A Causal, Theory-Informed Framework for Traffic Flow Forecasting

P. Fafoutellis[a,*], E. I. Vlahogianni[a]

[a] National Technical University of Athens, Athens, Greece
panfaf@mail.ntua.gr, elenivl@central.ntua.gr
* Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 29, 2024

Keywords: traffic forecasting, physics-informed neural network, granger causality

1 INTRODUCTION

Traffic forecasting using Deep Learning has been a remarkably active and innovative research field during the last decades (Vlahogianni et al., 2014; Yin et al., 2021). However, there are still several barriers to real-world, large-scale implementation of Deep Learning forecasting models, including their data requirements, limited explainability and low efficiency (Wang et al., 2019; Fafoutellis and Vlahogianni, 2023). In this paper, we propose a novel theory-driven framework that is based on a Granger causality-inspired feature selection method and a multitask LSTM to jointly predict two traffic variables. At the core of our methodology, there is a novel traffic flow theory-informed multitask neural network, which is used for the joint short-term forecasting of two traffic variables. For training the model, we propose a custom-made, theory-aware loss function, which incorporates the distance of the emerging multivariate forecast (pairs of traffic variables) from the fundamental diagram of the corresponding location. To enhance the model's performance and interpretability, network-level traffic information is selected from the most relevant locations, specific to each target location, using the Neural Granger adaptation of classic Granger causality.

The proposed approach is inspired by the concept of Physics- (or Theory-) Informed Deep Learning incorporates the advantages of the two main modeling approaches, namely analytical or physics-based and data-driven, in order to achieve better generalizability and higher accuracy (Di et al., 2023). Specifically, it can be deployed when a data-driven model includes, as input or output, one or more variables for which an analytical/mathematical relation is known from the corresponding scientific field. This relation is usually incorporated into the model's loss function or is used in an independent module and aims to adjust highly erroneous, unreasonable predictions of the model, towards what is "in theory" expected. In this case, the model's outcomes are assessed based on their distance from the observed (actual) outcome and the one expected by the relevant physics law (Usama et al., 2022).

2 METHODOLOGY

2.1 Neural Granger Causality

In this work, we exploit the concept of Neural Granger causality for the causal feature selection task, which is a Deep Learning adaptation of the classic Granger causality test, proposed by (Tank et al., 2021). The main advantage of this approach is that it overcomes the shortcomings of the classic one and, especially, the linear relations assumption, by using a neural network to simulate the relation between input and output variables. In our implementation, we use an LSTM network, which is more suitable for time series data, however, other architectures can be used as well.

The main idea behind Neural Granger is to train a neural network to estimate the target time series values getting, initially, as input all the other time series and, gradually, as the training process progresses, assign zero input weights to those time series that are not causally related with the target, using the following loss function:

$$\frac{1}{N}\sum^{N}(x_t - h(x_{i,<t}))^2 + \lambda\sum_{i=1}^{n}\|W_i^1\|) \qquad (3)$$

2.2 Traffic Flow Theory-Informed Loss Function

For the purposes of this work, we propose the Traffic Flow Theory-Informed loss function (TFTI loss), which combines the mean value of the MSEs' square root of both individual variables, which is proportional to the Euclidean distance of the predicted and the actual points, with the distance of the prediction from the closest point on the estimated fundamental diagram. Let $\hat{y}_i = (\hat{v}_i, \hat{s}_i)$ the prediction for a real pair $y_i = (v_i, s_i)$ and $g_j = (v_j^e, s_j^e)$ the closest point on the estimated fundamental graph which characterizes the location of interest, TFTI loss is defined as follows:

$$TFTI\ loss = \ a * \frac{1}{2}\sqrt{[MSE_v + MSE_s]} + (1 - a) * d(\hat{y}_i, g_j) \qquad (1)$$

where MSE_v MSE_s are the mean square error of the predictions of volume and speed respectively, $d(\hat{y}_i, g_j)$ is the Euclidean distance of the predicted pairs and the closest point of the fundamental diagram to the corresponding actual value. The factor a controls the significance of the second factor over the first, $a \in [0,1]$.

3 IMPLEMENTATION AND RESULTS

3.1 Data

The data used in this work were derived from a database that is maintained by the Traffic Management Center of the Region of Attica. The database includes daily measurements from about 1000 loop detectors that are installed in the road network of the broader area of Athens. The time resolution of the data is 6 minutes, i.e., the sensed data are aggregated in 6-minute intervals. The data refer to 40 days of measurements, between March 20th and April 30th of 2023. For the sake of brevity, 4 locations that lie close to the city center and significant economic/business areas were selected as the target locations.

To use the TFTI loss, one should first produce a functional form of the fundamental diagram by fitting a curve to the target location's measurements and develop a multitask prediction model, as both traffic volume and speed variables are considered in the loss function. More specifically, for the fundamental diagram of all locations (not only target locations), first, the (traffic volume, speed) pair measurements of each location are separated into congested and uncongested using the k-Means clustering algorithm. Then, a second-order polynomial curve was fitted at each cluster of each location, so, at each location, we have a fundamental diagram that consists of two curves (two-variate, discontinuous diagram).

3.2 Modeling Setup

The experiment aims to evaluate the effectiveness of the TFTI loss function in reducing the overall forecasting error (for both variables) but also increasing the trustworthiness of the model. The two approaches compared are training the model with (i)simple MSE error and (ii)proposed TFTI loss function, with parameter a=0.7, which corresponds to 70% weight for the first term (MSE) and 30% for the second (distance from the fundamental diagram). To have a valid and fair comparison, the two models that are developed for each target location (MSE and TFTI loss) are identical, except for the loss function: they have exactly the same architecture.

3.3 Detected Causalities

After applying the Neural Granger method on the available dataset, the locations that Granger-cause the target locations emerged. The average number of detected locations was 6.6 out of the 420 locations, which indicates that the dimensionality of the input space for each forecasting model can be vastly decreased, without, however, losing significant information. The locations that were found to Granger-cause the first 4 of the target locations are presented in Figure 1.

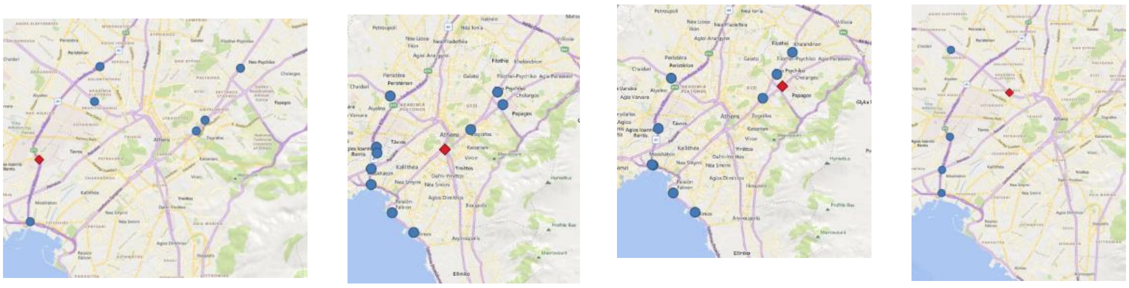

Figure 1 - Emerging locations (blue circle) that Granger-cause the target locations (red diamond)

3.4 Forecasting Results

In Table 1, the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE) of the predictions are presented for the two loss functions overall and for each variable separately for the four indicative locations.

Table 1 - Forecasting results comparison

	MSE loss	***TFTI loss***	*MSE loss*	***TFTI loss***
Location	**MS106**		**MS230**	
Volume MAPE	6.8%	9.1%	9.2%	10.5%
Speed MAPE	17.3%	6.4%	23.0%	9.4%
Overall MAPE	12.0%	**7.8%**	15.6%	**9.9%**
Location	**MS443**		**MS634**	
Volume MAPE	9.6%	13.0%	8.6%	9.7%
Speed MAPE	17.0%	9.3%	11.1%	6.7%
Overall MAPE	13.3%	**11.1%**	9.8%	**8.2%**

To further examine the trustworthiness of the models, we provide the fundamental diagrams emerging from the predictions. In Figure 2, the diagrams of an indicative location are provided.

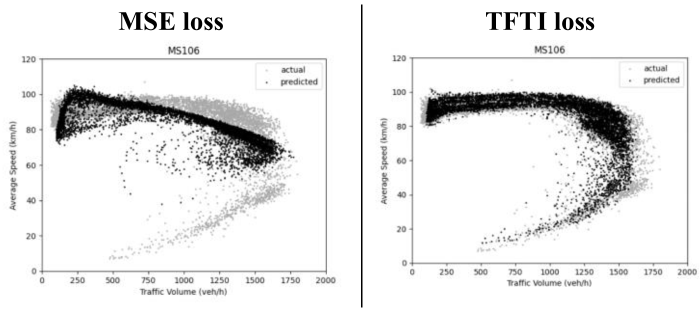

Figure 2 - Emerging predicted fundamental diagrams versus actual for MSE loss (left) and TFTI loss (right)

4 CONCLUSIONS

In this work, we proposed a holistic approach for actionable and trustworthy short-term traffic forecasting consisting of a causal Granger LSTM network and a multitask LSTM network, enhanced with the novel TFTI loss function. Findings showed that using the concept of Neural Granger for detecting causal features, the dimensionality of the input space can be significantly reduced. Moreover, the use of the TFTI loss was found to increase the overall accuracy of the two variables, as well as the trustworthiness of the predictions.

In the future, we aim to apply the proposed methodology to road networks of other cities, to examine whether similar results will emerge in terms of the forecasting performance and the detection of causal relations and associated traffic patterns. Furthermore, we will assess the transferability of the already trained models.

References

Di, X., Shi, R., Mo, Z., Fu, Y., 2023. Physics-Informed Deep Learning for Traffic State Estimation: A Survey and the Outlook. Algorithms 16, 305. https://doi.org/10.3390/a16060305

Fafoutellis, P., Vlahogianni, E.I., 2023. Unlocking the Full Potential of Deep Learning in Traffic Forecasting Through Road Network Representations: A Critical Review. Data Sci. Transp. 5, 23. https://doi.org/10.1007/s42421-023-00083-w

Tank, A., Covert, I., Foti, N., Shojaie, A., Fox, E., 2021. Neural Granger Causality. IEEE Trans. Pattern Anal. Mach. Intell. 1–1. https://doi.org/10.1109/TPAMI.2021.3065601

Usama, M., Ma, R., Hart, J., Wojcik, M., 2022. Physics-Informed Neural Networks (PINNs)-Based Traffic State Estimation: An Application to Traffic Network. Algorithms 15, 447. https://doi.org/10.3390/a15120447

Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2014. Short-term traffic forecasting: Where we are and where we're going. Transportation Research Part C: Emerging Technologies 43, 3–19. https://doi.org/10.1016/j.trc.2014.01.005

Wang, Y., Zhang, D., Liu, Y., Dai, B., Lee, L.H., 2019. Enhancing transportation systems via deep learning: A survey. Transportation Research Part C: Emerging Technologies 99, 144–163. https://doi.org/10.1016/j.trc.2018.12.004

Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., Yin, B., 2021. Deep Learning on Traffic Prediction: Methods, Analysis and Future Directions. IEEE Transactions on Intelligent Transportation Systems 1–15. https://doi.org/10.1109/TITS.2021.3054840