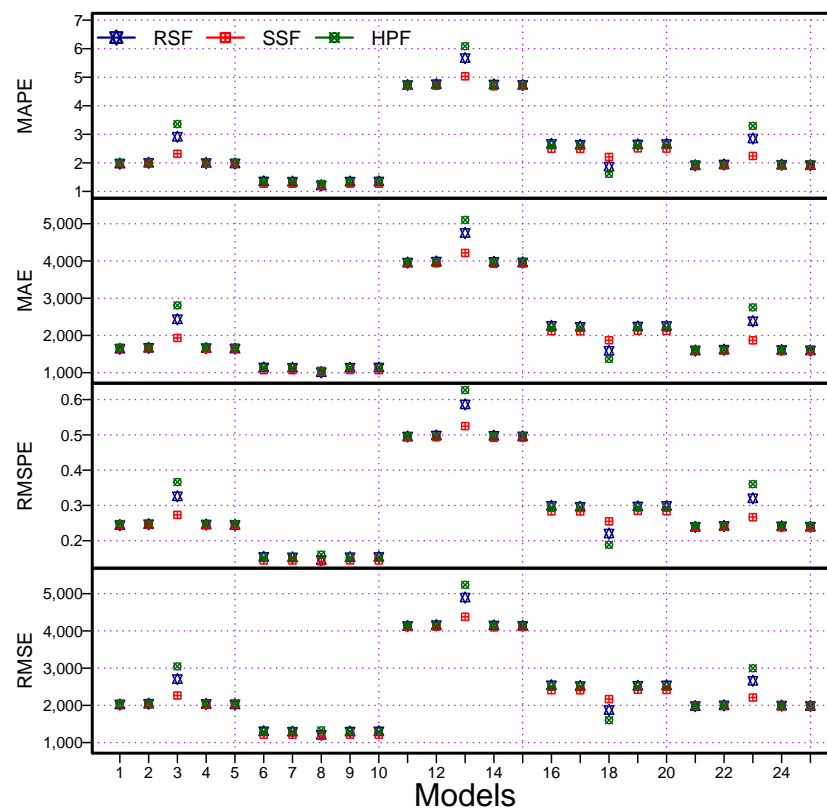


Figure 4. World Monkeypox Virus Data: accuracy measurement plots: MAPE (1st), MAE (2nd), RMAPE (3rd), and RMSE (4th), for all combination models using two proposed filters and a benchmark filter.

Finally, the graphical representations of the performance measures for all 75 models are also shown in Figure 4, for the MAPE (1st), MAE (2nd), RMSPE (3rd), and RMSE (4th). In these plots, we can see that the proposed filters produced the highest accuracy (MAPE, MAE, RMSPE, and RMSE) when compared with the considered benchmark filter (HPF). However, within the proposed filters, the RSF obtained the highest accuracy. Moreover, within all the possible combination models (75), the RSF₁² was declared as the best model. Therefore, from the descriptive statistics, statistical tests, and graphical results, we can conclude that the proposed forecasting methodology is highly accurate and efficient for the daily cumulative confirmed cases of MV forecasting. Additionally, the proposed filters

had high accuracy and resulted in efficient forecasts when compared with the considered benchmark filter. Within the set of proposed filters, the RS filter produced a more precise forecast when compared with the alternatives. Moreover, it was confirmed from these visualizations that the RSF_1^2 was statistically superior to the others (Figure 5).

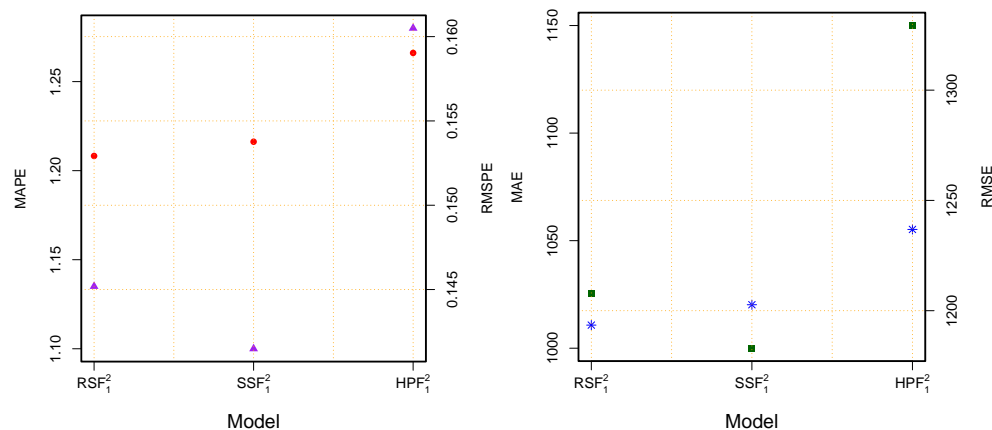


Figure 5. World Monkeypox Virus Data: accuracy measurement plots for the best three models, (left) (MAPE-red circle and RMSPE-purple triangle) and (right) (MAE-blue star and RMSE-green square) for the best three final models using the two proposed filtering methods and a benchmark filter.

4. Discussion

According to the results (the descriptive statistics, statistical test, and graphical analysis), the conclusion is that the final best combination model was the RSF_1^2 , which was highly accurate and efficient for the daily cumulative confirmed cases of MV. Hence, it is important to note that reported mean accuracy errors (RMSPE, MAPE, MAE, and RMSE) in this work were, for the best combination model (RSF_1^2), relatively lower than the considered times series and machine learning benchmark models. The considered benchmark models were the following: three standard time series models: the autoregressive (AR), nonparametric autoregressive (NPAR), and autoregressive moving integrated average (ARIMA) models, and four popular machine learning models: ANN, SVM, DT, and RF. An empirical comparison of the proposed work's best model with the other considered benchmarks (AR, ARIMA, NPAR, ANN, SVM, DT, and RF) models are presented numerically in Table 7. As we can see from Table 7, the proposed best combination model (RSF_1^2) in this work obtained significantly lower accuracy mean errors as compared to the time series and machine learning models. Additionally, to confirm the superiority of the proposed best combination model mentioned in Table 8, we performed a statistical test using the DM on each pair of models. The results (p -values) of the DM test are reported in Table 8, showing that the proposed models among all the considered time series and machine learning models were outperformed by our best model at the 5% significance level. To conclude, based on all of these results, the accuracy of the proposed forecasting methodology is comparatively high and efficient when compared with all the considered competitors.

Once the best models were assessed through descriptive statistics, statistical tests, and graphical analysis, we proceeded to future forecasting with the superior model. We used the RSF_1^2 for the confirmed cases of MV, and forecast from 15 February to 26 February 2023 (two weeks) for the cumulative confirmed cases of MV. The forecasted values of the cumulative and daily confirmed cases from the MV are tabulated in Table 9. This table reveals that the daily confirmed cases increased in the first two days but decreased in the following three consecutive days, while on the rest of the days, there was an average increase of 45 per day. Furthermore, we predicted that during the next fourteen days (15 February to 26 February 2023), a total of 642 new cases of MV would be added globally. However, these forecasts continued to support an overall declining trend in the number of newly infected cases of MV around the world. Finally, to justify the superiority of the

proposed final best model forecasting performance, we compared the cumulative infected cases of MV with the forecasted case by the proposed best model. To achieve this, we computed the percentage forecast error (PFE); the PFE is defined as $PFE = (| \text{forecasted value} - \text{actual value} | / | \text{actual value} |) \times 100$. The values of the PFE are listed in the last column of Table 9. From this column, one can see that the forecasted values were relatively close to the actual values in terms of a low PFE. Therefore, our results offer valuable information to policymakers to guide the continued allocation of resources and inform mitigation efforts. Moreover, the forecasting exercise will help to understand the spread and lead to an understanding of the risk, which may be used to prevent further spread and enable timely and effective treatment.

Table 7. World Monkeypox Virus Data: comparison of the proposed best model versus the considered time series and machine learning benchmark models: out-of-sample cumulative confirmed day-ahead mean forecast error.

S.No	Models	RMSPE	MAPE	MAE	RMSE
1	RSF ₁ ²	0.1452	1.2083	1010.7360	1207.9950
2	ANN	0.2470	2.0055	1669.6410	2045.9800
3	SVM	0.4868	4.6378	3882.9750	4059.1410
4	RF	0.2826	2.4879	2108.1850	2402.4950
5	DT	0.2400	1.9241	1601.2560	1987.3760
6	AR	0.2556	2.1090	1756.6180	2117.7440
7	NPAR	0.2482	2.0205	1682.1830	2056.4220
8	ARIMA	0.2502	2.0441	1702.1040	2072.7490

Table 8. World Monkeypox Virus Data: the best final model and the considered time series and machine learning benchmark models: results (*p*-value) of the DM test for the null hypothesis that the two models on the rows and columns are equally accurate and the alternative hypothesis that the model on the columns is more accurate than the model on the rows (using the loss square function).

Models	RSF ₁ ²	ANN	SVM	RF	DT	AR	NPAR	ARMA
RSF ₁ ²	-	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ANN	0.00	-	1.00	0.93	0.00	1.00	1.00	1.00
SVM	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00
RF	0.00	0.07	1.00	-	0.04	0.12	0.07	0.09
DT	0.00	1.00	1.00	0.96	-	1.00	1.00	1.00
AR	0.00	0.00	1.00	0.88	0.00	-	0.00	0.00
NPAR	0.00	0.00	1.00	0.93	0.00	1.00	-	1.00
ARMA	0.00	0.00	1.00	0.91	0.00	1.00	0.00	-

Table 9. World Monkeypox Virus Data: the forecasted cumulative and daily confirmed cases of the monkeypox virus using the best-proposed model over two weeks.

S.No	Date	CC	DC	PFE
1	11 February 2023	84063	77	2.129
2	12 February 2023	84,140	77	2.051
3	13 February 2023	84,189	49	1.994
4	14 February 2023	84,225	36	1.976
5	15 February 2023	84,264	39	1.956
6	16 February 2023	84,309	45	1.908
7	17 February 2023	84,356	47	1.857
8	18 February 2023	84,402	46	1.920
9	19 February 2023	84,448	46	1.866
11	20 February 2023	84,492	44	1.815
12	21 February 2023	84,537	45	1.769
13	22 February 2023	84,583	46	1.748
14	23 February 2023	84,628	45	1.777

Note: cumulative confirmed cases (CC), daily confirmed cases (DC), and percentage forecast error (PFE).

On the other hand, actual-time forecasting in epidemic emergencies provides actionable information that governments can use to anticipate medical needs and strategize the intensity and composition of public health interventions. This study provides and evaluates accurate real-time short-term forecasts of monkeypox cases worldwide, with much higher case numbers reported in most epidemics. Our model continues to predict a slowdown in MV incidence globally. Overall, our model has proven useful in making short-term forecasts and capturing slowdowns and peaks in growth with reasonable accuracy.

5. Conclusions

The main aim of this work was to forecast the short-term transmission rate of the monkeypox virus. For this purpose, we proposed a novel filtering and combination forecasting approach to accurately forecast the daily cumulative confirmed monkeypox virus cases using the publicly updated monkeypox virus dataset. To this end, we first filtered the original time series of the cumulative confirmed cases into new subseries, that is, the long-term trend series and residual series, using two proposed and one benchmark filter. Then, to predict the decomposed subseries, we considered five well-known machine learning models and all their possible combinations. Hence, we combined the individual predictive models directly to obtain the final predictions for newly confirmed cases one day ahead. To verify the performance of the proposed filtering and combination methodology, four mean errors—two absolute errors, two relative errors, and a statistical test—were considered. The experimental results showed the efficiency and accuracy of the proposed filtering and combination forecasting system. Overall, this work was compared within the proposed forecasting technique with three filters: two proposed and a benchmark filter. Among these three filtering methods, the regression smoothing spline filter outperformed the rest. On the other hand, within all the possible combination models, the (RSF_1^2) , (SSF_1^2) , and (HPF_1^2) were the first-, second-, and third-best models. In addition, the proposed final combination model (RSF_1^2) had superior results (accuracy mean errors, graphical analysis, and statistical test) when all the time series and machine learning models were considered. Finally, based on the best-selected model (RSF_1^2) , we provided a forecast of the next fourteen days (15 February to 26 February 2023), which will help to understand the spread and lead to understanding the risk, which may be used to prevent further spread and enable timely and effective treatment.

As this study used only cumulatively confirmed monkeypox data, it could be extended to other variables (such as daily new infected cases and daily and cumulative death counts) to evaluate the performance of the proposed forecasting technique. Further, it could be used for the short-term forecasting of daily and cumulative COVID-19 confirmed cases, death counts, and recovered cases. On the other hand, the proposed forecasting techniques used only machine learning models; in the future, they will be extended by time series models, such as parametric autoregressive, nonparametric autoregressive, autoregressive integrated moving average models, etc.

Author Contributions: Conceptualization, methodology, and software, H.I.; validation, H.I., M.K. (Murad Khan) and M.K. (Mehak Khan); formal analysis, H.I.; investigation, H.I., M.K. (Murad Khan) and M.S.K.; resources, M.S.K. and M.K. (Mehak Khan); data curation, H.I. and M.K. (Murad Khan); writing—original draft preparation, H.I., M.K. (Murad Khan) and M.K. (Mehak Khan); writing—review and editing, H.I. and M.K.; visualization, H.I., M.K. (Murad Khan) and M.K. (Mehak Khan); supervision, H.I. and M.K. (Mehak Khan); project administration, H.I., M.S.K. and M.K. (Mehak Khan); funding acquisition, M.K. (Mehak Khan) and M.S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be provided upon the request of the first author.

Acknowledgments: This research work was funded by the Department of Computer Science, AI Lab, Oslo Metropolitan University, P.O. Box 4St. Olavs plass, Oslo, 0130, Norway.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- World Health Organization. COVID-19. Available online: <https://covid19.who.int/> (accessed on 17 February 2023).
- Taha, M.J.; Abuawwad, M.T.; Alrubasy, W.A.; Sameer, S.K.; Alsafi, T.; Al-Bustanji, Y.; Abu-Ismael, L.; Nashwan, A.J. Ocular manifestations of recent viral pandemics: A literature review. *Front. Med.* **2022**, *9*, 101133. [[CrossRef](#)] [[PubMed](#)]
- Chowdhury, S.R.; Datta, P.K.; Maitra, S. Monkeypox and its pandemic potential: What the anaesthetist should know. *Br. J. Anaesth.* **2022**, *129*, e49–e52. [[CrossRef](#)] [[PubMed](#)]
- Cho, C.T.; Wenner, H.A. Monkeypox virus. *Bacteriol. Rev.* **1973**, *37*, 1–18. [[CrossRef](#)] [[PubMed](#)]
- Marennikova, S.S.; Šeluhina, E.M.; Mal’Ceva, N.N.; Čimiškjan, K.L.; Macevič, G.R. Isolation and properties of the causal agent of a new variola-like disease (monkeypox) in man. *Bull. World Health Organ.* **1972**, *46*, 599. [[PubMed](#)]
- Bunge, E.M.; Hoet, B.; Chen, L.; Lienert, F.; Weidenthaler, H.; Baer, L.R.; Steffen, R. The changing epidemiology of human monkeypox—A potential threat? A systematic review. *PLoS Negl. Trop. Dis.* **2022**, *16*, e0010141. [[CrossRef](#)]
- Meyer, H.; Perrichot, M.; Stemmler, M.; Emmerich, P.; Schmitz, H.; Varaine, F.; Shungu, R.; Tshioko, F.; Formenty, P. Outbreaks of disease suspected of being due to human monkeypox virus infection in the Democratic Republic of Congo in 2001. *J. Clin. Microbiol.* **2002**, *40*, 2919–2921. [[CrossRef](#)]
- Vivancos, R.; Anderson, C.; Blomquist, P.; Balasegaram, S.; Bell, A.; Bishop, L.; Brown, C.S.; Chow, Y.; Edeghere, O.; Florence, I.; et al. Community transmission of monkeypox in the United Kingdom, April to May 2022. *Eurosurveillance* **2022**, *27*, 2200422. [[CrossRef](#)]
- Mathieu, E.; Spooner, F.; Dattani, S.; Ritchie, H.; Roser, M. “Mpox (monkeypox)”. Published Online at OurWorldInData.org. 2022. Available online: <https://ourworldindata.org/monkeypox> (accessed on 10 February 2023).
- Mao, L.; Ying, J.; Selekon, B.; Gonofio, E.; Wang, X.; Nakoune, E.; Wong, G.; Berthet, N. Development and Characterization of Recombinase-Based Isothermal Amplification Assays (RPA/RAA) for the Rapid Detection of Monkeypox Virus. *Viruses* **2022**, *14*, 2112. [[CrossRef](#)]
- World Health Organization. Monkeypox. Available online: <https://www.who.int/publications/m/item/multi-country-outbreak-of-mpox--external-situation-report--22---11-may-2023/> (accessed on 11 May 2023).
- Yang, Z. Monkeypox: A potential global threat? *J. Med. Virol.* **2022**, *94*, 4034–4036. [[CrossRef](#)]
- Iftikhar, H.; Rind, M. Forecasting daily COVID-19 confirmed, deaths and recovered cases using univariate time series models: A case of Pakistan study. *MedRxiv* **2020**. [[CrossRef](#)]
- Bantan, R.A.; Ahmad, Z.; Khan, F.; Elgarhy, M.; Almaspoor, Z.; Hamedani, G.G.; El-Morshedy, M.; Gemeay, A.M. Predictive modeling of the COVID-19 data using a new version of the flexible Weibull model and machine learning techniques. *Math. Biosci. Eng.* **2023**, *20*, 2847–2873. [[CrossRef](#)] [[PubMed](#)]
- Ahmad, Z.; Almaspoor, Z.; Khan, F.; El-Morshedy, M. On predictive modeling using a new flexible Weibull distribution and machine learning approach: Analyzing the COVID-19 data. *Mathematics* **2022**, *10*, 1792. [[CrossRef](#)]
- Taimoor, M.; Ali, S.; Shah, I.; Muwanika, F. R. COVID-19 pandemic data modeling in Pakistan using time-series SIR. *Comput. Math. Methods Med.* **2022**, *2022*, 6001876. [[CrossRef](#)] [[PubMed](#)]
- Abbasimehr, H.; Paki, R.; Bahrini, A. A novel approach based on combining deep learning models with statistical methods for COVID-19 time series forecasting. *Neural Comput. Appl.* **2022**, *34*, 3135–3149. [[CrossRef](#)]
- Alshanbari, H.M.; Iftikhar, H.; Khan, F.; Rind, M.; Ahmad, Z.; El-Bagoury, A.A.A.H. On the Implementation of the Artificial Neural Network Approach for Forecasting Different Healthcare Events. *Diagnostics* **2023**, *13*, 1310. [diagnostics13071310](#) [[CrossRef](#)]
- Iftikhar, H.; Khan, M.; Khan, Z.; Khan, F.; Alshanbari, H.M.; Ahmad, Z. A Comparative Analysis of Machine Learning Models: A Case Study in Predicting Chronic Kidney Disease. *Sustainability* **2023**, *15*, 2754. [[CrossRef](#)]
- Khan, M.; Wang, H.; Riaz, A.; Elfatyany, A.; Karim, S. Bidirectional LSTM-RNN-based hybrid deep learning frameworks for univariate time series classification. *J. Supercomput.* **2021**, *77*, 7021–7045. [[CrossRef](#)]
- Khan, M.; Wang, H.; Nguetilbaye, A.; Elfatyany, A. End-to-end multivariate time series classification via hybrid deep learning architectures. *Pers. Ubiquitous Comput.* **2023**, *27*, 177–191. [[CrossRef](#)]
- Khan, M.; Wang, H.; Nguetilbaye, A. Attention-based deep gated fully convolutional end-to-end architectures for time series classification. *Neural Process. Lett.* **2021**, *53*, 1995–2028. [[CrossRef](#)]
- Patel, M.; Surti, M.; Adnan, M. Artificial intelligence (AI) in Monkeypox infection prevention. *J. Biomol. Struct. Dyn.* **2022**, 1–5. [[CrossRef](#)]
- Qureshi, M.; Khan, S.; Bantan, R.A.; Daniyal, M.; Elgarhy, M.; Marzo, R.R.; Lin, Y. Modeling and Forecasting Monkeypox Cases Using Stochastic Models. *J. Clin. Med.* **2022**, *11*, 6555. [[CrossRef](#)]
- Dada, E.G.; Oyewola, D.O.; Joseph, S.B.; Emebo, O.; Oluwagbemi, O.O. Ensemble Machine Learning for Monkeypox Transmission Time Series Forecasting. *Appl. Sci.* **2022**, *12*, 12128. [[CrossRef](#)]
- Long, B.; Tan, F.; Newman, M. Forecasting the Monkeypox Outbreak Using ARIMA, Prophet, NeuralProphet, and LSTM Models in the United States. *Forecasting* **2023**, *5*, 127–137. [[CrossRef](#)]

27. Eid, M.M.; El-Kenawy, E.S.M.; Khodadadi, N.; Mirjalili, S.; Khodadadi, E.; Abotaleb, M.; Alharbi, A.H.; Abdelhamid, A.A.; Ibrahim, A.; Amer, G.M.; et al. Meta-heuristic optimization of LSTM-based deep network for boosting the prediction of monkeypox cases. *Mathematics* **2022**, *10*, 3845. [[CrossRef](#)]
28. Priyadarshini, I.; Mohanty, P.; Kumar, R.; Taniar, D. Monkeypox Outbreak Analysis: An Extensive Study Using Machine Learning Models and Time Series Analysis. *Computers* **2023**, *12*, 36. [[CrossRef](#)]
29. Khan, M.I.; Qureshi, H.; Bae, S.J.; Awan, U.A.; Saadia, Z.; Khattak, A.A. Predicting Monkeypox incidence: Fear is not over! *J. Infect.* **2022**, *86*, 256–308. [[CrossRef](#)]
30. Mohbey, K.K.; Meena, G.; Kumar, S.; Lokesh, K. A CNN-LSTM-based hybrid deep learning approach to detect sentiment polarities on Monkeypox tweets. *arXiv* **2022**, arXiv:2208.12019.
31. Bleichrodt, A.; Dahal, S.; Maloney, K.; Casanova, L.; Luo, R.; Chowell, G. Real-time forecasting the trajectory of monkeypox outbreaks at the national and global levels, July–October 2022. *BMC Med.* **2023**, *21*, 19 [[CrossRef](#)]
32. Yan, W.; Du, M.; Qin, C.; Liu, Q.; Wang, Y.; Liang, W.; Liu, M.; Liu, J. Association between public attention and monkeypox epidemic: A global lag-correlation analysis. *J. Med. Virol.* **2023**, *95*, e28382. [[CrossRef](#)]
33. Islam, M.P.; Morimoto, T. Non-linear autoregressive neural network approach for inside air temperature prediction of a pillar cooler. *Int. J. Green Energy* **2017**, *14*, 141–149. [[CrossRef](#)]
34. Zendejboudi, A.; Baseer, M.A.; Saidur, R. Application of support vector machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.* **2018**, *199*, 272–285. [[CrossRef](#)]
35. Breiman, L. *Classification and Regression Trees*; Routledge: London, UK, 2017.
36. Bibi, N.; Shah, I.; Alsubie, A.; Ali, S.; Lone, S.A. Electricity spot prices forecasting based on ensemble learning. *IEEE Access* **2021**, *9*, 150984–150992. [[CrossRef](#)]
37. Liu, C.; Hu, Z.; Li, Y.; Liu, S. Forecasting copper prices by decision tree learning. *Resour. Policy* **2017**, *52*, 427–434. [[CrossRef](#)]
38. Kumar, R. Decision tree for the weather forecasting. *Int. J. Comput. Appl.* **2013**, *76*, 31–34. [[CrossRef](#)]
39. Zhang, G.P.; Qi, M. Neural network forecasting for seasonal and trend time series. *European J. Oper. Res.* **2005**, *160*, 501–514. [[CrossRef](#)]
40. Shah, I.; Iftikhar, H.; Ali, S.; Wang, D. Short-term electricity demand forecasting using components estimation technique. *Energies* **2019**, *12*, 2532. [[CrossRef](#)]
41. Shah, I.; Iftikhar, H.; Ali, S. Modeling and forecasting medium-term electricity consumption using component estimation technique. *Forecasting* **2020**, *2*, 163–179. [[CrossRef](#)]
42. Shah, I.; Iftikhar, H.; Ali, S. Modeling and forecasting electricity demand and prices: A comparison of alternative approaches. *J. Math.* **2022**, *2022*, 3581037. [[CrossRef](#)]
43. Iftikhar, H.; Bibi, N.; Canas Rodrigues, P.; López-Gonzales, J.L. Multiple Novel Decomposition Techniques for Time Series Forecasting: Application to Monthly Forecasting of Electricity Consumption in Pakistan. *Energies* **2023**, *16*, 2579. [[CrossRef](#)]
44. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **2002**, *20*, 134–144. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.