

Estimating motorway traffic states with data fusion and physics-informed deep learning

Felix Rempe¹, Allister Loder² and Klaus Bogenberger²

Abstract—Traffic state estimation is an essential task in traffic engineering. It requires observations of traffic that are, so far, even with emerging technologies, only partially available at large, as neither Eulerian nor Lagrangian observations are available everywhere at all times. We propose a methodology to fuse both observation types using physics informed deep learning that is based on the Lighthill-Whitham-Richards (LWR) model to estimate traffic states at locations without observations, in particular to infer traffic density. We use two types of fundamental diagrams: Greenshields' parabola and a differentiable version of the trapezoidal fundamental diagram in the estimation. In the latter, we estimate from the observations the collective impact of all, even immeasurable, factors that lead to a reduction in traffic performance. We apply it to real-world data from the German motorway A9, where we find that it provides an opportunity to improve the estimation and understanding of traffic density by data fusion.

I. INTRODUCTION

Deep learning has undoubtedly changed the world of data [1], [2], most notably due to the understanding of parameter estimation in large neural networks [3]. Human mobility has always been a field with large-scale availability of data (e.g., [4], [5]) that naturally benefits from the methodology advances in deep learning. An essential task in traffic engineering is the estimation of traffic states. A traffic state commonly refers to the information on vehicle density, speed and flow at a specific location in space and time. Its estimation usually relies on partial observations as full information is impossible or too costly to obtain. Various methods have been proposed to estimate full traffic states from partial observations (c.f. [6] for a comprehensive review). There are two categories of approaches for traffic state estimation. First, analytical models, e.g. [7], [8], [9]. Second, empirical models, e.g., [10], [11]. Another important aspect in traffic state estimation is the data source for the partial observations. As many different measurement techniques, e.g., loop detectors, floating car data, camera data, exist, the task of *data fusion* is to combine heterogeneous data sources to improve traffic state estimation. There are different data fusion methods present in literature, c.f. [12], [13] for reviews, however, the focus seems to be primarily on travel times or speeds, and less on traffic density estimation. Lately, Cvetek et al. [14] emphasized that data fusion for traffic state

*Allister Loder acknowledges support from the German Federal Ministry of Transport and Digital Infrastructure (BMVI) for the funding of the project KIVI (Artificial Intelligence in Ingolstadt's Transportation System)

¹BMW Group, Munich, Germany felix.rempe@bmw.de

²Chair of Traffic Engineering and Control, Department of Civil, Geo and Environmental Engineering, TU Munich, Munich, Germany allister.loder@tum.de klaus.bogenberger@tum.de

estimation has seen a shift from pure statistical approaches towards machine learning, in particular deep learning with neural networks, e.g., [15], [16].

A key feature of neural networks is that they can model almost any function. When working with physical systems, there are usually constraints attached, e.g., speeds cannot be negative or flows cannot be at their peak at jam density. These constraints are not reflected in the neural network structure per se, unless they are part of the learning data, but which can be noisy or biased. Recently, advances in numerical methods for physical system modeling with deep learning emerged [17], [18], [19], [20] that allow to regularize the learning process with the information embedded in the differential equations governing the physical system. Traffic flow is such a physical system [21]. Consequently, these methodological advances were applied to traffic state estimation [22], [23], [24], [25], where Shi et al. [26], [27] and Yuan et al. [24] have shown that a physics-informed deep learning (PIDL) can outperform conventional methods for short road segments from a speed perspective.

Importantly, although some previous works already implicitly discussed data fusion (loop detector and floating car data), they all assume that traffic density is directly measured or can be uniquely derived from either speed or flow. However, neither of both is possible. First, traffic density cannot be measured with widely used measurement equipment, only speeds and vehicle flows are. Second, empirical data shows that flow (or speed) and density relationships are not smooth and exhibit scatter, e.g., resulting from hysteresis and capacity drop phenomena [28], [29] or heterogeneous traffic [30], either as different vehicle types are present or vehicles drive at different desired speeds. As traffic density is a key variable in the most common traffic flow models, e.g., the Lighthill-Whitham-Richards (LWR) model [31], [32], a discussion of this issue is important. As the previous work already showed the general benefits of PIDL for traffic state estimation, the next step is therefore to address the estimation of traffic density in real-world problems, where data from different sources is available. Note that traffic state estimation as defined here is a posteriori and is not a prediction of future traffic states.

In this paper, we contribute with a physics-informed deep learning approach for traffic state estimation that addresses data fusion for traffic density estimation as well as the heterogeneity in traffic states. We build upon the advances by Raissi et al. [19] and its early application to traffic state estimation by Shi et al. [26], [27]. We apply the proposed approach to empirical data from a motorway section on the

A9 in Germany. Comparing it to the adaptive smoothing method (ASM) and an unconstrained neural network, we find that simple PIDL estimates traffic states as good as ASM, but provides an opportunity to improve traffic density estimation from data fusion, while ASM and PIDL outperform an unconstrained neural network.

II. METHODOLOGY

A. Model structure

We propose to estimate the traffic state on the motorway section using separate neural networks for speed and flow as shown in Figure 1. The entire space-time domain is discretized into a grid with G elements (points or cells), where each element is uniquely identified by (x, t) . Flow measurements are taken at $O_q \in G$, resulting in $N_q = |O_q|$ measurements in total. Similarly, speed measurements from floating car data are available at $O_v \in G$, resulting in $N_v = |O_v|$ measurements in total. Location i for both type of measurements is defined by (x^i, t^i) .

We use the flow observations $q(x, t)$ with $(x, t) \in O_q$ to learn parameters θ_q that predict flow \hat{q} at all locations (x, t) . These parameters are estimated or learnt in the common fashion by minimizing the loss between the observed and predicted values.

$$\mathcal{L}_q = \sum_{(x, t) \in O_q} |\hat{q}(x, t) - q(x, t)|^2 \quad (1)$$

Similarly, we learn the parameters θ_v to predict \hat{v} at all locations (x, t) from the speed observations $v(x, t)$ with $(x, t) \in O_v$ and minimizing the L2 norm between observed and predicted speeds at O_v .

$$\mathcal{L}_v = \sum_{(x, t) \in O_v} |\hat{v}(x, t) - v(x, t)|^2 \quad (2)$$

Importantly, these two *measurement models* are trained to predict speed and flows only at the locations where each of them is observed, i.e., O_q and O_v , respectively. However, the parameters θ_q and θ_v should be able to predict speed and flow at all points G in a physically consistent manner. We ensure physical consistency by forcing predictions \hat{q} and \hat{v} for all points G to comply the Lighthill, Witham and Richards (LWR) model [31], [32]. In a basic formulation, the LWR model is the continuity equation as given by Eqn. 3 together with an assumption of a density-flow relationship $q(\rho)$.

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial q(\rho(x, t))}{\partial x} = 0 \quad (3)$$

This model requires an estimation of traffic density $\hat{\rho}$ at all grid points G that is obtained from the fundamental equation of traffic flow $\hat{\rho} = \hat{q}/\hat{v}$. The predictions of flow \hat{q} and speed \hat{v} result from the neural networks. This logic is shown in Figure 1. The physical consistency of the parameter estimates θ_q and θ_v for all points G is ensured by adding Eqn. 3 as a constraint. As illustrated in Figure 1, the numerical gradients

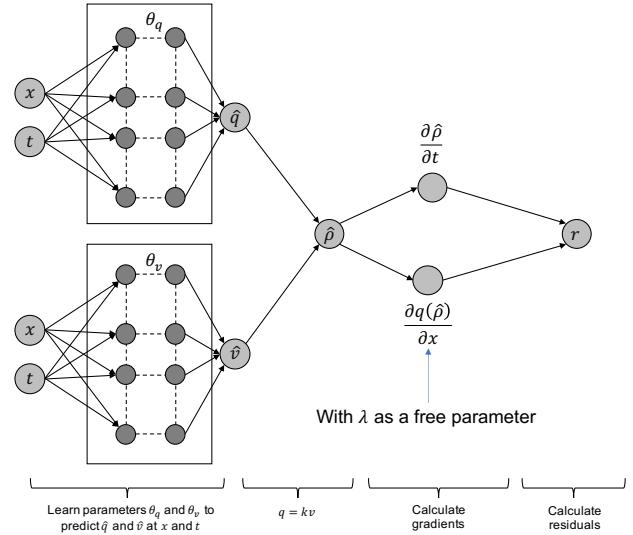


Fig. 1. Structure of the proposed model to fuse loop detector and floating car data using a physics informed neural network (PIDL).

$\partial \rho(x, t) / \partial t$ and $\partial q(\rho(x, t)) / \partial x$ are computed first. Then, the residual r of Eqn. 3 is derived as formulated by Eqn. 4.

$$r(x, t) = \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial q(\rho(x, t))}{\partial x} \quad (4)$$

The closer r is to zero, the better the estimates θ_q and θ_v satisfy the physical consistency formulated by the LWR model at all points G .

$$\mathcal{L}_{LWR} = \sum_{(x, t) \in G} |r(x, t)|^2 \quad (5)$$

B. Fundamental diagram

Previous work used for the fundamental diagram either Greenshields' seminal q - ρ parabola or a differentiable version of the triangular (piece-wise linear) flow-density relationship [26], [27]. Greenshields fundamental diagram as formulated in Eqn. 6 requires the road capacity q_{max} and the jam density ρ_{max} as parameters.

$$q(\rho) = \frac{4q_{max}}{\rho_{max}} \left(k - \frac{k^2}{\rho_{max}} \right) \quad (6)$$

Previous work did not use the piece-wise linear trapezoidal fundamental diagram as given in Eqn. 7 as used by [33] because it is not differentiable, although its physically meaningful parameters could be of value for learning. In addition to the parameters q_{max} and ρ_{max} already introduced for Eqn. 6, this fundamental diagram also considers the free flow speed u_{max} and the backward wave speed w on the link. Despite the additional parameters, this fundamental diagram is still a one-regime traffic model.

$$q(\rho) = \min\{\rho \cdot u_{max}, q_{max}, w(\rho_{max} - \rho)\} \quad (7)$$

We make this function differentiable and maintain its parameters' physical interpretation by using the λ approxi-

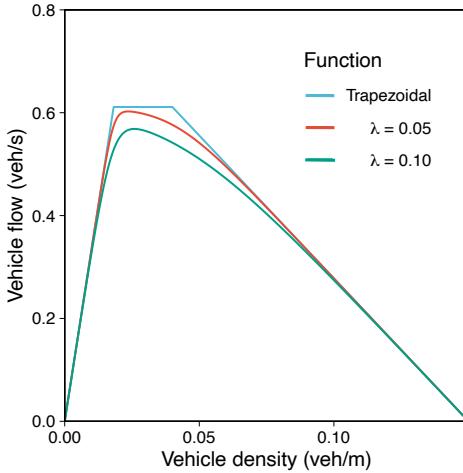


Fig. 2. The trapezoidal fundamental diagram with λ smoothing. The parameters for the trapezoidal fundamental diagram are $u_{max} = 33.6 [m/s]$, $q_{max} = 0.61 [veh/s]$, $\rho_{max} = 0.15 [veh/m]$ and $w = 5.55 [m/s]$.

mation as proposed by [34] for the macroscopic fundamental diagram. Eqn. 8 provides this smoothing to Eqn. 7.

$$q(\rho) = -\lambda \log \left(\exp \left(-\frac{\rho \cdot u_{max}}{\lambda} \right) + \exp \left(-\frac{q_{max}}{\lambda} \right) + \exp \left(-\frac{w(\rho_{max} - \rho)}{\lambda} \right) \right) \quad (8)$$

In Figure 2 we show the trapezoidal fundamental diagram as well as two smooth approximations with different values of λ . Note that in applications of this function, λ must not be constant and can vary over space and time according to traffic conditions. The parameter λ offers an additional degree of freedom, while the continuity equation for ρ is still satisfied. λ can capture heterogeneous traffic that results from the mixing of different vehicle classes or vehicles driving at different desired speeds. Importantly, λ can be estimated in the learning process, either as a constant value for the entire space-time domain or for small patches in space-time and thus can generate additional insights into the actual traffic behavior.

C. Loss function

Last, the loss function for the physical consistent learning process is the weighted sum of the three single loss elements as formulated in Eqn. 9. The parameters are estimated by minimizing the total loss \mathcal{L}_{total} . Note that we added λ to \mathcal{L}_{LWR} which means that λ is estimated in the entire learning procedure that minimizes \mathcal{L}_{total} . However, λ must not be present, either when Greenshields' fundamental diagram from Eqn. 6 is used or λ is fixed. The parameters α_q , α_v and α_{LWR} correspond to weights for the importance in the learning procedure.

$$\begin{aligned} \mathcal{L}_{total}(\theta_q, \theta_v, \lambda) &= \frac{\alpha_q}{|O_q|} \mathcal{L}_q(\theta_q) \\ &+ \frac{\alpha_v}{|O_v|} \mathcal{L}_v(\theta_v) + \frac{\alpha_{LWR}}{|G|} \mathcal{L}_{LWR}(\lambda) \end{aligned} \quad (9)$$

III. PIDL-BASED DATA FUSION APPLICATION

We test the proposed PIDL-based data fusion algorithm using data from a German motorway. We analyze its performance with respect to the two discussed fundamental diagram and compare the proposed algorithm to the adaptive smoothing method (ASM) (see [35] for details) and a simple neural network (NN) without physical constraint.

A. Data

We use traffic data from the northbound section of the A9 motorway from Munich to Nuremberg in Germany. This motorway section has a total length of 159 kilometers. In this analysis we focus on traffic data from May 29, 2019, on the first 30 kilometers north of Munich, where a rather high density of stationary detector location exists. The considered road section has a lane layout as shown in Figure 3a. The observed floating car and loop detector data is shown in 3c, respectively. A detailed description on the floating car data is given in [36] and for the loop detector data in [37]. We aligned the spatio-temporal resolution of both sources to have joint resolution of one minute in time and one hundred meters in space¹. Thus, the resulting space-time grid for the evaluation of the physical consistency has in total 119'498 points, where for May 29, 2019, speed measurements from floating car data is available at 45.9 % and flow measurements from loop detectors at 4.3 % of all points. Note that the available speed measurements provide the average speed of a subset of all vehicles, while the flow measurements are based on counting all vehicles.

B. Setup

For the parameters of the fundamental diagram we set $u_{max} = 130 [km/h]$, $\rho_{max} = 100 [veh/km]$, $q_{max} = 2500 [veh/h]$ and $w = 5 [km/h]$. These values are typical for German motorways [38]. Each of the two neural networks as shown in Figure 1 has three layers, each with 100 fully connected neurons. The hidden layers have a tanh activation function and the output is a softplus layer for strictly positive output. The learning procedure is implemented in python using the tensorflow framework, based on Raissi's seminal work [19]² using the ADAM optimizer and is initialized with Xavier. The iteration limit is set to 40'000 and the weights are $\alpha_q = 2$, $\alpha_v = 1$ and $\alpha_{LWR} = 100$. The ASM is parametrized according to [11].

To compare the four approaches, we divide the available data as shown in Figure 3 into training and a test samples.

¹The individual trajectories of a vehicle reporting location and speeds are converted into travel-times per grid cell of a 100m and 60s. In the case that multiple vehicles passed a grid cell, the respective mean of the travel times is used

²<https://github.com/maziarraissi/PINNs>

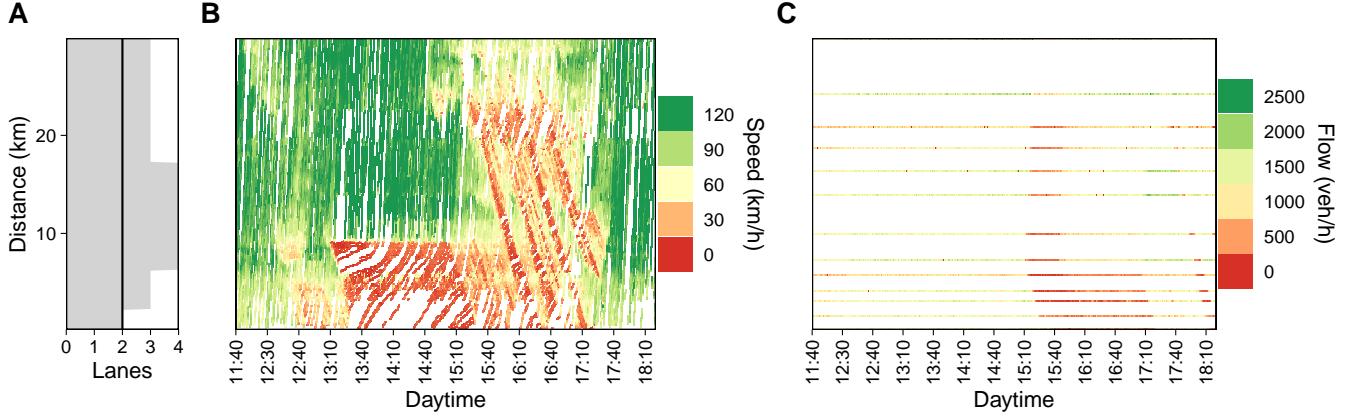


Fig. 3. Summary of the data used in this study from the German Motorway A9, bound from Munich to Nuremberg. Panel **a** shows the layout of the road section in terms of lane configuration; **b** depicts the available speed measurements from floating car data; **c** depicts the available flow measurements.

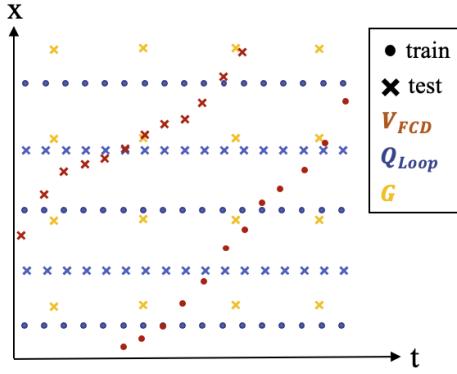


Fig. 4. Schematic representation of space-time distribution of data used for training (marked as circles, for both stationary loop detector and floating car data); and respective data for testing (marked as 'x') using raw speed and flow observations as well as regularly placed test locations (G) evaluating the physical compliance with the given FD.

The set of all available trajectories is split in half, using one half of the contained speeds for training of the network, and the other half for testing. Flow data of every second detector position along the road is used as test data. This gives the entire approach the structure as shown in Figure 4 with observation points as well as the grid evaluation points G for the physical compliance of the neural networks measurement models with the chosen fundamental diagram. We train the network on a machine with Intel Xeon CPU, 64GB RAM and an Nvidia Quattro RTX 4000. One training run takes approximately one hour.

C. Results

The estimation results of flow \hat{q} and speed \hat{v} in the entire considered space-time domain with the four methods (ASM, NN, PIDL-Greenhields and PIDL-trapezoidal) is shown in Figure 5. The results are based on the predictions on grid G using the trained network. Speed and flow errors are computed using the test data sets. The density error is computed using space-times for which test speeds as well as test flows are available, which allows to estimate

a density. Although the first impression suggests that overall all four methods result in similar results, differences are clearly visible. The traffic state estimation in Figure 5a-b based on ASM [35] clearly estimates traffic states along the propagation directions of free-flow and congested traffic states. The unconstrained neural network approach in Figure 5c-d that learns θ_q and θ_v without physical constraints by minimizing $\mathcal{L}_q + \mathcal{L}_v$ shows many inconsistencies in space-time, where locally large estimations errors could be expected. Contrary, the results of traffic state estimation with PIDL with Greenhields' fundamental diagram (Figures 5e-f) and with the trapezoidal fundamental diagram (Figures 5g-h) do not only show less inconsistencies in space-time, but also appear to have smoother transition phases. For the trapezoidal fundamental diagram from Eqn. 8, the learning based on minimizing Eqn. 9 finds $\lambda = 0.251$. Comparing ASM and PIDL, we find that especially for flow, PIDL appears to be less sensitive for small variations that lead to fine trails in ASM, but are dispersed in PIDL. Using the predictions for speed \hat{v} and flow \hat{q} , the predicted density $\hat{\rho}$ in the entire space-time domain can be computed. Figure 6 shows the estimation for the trapezoidal fundamental diagram. We find the expected, smoothed pattern, but importantly we can observe a remarkable difference between the two jams occurring at 10 km and 20 km that are not directly apparent from Figure 5. The first jam is associated with an accident near a motorway interchange, while the second stop-and-go jam is most likely resulting from the negotiations of the motorway interchange at 20 km. However, according to Figure 6, the mega jam at 10 km can be split into two phases from 13:10 to 14:10 and 14:10 to 15:10. Consequently, the PIDL-based data fusion as used here for traffic density estimation allows to advance the understanding and evaluation of congestion.

To understand the estimation results from another perspective, we derive the fundamental diagram for each estimation method for the entire space-time domain. The resulting fundamental diagrams are shown in Figure 7. While we find

Method	RMSE_v [km/h]	RMSE_q [1/min]	RMSE_ρ [1/km]
ASM	13.432	3.44799	12.975
Standard NN	14.465	6.43162	16.525
Greenshields FD	13.106	5.68028	13.239
Trapezoidal FD	12.673	4.98380	12.781

TABLE I

EVALUATION OF THE ESTIMATION ERRORS OF THE FOUR TRAFFIC ESTIMATION METHODS FOR SPEED v , FLOW q AND DENSITY ρ .

that all four methods recover a convincing free-flow branch from the data, the congested branch exhibits substantial differences. First, from all four methods, the ASM in Figure 7a recovered, as expected from its design, the upper limit of the congested branch convincingly. Second, the other three methods exhibit substantial scatter with the NN approach in Figure 7b having the most observations above the underlying fundamental diagram's upper limit. However, this is as expected as there is no constraint or limit imposed that could ensure such a behavior in the estimation. Third, the two PIDL approaches in Figures 7c-d still exhibit scatter. Future research will investigate likely causes for this behavior. It can be concluded that the PIDL approaches lead to a reduction in scatter compared to an unconstrained neural network and that in the present problem formulation the ASM's physical constraints seem to regularize the traffic state estimation too, if not better than the rather simple PIDL approaches.

Last, using the test data of stationary loop detector and floating car data (10 % of the entire sample), we can compare the accuracy of the reconstructed speed and flow in t and x as $\mathcal{L}_q^{\text{Test}}$ and $\mathcal{L}_v^{\text{Test}}$ of each estimation method. Table I summarizes the mean errors for flow, speed and density. We find that the standard NN approach has the worst performance, while ASM outperforms in the flow estimation and the two PIDL approaches provide the best speed and density estimates. The increased flexibility by the smoothed trapezoidal fundamental diagram lead to a decrease in the speed and density error. We expect that relaxing the constraint of a uniform λ in space and time will improve the estimation further. To summarize, we find that the proposed data fusion approach based on PIDL is at least as good as the existing the ASM approach, with chances to outperform especially in the speed and density estimation.

IV. CONCLUSIONS

In this paper we showed how advances in deep learning with physical constraints can be used for traffic state estimation, in particular for fusing stationary loop detectors (Eulerian observations) and floating car data (Lagrangian observations) to reconstruct traffic density along a long road section. We tested the physical-informed deep learning (PIDL) approach using Greenshields and a trapezoidal fundamental diagram to satisfy the continuity equation of traffic flow. We found that PIDL performs as least as good as ASM, but provides an opportunity to improve traffic density estimation. In particular, the integration of the LWR

model in deep learning shows that PIDLs integrates valuable information into the data fusion problem: qualitatively as the results are less prone to overfitting as well as the moving jams are reconstructed more accurately. This finding is confirmed by a quantitative evaluation of the estimated speed and flow values. Further, we found that using the smoothed trapezoidal fundamental diagram achieves higher values than the simpler Greenshields model. Arguably, using a flexible λ , the smoothed trapezoidal fundamental diagram is able to calibrate to the prevailing situation.

Future research will investigate how the traffic density estimation can be improved by allowing λ to vary over space and time. We expect that this will not only improve the quality of estimation, but also provides new ways to assess motorway performance. This can range from section control to effects of heterogeneous traffic. Further, the possibility to integrate existing standard data fusion approaches, e.g., [12], [13], deep convolutional neural nets, e.g., [39], or higher order macroscopic traffic flow models [21], [26] should be scrutinized. Finally, possibilities to use physics-informed deep learning for online traffic state prediction as well as congestion type identification should be explored as such an approach increases the compliance of results with physics.

In closing, we have shown in this paper that the opportunities provided by deep learning can prove useful for the analysis of traffic flow. In particular, as traffic density, almost immeasurable at large scale, can be inferred, but it has to be accepted that many parameters are estimated that bare any physical interpretation, at least for the time being.

REFERENCES

- [1] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. July, pp. 504–507, 2006.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.
- [3] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 9, pp. 534–546, 1986.
- [4] S. Çolak, A. Lima, and M. C. González, "Understanding congested travel in urban areas," *nature Communications*, vol. 7, p. 10793, 2016.
- [5] A. Loder, L. Ambühl, M. Menendez, and K. W. Axhausen, "Understanding traffic capacity of urban networks," *Scientific Reports*, vol. 9, no. 16283, 2019.
- [6] 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>
- [7] C. Nanthawichit, T. Nakatsui, and H. Suzuki, "Application of Probe-Vehicle Data for Real-Time Traffic-State Estimation and Short-Term Travel-Time Prediction on a Freeway," *Transportation Research Record*, no. 1855, pp. 49–59, 2003.
- [8] A. Aw and M. Rascle, "Resurrection of 'second order' models of traffic flow," *SIAM Journal on Applied Mathematics*, vol. 60, no. 3, pp. 916–938, 2000.
- [9] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: A general approach," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 141–167, feb 2005.
- [10] T. Schreiter, H. Van Lint, M. Treiber, and S. Hoogendoorn, "Two fast implementations of the adaptive smoothing method used in highway traffic state estimation," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 1202–1208, 2010.

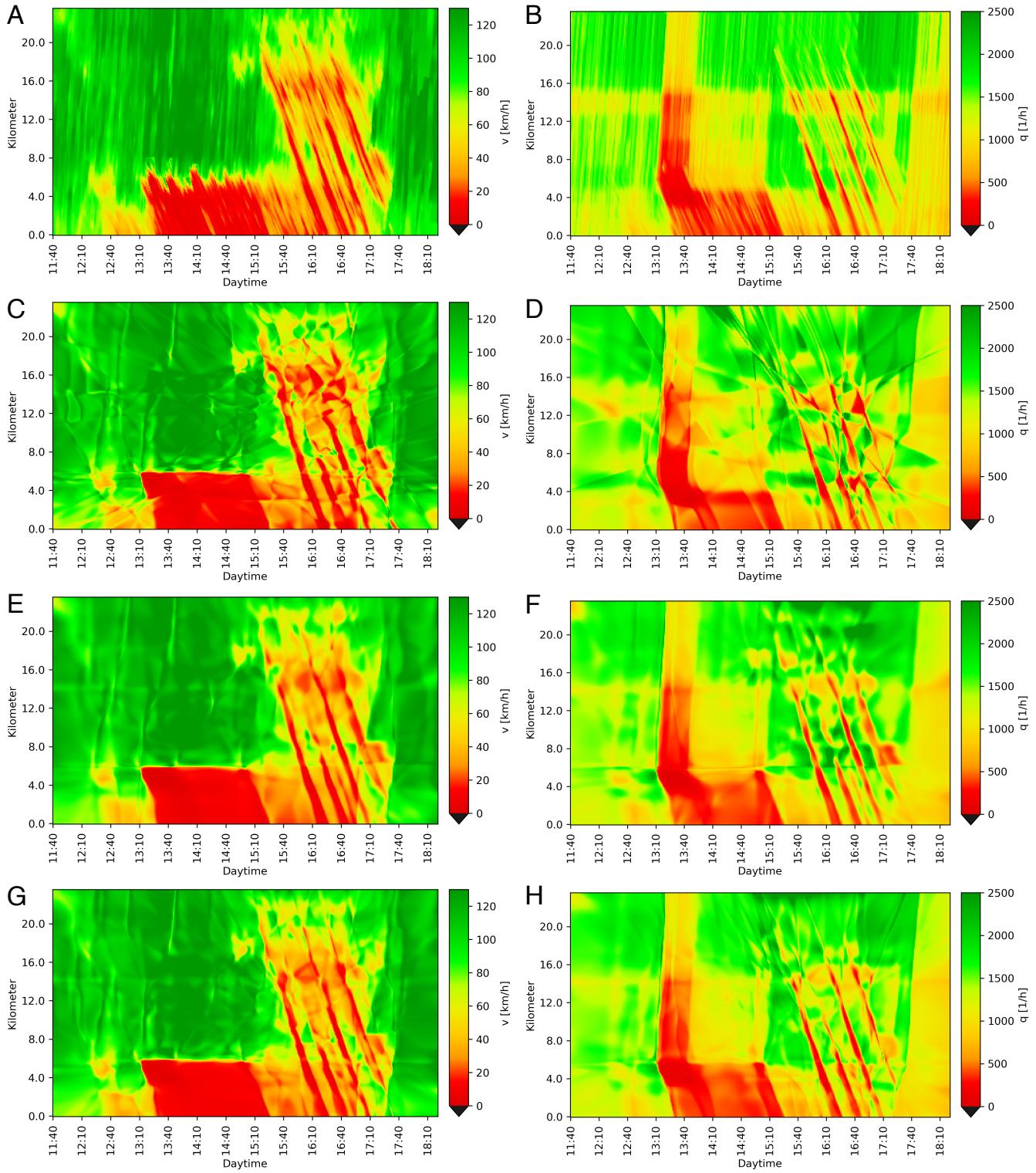


Fig. 5. Traffic state estimation results for speeds (left column) and flows (right column). Panels **a-b** result from the adaptive smoothing method (ASM); **c-d** from an unconstrained neural network; **e-f** from the PIDL using Greenshields' fundamental diagram; **g-h** from the PIDL using the trapezoidal fundamental diagram.

- [11] M. Treiber, A. Kesting, and R. E. Wilson, "Reconstructing the Traffic State by Fusion of Heterogeneous Data," *Computer-Aided Civil and Infrastructure Engineering*, vol. 26, no. 6, pp. 408–419, 2011.
- [12] C. Bachmann, B. Abdulhai, M. J. Roorda, and B. Moshiri, "A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 33–48, jan 2013.
- [13] N. E. E. Faouzi and L. A. Klein, "Data Fusion for ITS: Techniques and Research Needs," in *Transportation Research Procedia*, vol. 15.

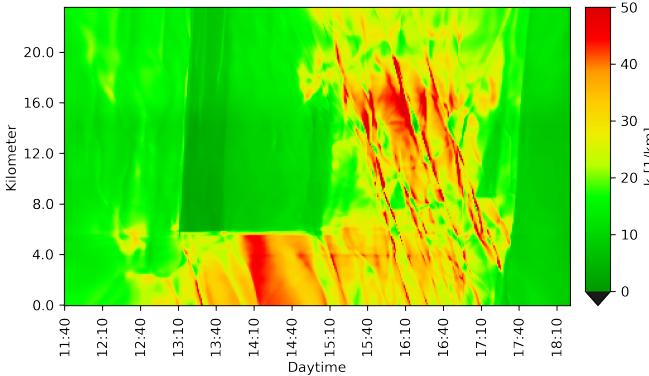


Fig. 6. Reconstructed traffic density field from the PIDL with the trapezoidal fundamental diagram with $\lambda = 0.251$.

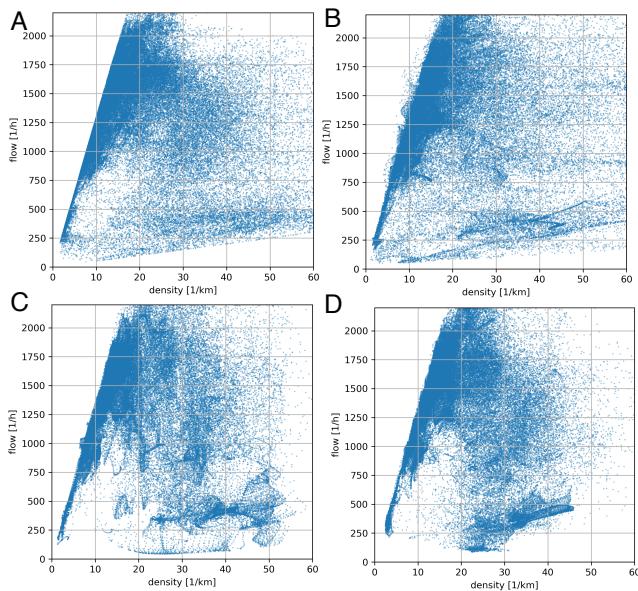


Fig. 7. Comparison of the resulting fundamental diagrams from the four traffic state estimation methods for the motorway section. Panel **a** shows the ASM-based fundamental diagram; **b** the standard, unconstrained neural network fundamental diagram; **c** the observed fundamental diagram from the PIDL with Greenshields fundamental diagram; **d** the observed fundamental diagram from the PIDL with the trapezoidal fundamental diagram.

Elsevier B.V., jan 2016, pp. 495–512.

- [14] D. Cvetek, M. Muštra, N. Jelušić, and L. Tišljarić, “A survey of methods and technologies for congestion estimation based on multisource data fusion,” *Applied Sciences (Switzerland)*, vol. 11, no. 5, pp. 1–19, 2021.
- [15] G. Yanjie, D. Fan, D. Hanxuan, P. Jiankun, and T. Huachun, “High Resolution Speed Estimation for Large-Scale Freeway Based on Data Fusion Technology,” pp. 654–664, mar 2021. [Online]. Available: <https://doi.org/10.1061/9780784482933.056>
- [16] J. Liu, J. Huang, R. Sun, H. Yu, and R. Xiao, “Data Fusion for Multi-Source Sensors Using GA-PSO-BP Neural Network,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–16, 2020.
- [17] 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.
- [18] ———, “Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations,” *arXiv preprint arXiv:1711.10566*, 2017.
- [19] ———, “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.

- [20] K. Um, R. Brand, Yun, Fei, P. Holl, and N. Thurey, “Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers,” jun 2020. [Online]. Available: <https://arxiv.org/abs/2007.00016>
- [21] M. Treiber and A. Kesting, *Traffic Flow Dynamics*. Berlin Heidelberg: Springer-Verlag, 2013.
- [22] 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
- [23] 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.
- [24] 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, feb 2021.
- [25] J. Liu, M. Barreau, M. Cičić, and K. H. Johansson, “Learning-based traffic state reconstruction using Probe Vehicles,” *arXiv*, 2020.
- [26] R. Shi, Z. Mo, K. Huang, X. Di, and Q. Du, “Physics-Informed Deep Learning for Traffic State Estimation,” jan 2021. [Online]. Available: <http://arxiv.org/abs/2101.06580>
- [27] 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.
- [28] M. Treiber, A. Kesting, and D. Helbing, “Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps,” *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 74, no. 1, pp. 1–10, 2006.
- [29] M. Saberi and H. Mahmassani, “Hysteresis and capacity drop phenomena in freeway networks: Empirical characterization and interpretation,” *Transportation Research Record*, no. 2391, pp. 44–55, 2013.
- [30] N. Webster and L. Elefteriadou, “A simulation study of truck passenger car equivalents (PCE) on basic freeway sections,” *Transportation Research Part B: Methodological*, vol. 33, no. 5, pp. 323–336, jun 1999.
- [31] M. J. Lighthill and G. B. Whitham, “On kinematic waves I. Flood movement in long rivers,” *Proceedings of the Royal Society (London)*, vol. A229, no. 1178, pp. 281–316, 1955.
- [32] P. I. Richards, “Shock Waves on the Highway,” *Operations Research*, vol. 4, no. 1, pp. 42–51, feb 1956. [Online]. Available: <https://doi.org/10.1287/opre.4.1.42>
- [33] 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.
- [34] 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.
- [35] 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.
- [36] F. Rempe, L. Kessler, and K. Bogenberger, “Fusing probe speed and flow data for robust short-term congestion front forecasts,” in *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 2017, pp. 31–36.
- [37] L. Kessler, B. Karl, and K. Bogenberger, “Congestion Hot Spot Identification using Automated Pattern Recognition,” *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*, 2020.
- [38] Forschungsgesellschaft für Straßen- und Verkehrswesen e.V., *Handbuch für die Bemessung von Straßenverkehrsanlagen: HBS*, 3rd ed. Köln: Forschungsgesellschaft für Straßen- und Verkehrswesen, 2009.
- [39] F. Rempe, P. Franeck, and B. K., “Estimating traffic speeds using probe data: A deep neural network approach,” *arXiv preprint*, 2021.