

Physics Informed Deep Learning for Traffic State Estimation

Jiheng Huang¹ and Shaurya Agarwal¹

Abstract—The challenge of traffic state estimation (TSE) lies in the sparsity of observed traffic data and the sensor noise present in the data. This paper presents a new approach — physics informed deep learning (PIDL) method — to tackle this problem. PIDL equips a deep learning neural network with the strength of the physical law governing traffic flow to better estimate traffic conditions. A case study is conducted where the accuracy and convergence-time of the algorithm are tested for varying levels of scarcely observed traffic density data — both in Lagrangian and Eulerian frames. The estimation results are encouraging and demonstrate the capability of PIDL in making accurate and prompt estimation of traffic states.

Index Terms—Physics Informed Machine Learning, Traffic State Estimation, Deep Learning, Sensor Placement

I. INTRODUCTION

Macroscopic traffic state variables, such as flow rate f , average vehicle speed v , and vehicle density ρ denote the traffic conditions on road segments in a traffic network. Through these indicators, transportation planners can perceive the congestion levels, recognize traffic demand, and even identify gridlocks and bottlenecks of a road network [1]. For instance, a severe decline in travel speed at a road section can implicate unusual events such as traffic incidents or dreadful congestion.

However, these key measurements needed for transportation planning and management are often sparse and potentially noisy [2]. Due to the combined factors such as cost of sensor installation, the accuracy of vehicle detection methods, and limitation on data storage and transmission, these traffic state variables are often observed partially [3] [4]. For example, traffic data would only be recorded at selected locations with the scattered deployment of vehicle detectors on a highway system [5]. Besides, this type of collected data is often compromised with various levels of inaccuracy, due to the existence of measurement noise in any type of detection and sensing devices [6]. Moreover, the data collected are routinely aggregated when transmitted from the sensors, worsening the time resolution of these measurements [7] [8].

Therefore, traffic state estimation (TSE) becomes an important task. It refers to the inference of traffic state variables of road segments using partially observed traffic data [9]. Accurate TSE in real-time is of the essence for efficient traffic management as control strategies are implemented

accordingly [10]. The practice of ramp control on freeways can illustrate the importance of TSE. It utilizes detected traffic flow data to recognize freeway traffic conditions, then alters the traffic signal timing to allow vehicles at the ramp to join the traffic on freeway according to the upstream and downstream flow levels [11].

Given the aforementioned impediment in collected traffic data and the importance of many TSE applications in transportation planning, oftentimes a priori knowledge is used by practitioners and researchers to estimate traffic states. These estimation approaches can be further categorized as model-driven approaches and data-driven approaches, based on the type of a priori knowledge they rely on [9].

Model-driven approaches adopt a traffic flow theory model in estimating traffic states. Models such as cell transmission model (CTM) [12] and switching mode model (SMM) [13] have been utilized to represent traffic flows. Kalman filter and its extensions [14] are commonly used to solve TSE efficiently and calibrate these models. The Kalman filter and its variations estimate the most probable state variables with respect to the observed data, assumed system model, and noises [9].

With the development of statistical methods and machine learning, data-driven TSE uses the power of historically observed traffic data for estimation tasks. This allows the discovery of the underlying structure of traffic data by the machine itself, eliminating the requirement of adopting a traffic flow model suitable for each road segment involved. Kernel regression [15], probabilistic principal component analysis [16], and deep learning [17] have been applied, among other data-driven techniques, in this kind of approach.

Comparing to the approach of blindly feeding traffic data into a deep learning neural network, the fundamental principles of traffic flow theory can aid in training the neural network and help uncover the implicit relationship between traffic state variables buried in the data. This builds a fast, resilient, and computation-friendly TSE approach.

This paper proposes a new approach in traffic state estimation, equipping a deep learning neural network with the physics of traffic flow theory, termed as *physics informed deep learning* (PIDL) for TSE. The relationship between the conventional TSE approaches and the proposed PIDL is shown in Fig. 1.

Outline: The rest of the paper is organized as follows: Section II reviews traffic flow theory and the conservation law, as the physical law in TSE. Section III provides the background of deep learning neural networks. Section IV

¹ Jiheng Huang and Shaurya Agarwal are with the Civil, Environmental & Construction Engineering Department at University of Central Florida, Orlando, FL, USA jihenghuang@knights.ucf.edu, shaurya.agarwal@ucf.edu

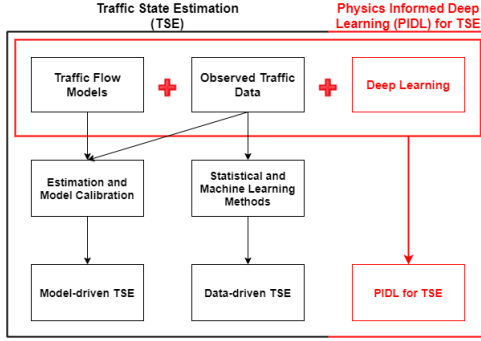


Fig. 1. Relationship between Traffic State Estimation Approaches

proposes the integration of physics and deep learning in the estimation of traffic states. Section V presents an experiment using the proposed method to verify its feasibility in TSE. Section VI details the experiment result and compares the performance of PIDL with a deep learning neural network in the absence of physical knowledge. Section VII concludes the paper and suggests the future work in PIDL for TSE.

II. PHYSICS OF THE TRAFFIC FLOW

The flow rate q indicates the number of vehicles that pass a set location in a unit time period. Average speed v is the mean value of speed among vehicles traveling on a road segment. Density ρ depicts the number of vehicles in a unit road space. Together, these time-series quantities show the temporal development of traffic conditions.

Cumulative flow $N(x, t)$ represents the number of vehicles which have passed a designated location x by the time t . Flow $q(x, t)$ is the partial differential of cumulative count $N(x, t)$ with respect to t . Density $\rho(x, t)$ is the partial differential of cumulative count $N(x, t)$ with respect to x . This relationship is exhibited in (1).

$$\begin{cases} q(x, t) = \frac{\partial N(x, t)}{\partial t} \\ \rho(x, t) = -\frac{\partial N(x, t)}{\partial x} \end{cases} \quad (1)$$

Lighthill-Whitham-Richards (LWR) model [18] relates flow $q(x, t)$, density $\rho(x, t)$ and the cumulative count of vehicles $N(x, t)$. When $N(x, t)$ is differentiable in both the time and space domain, (1) leads to the conservation law of traffic flow. The conservation law is given by (2), which is also referred to as LWR traffic flow PDE.

$$\frac{\partial q(x, t)}{\partial x} + \frac{\partial \rho(x, t)}{\partial t} = 0 \quad (2)$$

Fundamental diagram of traffic flow gives the relationship between traffic state variables q , v , and ρ , allowing us to hypothesize about the average driving behavior. Diverse representations of fundamental diagram exist and one of the widely used, due to its simplicity, is proposed by Greenshields [19].

In the Greenshields fundamental diagram, the relationship between traffic state variables can be summarized in (3),

where v_f is the free-flow speed and ρ_m is the maximum density (also known as jam density).

$$\begin{cases} v(\rho) = v_f \left(1 - \frac{\rho}{\rho_m}\right) \\ q(\rho) = \rho v_f \left(1 - \frac{\rho}{\rho_m}\right) \end{cases} \quad (3)$$

Aside from the fundamental diagram, traffic flow models are also critical components in traffic flow theory. Traffic flow models use empirical data and developed hypotheses to model traffic conditions [20]. Making use of traffic flow models, model-based TSE conveys the principles observed in the physical world about the relationship between traffic state variables such as flow and density.

If the adopted traffic flow model can accurately illustrate the relationship between traffic state variables observed in reality, the model-driven approach — based on the physics of traffic flow — can precisely predict traffic states in unobserved areas (with no data collection devices) and provide higher resolution of traffic state data (at data collection locations). However, the dependence on the traffic flow model brings the risk of adopting an unfit model in TSE. Using common and proven traffic flow models somewhat lowers this kind of risk, nevertheless, the robustness of estimation in the event of an anomaly such as a traffic accident or inclement weather condition leaves room for improvement.

III. DEEP LEARNING NEURAL NETWORK

The data-driven approach of traffic state estimation addresses the issue with traffic flow model accuracy from another source: the data collected on the roadways itself. Instead of relying on the formulated traffic flow models based on empirical evidence or historical experience, it applies the rapidly emerging machine learning techniques to recognize the relationship between traffic state variables. If deployed in real-time, the machine learning algorithm learns to adapt to the particular environment patterns such as congestion and rush-hour traffic based on the input data it gets. Therefore comparing to a model-driven approach, the estimation result of the data-driven approach is expected to be more reliable when a traffic anomaly is present.

Deep learning (DL) neural network is a subset learning method of machine learning techniques. The topology of a deep learning neural network consists of sequential layers of neurons, which are computation units resembling biological neurons in the human brain. The three common types of layers are *input layer*, where input data are fed into the neural network; *output layer*, where prediction is given as output; and *hidden layer*, where information is processed between the input and output layer. When building a deep learning neural network for TSE, some essential components that warrant contemplation are:

Network Size - Network size reflects the complexity of a DL neural network and has an impact on estimation accuracy. The network size is determined by the number of layers n_l and the number of hidden units n_h . Building a complex DL

may be beneficial in yielding accurate results, but it may be prone to over-fitting [21], and will generally take longer computation time, which will be discussed below.

Cost Function - Cost function is constructed to evaluate the inaccuracy of a DL neural network in estimation. In each training iteration, weights in neurons are adjusted to minimize the cost. Various measurements such as mean square error (MSE) can be utilized to build cost function. (4) is a cost function using MSE, in which N is the number of estimation outputs. At location x and time t , $\hat{\rho}(x, t)$ is the estimation of the genuine vehicle density $\rho(x, t)$.

$$J_{DL} = MSE_{(\rho(x,t), \hat{\rho}(x,t))} = \frac{1}{N} \sum_{i=1}^N |\rho(x, t) - \hat{\rho}(x, t)|^2 \quad (4)$$

Accuracy - Accuracy can also represent the DL performance in estimation. As the objective of minimizing cost drives the DL to search for finer weight configuration in the training process, accuracy is evaluated based on the estimation performance of DL on the testing dataset. A normalized accuracy measurement ACC using Frobenius norm is shown in (5). \mathbf{P} is the matrix form of vehicle density $\rho(x, t)$, and $\hat{\mathbf{P}}$ is the estimation of \mathbf{P} . ACC normalizes the cost and represents the relative error in estimation.

$$ACC = \frac{\|\mathbf{P} - \hat{\mathbf{P}}\|_F}{\|\mathbf{P}\|_F} = \frac{\sqrt{\sum_{x=1}^{X_m} \sum_{t=1}^{T_n} |\rho(x, t) - \hat{\rho}(x, t)|^2}}{\sqrt{\sum_{x=1}^{X_m} \sum_{t=1}^{T_n} |\rho(x, t)|^2}} \quad (5)$$

Computation Time - As the estimation of traffic states in real-time is preferable in traffic management [22], the evaluation of DL for TSE should also include examining the computation time. The long computation time of complex DL neural networks limits the scope of adoption.

Deep learning for TSE has the ability to swiftly modify its prediction based on real-time data. However, the looseness in the collected data may bring the sensor or detection bias into the neural network, resulting in less reliable estimation results.

IV. PHYSICS INFORMED DEEP LEARNING FOR TRAFFIC STATE ESTIMATION

Physics informed deep learning (PIDL) is a type of deep learning (DL) method in which a neural network is trained to solve learning tasks while *respecting the law of physics*, given by general nonlinear partial differential equations (PDE) [23]. With the inherent physical laws encoded as a priori knowledge, the resulting neural network forms data-efficient approximators to process input information and give prediction result [24]. When input data is inadequate or noisy, PIDL allows the neural network to make full use of the data with the aid of physics. It explains the underlying relationship between state variables in the data and improves the prediction result.

Therefore, given the unique advantage of PIDL in efficiently utilizing limited input data and the conservation law of traffic flow, we propose physics informed deep learning for traffic state estimation problem. It combines the strengths

of the underlying physical law of the system and the deep learning method in exploiting sparsely observed traffic data for TSE.

The conservation law of traffic flow described in (2) establishes the relationship of traffic state variables with respect to location x and time t . It needs to be integrated into the learning process of a PIDL neural network. When vehicle density $\rho(x, t)$ is used as training input, we establish two measures to evaluate the estimation performance: 1) DL-cost, J_{DL} , denoting the error in estimation while using just the DL; 2) physics-cost, J_{PHY} , denoting the disobedience of conservation law in the estimation. They are described as:

$$\begin{cases} J_{DL} = MSE_{(\rho(x,t), \hat{\rho}(x,t))} \\ J_{PHY} = MSE_{(0, \frac{\partial q(x,t)}{\partial x} + \frac{\partial \rho(x,t)}{\partial t})} \end{cases} \quad (6)$$

If ρ is the only available input data, the relationship between ρ and traffic flow q from Greenshields' fundamental diagram (3) transforms the conservation law (2) into (7).

守恒方程 \rightarrow

$$v_f(1 - \frac{2\rho(x, t)}{\rho_m}) \frac{\partial \rho(x, t)}{\partial x} + \frac{\partial \rho(x, t)}{\partial t} = 0 \quad (7)$$

Consequently, we can further formulate the DL-cost J_{DL} and physics-cost J_{PHY} as:

$$\begin{cases} J_{DL} = \frac{1}{N} \sum_{i=1}^N |\rho(x, t) - \hat{\rho}(x, t)|^2 \\ J_{PHY} = \frac{1}{N} \sum_{i=1}^N \left| v_f(1 - \frac{2\rho(x, t)}{\rho_m}) \frac{\partial \rho(x, t)}{\partial x} + \frac{\partial \rho(x, t)}{\partial t} \right|^2 \end{cases} \quad (8)$$

To incorporate the conservation law of traffic flow in training a PIDL neural network, the cost function of PIDL is comprised of the DL-cost J_{DL} and the physics-cost J_{PHY} . Parameter μ is introduced to adjust the weights of J_{DL} and J_{PHY} . The cost function \mathcal{J} of PIDL for TSE is given as:

$$\mathcal{J} = \mu * J_{DL} + (1 - \mu) * J_{PHY} \quad (9)$$

The PIDL algorithm for TSE is as follows: after splitting the collected traffic state dataset into training and testing datasets, the training dataset is fed into the neural network and PIDL makes an estimation. Both estimation cost and physics cost are computed based on the result. Consequently, the neural network gains the knowledge of the governing physical law by incorporating the non-compliance of conservation law J_{PHY} into the cost function \mathcal{J} . The training iteration is repeated if the sum of estimation cost and physics cost is greater than the designated threshold. Otherwise, the fine-tuned PIDL gives the estimation as output. A maximum of allowed learning iterations - i_{max} is set to prevent the learning process from running eternally, in the event of no change of total cost. These steps are graphically demonstrated in Fig. 2.

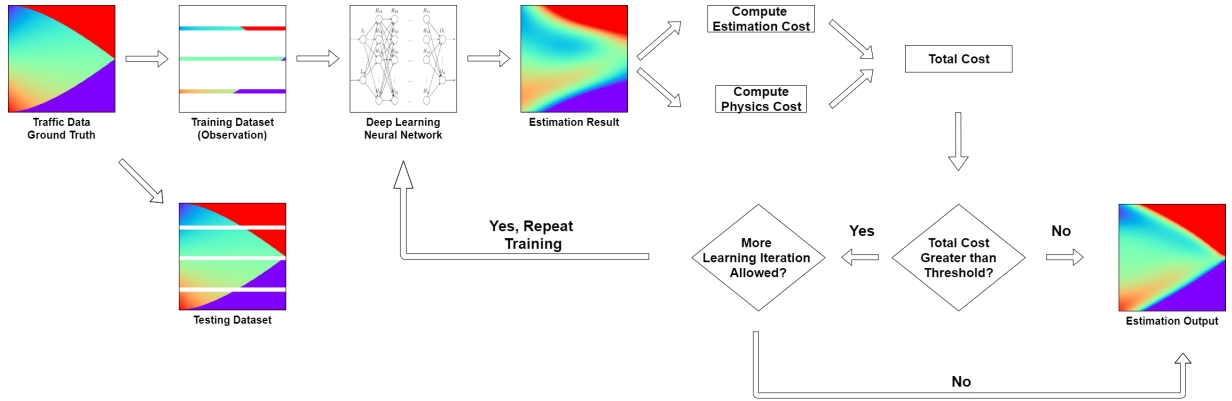


Fig. 2. Physics Informed Deep Learning for Traffic State Estimation

V. EXPERIMENTAL SETUP

An experiment is conducted to validate the capability of PIDL neural network for TSE. Given the nature of sparsity in traffic data, the ability to gain accurate perception with limited input is highly valuable. For comparison, a deep learning (DL) neural network with the same configuration is built without the knowledge of physics.

For this experiment, synthetic traffic data on a road segment is generated by using the Lax-Hopf method as described in [25], under a no upstream flow and no downstream flow condition. Traffic density value $\rho(x, t)$ is created for time $t \in [0, 50]$ seconds and location $x \in [0, 1000]$ meters. Both time and distance scales are evenly divided into 500 units. Therefore the time step is 0.1 seconds, and the distance step is 2 meters.

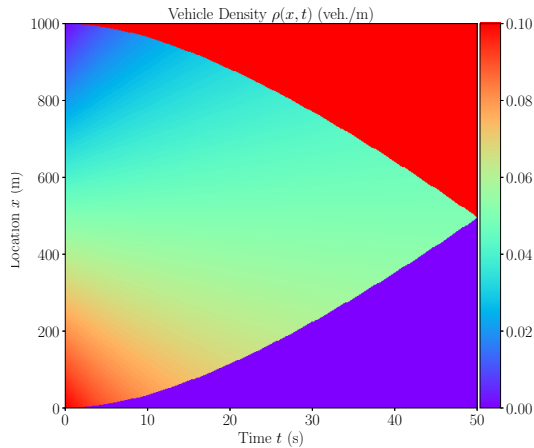


Fig. 3. Experiment Dataset - Vehicle Density $\rho(x, t)$

Because there is no upstream and downstream flow, the boundary conditions are formulated follows:

$$\frac{\partial q(0, t)}{\partial x} = \frac{\partial q(1000, t)}{\partial x} = 0 \quad (10)$$

We set free-flow speed v_f as 30 meters per second (67 miles per hour) and maximum density ρ_m as 0.1 vehicles per meter (161 vehicles per mile). The relationship between traffic states (speed v , flow q , and density ρ) is modeled using Greenshields' relationship as described in (3). The generated synthetic dataset is shown in Fig. 3.

In the experiment, we set the value of parameter μ in (9) as 0.5 to give equal weight to J_{DL} and J_{PHY} in calculating the cost \mathcal{J} . We utilize an 11-layer architecture with 50 neurons on each hidden layer to construct both PIDL and DL for the estimation task. The optimization algorithm used in both neural networks is limited-memory Broyden-Fletcher-Goldfarb-Shanno (LM-BFGS) method. PIDL and DL are built by using TensorFlow, and TSE tasks are conducted on an Intel Core i7-8700 CPU @ 3.20GHz, with a RAM size of 16 GB.

VI. RESULTS AND DISCUSSION

The estimation accuracy and computation time of the PIDL neural network and DL neural network are compared in this section. Both neural networks are built with the same structure — the same number of layers n_l and the same number of neurons in each hidden layers n_h^l , and are tasked with the same estimation task. The mere difference is that PIDL has the a priori knowledge of the underlying physics, i.e. conservation law of traffic flow. Physics-cost, J_{PHY} , is integrated into the cost function of PIDL.

Firstly, we investigate a scenario in which detected traffic input comes from sensors at fixed locations (Eulerian measurement frames). Secondly, we explore a scenario in which data from randomized locations are considered to compare the performance between PIDL and DL. The second scenario can be compared to the one where measurements are obtained in Lagrangian frame of reference using probe vehicles or GPS data of connected vehicles. The accuracy of the estimation result is evaluated by using an accuracy metric (ACC) defined in (5).

A. Fixed Sensor Location Scenario

Traffic data at 5 locations ($x = 2m, 250m, 500m, 750m, 1000m$) are selected to train the neural network. These locations include the boundaries of the road segment (2m and 1000m) and sample locations in between (250m, 500m, and 750m). Adding difficulty to the estimation task, we randomly chose 1000 samples from the 2500 samples at locations selected as training data. 1000 samples represent 0.4 percent of data from the entire road segment. The result is shown in Fig. 4, in which training samples are also marked.

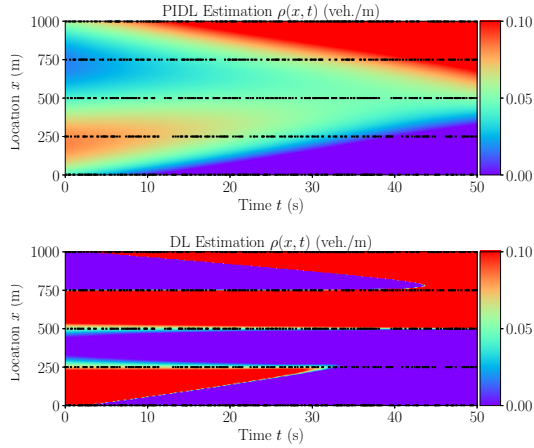


Fig. 4. PIDL & DL Prediction Result, Fixed Sensor Location

With the aid of physics, PIDL efficiently utilizes the training data and produces estimation result bearing the resemblance of the entire experiment dataset shown in Fig. 3. In contrast, the meager training samples are insufficient for the DL neural network to gain insightful knowledge of traffic conditions in the entire road segment. It produces an inscrutable estimation result in this scenario. Table I details the accuracy and computation time of PIDL and DL for comparison (bold font indicates better performance).

TABLE I
PIDL & DL PERFORMANCE COMPARISON, FIXED SENSOR LOCATION

Sensor Location	Computation Time (s)		Accuracy	
	PIDL	DL	PIDL	DL
Fixed	5.7	31.3	73.7%	37.0%

Varying Sample Size: We further inspect the impact of varying training sample size on the estimation result of PIDL and DL. Using a sample size from 250 (0.1 percent of data) to 2500 (1 percent of data) as training data, the accuracy and computation time of PIDL and DL estimation are shown in Fig. 5. The result on accuracy and computation time is given as the average of 5 repeated experiments with different (but of the same size) randomly selected training sets.

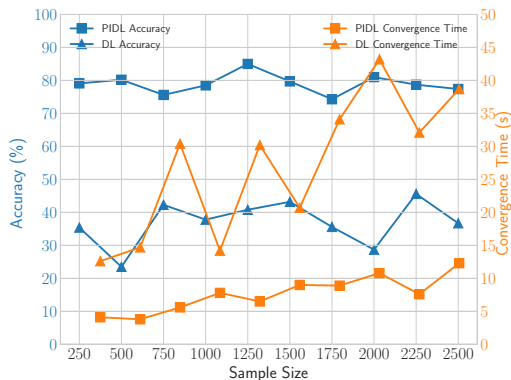


Fig. 5. Accuracy and Convergence Time Comparison, Fixed Sensor Location

Results show that PIDL exceeds the performance of DL at all sample levels by a large margin. It also shows that PIDL

produces the estimation result in a shorter period of time.

This scenario involved using input data from sensors at fixed locations and demonstrated the powerful capability of PIDL in utilizing extremely limited data to produce accurate estimation results of traffic states. Results highlight the value of PIDL in TSE applications where sensors such as loop detectors are often installed sparsely at predetermined locations on the road. It demonstrates that the knowledge of traffic flow theory (LWR in this case) can effectively aid a deep learning neural network in the task of estimating traffic states using limited sensor data.

B. Randomized Sensor Location Scenario

We admit 1000 training samples (same as the previous scenario) at random locations and time, which allow us to mimic another traffic data collection method: through floating cars or probe vehicles. The prediction result is shown in Fig. 6.

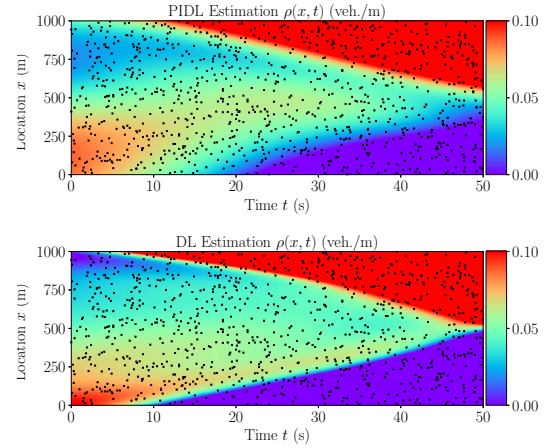


Fig. 6. PIDL & DL Prediction Result, Random Sensor Location

Both neural networks achieved acceptable estimation accuracy. Table II summarizes the performance of PIDL and DL in this scenario. Given the variance in time and location of training samples this time, it takes both neural networks a longer time to give estimation results. DL accomplishes better estimation accuracy and PIDL has the advantage of curtailing computation time.

TABLE II
PIDL & DL PERFORMANCE COMPARISON, RANDOM SENSOR LOCATION

Sensor Location	Computation Time (s)		Accuracy	
	PIDL	DL	PIDL	DL
Randomized	411.8	861.3	79.0%	86.1%

Note that Table II summarizes PIDL and DL performance under *one* sampling condition in which 1,000 training data were used. Before assuming DL achieved better accuracy, the following results using varying sample size should also be considered as they provide the fuller picture of performance comparison between PIDL and DL.

Varying Sample Size: The effect of varying training sample sizes in estimation performance is once again evaluated. The estimation accuracy and computation time of PIDL and

DL are shown in Fig. 7. It shows that both neural networks have the ability to acquire accurate estimation result and the difference is relatively insignificant. Again the result is the average of 5 experiments using diverse training sets randomly selected.

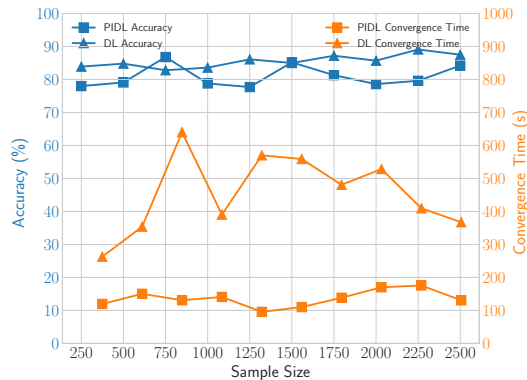


Fig. 7. Accuracy and Convergence Time Comparison, Random Sensor Location

The variety of observation locations where training data were obtained undoubtedly boosted the ability of DL in attaining more accurate estimation. Still, achieving similar estimation accuracy, PIDL converges at a faster speed. The shortened computation time gives PIDL a unique advantage in real-time applications during which timely estimation of traffic states is desired.

VII. CONCLUSION

In this paper, we presented a physics informed deep learning (PIDL) methodology for traffic state estimation (TSE). Combining the strength of underlying physical laws of traffic flow and deep learning techniques, we demonstrated the potent capability of PIDL in utilizing extremely limited traffic data for accurate and real-time TSE.

The experiment utilized a synthetic dataset generated by using the Lax-Hopf method as readily available field data sources were in an temporally-aggregated form, hindering the task of achieving meaningful results. Future work should resolve this issue by using field data in better temporal resolution. Further, depending on the field data, the exact relationship between the traffic parameters should be utilized along with an appropriate calibration process.

REFERENCES

- [1] 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.
- [2] N. Bekiaris-Liberis, C. Roncoli, and M. Papageorgiou, "Highway traffic state estimation with mixed connected and conventional vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3484–3497, 2016.
- [3] S. Agarwal, P. Kachroo, and S. Contreras, "A dynamic network modeling-based approach for traffic observability problem," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 1168–1178, 2015.
- [4] S. Agarwal and P. Kachroo, "Controllability and observability analysis for intelligent transportation systems," *Transportation in Developing Economies*, vol. 5, no. 1, p. 2, 2019.
- [5] S. Contreras, P. Kachroo, and S. Agarwal, "Observability and sensor placement problem on highway segments: A traffic dynamics-based approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 848–858, 2015.
- [6] S. Contreras, S. Agarwal, and P. Kachroo, "Quality of traffic observability on highways with lagrangian sensors," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 2, pp. 761–771, 2017.
- [7] 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.
- [8] S. Agarwal, E. E. Regentova, P. Kachroo, and H. Verma, "Multidimensional compression of its data using wavelet-based compression techniques," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1907–1917, 2016.
- [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.
- [10] E. F. Grumert and A. Tapani, "Traffic state estimation using connected vehicles and stationary detectors," *Journal of advanced transportation*, vol. 2018, 2018.
- [11] S. Agarwal, P. Kachroo, S. Contreras, and S. Sastry, "Feedback-coordinated ramp control of consecutive on-ramps using distributed modeling and godunov-based satisfiable allocation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2384–2392, 2015.
- [12] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994.
- [13] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez, "Piecewise-linearized cell transmission model and parameter calibration methodology," *Transportation Research Record*, vol. 1965, no. 1, pp. 183–191, 2006.
- [14] Y. Yuan, F. Scholten, and H. van Lint, "Efficient traffic state estimation and prediction based on the ensemble kalman filter with a fast implementation and localized deterministic scheme," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 477–482.
- [15] W. Yin, P. Murray-Tuite, and H. Rakha, "Imputing erroneous data of single-station loop detectors for nonincident conditions: Comparison between temporal and spatial methods," *Journal of Intelligent Transportation Systems*, vol. 16, no. 3, pp. 159–176, 2012.
- [16] L. Li, Y. Li, and Z. Li, "Efficient missing data imputing for traffic flow by considering temporal and spatial dependence," *Transportation research part C: emerging technologies*, vol. 34, pp. 108–120, 2013.
- [17] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 1–17, 2017.
- [18] P. I. Richards, "Shock waves on the highway," *Operations research*, vol. 4, no. 1, pp. 42–51, 1956.
- [19] B. Greenshields, W. Channing, H. Miller *et al.*, "A study of traffic capacity," in *Highway research board proceedings*, vol. 1935. National Research Council (USA), Highway Research Board, 1935.
- [20] D. B. Work, O.-P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton, "An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices," in *2008 47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 5062–5068.
- [21] C. Dumitru and V. Maria, "Advantages and disadvantages of using neural networks for predictions," *Ovidius University Annals, Series Economic Sciences*, vol. 13, no. 1, 2013.
- [22] 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, 2005.
- [23] 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.
- [24] —, "Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations," *arXiv preprint arXiv:1711.10566*, 2017.
- [25] P.-E. Mazaré, A. H. Dehwhah, C. G. Claudel, and A. M. Bayen, "Analytical and grid-free solutions to the lighthill-whitham-richards traffic flow model," *Transportation Research Part B: Methodological*, vol. 45, no. 10, pp. 1727–1748, 2011.