Article

Time-Series Analysis for the Number of Foot and Mouth Disease Outbreak Episodes in Cattle Farms in Thailand Using Data from 2010–2020

Veerasak Punyapornwithaya [1,2], Pradeep Mishra [3], Chalutwan Sansamur [4,5], Dirk Pfeiffer [6,7], Orapun Arjkumpa [8], Rotchana Prakotcheo [9], Thanis Damrongwatanapokin [4,5,*] and Katechan Jampachaisri [10,*]

1   Center of Excellence in Veterinary Public Health, Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai 50100, Thailand; veerasak.p@cmu.ac.th
2   Excellence Center in Veterinary Bioscience, Chiang Mai University, Chiang Mai 50100, Thailand
3   College of Agriculture, Jawaharlal Nehru Krishi Vishwavidyalaya, Powarkheda, Narmadapuram 461110, India; pradeepjnkvv@gmail.com
4   Akkhraratchakumari Veterinary College, Walailak University, Nakorn Si Thammarat 80160, Thailand; chakream@gmail.com
5   Centre for One Health, Walailak University, Nakhon Si Thammarat 80161, Thailand
6   Centre for Applied One Health Research and Policy Advice, City University of Hong Kong, Hong Kong SAR, China; pfeiffer@cityu.edu.hk
7   Department of Pathobiology and Population Sciences, Royal Veterinary College, London, NW1 0TU, UK
8   The 4th Regional Livestock Office, Mueang Khon Kaen, Khon Kaen 40260, Thailand; arjkumpa@hotmail.com
9   Bureau of Disease Control and Veterinary Services, Department of Livestock Development, Bangkok 10400, Thailand; dr_sam339@hotmail.com
10 Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand
*   Correspondence: thanis.da@wu.ac.th (T.D.); katechanj@nu.ac.th (K.J.)

Abstract: Thailand is one of the countries where foot and mouth disease outbreaks have resulted in considerable economic losses. Forecasting is an important warning technique that can allow authorities to establish an FMD surveillance and control program. This study aimed to model and forecast the monthly number of FMD outbreak episodes (n-FMD episodes) in Thailand using the time-series methods, including seasonal autoregressive integrated moving average (SARIMA), error trend seasonality (ETS), neural network autoregression (NNAR), and Trigonometric Exponential smoothing state–space model with Box–Cox transformation, ARMA errors, Trend and Seasonal components (TBATS), and hybrid methods. These methods were applied to monthly n-FMD episodes (n = 1209) from January 2010 to December 2020. Results showed that the n-FMD episodes had a stable trend from 2010 to 2020, but they appeared to increase from 2014 to 2020. The outbreak episodes followed a seasonal pattern, with a predominant peak occurring from September to November annually. The single-technique methods yielded the best-fitting time-series models, including $SARIMA(1,0,1)(0,1,1)_{12}$, $NNAR(3,1,2)_{12}$, $ETS(A,N,A)$, and $TBATS(1,\{0,0\},0.8,\{<12,5>\})$. Moreover, SARIMA-NNAR and NNAR-TBATS were the hybrid models that performed the best on the validation datasets. The models that incorporate seasonality and a non-linear trend performed better than others. The forecasts highlighted the rising trend of n-FMD episodes in Thailand, which shares borders with several FMD endemic countries in which cross-border trading of cattle is found common. Thus, control strategies and effective measures to prevent FMD outbreaks should be strengthened not only in Thailand but also in neighboring countries.

Keywords: foot and mouth disease; outbreak; time-series model; forecast; trend; seasonality; endemic settings

1. Introduction

The World Organization for Animal Health (OIE) lists foot and mouth disease (FMD) as an important transboundary disease [1]. The disease is caused by the FMD virus (FMDV

belonging to the genus *Aphthovirus* of the *Picornaviridae* family), which affects cloven-hoof animals [2]. FMD results in economic losses due to decreased production of cattle in endemic regions and trade restrictions on disease-affected countries imposed by disease-free countries [3,4]. FMD outbreaks have been reported on several continents, including Europe, Africa, South America, and Asia [2,5–11]. FMD is endemic in several countries in Asia. The main serotypes of FMDV circulating in this region are O, A, and Asia 1 [12,13]. Numerous countries in Southeast Asia, including Cambodia, Lao PDR, Malaysia, Myanmar, Thailand, and Vietnam, are endemic to FMD [9,13,14]. Between 2007 and 2017, a total of 4961 FMD outbreaks were reported in these countries [15]. In Thailand, FMD has persisted for more than 60 years and spread periodically and temporarily in several regions of the country [16–21]. More than 1200 FMD outbreaks in cattle farms across several regions were officially reported from 2008 to 2019 [16]. In Southeast Asia, Thailand plays an important part in the regional FMD epidemiology due to the movements of livestock by trading with several countries, including Laos PDR, Cambodia, Vietnam, and China [22].

Time-series modeling has demonstrated its usefulness for event prediction [23–25]. At present, a wide range of time-series methods has been used in several epidemiological studies, resulting in predictive models that can be applied to a variety of data [25–28]. The time-series models were considered in this study for their capability to manage seasonality time-series data due to seasonal FMD outbreaks in Thailand [16]. Accordingly, (i) seasonal autoregressive integrated moving average (SARIMA) [29], (ii) error trend and seasonality (ETS) [30], and (iii) trigonometric exponential smoothing state–space model with Box–Cox transformation and an autoregressive moving average error, trend, and seasonality (TBATS) [30,31] methods were used together in the present study. Moreover, the application of neural network nonlinear autoregression (NNAR), a widely used machine learning method for nonlinear time-series data, was explored because FMD outbreak data may contain nonlinear patterns.

The SARIMA modeling approach is widely used in various disciplines and is regarded as one of the most efficient methods of modeling time-series data with a seasonal trend [29]. SARIMA has the advantage of requiring few model parameters for modeling and prediction. Similar to SARIMA, the ETS can deal with seasonality data. When dealing with data that exhibits heterogeneity and non-linearity, the ETS offers an advantage over SARIMA. The artificial neural network is a machine learning algorithm based on the idea of a neuron from a human brain [32–34]. As a result, the computer machine is able to learn and make decisions in a human-like fashion [35]. Nonlinear autoregression (NNAR) is a type of artificial neural network that uses lagged values of a time series as input to a neural network. This method incorporates the most recent observations from the same season when used with seasonal data [29]. The main difference between NNAR and SARIMA models is that NNAR can manage nonlinear sequences of time-series data better than ARIMA [33], whereas SARIMA has a better capability to take account of a linear pattern in data than NNAR. Furthermore, it is possible to model time series with different seasonality using TBATS [30,31]. With this method, a Fourier series-based trigonometric representation is used to model seasonality. The advantage of this method is the capacity to model seasonal effects of non-integer lengths [36]. Aside from simple time-series models, hybrid models that integrate forecasting methods are increasingly being used in a variety of disciplines because they typically outperform single methods [28,34,37]. In the current study, the predictive abilities of hybrid models based on the combination of SARIMA, NNAR, ETS, and TBATS, such as SARIMA-NNAR, ETS-TBATS, and NNAR-TBATS, were assessed. Supplementary Materials Table S1 provides a brief overview of the models used in this study, including their applications, assumptions, and limitations.

In epidemiology, the development of a reliable method for the explanation and prediction of FMD outbreaks is of utmost importance. However, only a few studies have used time-series methods to explore and forecast FMD outbreaks [7,38–40]. A lack of outbreak prediction may pose some challenges in the formulation of an effective FMD surveillance and control program in Thailand. Furthermore, a better understanding of

FMD outbreaks in one country can support other countries in developing an effective control program to prevent FMD from spreading across borders. Thus, given Thailand's proximity to other FMD-endemic countries, the findings of this study, which are based on 10 years of retrospective data, will be useful to livestock authorities and stakeholders in those boundary countries.

The objectives of this study were to examine FMD epidemic episodes (n-FMD episodes) in Thailand over a 10-year period for the presence of trend and seasonality using time-series methods and to compare the prediction performances of time-series forecasting models for prospective estimation of FMD outbreak occurrences.

2. Materials and Methods

2.1. Outbreak Episode Definition, Data, and Data Analysis Steps

An outbreak episode of FMD was defined as an official report of an FMD outbreak in an outbreak area where disease investigation and FMD disease confirmation were undertaken by an authority from the Department of Livestock and Development (DLD) [19]. Notably, the episode of FMD outbreak defined as the unit of analysis in this study is similar to the unit utilized in several previous studies [7,9,41].

Data of n-FMD episodes in Thailand from 2010 to 2020 (n = 1209 episodes) were obtained from the DLD. Time-series data were generated from the original data. Further, the data was divided into two parts: data from 2010 to 2019 (Data: 2010–2019; training data) and data from 2020 (Data: 2020; validation data), with the former being used for model development and the latter for model validation, respectively.

Data analysis and time-series modeling were performed using R statistical software [42] and "xts", "tsbox", "plotly", "TSstudio", "forecast", and "forecast Hybrid" packages. The step of time-series data analysis and modeling includes (i) decomposition of time-series data into several components, (ii) determination of the final model through the model development process, (iii) evaluation of the performances of the final model developed, and (iv) forecasting the n-FMD episodes from the final model. These steps are depicted in Figure 1.

2.2. Decomposition of Time-Series Data

An additive decomposition of the FMD time series was carried out to describe the trends and seasonality components using R statistical software version 4.0.4 [42]. At this step, time-series data were decomposed into trend, seasonal, and residual components. The model is given as follows:

$$Y_t = T_t + S_t + I_t \tag{1}$$

where Y_t is the n-FMD episode at time t, T_t is the trend-cycle component at time t, S_t is the seasonal component at time t, and I_t is the irregular component at time t.

2.3. Model Development

2.3.1. SARIMA Model

The SARIMA is composed of seasonal and non-seasonal components. The general SARIMA model [29] has the following form:

$$\varphi(X)\Phi\left(X^{12}\right)\Delta^d\Delta_{12}^D Z_t = \theta(X)\Theta(X)a_t \tag{2}$$

where $\varphi(X)$ = non-seasonal autoregressive operator of order p, $\theta(X)$ = non-seasonal moving average operator of order q, $\Phi(X^{12})$ = the autoregressive seasonal operator of order P, $\Theta(X)$ = the moving averages operator of order Q, Δ^d = the operator difference, Δ_{12}^D = operator seasonal difference, and a_t = white noise.
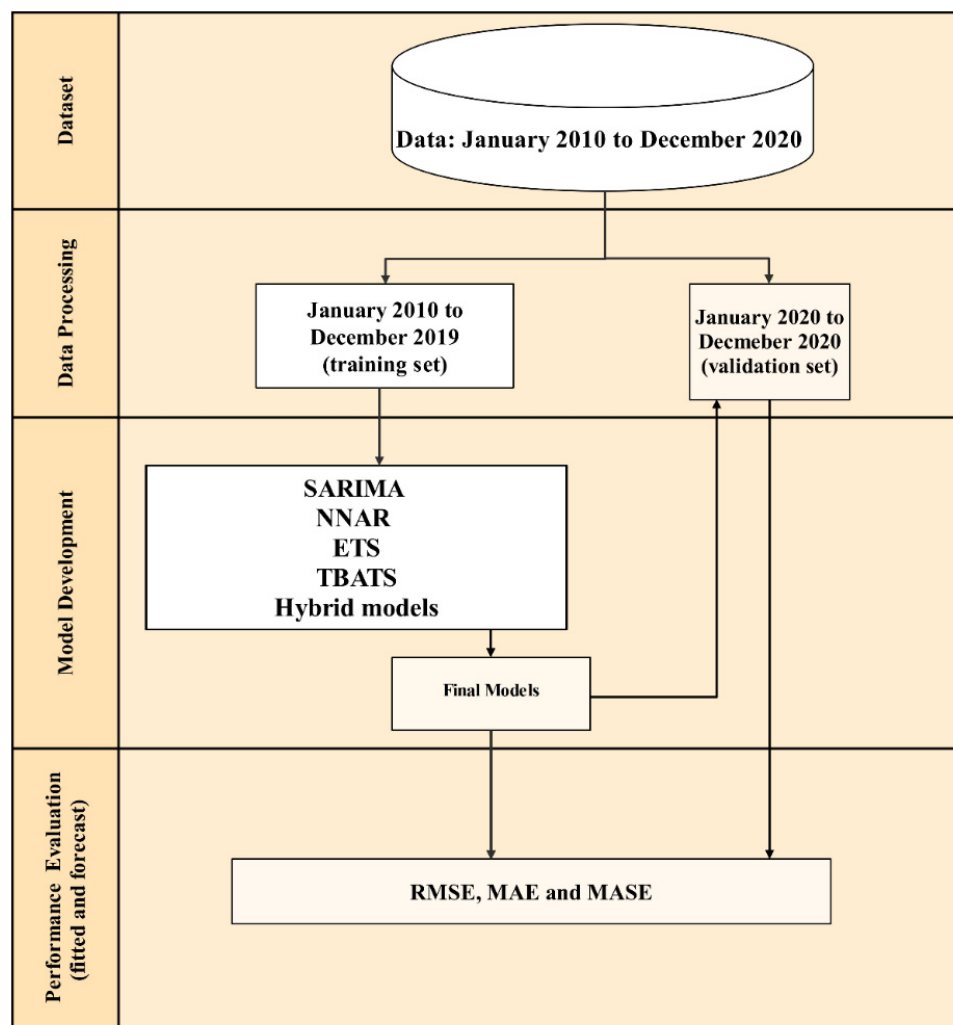
Figure 1. Time-series modeling procedure. A full dataset of the number of FMD outbreak episodes was split into training and validation datasets. Forecast models were developed using seasonal autoregressive integrated moving average (SARIMA), error trend seasonality (ETS), neural network autoregression (NNAR), and Trigonometric Exponential smoothing state–space model with Box–Cox transformation, ARMA errors, Trend and Seasonal components (TBATS), and hybrid methods. With the validation data, error measures, including root mean squared error (RMSE), mean absolute error (MAE), and mean absolute scaled error (MASE), were determined in order to compare the performances of prediction models.

The expression of SARIMA is given as

$$ARIMA(p,d,q)(P,D,Q)_s \tag{3}$$

where p = autoregressive order in non-seasonality, d = difference in non-seasonality, q = the non-seasonal moving average order, P = autoregressive order in seasonality, D = differences in seasonality, Q = moving average order in seasonality, and s = length of the seasonal pattern (s = 12 in this study).

The development of SARIMA models was carried out using *auto.arima*( ) function from a "forecast" package. This function determined the p, d, q, P, D, Q for several candidate models. Akaike information criterion (AIC) was used as criteria to verify the model that best fit the data.

The *auto.arima* function utilized the Hyndman–Khandarkar algorithm for automatic autoregressive integrated moving average (ARMA) modeling. As previously described, this algorithm performed various steps in the model selection procedure [29,43]. Based on

such procedures, the final model was defined as the model with the lowest AIC. Additionally, the model assumptions were checked by examining the standardized residual plot, autocorrelation function plot of residuals, and the *p*-values plot from the Ljung–Box test.

2.3.2. NNAR Model

The NNAR used lagged values of the time series as input to a neural network. A network of three layers of functioning is linked by acyclic linkages. The equation of the NNAR is as follows [44]:

$$y_t = \omega_0 + \sum_{j=1}^{Q} \omega_g g \left(\omega_{oj} + \sum_{i=1}^{p} \omega_{i,j} y_{t-1} \right) + e_t \tag{4}$$

where y_t and $(y_{t-i,...,}y_{t-p})$ are the output and the input, $\omega_{i,j}(i = 0, 1, 2, ..., P, \; j = 1, 2, ..., Q)$ and $\omega_j(j = 0, 1, 2, ..., Q)$ are model parameters, which are known as connection weights; the number of input nodes is represented by P while the number of hidden nodes is indicated by Q.

For seasonal data, the NNAR model [29] can be written as

$$NNAR(p, P, K)_m \tag{5}$$

where p = the last observed values from the same season using as inputs, P = lagged inputs, K = number of neurons (nodes) in the hidden layer, and m = the number of months.

A function *nnetar* from "forecast" package was used to fit $NNAR(p, P, K)_m$ model. For seasonal time series, $P = 1$ was set as default value and p was chosen from the optimal linear model fitted to the seasonally adjusted data. We used the default setting, and thus $K = (p + P + 1)/2$ was set using the *nnetar*( ) function [29].

2.3.3. ETS Model

Regarding a state–space framework, the ETS combines error (E), trend (T), and seasonal (S) components in a smoothing calculation. The ETS offers a total of 30 possible ETS combinations; thus, it has the ability to analyze a variety type of time-series data, even with both heterogeneity and non-linearity. Within a state–space framework, the E component is either additive (A) or multiplicative (M), T and S components may be A, M, or inexistent (N), and T is referred to as dampened additively (A_d) or multiplicatively (M_d) [45]. The state–space equations can be written as [46]

$$y_t = w(x_{t-1}) + r(x_{t-1})\varepsilon_t \tag{6}$$

$$x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t \tag{7}$$

where w, f, and g are coefficient while ε_t represents the Gaussian white noise series. Equation (6) is known as the observation that describes the relationship between the observation x_{t-1} and y_t. Equation (7) is the transitional equation describing the evolution of states over time. The ETS model was performed using the *ets*( ) function from the "forecast" package.

2.3.4. TBATS Model

Multiple seasonal incorrect cycles can be accommodated by TBATS. A combination of Fourier with an exponential smoothing state–space model and a Box–Cox transformation is represented by the TBATS models in this study. The basic equation of the TBATS took the following form [30,47]:

$$y_t^{(\omega)} = l_{t-1} + \phi b_{t-1} + \sum_{t=1}^{T} s_{t-m_i}^{(i)} + d_t \tag{8}$$

where $y_t^{(\omega)}$ indicates the Box–Cox transformation parameter (ω) and y_t is the observation at time t, l_t is the local level, ϕ denotes the damped trend, b denotes the long-run trend, T indicates the seasonal pattern, $s_{t_i}^{(i)}$ denotes the ith seasonal component, m_i indicates the seasonal periods, and d_t is an $ARMA(p,q)$ process for residuals. The TBATS model was identified by $tbats(\ )$ function included in the "forecast" package.

2.3.5. Hybrid Model

Time-series hybrid models including SARIMA-NNAR, SARIMA-ETS, SARIMA-TBATS, NNAR-ETS, NNAR-TBATS and ETS-TBATS were developed using functions from the "forecast Hybrid" package and R. The functions offer the automated procedure to obtain the best fitting model. The detail of model development from the functions was previously described [8,48].

2.4. Forecast and Model Performances

Time-series models developed from Data: 2010–2019 were used to predict n-FMD episodes in 2020. By comparing forecast and actual n-FMD episode values, the forecasting ability was evaluated.

We used several evaluation error metrics, including mean absolute error (MAE), root mean squared error (RMSE), and mean absolute scaled error (MASE) [29,49], to determine prediction performances among models developed from the full and training datasets [29]. It was important to note that a mean absolute percentage error (MAPE), which was generally reported, was unable to be determined as our data contained zero counts for some months of time-series data. Therefore, the scaled errors such as MASE were proposed as an alternative to using MAPE. Similar to MAPE, the MASE is independent of the scale of the data [29]. The MASE is considered the most versatile and reliable measure of forecast accuracy and can be used to compare forecast accuracy both on single and multiple time series [30].

The MAE and RMSE are expressed as follows:

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |Y_t - \hat{Y}_t| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} |Y_t - \hat{Y}_t|^2} \tag{10}$$

$$MASE = \frac{1}{n} \sum_{t=1}^{n} \left(|y_t - \hat{y}_t| \Big/ \frac{1}{n-1} \sum_{t=2}^{n} |y_t - \hat{y}_t - 1| \right) \tag{11}$$

where y_t denotes the actual values, $\hat{y}_t$ represent the predicted values, and n indicates the number of observations.

It is generally accepted that the lower the error metrics, the better the method [30,49]. All error metric values were calculated for models developed from Data: 2010–2019 to assess the accuracy of such models for in-sample data (training data), whereas these values were also determined to assess how the models developed from the training data performed on out-of-sample data (validation data; Data: 2020).

In addition, we also forecasted the n-FMD episodes for the period of 2021–2023 based on 2010–2020 FMD data.

3. Results

3.1. Trends and Seasonality

From 2013 to 2019, there was a seasonal increase in the number of n-FMD episodes. The year with the most episodes was 2016 (Figure 2). There was a downward trend in n-FMD episodes from mid-2016 to mid-2017. From late 2018 to mid-2019, the n-FMD episodes appeared to be stable. Following that, there was an upward trend in the number of n-FMD episodes.
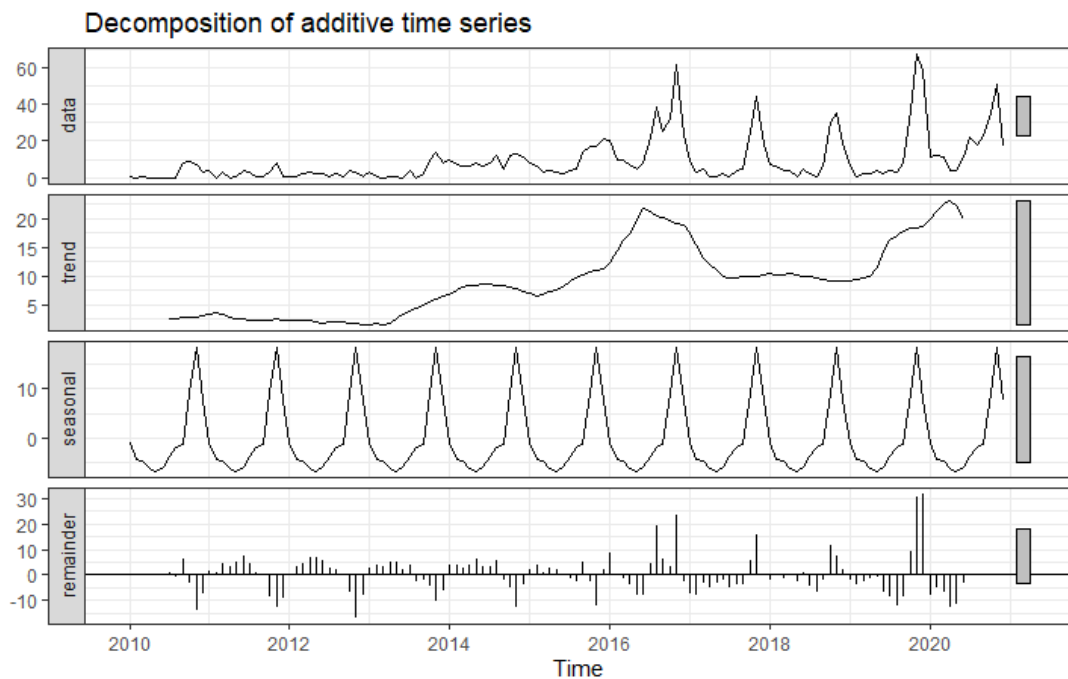
Figure 2. Decomposition of the number of time-series FMD outbreak episodes from January 2010 to December 2020 in actual (data), trend, decomposed seasonal trait (seasonal), and random fluctuation (remainder) of FMD outbreak episodes were illustrated.

The seasonal pattern demonstrated that the peak of n-FMD episodes was mainly observed from September to December annually.

3.2. Fitted Time-Series Models, Model Performances, and Forecasts

For the entire dataset, the program generated and tested 192 potential models, $SARIMA(1,0,1)(0,1,1)_{12}$ was the best-fitting model with the lowest AIC. Assumptions for model residuals were tested before this model could be used to forecast (Figure 3). The ACF of residuals showed no significant deviation from a zero-mean white noise process, and all *p* values for the Ljung–Box statistic were large ($p > 0.05$), indicating no violations. The best forecasting model for NNAR was found to be $NNAR(3,1,2)_{12}$, whereas the fitted model for ETS was $ETS(A,N,A)$. The fitted model for TBATS was $TBATS(1,\{0,0\},0.8,\{<12,5>\})$.

Figure 3 shows the actual and fitted values for each final model derived from the training data (Data: 2010–2019). Furthermore, actual and forecast values from the final models tested with validation data (Data: 2020) were illustrated (Figure 4). The projection appearance of actual and fitted values among models appears to be similar graphically (Figure 3), whereas actual and forecast values of the n-FMD episodes appear to differ among models (Figure 4). Additionally, the forecast values of n-FMD episode from each model are presented in Supplementary Materials Table S2.

In terms of RMSE, MAE, and MASE (Table 1), it was found that using NNAR for training data produced results that outperformed those obtained from other single-technique methods. On the other hand, the lowest values of the RMSE, MAE, and MASE from testing data demonstrate that SARIMA is more accurate than other competing models in terms of accuracy. In hybrid models, the SARIMA-NNAR and NNAR-TBATS models outperformed the others.
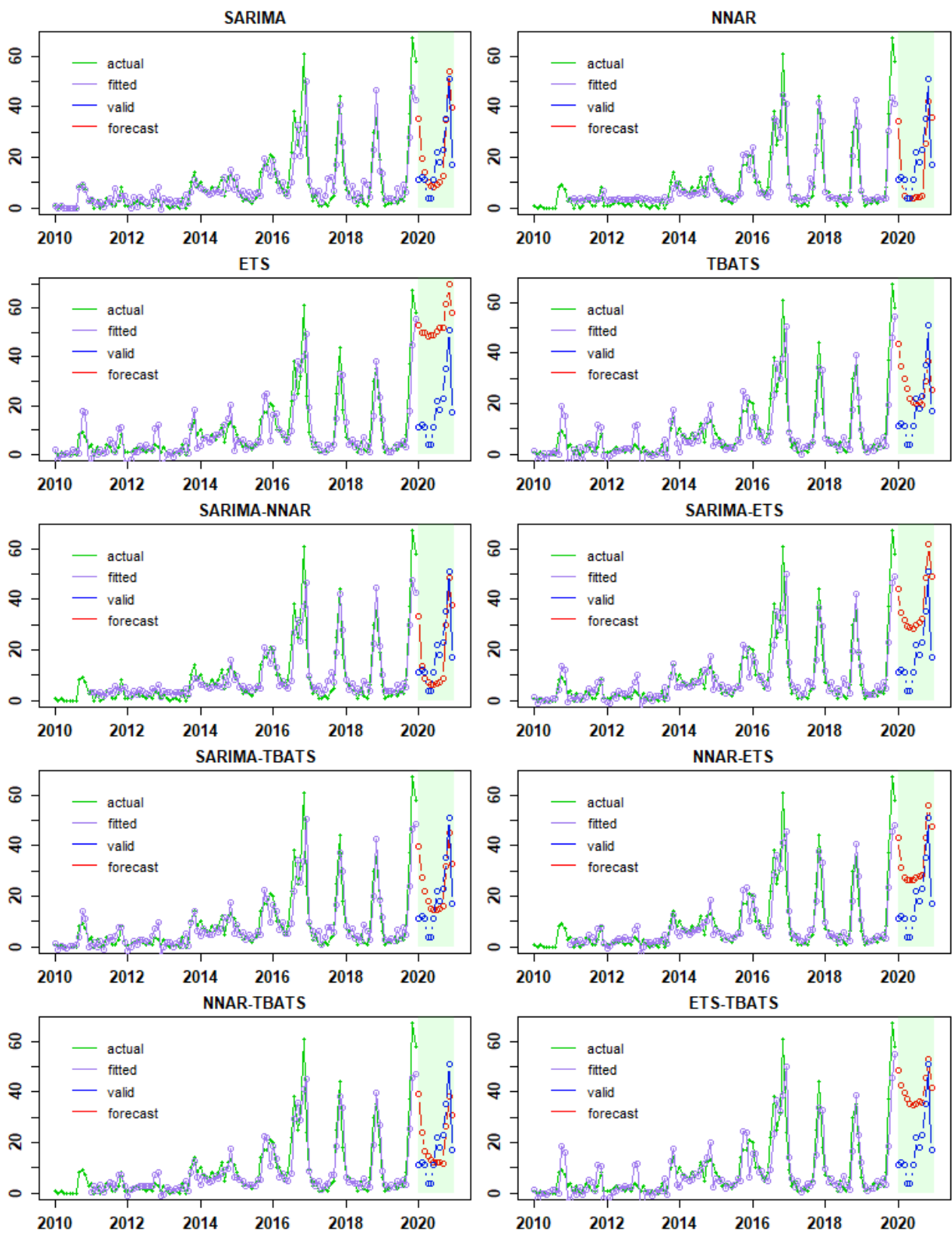
Figure 3. Actual, fitted, and forecast value from SARIMA, NNAR, ETS, TBATS. SARIMA-NNAR, SARIMA-ETS, SARIMA-TBATS, NNAR-ETS, NNAR-TBATS, and ETS-TBATS models. The x-axis and y-axis are the year and number of FMD outbreak episodes, respectively.
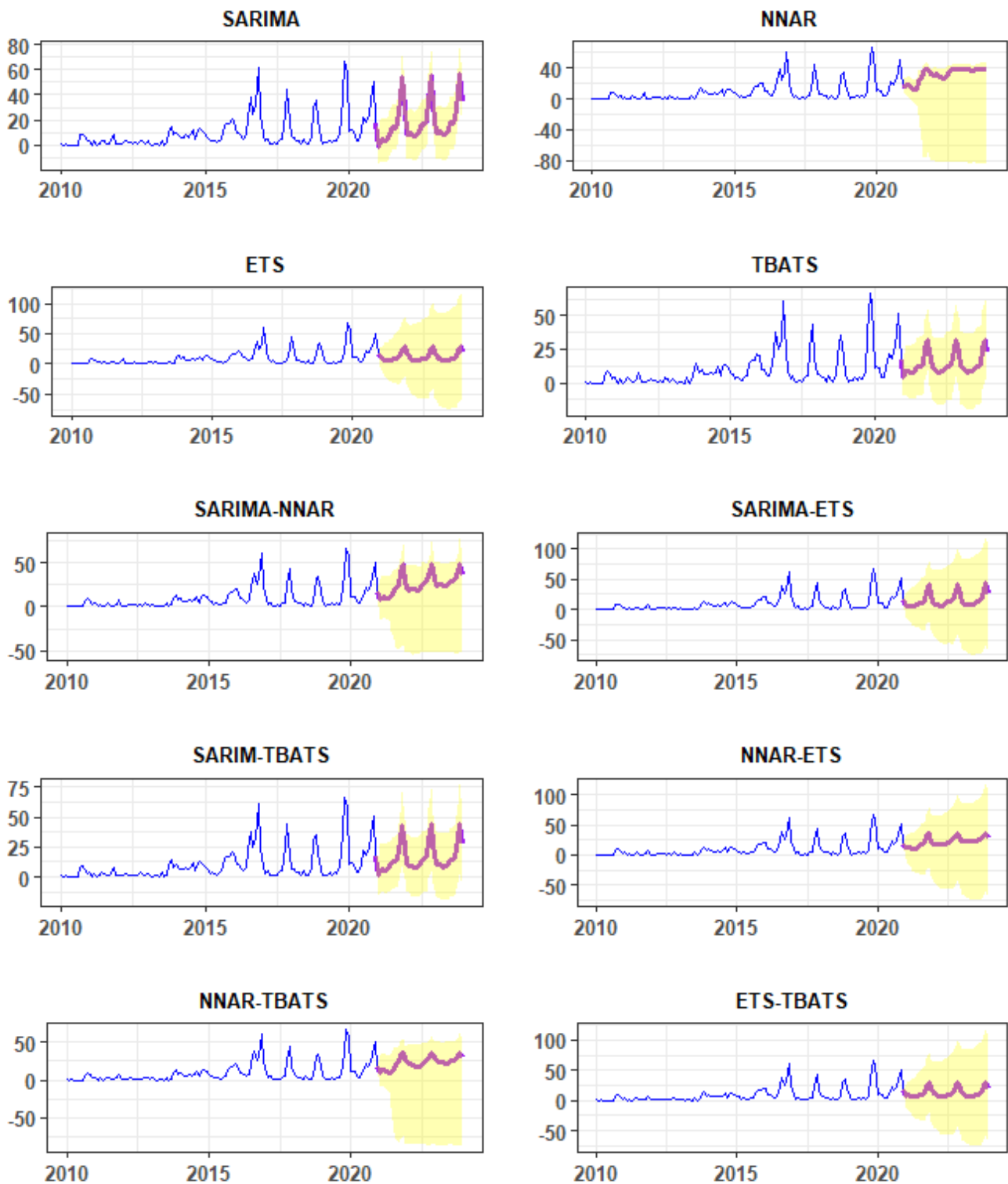
Figure 4. Forecasts of the number of FMD episodes (red line) from SARIMA, NNAR, ETS, TBATS, SARIMA-NNAR, SARIMA-ETS, SARIMA-TBATS, NNAR-ETS, NNAR-TBATS, and ETS-TBATS models. The x-axis and y-axis are the year and number of FMD outbreak episodes, respectively. The yellow band indicates a 95% confidence interval of forecast values.

Table 1. Error matrices for time-series models applied to training and validation datasets.

Model [1]	Training Data (Data: 2010–2019)			Validation Data (Data: 2020)		
	RMSE [2]	MAE	MASE	RMSE	MAE	MASE
SARIMA	6.21	3.81	0.64	11.47	8.75	1.47
NNAR	5.21	3.39	0.56	12.61	10.46	1.76
ETS	6.59	4.44	0.75	36.11	35.26	5.94
TBATS	6.38	4.26	0.89	16.24	13.38	2.25
SARIMA-NNAR	5.59	3.63	0.61	11.43	8.79	1.48
SARIMA-ETS	5.94	3.82	0.64	20.89	19.2	3.23
SARIMA-TBATS	5.91	3.75	0.63	12.56	10.4	1.75
NNAR-ETS	5.38	3.52	0.60	17.49	15.17	2.60
NNAR-TBATS	5.40	3.53	0.60	11.41	10.02	1.69
ETS-TBATS	6.42	4.29	0.72	24.44	22.14	3.73

[1] SARIMA = seasonal autoregressive integrated moving average, ETS = error trend seasonality, NNAR = neural network autoregression, TBATS = Trigonometric Exponential smoothing state-space model with Box–Cox transformation, autoregressive integrated moving errors, Trend and Seasonal components. [2] RMSE = root mean squared error, MAE = mean absolute error and MASE = mean absolute scaled error.

4. Discussion

In the present study, we applied the time-series method to FMD outbreak data to determine epidemiological features and explore the predictivity in forecasting the n-FMD episodes through various time-series models. Based on the most recent outbreak data used in this study, forecasts of n-FMD episodes for the following three years were also projected.

From 2014 to 2019, there was an increasing trend of n-FMD episodes in Thailand, which was in agreement with a previous report [16]. The increasing trend could be attributed to the presence of more than one serotype of FMD virus causing outbreaks. In general, FMD virus serotype O was responsible for the majority of the FMD outbreaks in Thailand each year [16,17]. However, FMD virus serotype A has become more prevalent since 2013. During the years 2013–2015, both serotypes A and O were responsible for a number of FMD outbreaks in different parts of the country [16]. It has been discussed that the presence of two serotypes causing outbreaks nationwide might pose a difficulty in preventing and controlling the disease [17,50]. The variation in the genetics of FMD virus serotype A is a major concern because the ratio of antibody titer against the heterologous field strain and antibody titer against the homologous vaccine strain or r-value [51] is not high, indicating that a moderate-level matching is observed between FMD vaccine strains and FMD outbreak strains, especially in 2016, 2017, and 2019 [17]. Given that, it is likely that vaccines used in some areas could not adequately provide disease protection [17]. Another explanation for the increase in n-FMD episodes is the enhancement of surveillance programs by DLD, which encourages farmers to report FMD outbreaks to DLD officers [50].

Our findings revealed a seasonal variation, indicating that the FMD outbreak in Thailand occurred on a regular basis. A predominant peak of reported FMD outbreaks was observed between October to December annually. This finding could be due to the possibility that climatic influences cause stress in cattle as they transition from the rainy season to the winter season, and thus, they are more susceptible to the disease [50]. Moreover, the high degree of trade activity throughout the last 3 months of the year might potentially have a role in the spreading of the disease across locations [50]. The finding of the seasonality of n-FMD episodes in this study was in agreement with several studies conducted in Thailand [52,53], Vietnam [9], and Colombia [41] but different from a report in Ethiopia that showed no seasonal trend for FMD outbreaks [7].

There is a lot of ambiguity and nonlinearity in FMD epidemic data. Unsatisfactory results may be obtained by predictive models that fail to account for seasonal fluctuation and nonlinearity [54]. Thus, SARIMA, ETS, and TBATS methods that account for seasonal variation, the nonlinear NNAR approach that incorporates seasonality and nonlinearity in its algorithm, and hybrid models that combine some of the advantages of these methods were utilized. According to our findings, NNAR outperformed other single-technique

methods. Better performance of NNAR may be due to its ability to take account of the nonlinear characteristics of n-FMD episodes exhibited during some parts of the study period. However, with validation data, the SARIMA model outperformed other single-technique methods in terms of validity. This could be because SARIMA appears to be an effective model for short-term forecasting [26,27]. However, a model's performance changes depending on the data it utilizes. Thus, we propose examining both SARIMA and NNAR predictions together for predicting future n-FMD episodes because they each have their own set of advantages. Furthermore, our findings demonstrated that the prediction ability from the single-technique methods and hybrid methods was likely to be similar. Nevertheless, it should not be extrapolated that hybrid models will not improve prediction ability for other FMD outbreak data or other data types because several studies from various fields have shown that hybrid models outperform non-hybrid models [28,33,44,55]. As addressed above, the performance of time-series methods is varied depending on characteristics and nature of the data. Thus, we suggest building multiple models from the same dataset to get better predictive insights.

Furthermore, it is generally recommended that the time-series method should be used on an updated dataset [40]. Hence, we recommend developing a time-series model to forecast n-FMD episodes on a regular basis, such as every six months or every year, for continuous projections and better prediction.

This study has some limitations to be considered. First, the FMD dataset is likely to be subjected to underreporting bias, as also indicated by other studies conducted in Thailand [56] and elsewhere [57]. Second, we did not determine temporary patterns for n-FMD episodes at provincial or regional levels due to the fact that the majority of the monthly FMD episodes data would contain zero values, which may not be appropriate for SARIMA and NNAR. For follow-up studies, other time-series methods should be investigated. In addition, FMD outbreak data from several countries in the region should be analyzed.

There is evidence that the same FMDV strains (e.g., O/EA/Mya-98, O/Middle East-South Asia, and O/ME-SA/Ind-2001) circulate in different Southeast Asian countries [8], suggesting that regional cooperation is required to control disease spread across borders. The lack of effective outbreak alerts and FMD outbreak data sharing among Southeast Asian endemic countries is one of the most difficult challenges for FMD prevention [58]. The analysis of retrospective time-series data and the prediction of future FMD outbreak occurrences in Thailand in this study would provide important information and warning messages to FMD endemic neighboring countries. The prediction of future n-FMD episodes would support the existing livestock surveillance system [16]. Livestock authorities can use the prediction from time-series models to develop an additional strategic plan for relevant control strategies for future FMD outbreaks, such as an implementation of rigorous animal movement control during the last three months of the year. Furthermore, in the Southeast Asia region, strict quarantine procedures should be implemented when cattle are transported across international borders. We also recommended greater regional cooperation in enhancing FMD prevention and controls.

5. Conclusions

To our knowledge, this is the first study to use the SARIMA, ETS, NNAR, TBATS, and hybrid models for the n-FMD episodes in Thailand to describe trends and seasonality as well as forecast the future n-FMD episodes. This study provides a better understanding of FMD outbreaks in Thailand in terms of trend, seasonality, and forecast based on 10-year outbreak data. The approaches demonstrated in this study could be used by livestock authorities to forecast FMD outbreak episodes in the short and long term and thus design or improve the FMD control strategies to prevent future outbreaks. The results of the study were not limited to Thailand; neighboring countries could use our findings and predictions as basic information to develop a regional strategy to mitigate FMD outbreaks in the region.

References

1. OIE. Foot and Mouth Disease. 2022. Available online: https://www.oie.int/en/disease/foot-and-mouth-disease/ (accessed on 23 December 2021).
2. Brito, B.; Rodriguez, L.; Hammond, J.; Pinto, J.; Perez, A. Review of the global distribution of foot-and-mouth disease virus from 2007 to 2014. *Transbound. Emerg. Dis.* **2017**, *64*, 316–332. [CrossRef] [PubMed]
3. Govindaraj, G.; Krishnamohan, A.; Hegde, R.; Kumar, N.; Prabhakaran, K.; Wadhwan, V.M.; Kakker, N.; Lokhande, T.; Sharma, K.; Kanani, A. Foot and Mouth Disease (FMD) incidence in cattle and buffaloes and its associated farm-level economic costs in endemic India. *Prev. Vet. Med.* **2021**, *190*, 105318.
4. Knight-Jones, T.J.; Robinson, L.; Charleston, B.; Rodriguez, L.; Gay, C.; Sumption, K.J.; Vosloo, W. Global Foot-and-Mouth Disease Research Update and Gap Analysis: 1-Overview of Global Status and Research Needs. *Transbound. Emerg. Dis.* **2016**, *63*, 3–13. [CrossRef] [PubMed]
5. Alhaji, N.; Amin, J.; Aliyu, M.; Mohammad, B.; Babalobi, O.; Wungak, Y.; Odetokun, I. Economic impact assessment of foot-and-mouth disease burden and control in pastoral local dairy cattle production systems in northern Nigeria: A cross-sectional survey. *Prev. Vet. Med.* **2020**, *177*, 104974. [CrossRef] [PubMed]
6. Ahmar, A.S.; Boj, E. Application of Neural Network Time Series (NNAR) and ARIMA to Forecast Infection Fatality Rate (IFR) of COVID-19 in Brazil. *JOIV Int. J. Inform. Vis.* **2021**, *5*, 8–10. [CrossRef]
7. Jemberu, W.; Mourits, M.; Sahle, M.; Siraw, B.; Vernooij, J.; Hogeveen, H. Epidemiology of foot and mouth disease in E thiopia: A retrospective analysis of district level outbreaks, 2007–2012. *Transbound. Emerg. Dis.* **2016**, *63*, 246–259. [CrossRef]
8. Gunasekera, U.C.; Sivasothy, A.; Wedasingha, N.; Thayaparan, S.; Rotewewa, B.; Muralithas, M.; Baumann, M.P.O.; Punyapornwithaya, V. Analyzing the Foot and Mouth Disease outbreak as from 2008 to 2014 in cattle and buffaloes in Sri Lanka. *Prev. Vet. Med.* **2017**, *148*, 78–88. [CrossRef]
9. Madin, B. An evaluation of Foot-and-Mouth Disease outbreak reporting in mainland South-East Asia from 2000 to 2010. *Prev. Vet. Med.* **2011**, *102*, 230–241. [CrossRef]
10. Picado, A.; Guitian, F.; Pfeiffer, D. Space–time interaction as an indicator of local spread during the 2001 FMD outbreak in the UK. *Prev. Vet. Med.* **2007**, *79*, 3–19. [CrossRef]
11. Bouma, A.; Elbers, A.; Dekker, A.; De Koeijer, A.; Bartels, C.; Vellema, P.; Van der Wal, P.; Van Rooij, E.; Pluimers, F.; De Jong, M. The foot-and-mouth disease epidemic in The Netherlands in 2001. *Prev. Vet. Med.* **2003**, *57*, 155–166. [CrossRef]
12. Upadhyaya, S.; Mahapatra, M.; Mioulet, V.; Parida, S. Molecular Basis of Antigenic Drift in Serotype O Foot-and-Mouth Disease Viruses (2013–2018) from Southeast Asia. *Viruses* **2021**, *13*, 1886. [CrossRef] [PubMed]
13. Gunasekara, U.; Bertram, M.R.; Dung, D.H.; Hoang, B.H.; Phuong, N.T.; Hung, V.V.; Long, N.V.; Minh, P.Q.; Vu, L.T.; Dong, P.V. Use of Slaughterhouses as Sentinel Points for Genomic Surveillance of Foot-and-Mouth Disease Virus in Southern Vietnam. *Viruses* **2021**, *13*, 2203. [CrossRef] [PubMed]
14. Bo, L.L.; Lwin, K.S.; Ungvanijban, S.; Knowles, N.J.; Wadsworth, J.; King, D.P.; Abila, R.; Qiu, Y. Foot-and-mouth disease outbreaks due to an exotic serotype Asia 1 virus in Myanmar in 2017. *Transbound. Emerg. Dis.* **2019**, *66*, 1067–1072. [CrossRef] [PubMed]

15. Blacksell, S.D.; Siengsanan-Lamont, J.; Kamolsiripichaiporn, S.; Gleeson, L.J.; Windsor, P.A. A history of FMD research and control programmes in Southeast Asia: Lessons from the past informing the future. *Epidemiol. Infect.* **2019**, *147*, e171. [CrossRef] [PubMed]

16. Arjkumpa, O.; Yano, T.; Prakotcheo, R.; Sansamur, C.; Punyapornwithaya, V. Epidemiology and National Surveillance System for Foot and Mouth Disease in Cattle in Thailand during 2008–2019. *Vet Sci.* **2020**, *7*, 99. [CrossRef] [PubMed]

17. Seeyo, K.B.; Nishi, T.; Kawaguchi, R.; Ungvanijban, S.; Udon, R.; Fukai, K.; Yamakawa, M.; Rukkwamsuk, T. Evolution of antigenic and genetic characteristics of foot-and-mouth disease virus serotype A circulating in Thailand, 2007–2019. *Virus Res.* **2020**, *290*, 198166. [CrossRef]

18. Arjkumpa, O.; Picasso-Risso, C.; Perez, A.; Punyapornwithaya, V. Subdistrict-Level Reproductive Number for Foot and Mouth Disease in Cattle in Northern Thailand. *Front. Vet. Sci.* **2021**, *8*, 757132. [CrossRef]

19. Sangrat, W.; Thanapongtharm, W.; Poolkhet, C. Identification of risk areas for foot and mouth disease in Thailand using a geographic information system-based multi-criteria decision analysis. *Prev. Vet. Med.* **2020**, *185*, 105183. [CrossRef]

20. Chanchaidechachai, T.; de Jong, M.C.; Fischer, E.A. Spatial model of foot-and-mouth disease outbreak in an endemic area of Thailand. *Prev. Vet. Med.* **2021**, *195*, 105468. [CrossRef]

21. Arjkumpa, O.; Sansamur, C.; Sutthipankul, P.; Inchaisri, C.; Na Lampang, K.; Charoenpanyanet, A.; Punyapornwithaya, V. Spatiotemporal analyses of foot and mouth disease outbreaks in cattle farms in Chiang Mai and Lamphun, Thailand. *BMC Vet. Res.* **2020**, *16*, 170. [CrossRef]

22. Gleeson, L. A review of the status of foot and mouth disease in South-East Asia and approaches to control and eradication. *Rev. Sci. Tech. Off. Int. Epiz.* **2002**, *21*, 465–472. [CrossRef] [PubMed]

23. Lee, H.S.; Her, M.; Levine, M.; Moore, G.E. Time series analysis of human and bovine brucellosis in South Korea from 2005 to 2010. *Prev. Vet. Med.* **2013**, *110*, 190–197. [CrossRef] [PubMed]

24. Sato, R.C. Disease management with ARIMA model in time series. *Einstein* **2013**, *11*, 128. [CrossRef] [PubMed]

25. Ward, M.; Iglesias, R.; Brookes, V.J. Autoregressive models applied to time-series data in veterinary science. *Front. Vet. Sci.* **2020**, *7*, 604. [CrossRef]

26. Tian, C.; Wang, H.; Luo, X. Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. *Epidemiol. Infect.* **2019**, *147*, e82. [CrossRef]

27. Weng, R.; Fu, H.; Zhang, C.; Ye, J.; Hong, F.; Chen, X.; Cai, Y. Time series analysis and forecasting of chlamydia trachomatis incidence using surveillance data from 2008 to 2019 in Shenzhen, China. *Epidemiol. Infect.* **2020**, *148*. [CrossRef]

28. Wang, K.; Deng, C.; Li, J.; Zhang, Y.; Li, X.; Wu, M. Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiol. Infect.* **2017**, *145*, 1118–1129. [CrossRef]

29. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Melbourne, Australia, 2018.

30. Perone, G. Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy. *Eur. J. Health Econ.* **2021**, *4*, 1–24. [CrossRef]

31. Yu, C.; Xu, C.; Li, Y.; Yao, S.; Bai, Y.; Li, J.; Wang, L.; Wu, W.; Wang, Y. Time series analysis and forecasting of the hand-foot-mouth disease morbidity in China using an advanced exponential smoothing state space TBATS model. *Infect. Drug Resist.* **2021**, *14*, 2809. [CrossRef]

32. Azeez, A.; Obaromi, D.; Odeyemi, A.; Ndege, J.; Muntabayi, R. Seasonality and trend forecasting of tuberculosis prevalence data in Eastern Cape, South Africa, using a hybrid model. *Int. J. Environ. Res. Public Health* **2016**, *13*, 757. [CrossRef]

33. Wang, Y.; Xu, C.; Li, Y.; Wu, W.; Gui, L.; Ren, J.; Yao, S. 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. [CrossRef] [PubMed]

34. Zhou, L.; Xia, J.; Yu, L.; Wang, Y.; Shi, Y.; Cai, S.; Nie, S. Using a hybrid model to forecast the prevalence of schistosomiasis in humans. *Int. J. Environ. Res. Public Health* **2016**, *13*, 355. [CrossRef] [PubMed]

35. Zou, J.; Han, Y.; So, S.-S. Overview of artificial neural networks. In *Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 14–22.

36. Abotaleb, M.; Makarovskikh, T.; Yonar, A.; Badr, A.; Mishra, P.; Williams, A.; Yonar, H. State of the art in wind speed in England using BATS, TBATS, Holt's Linear and ARIMA model: Wind speed in England using BATS, TBATS, Holt's Linear and ARIMA model. *Mausam* **2022**, *73*, 129–138. [CrossRef]

37. Zhou, L.; Yu, L.; Wang, Y.; Lu, Z.; Tian, L.; Tan, L.; Shi, Y.; Nie, S.; Liu, L. A hybrid model for predicting the prevalence of schistosomiasis in humans of Qianjiang City, China. *PLoS ONE* **2014**, *9*, e104875. [CrossRef] [PubMed]

38. Peralta, E.A.; Carpenter, T.E.; Farver, T.B. The application of time series analysis to determine the pattern of foot-and-mouth disease in cattle in Paraguay. *Prev. Vet. Med.* **1982**, *1*, 27–36. [CrossRef]

39. Ferguson, N.M.; Donnelly, C.A.; Anderson, R.M. The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. *Science* **2001**, *292*, 1155–1160. [CrossRef] [PubMed]

40. Souley Kouato, B.; De Clercq, K.; Abatih, E.; Dal Pozzo, F.; King, D.P.; Thys, E.; Marichatou, H.; Saegerman, C. Review of epidemiological risk models for foot-and-mouth disease: Implications for prevention strategies with a focus on Africa. *PLoS ONE* **2018**, *13*, e0208296. [CrossRef]

41. Gallego, M.; Perez, A.; Thurmond, M. Temporal and spatial distributions of foot-and-mouth disease under three different strategies of control and eradication in Colombia (1982–2003). *Vet. Res. Commun.* **2007**, *31*, 819–834. [CrossRef]

42. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2021.

43. Azad, A.S.; Sokkalingam, R.; Daud, H.; Adhikary, S.K.; Khurshid, H.; Mazlan, S.N.A.; Rabbani, M.B.A. Water Level Prediction through Hybrid SARIMA and ANN Models Based on Time Series Analysis: Red Hills Reservoir Case Study. *Sustainability* **2022**, *14*, 1843. [CrossRef]

44. Yusof, F.; Kane, I. Modelling monthly rainfall time series using ETS state space and SARIMA models. *Int. J. Curr. Res.* **2012**, *4*, 195–200.

45. Sun, Z. Comparison of Trend Forecast Using ARIMA and ETS Models for S&P500 Close Price. In Proceedings of the 2020 the 4th International Conference on E-Business and Internet, Singapore, 9–11 October 2020; pp. 57–60.

46. De Livera, A.M.; Hyndman, R.J.; Snyder, R.D. Forecasting time series with complex seasonal patterns using exponential smoothing. *J. Am. Stat. Assoc.* **2011**, *106*, 1513–1527. [CrossRef]

47. Shaub, D. Fast and accurate yearly time series forecasting with forecast combinations. *Int. J. Forecast.* **2020**, *36*, 116–120. [CrossRef]

48. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]

49. Yano, T.; Premashthira, S.; Dejyong, T.; Tangtrongsup, S.; Salman, M.D. The Effectiveness of a Foot and Mouth Disease Outbreak Control Programme in Thailand 2008–2015: Case Studies and Lessons Learned. *Vet. Sci.* **2018**, *5*, 101. [CrossRef] [PubMed]

50. Samuel, A.; Ouldridge, E.; Arrowsmith, A.; Kitching, R.; Knowles, N. Antigenic analysis of serotype O foot-and-mouth disease virus isolates from the Middle East, 1981 to 1988. *Vaccine* **1990**, *8*, 390–396. [CrossRef]

51. Cleland, P.C.; Baldock, F.C.; Chamnanpood, P. Village level risk factors for foot-and-mouth disease in northern Thailand. *Prev. Vet. Med.* **1996**, *26*, 253–261. [CrossRef]

52. Yano, T. Study of foot and mouth disease outbreak in Chiang Mai and Lamphoon. *ChiangMai V. J.* **2011**, *11*, 277–287.

53. Wang, Y.; Xu, C.; Wu, W.; Ren, J.; Li, Y.; Gui, L.; Yao, S. Time series analysis of temporal trends in hemorrhagic fever with renal syndrome morbidity rate in China from 2005 to 2019. *Sci. Rep.* **2020**, *10*, 9609. [CrossRef]

54. Xiao, Y.; Xiao, J.; Wang, S. A hybrid model for time series forecasting. *Hum. Syst. Manag.* **2012**, *31*, 133–143. [CrossRef]

55. Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [CrossRef]

56. Sansamur, C.; Wiratsudakul, A.; Charoenpanyanet, A.; Punyapornwithaya, V. Estimating the number of farms experienced foot and mouth disease outbreaks using capture-recapture methods. *Trop. Anim. Health Prod.* **2020**, *53*, 12. [CrossRef] [PubMed]

57. Gunarathne, A.; Kubota, S.; Kumarawadu, P.; Karunagoda, K.; Kon, H. Is hiding foot and mouth disease sensitive behavior for farmers? A survey study in Sri Lanka. *Asian-Australas. J. Anim. Sci.* **2016**, *29*, 280. [CrossRef] [PubMed]

58. OIE. SEACFMD Roadmap. 2016. Available online: https://rr-asia.oie.int/en/projects/foot-and-mouth-disease-fmd/seacfmd-campaign/seacfmd-roadmap/ (accessed on 23 December 2021).

59. Al Khatib, A.M.G.; Yonar, H.; Abotaleb, M.; Mishra, P.; Yonar, A.; Karakaya, K.; Badr, A.; Dhaka, V. Modeling and forecasting of egg production in India using time series models. *Eurasian J. Vet. Sci.* **2021**, *37*, 265–273. [CrossRef]

60. Mishra, P.; Al Khatib, A.M.G.; Sardar, I.; Mohammed, J.; Ray, M.; Manish, K.; Rawat, D.; Pandey, S.A.; Dubey, A.; Feys, J.; et al. Modelling and forecasting of COVID-19 in India. *J. Infect. Dis. Epidemiol.* **2020**, *6*, 1–11. [CrossRef]

61. Mishra, P.; Alakkari, K.M.; Lama, A.; Ray, S.; Singh, M.; Shoko, C.; Abotaleb, M.; Al khatib, A.M.G.; Karakaya, K. Modeling and Forecasting of Sugarcane Production in South Asian Countries. *Curr. Appl. Sci. Technol.* **2023**, *23*, 1–15. [CrossRef]

62. Raghav, Y.S.; Mishra, P.; Alakkari, K.M.; Singh, M.; Al Khatib, A.M.G.; Balloo, R. Modelling and forecasting of pulses production in south asian countries and its role in nutritional security. *Legume Res.-Int. J.* **2022**, *45*, 454–461. [CrossRef]

63. Rahman, U.H.; Ray, S.; Al Khatib, A.M.G.; Lal, P.; Mishra, P.; Fatih, C.; Williams, A.J.; Karakaya, K.; Shrivastri, S.; Alakkari, K. State of Art of SARIMA Model in Second Wave on COVID-19 in India. *Int. J. Agric. Stat. Sci.* **2022**. Available online: https://connectjournals.com/03899.2022.18.141 (accessed on 28 March 2022).

64. Schreiber, M.A.; Cavanaugh, C.H. The Application of Machine Learning Methods to Time Series Forecasting. 2020. Available online: https://research-api.cbs.dk/ws/portalfiles/portal/66025537/Cavanaugh_and_Schreiber.pdf (accessed on 28 March 2022).

65. Yonar, H.; Yonar, A.; Mishra, P.; Abotaleb, M.; Al Khatib, A.M.G.; Makarovskikh, T.; Cam, M. Modeling and forecasting of milk production in different breeds in Turkey. *Indian J. Anim. Sci.* **2022**, *92*, 105–111.

66. Abotaleb, M.; Ray, S.; Mishra, P.; Karakaya, K.; Shoko, C.; Al khatib, A.M.G.; Balloo, R. Modelling and forecasting of rice production in south Asian countries. *AMA Agric. Mech. Asia Afr. Lat. Am.* **2021**, *51*, 1611–1627.