PLOS ONE

A comparative study of two methods to predict the incidence of hepatitis B in Guangxi, China

Yanling Zheng [1]*, **Liping Zhang**[1], **XiXun Zhu**[2], **Gang Guo**[3]*

1 College of Medical Engineering and Technology, Xinjiang Medical University, Urumqi, People's Republic of China, 2 School of Computer Engineering, Jingchu University of Technology, Jingmen, People's Republic of China, 3 State Key Laboratory of Pathogenesis, Prevention and Treatment of High Incidence Diseases in Central Asia, Clinical Medicine Institute, the First Affiliated Hospital of Xinjiang Medical University, Urumqi, People's Republic of China

* zhengyl_math@sina.cn (YZ); zhub168_math@sina.com (GG)

Abstract

In recent years, the incidence of hepatitis B (HB) in Guangxi is higher than that of the national level; it has been increasing, so it is urgent to do a good predictive research of HB incidence, which can help analyze the early warning of hepatitis B in Guangxi, China. In the study, the feasibility of predicting HB incidence in Guangxi by autoregressive integrated moving average (ARIMA) model method and Elman neural network (ElmanNN) method was discussed respectively, and the prediction accuracy of the two models was compared. Finally, we established the ARIMA (0, 1, 1) model and ElmanNN with 8 neurons. Both ARIMA (0, 1, 1) model and ElmanNN model had good performance, and their prediction accuracy were high. The fitting and prediction root-mean-square error (RMSE) and mean absolute error (MAE) of ElmanNN were smaller than those of ARIMA (0, 1, 1) model, which indicated that ElmanNN was superior to ARIMA (0, 1, 1) model in predicting the incidence of hepatitis B in Guangxi. Based on the ElmanNN, the HB incidence from September 2019 to December 2020 in Guangxi was predicted, the predicted results showed that the incidence of HB in 2020 was slightly higher than that in 2019 and the change trend was similar to that in 2019, for 2021 and beyond, the ElmanNN model could be used to continue the predictive analysis.

Introduction

Hepatitis B (HB) is a chronic infectious disease caused by HB virus infection in liver tissue. If HB is not treated in time, a small number of patients with liver tissue will further appear inflammatory aggravation, liver cirrhosis, liver cancer [1]. At present, there are about 300 million people living with chronic HB virus (including chronic HB patients and HB virus carriers), mainly in East Asia, South Asia, sub-Africa and other regions [2]. Due to the large number of people living with chronic HB virus, it has become an important public health problem in the world. China is the country with the heaviest burden of HB infection. At

present, there are about 100 million chronic HB virus infections in China. However, the new cases of liver cancer in China account for about 59% of the world every year, most of which are caused by hepatitis B virus infection and hepatitis C virus infection [3,4]. Therefore, the current situation of HB prevention and control in China is still very serious.

In order to help the prevention and control of hepatitis B, research from different perspectives should be done, such as the early warning research, the analysis of socio-economic factor, demographic factors and meteorological environment factors of affecting HB, etc, among them, reasonable prediction and early warning links can help to do a good job of resource demand planning in advance, so as to formulate the corresponding prevention and control preparations.

The prediction analysis includes qualitative prediction and quantitative prediction, qualitative prediction is fuzzy and usually used for descriptive expression, while quantitative prediction is more precise and usually used to predict more accurate future situation by data analysis. At present, quantitative prediction is widely used in various fields, and the model methods used are various, such as: neural network models were used to predict crash frequency [5,6], and Bayesian spatial generalized ordered logit model was developed for predicting the crash severity [7]; the integrated approach of belief rule base and deep learning was used to predict air pollution [8]; machine-learning techniques was used to accurately predict battery life [9]; regression model was used to energy performance of a net-zero energy building [10], etc, the results of these studies can help early analysis for the future situation of these research field, so as to do the planning work well in advance and reduce loss. In recent years, many mathematical model methods were used to predict the incidence of infectious diseases, such as linear model [11,12], dynamics model [13,14], grey model [15], time series ARIMA model, neural network model, and so on. Since the time series of infectious diseases often have the characteristics of trend and randomness, ARIMA model and neural network model can capture the regularity of such data well, so they were most widely used and obtained good prediction performance and high prediction accuracy [16–24].

In China, the incidence of HB in Guangxi Province is higher than that of the national level, and it has an upward trend in recent years, so the task of prevention and control of HB in Guangxi Province is very serious. This study attempts to help analyze the early warning of hepatitis B. Based on the data of HB incidence from January 2012 to August 2019 in Guangxi, we studied the feasibility of predicting HB incidence in Guangxi by ARIMA model and Elman neural network model, and compared the prediction effects of the two models and did predictive analysis, so as to provide scientific reference for the prevention and control of HB in Guangxi.

Materials and methods

Study area and data sources

Guangxi is located in the south of China, between 20˚54' to 26˚24' N and 104˚26' to 112˚04' E. It covers an area of 236,700 square kilometers and has a population of 49.26 million at the end of 2018. This study collected the data of hepatitis B cases in Guangxi and the population data of Guangxi from January 2012 to August 2019. The reported case data of hepatitis B (HB) in Guangxi came from the website of Health Commission of Guangxi, China, and the population data came from Guangxi Bureau of Statistics. The incidence data of HB were calculated by using the case data and population data (see S1 File). Matlab2016b and Eviews7 were used for data analysis.

ARIMA model

ARIMA (p, d, q) model is an important time series analysis and prediction model, which is also called Autoregressive Integrated Moving Average Model [25,26]. Because the model can

capture the trend and randomness of data, it is widely used in the prediction of infectious diseases, and has achieved good prediction results, such as, Wang et al. [16] found that ARIMA model could predict the morbidity of influenza in Ningbo, China, 2006–2014, successfully; Shen et al. [18] analyzed that ARIMA model was successful in predicting hemorrhagic fever with renal syndrome in China; Anokye et al. [21] found that ARIMA model had good performance in predicting malaria incidence; etc [17,19,20].

The ARIMA (p, d, q) model is established on the basis of stationary time series, so the stationarity of time series is an important prerequisite for modeling. The Augmented Dickey-Fuller (ADF) unit root test method is generally used to test the stability of time series. If the significant level P is less than 0.05, the data is stable, otherwise, the time series can also be stable through some operations (for example, logarithm, difference, if the n order difference is done, d = n). When d is 1, its mathematical expression is as follows

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q},$$

where, X_t is the incidence of t time, $y_t = X_t\text{-}X_{t-1}$, $\phi_i, i = 1,2,\cdots,p$ and $\theta_i, i = 1,2,\cdots,q$ are parameters, p is the autoregressive order and q is the moving average order.

The modeling process includes model identification, parameter estimation, hypothesis testing and residual analysis. Firstly, the ACF and PACF diagrams of stationary data are made to determine the possible values of p and q, and then the parameter estimation and parameter hypothesis test are made. Next, the residual analysis of the model passed the parameter test is carried out, if the self-correlation and the partial correlation coefficient of the model residue are within the standard deviation of two times, then the model residual is white noise, which indicates that the established model has good performance and can be used for prediction and analysis.

Elman neural network model [27,28]

Elman neural network (ElmanNN) was first proposed by J. L. Elman in 1990 to solve the problem of speech processing. It is a typical local regression network. Elman network can be regarded as a recurrent neural network with local memory unit and local feedback connection. It is a dynamic feedback network, which can feedback, store and use the output information of the past time. It can not only realize the modeling of static system, but also realize the mapping of dynamic system and directly reflect the dynamic characteristics of the system, it is better than propagation neural network in computing power and network stability. Elman neural network is very suitable for time-series prediction, in many fields of prediction research; it has achieved many successful results, such as Liu et al. [29] did the analysis of wind speed prediction using ElmanNN; Ismael et al. [30] predicted the data consumption for power demands by ElmanNN successfully; Li et al. [31] did the prediction of Urban Rail Transit Sectional Passenger Flow Based on ElmanNN successfully; etc [32–34].

It consists of input layer, hidden layer, bearing layer and output layer. The mathematical expressions of Elman network structure used in this study are as follows:

$$x(k) = f(w^1 x_c(k) + w^2 u(k-1)),$$

$$x_c(k) = x(k-1),$$

$$y(k) = g(w^3 x(k)).$$

Where, y, x, u and x_c represent m-dimensional output vector, n-dimensional implicit layer node vector, r-dimensional input vector and n-dimensional feedback state vector, respectively.

w^1, w^2 and w^3, respectively, represent the connection weight from the receiving layer to the hidden layer, the connection weight from the input layer to the hidden layer, and the connection weight from the hidden layer to the output layer. $g(w^3 x(k))$ is the transfer function of the output neuron, which is the linear combination of the hidden layer output. $f(w^1 x_c(k) + w^2 u(k-1))$ is the transfer function of hidden layer neurons.

In general, if the number of neurons in the hidden layer is too small, the network will not be able to establish a complex mapping relationship, so that the network cannot be trained, or cannot identify the previous samples, and the fault tolerance is poor, if the number of neurons is too large, then the network learning time is too long, and the error is uncertain the smallest. In this paper, the optimal number of neurons is selected from 1 to 20 according to the value of training sample and its fitting mean square error (MSE), and the number of neurons corresponding to the minimum MSE value is the best number of neurons. In this study, the Elman network establishment process is shown in S2 File.

Model comparison

Root-mean-square error (RMSE) and mean absolute error (MAE) are used to compare the performance of two models, the smaller the two values, the better the model.

Results

Establishment of ARIMA model

From January 2012 to August 2019, 486983 cases were reported in Guangxi and 183 cases died, from 2016 to 2019, the incidence of hepatitis B in Guangxi showed an upward trend, especially in 2018. See Fig 1.

The data from January 2012 to August 2019 were divided into two parts, the part from January 2012 to December 2018 was used to build the model, and the part from January 2019 to August 2019 was used to test the model. The ARIMA model requires the data to be stable. Firstly, the stability of the modeling part of the data was checked, the ADF test result P was more than 0.05, which indicated that the data were not stable. In order to make the data stable, the first order difference of the data was done, and then the ADF test was performed. The test results were shown in Table 1. P was less than 0.05, which indicated that the data were stable after the first order difference, and d equaled 1. Draw ACF and PACF diagrams for stationary data, look at Fig 2, from which we determined p, q take 0 or 1, so the tentative models were ARIMA (1, 1, 1), ARIMA (0, 1) and ARIMA (1, 1, 0). The least square method was used to test the parameters of the three models, and only the ARIMA (0, 1, 1) model passed the test, then the residual analysis of the ARIMA (0, 1, 1) model was carried out. From the residual ACF and PACF graph (Fig 3), it can be seen that the self-correlation and the partial correlation coefficient of the model residue were within the standard deviation of two times, which indicated that the model residual was white noise, so, the ARIMA (0, 1, 1) was feasible. Based on this model, the incidence data of HB in Guangxi from January 2012 to December 2018 were fitted. The fitting RMSE was 0.95, the fitting MAE was 0.7; the incidence data of HB in Guangxi from January 2019 to August 2018 were predicted, the prediction RMSE was 0.94 and the prediction MAE was 0.81.

Establishment of ElmanNN model

After careful observation of Fig 1, we could find that there was a certain similarity in the sequence change map of HB incidence each year; therefore, we created twelve time-lagged variables as input features. The incidence data of hepatitis B in Guangxi were processed as follows
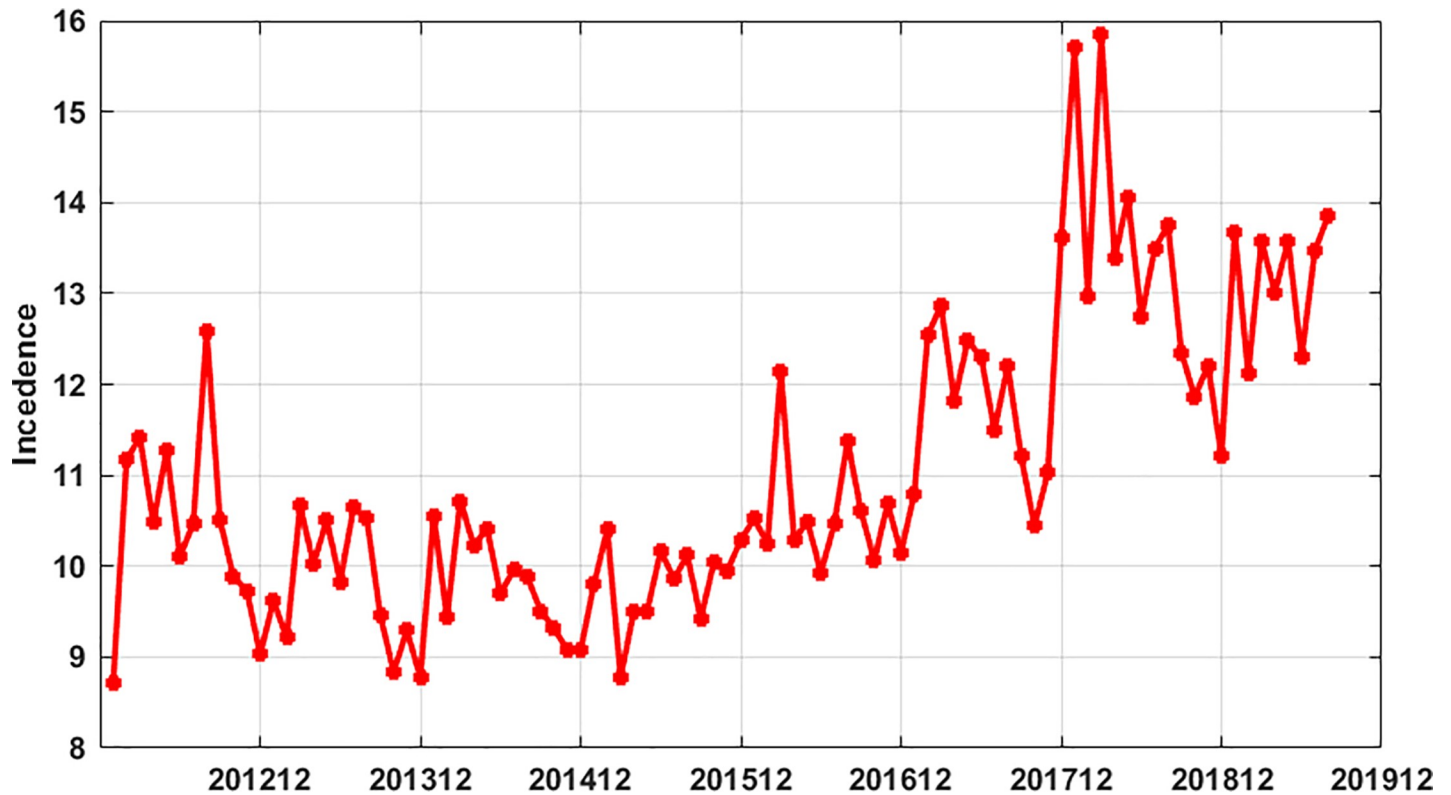
Fig 1. Incidence of hepatitis B in Guangxi from January 2012 to August 2019.

input and output matrix,

$$input\ matrix = \begin{bmatrix} x_1 & x_2 & \cdots & x_i \\ x_2 & x_3 & \cdots & x_{i+1} \\ & \cdots\cdots\cdots \\ x_N & x_{N+1} & \cdots \\ x_{N+1} & x_{N+2} & \cdots \end{bmatrix}, output\ matrix = \begin{bmatrix} x_{N+1} \\ x_{N+2} \\ \cdots \\ x_{N+m} \end{bmatrix}$$

where, x_i represents the incidence of HB at time i, N = 12. m is the number of input variables, and the first input variable is $(x_1, x_2, \ldots\ldots, x_i)$, the second input variable is $(x_2, x_3, \ldots\ldots, x_{i+1})$, and so on. In this study, m was 12, and the matlab cyclic structure was used to find the optimal neurons from 1 to 20. Finally, when the number of neurons was 8, the MSE was the smallest (see Fig 4). The Elman network structure with neuron 8, 12 input variables and 1 output variable was constructed, and the maximum number of iterations was 2000 (the error drop of the

Table 1. The ADF test of the modeling data.

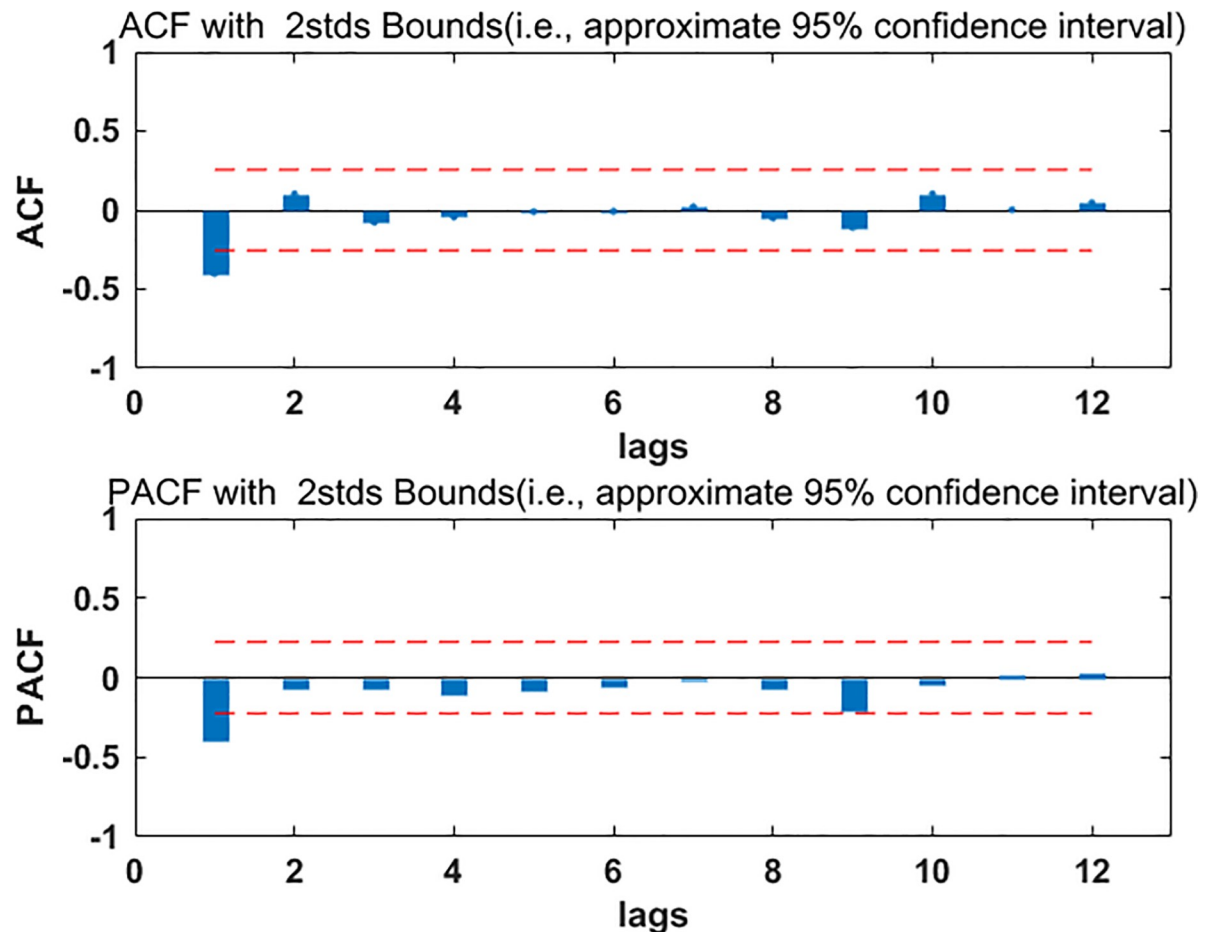|		t-Statistic	P
Augmented Dickey-Fuller test statistic		-14.17	0.0001
Test critical values:	1% level	-3.51	
	5% level	-2.9	
	10% level	-2.59	

Fig 2. The ACF and PACF diagrams of stationary data. ACF = autocorrelation function, PACF = partial autocorrelation function.

network was shown in Fig 5). The fitting RMSE of ElmanNN was 0.37, and the fitting MAE was 0.27. Using the ElmanNN to predict the incidence of HB in Guangxi from January 2019 to August 2019, the prediction RMSE was 0.89 and the prediction MAE was 0.70.

Comparison between ARIMA model and ElmanNN model

RMSE and MAE indexes were used to compare ARIMA model with ElmanNN. The fitting and prediction index values of the two models were shown in Table 2. The fitting and prediction of the actual incidence of HB by the two models were shown in Fig 6.

Discussion

Chronic hepatitis B, as a widely spread infectious disease, has brought great burden to the health cause all over the world. The World Health Organization has proposed that HB should be eliminated in 2030 [2]. China is the country with the heaviest burden of HB infection in the world. In order to achieve the goal of eliminating HB by 2030, it is necessary not only to do a good job in the treatment of HB, to improve the cure rate of HB, but also to do a good job in the prevention and control of HB. The incidence of HB in Guangxi is higher than that of the national level in China. It can be seen from the Fig 1 that the incidence of HB in Guangxi has shown an upward trend in recent years, which needs to be paid more attention to by the
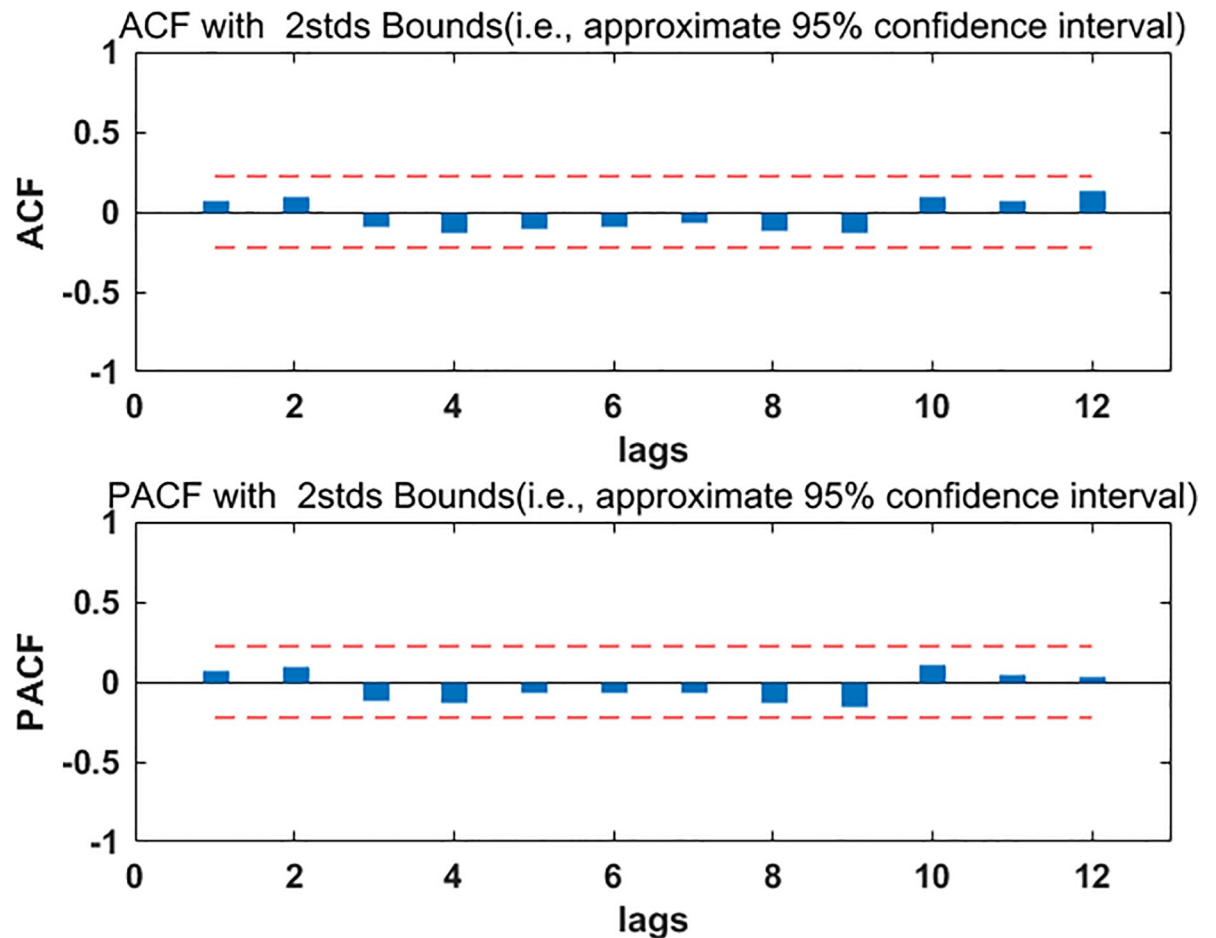
Fig 3. The ACF and PACF diagrams of residual errors of ARIMA (0, 1, 1) model. ACF = autocorrelation function, PACF = partial autocorrelation function.

prevention and control departments, and analyzes the cause of increasing incidence and take necessary measures to prevent and control it. However, there was little research on the prediction of HB incidence in Guangxi at present; therefore, this study aimed to do prediction analysis of HB incidence in Guangxi by two widely popular prediction model methods, so as to provide scientific reference for prevention and control work.

The data of HB incidence from January 2012 to August 2019 were analyzed, the data from January 2012 to December 2018 were used as modeling data and the data from January 2019 to August 2019 as test model data.

First of all, we analyzed the feasibility of ARIMA model prediction. Through the data analysis of the modeling part, it was found that the data were stable after the first order difference. The possible p, q values in ARIMA (p, 1, q) were determined by drawing ACF and PACF maps on the stationary data. Finally, only the ARIMA (0, 1, 1) model passed the parameter test, and the residual of the model was also white noise, which indicated that the model has good performance and could be used to predict the incidence of HB in Guangxi. Using this model to predict the incidence data of HB in Guangxi from January 2019 to August 2019, the prediction RMSE was 0.94 and the prediction MAE was 0.81, the two values were relatively small, which indicated that the prediction accuracy of the model was high. Secondly, we analyzed the feasibility of ElmanNN prediction, and constructed the input and output matrix of
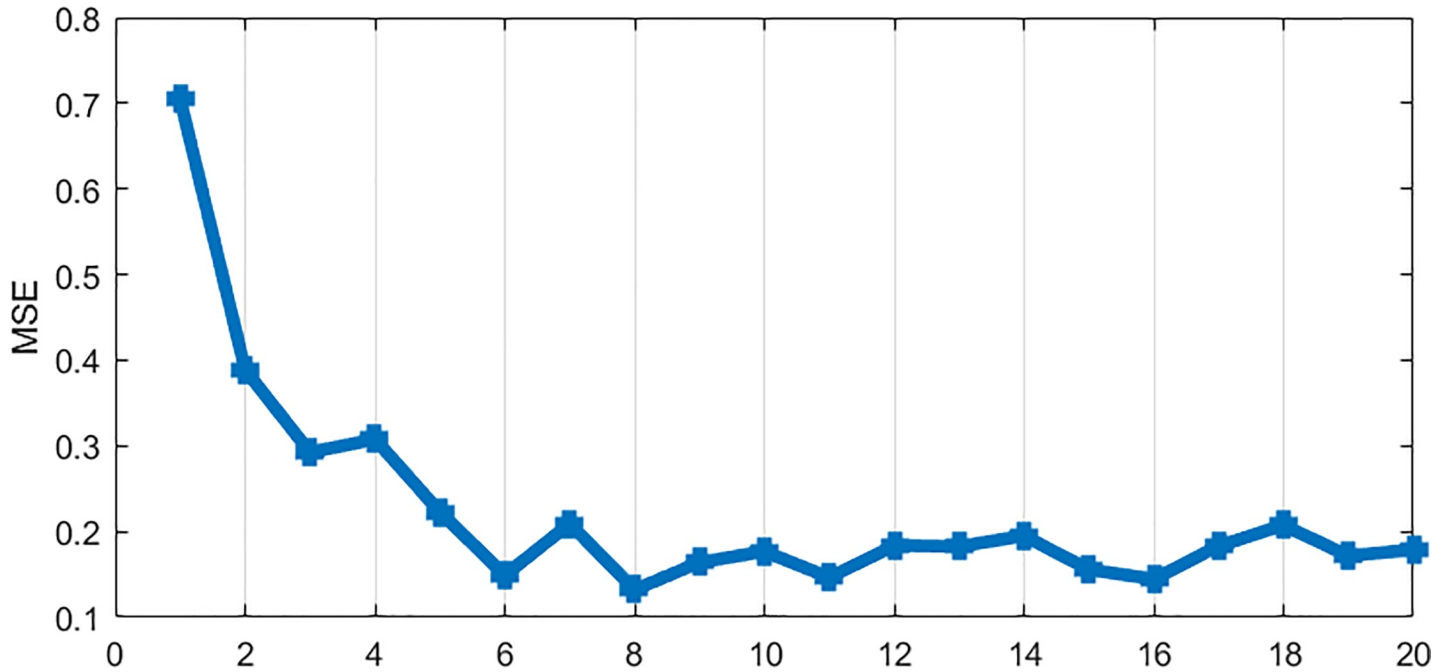
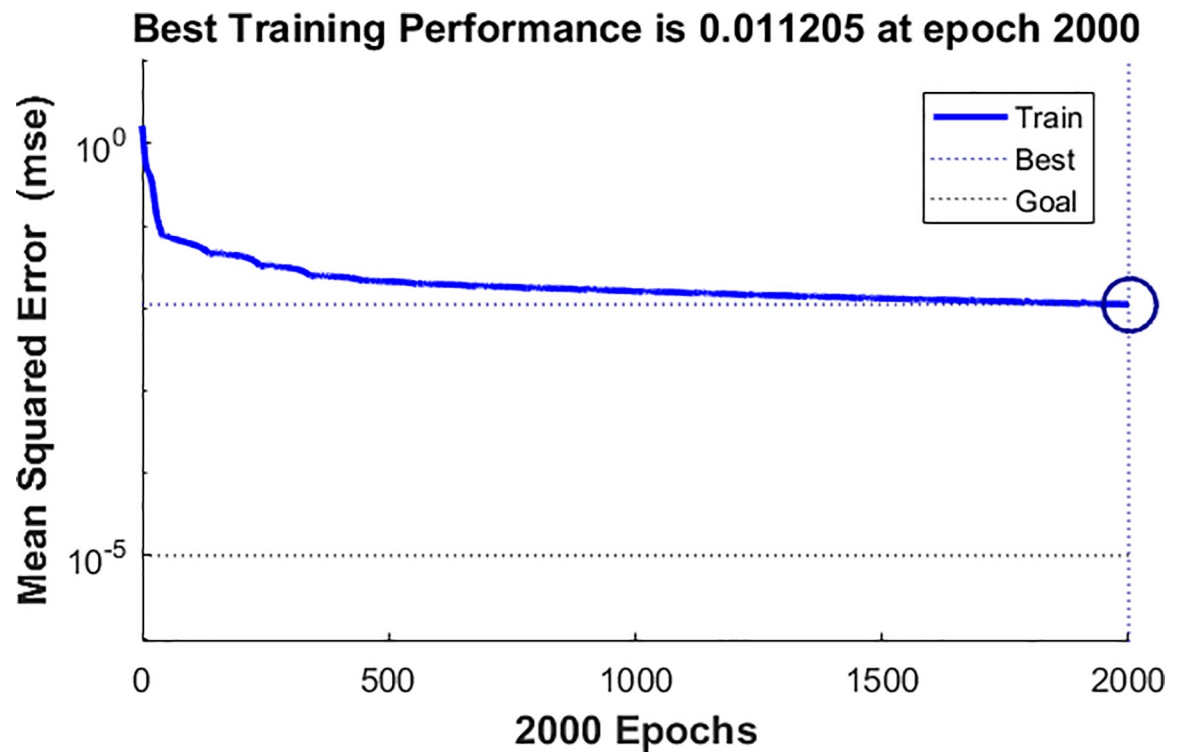**Fig 4. The numbers of neurons in ElmanNN and corresponding MSE values.**

Fig 5. The error drop curve of Elman network.

Table 2. The fitting and prediction RMSE and MAE values of ARIMA (0, 1, 1) model and ElmanNN model.

	ARIMA(0,1,1)fitting	ARIMA(0,1,1)prediction	ElmanNN fitting	ElmanNN prediction
RMSE	0.95	0.94	0.37	0.89
MAE	0.7	0.81	0.27	0.70

ElmanNN model based on Guangxi HB incidence data from January 2012 to August 2019. When the number of neurons was 8, the network was optimal. We used the optimal ElmanNN network to fit and predict the HB incidence data in Guangxi,China, the prediction RMSE was 0.89, and the prediction MAE was 0.70, these two values were relatively small, which indicated that the prediction accuracy of the model was high. The prediction performance of ARIMA model and Elman model was compared by RMSE and MAE. From Table 2, it was found that the RMSE and MAE values of Elman model were smaller than those of ARIMA model, which indicated that Elman model had better prediction performance. Using ARIMA model and Elman model to fit the observed data of hepatitis B incidence (see Fig 6); we could see that the fitting effect of Elman model was better than that of ARIMA model, especially from December 2016 to December 2018. Based on the ElmanNN model, we made a prediction analysis of the HB incidence in Guangxi from September 2019 to December 2020(see Fig 6), and the forecast results showed that the trend of HB incidence in 2020 was similar to that in 2019 and slightly higher than that in 2019. According to the forecast results of this study, the department of disease prevention and control should make the optimal allocation of resources in advance, so as to better serve the health protection work of the HB.

Conclusions

The World Health Organization has put forward the goal of eliminating hepatitis B in 2030. However, in recent years, the incidence of hepatitis B in Guangxi, China, has shown an upward trend, which needs to attract the attention of disease prevention and control departments, some departments should take joint actions to reduce the incidence of HB, and research from different angles should also be carried out. However, at present, there is little research on
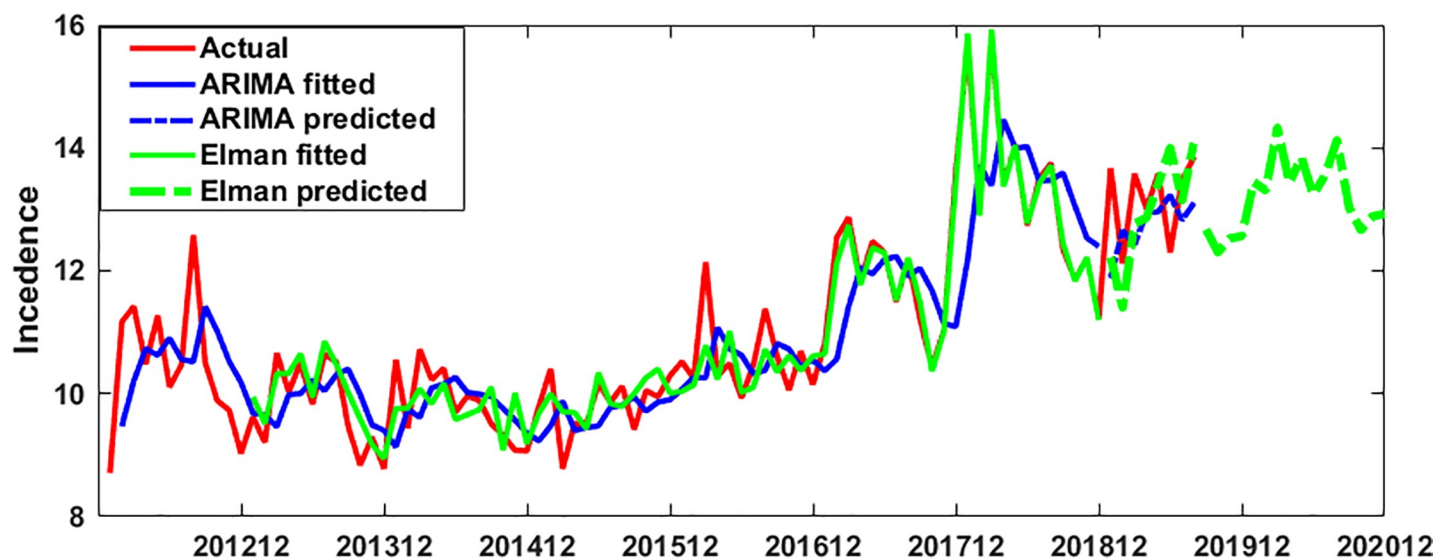

Fig 6. The fitting and prediction of HB incidence by ARIMA (0, 1, 1) and ElmanNN model.

quantitative prediction of Hepatitis B in Guangxi. Therefore, in order to do a good job in prevention and resource allocation, it is necessary to give prediction and analysis in advance. Hence, this study used the widely used ARIMA model and Elman neural network to predict the incidence of HB in Guangxi. The results showed that ARIMA (0, 1, 1) model and ElmanNN model could be used to predict the incidence of HB in Guangxi, and the prediction performance was good, and the prediction accuracy of the ElmanNN model was higher than that of the ARIMA model. Based on the ElmanNN model, we predicted the HB incidence from September 2019 to December 2020, the predicted results showed that the incidence of HB in 2020 was slightly higher than that in 2019 and the change trend was similar to that in 2019, for 2021 and beyond, the ElmanNN model could be used to continue the predictive analysis of HB incidence in Guangxi,China.

At present, under the severe situation of rising HB incidence in Guangxi, some institutions need joint action to help do a good job in the prevention and control of HB. This study only made an analysis from the aspect of incidence prediction, and provided some reference for the optimal allocation of health resources in advance. Some studies had shown that many factors were related to the high incidence of diseases, such as socio-economic factors [34,35], demographic factors [36,37] meteorological factors [38–41], etc. The related research on the influencing factors of HB incidence should also be carried out in time, which can also help the prevention and control of HB. In further research work, we will use appropriate statistical methods to explore the key factors affecting HB incidence.

Supporting information

S1 File. Data set of HB incidence used in the manuscript, as shown in Fig 1.
(XLSX)

S2 File. The flowchart of Elman network establishment process.
(TIF)

Acknowledgments

I would like to express my gratitude to anonymous peer reviewers for carefully revising our manuscript and for his or her useful comments.

Author Contributions

Conceptualization: Yanling Zheng, Liping Zhang, XiXun Zhu, Gang Guo.

Formal analysis: Yanling Zheng.

Funding acquisition: Liping Zhang.

Investigation: Yanling Zheng, Liping Zhang, XiXun Zhu, Gang Guo.

Methodology: Yanling Zheng.

Project administration: Yanling Zheng, Gang Guo.

Resources: XiXun Zhu.

Supervision: Gang Guo.

Validation: Liping Zhang, XiXun Zhu, Gang Guo.

Visualization: Yanling Zheng, Liping Zhang, XiXun Zhu, Gang Guo.

Writing – original draft: Yanling Zheng.

Writing – review & editing: Yanling Zheng, Liping Zhang, XiXun Zhu, Gang Guo.

References

1. Devin Razavi-Shearer Ivane Gamkrelidze, Mindie HNguyen Ding-Shinn Chen, Pierre Van Damme ZaighamAbbas, et al. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. The Lancet Gastroenterology & Hepatology, 2018, 3(6):383–403.

2. Dusheiko G, Agarwal K. Delineating the global challenges of hepatitis B virus infection. The Lancet Gastroenterology & Hepatology, 2018, 3(6):372–373.

3. World Health Organization. Global Hepatitis Report, 2017.

4. Rongshou ZhengChunfeng, Zhang QuSiwei, Zeng Hongmei, Sun Kexin, Gu Xiuying, et al. Liver cancer incidence and mortality in China: Temporal trends and projections to 2030. Cancer Research in China: English version, 2018, 30(6):571–579.

5. Zeng Qiang, Huang Helai, XinPei, Wong S.C. Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks. Analytic Methods in Accident Research, 2016, 10: 12–25.

6. Zeng Qiang, Huang Helai, Pei Xin, Wong S.C., Mingyun Gao. Rule extraction from an optimized neural network for traffic crash frequency modeling. Accident Analysis and Prevention, 2016, 97: 87–95.

7. Zeng Qiang, Hao Wei, Lee Jaeyoung, Chen Feng. Investigating the Impacts of Real-Time Weather Conditions on Freeway Crash Severity: A Bayesian Spatial Analysis, International Journal of Environmental Research and Public Health, 2020, 17(8): 2768.

8. Kabir Sami, Raihan Ul Islam Mohammad Shahadat Hossain, Andersson Karl. An Integrated Approach of Belief Rule Base and Deep Learning to Predict Air Pollution. Sensors, 2020, 20(7):1956.

9. Berecibar M. Machine-learning techniques used to accurately predict battery life. Nature, 2019, 568 (7752):325–326. https://doi.org/10.1038/d41586-019-01138-1 PMID: 30980036

10. Kneifel J, Webb D. Predicting energy performance of a net-zero energy building: A statistical approach. Applied Energy, 2016, 178(sep.15):468–483.

11. Heathfield Laura, Lacerda Miguel, Nossek Christel, Roberts Lisa, Rajkumar S Ramesar. Stargardt disease: towards developing a model to predict phenotype. European Journal of Human Genetics Ejhg, 2013, 21(10):1173–1176. https://doi.org/10.1038/ejhg.2013.92 PMID: 23695285

12. Song L, Horvath S. Predicting COPD status with a random generalized linear model. Systems Biomedicine, 2013, 1(4):261–267.

13. Brittin J, Araz O M, Nam Y, Huang TT-K. A system dynamics model to simulate sustainable interventions on chronic disease outcomes in an urban community. Journal of Simulation, 2015, 9(2):140–155.

14. Tomaskova Hana, Kuhnova Jitka, Cimler Richard, Dolezal Ondrej, and Kuca Kamil. Prediction of population with Alzheimer's disease in the European Union using a system dynamics model. Neuropsychiatric Disease & Treatment, 2016, 12(Issue 1):1589–1598.

15. Yang X, Zou J, Kong D, Jiang G. The analysis of GM (1, 1) grey model to predict the incidence trend of typhoid and paratyphoid fevers in Wuhan City, China. Medicine, 2018, 97(34):e11787. https://doi.org/10.1097/MD.0000000000011787 PMID: 30142765

16. Chunli Wang, Yongdong Li, Wei Feng, Kui Liu, Shu Zhang, Fengjiao Hu, et al. Epidemiological Features and Forecast Model Analysis for the Morbidity of Influenza in Ningbo, China, 2006–2014. Int J Environ Res Public Health, 2017, 14(6):559.

17. Khakban A, Sin D D, Fitzgerald J M. The Projected Epidemic of Chronic Obstructive Pulmonary Disease Hospitalizations over the Next 15 Years. A Population-based Perspective. American Journal of Respiratory & Critical Care Medicine, 2017, 195(3):287.

18. Shen Z Z, Jiang Y. Comparison of autoregressive integrated moving average model and generalised regression neural network model for prediction of haemorrhagic fever with renal syndrome in China: a time-series study. BMJ Open, 2019, 9(6):e025773. https://doi.org/10.1136/bmjopen-2018-025773 PMID: 31209084

19. Hasan A H M I, Haddawy P, Lawpoolsri S. A Comparative Analysis of Bayesian Network and ARIMA Approaches to Malaria Outbreak Prediction. 2017.

20. Zhou Lingling, Zhao Ping, Wu Dongdong, Cheng Cheng, Huang Hao. Time series model for forecasting the number of new admission inpatients. BMC Medical Informatics & Decision Making, 2018, 18(1):39.

21. Anokye Reindolf, Acheampong Enoch, Owusu Isaac, Edmund Isaac Obeng. Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. Nephron Clinical Practice, 2018, 4(1):-.

22. Ren Yafeng, Fei Hao, Liang Xiaohui, Ji Donghong,Cheng Ming. A hybrid neural network model for predicting kidney disease in hypertension patients based on electronic health records. BMC Medical Informatics and Decision Making, 2019, 19(S2).

23. Alkhasawneh M S. Hybrid Cascade Forward Neural Network with Elman Neural Network for Disease Prediction. Arabian Journal for Science and Engineering, 2019, 44(11):9209–9220.

24. Jamshid Yazdani Charati Ghasem Janbabaei, Alipour Nadia, Mohammadi Soraya, Somayeh Ghorbani Gholiabad Afsaneh Fendereski. Survival prediction of gastric cancer patients by Artificial Neural Network model. Gastroenterology & Hepatology from Bed to Bench, 2018, 11(2):110–117.

25. Govind Eknath Kulkarni Aniket Avinash Muley, Nilesh Kailas Deshmukh Parag Upendra Bhalchandra. Autoregressive integrated moving average time series model for forecasting air pollution in Nanded city, Maharashtra, India. Modeling Earth Systems & Environment, 2018, 4(4):1435–1444.

26. He Z, Tao H. Epidemiology and ARIMA Model of Positive-Rate of Influenza Viruses among Children in Wuhan, China: A Nine-Year Retrospective Study. International Journal of Infectious Diseases, 2018, 74:61–70. https://doi.org/10.1016/j.ijid.2018.07.003 PMID: 29990540

27. Orva N L V, Budi W, Asih I M D. Comparison of ARCH / GARCH model and Elman Recurrent Neural Network on data return of closing price stock. Journal of Physics: Conference Series, 2018, 1025:012103-.

28. Ganesh S S, Arulmozhivarman P, Tatavarti V S N R. Prediction of PM 2.5 using an ensemble of artificial neural networks and regression models. Journal of Ambient Intelligence & Humanized Computing, 2018(8):1–11.

29. Liu H, Tian H Q, Liang X F, Yan F L. Wind speed forecasting approach using secondary decomposition algorithmn, and Elman neural networks. Applied Energy, 2015, 157(NOV.1):183–194.

30. Ismael L, Issa A, Alkurawy L. Prediction the data consumption for power demands by Elman neural network. International Journal of Electrical & Computer Engineering, 2019, 9(5).

31. Li Qian, Qin Yong, Zi Yang Wang Zhong Xin Zhao, Ming Hui Zhan Yu Liu. Prediction of Urban Rail Transit Sectional Passenger Flow Based on Elman Neural Network. Applied Mechanics & Materials, 2014, 505–506:1023–1027.

32. Lin W, Yongqing W, Hua F, Yuping Y. Modeling and Simulation of Gas Emission Based on Recursive Modified Elman Neural Network. Mathematical Problems in Engineering, 2018, 2018 (PT.2):9013839.1–9013839.10.

33. Kelo Sanjay, Dudul Sanjay. A wavelet Elman neural network for short-term electrical load prediction under the influence of temperature. International Journal of Electrical Power & Energy Systems, 2012, 43(1):p.1063–1071.

34. Wang Jie, Wang Jun, Fang Wen, Niu Hongli. Financial Time Series Prediction Using Elman Recurrent Random Neural Networks. Computational Intelligence & Neuroscience, 2016, 2016:1–14.

35. Zuin M., Caserta C., Romanò L., Mele A., Zanetti A., Cannatelli R., et al. Seroepidemiology of HEV and HAV in two populations with different socio-economic levels and hygienic/sanitary conditions. European journal of clinical microbiology & infectious diseases, 2017, 36 (3): 479–485.

36. Simoes E A F. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease. Journal of Pediatrics, 2003, 143(5 Suppl):S118–26.

37. Johanna Haapamäki Ulla Turunen, Roine Risto P., Martti A. Färkkilä, Perttu E. T. Arkkila. Impact of demographic factors, medication and symptoms on disease-specific quality of life in inflammatory bowel disease. Quality of Life Research, 2009, 18(8):961–969. https://doi.org/10.1007/s11136-009-9514-y PMID: 19629750

38. Lelieveld J, Klingmüller K, Pozzer A, Pöschl U, Fnais M, Daiber A, et al. Cardiovascular disease burden from ambient air pollution in Europe reassessed using novel hazard ratio functions. Eur Heart J, 2019, 40(20):1590–1596. https://doi.org/10.1093/eurheartj/ehz135 PMID: 30860255

39. Mohd Danish Khan Hong Ha Thi Vu, Quang Tuan Lai Ji Whan Ahn. Ahn JW. Aggravation of Human Diseases and Climate Change Nexus. Int J Environ Res Public Health, 2019, 16(15):2799.

40. Anna M Stewart-Ibarra, Moory Romero, Hinds Avery Q. J., Rachel Lowe, Roché Mahon, Cedric J. Van Meerbeeck, et al. Co-developing climate services for public health: Stakeholder needs and perceptions for the prevention and control of Aedes-transmitted diseases in the Caribbean. PLoS Negl Trop Dis, 2019, 13(10):e0007772. https://doi.org/10.1371/journal.pntd.0007772 PMID: 31658267

41. Orioli R, Cremona G, Ciancarella L, Solimini AG. Association between PM10, PM2.5, NO2, O3 and self-reported diabetes in Italy: A cross-sectional, ecological study. PLoS One, 2018, 13(1):e0191112.