PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Predictive study of tuberculosis incidence by time series method and Elman neural network in Kashgar, China
AUTHORS	Zheng, Yanling; Zhang, Xueliang; Wang, Xijiang; Wang, Kai; Cui, Yan

VERSION 1 – REVIEW

REVIEWER	JIANMING WANG School of Public Health, Nanjing Medical University, China
REVIEW RETURNED	25-Jul-2020

GENERAL COMMENTS	The authors have constructed a single Box-Jenkins model and Box-Jenkins model combined with Elman neural network model to do prediction analysis of TB incidence in Kashgar with the highest TB incidence in Xinjiang province of China. It is an important work in using the existing surveillance data to predict the possible epidemic trend and provide reference information for the prevention and control of the occurrence and epidemic of TB. However, careful revision is needed. 1. Page 3, line 21-23, the fitting and predicting RMSEs and MAEs of the single AR model should also be listed. 2. Page 4, line 50-52, according to the global tuberculosis report 2019, the number of new TB cases in China ranks second in the world instead of third, second only to India. Please correct it. 3. Some abbreviations should be written in full when they first appear, such as AIC, SC and R2. 4. Page 9, line 4, the value range of □ should be “0 ≤ α < 1” instead of “1 □ □ □ 0”. Please correct. 5. Page 9, line 10-11, how did the authors standardize the data? Please clarify. 6. Page 9, line 12-13, how did the authors determine the input layer and the output layer? Please clarify. 7. The determination of parameters is quite crucial for the construction of a neural network, such as training epochs and goals. How did the author set the parameters of the Elman model? Please clarify. 8. Neural networks are prone to overfitting. How did the authors deal with it? Also, external validation, at least internal validation, is necessary. 9. Page 12 line 25--Page 13 line 16, the discussion should not be a simple repetition of results.

REVIEWER	Yongbin Wang Xinxiang Medical University
REVIEW RETURNED	09-Aug-2020

GENERAL COMMENTS	In China, though the TB incidence was decreasing at a rate of around 3% annually in recent years, owing to a growing transient population, drug-resistant TB, co-infections of HIV-TB and other TB co-morbidities such as diabetes mellitus, hypertension, and immune-compromising disorders, which have caused a risk of recurrence in the TB incidence in some areas, and have also flung down a challenge for goal of ending TB with milestones by 2020 and 2025 and targets by 2030 and 2035 in China. Therefore, as stated by the author that a reliable forecasting approach with robust accuracy and precision to track the epidemic patterns of TB is required. This is also an interesting topic performing forecasting for infectious diseases. This manuscript used the most classical Box-Jenkins model to TB morbidity in Kashgar, where has the highest TB incidence in China. But there are a number of issues that need be addressed by major revision. Major concerns: I have experience use of Box-Jenkins modelling and there seems to be an element of selecting what works best, sometimes without full understanding of the reason a particular version is more successful. The Box-Jenkins method assuming time series to be stationary is the most popular approach for time series estimation. But it is well known that the morbidity data of infectious diseases are commonly affected and constrained by the time-varying trends, cyclicity, seasonal variation and random fluctuation. These facets make the data show complex linear and nonlinear interactions. Namely, the seasonal ARIMA (SARIMA) ARIMA(P,D,Q)(p,d,q) model needs to be established to model the TB incidence, whereas the author only used a simple ARIA(p,d,q). According to my experience, the simple ARIMA fails to unearth the seasonal effects, which may result in a low forecasting level. The discussion was poor, which lacks comparisons of the prediction results to literature, please add Detailed comments: There are many grammar editing errors that need to be corrected (e.g., B incidence in Kashgar. Root mean squared error (RMSE), mean absolute error (MAE) Lines 17-19 this should be linked with ‘and’ ; combined..with..., this should corrected as ‘combined..and...’;Line 35, it should be ‘, respectively’; page10 Line22 ‘to found’? …and so on Name must be capitalized, p-values is more common written in lower-case 'p'… Based on the goal and/or objectives of this study. Please add a paragraph to the introduction to clearly tell the reader the study progress in the methodologies used to model the infectious diseases. Of these methods, why do you use SARIMA model, why not others, this should be written clearly in introduction instead of in discussion). See STROBE reporting guidelines for more information on the importance of doing this. The methods are very heavy in statistical analyses, with inadequate epidemiological details. The following information is required for transparency, replicability, and for readers to evaluate the rigour, merit, and potential biases in the study design.

	According to my experience, the performance evaluation indices (e.g. MAE, MAPE, MSE, RMSE and the like) should be utilized to evaluate the forecasting robustness instead of MAE and RMSE. Particularly the MAPE is a very important index to estimate the fitting and forecasting powers Add one Table reporting all the estimated performance evaluation indices to the main text. In Figs: what do the symbols/numbers/processes mean? please add to the caption as well as in the method descriptions. The same for all Figs. Please take more care for all the figures' caption.

VERSION 1 – AUTHOR RESPONSE

Responses to Reviewers

To Reviewer 1:

The authors have constructed a single Box-Jenkins model and Box-Jenkins model combined with Elman neural network model to do prediction analysis of TB incidence in Kashgar with the highest TB incidence in Xinjiang province of China. It is an important work in using the existing surveillance data to predict the possible epidemic trend and provide reference information for the prevention and control of the occurrence and epidemic of TB. However, careful revision is needed.

From your comments we can feel that you are a very excellent expert in Elman neural network methods. We are very lucky to be able to get your guidance.

Thank you very much for giving us a positive evaluation, which is very encouraging, and thank you very much for your excellent and professional revision of our manuscript. We have studied your comments carefully and have made correction which we hope to meet with your approval.

Q1: Page 3, line 21-23, the fitting and predicting RMSEs and MAEs of the single AR model should also be listed.

Thank you very much for your professional suggestion. We have added the precision index values of AR model fitting and prediction, such as:

For the fitting dataset, the RMSE, MAE and MAPE were 6.15, 4.33 and 0.2858, respectively, for the AR((1,2,8)) model. For the forecasting dataset, the RMSE, MAE and MAPE were 10.88, 8.75 and 0.2029, respectively, for the AR((1,2,8)) model. (please see the Abstract section, line 14-18)

Q2: Page 4, line 50-52, according to the global tuberculosis report 2019, the number of new TB cases in China ranks second in the world instead of third, second only to India. Please correct it.

Thank you very much for your careful preview of this article, we have changed the sentence as follows(please see introduction section, line 53-54),such as：

According to the global tuberculosis report 2019, China has the second highest number of TB cases in the world.

Q3: Some abbreviations should be written in full when they first appear, such as AIC, SC and R2.

We agree with you. All the abbreviations in this article have been written in full when they first appear (please see revised manuscript).

Q4: Page 9, line 4, the value range of α should be “$0 \leq \alpha < 1$” instead of “$1 \leq \alpha < 0$”. Please correct.

We are ashamed that we have made such a mistake. Thank you very much for your careful preview of this article, we've corrected our mistake (Please see Elman neural network model section, line 169), and carefully checked and modified other errors in the article.

Q5: Page 9, line 10-11, how did the authors standardize the data? Please clarify.

Data standardization is scaling the data to a small specific interval. In order to remove the unit limit of the data and convert it into dimensionless pure value, it is convenient for the index of different units or order of magnitude to be compared and weighted In our study, we used function

package mapminmax() to standardize the data, standardized data was in [-1,1] range. (We have added the explanation, please see line 171-176).
The algorithm for this mapminmax() function package is:
It is assumed that x has only finite real values, and that the elements of each row are not all equal.
y=(ymax-ymin)*(x-xmin)/(xmax-xmin)+ymin;
where, ymax=1 and ymin=-1, xmax is the maximum in a row of x, xmin is the minimum in a row of x, and y is the normalized data.
Q6: Page 9, line 12-13, how did the authors determine the input layer and the output layer? Please clarify.
According to the characteristics of the data and the needs of our analysis, this study determined the input and output layer (please see Establishment of AR-Elman Model section, line 246-253). Such as: Due to the a little similarity of the annual trend of TB incidence in Kashgar (see Figure 2),we selected twelve as the number of input layers of Elman and one as the number of output layers representing the forecast value. Supposing that x_t represented the TB incidence at time t, and then the input matrix and the output matrix of the modeling data set used in this study were designed as follows (N=12):
,
Q7: The determination of parameters is quite crucial for the construction of a neural network, such as training epochs and goals. How did the author set the parameters of the Elman model? Please clarify.
We agree with you, and we've added parameter settings (please see Elman neural network model section, line 180-182), such as:
In our study, training epochs and goals of Elman neural network were set to 2000 and 0.00001, respectively. The training codes are set as follows:
net.trainParam.epochs=2000;
net.trainParam.goal=0.00001;
Q8: Neural networks are prone to overfitting. How did the authors deal with it? Also, external validation, at least internal validation, is necessary.
There are many ways to prevent overfitting of neural networks, among which the best way is to use more training data to train the network, we adopted this method.
In our study, we used two modeling methods: Box-Jenkins method and Elman neural network method. In general, in Box-Jenkins method modeling, if monthly time data are more than 50 months, then the data are enough, but in order to avoid neural network overfitting, we tried to collect more data, and finally within our capabilities, we collected 156 months of data for modeling analysis.

We are sorry that we did not express clearly the part of verification that we did in the original article, we explain to you the following:
Based on data from January 2005 to December 2016, we developed the AR((1,2,8)) model, using which we fitted the Kashgar tuberculosis incidence data from January 2006 to December 2016 and predicted the Kashgar tuberculosis incidence from January 2017 to December 2017. Using these data, we constructed a 13 * 132 data Matrix A （according to the form of matrix X below）, and using actual data on the TB incidence in Kashgar from January 2006 to December 2016, we constructed a 13 * 132 data Matrix B, the data of Matrix A and B were used to train and validate the Elman network.

Where, N is 12. The code for training and validation data is as follows:
The training data of Elman network:
trainx=A(1:12, 1:120);
trainy=B(13, 1:120);
The test data of Elman network:
testx = A(1:12, 121:end);
testy =B(13, 121:end);

Finally, we used the constructed AR-Elman hybrid model to fit the tuberculosis incidence from January 2006 to December 2016, and predict the tuberculosis incidence in Kashgar from January 2017 to December 2017(verification part). The accuracy indexes of the model fitting and verification were shown in Table 3.

To help readers better understand our article, we added Table 3 to the revised manuscript (please see revised manuscript, line 261-264, and 490-492).

Table 3. Comparison results of in-sample fitting and out-of-sample forecasting performance for the selected models

Models	Fitted efficacy			Models	Forecasted efficacy		
	RMSE	MAE	MAPE		RMSE	MAE	MAPE
AR((1,2,8))	6.15	4.33	0.2585	AR((1,2,8))	10.88	8.75	0.2029
AR-Elman	3.78	3.38	0.1837	AR-Elman	8.86	7.29	0.2006

Q9: Page 12 line 25--Page 13 line 16, the discussion should not be a simple repetition of results.

Thank you very much for your guidance. We have revised the discussion section, (please see the Discussion section of the revised manuscript, line 287-321). The main changes are as follows:

Many studies have found that Box-Jenkins method has a good ability of fitting and forecasting. For stationary time series that do not contain seasonality, it is more suitable to use the ARMA model of the Box-Jenkins method to do prediction analysis [35], for non-stationary time series of infectious diseases with obvious seasonality, it is more suitable to use seasonal autoregressive integrated moving average (SARIMA) model of the Box-Jenkins method for prediction analysis.[9-12] In our study, from Figure 2, we could see that the seasonality of the TB incidence in Kashgar was not obvious, and we found that the time series of TB incidence was stable by ADF unit root test, and the autocorrelation and partial correlation coefficients of modeling data at lag 12, 24 were not obviously large, therefore, for our research data, we used ARMA model to do forecast analysis, finally, we established AR((1,2,8)) model of the Box-Jenkins method, it has a good performance to fit and predict the TB incidence of Kashgar in Xinjiang. From Figure 2, we can also see that the time series of TB incidence has strong non-linear, the established AR((1,2,8)) model mainly extracted the linear information of data, considering that the neural network can capture the non-linear information of data well, in order to improve the prediction accuracy of TB incidence rate in Kashgar, we used AR((1,2,8)) model and Elman neural network model to establish AR-Elman hybrid model. Many studies have found that the combination model can improve the accuracy of prediction, such as, Wang et al.[28] found that SARIMA-NAR hybrid model has an outstanding ability to improve the prediction accuracy relative to SARIMA model and nonlinear autoregressive network (NAR) model when they were used to predict pertussis incidence in China. Li et al.[27] found ARIMA-GRNN hybrid model was shown to be superior to the single ARIMA model in predicting the short-term TB incidence in the Chinese population. Our research was consistent with these literatures that our AR-Elman hybrid model was more accurate than the single AR((1,2,8)) model.

The incidence of tuberculosis in Kashgar, Xinjiang is very high, and the relevant departments of disease prevention and control in Xinjiang have also done a lot of effective work. Our research was mainly to build a high-precision prediction model to help early warning of tuberculosis in Kashgar. Finally, we established the AR-Elman hybrid model, which had high fitting and prediction accuracy of TB incidence in Kashgar, Xinjiang. Because the development of any event may be affected by many factors, such as social, economic, political, demographic factors, so the long-term forecast accuracy of the AR-Elman hybrid model may decline.

To Reviewer 2:

In China, though the TB incidence was decreasing at a rate of around 3% annually in recent years, owing to a growing transient population, drug-resistant TB, co-infections of HIV-TB and other TB co-

morbidities such as diabetes mellitus, hypertension, and immune-compromising disorders, which have caused a risk of recurrence in the TB incidence in some areas, and have also flung down a challenge for goal of ending TB with milestones by 2020 and 2025 and targets by 2030 and 2035 in China. Therefore, as stated by the author that a reliable forecasting approach with robust accuracy and precision to track the epidemic patterns of TB is required. This is also an interesting topic performing forecasting for infectious diseases. This manuscript used the most classical Box-Jenkins model to

TB morbidity in Kashgar, where has the highest TB incidence in China. But there are a number of issues that need be addressed by major revision.

From your comments we can feel that you are a very excellent expert in Box-Jenkins methods.

We are very lucky to be able to get your guidance.

Thank you very much for giving us a positive evaluation, which is very encouraging, and thank you very much for your excellent and professional revision of our manuscript. We have studied your comments carefully and have made correction which we hope to meet with your approval.

Major concerns:

Q1. I have experience use of Box-Jenkins modelling and there seems to be an element of selecting what works best, sometimes without full understanding of the reason a particular version is more successful. The Box-Jenkins method assuming time series to be stationary is the most popular approach for time series estimation. But it is well known that the morbidity data of infectious diseases are commonly affected and constrained by the time-varying trends, cyclicity, seasonal variation and random fluctuation. These facets make the data show complex linear and nonlinear interactions. Namely, the seasonal ARIMA (SARIMA) ARIMA(P,D,Q)(p,d,q) model needs to be established to model the TB incidence, whereas the author only used a simple ARIMA(p,d,q). According to my experience, the simple ARIMA fails to unearth the seasonal effects, which may result in a low forecasting level.

Thank you very much for your guidance. In the initial conception of this article, we were very careful, and we first thought of using the SARIMA model for predictive analysis, as you thought, but when we analyzed the modeling data, the unit root test results show that the

data was stationary (please see Table 1) and the autocorrelation and partial correlation coefficients of modeling data at lag 12, 24 and 36 were not obviously large(please see Figure below, which indicated that the seasonality of the data was not apparent), and the time series of tuberculosis incidence in Kashgar can not show the obvious seasonal characteristics either(please see Figure 2), so we considered doing ARMA model analysis (ARMA model is often used to do prediction analysis of the stationary data).

The autocorrelation and partial correlation graphs of the data showed that the

Autocorrelation coefficients of the lag orders were obviously trailed, and the partial correlation coefficients were truncated, therefore, the AR model was used in this study. The partial correlation coefficients were almost a second-order truncated distribution, only at lag 7, 8 and 9, the correlation coefficients were a little large, based on this situation and our experience, we considered establishing four models, such as AR(2), AR((1,2,7)), AR((1,2,8)) and AR((1,2,9)), where, AR((1,2,7), AR((1,2,8) and AR((1,2,9) were sparse models. Through the determination of these model parameters and the comparison of their AIC and SC (see Table 2), we finally established the AR((1/2/8) model.

Figure. Autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs of modeling data. As the delay of the lag order, the autocorrelation coefficients were trailing and the partial correlation coefficients were truncated, so it was suitable to establish the AR model.

Table. Autocorrelation and partial autocorrelation feature tables.

Models	AR(p)	MA(q)	ARMA(p,q)

Autocorrelation function (ACF)	trailing tail	Truncation	trailing tail
Partial autocorrelation function (PACF)	Truncation	trailing tail	trailing tail

Q2. The discussion was poor, which lacks comparisons of the prediction results to literature, please add.

Thank you very much for your guidance, we've enriched the discussion and added some comparisons of the prediction results to literature (please see the Discussion section of the revised manuscript, line 287-321). The main changes are as follows:

Many studies have found that Box-Jenkins method has a good ability of fitting and forecasting. For stationary time series that do not contain seasonality, it is more suitable to use the ARMA model of the Box-Jenkins method to do prediction analysis [35], for non-stationary time series of infectious diseases with obvious seasonality, it is more suitable to use seasonal autoregressive integrated moving average (SARIMA) model of the Box-Jenkins method for prediction analysis.[9-12] In our study, from Figure 2, we could see that the seasonality of the TB incidence in Kashgar was not obvious, there was only a certain seasonality from 2015 to 2017, and we found that the time series of TB incidence was stable by ADF unit root test, and the autocorrelation and partial correlation coefficients of modeling data at lag 12, 24 were not obviously large, therefore, for our research data, we used ARMA model to do forecast analysis, finally, we established AR((1,2,8)) model of the Box-Jenkins method, it has a good performance to fit and predict the TB incidence of Kashgar in Xinjiang. From Figure 2, we can also see that the time series of TB incidence has strong non-linear, the established AR((1,2,8)) model mainly extracted the linear information of data, considering that the neural network can capture the non-linear information of data well, in order to improve the prediction accuracy of TB incidence rate in Kashgar, we used AR((1,2,8)) model and Elman neural network model to establish AR-Elman hybrid model. Many studies have found that the combination model can improve the accuracy of prediction, such as, Wang et al.[28] found that SARIMA-NAR hybrid model has an outstanding ability to improve the prediction accuracy relative to SARIMA model and nonlinear autoregressive network (NAR) model when they were used to predict pertussis incidence in China. Li et al.[27] found ARIMA-GRNN hybrid model was shown to be superior to the single ARIMA model in predicting the short-term TB incidence in the Chinese population. Our research was consistent with these literatures that our AR-Elman hybrid model was more accurate than the single AR((1,2,8)) model.

The incidence of tuberculosis in Kashgar, Xinjiang is very high, and the relevant departments of disease prevention and control in Xinjiang have also done a lot of effective work. Our research was mainly to build a high-precision prediction model to help early warning of tuberculosis in Kashgar. Finally, we established the AR-Elman hybrid model, which had high fitting and prediction accuracy of TB incidence in Kashgar, Xinjiang. Because the development of any event may be affected by many factors, such as social, economic, political, demographic factors, so the long-term forecast accuracy of the AR-Elman hybrid model may decline.

Detailed comments:

Q1. There are many grammar editing errors that need to be corrected (e.g., B incidence in Kashgar. Root mean squared error (RMSE), mean absolute error (MAE) Lines 17-19 this should be linked with ‘and’; combined..with..., this should corrected as ‘ combined..and... ’ ;Line 35, it should be ‘ , respectively’; page10 Line22 ‘to found’? …and so on.

Thank you very much for reading our article carefully. We are ashamed of the mistakes we made. We have carefully corrected the mistakes in the article, and we have asked a native English-speaking colleague to help us checking the language and spelling mistakes of our article seriously.(please see revised manuscript)

Q2. Name must be capitalized, p-values is more common written in lower-case 'p'…

Thank you very much for your professional advice, we have made the corresponding changes according to your suggestion. Since there were p in the ARMA (p,q) model, in order to distinguish clearly, we have replaced the original Prob in this paper by p-value(please see revised manuscript, line 134,141,220,484 and 488).

.

Q3. Based on the goal and/or objectives of this study. Please add a paragraph to the introduction to clearly tell the reader the study progress in the methodologies used to model the infectious diseases. Of these methods, why do you use SARIMA model, why not others, this should be written clearly in introduction instead of in discussion). See STROBE reporting guidelines for more information on the importance of doing this.

Thank you very much for your professional advice, we have added a paragraph (please see revised manuscript, line 73-88), such as:

For study of quantitative prediction of infectious diseases, there are many methods, such as grey prediction method [5], exponential smoothing prediction method [6,7], dynamic model prediction method [8], Box-Jenkins method [9], neural network method [10], etc., with the deepening of prediction research, more and more scholars like to use the Box-Jenkins method [11-21], there are many different models in this method, and if appropriate models are established according to the characteristics of time series, high prediction ability often can be obtained. Neural network has strong nonlinear mapping ability, in which Elman neural network is composed of input layer, hidden layer, connection layer and output layer. The Elman network has dynamic memory function, and it is very suitable for time series prediction. At present, the Elman network is widely used in various fields, and has achieved successful prediction results. [22-26] Sometimes, the prediction effect of a single model is not ideal, in order to further improve the prediction accuracy, many studies adopt the combined model prediction method [27-29], the combined model can absorb the advantages of two or more methods so as to achieve a higher prediction accuracy.

Q4. The methods are very heavy in statistical analyses, with inadequate epidemiological details. The following information is required for transparency, replicability, and for readers to evaluate the rigour, merit, and potential biases in the study design.

According to your suggestion, we answered the questions Q4.1、Q4.2 and Q4.3 carefully, please see the answers to these questions below.

Q4.1. According to my experience, the performance evaluation indices (e.g. MAE, MAPE, MSE, RMSE and the like) should be utilized to evaluate the forecasting robustness instead of MAE and RMSE. Particularly the MAPE is a very important index to estimate the fitting and forecasting powers

Thank you very much for your professional advice, we have added MAPE as the performance evaluation index, (please see revised manuscript, line,10-11,15,17,196, 201, and 481-492)

Q4.2. Add one Table reporting all the estimated performance evaluation indices to the main text.

We have added table 3 reporting all the estimated performance evaluation indices(please see revised manuscript, line 261-262, and 490-492). Such as:

Table 3. Comparison results of in-sample fitting and out-of-sample forecasting performance for the AR((1,2,8)) model and AR-Elman model

Models	Fitted efficacy			Models	Forecasted efficacy		
	RMSE	MAE	MAPE		RMSE	MAE	MAPE
AR((1,2,8))	6.15	4.33	0.2585	AR((1,2,8))	10.88	8.75	0.2029
AR-Elman	3.78	3.38	0.1837	AR-Elman	8.86	7.29	0.2006

Q4.3. In Figs: what do the symbols/numbers/processes mean? please add to the caption as well as in the method descriptions. The same for all Figs. Please take more care for all the figures' caption.

Thank you very much for your professional advice, we have added captions of all the Figures (please see captions of all Figures in the revised article), such as:

Figure 1. The red part of this picture is the location of kashgar in Xinjiang, China.

Kashgar is located in the south of Xinjiang, and it has a very high incidence of tuberculosis.

Figure 2. Graph of the tuberculosis(TB) incidence in Kashgar from January 2005 to December 2017. The curve of TB incidence showed strong nonlinear characteristics from 2005 to 2014, and the TB incidence increased significantly from 2015 to 2017.

Figure 3. Autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs of modeling data. As the delay of the lag order, the autocorrelation coefficients were trailing and the partial correlation coefficients were truncated, so it was suitable to establish the AR model.

Figure 4. Autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs of residuals of AR((1,2,8)) model. Autocorrelation coefficients and partial correlation coefficients were almost in 95% confidence interval, so AR ((1,2,8)) model could extract the information of original data well.

Figure 5. The numbers of neurons in AR-Elman model and the corresponding root mean squared error (RMSE). When the number of neuron was 6, the RMSE was the smallest, and the AR-Elman model fitting ability was the strongest.

Figure 6. Fitting comparison graph of AR ((1,2,8)) model and AR-Elman model.
Red line stands for the original tuberculosis (TB) incidence curve, green line stands for AR((1,2,8)) model fitting curve, blue line stands for AR-Elman model fitting curve. The fitting ability of AR-Elman hybrid model was slightly better than that of the single AR((1,2,8)).

VERSION 2 – REVIEW

REVIEWER	Yongbin Wang Department of Epidemiology and Health Statistics, School of Public Health, Xinxiang Medical University
REVIEW RETURNED	07-Oct-2020

GENERAL COMMENTS	I can see that the authors have put exceptional effort into addressing the comments raised by reviewers. At this time, the paper has well improved. I am sorry but some minor concerns or comments of mine are needed to further improve the paper as I have many interests and experience use of ANNs. Certainly, these minor concerns can help to add a beautiful thing to a contrasting beautiful thing. Although this is a minor revision, I would still love to see how the author to address these concerns. So I would still love to review the revised version of the current manuscript again. 1. The author thinks that the unit root test results show that their data was stationary. I would love to know how the authors distinguish between trend-stationary and non-stationary. This is a very important issue as the trend-stationary is not the real 'stationarity' in the time series, namely, which often produces pseudo-regression, in fact, the ARIMA model belongs to a regression in nature. 2. Add a Figure displaying the Structure diagram of Elman 3. I did not find any comment on the limitations of the study. Include them in the Discussion 4. This study looks interesting to me, partially because of the worsening incidence in their study region. I would love to see some discussion on the possible reasons for the increasing incidence; however, it is a pity that the authors have not said anything about it, except for figures and charts in the Results section

	5. Please develop a way to add confidence intervals for the forecasts or the testing set. Otherwise, I think it is impractical to use these types of forecasts to guide public health decisions.

VERSION 2 – AUTHOR RESPONSE

Responses to Reviewers

To Reviewer 2:

I can see that the authors have put exceptional effort into addressing the comments raised by reviewers. At this time, the paper has well improved. I am sorry but some minor concerns or comments of mine are needed to further improve the paper as I have many interests and experience use of ANNs. Certainly, these minor concerns can help to add a beautiful thing to a contrasting beautiful thing. Although this is a minor revision, I would still love to see how the author to address these concerns. So I would still love to review the revised version of the current manuscript again.

From your comments we can feel that you are a very excellent expert of ANNs and ARIMA methods. We are very lucky to be able to get your guidance. We admire your rigorous academic attitude.

Thank you very much for giving us a positive evaluation, which is very encouraging, and thank you very much for your excellent and professional revision of our manuscript again.

We have studied your comments carefully and have made correction which we hope to meet with your approval.

Q1. The author thinks that the unit root test results show that their data was stationary. I would love to know how the authors distinguish between trend-stationary and non-stationary. This is a very important issue as the trend-stationary is not the real 'stationarity' in the time series, namely, which often produces pseudo-regression, in fact, the ARIMA model belongs to a regression in nature.

Thank you very much for your professional question. Our model sample data was TB incidence data from January 2005 to December 2016. As can be seen from Figure 3, the trend of TB incidence from January 2005 to December 2014 was not obvious, only one increase in 2015, and then the TB incidence from 2015 to 2016 fluctuated at a higher level, so the trend of overall sample data was not obvious, and the results of unit root test showed that p-value (0.0102) was less than 0.05(please see the follow Figure s), according to the above analysis, we judged that the data was stationary.

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.471272	0.0102
Test critical values:	1% level	-3.476805	
	5% level	-2.881830	
	10% level	-2.577668	

*MacKinnon (1996) one-sided p-values.

Figure s. The results of unit root test of sample data

If we can get the latest TB incidence data in future, and there is a trend in TB incidence data, then the data is non-stationary. For the non-stationary data, differential processing can be done to make the data stationary, then to establish ARIMA model or SARIMA model.

Teacher Wang's question is very professional, we are very sorry, our answer may not be perfect, but we still hope to get teacher Wang's support. We will continue to work hard in the field of statistics. We also hope that you can give us more guidance in the future study.

Q2. Add a Figure displaying the Structure diagram of Elman

Thank you very much for your professional suggestion, we have added the structure diagram of Elman, please see Figure 2.

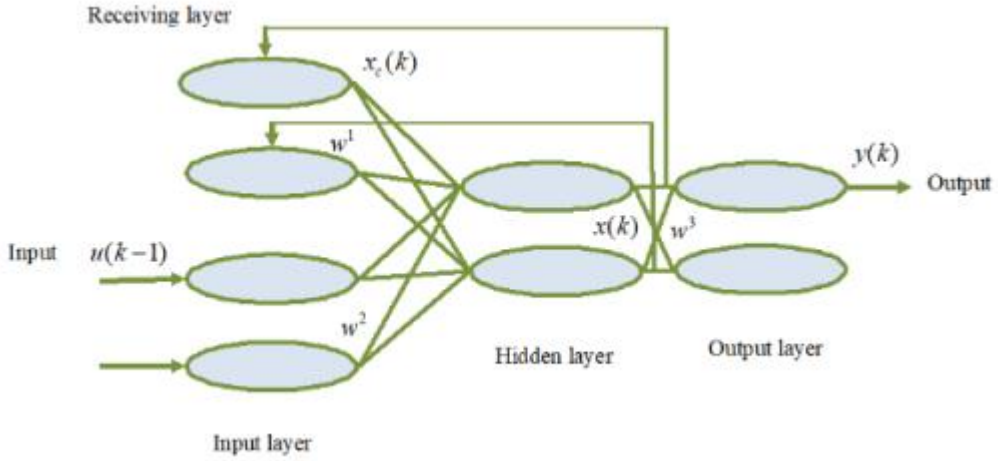

Figure 2. The Structure diagram of Elman neural network. w^1, w^2 and w^3 are the connection weight matrixes. $x_c(k)$ and x(k) represent the output of the contact unit and the hidden layer unit, respectively, y(k) represents the output of the output unit, u(k-1) represents the input of the input unit.

Q3. I did not find any comment on the limitations of the study. Include them in the Discussion

Thank you very much for your guidance. We agree with you, we have added any comment on the limitations of the study, such as:

Our study found that Box-Jenkins and Elman neural network hybrid method was an effective method for predicting the incidence of TB in Kashgar, it could provide a scientific reference for prediction analysis of TB incidence. However, our study also has some limitations: our method is only suitable for short-term prediction, long-term prediction performance will decline, two main reasons: first, our model was based on historical data characteristics; second, climatic factors, environmental factors, demographic factors and political issues may have certain impacts on the change of incidence. Therefore, if the established model becomes old and people want to obtain more accurate prediction results, it will be needed to adjust the model parameters, update the model based on the new modeling sample data, and then to do prediction analysis (please see the Discussion section, line 325-335).

Q4. This study looks interesting to me, partially because of the worsening incidence in their study region. I would love to see some discussion on the possible reasons for the increasing incidence; however, it is a pity that the authors have not said anything about it, except for figures and charts in the Results section

We are very sorry for your pity. In the revised manuscript, We have added some reasons for the increasing TB incidence and current situation of prevention and control TB in Xinjiang,such as：

In the past few years, Xinjiang's economic development was relatively backward, medical resources were scarce, diagnosis and treatment were delayed, the continuous spread of TB has become a difficult problem in the control of TB in Xinjiang. In recent years, Xinjiang has introduced many new policies to increase investment in TB prevention and control, and the relevant departments of disease prevention and control in Xinjiang have also done a lot of effective work, which has helped to control effectively the rapid increase of the TB incidence in Xinjiang. In order to do a good job in the prevention and control of TB in Xinjiang, many departments need to make joint efforts. Our research was mainly to build a high-precision prediction model to help early warning and prediction analysis of tuberculosis in Kashgar. (please see the Discussion section, line 313-322)

Q5. Please develop a way to add confidence intervals for the forecasts or the testing set. Otherwise, I think it is impractical to use these types of forecasts to guide public health decisions.

Thank you very much for your guidance, we agree with you. we have revised Figure 7, in Figure 7,we have added confidence intervals for the test set of AR-Elman hybrid model, please see the revised Figure7.

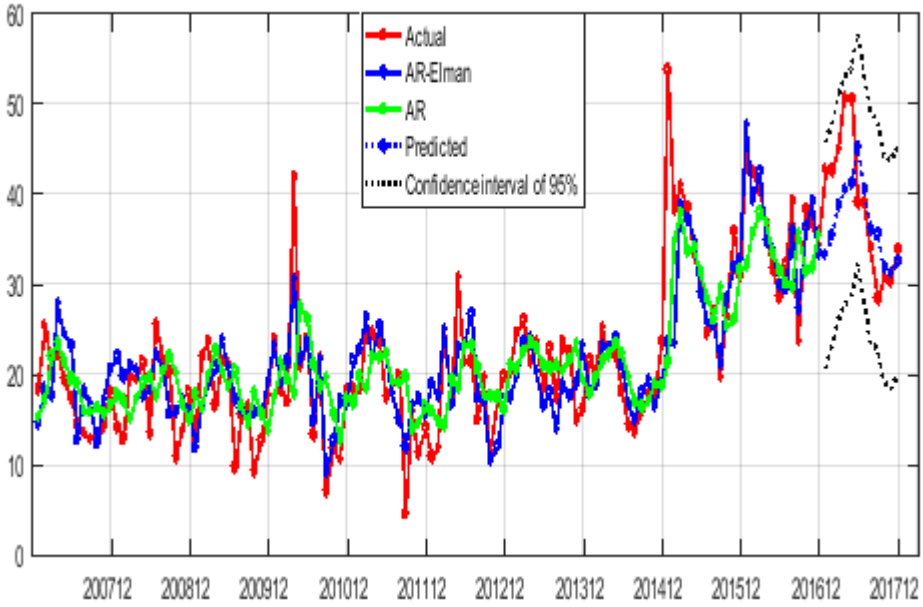

Figure 7. The fitting curves of AR((1,2,8)) model and AR-Elman model, and the prediction curve of AR-Elman model. Red line stands for the original tuberculosis (TB) incidence curve, green line stands for AR((1,2,8)) model fitting curve, blue line stands for AR-Elman model fitting curve. Blue dotted line stands for prediction curve of AR-Elman model, black dotted line stands for predicted curve of confidence intervals. The fitting ability of AR-Elman hybrid model was slightly better than that of the single AR((1,2,8)).