TRAFFIC STATE ESTIMATION WITH LOSS CONSTRAINT

Victoria Dahmen[1], Allister Loder[2], Gabriel Tilg[2], Alexander Kutsch[2] and Klaus Bogenberger[2]

***Abstract*— Traffic state estimation is relevant for real-time traffic control, providing travel information as well as for ex-post analysis of traffic patterns. While the output is usually the average speed and vehicle flow along street segments, the type of input data and the existing methods to obtain the output are diverse. Recently, physics-informed data-driven approaches started to emerge that enrich the estimation process with information taken from physical models. In traffic, so far, these have been the continuity equation and the fundamental diagram, designed to describe fully the traffic dynamics along links and corridors. In this paper, we propose a simpler and practice-ready physics-informed machine learning approach that informs the estimation through the well-established fundamental diagram in a loss constraint. It is designed for a link-level analysis where traffic homogeneity along the considered link is assumed. We apply the proposed method to full-trajectory drone data from Athens, Greece, demonstrate the applicability of our proposed approach, and point out its potential to future applications, e.g., a filter for control algorithms.**

I. INTRODUCTION

Traffic state estimation (TSE) is an essential part in transport planning and operations since it is required for effective design, evaluation, and selection of measures. The general aim is to infer the macroscopic traffic variables of flow, speed and density for a given point in time and space, by processing the available information from sensors such as loop detectors or probe vehicles.

Classical TSE was first developed for freeways. A well-known method is the adaptive smoothing method [1] that estimates traffic states between observation points and reduces the measurement noise. Further approaches are based on the extended Kalman Filter (EKF) and loop detector data (LDD) [2]; an EKF, LDD, and a multi-class first-order traffic flow model [3]; a Kalman Filter and floating car data (FCD) [4]; and an EKF exploiting both FCD and LDD using the Godunov scheme [5]. Also, data-driven approaches with clustering and time series analysis were applied for state estimation and prediction [6]. EKF and particle filtering has also been used for online TSE for multimodal traffic [7]. In recent years, advances in machine learning and rich data availability enabled the deployment of powerful methods such as deep learning for TSE, e.g. [8]. A review on highway state estimation can be found in Seo et al. [9].

*Allister Loder further acknowledges the support from the German Federal Ministry for Digital and Transport for the funding of the project TEMPUS, grant no. 01MM20008K. Data source: open-traffic.epfl.ch (*pNEUMA*)

[1]Department of Informatics, Technical University of Munich, Garching, Germany v.dahmen@tum.de

[2]Chair of Traffic Engineering and Control, Department of Mobility Systems Engineering, School of Engineering and Design, Technical University of Munich, Munich, Germany allister.loder@tum.de gabriel.tilg@tum.de klaus.bogenberger@tum.de

In contrast to highways, urban areas are usually associated with substantially higher complexities such as traffic lights, public transport operation, slower road users, and the network structure itself. These aspects increase the complexity of TSE for urban areas. Similar to highways, the main data sources are still LDD and FCD (e.g., taxis and public transport vehicles). Respective methods were developed based on the Lighthill-Whitham-Richards model using artificial FCD [10], the Markovian queueing model [11], speed transition matrices [12], and multiple regression [13]. Recently, data from drones has become available which seems promising for TSE, e.g. [14]. A number of data-driven TSE methods for congestion detection are also reviewed in [15].

Lately, progress has been made in numerical methods for modeling physical systems with deep learning [16]–[19]. These advances allow the learning process to be regularized with the information contained in the differential equations of the physical system. As traffic flow dynamics can be described by differential equations, these new insights were applied to TSE [20]–[23] and it was shown that they can outperform comparable conventional methods in performance [24], [25]. Also, the fusion of LDD and FCD can be further improved using physics-informed TSE [26].

Current physics-informed approaches are a promising way to combine the best of both worlds - physical model-based and data-driven TSE techniques. The state of the art is heavily focused on integrating partial differential equations into data-driven estimation. However, these approaches' requirements likely stand against scaling it to entire networks. In this paper we contribute a simple approach that combines well-established physical traffic flow concepts and widely used data-driven techniques to develop an explainable physics-informed and data-driven TSE approach based on loss constraints. More specifically, we combine the fundamental diagram of traffic flow (FD) with multiple linear regression and neural networks. While existing methods describe the full complexity of traffic dynamics along links and corridors, we assume traffic homogeneity along links that simplifies computational complexity. Drone data providing a multimodal ground truth traffic state is used to mimic LDD and FCD data used in the TSE procedure, which ensures the transferability of our approach. As the used methods are well-known and data becomes increasingly available, our proposed approach is highly relevant for practice and science.

II. METHODOLOGY

A. Mapping function

As the idea of a loss constraint has received limited attention in the field of traffic engineering, we use two simple

mapping functions to predict vehicle flow $\hat{q}$ and vehicle density $\hat{k}$. More specifically, we introduce the idea of loss constraint TSE for a multi-linear regression (MLR) and a simple neural network (NN). However, the integration of a loss constraint-based TSE is possible for various machine learning methods.

The MLR models a linear relationship between the independent (input) variables x_i and the dependent (output) variable y. Each input variable is weighted by a factor β_i, resulting in $y = \beta_0 + \sum_i \beta_i x_i + \varepsilon$, where β_0 equates to the bias and ε is the error term. The MLR is implemented as a linear NN with no hidden layers.

The NN approach introduces non-linearity to the estimation. Instead of just one transformation from inputs to outputs, such as for the MLR, there are multiple layers that are separated by (non-)linear functions (activation functions). All weighting factors (weights) are learned from the losses incurred when training with a designated data set. Several parameters such as the learning rate, activation function, optimizer, and the architecture, i.e., number of layers and nodes, must be selected such that the network can best learn to model the inherent relationships reflected in the data.

B. Loss constraint

Loss constraint in general refers to a customized loss function that constrains the learning of the parameters of the mapping function. Physics-informed networks typically include the 'physics' component by modifying the loss function. This can be in the form of an equation, regularizing terms, or a constraint, as explained by Alber et al. [27]. For example, McGowan [28] tested a variety of physics-informed custom loss functions by adding formulae known from additive manufacturing to the regular loss. Jin et al. [29] also added a physical loss to the regular loss function to enforce boundary and initial conditions of the Navier-Stokes equation when modeling incompressible fluids. So far, their application in transportation research are very limited.

In this paper, we implement a loss constraint that aims to improve physical consistency in data-driven TSE by building on the well-known FD, e.g., [30]. Typically, we would expect data points to deviate from the theoretical FD and show scatter, as the stationary states described in the FD are rarely observed in complex urban settings. Yet, we would not expect a very strong deviation in the positive direction, as disturbances generally reduce rather than increase the traffic speed and flow. We use this knowledge to our advantage to improve the purely data-driven TSE by the MLR and the NN. Imposing a loss constraint could achieve fewer physically unreasonably high speeds, i.e., points with a very steep slope in Fig. 1, or flows, i.e., points that deviate substantially from the FD, which stem from noise.

We test the loss constraint approach with three common representations of the FD: the triangular, trapezoidal [31], and Greenshield's [30] model. Their functional form is shown schematically in Figure 1. We only explain in detail the loss constraint formulation for the triangular FD, as the general procedure for the other two representations is highly similar. The triangular FD $Q(K)$ can be described by Eqn. 1, where u_f is the free flow speed, K_c is the critical density, and w is backward wave speed. These parameters must be provided as input parameters to the model.

$$Q(K) = \begin{cases} u_f * K & : K < K_c \\ w * (K - K_c) + u_f * K_c & : K \geq K_c \end{cases} \quad (1)$$

The loss is implemented differently for the MLR and the NN. For the MLR, two models are trained in parallel to predict $\hat{q}$ and $\hat{k}$, whilst the observed values q and k are used to determine the losses. Each model only receives the loss relating to its output variable, i.e., flow ($\mathcal{L}^q_{MLR}$) or density ($\mathcal{L}^k_{MLR}$). In the NN, one model simultaneously predicts $\hat{q}$ and $\hat{k}$, thus they share a loss function, $\mathcal{L}_{NN}$. This function has two components as shown in Eqn. 2 and 6: the mean squared error $\mathcal{L}_{MSE}$ and the physical constraint $\mathcal{L}_{PHY}$. This loss function is modified for the subset of points in the sample where the predicted values, $\hat{q}$ and $\hat{k}$, and its corresponding observed values, q and k, are above $Q(K)$. The loss function is adjusted such that the corresponding point on the function is (partially) used in place of the observed value when calculating the Mean Square Error (MSE). The rationale is to encourage a state where these points are below $Q(\hat{k})$. This is expressed in Eqn. 2 for the MLR function, where the parameter γ signifies the amount by which the loss constraint $\mathcal{L}^q_{PHY}$ replaces the regular MSE loss $\mathcal{L}^q_{MSE}$. The loss functions for density $\mathcal{L}^k$ are formulated in the same way as for the flow.

$$\mathcal{L}^q_{MLR} = (1 - \gamma) * \mathcal{L}^q_{MSE} + \gamma * \mathcal{L}^q_{PHY} \quad (2)$$

Here, $\mathcal{L}^q_{MSE}$ is defined in the usual MSE formulation as shown in Eqn. 3.

$$\mathcal{L}^q_{MSE} = (q - \hat{q})^2 \quad (3)$$

The physics-informed loss $\mathcal{L}^q_{PHY}$ is defined as the squared difference between the flow on the FD $Q(\hat{k})$ at the predicted density $\hat{k}$ and the predicted flow $\hat{q}$ as shown in Eqn. 4, and similarly Eqn. 5 for $\mathcal{L}^k_{PHY}$.

$$\mathcal{L}^q_{PHY} = \left(Q\left(\hat{k}\right) - \hat{q}\right)^2 \quad (4)$$

$$\mathcal{L}^k_{PHY} = \left(K\left(\hat{q}\right) - \hat{k}\right)^2 \quad (5)$$

In the NN's loss function $\mathcal{L}_{NN}$, the single losses of flow and density for the MSE as well as the physics-informed loss are combined as shown in Eqn. 6.

$$\mathcal{L}_{NN} = \frac{1}{2n}\left(\sum_{i=1}^{n}\left((1-\gamma) * \mathcal{L}^q_{MSE,i} + \gamma * \mathcal{L}^q_{PHY,i}\right) + \sum_{i=1}^{n}\left((1-\gamma) * \mathcal{L}^k_{MSE,i} + \gamma * \mathcal{L}^k_{PHY,i}\right)\right) \quad (6)$$

where n is the batch size. As the network output is a tensor, which takes the form of $[q\ k]$, it is split into two flow and density tensors to calculate the individual losses.

All points in the sample that are below the FD are not subject to $\mathcal{L}_{PHY}$, i.e., $\gamma = 0$, as summarized in Table 1.

γ	**Predictions below curve**	**Predictions above curve**
Observed values below curve	0	0
Observed values above curve	0	[0,1]

TABLE 1: The factor γ is only applied to a subset of points.

Points that are above the FD and have their observed values below the FD, do not require a physics-informed loss, hence $\gamma = 0$ in Eqn. 2. In other words, the FD parameters must be defined in such a way that the FD presents the physically meaningful information that is required for the respective TSE application, e.g., filter for control algorithms. However, this paper focuses on showing the applicability of loss-constraining, while it is clear that the parameters of the FD and γ must be calibrated to each individual situation. Further points of research are suggested in the conclusion. In case of the piece-wise linear triangular FD, which is non-differentiable at the capacity, $\mathcal{L}_{PHY}$ will be calculated separately either for the case $k < K_c$ or $k \geq K_c$ as defined in Eqn. 1.

C. Alternative FD functions for the loss constraint

As previously mentioned, the loss constraint approach is also applied to Greenshields' quadratic FD and the trapezoidal FD. Both representations can be included in the loss constraint model in a similar fashion as the triangular FD. While the loss for the triangular FD is split into two differentiable segments, the trapezoidal FD is split into three. The Greenshields' approach is fully differentiable and thus does not require a case distinction with respect to density. However, two separate flow cases must be considered. When the flow is greater than the capacity (maximum flow), there is no corresponding density on the quadratic curve, thus the critical density is used. The issue of an FD not being fully differentiable could alternatively be resolved, e.g., by using a smooth maximum function [32] or a barrier function where applicable.

III. CASE STUDY

A. Data set

We apply the proposed loss constraint TSE method to empirical data to demonstrate its applicability in complex urban traffic situations. To do so, we use the pNEUMA data [14]. It stems from video recordings of a swarm of drones

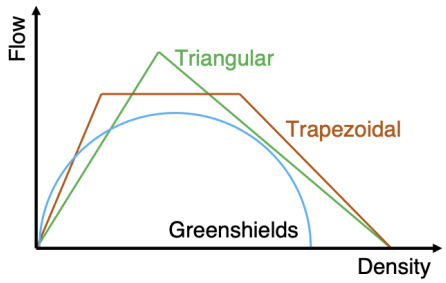

Fig. 1: The three shapes of the FD used in this analysis.

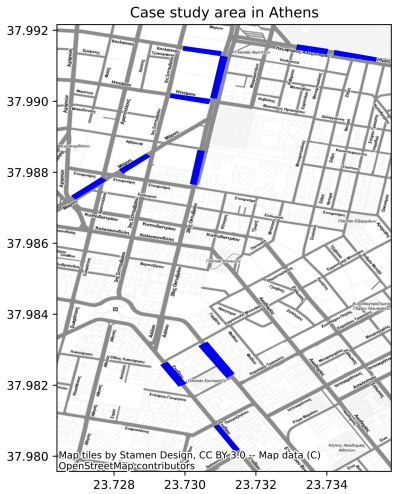

Fig. 2: Map of the case study area in Athens. The selected links are highlighted in blue.

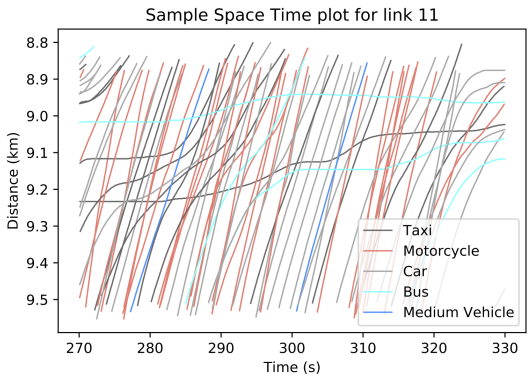

Fig. 3: A space-time plot of a short time interval for one of the road segments. It illustrates the speed variation between vehicle classes.

of an area of 1.3 km^2 in the congested city center of Athens. The resulting processed data set contains the trajectories of all detected vehicles including their speeds, vehicle class, and vehicle id. For our case study, we selected eleven one-way links from the recorded network, which are defined as road segments between two intersections. This is illustrated in Figure 2.

The observed traffic states include substantial scatters that reflect the complex dynamics and multi-modal heterogeneity of urban traffic during rush hours as seen in the time-space diagram in Figure 3. This ground truth data has the advantage over LDD or FCD that the traffic states and their distribution are entirely known, which facilitates the comprehensive understanding of the loss constraint mechanism.

B. Data pre-processing

The data was filtered to ensure that the selected links do not contain vehicles moving against the direction of travel. Additionally, any parking vehicles or objects classified as pedestrians or cyclists are removed. The traffic states are

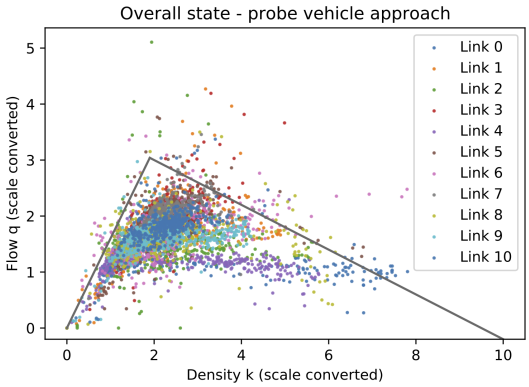

Fig. 4: A density-flow plot (probe vehicle approach) of the data set after link-specific scaling.

aggregated at 3-minute intervals to consider the general current state and to incorporate the effect of the signals.

C. Extracted variables

The observed outcome variables of the MLR and NN models, i.e., the traffic states (q, k), are computed using Edie's definition [33]. The speed is calculated using the fundamental equation of traffic flow $q = k * v$. In addition to the overall q and k, the mode-specific values are determined for the taxis and cars. Lastly, a virtual loop detector is placed on each link and the occupancy time and flow are used to calculate the density and flow at the LD's location along the link.

The characteristics of the considered links vary regarding the number of lanes and bus stops, and the presence of dedicated bus lanes. To account for these effects, a transformation is applied to the data set to collapse all subsets (one per link) into a single FD. First, the observed density and flow are normalized by dividing by the respective mean values: $k_s = k/\overline{k}$ and $q_s = q/\overline{q}$. Then the normalized flow values are divided by the link-specific free flow speed: $q_{s2} = q_s/u_f$. These three factors are later applied to the results of the MLR and NN. The resulting data is shown in Fig. 4.

The scatter in Figure 4 is informative. We find that it is common to a group of outliers above the FD that they correspond to traffic states with a high share of motorcycles. This points out to another application of the proposed method as it allows to project multimodal traffic states to a standardized unimodal FD, e.g., for traffic control purposes.

D. Model specification and estimation

The model specification of the MLR and NN includes the following input features: the existence of a dedicated bus lane, the number of bus stops and lanes, the link length, and the rank of the link. The latter describes the functional level of the link within the overall road network. Additionally, the observed flow and density (q and k) of the virtual LD, taxis, and cars at 5 % probe penetration are used. The outputs is the $\hat{q}$ and $\hat{k}$ of the overall traffic on a link.

The architecture of the NN contains seven hidden layers with 64 nodes each. The input layer has 11 nodes, while the output layer has two. The following parameters are used: learning rate [0.001], optimizer [Adam], activation function [LeakyReLU(0.01)], weight initialization [Kaiming], train-val-test split [0.7,0.15,0.15]. Furthermore, the training data was processed in batches of 256 rather than learning from each sample individually.

IV. RESULTS AND DISCUSSION

A. Results

The impact of the γ parameter of the physics-informed loss $\mathcal{L}_{PHY}$ is explored by testing values between 0 and 1, as it reflects the amount by which $\mathcal{L}_{PHY}$ replaces the MSE loss. Other sensitivity factors are left for further exploration, as we here focus on the feasibility of the proposed method. The higher the γ value, the more the model will avoid predicting values above the FD, as depicted in Figure 5, where the FD curve is set lower than in Figure 4 to emphasize the effect of changing γ. When it equals zero, the model will only train based on the MSE loss.

The evaluation of the average results across 20 runs (experiments) will be shown in this section. There is no notable change in the accuracy of the immediate model outputs (flow and density), as there are relatively few points above the FD when considering the whole data set, i.e., affected by the physics loss. A larger difference is, however, observed for the speeds associated with those flow and density values. The speeds derived from the model predictions ($\hat{v} = \hat{q}/\hat{k}$) are compared, first, to the observed speed (v), which reflects the mean actual speed on a link of the last 3-minutes. Second, they are compared to $v_{target} = q/k$, which reflects the speed obtained from the models' input flow and density. v and v_{target} differ slightly, as the data is resampled from 30 seconds to 3 minutes to counteract the effect of traffic lights. We expect the model to estimate v_{target} better, yet v reflects the actual traffic state more accurately.

In the case of the MLR model using the triangular FD, the RMSE of the v_{target} and v improves by 2.3 % and 1.7 % for a γ value of 0.6 compared to a value of 0, respectively. The slight improvement in the R2 value and the expected reduction in the number of predictions above the FD are also shown in Table 2. More γ values were tested and consistent with the above trends; this also applies to the NN.

The Greenshield approach improves the RMSE of the speed estimates most for $\gamma = 0.6$, by about 7.0 % for v_{target} and 7.4 % for v. For the trapezoidal FD, there is an improvement of 6.5 % for v_{target} and 5.6 % for v for $\gamma = 0.6$. The number of points above the FD is reduced, as expected.

γ	**RMSE** v_{target}	**RMSE** v	**R2** v_{target}	**R2** v	**Max.** $\hat{v}$	**Num. preds above FD**
0	3.89	4.29	0.891	0.855	58.8	12.2
0.3	3.84	4.20	0.892	0.857	58.9	10.4
0.6	3.80	4.22	0.894	0.858	53.7	9.0

TABLE 2: Results for MLR with triangular FD. Mean over 20 runs.

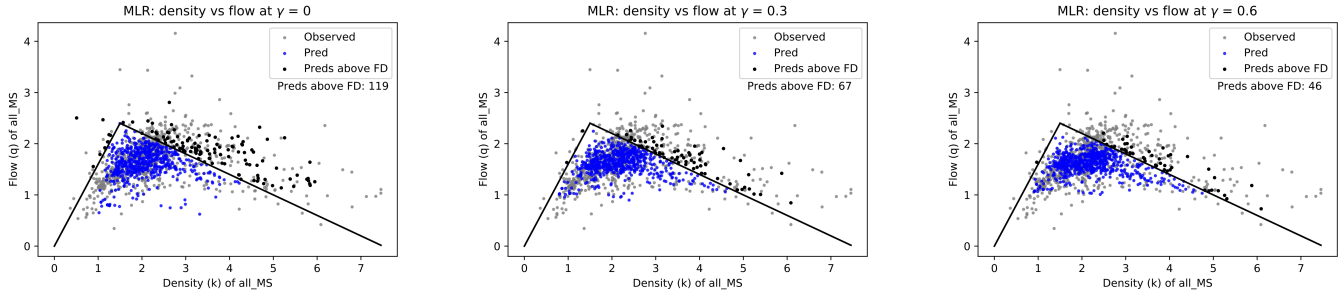

Fig. 5: Sample plots of the output and observed flow and density for different γ values. The FD is set much lower in this example to emphasize the effect and behavior of the proposed physics loss.

With regards to the NN, the triangular FD shows the most promising results. These are displayed in Table 3. The mean maximum predicted speed per run is lowest for $\gamma = 0.6$, decreasing from 76 km/h to 54 km/h. This is also where the improvement for the RMSE of v is greatest at 10.4 %. The error of v_{target} is reduced by 10.5 %. Improvements are also observed for the R2 values (3 % each) and the number of points above the FD.

The Greenshield method shows a 9.4 % and 9.9 % improvement for the RMSE of v_{target} and v at $\gamma = 0.3$, and has a reduced number of predictions above the FD. The R2 also slightly improves for both speeds by 2 %. Lastly, when using the trapezoidal FD, for a γ of 0.3 there is an improvement of 3.2 % for the RMSE of v_{target}. The other RMSE remains roughly the same, yet the R2 improves. As in all the other cases, the number of points above the curve is reduced, on average from 29 to 16.

B. Discussion

During the in-depth analysis it is observed that the RMSE values improve and extremely high speeds occur less frequently when the physics-informed loss is applied. Points above the left-hand-side of the triangular FD are associated with high speeds and discouraged, as they are subject to the physics-informed loss. As this loss is a soft constraint, rather than a strict constraint that would force all points to be below the FD, it is still possible that (high) speeds above the FD occur.

In this case study the FD was set at a specific, but arbitrary location. In a calibrated model, one would want to stretch the actual FD curve, as the FD is typically lower than the curve illustrated in Fig. 4.

The most suitable γ value will vary for each case, too. However, it was observed that a $\gamma = 0.3$ improved the RMSE and R2 in all studied cases, a value of 0.6 only offered further improvement for some. Thus, our method outperforms the reference case where $\gamma = 0$, i.e., no physical loss constraint is applied. These values will also depend on whether the focus of the application is more on improving speeds or dealing with multi-modal traffic states. The latter is likely to benefit from a higher value, as this would further reduce the number of points that are above the curve.

γ	**RMSE** v_{target}	**RMSE** v	**R2** v_{target}	**R2** v	**Max.** $\hat{v}$	**Num. preds above FD**
0	3.67	4.0	0.903	0.880	76.8	57.1
0.3	3.60	3.93	0.908	0.883	67.2	34.4
0.6	3.29	3.58	0.932	0.908	54.6	32.6

TABLE 3: Results for NN with triangular FD. Mean over 20 runs.

In closing, the proposed method evidently improves the speed predictions' RMSE and R2 values, and also reduces the number of predictions above the FD, compared to the non-physics-informed case of $\gamma = 0$.

V. CONCLUSIONS

In this paper, we proposed a novel approach of physics-informed data-driven TSE. Using traffic data recorded during the pNEUMA experiment in Athens [14], we showed that the speed prediction can be improved when the FD is incorporated in the loss function. Compared to the idea of integrating both the partial differential equation and the FD in the loss function, the presented approach is a more direct and parsimonious one, as it considers only the latter concept. This has several advantages:

- The FD alone is a simple, widely acknowledged concept and explainable function, which can be estimated for almost every street using publicly available data and guidelines.
- It scales easily to an entire network as it assumes traffic homogeneity at the link level, which reduces the computational complexity of estimating gradients in space-and-time for every link.

It must further be explored how and with which sensitivity the parameters of the FD, the γ, along with the LD location and penetration rate, affect the exact behavior of the proposed model. Additionally, it can be compared to conventional TSE methods (e.g. EKF [34], [35]) other than the regular NN (when $\gamma = 0$).

Possible applications of the proposed method are its use as a filter for traffic control algorithms, where physics-informed input data is required for performance, or for improved detection of the critical density and flow. The loss constraint can, however, accommodate and communicate various types of information. For example, in terms of online or real-time estimation of (macroscopic) FDs, the patterns and time-series

of residuals could be informative of the prevailing traffic or congestion patterns [36]. This can be useful for the control of larger networks. Future research refers to the extension of the method to motorways, the incorporation of a multi-class FD that can explain the observed scatter inside the TSE algorithm, the consideration of loss constraints in other TSE methods, as well as the integration of the proposed method into real applications.

REFERENCES

[1] M. Treiber and D. Helbing, "An Adaptive Smoothing Method for Traffic State Identification from Incomplete Information," in *Interface and Transport Dynamics*, H. Emmerich, B. Nestler, and M. Schreckenberg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 343–360.

[2] Y. Wang, M. Papageorgiou, and A. Messmer, "Real-time freeway traffic state estimation based on extended kalman filter: A case study," *Transportation Science*, vol. 41, no. 2, pp. 167–181, 2007.

[3] J. Van Lint, S. P. Hoogendoorn, and A. Hegyi, "Dual EKF state and parameter estimation in multi-class first-order traffic flow models," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 14 078–14 083, 2008, 17th IFAC World Congress.

[4] J. C. Herrera and A. M. Bayen, "Incorporation of Lagrangian measurements in freeway traffic state estimation," *Transportation Research Part B: Methodological*, vol. 44, no. 4, pp. 460–481, 2010.

[5] Y. Yuan, "Lagrangian multi-class traffic state estimation," Ph.D. dissertation.

[6] C. Antoniou, H. N. Koutsopoulos, and G. Yannis, "Dynamic data-driven local traffic state estimation and prediction," *Transportation Research Part C: Emerging Technologies*, vol. 34, pp. 89–107, 2013.

[7] D. Ngoduy, "Applicable filtering framework for online multiclass freeway network estimation," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 2, pp. 599–616, 2008.

[8] Y. Han and S. Ahn, "Estimation of traffic flow rate with data from connected-automated vehicles using bayesian inference and deep learning," *Frontiers in Future Transportation*, vol. 2, p. 2, 2021.

[9] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura, "Traffic state estimation on highway: A comprehensive survey," *Annual Reviews in Control*, vol. 43, pp. 128–151, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.arcontrol.2017.03.005

[10] G. Hiribarren and J. C. Herrera, "Real time traffic states estimation on arterials based on trajectory data," *Transportation Research Part B: Methodological*, vol. 69, pp. 19–30, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S019126151400126X

[11] S. E. Jabari and H. X. Liu, "A stochastic model of traffic flow: Gaussian approximation and estimation," *Transportation Research Part B: Methodological*, vol. 47, pp. 15–41, 2013.

[12] L. Tišljarić, T. Carić, B. Abramović, and T. Fratrović, "Traffic state estimation and classification on citywide scale using speed transition matrices," *Sustainability*, vol. 12, no. 18, 2020. [Online]. Available: https://www.mdpi.com/2071-1050/12/18/7278

[13] L. Pun, P. Zhao, and X. Liu, "A multiple regression approach for traffic flow estimation," *IEEE Access*, vol. 7, pp. 35 998–36 009, 2019.

[14] E. Barmpounakis and N. Geroliminis, "On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment," *Transportation Research Part C: Emerging Technologies*, vol. 111, no. November 2019, pp. 50–71, 2020.

[15] N. Kumar and M. Raubal, "Applications of deep learning in congestion detection, prediction and alleviation: A survey," *Transportation Research Part C: Emerging Technologies*, vol. 133, p. 103432, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X21004241

[16] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations," *arXiv preprint arXiv:1711.10561*, 2017.

[17] ——, "Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations," *arXiv preprint arXiv:1711.10566*, 2017.

[18] ——, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[19] K. Um, R. Brand, Yun, Fei, P. Holl, and N. Thuerey, "Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers," 2020. [Online]. Available: https://arxiv.org/abs/2007.00016

[20] N. A. Akwir, J. C. Chedjou, and K. Kyamakya, "Neural-Network-Based Calibration of Macroscopic Traffic Flow Models BT - Recent Advances in Nonlinear Dynamics and Synchronization: With Selected Applications in Electrical Engineering, Neurocomputing, and Transportation," K. Kyamakya, W. Mathis, R. Stoop, J. C. Chedjou, and Z. Li, Eds. Cham: Springer International Publishing, 2018, pp. 151–173. [Online]. Available: https://doi.org/10.1007/978-3-319-58996-1_7

[21] J. Huang and S. Agarwal, "Physics Informed Deep Learning for Traffic State Estimation," *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*, 2020.

[22] Y. Yuan, X. T. Yang, Z. Zhang, and S. Zhe, "Macroscopic traffic flow modeling with physics regularized Gaussian process: A new insight into machine learning applications," pp. 88–110, 2021.

[23] J. Liu, M. Barreau, M. Čičić, and K. H. Johansson, "Learning-based traffic state reconstruction using Probe Vehicles," *arXiv*, 2020.

[24] R. Shi, Z. Mo, K. Huang, X. Di, and Q. Du, "Physics-Informed Deep Learning for Traffic State Estimation," 2021. [Online]. Available: http://arxiv.org/abs/2101.06580

[25] R. Shi, Z. Mo, and X. Di, "Physics-Informed Deep Learning for Traffic State Estimation : A Hybrid Paradigm Informed By Second-Order Traffic Models," 2021.

[26] F. Rempe, A. Loder, and K. Bogenberger, "Estimating motorway traffic states with data fusion and physics-informed deep learning," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 2208–2214.

[27] M. Alber, A. Buganza Tepole, W. R. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W. W. Lytton, P. Perdikaris, L. Petzold, and E. Kuhl, "Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences," *npj Digital Medicine*, vol. 2, no. 1, 2019. [Online]. Available: http://dx.doi.org/10.1038/s41746-019-0193-y

[28] E. McGowan, V. Gawade, and W. Guo, "A Physics-Informed Convolutional Neural Network with Custom Loss Functions for Porosity Prediction in Laser Metal Deposition," *Sensors*, vol. 22, no. 2, 2022.

[29] X. Jin, S. Cai, H. Li, and G. E. Karniadakis, "NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations," *Journal of Computational Physics*, vol. 426, no. Hui Li, 2021.

[30] B. D. Greenshields, "A study of highway capacity," *Highway Research Board Proceedings*, vol. 14, pp. 448–477, 1935.

[31] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B*, vol. 28, no. 4, pp. 269–287, 1994.

[32] L. Ambühl, A. Loder, M. C. Bliemer, M. Menendez, and K. W. Axhausen, "A functional form with a physical meaning for the macroscopic fundamental diagram," *Transportation Research Part B: Methodological*, vol. 137, pp. 119–132, 2020.

[33] L. Edie, "Discussion of traffic stream measurements and definitions," in *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, 1963, p. 139–154.

[34] C. Tampère and B. Immers, "An extended kalman filter application for traffic state estimation using ctm with implicit mode switching and dynamic parameters," 09 2007, pp. 209 – 216.

[35] C. P. I. J. van Hinsbergen, T. Schreiter, F. S. Zuurbier, J. W. C. van Lint, and H. J. van Zuylen, "Localized extended kalman filter for scalable real-time traffic state estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 385–394, 2012.

[36] L. Ambühl, A. Loder, L. Leclercq, and M. Menendez, "Disentangling the city traffic rhythms: A longitudinal analysis of MFD patterns over a year," *Transportation Research Part C: Emerging Technologies*, vol. 126, p. 103065, 2021.