A Hybrid Physics Machine Learning Approach for Macroscopic Traffic State Estimation
Zhao Zhang, Ding Zhao, Xianfeng Terry Yang

Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT 84112, USA
Mechanical Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Full-field traffic state information (i.e., flow, speed, and density) is critical for the successful operation of Intelligent Transportation Systems (ITS) on freeways. However, incomplete traffic information tends to be directly collected from traffic detectors that are insufficiently installed in most areas, which is a major obstacle to the popularization of ITS. To tackle this issue, this paper introduces an innovative traffic state estimation (TSE) framework that hybrid regression machine learning techniques (e.g., artificial neural network (ANN), random forest (RF), and support vector machine (SVM)) with a traffic physics model (e.g., second-order macroscopic traffic flow model) using limited information from traffic sensors as inputs to construct accurate and full-field estimated traffic state for freeway systems. To examine the effectiveness of the proposed TSE framework, this paper conducted empirical studies on a real-world data set collected from a stretch of I-15 freeway in Salt Lake City, Utah. Experimental results show that the proposed method has been proved to estimate full-field traffic information accurately. Hence, the proposed method could provide accurate and full-field traffic information, thus providing the basis for the popularization of ITS.

Keywords: Second-order traffic flow model; traffic state estimation; hybrid physics machine learning; Traffic sensor data

1. Introduction

In the last several decades, Intelligent Transportation Systems (ITSs) have been widely deployed on freeway systems for improving travel safety and offering more efficient travel choices (e.g., departure time, route, and mode) to travelers (Zhang et al., 2011). The effectiveness of an ITS depends on the quality of obtained traffic information, especially for the Advanced Traffic Management System (ATMS) and Advanced Traveler Information System (ATIS) (Ma et al., 2015). Hence, the success of ITSs on freeways significantly relies on whether the system could offer accurate and full-field historical traffic state. The accurate and full-field historical and future traffic state can help travelers to preplan and reschedule trips, traffic engineers to find better countermeasures to mitigate traffic congestion, and transportation agencies to improve traffic operational efficiency and safety (Lv et al., 2015; Ma et al., 2015; Wang et al., 2019). TSE is a method that can infer traffic state (e.g., flow, speed, density, etc.) using partially observed data from traffic sensors on the roadway system (Seo et al., 2017). It plays a key role in traffic control and operations of ITS since accurate network-wide is usually hard to obtain.

In practice, traffic information is collected by multiple sensor sources (e.g., inductive loops, radars, cameras, Global Positioning System-GPS, etc.). However, the limitations of traffic information which is directly obtained from those sources include: (a) roadside sensors, it could provide accurate traffic information but only available on sparse locations with the sensors installed, and (b) probe vehicles, traffic data estimated by using probe vehicle trajectories or GPS
information, which can provide full-field traffic information, but the estimation are often biased due to the low penetration rate of probe vehicles, and it is only available for few transportation systems in the city. Therefore, a reliable traffic state estimation (TSE) system, which can estimate accurate, timely, and full-field traffic information for freeway system with existing collected traffic data, is strongly needed. The reliable TSE systems can potentially promote ITSs development, making the ITSs more effective and reliable.

Data-driven and model-driven approaches are the two most common TSE methods on freeways (Seo et al., 2017). Both approaches have their advantages and limitations, considering the advantage of the traffic physics model and data dependence limitation of machine learning (ML) models, a hybrid physics ML (HPML) that using the output of the traffic flow model with traffic sensor data as input for ML models will produce a better alternative to address the TSE challenges. This TSE framework consists of base machine learning models (e.g., artificial neural network (ANN), random forest (RF), support vector machine (SVM)) and a traffic flow model – second ordered traffic flow model. It combines the advantages of both approaches to construct accurate and full-field traffic information on freeway systems.

The remainder of the paper is organized as follows: Section 2 reviews related existing research about TSE; Section 3 introduces the framework of hybrid physics machine learning model for TSE; Section 4 performs results comparison and error analysis for demonstrating the effectiveness of the proposed method. Conclusions and future research directions are discussed in the last section.

2. Literature review

Having an accurate and full-field estimation of traffic state (e.g., flow, speed, and density) is one of the most essential components in Intelligent Transportation Systems (ITS) (e.g., Advanced Traffic Management systems (ATMS), Advanced Traveler Information Systems (ATIS)) (Ma et al., 2015; Vlahogianni et al., 2014; Wang, 2010). The accuracy of traffic state prediction relies heavily on traffic data availability and coverage due to the stochastic characteristics of traffic flow. Most of the existing studies on traffic state predictions are dependent on historical traffic information collected by various traffic sensors (Lv et al., 2015; Ma et al., 2015; Tian et al., 2018; Zhang and Ge, 2013; Zou et al., 2014), such as inductive loop detectors, radars, cameras, and Global Positioning System (GPS). The implementation of traffic state prediction tends to limit by the location where traffic sensors are installed. Therefore, a reliable traffic state estimation (TSE) method that could resolve the limitation of traffic data availability and coverage, is strongly needed for better traffic state prediction. Traffic state estimation (TSE) has the potential to solve this limitation (Yuan et al., 2020; Zhang et al., 2020). Such a method will promote the implementation of more successful traffic predictions on the freeway system, and thereby providing the transportation agencies more useful information in terms of making timely countermeasures and creating a much safer and efficient traffic environment.

The effectiveness of data fusion techniques for improving travel time estimation accuracy was proved by some research (Anusha et al., 2012; Zhu et al., 2018). Zhang and Yang (Zhang and Xianfeng Yang, 2019) fused probe data and sensor detector data to accurately estimate freeway traffic speed by using regression machine learning models. The hybrid data-driven and model-based approach for traffic time estimation and forecasting were implemented and evaluated by another group of studies (Allström et al., 2016; Hofleitner et al., 2012; Kumar et al., 2017; You and Kim, 2000; Yu et al., 2010) For example, You and Kim (You and Kim, 2000) proposed a hybrid nonparametric regression model with geographic information systems (GIS) information
to forecast link travel times in congested road networks. Yu et al. (You and Kim, 2000) proposed a hybrid model that combines SVM and Kalman filtering (KF) for effectively predicting bus arrival time. Hofleitner et al. (Hofleitner et al., 2012) presented a hybrid modeling framework, that combines the advantages of pure statistical and traffic flow models, to forecast travel time on local arterial. Allström et al. (Allström et al., 2016) applied a hybrid approach that adopted ANN with output from the cell transmission model (CTM) for short-term traffic state and travel time prediction. Kumar et al. (Kumar et al., 2017) proposed an integrated method, which uses exponential smoothing (ES) and Kalman filter (KF) to estimate travel time as a new observation for ARIMA models, to estimate bus travel time. Sharmila et al. (Sharmila et al., 2019) proposed a hybrid model that combines data-driven approach and model-based approach for corridor-level travel time estimation. These studies demonstrated that data expansion, data fusion, and hybrid approaches could improve the estimation and prediction accuracy of traffic measures.

Data-driven and model-driven approaches are the two most common TSE methods on freeways (Seo et al., 2017). These approaches are designed to simulate traffic dynamics, capture data noise, and predict unobserved spatiotemporal traffic states. With the advance of data collecting, processing, and computation technologies in recent years, data-driven approaches have been widely developed and implemented for the TSE problem. Machine learning (ML) models are prevailing in capturing the stochastic characteristics of traffic flow with a massive amount of data (Duan et al., 2016; Li et al., 2013; Ni and Leonard, 2005; Polson and Sokolov, 2018; Tak et al., 2016; Tan et al., 2014; Tang et al., 2015; Wu et al., 2018; Yuan et al., 2020; Zhang et al., 2020; Zhong et al., 2004). The performance of ML models heavily relies on high-quality data due to the driven nature. However, the traffic data on freeway systems usually have two constraints (Zhang et al., 2020): (1) traffic sensors could offer accurate traffic information, but they are only available on the locations where the traffic sensors have been installed, and (2) estimated probe vehicle data could offer full-field traffic information for highway system, but it is only available in few transportation systems in the city and has data bias problem. The model-driven approaches are referred to the traffic physic models, which can be classified into three categories: (1) the first-order Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), (2) the second-order Payne-Whitham (PW) model (PAYNE, 1971; Whitham, 1975), and (3) the second-order Aw-Rascle-Zhang (ARZ) model (Aw and Rascle, 2000; Zhang, 2002). The LWR is developed under ideal theoretical conditions, thus it cannot reproduce complicated traffic phenomena. The PW and ARZ models could deal with more complicated traffic conditions by the proposed momentum equation, but they still require tremendous computation efforts. Papageorgiou et al. (Papageorgiou et al., 1989a) proposed a discrete PW-like TSE model, named METANET, which can overcome the limitation of the PW model. METANET has been successfully applied in many studies (Wang and Papageorgiou, 2005; Zhang et al., 2020). Therefore, this study utilized a discrete macroscopic traffic flow model to simulate traffic state on the freeway.

This paper contributes to the literature in the following aspects:
(1) The traffic flow model increases the availability of data that TSE is more resistant to the current availability constraints of traffic data.
(2) The proposed hybrid ML brings transportation theoretical foundations to ML-based methodologies combine the advantages of both model-driven and data-driven approach for TSE.
(3) Data from traffic sensors on interstate freeway segments are used to validate the effectiveness and the applicability of the proposed method.
3. Methodology

Framework of Hybrid Physics Machine Learning Model for TSE

In this part, the framework of HPML for TSE are introduced, which includes two main components: (1) second-order macroscopic traffic flow model to create simulated traffic state; (2) regression machine learning models that can capture the stochastic characteristics of traffic state. The applied regression machine learning models include SVM, RF, and ANN.

Fig. 1 shows the diagram of proposed HPML model for TSE problem. The HPML model is a combination of data-driven and model-driven approaches, which uses the output of a physics-based model as input to an ML model. The second-order macroscopic traffic flow model is used as base traffic flow model (TFM) in the HPML approach. The basic logic of proposed HPML approach for TSE is presented in Alg. 1.

![Diagram of proposed HPML model for TSE problem](image)

Figure 1: Diagram of proposed HPML model for TSE problem

Fig. 1 shows the diagram of proposed HPML model for TSE problem. The HPML model is a combination of data-driven and model-driven approaches, which uses the output of a physics-based model as input to an ML model. The second-order macroscopic traffic flow model is used as base traffic flow model (TFM) in the HPML approach. The basic logic of proposed HPML approach for TSE is presented in Alg. 1.

|Stationary Data($q(d), u(d))$| Flow direction|
|---|---|
|Downstream|Upstream|

|Traffic Flow Model| HPML|
|---|---|

|Physics Model Estimates($q_{phy}, u_{phy}$)|
|---|

|Traffic State Estimates ($Y_{est}$)|
|---|

Let $\mathcal{D}$ denotes the discrete spatiotemporal traffic data points for whole freeway segments: $\mathcal{D} = \{x_{i,k}|i \in [0,n], \forall k \in [0,t]\}$ ($i$ denotes the segment number; $k$ denotes the time interval (5-min)). The traffic information from stationary data points is denoted by $\mathcal{S} = \{x_{i,k}^s|i = 1, ..., n_s\}$ (e.g., traffic information from stationary traffic sensors). Based on the upstream and down stream stationary data $[x_{i,k}^s, x_{n,k}^s]$, the TFM model points $\mathcal{T} = \{x_{j,k}^{tfm}|j = 1, ..., n_T\}$ could be generated by TFM model. Therefore, the training data for HPML model consist of (1) stationary data points is denoted by $\mathcal{S} = \{x_{i,k}^s|i = 1, ..., n_s\}$; (2) TFM data points $\mathcal{T} = \{x_{j,k}^{tfm}|i = 1, ..., n_T\}$; (3) target values $\mathcal{Y} = \{y_{i,k}|i = 1, ..., n_s\}$ (i.e., the true traffic states at the stationary points). $i$ and $j$ are the indexes of stationary points and TFM data points. $\mathcal{S}$ and $\mathcal{T}$ are subsets of the discrete points $\mathcal{D}$ (e.g., $O \in \mathcal{D}$ and $P \in \mathcal{D}$). $\mathcal{S}$ and $\mathcal{Y}$ have the same index $i$ in the case of the target value $\mathcal{Y}$ paired with stationary point $\mathcal{S}$. In experiments, the stationary data points are usually limited by the availability of traffic sensors (e.g., probe and sensor detectors). Hence, the traffic state can be observed only in limited locations with traffic sensors installed. Therefore, an estimation method is needed to infer the unknown traffic state on the locations without traffic sensors. 

\[\text{Estimated Traffic State} = \mathcal{Y}_{est} = \mathcal{Y}(\mathcal{S}, \mathcal{T})\]
sensors installed. In this study, the traffic physics model is utilized to construct a traffic state for locations without traffic sensors installed based on limited traffic information from $O$, named TFM data points $T$. TFM points $T$ could overcome both location limitations and measurement limitations, which also can reflect the real traffic physical truth and is quite similar with the probe data. Hence, the HPML model is proposed to reach equivalent estimation precision with probe-ML with probe data.

Algorithm 1: HPML algorithm

**Result:** Estimated traffic states

Traffic physics simulation
1. Set the length of each sub-segment to 500m
2. **for** sub-segment $i$ **do**
3. Run TFM model: $[x_{1,t}^S, x_{n,t}^S]_{t \times 2}^{TFM} \rightarrow [x_{1,t}^S, x_{2,t}^S, \ldots, x_{n,t}^S]$ 
4. **end for**

Traffic flow modeling by HPML
5. Group the training dataset: $[t_{n,t}, d_{n,t}, q_{n,t}^{on}, q_{n,t}^{off}] \rightarrow [q_{n,t}, u_{n,t}]$
6. **for** sub-segment $i = 1, \ldots, N$ **do**
7. while sub-segment $i$ without point data **do**
8. Train HPML model with grouped dataset: $Loss = \alpha \cdot Loss(Y, \hat{Y}) + \lambda \cdot R(f)$
9. Use the trained model to estimate traffic state for sub-segment $i$, $[u_{n,t}^{hpm}, q_{n,t}^{hpm}]$
10. **end while**
11. **end for**

The schematic of the HPML model is shown in Fig. 1. The detailed introduction of HPML model is described in Chapter 4. The input of TFM model includes speed and flow from upstream and downstream segments. The unobserved traffic states could be estimated by TFM for all segments from the Eq. 1. The ANN, RF, and SVM are used as base machine learning model for constructing HPML model, then we can predict the unobserved traffic states by HPML in Eq. 4 from the samples in Eqs. 2 – 3. The input $X$ represents time $t$, distance $d$, TFM speed $u$ and flow $q$, the output $Y$ represents the corresponding vector of flow and speed.

$$\begin{bmatrix} x_{1,t}^S & x_{n,t}^S \\ x_{1,t}^S & x_{n,t}^S \\ \vdots & \vdots \\ x_{1,t}^S & x_{n,t}^S \end{bmatrix}_{t \times 2}^{TFM} \rightarrow \begin{bmatrix} x_{1,1}^T & x_{2,1}^T & \cdots & x_{n,1}^T \\ x_{1,2}^T & x_{2,2}^T & \cdots & x_{n,2}^T \\ \vdots & \vdots & \vdots & \vdots \\ x_{1,t}^T & x_{2,t}^T & \cdots & x_{n,t}^T \end{bmatrix}_{t \times n}$$ (1)
Second-order Macroscopic Traffic Flow Model

As an influential study in the literature, Papageorgiou et al. (Papageorgiou et al., 1989b) proposed the second-order macroscopic traffic flow model, METANET, which divide the target freeway segment into N subsegments with a unit length of ΔL (500m). For each subsegment \( i \), the mean density, \( d_j(k) \), can be determined by the difference between the input and output flows as follows:

\[
d_i(k + 1) = d_i(k) + \frac{ΔT}{λ_i ΔL} [q_{i-1}(k) - q_i(k) + σ_i(k) - s_i(k)]
\]

The dynamics of the speed can be represented by Eq. 2:

\[
\begin{align*}
    u_i(k + 1) &= u_i(k) + \frac{ΔT}{τ_i} [V[d_i(k)] - u_i(k)] + \frac{ΔT}{L_i} u_i(k)[u_{i-1}(k) - u_i(k)] \\
    &= \frac{γ_i ΔT [d_{i+1}(k) - d_i(k)]}{τ_i ΔL} \frac{ΔT}{L_i} \left[ r_i(k) - v_i(k) \right]
\end{align*}
\]

where \( V[d_i(k)] \) is the static speed for segment \( i \) at time \( k \) with respect to the density \( d_i(k) \):

\[
V[d_i(k)] = u_f \exp \left( -\frac{1}{a} \left( \frac{d_i(k)}{d_{cr}} \right)^{a} \right)
\]
The relationship between flow, density, and speed is given by the following:

\[ q_i(k) = d_i(k)u_i(k)\lambda_i \]  

(8)

where, \( i \) denotes the index of sub-sections of a freeway subsegment; \( k \) denotes the index of the time intervals; \( q_i(k) \) denotes the transition flow rate entering subsegment \((i+1)\) from \( i \) during interval \( k \); \( r_i(k) \) denotes the on-ramp flow rate entering subsegment \( i \) during interval \( k \); \( s_i(k) \) denotes the off-ramp flow rate leaving subsegment \( i \) during interval \( k \); \( d_i(k) \) denotes the average traffic density per lane in the subsegment \( i \) during interval \( k \); \( u_i(k) \) denotes the average speed in subsegment \( i \) during interval \( k \); \( \gamma, \tau, \delta, \kappa, \alpha \) are traffic state model parameters; \( \Delta L \) denotes the length of each freeway subsegment; and \( \lambda_i \) denotes the number of lanes in subsegment. Using the traffic flow and speed from traffic sensors at upstream and downstream stations, on-ramps, and off-ramps, one can directly use Eqs. 5-8 to construct the traffic state evolution on the target freeway section.

**Applied Regression Machine Learning Model for TSE**

- **Support Vector Machine:** The SVM algorithm is a supervised artificial intelligence (AI) approach developed by Vapnik with colleagues (Boser et al., 1992; Smola and Schölkopf, 2004; Vapnik, 2013). In the literature, the SVM method is considered as an effective and efficient algorithm for regression and forecasting. The SVM estimates the regression based on a series of kernel functions, has an ability to convert the lower-dimensional input data to a higher dimensional feature space via a nonlinear relationship and then perform linear regression within this space (Smola and Schölkopf, 2004). Time series and regression problems can be effectively modeled by SVM-based traffic models, which have been proved by several studies (Asif et al., 2014; Ma et al., 2015; Wu et al., 2004; Zhang and Liu, 2009).

- **Random Forest:** The RF model is an ensemble technique that can be performed in both regression and classification problems (Cutler et al., 2012), also called random decision forests model. The RF model can randomly select features to avoid the overfitting problem by utilizing Breiman’s “basgging” ideas. The application of the RF model has been successfully implemented in traffic prediction (Cutler et al., 2012; Hamner, 2010; Leshem and Ritov, 2007). Zhao et al. (Zhang et al., 2020) has applied RF for TSE with sensor and probe data on freeway and proved that RF could obtain an acceptable estimation accuracy.

- **Artificial Neural Network:** ANN is a popular AI approach which has been widely implemented to a variety of transportation problems. In the literature, most implemented ANN models are the multi-layer perceptron (MLP) which can be expressed as follow:

\[ y = h \left( \varphi_0 + \sum_{j=1}^{N} \varphi_j g \left( \sum_{i=1}^{M} \theta_i x_i \right) \right) \]  

(9)

where, \( M \) and \( N \) denote the number of neurons in the input layer and hidden layer, respectively; \( g \) and \( h \) represent the transfer functions for the input layer and hidden layer; and the vector matrices of \( \theta \) and \( \varphi \) denote the weight values for neurons in both input layer and hidden layer, respectively.
4. Result Analysis and Comparison

Experiments Setup

To evaluate the effectiveness of the proposed TSE framework, we applied three HPML models to estimate the traffic state in a stretch of interstate freeway I-15, Salt Lake City, Utah, U.S. The studied freeway stretch is presented in Fig. 2, where the seven blue rectangles (three detectors located on normal segment, two detectors located on off-ramp, and two detectors located on on-ramp) represent the detectors for training and the three yellow rectangles (one detector located on normal segment, one detector located on off-ramp segment, and one detector located on on-ramp segment) represent the detectors for testing.

- N: Normal segment; - Off: Off-ramp segment; - On: On-ramp segment

Figure 2: The deployment freeway corridor and stations

Three benchmark pure machine learning (pure-ML) models and three pure-ML with probe data and three HPML models are conducted for validating the performance of HPML models for TSE: 1) pure-ML: spatiotemporal and geometry information as input and observed traffic state as label; 2) pure-ML with probe data (termed as pure-ML with probe): spatiotemporal information, geometry information, and probe data as input and observed traffic state as label; 3) HPML: spatiotemporal information, geometry information, and TFM data as input and observed traffic state as label for training ML models. The pure-ML is the common scenario for TSE problem, which could be used as benchmark model for evaluating the performance of proposed model. The pure-ML with probe data evaluates the TSE accuracy when the estimated probe data is available as additional training variables. The hybrid-ML is conducted for evaluating the effectiveness of TFM data for TSE compared with pure-ML with probe. The utilized data in this study is obtained from the Performance Measurement System (PeMS) and Utah iPeMS databases managed by UDOT for model development and validation. The traffic information of PeMS is collected by installed sensors on every a few miles along the freeway. These sensors are helpful and accurate in terms of counting the number of vehicles in a certain time period and measuring the vehicle speed, but these data can only be obtained from the location that already installed sensors. iPeMS database, which utilizes probe data, is a sample of information collected from vehicle navigation systems, cell phone applications, and fleet vehicles. Notably, even though it can obtain traffic information (e.g., traffic speed and flow) on all segments of statewide freeways, the resolution is not ideal and could affect the accuracy of further traffic prediction. The real-time traffic data and road conditions are available online and can be accessed by the public. For model evaluation, the time range of the used data between 01/04/2021 and 01/10/2021. The input variables for HPML include traffic information (i.e., traffic flow and speed), the location coordinates of each sensor
(i.e., distance, on-ramp, and off-ramp information), and the time of each read. The Estimation results will be compared with the observed traffic state collected by traffic detectors. In this study, the calibrated initial traffic flow model parameters are listed in Table 1.

Table 1: The parameters of traffic physics model

| Parameter | Value         |
|-----------|---------------|
| n         | 9             |
| $\lambda_i$ | 4            |
| $\Delta T$ | 1/360 (h)    |
| $u_f$     | 120 (km/h)   |
| $\gamma$  | 35 (km$^2$/h)|
| $\Delta L$| 0.5 (km)     |
| $\tau$    | 0.05 (h)     |
| $\alpha$  | 1.4324        |
| $d_{cr}$  | 36.85 (veh/km)|
| $\kappa$  | 13 (veh/km)  |

Model Performance Measurement

To qualify the accuracy of outputs, three commonly statistical indicators, including Root Meaning Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). The RMSE, MAPE and MAE are defined as following:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y} - y_i)^2}$$  \hspace{1cm} (10)

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} |\frac{\hat{y} - y_i}{y_i}| \times 100\%$$  \hspace{1cm} (11)

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y} - y_i|$$  \hspace{1cm} (12)

where, $y_i$ is the observed traffic speed and $\hat{y}$ is the estimated or predicted traffic speed and flow.

Results Analysis and Comparison

Table 2 presents the TSE results from TFM, pure-ML, pure-ML with probe, and HPML models of freeway normal segment, on-ramp, and off-ramp segment. The HP-RF model yields lowest RMSE (43.25 veh/5-min), lowest MAPE (10.67%), and lowest MAE (29.46 veh/5-min) for flow and HP-SVM generates the lowest RMSE (2.28 mph), lowest MAPE (1.62%), and lowest MAE (1.08 mph) for speed. It indicates that HPML could obtain accurate TSE results. Compared with the results generated by pure-ML with probe models, the performance of HPML models has been greatly improved, especially for the flow. It demonstrated that the TFM data is more effective for ML models to improve TSE accuracy. In this case, the performance of TFM results is higher than pure-ML and pure-ML with probe models, especially for the flow. This shows that it is difficult for ML models to obtain an accurate estimation with limited information. To better illustrate the performance improvement by HPML models, Figure 5 (a)-(b) shows the comparison of the best estimation results obtained by TFM, pure-RF, pure-RF with probe, and HP-ANN models for flow and by TFM, pure-SVM, pure-SVM with probe, and HP-SVM for speed. The overall line pattern
of pure-ANN is roughly similar to ground truth, but the changes in details are not captured. For TFM and pure-ML with probe, the line pattern is almost the same as the ground truth, but it contains too many noises. It can be clearly seen that the line of HP-ANN better fits the ground truth, which demonstrates that three HPML models could accurately estimate speed and flow. To further demonstrate the variation of MAPEs of different models, Figure 6 shows the violin plots of MAPEs on test set by different models. For each ‘violin’ in the figure, a box plot is drawn inside, and its margin shows the Gaussian distribution of dataset. In Figure 6, it is noted that HPML model generally has lower MAPE value. It further proved that HPML could achieve relatively higher TSE accuracy.

Table 2: Estimation results of freeway

| Models                     | Flow RMSE | Flow MAPE | Flow MAE | Speed RMSE | Speed MAPE | Speed MAE |
|----------------------------|-----------|-----------|----------|------------|------------|-----------|
| **Normal Segment**         |           |           |          |            |            |           |
| TFM                        | 56.28     | 14.80%    | 38.14    | 2.40       | 1.91%      | 1.30      |
| Pure-SVM                   | 106.73    | 26.09%    | 68.08    | 2.56       | 2.04%      | 1.38      |
| Pure-RF                    | 103.25    | 26.50%    | 71.19    | 2.80       | 2.57%      | 1.78      |
| Pure-ANN                   | 95.30     | 31.17%    | 65.86    | 2.68       | 2.43%      | 1.68      |
| Pure-SVM with probe        | 99.16     | 31.17%    | 64.81    | 2.39       | 1.96%      | 1.33      |
| Pure-RF with probe         | 71.04     | 16.82%    | 48.71    | 2.39       | 2.09%      | 1.46      |
| Pure-ANN with probe        | 84.88     | 31.34%    | 58.87    | 2.39       | 2.29%      | 1.59      |
| HP-SVM                     | 46.92     | 12.25%    | 31.89    | 2.28       | 1.62%      | 1.08      |
| HP-RF                      | 43.25     | 10.67%    | 29.46    | 2.36       | 1.85%      | 1.26      |
| HP-ANN                     | 45.10     | 12.46%    | 31.82    | 2.38       | 1.74%      | 1.17      |
| **On-ramp Segment**        |           |           |          |            |            |           |
| TFM                        | 67.47     | 21.74%    | 51.12    | 1.92       | 2.02%      | 1.46      |
| Pure-SVM                   | 72.74     | 24.89%    | 50.20    | 1.99       | 2.17%      | 1.56      |
| Pure-RF                    | 50.82     | 20.01%    | 33.66    | 3.85       | 2.81%      | 2.03      |
| Pure-ANN                   | 55.08     | 26.91%    | 40.83    | 2.22       | 2.39%      | 1.72      |
| Pure-SVM with probe        | 69.52     | 24.83%    | 48.60    | 1.96       | 2.19%      | 1.58      |
| Pure-RF with probe         | 45.75     | 15.98%    | 30.48    | 2.14       | 1.99%      | 1.44      |
| Pure-ANN with probe        | 51.81     | 22.99%    | 36.69    | 2.02       | 1.99%      | 1.45      |
| HP-SVM                     | 45.01     | 15.82%    | 31.82    | 1.80       | 1.99%      | 1.44      |
| HP-RF                      | 36.06     | 11.91%    | 24.74    | 2.03       | 1.88%      | 1.35      |
| HP-ANN                     | 35.04     | 11.71%    | 24.03    | 1.82       | 1.99%      | 1.44      |
| **On-ramp Segment**        |           |           |          |            |            |           |
| TFM                        | 40.01     | 12.73%    | 26.82    | 2.04       | 1.98%      | 1.44      |
| Pure-SVM                   | 57.95     | 26.53%    | 41.70    | 2.00       | 1.67%      | 1.19      |
| Pure-RF                    | 52.00     | 20.37%    | 34.71    | 2.64       | 1.92%      | 1.37      |
| Pure-ANN                   | 55.87     | 24.06%    | 40.68    | 2.23       | 1.96%      | 1.41      |
| Pure-SVM with probe        | 55.30     | 26.74%    | 40.38    | 1.81       | 1.53%      | 1.09      |
| Pure-RF with probe         | 48.33     | 16.33%    | 34.50    | 1.79       | 1.71%      | 1.25      |
| Pure-ANN with probe        | 50.33     | 23.96%    | 35.71    | 1.82       | 1.63%      | 1.17      |
| HP-SVM                     | 36.39     | 13.32%    | 25.04    | 1.71       | 1.33%      | 0.94      |
| HP-RF                      | 35.85     | 12.96%    | 25.09    | 1.73       | 1.42%      | 1.02      |
| HP-ANN                     | 34.99     | 12.86%    | 23.89    | 1.75       | 1.54%      | 1.10      |

* Flow unit: (veh/5-min); Speed unit: mph
(a) Flow of normal segment

(b) Speed of normal segment

Figure 1: TSE estimates v.s. ground truth

Figure 2: MAPE distribution of different models on test set

The TSE results from TFM, pure-ML, pure-ML with probe, and HPML models of freeway on-ramp and off-ramp segment are shown in Table 2. Compared with normal segment, the conclusion can be reached that the performance of pure-ML, pure-ML with probe, and HPML models on on-ramp and off-ramp segment are better than normal segment. Overall, similar results pattern that the HPML models are perform better than other models was found on these three types of segments. It indicates that TSE accuracy can be significantly enhanced by HPML models on normal, on-ramp, and off-ramp segment. Fig. 4 shows the comparison of estimation results of TFM, pure-ML, pure-ML with probe, and HPML models with ground truth for on-ramp and off-ramp segment.
The lines of TFM, pure-ML and pure-ML with probe can’t fit the ground truth well. It can be clearly seen that the line of HPML model better fit the ground truth, which further indicates that HPML could obtain the best performance on all segments.

Figure 4: TSE estimates v.s. ground truth
The performance of different models in peak hours and off-peak hours are also compared in Fig. 5. Peak hour includes six hours (7am – 10am and 16pm – 19pm) and off-peak hour includes eighteen hours (midnight – 7am, 10am – 16pm, and 16pm – midnight). As shown in the figure, the RMSE of all HPML models in peak hours are higher than off-peak hours. It indicates that the HPML model could obtain better estimation accuracy under low traffic volume condition. The HPML model performance degradation during peak hour may be due traffic congestion or traffic crash.

Figure 5: HPML model performance comparison between peak hours and off-peak hours

5. Conclusions and Future Research Directions

Accurate TSE plays a critical role in the success of ITS on freeways. The accuracy of TSE tends to be affected by the limitation of data quality and quantity. To overcome these issues, this paper developed an innovative HPML approach, which uses the output of second-order macroscopic traffic flow model as additional training inputs of ML models to improve TSE accuracy. To evaluate the effectiveness of HPML model, it was deployed on interstate freeway I-15 in Salt Lake City, Utah. The proposed HPML model could offer higher resolution, network-wide and more accurate traffic information, which can concurrently offset the limitations of both point and probe data. In addition, the proposed hybrid ML brings transportation theoretical foundations to ML-based methodologies combine the advantages of both model-driven and data-driven approach for TSE. It could be used as a basis for similar traffic research concerns in the future. Moreover, this method combines machine learning models with hybrid data from a physics-based traffic model, which do not require a massive amount of data from traffic sensors. This will allow the government to save a huge amount of money on traffic detector installation.

The effectiveness of the proposed HPML TSE framework has been approved. Future research directions include: 1) more efficient machine learning algorithm and traffic physical model and its application on urban freeway network are worth studying; 2) the proposed method can be used for missing data supplement and validation of traffic detectors data.
Acknowledgement

This research is supported by the grant “CAREER: Physics Regularized Machine Learning Theory: Modeling Stochastic Traffic Flow Patterns for Smart Mobility Systems” sponsored by the National Science Foundation and grant “MPC-657 - Knowledge-Based Machine Learning for Freeway COVID-19 Traffic Impact Analysis and Traffic Incident Management” funded by the US Department of Transportation Mountain-Plains Consortium (MPC).

References

Allström, A., Ekström, J., Gundlegård, D., Ringdahl, R., Rydnergren, C., Bayen, A.M., Patire, A.D., 2016. Hybrid approach for short-term traffic state and travel time prediction on highways. Transp. Res. Rec. 2554, 60–68. https://doi.org/10.3141/2554-07

Anusha, S.P., Anand, R.A., Vanajakshi, L., 2012. Data fusion based hybrid approach for the estimation of urban arterial travel time. J. Appl. Math. 2012. https://doi.org/10.1155/2012/587913

Asif, M.T., Dauwels, J., Goh, C.Y., Oran, A., Fathi, E., Xu, M., Dhanya, M.M., Mitrovic, N., Jaillet, P., 2014. Spatiotemporal patterns in large-scale traffic speed prediction. IEEE Trans. Intell. Transp. Syst. 15, 794–804. https://doi.org/10.1109/TITS.2013.2290285

Aw, A., Rascle, M., 2000. Resurrection of “second order” models of traffic flow 60, 916–938.

Boser, B.E., Vapnik, V.N., Guyon, I.M., 1992. Training Algorithm Margin for Optimal Classifiers. Perceptron 144–152.

Cutler, A., Cutler, D.R., Stevens, J.R., 2012. Random forests. Ensemble Mach. Learn. Methods Appl. 157–175. https://doi.org/10.1007/9781441993267_5

Duan, Y., Lv, Y., Liu, Y., Wang, F., 2016. An efficient realization of deep learning for traffic data imputation. Transp. Res. Part C 72, 168–181. https://doi.org/10.1016/j.trc.2016.09.015

Hammer, B., 2010. Predicting travel times with context-dependent random forests by modeling local and aggregate traffic flow. Proc. - IEEE Int. Conf. Data Mining, ICDM 1357–1359. https://doi.org/10.1109/ICDMW.2010.128

Hofleitner, A., Herring, R., Bayen, A., 2012. Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning. Transp. Res. Part B Methodol. 46, 1097–1122. https://doi.org/10.1016/j.trb.2012.03.006

Kumar, S. V., Chaitanya Dogiparthi, K., Vanajakshi, L., Subramanian, S.C., 2017. Integration of exponential smoothing with state space formulation for bus travel time and arrival time prediction. Transport 32, 358–367.

Leshem, G., Ritov, Y., 2007. Traffic flow prediction using adaboost algorithm with random forests as a weak learner. Int. J. Math. Comput. Sci. 1, 193–198. https://doi.org/10.1007/s11117-011-0122-z

Li, L., Li, Y., Li, Z., 2013. Efficient missing data imputing for traffic flow by considering temporal and spatial dependence. Transp. Res. Part C Emerg. Technol. 34, 108–120.
Lighthill, M. J., Whitham, G. B., 1955. On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. Proc. R. Soc. London. Ser. A, Math. Phys. Sci. 229, 317–345.

Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.Y., 2015. Traffic Flow Prediction with Big Data: A Deep Learning Approach. IEEE Trans. Intell. Transp. Syst. 16, 865–873. https://doi.org/10.1109/TITS.2014.2345663

Ma, X., Tao, Z., Wang, Yinhai, Yu, H., Wang, Yunpeng, 2015. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp. Res. Part C Emerg. Technol. 54, 187–197. https://doi.org/10.1016/j.trc.2015.03.014

Ni, D., Leonard, J.D., 2005. Markov chain Monte Carlo multiple imputation using Bayesian networks for incomplete intelligent transportation systems data. Transp. Res. Rec. 57–67. https://doi.org/10.3141/1935-07

Papageorgiou, M., Blosseville, J.M., Hadj-Salem, H., 1989a. Macroscopic modelling of traffic on the boulevard peripherique in paris. Transp. Res. Part B Methodol. 23, 29–47.

Papageorgiou, M., Blosseville, J.M., Hadj-Salem, H., 1989b. Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris. Transp. Res. Part B Methodol. 23, 29–47.

PAYNE, H.J., 1971. Models of freeway traffic and control. Mathematical Models of Public Systems. Simulation Councils.

Polson, N., Sokolov, V., 2018. Bayesian Particle Tracking of Traffic Flows. IEEE Trans. Intell. Transp. Syst. 19, 345–356. https://doi.org/10.1109/TITS.2017.2650947

Richards, P.I., 1956. Shock Waves on the Highway. INFORMS 4, 42–51.

Seo, T., Bayen, A.M., Kusakabe, T., Asakura, Y., 2017. Traffic state estimation on highway: A comprehensive survey. Annu. Rev. Control 43, 128–151. https://doi.org/10.1016/j.arcontrol.2017.03.005

Sharmila, R.B., Velaga, N.R., Kumar, A., 2019. SVM-based hybrid approach for corridor-level travel-time estimation. IET Intell. Transp. Syst. 13, 1429–1439.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Tak, S., Woo, S., Yeo, H., 2016. Data-Driven Imputation Method for Traffic Data in Sectional Units of Road Links. IEEE Trans. Intell. Transp. Syst. 17, 1762–1771. https://doi.org/10.1109/TITS.2016.2530312

Tan, H., Wu, Y., Cheng, B., Wang, W., Ran, B., 2014. Nonnegativity and Road Capacity 2014.

Tang, J., Zhang, G., Wang, Y., Wang, H., Liu, F., 2015. A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation. Transp. Res. Part C Emerg. Technol. 51, 29–40. https://doi.org/10.1016/j.trc.2014.11.003

Tian, Y., Zhang, K., Li, J., Lin, X., Yang, B., 2018. LSTM-based traffic flow prediction with
missing data. Neurocomputing 318, 297–305. https://doi.org/10.1016/j.neucom.2018.08.067

Vapnik, V., 2013. The nature of statistical learning theory. Springer science & business media.

Vlahogianni, E.I., Karlaftis, M.G., Golas, J.C., 2014. Short-term traffic forecasting: Where we are and where we’re going. Transp. Res. Part C Emerg. Technol. 43, 3–19. https://doi.org/10.1016/j.trc.2014.01.005

Wang, F.Y., 2010. Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications. IEEE Trans. Intell. Transp. Syst. 11, 630–638. https://doi.org/10.1109/TITS.2010.2060218

Wang, J., Chen, R., He, Z., 2019. Traffic speed prediction for urban transportation network: A path based deep learning approach. Transp. Res. Part C Emerg. Technol. 100, 372–385. https://doi.org/10.1016/j.trc.2019.02.002

Wang, Y., Papageorgiou, M., 2005. Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. Transp. Res. Part B Methodol. 39, 141–167. https://doi.org/10.1016/j.trb.2004.03.003

Whitham, G.B., 1975. Linear and nonlinear waves. Modern Book Incorporated.

Wu, C., Ho, J., Lee, D.T., 2004. Travel-Time Prediction With Support Vector Regression 5, 276–281.

Wu, Y., Tan, H., Qin, L., Ran, B., Jiang, Z., 2018. A hybrid deep learning based traffic flow prediction method and its understanding. Transp. Res. Part C Emerg. Technol. 90, 166–180. https://doi.org/10.1016/j.trc.2018.03.001

You, J., Kim, T.J., 2000. Development and evaluation of a hybrid travel time forecasting model. Transp. Res. Part C Emerg. Technol. 8, 231–256. https://doi.org/10.1016/S0968-090X(00)00012-7

Yu, B., Yang, Z.Z., Yu, K.C.B., 2010. Hybrid model for prediction of bus arrival times at next station. J. Adv. Transp. 44, 193–204. https://doi.org/DOI: 10.1002/atb.136

Yuan, Y., Yang, X.T., Zhang, Z., Zhe, S., 2020. Macroscopic traffic flow modeling with physics regularized gaussian process: A new insight into machine learning applications.

Zhang, H.M., 2002. A non-equilibrium traffic model devoid of gas-like behavior. Transp. Res. Part B 36, 275–290.

Zhang, J., Wang, F.Y., Wang, K., Lin, W.H., Xu, X., Chen, C., 2011. Data-driven intelligent transportation systems: A survey. IEEE Trans. Intell. Transp. Syst. 12, 1624–1639. https://doi.org/10.1109/TITS.2011.2158001

Zhang, Y., Ge, H., 2013. Freeway travel time prediction using takagi-sugeno-kang fuzzy neural network. Comput. Civ. Infrastruct. Eng. 28, 594–603. https://doi.org/10.1111/mice.12014

Zhang, Y., Liu, Y., 2009. Traffic forecasting using least squares support vector machines. Transportmetrica 5, 193–213. https://doi.org/10.1080/18128600902823216

Zhang, Z., Xianfeng Yang, 2019. Freeway Traffic Speed Estimation by Regression Machine
Learning Techniques Using Probe Vehicle and Sensor Detector Data. J. Transp. Eng. Part A Syst.

Zhang, Z., Yuan, Y., Yang, X.T., 2020. A Hybrid Machine Learning Approach for Freeway Traffic Speed Estimation. https://doi.org/10.1177/0361198120935875

Zhong, M., Lingras, P., Sharma, S., 2004. Estimation of missing traffic counts using factor, genetic, neural, and regression techniques. Transp. Res. Part C Emerg. Technol. 12, 139–166. https://doi.org/10.1016/j.trc.2004.07.006

Zhu, L., Guo, F., Polak, J.W., Krishnan, R., 2018. Urban link travel time estimation using traffic states-based data fusion. IET Intell. Transp. Syst. 12, 651–663. https://doi.org/10.1049/iet-its.2017.0116

Zou, Y., Zhu, X., Zhang, Y., Zeng, X., 2014. A space-time diurnal method for short-term freeway travel time prediction. Transp. Res. Part C Emerg. Technol. 43, 33–49. https://doi.org/10.1016/j.trc.2013.10.007