Iran J Public Health, Vol. 52, No. 10, Oct 2023, pp.2063-2072

Original Article

Research of Combined ES-BP Model in Predicting Syphilis Incidence 1982-2020 in Mainland China

***Daren Zhao**

Department of Medical Administration, Sichuan Provincial Orthopedics Hospital, Chengdu, Sichuan, P.R. China

***Correspondence:** Email: cdzhaodaren@yeah.net

(Received 12 Sep 2022; accepted 18 Dec 2022)

Abstract

Background: Syphilis remains a major public health concern in China. We aimed to construct an optimum model to forecast syphilis epidemic trends and provide effective precautionary measures for prevention and control.

Methods: Data on the incidence of syphilis between 1982 and 2020 were obtained from the China Health Statistics Yearbook. An exponential smoothing model (ES model) and a BP neural network model were constructed, and on this basis, the ES-BP combination model was created. The prediction performance was assessed to compare the MAE (Mean Absolute Error), MSE (Mean Squared Error), MAPE (Mean Absolute Percentage Error), and RMSE (Root Mean Square Error).

Results: The optimum ES model was Brown's linear trend model, which had the lowest MAE and MAPE values, and its residual was a white noise sequence (P=0.359). The optimum BP neural network model had three layers with the number of nodes in the input, hidden, and output layers set to 5, 11, and 1, and the mean values of MAE, MSE, and RMSE by five-fold cross-validation were 1.519, 6.894, and 1.969, respectively. The ES-BP combination model had three layers, with model nodes 1, 4, and 1. The lowest mean values of MAE, MSE, and RMSE obtained by five-fold cross-validation were 1.265, 5.739, and 2.105, respectively.

Conclusion: The ES, BP neural network, and ES-BP combination models can be used to predict syphilis incidence, but the prediction performance of the ES-BP combination model is better than that of a basic ES model and a basic BP neural network model.

Keywords: Syphilis; Exponential smoothing; BP neural network; Incidence; China

Introduction

Syphilis is a common sexually transmitted disease that occurs when syphilis spirochetes infect the body (1,2). The incidence of syphilis is estimated to be 10.6 million each year all over the world (3). Syphilis is a global public health issue, especially among men who have sex with men (MSM) in high- and middle-income countries (4,5). Syphilis also leads to hundreds of thousands of stillbirths and neonatal deaths annually in developing countries (4).

Over the past two decades, the incidence of syphilis has increased in China (6). Under the Law of the People's Republic of China on the Prevention and Treatment of Infectious Diseases, syphilis is classified as a Class B infectious disease, and its incidence has always been among the

Copyright © 2023 Zhao. Published by Tehran University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license.
(https://creativecommons.org/licenses/by-nc/4.0/). Non-commercial uses of the work are permitted, provided the original work is properly cited

highest in the list of Class B infectious diseases with a high disease burden (7). Therefore, strengthening the prevention and control of syphilis infection is an urgent and important task for public health in China.

To reduce the incidence of syphilis and the economic burden of this infectious disease, it is important to strengthen its surveillance and forecasting. Time series prediction and machine learning models have been used to predict syphilis incidence. For time-series models, Zhang et al. (6) used the ARIMA and ARIMAX models to predict the incidence of syphilis in mainland China from 2005 to 2012. Guo et al. (8) used the conventional GM (1,1) model and GM (1,1) model with the self-memory principle (SMGM (1,1) model) to forecast the incidence of syphilis in mainland China. Li et al. (9) explored the feasibility of using the grey GM (1,1) model to predict the incidence of syphilis. Ibáñez-Cervantes et al. (10) utilized the seasonal autoregressive integrated moving average (ARIMA) model to forecast syphilis cases in Mexico. For machine learning models, Zhang et al. (11) used regression, exponential smoothing, the ARIMA model, and a support vector machine (SVM) to forecast nine types of infectious diseases, including syphilis, and showed that the prediction performance of the SVM model outperformed the other models.

However, no studies have focused on combination models to predict the worldwide incidence of syphilis. This study aimed to explore a new combination model for predicting the incidence of syphilis in mainland China from 1982 to 2020. An exponential smoothing (ES) model and a BP neural network model were created, and an ES-BP combination model was developed. This is the first study to construct an ES-BP combination model for the prediction of syphilis incidence in mainland China. We hope that the study methodology will provide new ideas for predicting the incidence of syphilis, and that the results will serve as effective guidance for the prevention and control of its epidemics.

Methods

Data Source

Data on syphilis incidence from 1982 to 2020 were obtained from the China Health Statistics Yearbook. In China, syphilis is a Class B infectious disease. Once a patient has been clinically diagnosed and laboratory-tested for syphilis, the clinician at the medical institution must report it to the local Center for Disease Control and Prevention by the National Infectious Surveillance System within 24 hours. The National Health Commission of the People's Republic of China publishes the China Health Statistics Yearbook annually, which includes data on syphilis morbidity and mortality for the previous year in Chapter 10, Diseases and Public Health.

In this study, 39 data points on syphilis incidence in mainland China from 1982 to 2020 were collected, of which 33 data points from 1982 to 2014 were used as the training set to construct the model and six data points from 2015 to 2020 were used as the test set to evaluate the generalization ability of the model.

Ethics statement

Data on syphilis incidence in mainland China from 1982 to 2020 are publicly available and can be obtained from the China Health Statistics Yearbook. Meanwhile, this infectious disease has annual data and excludes patient information; therefore, no ethical review was required for this study.

Exponential smoothing model

The basic idea of exponential smoothing is that forecasts are weighted by historical observations, and it assigns different weights to different historical data, with new data assigned a larger weight and old data assigned a smaller weight (12). One-time exponential smoothing was suitable for a series with no trend and no seasonality; the quadratic exponential smoothing method is suitable for a series with a trend but no seasonality; and the cubic exponential smoothing method is suitable for a series with both trend and seasonality (13).

In our study, data on the incidence of syphilis were annual data, showing a trend but no seasonality; therefore, we chose a quadratic exponential smoothing model for the prediction.

The equation is as follows (13):

$$S_t^{(2)} = \alpha S_t^{(1)} + (1-\alpha) S_{t-1}^{(2)} \quad [1]$$

where α is the smoothing coefficient, $S_t^{(1)}$ is the one-time exponential smoothing value at time t, $S_t^{(2)}$ is the quadratic exponential smoothing value at time t, and $S_{t-1}^{(2)}$ is the cubic exponential smoothing value at time t-1. The optimum model was evaluated as the highest value of R-squared, the smallest values of Bayesian information criterion (BIC), mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and white noise series for the Lung-box residual test (14-16).

BP neural network

A BP neural network is a multilayer feedforward network (17). The basic idea of a BP neural network is that it uses gradient descent to train the range of weights of the neural network by backpropagation and finally minimizes the global error coefficient (17). The BP neural network comprises an input layer, hidden layer, and output layer, with interconnected neurons between layers and no connections between the same layers (18). The construction of the BP neural network includes the following five steps (19): ① setting the network structure, ② normalization of the data, ③ calculation and back propagation of the error, ④ parameter setting of the learning algorithm, and ⑤ output of the prediction results after training the model and inverse normalization.

The number of nodes in the input layer is *n*, the number of nodes in the hidden layer is ***l***, and the number of nodes in the output layer is *m*. The output of the hidden layer (H_j) was calculated using Equation (19).

$$H_j = g\left(\sum_{i=1}^{n} w_{ij} x_i - a_j\right) \quad [2]$$

$$g(x) = \frac{1}{1+e^{-x}} \quad [3]$$

W_{ij} is the connection weight between the input layer node i and hidden layer node j, a_j is the threshold of the hidden node j, and g(x) is the activation function.

ES-BP combination model

First, data from 1982 to 2014 were used as the training set to construct the ES model, and predictive values were obtained. Second, the predictive values from the ES model were used as input variables, and the observed values were used as output variables to construct the BP neural network model. Third, the predictive values from the ES-BP combination model were used to compare and evaluate the predictive performance.

Model evaluation

Indices, such as MAE, MSE, MAPE, and RMSE, are commonly used to evaluate the fitting performance of these models (20). The smaller the error, the better is the fitting performance. The formula can be expressed as:

$$MAE = \frac{\sum_{t=1}^{n} \left| X_t - \hat{X}_t \right|}{n} \quad [4]$$

$$MSE = \frac{1}{n}\sqrt{\sum_{t=1}^{n} (X_t - \hat{X}_t)^2} \quad [5]$$

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{X_t - \hat{X}_t}{X_t} \right| \times 100\%}{n} \quad [6]$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (X_t - \hat{X}_t)^2}{n}} \quad [7]$$

Where, $\hat{X}_t$ is the predicted value, X_t is the observed value, and *n* is the sequence sample size.

Statistical analysis

Excel 2010 was used to collect the data, and SPSS version 23.0 (IBM Corp., Armonk, NY, USA) was used to construct the ES model. The neural network toolbox of MATLAB 2020a was used to construct the BP neural networks. The level of significance was set at $P<0.05$.

Results

ES model

The incidence of syphilis was annual data, showing a trend with non-seasonality; thus, quadratic exponential smoothing was chosen as the model for prediction (Fig. 1). The non-seasonality model type, including Simple, Holt's linear trend, Brown's linear trend, and Damped trend, was selected to build the exponential smoothing model, respectively.

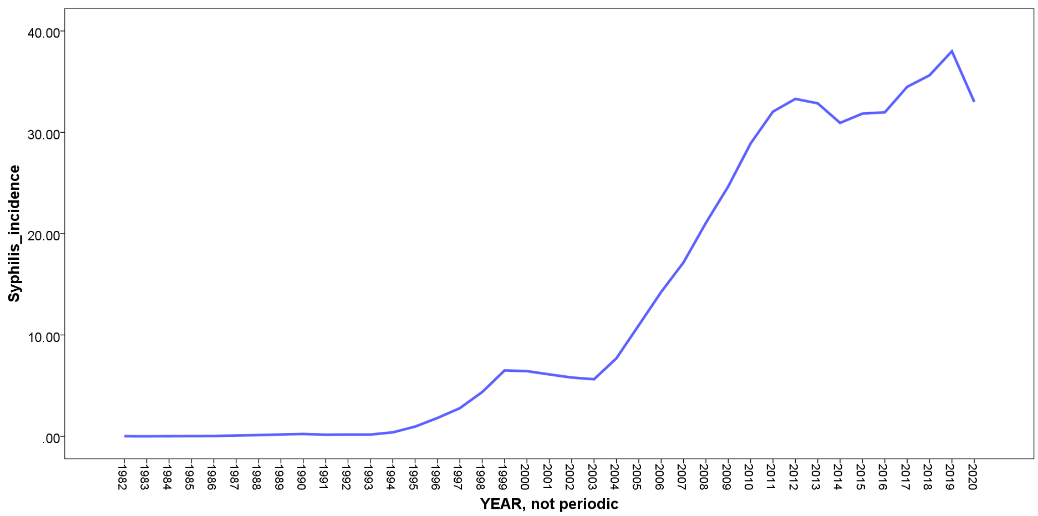

Fig.1: Time-series of syphilis incidence from 1982 to 2020 in mainland China

The optimum model was determined by the highest R-squared value and the lowest MAE, PAPE, and RMSE values as well as the residual white noise sequence. The difference between the R-squared values of the four models was not significant; however, the RMSE, MAPE, and MAE values differed slightly (Table 1). Moreover, the residuals of all four models passed the white noise test (P_{Simple}=0.077, P_{Holt}=0.281, P_{Brown}=0.359, P_{Damped}=0.360). Therefore, the optimum model was Brown's linear trend model with the lowest BIC, MAE, and MAPE values, and its residual was a white noise sequence. In addition, the model statistics showed that only Brown's linear trend model was statistically significant (t=7.592, P=0.000); the results are shown in Table 2 and Fig. 2.

Table 1: Fitting performance of the exponential smoothing model

Model	*R-squared*	*RMSE*	*MAPE*	*MAE*	*Normalized BIC*	*Ljung-Box Q*	
						Statistics	*P*
Simple	0.981	1.948	22.794	1.273	1.428	25.838	0.077
Holt's linear trend	0.988	1.576	19.115	0.832	1.097	18.775	0.281
Brown's linear trend	0.988	1.567	18.915	0.851	0.993	18.480	0.359
Damped trend	0.989	1.524	18.956	0.854	1.125	16.330	0.360

Table 2: Model Statistics of four exponential smoothing models

Models		*Estimate*	*SE*	*t*	*P*
Simple	Alpha (Level)	1.000	0.178	5.605	0.000
Holt's linear trend	Alpha (Level)	0.900	0.250	3.603	0.001
	Gamma (Trend)	1.000	0.526	1.900	0.065
Brown's linear trend	Alpha (Level)	0.895	0.118	7.592	0.000
Damped trend	Alpha (Level)	0.906	0.448	2.024	0.050
	Gamma (Trend)	0.999	1.288	0.776	0.443
	Phi (Trend damping factor)	0.762	0.232	3.279	0.002

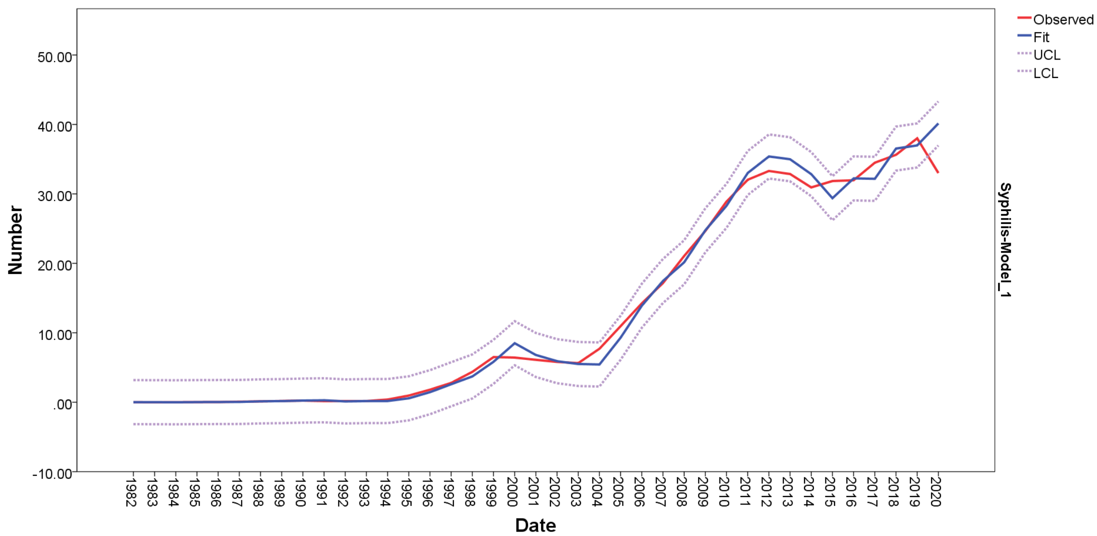

Fig.2: Comparison of actual and predicted values of Brown's linear trend model

BP neural network model

Before building the BP neural network, the data must be preprocessed. The historical values at every 5th moment (year) influence the values at the latter moment (year), the input and output values of the BP neural network were determined. Therefore, the input and output values of the BP neural network were five (assuming X_1-X_5) and one (assuming Y), respectively. Because of the data sliding method, the initial six data were missing for both the input and output variables.

We then divided 70% of the preprocessed data into a training set and 30% into a test set. The training set was used to build the model and the test set was used to evaluate the performance of the model. By trial and error, we set the number of nodes in the input, hidden, and output layers to 5, 11, and 1, respectively, while setting the number of epochs to 1000, the learning rate to 0.01, and the minimum error of the training target to 0.000001 to build the BP neural network model. Owing to the small sample size in our study, five-fold cross-validation was chosen to evaluate the predictive performance of the model. The results showed that the goodness of fit R-squared of the model was 0.9859, and the mean values of MAE, MSE, and RMSE of the five-fold cross-validation were 1.519, 6.894, and 1.969, respectively.

ES-BP combination model

First, the predictive values from Brown's linear trend model were used as input variables, and the observed values were used as output variables to construct the BP neural network model. Second,

70% of the data was divided into a training set and 30% into a test set. An optimum BP neural network model was constructed by trial and error. The number of nodes in the input, hidden, and output layers was set before starting the training.
The BP neural network model parameters, such as the number of nodes of the input, hidden, and output layers, were 1, 4, and 1, respectively. We set the epoch, learning rate, and minimum error of the training target to 1000, 0.01, and 0.000001, respectively. Subsequently, an ES-BP neural network model was constructed. Similarly, we conducted five-fold cross-validation to evaluate the predictive performance of the model. The goodness-of-fit R-squared of the ES-BP neural network model was 0.9775, and the mean values of MAE, MSE, and RMSE were 1.265, 5.739, and 2.105, respectively. Finally, we compared the observed value with the predicted value fitted by the ES-BP neural network model for the training and test sets (Figs. 3-4).

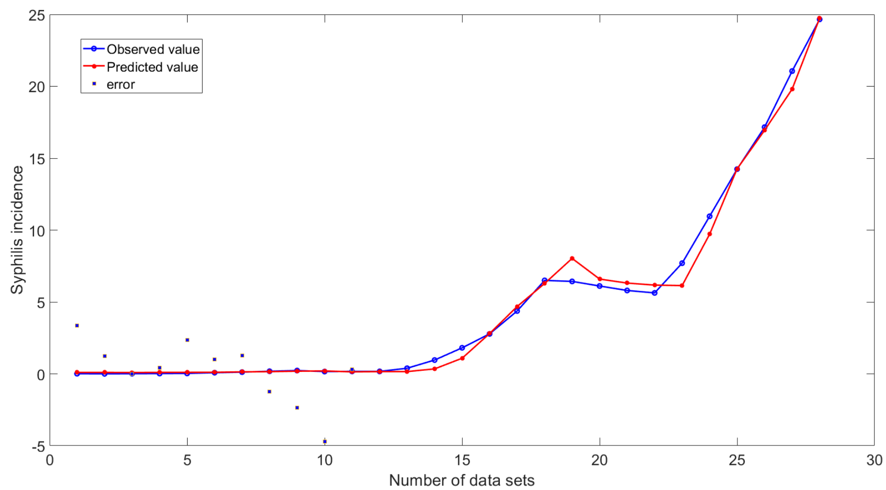

Fig.3: Comparison of predicted value and observed value of ES-BP Combination model of training set

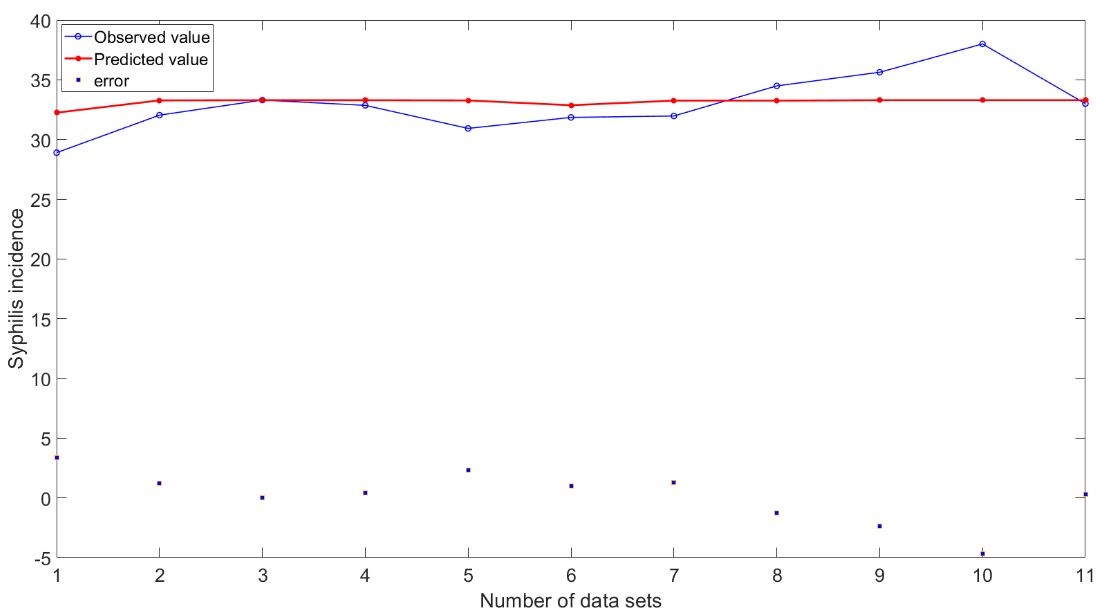

Fig.4: Comparison of predicted value and observed value of ES-BP Combination model test set

Model evaluation

Based on the prediction values of the ES, BP neural network, and ES-BP combination models, equations [4-7] were used to calculate the values of MAE, MSE, MAPE, and RMSE, respectively. Because the data sliding method was adopted in the construction of the BP neural network model, six data were lost, and 33 data were finally included in the model evaluation. The ES-BP combination model presented superior prediction performance in both the fitting and forecasting parts, with the smallest values of MAE, MSE, MAPE, and RMSE (Table 3).

Table 3: Evaluation of the prediction performance of the ES, BP neural network and ES-BP Combination models

Indices	*Fitting part*			*Forecasting part*		
	ES	BP	ES-BP	ES	BP	ES-BP
MAE	0.6804	0.7721	0.6132	2.3517	3.4280	1.8150
MSE	2.6882	3.2141	2.6671	4.0158	4.6177	2.8278
MAPE	0.1591	0.7072	0.1521	0.0702	0.0957	0.0511
RMSE	3.8017	4.5455	3.7718	5.6792	6.5304	3.9991

Discussion

Monitoring the prevalence of infectious illnesses is of great importance for prevention, control, and conventional health education (21). In this study, we collected data on syphilis incidence from 1982 to 2020 and constructed ES, BP neural network, and ES-BP combination models to forecast the prevalence trends of this infectious disease. The prediction results will not only provide timely monitoring information to the relevant authorities but also help decision-makers to know the epidemic trajectories of syphilis in mainland China. Furthermore, it can also provide a basis for the rational allocation of limited medical resources in mainland China.

Different prediction models have different characteristics and advantages; therefore, they are suitable for different scenarios. The exponential smoothing model is a special type of weighted moving average method (15), and it does not abandon previous data but gradually reduces the weight of historical information (12). Therefore, the exponential smoothing model is unsuitable for long-term forecasting. However, the exponential smoothing model has some advantages; for example, it is relatively simple to construct and has a wide range of applications (22).

The BP neural network model, a type of machine learning model (19), is particularly suitable for solving problems with complex internal mechanisms, which have strong self-learning ability, fast training speed, and good convergence ability (23-26). Therefore, these advantages make the BP neural network model attractive for infectious disease prediction applications. However, the BP neural network construction depends on the setting of the model parameters, and different parameters cause the network to converge at different speeds, leading to a prediction result having a large deviation or even no result to some extent (27). A three-layer BP network model can complete any mapping from n to m dimensions (28). The number of nodes in the input and output layers depends on the distribution and characteristics of the data (21). The number of nodes in the hidden layer depends on the complexity of the data and should not be too many or too few (21). If the number of neurons is too small, it will lead to under-fitting; conversely, too many neurons will lead to over-fitting (21).

Different models are suitable for different data characteristics and sample size (15,29). In our study, based on the distribution and characteristics of syphilis incidence data, we set up a 3-layer neural network, including an input layer, a hidden layer, and output layers. Meanwhile, the number of nodes in the input, hidden, and output layers was set by repeated attempts to avoid underfitting and overfitting. We combined the characteristics of the data and repeated the experiments to find the most appropriate number of hidden layer neurons to fit an optimal BP neural network. Moreover, five-fold cross-validation was conducted to obtain a reliable and stable model. In this study, five-fold cross-validation confirmed that the BP neural network model displayed a good prediction performance.

We found that the ES-BP combination model had the smallest MAE, MSE, MAPE, and RMSE values in both the fitting and forecasting parts, indicating that it achieved a better prediction performance, which is consistent with similar previous studies on infectious disease prediction. For example, the new SARIMA-NARNNX combined model was better than other methods, which was suitable for forecasting the long-term epidemic patterns of tuberculosis morbidity in mainland China (30). The prediction performance of the ARIMA-NARNN hybrid model was more effective than other models, and the predictions can improve our understanding of the epidemic characteristics of hemorrhagic fever with renal syndrome (HFRS) (31). In addition, Jia et al. (27) used a BP neural network, GM (1,1) grey model, and the combination models to forecast mumps incidence and found that the efficiency of the prediction combination model was better than that of a basic BP neural network and a basic GM (1,1) grey model. These studies have confirmed the advantages of combination models and their application value in the prediction of infectious diseases.

This finding may be attributed to several possible reasons. First, the problem of predicting an infectious disease exhibits complex characteristics (32) and contains both linear and nonlinear relationships in the real world. Therefore, a single model may not be able to predict time series with such complex characteristics, and a combination model is more suitable for forecasting infectious disease incidence. The combination model can better reflect the internal regularities of objectives in the prediction of infectious diseases (30). Second, the ES model was constructed to explore the linear pattern of the syphilis time series, and the BP neural network model was used to mine the nonlinear relationship of this time series. The ES-BP combination model can better explain the linear and nonlinear relationships in complex data features of the syphilis incidence time series, resulting in more robust and accurate predictions.

However, our study differs from the previous findings. Zhu et al. (33) used ARIMA, Elman neural network (ERNN), ARIMA-ERNN hybrid, and long short-term memory (LSTM) models to forecast syphilis in mainland China, and their findings showed that the prediction performance of LSTM outperformed ARIMA, ERNN, and ARIMA-ERNN hybrid models. This may be related to data characteristics and sample size. Because they collected monthly syphilis incidence data from 2011 to 2021 with 132 observations, only 39 annual data points between 1982 and 2020 were collected in our study. Therefore, different sample sizes and data characteristics, as well as the use of different models to predict syphilis incidence, have led to different conclusions. In practice, the appropriate combination model was selected based on the characteristics of data and sample size by trial and error to obtain more accurate prediction results.

Conclusion

The ES-BP combination model achieved better prediction performance. Our findings not only reasonably reflect the internal regularities of objectives in the prediction of syphilis incidence but can also help decision-makers monitor and promptly provide effective preventive measures in mainland China.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

This study was supported by the Sichuan Provincial Primary Health Service Development Research Center (Grant No. SWFZ21-Q-59), and Sichuan Provincial Orthopedics Hospital (Grant No. 2021GL01) for funding this study.

Conflict of interest

The authors report no conflicts of interest in this study.

References

1. Pham MD, Ong JJ, Anderson DA, et al (2022). Point-of-Care Diagnostics for Diagnosis of Active Syphilis Infection: Needs, Challenges and the Way Forward. *Int J Environ Res Public Health,* 19(13):8172.
2. Wan Z, Zhang H, Xu H, et al (2020). Maternal syphilis treatment and pregnancy outcomes: a retrospective study in Jiangxi Province, China. *BMC Pregnancy Childbirth,* 20(1):648.
3. Tuddenham S, Ghanem KG (2015). Emerging trends and persistent challenges in the management of adult syphilis. *BMC Infect Dis*, 15:351.
4. Ghanem KG, Ram S, Rice PA (2020). The Modern Epidemic of Syphilis. *N Engl J Med*, 382(9):845-854.
5. Peeling RW, Mabey D, Kamb ML, et al (2017). Syphilis. *Nat Rev Dis Primers,* 3:17073.
6. Zhang X, Zhang T, Pei J, et al (2016). Time Series Modelling of Syphilis Incidence in China from 2005 to 2012. *PLoS One*, 11(2):e0149401.
7. Chen X, Li G, Gan Y, et al (2019). Availability of benzathine penicillin G for syphilis treatment in Shandong Province, Eastern China. *BMC Health Serv Res*, 19(1):188.
8. Guo X, Liu S, Wu L, et al (2014). Application of a novel grey self-memory coupling model to forecast the incidence rates of two notifiable diseases in China: dysentery and gonorrhea. *PLoS One*, 9(12):e115664.
9. Li R H, Huang J, Luo S Y (2020). Prediction of syphilis incident rate and number in China based on the GM (1, 1) grey model. *Food Therapy and Health Care*, 2(4): 170-175.
10. Ibáñez-Cervantes G, León-García G, Vargas-De-León C, et al (2020). Epidemiological behavior and current forecast of syphilis in Mexico: increase in male population. *Public Health*, 185:386-393.
11. Zhang X, Zhang T, Young AA, et al (2014). Applications and comparisons of four time series models in epidemiological surveillance data. *PLoS One*, 9(2):e88075.
12. Guan P, Wu W, Huang D (2018). Trends of reported human brucellosis cases in mainland China from 2007 to 2017: an exponential smoothing time series analysis. *Environ Health Prev Med,* 23(1):23.
13. Wang XY, Yu Y, Hu Y, et al (2020). COVID-19 analysis and forecast based on Exponential Smoothing Model in Hubei Province. *J Public Health Pre Med,* (6): 1-4. (in Chinese)
14. Rostami M, Jalilian A, Poorolajal J, et al (2019). Time Series Analysis of Monthly Suicide Rates in West of Iran, 2006-2013. *Int J Prev Med*, 10:78.
15. Zhang YQ, Li XX, Li WB, et al (2020). Analysis and predication of tuberculosis registration rates in Henan Province, China: an exponential smoothing model study. *Infect Dis Poverty*, 9(1):123.
16. Shang Y, Zhang TT, Wang ZF, et al (2022). Spatial epidemiological characteristics and exponential smoothing model application of tuberculosis in Qinghai Plateau, China. *Epidemiol Infect,* 150:e37.
17. Liu T, Zou G (2021). Evaluation of Mechanical Properties of Materials Based on Genetic Algorithm Optimizing BP Neural Network. *Comput Intell Neurosci*, 2021:2115653.
18. Fan K, Gu X (2021). Image Quality Evaluation of Sanda Sports Video Based on BP Neural Network Perception. *Comput Intell Neurosci*, 2021:5904400.
19. Liu W, Bao C, Zhou Y, er al (2019). Forecasting incidence of hand, foot and mouth disease using BP neural networks in Jiangsu province, China. *BMC Infect Dis*, 19(1):828.
20. Alim M, Ye GH, Guan P (2020). Comparison of ARIMA model and XGBoost model for prediction of human brucellosis in mainland China: a time-series study. *BMJ Open*, 10(12):e039676.
21. Li Z, Li Y (2020). A comparative study on the

prediction of the BP artificial neural network model and the ARIMA model in the incidence of AIDS. *BMC Med Inform Decis Mak*, 20(1):143.
22. Wu L, Liu S, Yang Y (2016). Grey double exponential smoothing model and its application on pig price forecasting in China. *Applied Soft Computing*, 39: 117-123.
23. Chen L, Jagota V, Kumar A (2021). Research on optimization of scientific research performance management based on BP neural network. *Int J Syst Assur Eng Manag* , 1-10.
24. Lyu J, Zhang J (2019). BP neural network prediction model for suicide attempt among Chinese rural residents. *J Affect Disord*, 246:465-473.
25. Lu S (2021). Research on GDP Forecast Analysis Combining BP Neural Network and ARIMA Model. *Comput Intell Neurosci*, 2021:1026978.
26. Feng W, Feng F (2022). Research on the Multimodal Digital Teaching Quality Data Evaluation Model Based on Fuzzy BP Neural Network. *Comput Intell Neurosci*, 2022:7893792.
27. Jia J, Liu M, Xue Z, et al (2021). Prediction of Mumps Incidence Trend in China Based on Difference Grey Model and Artificial Neural Network Learning. *Iran J Public Health*, 50(7):1405-1415.
28. Zhou H, Zhang H, Yang C (2020). Hybrid-model-based intelligent optimization of iron-making process. *IEEE Transactions on Industrial Electronics,* 67(3):2469-2479.
29. Zhao D, Zhang H, Cao Q, er al (2022). The research of ARIMA, GM(1,1), and LSTM models for prediction of TB cases in China. *PLoS One*, 17(2): e0262734.
30. Wang Y, Xu C, Zhang S, et al (2019). Temporal trends analysis of tuberculosis morbidity in mainland China from 1997 to 2025 using a new SARIMA-NARNNX hybrid model. *BMJ Open*, 9(7):e024409.
31. Wu W, Guo J, An S, et al (2015). Comparison of Two Hybrid Models for Forecasting the Incidence of Hemorrhagic Fever with Renal Syndrome in Jiangsu Province, China. *PLoS One,* 10(8):e0135492.
32. Wang Y, Xu C, Wang Z, et al (2018). Time series modeling of pertussis incidence in China from 2004 to 2018 with a novel wavelet based SARIMA-NAR hybrid model. *PLoS One*, 13(12):e0208404.
33. Zhu Z, Zhu X, Zhan Y, et al (2022). Development and comparison of predictive models for sexually transmitted diseases-AIDS, gonorrhea, and syphilis in China, 2011-2021. *Front Public Health*, 10:966813.