**ANSWER TO REVIEWER COMMENTS**

**TTRA-2018-0169**

The authors would first of all like to thank the reviewers for the positive and constructive feedback on our original manuscript. We believe that y our comments have enabled us to more clearly state our scope, results and conclusions and we hope we have reflected this to your satisfaction in our revised manuscript. The tables below address the specific comments. Correspondingly, some changes to the text of the original submission have been made and such changes are marked with red color.

**REVIEWER #1**

|  |  |  |
| --- | --- | --- |
| **#** | **Comment** | **Answer** |
| **1** | #1. This paper does not provide very convinced literature review to indicate the background why the CNN model is suitable to forecast the traffic flow, is only that the CNN right now a popular tools? Please provide more convinced relevant literature review to support the very reason why CNN is suitable to be employed to traffic flow forecast. | Thank you for pointing out this issue.  We have added more literature review on CNN for traffic prediction.  As a Representative deep learning method, CNN is widely used in computer vision and image classification (Krizhevsky et al., 2012). In transportation domain, Ma et al. (2017) has demonstrated that convolutional neural network (CNN) is suitable for traffic speed prediction, and after that, Du et al. (2018) uses CNN as a first layer to capture the features of different modality traffic data.  Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. *Sensors*, *17*(4), 818.  Du, S., Li, T., Gong, X., Yu, Z., & Horng, S. J. (2018). A Hybrid Method for Traffic Flow Forecasting Using Multimodal Deep Learning. *arXiv preprint arXiv:1803.02099*. |
| **2** | #2. Section 2.1 should move to the relevant literature review section, due to it divides the traffic flow conditions have temporal and spatial characteristics. | Thank you for pointing out this issue. Section 2.1 had moved to the relevant literature review section and some of statements have been improved. |
| **3** | #3. In addition, in page 6, please provide more statements for the contributions of this paper, at least 150 words for each contribution point. | Thank you for pointing out this issue.  In light of the comments of other reviewers, we restated the contributions of this paper as follows:   1. This paper analyses the spatio-temporal relationship of traffic flow and proposes a spatio-temporal feature selection algorithm. 2. We propose a CNN predictive model based on spatio-temporal correlations.   For the first point, as described in the introduction section of the article, the traffic flow contains rich spatio-temporal characteristics, and how to effectively use time and space traffic flow data is worth studying. The prediction algorithm considering single dimensional data will lose the distribution and transition relationship of the traffic flow, which limits the improvement of prediction accuracy. Most of the research is only for the establishment and adjustment of the prediction algorithm model, and there is little research on the influence of the input data selection on the prediction accuracy. In this paper, by verifying the influence of input spatiotemporal data on prediction accuracy, an algorithm for input spatiotemporal data selection is proposed based on traffic spatiotemporal data analysis. It can maximize the accuracy of the model prediction and the effectiveness of the prediction algorithm within an acceptable computation time.  For the second point, the deep learning method can efficiently deal with the multi-dimensional, large-scale, complex nonlinear relationship of traffic data and has strong learning ability. Therefore, in recent years it has attracted a lot of attention in the field of transportation. As a representative deep learning method, convolutional neural networks and their variants have received continuous attention and sufficient applications. In this paper, we transform spato-temporal traffic data into grid data suitable for convolutional neural network processing. In addition, we elaborated on the loss function construction method for noisy traffic flow data. The operation process of the convolutional neural network and the selection of the model hyper-parameters are analyzed, and the detailed parameter configuration in the proposed algorithm is given. Finally comparing to other baseline models, the prediction accuracy of the proposed model turns out to be more efficient to learn spatio-temporal feature from the dataset. |
| **4** | #4. By the way, for the citation problem, line 31 on page 3, Yang and Zhu (1999) is lost in the reference list, please check it carefully. For the reference list problems, (1) please avoid citing working papers, such as Ermagun and Levinson (2016); (2) please also avoid citing meeting conclusion, such as Zhang et al. (2016); (3) please provide complete citation information, such as Kingma and Ba (2014), Wu et al. (2015), and Wu and Tan (2016); (4) please provide relevant DOI for those papers in press status, such as Xu et al. (2018) and Zhang et al. (2018). | Thank you for pointing out this issue.  We have fixed all the issues mentioned and checked the full text in the new manuscript. |
| **5** | #5. Authors should provide the very details illustrating how the proposed model is working in the experimental results section, i.e., lacking of some essential brief explanation vis-à-vis the text to indicate how the proposed methodology (Figure 1 in page 10) is working in the experimental results section. In addition, for Figure 3 in page 14, please also provide some necessary wordings to guide readers to understand what authors have done and obtained from Figure 3. For Eq. (11) in page 18, please explain why this equation should be involved to be as the loss function; similarly, for Eq. (12) in page 19. Algorithm 1 in pages 26 and 27 is not introduced in text, please provide some necessary illustrations for Algorithm 1. | Thank you for pointing out this issue.  Since we have reorganized the full text, the figure number and formula number of the latest manuscript may differ from the original manuscript.  We have added the necessary explanations for the figures, tables, etc. in the articles mentioned. Figures 1 and 3 in the original manuscript have been changed to Figures 2 and 5. We have added a concise text narrative in the text and marked it in red.  For Eq. (11), Eq. (12) and Algorithm 1, we also added more explanatory text and marked it in red in the updated manuscript. We checked other figures, tables and formulas that were not mentioned. |
| **6** | #6.  For Table 5 in page 31, authors should try to conduct some statistical test to verify the significance of the forecasting performance from the proposed approach. Without the significant test, this paper only has minor contribution. Please refer Diebold and Mariano (1995) and Derrac et al. (2011). F. X. Diebold and R. S. Mariano, “Comparing predictive accuracy,” Journal of Business & Economic Statistics, vol. 13, No. 3, pp. 134-144, 1995. Derrac, J.; García, S.; Molina, D.; Herrera, F. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. Swarm Evolutionary Computation 2011, 1, 3–18. | Thank you for indicating this issue.  Our foregoing analysis dose not answer the question of how much statistical confidence we have in saying our proposed method is “better” than other methods on the basis of MAPE. So we perform a two-tailed binomial test pairwise comparison between the proposed algorithm and other algorithms. An *α*=0.05 significance level was used for hypothesis testing. The two-tailed binomial test was carried out in the statistical software R. The detailed process is referred to Derrac et al (2011) and given in the last paragraph of Section 3.4 of the updated manuscript.  Derrac, J.; García, S.; Molina, D.; Herrera, F. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. Swarm Evolutionary Computation 2011, 1, 3–18. |

**REVIEWER #2**

|  |  |  |
| --- | --- | --- |
| **#** | **Comment** | **Answer** |
| **1** | #1. First, the paper is unnecessarily lengthy, and the authors should consolidate the presentations greatly for concisely demonstrating the studies. | Thank you for pointing out this issue.  We have simplified the language of the full text, and deleted the redundant part. In order to explain the process of research in more detail, we added the necessary explanation for the diagram and table. |
| **2** | **2.   #**2. Second, the proposed methods and the comparative methods should be given in separate sections for clear presentation. | Thank you for pointing out this issue.  A brief introduction to each comparative model and detailed parameter settings have separately given in Section 3.2. |
| **3** | 3.     #3. Third, computational efficiency is important for online traffic prediction, and this should be investigated in this paper. | Thank you for pointing out this issue.  We analyzed the training time and online prediction time of the proposed model and comparison models, and presented in Section 3.4 of the revised manuscript. |
| **4** | 4.    #4. Finally, seasonal time series model and k-nearest neighbor model are two conventional short term traffic prediction models, and should be selected in this paper as comparative methods. | Thank you for pointing out this issue.  SARIMA is a linear model that is most widely used in the forecasting field, and KNN is also a representative non-parametric model. Both models are widely used in traffic prediction. We added these two models as a comparison model in the revised submission. The construction process of the model is referred to Williams and Hoel (2003) and Habtemichael and Cetin (2016) and detailed parameter settings have separately given in Section 3.2.  Williams, B. M., & Hoel, L. A. (2003). Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *Journal of transportation engineering, 129*(6), 664-672.  Habtemichael, F. G., & Cetin, M. (2016). Short-term traffic flow rate forecasting based on identifying similar traffic patterns. *Transportation Research Part C: Emerging Technologies, 66*, 61-78. |

**REVIEWER #3**

|  |  |  |
| --- | --- | --- |
| **#** | **Comment** | **Answer** |
| **1** | #1. The contributions of the work in the context of existing literature is not clear. Presently, the contributions stated are:  *(1) A framework for traffic flow prediction, (2) an optimal input data selection algorithm, and (3) a CNN prediction model based on spatial-temporal correlations.*  (1) & (2) are not contributions to knowledge. (3) can be a contribution provided it is established through literature. So, please rephrase and be specific about the contributions. | Thank you for pointing out this issue.  Contributions are updated as below:  (1) Traffic flow spatio-temporal relationship analysis and proposed a spatio-temporal feature selection algorithm.  (2) A CNN prediction model based on spatio-temporal correlations.  We have added more literature support for contribution point 1, and strengthened the theoretical basis and simulation analysis in the choice of spatiotemporal data. We describe the analysis and implementation of the algorithm in detail in Section 2.2 of the updated manuscript. It can be considered as a contribution of this paper. |
| **2** | #2. The matrix presentation methods described for choosing the temporal and spatial input points seems heuristic /arbitrary. The ACF plot presented is for non-stationary data and 0.68 cannot be arbitrarily chosen as a cut-off! The spatial correlation method described is quite vague. How are the datasets are actually matched? Pearson’s correlation? What about lagged spatial correlation?? | Thank you for pointing out this issue.  In a practical rode traffic prediction task, there are two problems we must face with: (1) Our data sets is only a sample of real data and the actual data full distribution is not available to the predictive model; (2) There are too many combinations of input features, and most practical algorithms attempt to find a hypothesis distribution by approximating NP-hard optimization problems. Due to above reasons, we propose a spatio-temporal feature selection algorithm with respect to a particular prediction algorithm, taking into account its heuristics, biases, and trade-offs.  This is true, 0.68 cannot be selected as a truncation, and we have re-described the correlation analysis in the updated manuscript.  The cross-correlation analysis in the original manuscript does not reflect the correlation between the different observation points well, so we replaced it with the heap map of Figure 1(b). For the spatial dimension matching problem, if we do not found a significant prediction error reduction after *P* expansions length of spatial data, then we stop searching in that dimension, otherwise we will increase the data length in that dimension.  In our analysis, we do not consider the lagged spatial correlation, but assume that the CNN can handle the lagged spatial correlation, because a lagged spatial correlation between two adjacent observation points is represented in the spatiotemporal traffic flow matrix as two diagonally adjacent elements. |
| **3** | #3. In figure one the prediction methods contain a few boxes which do not show any sequence details of the process. It is necessary to provide more information in significantly more details on how the boxes are linked and what happens inside a box. For eg. ‘Determine network structure’ Does not give information about which network, how to structure is determined, what elements of the structures are determined at every step of prediction or during the training stage. It is also unclear whether all these steps are necessary to be followed during the production process or just during the training of the network. | Thank you for pointing out this issue.  Since we have reorganized the full text, the figure number and formula number of the latest manuscript may differ from the original manuscript.  We have added the necessary explanations for the figures, tables, etc. in the articles mentioned. Figures 1 in the original manuscript have been changed to Figures 2. The inner boxes are the process of determining network hyper-parameters that can affect the results of neural network learning and the final regression prediction. The predictive model building process and the specific hyper-parameter configuration are detailed in Section 2.3.  We checked other numbers, tables and formulas that were not mentioned to avoid barriers to the reader. |
| **4** | #4. How is the initial correlation analysis for matrix presentation is linked with STFSA? Does it provide/generate Rinitial? Why two steps are necessary? Why not start with STFSA directly rather than using arbitrary correlation analysis? | Thank you for pointing out this issue.  We really don't need to add an arbitrary correlation analysis to our STFSA. We have revisited this part in the updated manuscript and added a detailed explanation.  In the original manuscript we just want to compare the input data without STFSA and the STFSA, but for now, the initial input data selection of 0.68 seems to be unreasonable. We reanalyzed and designed the STFSA algorithm and detailed in Section 2.3. |
| **5** | #5. Fig 6 presents some confusing results. It shows that with longer length of data the prediction accuracy decreases. This is counterintuitive. The MAPE should show that prediction accuracy increases when input data length increases between 5-20mins and then it levels out for further data points. Again, if the input length in increased significantly to up to 24 hrs or longer, we should see another improvement in prediction accuracy. Longer length of data should not affect the prediction accuracy. For spatial sections, it was initially described only 8 neighboring detectors were used…however the results are presented for nearly up to 30 detectors. Please explain. | Thank you for pointing out this issue.  First of all, we should explain that the first submission version does not clearly explain this part, which may lead to readers' misunderstanding. The conclusion of original manuscript in Figure 6 corresponds to the 5-minute traffic prediction problem for different prediction problems. If the predicted length changes, data selection and model training should be repeated.  For the accuracy decreases, as can be seen from Figure 3(b) of Chapter 2.2, taking NO 95 observation point as an example, for the 5-minute traffic flow forecast, the extended time lag does not bring about an effective improvement in improving the prediction effect for different spatial length. Further, if the time interval is increased, the prediction accuracy can be a bit reduced. Similar research phenomena can be found in Figure 8 of Du et al (2018). For a short-term traffic flow prediction problem, researchers often empirically choose a relatively short observation time such as Ma et al. (2017) and Polson and Sokolov (2017) for less than 40 minutes.  For other data sets, prolonging the length of the observations to more than 24 hours may result in an increase in accuracy, which may be due to the seasonality of the traffic data. But as Polson and Sokolov said in discussion, future traffic conditions are more similar to current ones as compared to those from previous days. Thus, allowed us to develop a powerful model by using recent observations as model features. Besides, our proposed STFSA can extend the prediction time lag up to 24h or even longer if the added data can bring error reduction.  D:\桌面\论文资料\文章用图\number of detector analysis.png  The picture above is Figure 6 in the original text. This is a curve in which the prediction error increases with the number of spatial nodes. We select 8 detection coils in advance in order to compare the prediction error of 8 detection coils with the prediction error after feature selection, and then explain the significance of feature selection. However, the selection process of 0.68 is not reasonable, and this part has been changed in the updated manuscript.  Du, S., Li, T., Gong, X., Yu, Z., & Horng, S. J. (2018). A Hybrid Method for Traffic Flow Forecasting Using Multimodal Deep Learning. *arXiv preprint arXiv:1803.02099*.  Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. *Sensors*, *17*(4), 818.  Polson, N. G., & Sokolov, V. O. (2017). Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies*, *79*, 1-17. |
| **6** | #6. Apart from the aforementioned issues, the rest of the results seem appropriate. But the STFSA which is claimed to be the main contribution of the work, does not seem to make a major difference to the MAPE error of the ANN or SVM accuracies published elsewhere. I understand that the model accuracy is dependent on the data characteristics to some extent, however STFSA framework is not making any significant difference. | Thank you for pointing out this issue.  It can be seen from Figure 3 in Chapter 2.2 that the temporal length and spatial length does have a large impact on the final prediction error value. We analyzed the short-term traffic flow prediction for 5-20 minutes, and finally we came to In conclusion, for different forecasting tasks, the impact of changes in time and spatial dimensions on the final forecast is unbalanced. For example, we found 5 minutes of traffic in different time and space data organization methods at different intervals of 20 intersections. The main source of influence of the flow prediction process error is the amount of spatial data. However, this situation is not applicable to the 10-15 minute traffic prediction problem, so we finally found that the necessary data selection process for different prediction tasks. It is indeed effective in reducing the error of prediction. We used six baseline models for comparative analysis during the experiment. The significance test results show that the proposed algorithm is effective. |
| **7** | #7. The input data size for ANN, SVR and CNN as it is needs to be presented and compared with ANN+STFSA, SVR+STFSA and CNN+STFSA. Computational time and efficiency needs to be discussed. | Thank you for pointing out this issue.  We conducted experiments on the traffic flow prediction of multiple road sections to verify the applicability of the proposed method to different road sections. In Chapter 3.3, we took the analysis coil of No. 95 as an example to give an analysis process for a specific road section. Parameter configuration, final use data for models and comparison models are detailed.  We analyzed the training time and online prediction time of the proposed model and comparison models, and presented in Figure 9 in Section 3.4 of the revised manuscript. |
| **8** | #8. The paper should be organized better. STFSA should be presented before other methods. Also, no need to give so much details of CNN, ANN or SVR. Provide appropriate reference. | Thank you for pointing out this issue.  We optimized the structure of the full text in the revised manuscript, and streamlined the linguistic expression of the full text, removed some of the redundancy, and added more explanations to the charts and tables to guide the reader through the work done by the author. In addition, the SARIMA and KNN models are added as two baseline models. A brief introduction and detailed parameter settings for each comparison algorithm are given in Section 3.2, and an appropriate reference is provided for each comparison algorithm. Besides, We believe that it is also important to deploy CNN in the transportation field, so the description of CNN is retained. |
| **9** | #7. The phrase ‘time correlation’ should be changed to ‘temporal correlation’. | Thank you for pointing out this issue.  We checked the full text and corrected the statement. |
| **10** | #10. Figure 1: There are some typos in this figure. The word person should be changed to Pearson. | Thank you for pointing out this issue.  We checked the full text and corrected the statement. |