frontiers | Frontiers in Veterinary Science

TYPE Original Research
PUBLISHED 29 November 2023
DOI 10.3389/fvets.2023.1294049

Check for updates

OPEN ACCESS

EDITED BY
Javier Caballero Gómez,
Maimonides Biomedical Research Institute of Cordoba (IMIBIC), Spain

REVIEWED BY
A. K. M. Anisur Rahman,
Bangladesh Agricultural University, Bangladesh
Rouzbeh Bashar,
Pasteur Institute of Iran, Iran

*CORRESPONDENCE
Veerasak Punyapornwithaya
✉ veerasak.p@cmu.ac.th
Orapun Arjkumpa
✉ arjkumpa@hotmail.com

RECEIVED 15 September 2023
ACCEPTED 08 November 2023
PUBLISHED 29 November 2023

CITATION
Punyapornwithaya V, Thanapongtharm W, Jainonthee C, Chinsorn P, Sagarasaeranee O, Salvador R and Arjkumpa O (2023) Time series analysis and forecasting of the number of canine rabies confirmed cases in Thailand based on national-level surveillance data. *Front. Vet. Sci.* 10:1294049. doi: 10.3389/fvets.2023.1294049

COPYRIGHT
© 2023 Punyapornwithaya, Thanapongtharm, Jainonthee, Chinsorn, Sagarasaeranee, Salvador and Arjkumpa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Time series analysis and forecasting of the number of canine rabies confirmed cases in Thailand based on national-level surveillance data

Veerasak Punyapornwithaya[1,2,3]*, Weerapong Thanapongtharm[4], Chalita Jainonthee[1,2], Pornpiroon Chinsorn[5], Onpawee Sagarasaeranee[5], Roderick Salvador[6] and Orapun Arjkumpa[7]*

[1]Research Center for Veterinary Biosciences and Veterinary Public Health, Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai, Thailand, [2]Veterinary Public Health and Food Safety Centre for Asia Pacific, Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai, Thailand, [3]Department of Veterinary Biosciences and Veterinary Public Health, Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai, Thailand, [4]Department of Livestock Development, Bangkok, Thailand, [5]Companion Disease Control Division, Bureau of Disease Control and Veterinary Services, Department of Livestock Development, Bangkok, Thailand, [6]College of Veterinary Science and Medicine, Central Luzon State University, Science City of Muñoz, Nueva Ecija, Philippines, [7]The 4th Regional Livestock Office, Department of Livestock Development, Khon Kaen, Thailand

Introduction: Rabies, a deadly zoonotic viral disease, accounts for over 50,000 fatalities globally each year. This disease predominantly plagues developing nations, with Thailand being no exception. In the current global landscape, concerted efforts are being mobilized to curb human mortalities attributed to animal-transmitted rabies. For strategic allocation and optimization of resources, sophisticated and accurate forecasting of rabies incidents is imperative. This research aims to determine temporal patterns, and seasonal fluctuations, and project the incidence of canine rabies throughout Thailand, using various time series techniques.

Methods: Monthly total laboratory-confirmed rabies cases data from January 2013 to December 2022 (full dataset) were split into the training dataset (January 2013 to December 2021) and the test dataset (January to December 2022). Time series models including Seasonal Autoregressive Integrated Moving Average (SARIMA), Neural Network Autoregression (NNAR), Error Trend Seasonality (ETS), the Trigonometric Exponential Smoothing State-Space Model with Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS), and Seasonal and Trend Decomposition using Loess (STL) were used to analyze the training dataset and the full dataset. The forecast values obtained from the time series models applied to the training dataset were compared with the actual values from the test dataset to determine their predictive performance. Furthermore, the forecast projections from January 2023 to December 2025 were generated from models applied to the full dataset.

Results: The findings revealed a total of 4,678 confirmed canine rabies cases during the study duration, with apparent seasonality in the data. Among the models tested with the test dataset, TBATS exhibited superior predictive accuracy, closely trailed by the SARIMA model. Based on the full dataset, TBATS projections suggest an annual average of approximately 285 canine rabies cases for the years 2023 to 2025, translating to a monthly average of 23 cases (range: 18–30). In contrast, SARIMA projections averaged 277 cases annually (range: 208–214).

Discussion: This research offers a new perspective on disease forecasting through advanced time series methodologies. The results should be taken into consideration when planning and conducting rabies surveillance, prevention, and control activities.

KEYWORDS

rabies, confirmed cases, time series model, forecasting, Thailand

1 Introduction

Rabies is one of the most significant public health problems in several countries worldwide (1, 2). This disease is caused by the rabies virus belonging to the genus *Lyssavirus* of the family *Rhabdoviridae*. Rabies is a fatal disease in humans and is considered to be somewhat neglected (3). Based on World Health Organization (WHO) data, it is estimated that ∼59,000 people die from dog-mediated rabies each year, mostly in Asia and Africa (4). Due to the global burden of rabies, the WHO, in cooperation with the Global Alliance for Rabies Control, the Food and Agriculture Organization of the United Nations (FAO), and the World Organization for Animal Health (OIE) have set a goal to reduce human rabies deaths to zero by 2030 (5).

In Thailand, rabies is considered an important notifiable disease (6, 7). From 2010 to 2015, rabies claimed 46 lives, and annually, over 600,000 individuals undergo post-exposure prophylaxis treatment. Dogs are identified as the primary reservoirs for the rabies virus and play a crucial role in transmitting the disease to humans and other animals (8, 9). Notably, a significant proportion of confirmed rabies cases in Thailand are linked to dogs (9, 10). As a result, rabies prevention efforts predominantly focus on curbing the transmission from dogs (11, 12). Collaborative efforts across various organizations in Thailand aim to reduce rabies incidence in both humans and animals. Within the animal sector, the Department of Livestock Development (DLD) assumes a pivotal role in overseeing rabies trends and executing control strategies to diminish animal rabies cases. The surveillance of animal rabies in the country encompasses both active and passive methodologies (9, 10).

Accurate prediction of infectious disease trends is pivotal for optimizing resource allocation and strategizing future prevention and control measures. Essentially, predictions about future events are often based on historical data (13, 14). In the field of infectious disease epidemiology, forecasting the number of prospective cases or fatalities is a predominant focus, particularly when the magnitude of the susceptible population is not clearly defined (15–19). At present, several advanced time series methods are available, providing a wide range of techniques to work with various types of data with high-level predictability (14, 18).

Time series analysis is universally acknowledged as a cornerstone for forecasting across various fields, including economics (20), medicine (21, 22), veterinary science (23, 24), environmental studies (25), and agriculture (26–28). For instance, many recent studies have employed time series analysis to project COVID-19 case numbers, aiding in the formulation of disease control strategies and evaluating intervention efficacy (29–31). In the context of animal health, time series analysis has been instrumental in forecasting trends in diseases like rabies (24) and infectious disease in livestock (23, 32).

This research evaluates the efficacy of the seasonal autoregressive integrated moving average (SARIMA) in forecasting the number of confirmed canine rabies cases, representing classical time series modeling. Additionally, we explored advanced time series models (13), encompassing error trend seasonality (ETS), the trigonometric exponential smoothing state-space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS), and seasonal trend decomposition procedures based on loess (STL). The study also incorporates the neural network autoregression (NNAR), a common method aligned with machine learning approaches.

To our knowledge, prior studies have not employed advanced time series techniques to analyze and project the incidence of canine rabies cases in Thailand based on national surveillance data. This presents a significant gap in predictive knowledge, which is imperative for livestock authorities and relevant stakeholders to devise effective strategies against canine rabies. Therefore, this study aims to determine the patterns, seasonality, and to forecast the incidence of canine rabies using time series methodologies, including SARIMA, NNAR, ETS, TBATS, and STL models.

2 Materials and methods

2.1 Data and time series decomposition

2.1.1 Rabies cases data

In this study, we utilized data on confirmed canine rabies cases sourced from the Department of Livestock Development (DLD). This data was gathered through both active and passive rabies surveillance initiatives done by the Thai government (9). Passive surveillance primarily involved the laboratory submission of animal samples, either carcasses or heads, suspected of rabies. These samples were typically provided by veterinarians or animal owners. Meanwhile, active surveillance, done by veterinary services, focused on animals that succumbed to ambiguous symptoms. A baseline level of active sampling is also maintained, ensuring the collection of at least one sample from every subdistrict annually (9). Diagnostic procedures for all suspected animal rabies cases were conducted in nine DLD-accredited laboratories and one affiliated with the Queen Saovabha Memorial Institute. Diagnostic methods employed included the fluorescent antibody test and the mouse inoculation method (9, 10). All laboratory findings were subsequently uploaded to the centralized "ThaiRabies.net" platform (9). The dataset evaluated spans canine rabies cases recorded from January 2013 to December 2021.

2.1.2 Time series decomposition and determination of seasonality in the time series data

The Ollech and Webel's combined seasonality test (WO-test) was used to ascertain the presence of seasonality in the rabies dataset. The WO-test combines the results of both the QS test and the KW-test, which are computed based on the residuals of an automated non-seasonal ARIMA. A time series is classified as having seasonality by the WO-test if the *p*-value of the QS test is <0.01 or if the *p*-value of the KW-test is <0.002 (33).

The WO-test was performed using the R statistical software, leveraging the "*seastests*" package via the "*combined_test*" function (33). Comprehensive details of the test can be found in the package's manual. Additionally, the ETS model procedure was utilized to evaluate the seasonal component present in the data. Further details about the ETS can be found in the next section.

2.2 Time series analysis and forecasts

2.2.1 Analytical and modeling procedure

This study encompassed two primary phases: (i) Identification of the most efficacious forecasting model by finalizing a model from each time series method and subsequently evaluating its performance using a validation dataset, and (ii) Utilization of the superior time series method, as determined from the previous phase, to predict future canine rabies cases based on a comprehensive dataset. Forecasts derived from other time series models, aside from the top-performing one, were also examined for comparison.

This study involved two different approaches. First, the full dataset (January 2013 to December 2022) was divided into a training dataset (January 2013 to December 2021) and a test dataset (January to December 2022). Forecasts for January 2022 to December 2022 were then generated from the training dataset and compared with the actual values within the same period in the test dataset to assess the forecasting model's performance. Second, in an effort to ensure up-to-date and accurate forecasts, it is suggested that the most current data be used for prediction (28, 34).Therefore, all time series models were applied to the full dataset to generate forecasts for the following 3 years (January 2023–2025). A schematic representation of the modeling procedure is provided in Figure 1 for clarity.

2.2.2 Time series models

2.2.2.1 SARIMA model

The SARIMA is an extension of the autoregressive integrated moving average (ARIMA) that explicitly supports univariate time series data with a seasonal component. It adds three new hyperparameters to specify the autoregression, differencing, and moving average for the seasonal component of the series, as well as an additional parameter for the period of seasonality. The form of SARIMA is written as (35).

$$\phi_P(B)\,\Phi_p\left(B^s\right)(1-B)^d\left(1-B^s\right)^D x_t = \theta_Q(B)\,\Theta_Q\left(B^s\right)\omega_t \quad (1)$$

where φ and θ are the parameters of the autoregressive and moving average, respectively. The terms Φ and Θ represent the parameters of the seasonal autoregressive and seasonal moving average. Additionally, SARIMA can be defined as SARIMA $(p,d,q)(P,D,Q)$. The terms p, d, and q denote the order of autoregression, degree of differencing, and order of moving average, respectively. Meanwhile, P, D, and Q represent the orders of the seasonal autoregression, degree of seasonal differencing, and order of seasonal moving average, respectively. The term S refers to seasonal periodicity. The parameters were estimated using the maximum likelihood method. The best fit model was identified based on the minimum value of the corrected Akaike's Information Criterion (AIC) (13).

2.2.2.2 NNAR model

The NNAR model can be thought of as a network of neurons or nodes displaying complex non-linear relationships and functional forms. The term NNAR (p, k) is defined to indicate that there are (p) lagged inputs and (k) nodes in the hidden layer. With seasonal data, the NNAR model can be written as

$$\text{NNAR}\ \left(p,P,k\right)_m \quad (2)$$

where p is the number of non-seasonal lagged inputs for the linear autoregressive process (AR), P indicates the seasonal lags for the AR process, k denotes the number of neurons in the hidden layer, and m represents the seasonal period.

The forecasting process can be divided into two phases. Initially, the order of autoregression is determined. Subsequently, with the training dataset, the neural network is trained in accordance with the previously determined autoregression order. The number of input nodes or time series lags within the neural network is determined from this autoregression order.

2.2.2.3 ETS model

The state-space equations can be written as follows (36).

$$y_t = w\left(x_{t-1}\right) + r\left(x_{t-1}\right)\varepsilon_t \quad (3)$$

$$x_t = f\left(x_{t-1}\right) + g\left(x_{t-1}\right)\varepsilon_t \quad (4)$$

where w, f, r and g are coefficients while ε_t denotes the Gaussian white noise series. Equation (3) is known as the measurement equation, describing the relationship between the unobserved states x_{t-1} and the observation y_t. Equation (4) is the transitional equation, describing the evolution of states over time. The use of identical errors in these two equations makes it an innovative state-space model (13).

The final ETS model is represented by a three-character string (Z, Z, Z), where the first, second, and third characters represent the error (A or M), type of trend (A or Ad or N), and the type of season (A or N or M), respectively. The letters A, Ad, N, and M denote additive, additive damped, none, and multiplicative (36). Additionally, the ETS forecasts are based on a weighted average of past observations, and the weight decays exponentially over time. As a result, the final observations have a greater weight than the earlier ones.

2.2.2.4 TBATS model

The TBATS is an advanced adaptation of the BATS model, designed to accommodate multiple seasonal cycles. This model is

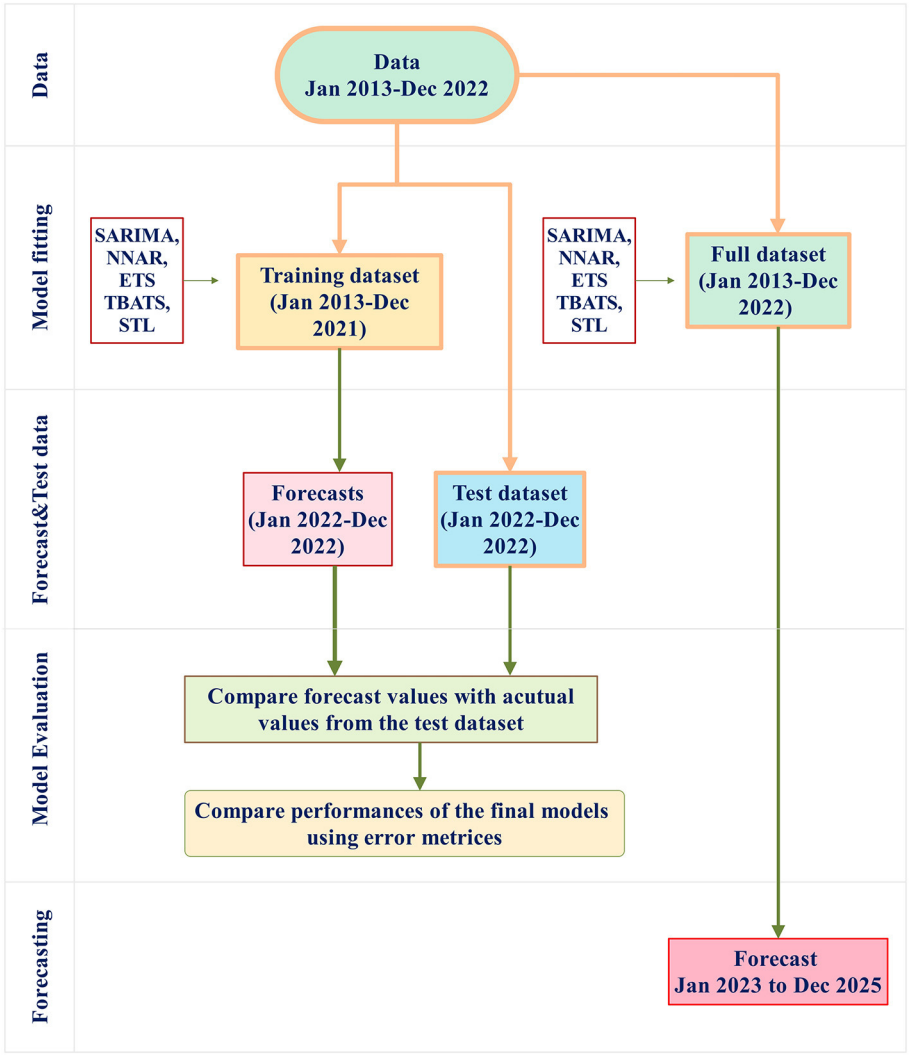

FIGURE 1
The modeling procedure involves several steps. First, the full dataset (January 2013 to December 2022) is divided into a training dataset (January 2013 to December 2021, shown as the orange box) and a test dataset (January to December 2022, shown as the blue box). The training dataset is used for model fitting using SARIMA, NNAR, ETS, TBATS, STL, and THETA methods. The final models from these methods are then evaluated for performance by applying them to the test dataset. Error metrics from these models are compared. Furthermore, the full dataset is used to build forecast models and implement the forecasting procedure to forecast the number of rabies cases in the period of January 2023 to December 2025.

supported by a trigonometric framework adept at navigating the intricacies of seasonality within a time series (19, 37).

The TBATS models is represented as $TBATS(\omega, p, q, \varphi, \{m_1, k_1\}, \{m_2, k_2\}, ..., \{m_T, k_T\})$. This formulation leverages a trigonometric representation of seasonal characteristics, drawing from the Fourier series. Within this model, the parameters p and q are associated with the ARMA process. The terms $m_{1,....}m_T$ specify the respective seasonal periods. The parameter k represents the number of harmonics designated for the seasonal characteristic. Additionally, ω is indicative of the Box-Cox transformation, and φ represents the dampening parameter value.

2.2.2.5 STL model

The STL model employs a locally weighted regression method to partition a time series into its constituent trend, seasonal, and remainder components. The trend component is estimated through LOESS regression. In contrast, the seasonal component is typically assessed via SARIMA or ETS models, as detailed in (38). A notable strength of the STL model is its adaptability to shifts in the series trend. Furthermore, it exhibits robust resistance to outliers present within the series. Another salient feature of the STL model is its proficiency in managing seasonal frequencies that exceed one, as highlighted in (39).

2.2.3 Model performance evaluation and forecasting

In this study, we employed a range of evaluation error metrics to assess prediction performances across models developed from both the full and training datasets. These metrics included the mean absolute error (MAE), mean absolute percent error (MAPE),

mean absolute scaled error (MASE), and root mean squared error (RMSE) as detailed in (13).

To ensure the utilization of the most recent data for forecasting, we applied all the final time series models to the full dataset. This approach facilitated the modeling and forecasting of canine rabies cases for the three subsequent years post the last observation data incorporated in this study, as described in (28).

Data organization, decomposition, and segmentation of the time series were executed using the R statistical software, leveraging the "*xts*" and "*TSstudio*" packages. The "*forecast*" and "*forecastHybrid*" packages (13) provided a suite of functions instrumental in the development of the time series model and the subsequent evaluation of its predictive performance. These functions were also pivotal in forecasting the canine rabies cases for the upcoming 3 years. Specifically, the functions *auto.arima*(), *nnetar*(), *ets*(), *tbats*(), and *stl*() were employed to facilitate the SARIMA, NNAR, ETS, TBATS, and STL models, respectively. A comprehensive description of these functions can be found in the package manual (40). For graphical representation, we utilized packages such as "*ggplots*", "*plotly*", "*scales*", and "*ggsci*".

3 Results

3.1 Descriptive and time series components

A total of 4,678 reported cases of canine rabies were reported, affecting 1,908 females and 2,770 males. Among these cases, 2,405 dogs were linked to individual or public owners. The remaining cases pertained to either stray dogs or dogs without a traceable owner history. From a geographical perspective, positive rabies cases were observed in 64 out of 78 provinces. Notably, Chon Buri reported the highest number of rabies cases, totaling 463, followed by Songkhla with 347 cases, Surin with 278 cases, Roi-Et with 211 cases, and Bangkok with 203 cases. A map depicting these provinces is available in the Supplementary Figure S1.

Figure 2 presents the confirmed positive and negative canine rabies cases as documented by government laboratories. Over the study period, a total of 4,678 canine rabies cases were recorded. Between 2013 and 2017, there was a consistent rise in the number of cases, ending in a peak in 2018. Subsequently, a decline was observed from 2020 to 2021, with a resurgence in 2022.

Figure 3A delineates the annual and monthly distribution of canine rabies cases. Figure 3B depicts the monthly number of rabies cases, including the mean and standard error values (mean and error bar). On a monthly average, 39 dogs were identified as rabies-positive. The peak of canine rabies cases was recorded in March 2018, closely followed by March 2017. Notably, the period from February to April consistently registered the highest number of cases each year (Figure 3B). Yet, certain years also witnessed spikes in June, October, November, and December.

To effectively visualize the raw time series data, which comprises multiple components, data decomposition becomes imperative. Figure 4 decomposes the time series data of canine rabies cases into its trend, seasonality, and residual (or error) components. The trend analysis reveals a growth trajectory from January 2013 to December 2017, succeeded by a decline from

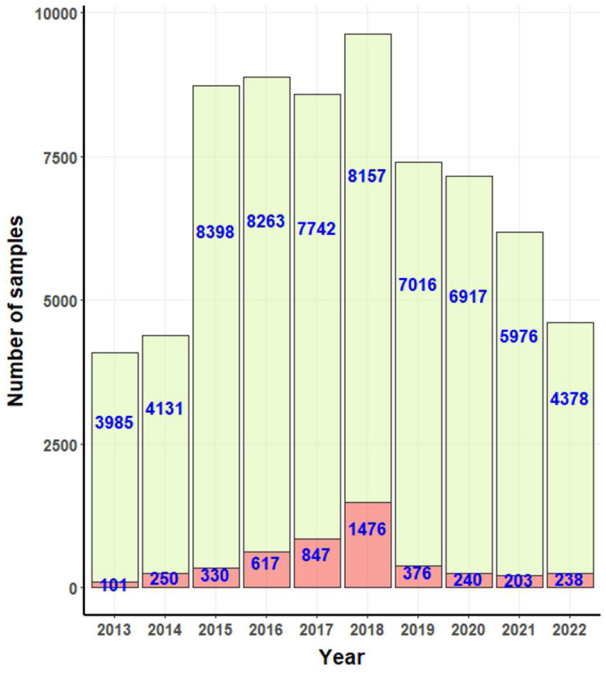

FIGURE 2
Bar graphs for number of positive (light red bar) and negative (light green bar) rabies cases confirmed by government laboratories.

August 2018 to January 2020. Post this period, the trend stabilizes. The seasonality inherent in the dataset is further confirmed by the WO-test, consistent with the ETS result. The model, with the form ETS (M, A, M), indicates the presence of a seasonal component in the data.

Figure 3B portrays the mean and standard deviations of canine rabies cases spanning 2013 to 2022. A pronounced surge in cases is evident during January and April relative to other months. It is pertinent to note that the heightened cases in January and April can be linked to the pronounced incidence of canine rabies in specific years, such as 2018.

3.2 Prediction performance

Figure 5 provide visual depictions of the observed vs. predicted canine rabies cases, derived from the testing and training datasets, respectively. A comparative analysis of the TBATS model against other models reveals a notable alignment of its forecasted values with the actual data. While the predictions from the SARIMA and NNAR models remain stable with slight variations, the ETS and STL models consistently project an upward trajectory. Notably, the STL model's forecasts display a mix of increasing and decreasing trends, with certain predictions deviating significantly from the actual values, as illustrated in Figure 5.

Table 1 presents the error metrices for the time series models applied to the dataset. Among these, the NNAR, THETA, and SARIMA models exhibited commendable performance on the training dataset. The NNAR model, in particular, showcased

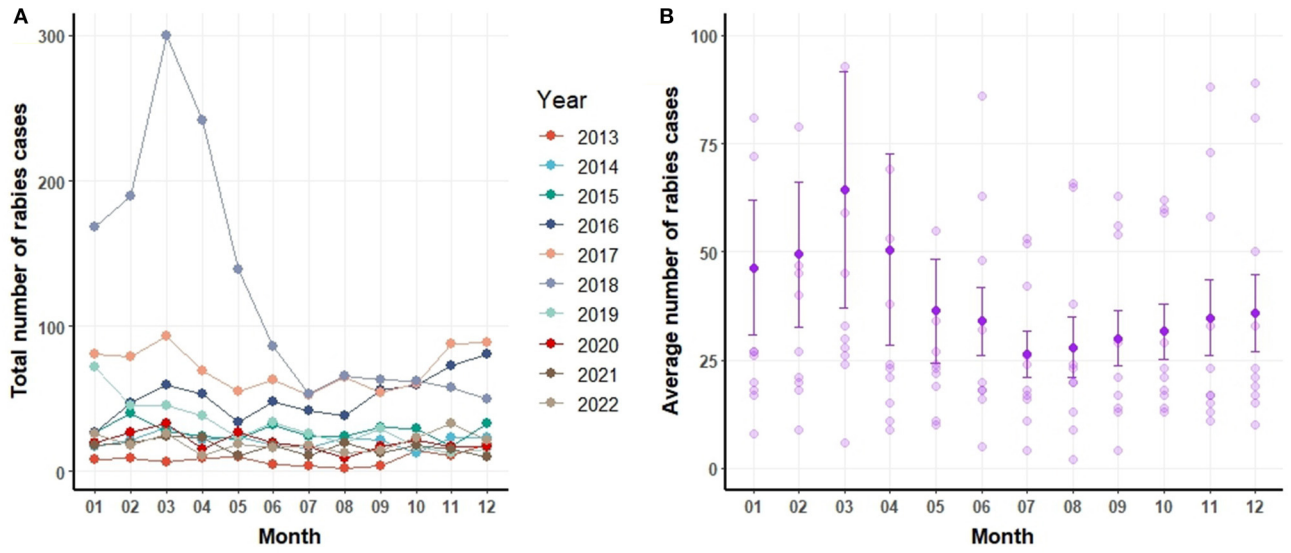

FIGURE 3
Number of canine rabies cases by year and month **(A)** and mean ± standard error of the mean for canine rabies cases by month based on data from 2013 to 2022 **(B)**.

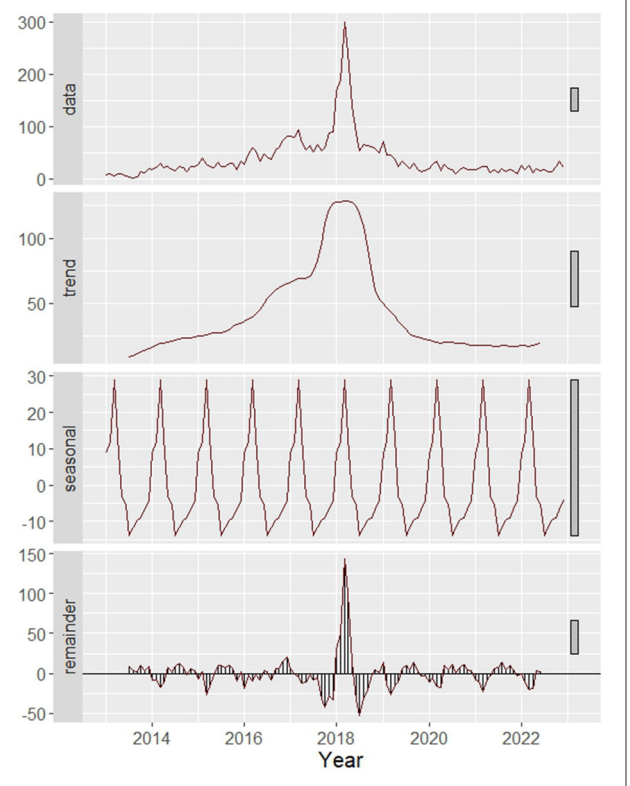

FIGURE 4
Decomposition of data on time series rabies into trend, seasonality, and remainder.

superior performance on the test data. However, it displayed signs of overfitting, as evidenced by its strong performance on the training data but suboptimal results on the test data. In the test dataset, it was observed that the error metrics values for the TBATS model were comparatively lower than those of other models. This suggests that the TBATS model exhibited superior performance in comparison to the other models.

3.3 Forecasting

Figure 6 illustrates the forecasted number of canine rabies cases for the years 2023 to 2025, as predicted by TBATS and SARIMA models applied to the full dataset. The forecasts from all models are further detailed in Supplementary Table S1 and Supplementary Figure S2. The aggregated predictions across all models estimate the average canine rabies cases to be 278, 288, and 300 for the years 2023, 2024, and 2025, respectively.

The TBATS model, which exhibited the highest predictive accuracy, estimates an annual average of ∼285 canine rabies cases for the period 2023 to 2025. This translates to a monthly average of 23 cases, with a range of 18–30 cases. In contrast, the SARIMA model, the subsequent top performer, projects a gradual increase in cases: 207 in 2023, 212 in 2024, and 213 in 2025 (Figure 6). The NNAR model's forecasts suggest consistent values across months, while the STL model's predictions indicate variability, with peaks in February and March and troughs in July and September each year (Supplementary Figure S2).

4 Discussion

For an effective rabies intervention program and efficient resource allocation, it is crucial to have a deep understanding and precise forecasting of rabies trends. The primary objective of this study was to analyze and forecast the trends of canine rabies cases

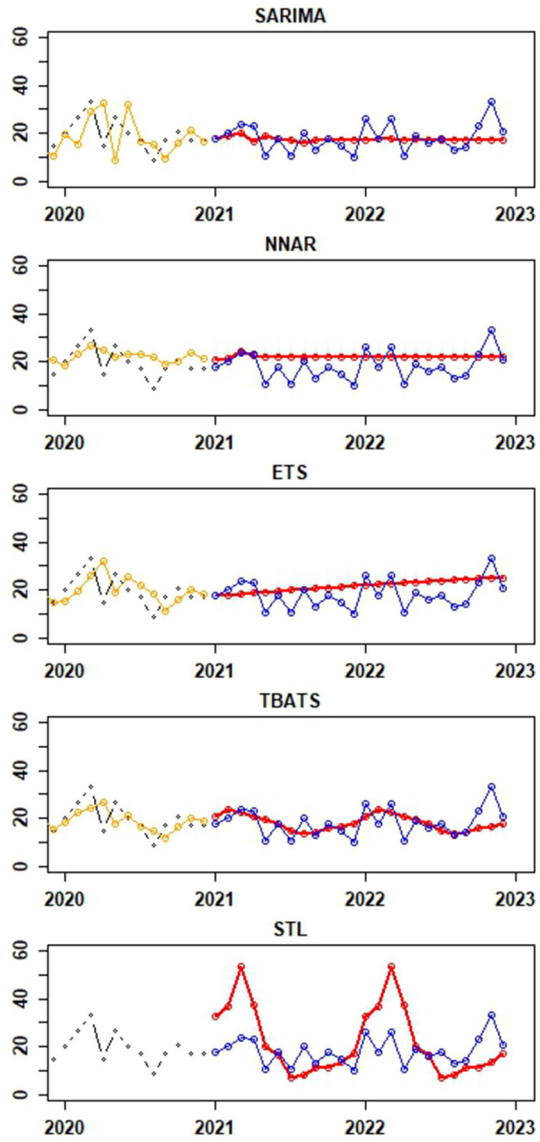

FIGURE 5
Depictions of the actual values of canine rabies cases (blue dots) from the test dataset (January to December 2022) and forecast values for the period of January to December 2022 (red dots) generated from time series forecast models applied to the training dataset (January 2013 to December 2021). The black and orange dots represent actual and fitted values from the last 24 observations of the training dataset, respectively. Notably, the fitted values from the STLM are not available for presentation.

in Thailand. Drawing on a decade's worth of national data, various time series methodologies were employed to achieve this.

Historically, there was a noticeable rise in canine rabies cases from 2013 to 2017, reaching its peak in 2018. However, a subsequent decline was observed until early 2020. This decrease is potentially attributable to the concerted efforts of the government to curtail human rabies cases. In fact, the last decade has witnessed an increase in rabies surveillance, leading to a higher number of reported canine rabies cases (41–43).

The post-2018 decline might be proof of the collaborative endeavors of both the public and private sectors. These strategies encompassed a range of measures, from enhanced surveillance and community engagement to robust public relations and the dissemination of knowledge. Approach to rabies control includes (i) setting up animal rabies monitoring, (ii) overseeing animal shelters, (iii) strengthening human rabies surveillance, (iv) assisting local communities in rabies prevention, (v) prioritizing public awareness, (vi) facilitating data sharing on rabies, (vii) concentrating on monitoring rabies cases, and (viii) disseminating knowledge and technological advancements (9).

Furthermore, the data also revealed apparent seasonal patterns in rabies occurrences, with pronounced spikes in March, June, and December. Such findings emphasize the importance of sustained rabies awareness throughout the year, rather than confining it to specific seasons.

In the aspect of prediction, the TBATS model stood out as the top performer. One of the primary reasons for its superior performance could be its proficiency in managing the complexities associated with seasonal variations and the non-linear nature of the data (19, 37, 44). This model is specifically designed to address such complexities, making it particularly suited for the task. Similarly, the SARIMA model showcased impressive results. Like the TBATS, SARIMA is equipped to handle datasets that have strong seasonal patterns, which is evident from its performance (45–47). Both these models, TBATS and SARIMA, have proven their capability in dealing with data that has a high degree of seasonal variation, making them invaluable tools in this context. On the other hand, the NNAR model, despite its merits, faced challenges when it came to managing data with intricate seasonal patterns. While it is a robust model in many scenarios, its limitations became evident in the face of complex seasonality, which affected its forecasting accuracy (13, 32).

The implications of this study are diverse. It explains how diverse forecasting models can be instrumental in shaping policy decisions. These forecasts can act as pivotal reference points, enabling authorities to set aspirational targets for the future. For instance, a tangible objective might be to ensure that the actual number of rabies cases remains below the forecasted figures. Furthermore, forecasting techniques should be integrated into existing rabies surveillance systems to assure that projections are based on the most recent data available.

However, it is imperative to acknowledge certain limitations. The data predominantly came from passive surveillance, which hinges on samples dispatched to labs either by individuals or pertinent authorities. This method might inadvertently lead to underreporting, given that some potential rabies cases might remain untested (6). However, such underreporting is a common challenge in epidemiological data (48–51). To address this issue, it is essential to intensify public awareness campaigns to encourage more extensive testing of suspected cases. Furthermore, the models deployed were somewhat myopic, focusing solely on the number of rabies cases and the corresponding months, without exploring into other potential determinants of rabies trends.

In this study, we employed time series models to analyze data available up to 2022. Should more recent data, such as from 2023 onward, become accessible, these models can be adapted accordingly. Future research should consider the development and evaluation of hybrid time series models (32, 52), as well as other time series models (23, 53), in order to determine their potential

TABLE 1 Error metrics for time series models applied to training data (January 2013 to December 2021) to forecast canine rabies cases for January to December 2022, compared with actual values present in the test dataset (January to December 2022).

Method	Training set				Testing set			
	MAE	MAPE	MASE	RMSE	MAE	MAPE	MASE	RMSE
SARIMA	11.91	30.69	0.36	21.01	4.27	25.32	0.13	5.60
NNAR	3.86	14.69	0.12	5.04	5.78	40.64	0.17	7.03
ETS	12.41	34.97	0.37	23.08	14.52	94.74	0.44	16.95
TBATS	10.56	25.69	0.32	20.61	4.15	24.87	0.12	5.58
STL	12.00	40.25	0.36	19.46	10.41	57.46	0.31	13.51

SARIMA, seasonal autoregressive integrated moving average; NNAR, neural network autoregression; ETS, error trend seasonality; TBATS, trigonometric exponential smoothing state-space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components; STL, seasonal and trend decomposition using LOESS.

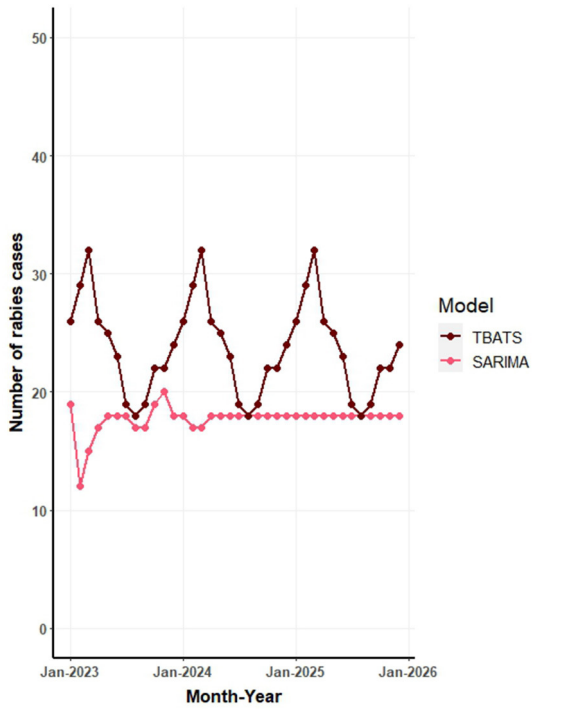

FIGURE 6
Forecast of monthly canine rabies cases in Thailand from 2023 to 2025 based on seasonal autoregressive integrated moving average (SARIMA) and trigonometric exponential smoothing state-space model with Box-Cox transformation (TBATS) models.

for enhancing the accuracy of canine rabies cases predictions. Additionally, while the primary focus of this study is on canine rabies, a comprehensive examination of the relationship between human and canine rabies cases at a national scale is recommended for subsequent investigations.

5 Conclusion

In a pioneering effort, this study utilized time series modeling to analyze and predict canine rabies cases based on national data. Observations revealed a decline in rabies cases from 2017 to 2019, stabilizing between 2020 and 2021. Of the models tested, the TBATS emerged as the most accurate predictor. Forecasts suggest that rabies cases will likely remain consistent in the near future, underscoring the need for intensified efforts to reduce these numbers. These predictions can guide authorities in strategic planning and resource allocation for rabies prevention and control. Ultimately, the methodologies showcased here offer a valuable tool for anticipating canine rabies trends in the years ahead.

Data availability statement

The datasets presented in this article are not readily available because the data used in this study can be obtained from the Department of Livestock Development (DLD) in Thailand. An official letter must be sent to foreign@dld.go.th in order to request access. The DLD has granted approval for the use of this data in our research, as indicated by Approval number: 1601(01)560. Requests to access the datasets should be directed to foreign@dld.go.th.

Author contributions

VP: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review & editing. WT: Conceptualization, Project administration, Resources, Supervision, Validation, Writing—review & editing. CJ: Data curation, Visualization, Writing—review & editing. PC: Data curation, Investigation, Resources, Validation, Visualization, Writing—review & editing. OS: Conceptualization, Data curation, Investigation, Resources, Visualization, Writing—review & editing. RS: Data curation, Writing—review & editing. OA: Conceptualization, Data curation, Formal analysis, Investigation, Resources, Validation, Visualization, Writing—review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The authors are grateful for research funding from Chiang Mai University (Grant: R66IN00356).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fvets.2023.1294049/full#supplementary-material

References

1. Riccardi N, Giacomelli A, Antonello RM, Gobbi F, Angheben A. Rabies in Europe: an epidemiological and clinical update. *Eur J Intern Med.* (2021) 88:15–20. doi: 10.1016/j.ejim.2021.04.010

2. Singh R, Singh KP, Cherian S, Saminathan M, Kapoor S, Manjunatha Reddy G, et al. Rabies–epidemiology, pathogenesis, public health concerns and advances in diagnosis and control: a comprehensive review. *Vet Q.* (2017) 37:212–51. doi: 10.1080/01652176.2017.1343516

3. Bourhy H, Dautry-Varsat A, Hotez PJ, Salomon J. Rabies, still neglected after 125 years of vaccination. *PLoS Negl Trop Dis.* (2010) 4:e839. doi: 10.1371/journal.pntd.0000839

4. Ghosh S, Rana M, Islam M, Chowdhury S, Haider N, Kafi MAH, et al. Trends and clinico-epidemiological features of human rabies cases in Bangladesh 2006–2018. *Sci Rep.* (2020) 10:1–11. doi: 10.1038/s41598-020-59109-w

5. World Health Organization, Food and Agriculture Organization of the United Nations, World Organisation for Animal Health. *Zero by 30: The Global Strategic Plan to End Human Deaths From Dog-Mediated Rabies by 2030: United Against Rabies Collaboration: First Annual Progress Report: Global Strategic Plan to End Human Deaths From Dog-Mediated Rabies by 2030*. Geneva: World Health Organization (2019). Available online at: https://apps.who.int/iris/handle/10665/328053 (accessed January 12, 2023).

6. Kanankege KST, Errecaborde KM, Wiratsudakul A, Wongnak P, Yoopatthanawong C, Thanapongtharm W, et al. Identifying high-risk areas for dog-mediated rabies using Bayesian spatial regression. *One Health.* (2022) 15:100411. doi: 10.1016/j.onehlt.2022.100411

7. Mitmoonpitak C, Tepsumethanon V, Wilde H. Rabies in Thailand. *Epidemiol Infect.* (1998) 120:165–9. doi: 10.1017/S0950268897008601

8. Leelahapongsathon K, Kasemsuwan S, Pinyopummintr T, Boodde O, Phawaphutayanchai P, Aiyara N, et al. Humoral immune response of Thai dogs after oral vaccination against rabies with the SPBN GASGAS vaccine strain. *Vaccines.* (2020) 8:573. doi: 10.3390/vaccines8040573

9. Thanapongtharm W, Suwanpakdee S, Chumkaeo A, Gilbert M, Wiratsudakul A. Current characteristics of animal rabies cases in Thailand and relevant risk factors identified by a spatial modeling approach. *PLoS Negl Trop Dis.* (2021) 15:e0009980. doi: 10.1371/journal.pntd.0009980

10. Premashthira S, Suwanpakdee S, Thanapongtharm W, Sagarasaeranee O, Thichumpa W, Sararat C, et al. The impact of socioeconomic factors on knowledge, attitudes, and practices of dog owners on dog rabies control in Thailand. *Front Vet Sci.* (2021) 8:699352. doi: 10.3389/fvets.2021.699352

11. Kahn S, Stuardo L, Rahman S. OIE guidelines on dog population control. *Dev Biol.* (2008) 131:511–6.

12. Kawaya EK, Marcotty T, Mfumu-Kazadi LM, Van Gucht S, Kirschvink N. Factors of maintenance of rabies transmission in dogs in Kinshasa, Democratic Republic of the Congo. *Prev Vet Med.* (2020) 176:104928. doi: 10.1016/j.prevetmed.2020.104928

13. Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*, 2nd ed. Melbourne, VIC: OTexts (2018).

14. Ward MP, Iglesias RM, Brookes VJ. Autoregressive models applied to time-series data in veterinary science. *Front Vet Sci.* (2020) 7:604. doi: 10.3389/fvets.2020.00604

15. Abotaleb M, Makarovskikh T. System for forecasting COVID-19 cases using time-series and neural networks models. *Eng Proc.* (2021) 5:46. doi: 10.3390/engproc2021005046

16. Khan FM, Gupta R, ARIMA. and NAR based prediction model for time series analysis of COVID-19 cases in India. *J Saf Sci Resil.* (2020) 1:12–8. doi: 10.1016/j.jnlssr.2020.06.007

17. Mavragani A, Gkillas K. COVID-19 predictability in the United States using Google Trends time series. *Sci Rep.* (2020) 10:1–12. doi: 10.1038/s41598-020-77275-9

18. Talkhi N, Fatemi NA, Ataei Z, Nooghabi MJ. Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: a comparison of time series forecasting methods. *Biomed Signal Process Control.* (2021) 66:102494. doi: 10.1016/j.bspc.2021.102494

19. Yu C, Xu C, Li Y, Yao S, Bai Y, Li J, et al. 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. doi: 10.2147/IDR.S304652

20. Gandhi N, Armstrong LJ. A review of the application of data mining techniques for decision making in agriculture. In: *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*. Greater Noida (2016). p. 1–6. doi: 10.1109/IC3I.2016.7917925

21. Pope III CA, Schwartz J. Time series for the analysis of pulmonary health data. *Am J Respir Crit Care Med.* (1996) 154:S229. doi: 10.1164/ajrccm/154.6_Pt_2.S229

22. He J, Wei X, Yin W, Wang Y, Qian Q, Sun H, et al. Forecasting scrub typhus cases in eight high-risk counties in China: evaluation of time-series model performance. *Front Environ Sci.* (2022) 9:783864. doi: 10.3389/fenvs.2021.783864

23. Punyapornwithaya V, Arjkumpa O, Buamithup N, Kuatako N, Klaharn K, Sansamur C, et al. Forecasting of daily new lumpy skin disease cases in Thailand at different stages of the epidemic using fuzzy logic time series, NNAR, and ARIMA methods. *Prev Vet Med.* (2023) 217:105964. doi: 10.1016/j.prevetmed.2023.105964

24. Ortega-Sánchez R, Bárcenas-Reyes I, Cantó-Alarcón GJ, Luna-Cozar JER-A, Contreras-Magallanes YG, González-Ruiz S, et al. Descriptive and time-series analysis of rabies in different animal species in Mexico. *Front Vet Sci.* (2022) 9:800735. doi: 10.3389/fvets.2022.800735

25. Blanc E, Reilly J. Approaches to assessing climate change impacts on agriculture: an overview of the debate. *Rev Environ Econ Policy.* (2020) 11:247–57.

26. Ding Y, Wang L, Li Y, Li D. Model predictive control and its application in agriculture: a review. *Comput Electron Agric.* (2018) 151:104–17. doi: 10.1016/j.compag.2018.06.004

27. Mishra P, Al Khatib AMG, Lal P, Anwar A, Nganvongpanit K, Abotaleb M, et al. An overview of pulses production in India: retrospect and prospects of the future food with an application of hybrid models. *Natl Acad Sci Lett.* (2023) 46:367–74. doi: 10.1007/s40009-023-01267-2

28. Punyapornwithaya V, Jampachaisri K, Klaharn K, Sansamur C. Forecasting of milk production in northern Thailand using seasonal autoregressive integrated moving average, error trend seasonality, and hybrid models. *Front Vet Sci.* (2021) 8:775114. doi: 10.3389/fvets.2021.775114

29. Molefi M, Tlhakanelo JT, Phologolo T, Hamda SG, Masupe T, Tsima B, et al. The impact of China's lockdown policy on the incidence of COVID-19: an interrupted time series analysis. *Biomed Res Int.* (2021) 2021:9498029. doi: 10.1155/2021/9498029

30. Iwata K, Doi A, Miyakoshi C. Was school closure effective in mitigating coronavirus disease 2019 (COVID-19)? Time series analysis using Bayesian inference. *Int J Infect Dis.* (2020) 99:57–61. doi: 10.1016/j.ijid.2020.07.052

31. Siedner MJ, Harling G, Reynolds Z, Gilbert RF, Haneuse S, Venkataramani AS, et al. Social distancing to slow the US COVID-19 epidemic: longitudinal pretest-posttest comparison group study. *PLoS Med.* (2020) 17:e1003244. doi: 10.1371/journal.pmed.1003244

32. Punyapornwithaya V, Mishra P, Sansamur C, Pfeiffer D, Arjkumpa O, Prakotcheo R, et al. Time-series analysis for the number of foot and mouth disease outbreak

episodes in cattle farms in Thailand using data from 2010-2020. *Viruses.* (2022) 14:1367. doi: 10.3390/v14071367

33. Ollech D. *Package 'seastests'*. (2022). Available online at: https://cran.r-project.org/web/packages/seastests/seastests.pdf (accessed July 3, 2023).

34. Liao Z, Zhang X, Zhang Y, Peng D. Seasonality and trend forecasting of tuberculosis incidence in Chongqing, China. *Interdiscip Sci.* (2019) 11:77–85. doi: 10.1007/s12539-019-00318-x

35. Zhai M, Li W, Tie P, Wang X, Xie T, Ren H, et al. Research on the predictive effect of a combined model of ARIMA and neural networks on human brucellosis in Shanxi province, China: a time series predictive analysis. *BMC Infect Dis.* (2021) 21:1–12. doi: 10.1186/s12879-021-05973-4

36. Hyndman R, Koehler AB, Ord JK, Snyder RD. *Forecasting With Exponential Smoothing: The State Space Approach*. Berlin: Springer (2008).

37. De Livera AM, Hyndman RJ, Snyder RD. Forecasting time series with complex seasonal patterns using exponential smoothing. *J Am Stat Assoc.* (2011) 106:1513–27. doi: 10.1198/jasa.2011.tm09771

38. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: a seasonal-trend decomposition procedure based on Loess. *J Off Stat.* (1990) 6:3–73.

39. Theodosiou M. Forecasting monthly and quarterly time series using STL decomposition. *Int J Forecast.* (2011) 27:1178–95. doi: 10.1016/j.ijforecast.2010.11.002

40. Hyndman RJ, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, et al. *Package 'forecast'*. (2020). Available online at: https://cran.r-project.org/web/packages/forecast/forecast.pdf (accessed July 3, 2023).

41. Komol P, Sommanosak S, Jaroensrisuwat P, Wiratsudakul A, Leelahapongsathon K. The spread of rabies among dogs in Pranburi district, Thailand: a metapopulation modeling approach. *Front Vet Sci.* (2020) 7:570504. doi: 10.3389/fvets.2020.570504

42. Department of Disease Control. *Annual Epidemiological Surveillance Report 2017*. Bangkok: Ministry of Public Health (2018).

43. Department of Disease Control. *Annual Epidemiological Surveillance Report 2018*. Bangkok: Ministry of Public Health (2019).

44. Naim I, Mahara T, Idrisi AR. Effective short-term forecasting for daily time series with complex seasonal patterns. *Procedia Comput Sci.* (2018) 132:1832–41. doi: 10.1016/j.procs.2018.05.136

45. Assimakopoulos V, Nikolopoulos K. The theta model: a decomposition approach to forecasting. *Int J Forecast.* (2000) 16:521–30. doi: 10.1016/S0169-2070(00)00066-2

46. Ebhuoma O, Gebreslasie M, Magubane L. A seasonal autoregressive integrated moving average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa. *S Afr Med J.* (2018) 108:573–8. doi: 10.7196/SAMJ.2018.v108i7.12885

47. Mao Q, Zhang K, Yan W, Cheng C. Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *J Infect Public Health.* (2018) 11:707–12. doi: 10.1016/j.jiph.2018.04.009

48. Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, et al. Estimating the global burden of endemic canine rabies. *PLoS Negl Trop Dis.* (2015) 9:e0003709. doi: 10.1371/journal.pntd.0003709

49. Dürr S, Naïssengar S, Mindekem R, Diguimbye C, Niezgoda M, Kuzmin I, et al. Rabies diagnosis for developing countries. *PLoS Negl Trop Dis.* (2008) 2:e206. doi: 10.1371/journal.pntd.0000206

50. 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.* (2021) 53:1–9. doi: 10.1007/s11250-020-02452-x

51. Traoré A, Keita Z, Léchenne M, Mauti S, Hattendorf J, Zinsstag J. Rabies surveillance-response in Mali in the past 18 years and requirements for the future. *Acta Trop.* (2020) 210:105526. doi: 10.1016/j.actatropica.2020.105526

52. 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) 357:1–24. doi: 10.2139/ssrn.3716343

53. Nason GP, Sachs Rv. Wavelets in time-series analysis. *Philos Trans Royal Soc A.* (1999) 357:2511–26. doi: 10.1098/rsta.1999.0445