Learning "Look-Ahead" Nonlocal Traffic Dynamics in a Ring Road

Chenguang Zhao   CZHAO704@CONNECT.HKUST-GZ.EDU.CN
Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou)

Huan Yu[*]   HUANYU@UST.HK
Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou)
Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology

Editors: A. Abate, M. Cannon, K. Margellos, A. Papachristodoulou

Abstract

The macroscopic traffic flow model is widely used for traffic control and management. To incorporate drivers' anticipative behaviors and to remove impractical speed discontinuity inherent in the classic Lighthill–Whitham–Richards (LWR) traffic model, nonlocal partial differential equation (PDE) models with "look-ahead" dynamics have been proposed, which assume that the speed is a function of weighted downstream traffic density. However, it lacks data validation on two important questions: whether there exist nonlocal dynamics, and how the length and weight of the "look-ahead" window affect the spatial temporal propagation of traffic densities. In this paper, we adopt traffic trajectory data from a ring-road experiment and design a physics-informed neural network to learn the fundamental diagram and look-ahead kernel that best fit the data, and reinvent a data-enhanced nonlocal LWR model via minimizing the loss function combining the data discrepancy and the nonlocal model discrepancy. Results show that the learned nonlocal LWR yields a more accurate prediction of traffic wave propagation in three different scenarios: stop-and-go oscillations, congested, and free traffic. We first demonstrate the existence of "look-ahead" effect with real traffic data. The optimal nonlocal kernel is found out to take a length of around 35 to 50 meters, and the kernel weight within 5 meters accounts for the majority of the nonlocal effect. Our results also underscore the importance of choosing a priori physics in machine learning models.

Keywords: Traffic flow model, Nonlocal traffic dynamics, Physics-constrained learning

1. Introduction

Macroscopic traffic flow models describe the dynamics of aggregated traffic states, density, speed, and flow, by partial differential equations (PDEs). It serves as the basis for various traffic management designs, such as those alleviating congestion (Yu and Krstic, 2022), improving throughput (Smith et al., 2019), and reducing emissions (Rodriguez-Rey et al., 2021). The first and widely used traffic flow model is the Lighthill–Whitham–Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), which describes the dynamics of density as a first-order hyperbolic PDE. It is derived from the conservation law and assumes that speed is decided by a static function of local density, also known as the fundamental diagram. Despite the simplicity, it causes shock waves in finite time with smooth initial conditions. The discontinuity of speed is not realistic in real traffic as the acceleration becomes infinite at the shock wave. To avoid discontinuity, the nonlocal LWR model has been proposed in which the speed is decided by a weighted average of downstream or upstream traffic density within a finite length (Blandin and Goatin, 2016). Human drivers have "look-ahead" of downstream traffic because they can anticipate and react not only to local traffic as in preceding leader vehicles, but may also to traffic conditions further downstream within the field of

---

[*] corresponding author

vision. The nonlocal effect can be amplified for traffic on a ring road, as shown in Fig. 1, where the span of drivers' vision is largely enhanced due to the curvature of the ring road, compared to straight roads. When there are connected automated vehicles that are controlled based on upstream vehicle information received from vehicle-to-vehicle communication, the traffic will present "look-behind" nonlocal property.

Theoretical analysis has proved that both "look-ahead" and "look-behind" nonlocal extensions of LWR generate smooth solutions under certain assumptions on the initial condition, boundary condition, and model parameters (Karafyllis et al., 2022). Besides, "look-behind" nonlocal controllers will increase traffic capacity (Karafyllis et al., 2022). Some analytical properties of the nonlocal LWR model have also been discussed, such as well-posedness (Goatin and Scialanga, 2016), controllability and stability (Bayen et al., 2021; Huang and Du, 2022), the vanishing nonlocality limit, (Colombo et al., 2019; Keimer and Pflug, 2019). However, to the best of our knowledge, there has been little research characterizing nonlocal traffic dynamics with field data. It remains an open question whether nonlocal dynamics exist in real traffic, and if so, what form the nonlocal function and the fundamental diagram can be validated with data. In this paper, we carefully choose the trajectory data of human drivers collected from a ring-road setting and present the first result on analyzing the nonlocal traffic dynamics, particularly "look-ahead" from the data.

Real traffic data has been collected by many researchers, such as Lagrangian data collected by vehicle sensors (Sun et al., 2020; Zheng et al., 2021) and Euclidean data collected from loop detectors or surveillance cameras (FHWA, 2007; Stern et al., 2018; Gloudemans et al., 2023; Krajewski et al., 2018). The data has been used to calibrate local LWR models via various methods, such as least square (Dervisoglu et al., 2009), Nelder-Mead method (Kontorinaki et al., 2017; Nelder and Mead, 1965), and genetic algorithm (Mohammadian et al., 2021). This paper conducts the calibration of nonlocal LWR models with ring-road traffic data. It is a more challenging task than the local ones. In local LWR models, the fundamental diagram reflects the dependence of speed and local density, which are directly measurable from detectors. While in nonlocal LWR models, the nonlocality results in coupling between nonlocal kernel function and local density-speed relation. The coupling is then embedded into a dynamic spatial-temporal propagation of traffic density, which makes it very difficult to identify from the data. Recently developed Physics-informed deep learning (PIDL) (Raissi et al., 2019), a physics-constrained learning approach, has been adopted to reinvent the local LWR model with data enhancement (Shi et al., 2021; Zhao and Yu, 2023). The PIDL traffic model achieves more accurate prediction than the pure physical model, since a trade-off between data and model can be achieved with a physics-uninformed neural network (NN) to learn the mapping from spatial-temporal location to system state, and a physics-informed NN to learn the system dynamics. In this paper, we tackle the nonlocality with PIDL and design an NN with physics constraints to learn the optimal fundamental diagram and kernel function.

The main contribution of this paper lies in first characterizing, learning, and analyzing the "look-ahead" traffic dynamics that best fit experiment traffic data. We adopt PIDL to learn the optimal fundamental diagram and kernel function. The NN is optimized to minimize a loss function of three components: a data loss that reflects the discrepancy between learned density and ground truth measurement, a physics dynamics loss that evaluates that discrepancy between learned dynamics and the nonlocal LWR model, and a physics static loss designed to satisfy constraints on the fundamental diagram the kernel function for well-posedness of the nonlocal LWR model. Based on the learned kernel function and fundamental diagram, we find that the nonlocal LWR yields a more accurate estimation of traffic dynamics, i.e., the propagation of traffic waves.
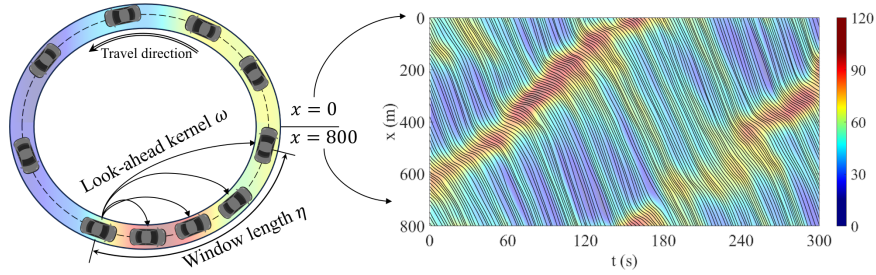
Figure 1: Ring road traffic dynamics. Left: ring road experiment setting. Right: vehicle trajectory (black line) and spatial-temporal propagation of density (veh per km).

2. Ring-road data and nonlocal traffic flow model

2.1. Ring-road traffic data

Several experiments have been conducted to collect traffic data on ring-roads. In (Sugiyama et al., 2008), the authors arrange 22 vehicles on a ring road with a circumference of 230 m. In (Stern et al., 2018), experiments are conducted with 22 vehicles on a ring road with a circumference of 230 m. The most up-to-date experiment (Zheng et al., 2021) collects data from 40 vehicles traveling on an 800-meter circumference ring road, which we will use in this paper. Each vehicle is equipped with high precision GPS to record its location and speed with a frequency of 10 Hz. Measurement error via the GPS is of ± 1 m for location and ± 1 km/h for velocity. To reconstruct the macroscopic density and speed from recorded vehicle trajectories, we adopt kernel density estimation (KDE) with a Gaussian kernel (Fan and Seibold, 2013; Parzen, 1962). The reconstructed states are in a discrete domain $\mathcal{G} = \{(t_i, x_j)|i = 0, \cdots, N_t - 1; x = 0, \cdots, N_x - 1\}$. We take the cell interval being $\Delta t = 1$ s and $\Delta x = 1$ m. The reconstructed traffic density is visualized in Fig. 1.

2.2. Nonlocal traffic flow model

We consider the macroscopic traffic dynamics on a ring road of length $L > 0$. Based on the conservation law of vehicle numbers, we have the following hyperbolic PDE:

$$\partial_t \rho(x,t) + \partial_x (\rho(x,t)v(x,t)) = 0, \tag{1}$$

where $\rho(x,t)$ and $v(x,t)$ represent the density and speed at time $t \geq 0$ and location $0 \leq x \leq L$ respectively. On a ring road, we have periodic boundary conditions, i.e., $\rho(x + L, t) = \rho(x, t)$. The LWR model assumes that speed $v(x,t)$ is dependent on the local density $\rho(x,t)$ by the fundamental diagram $V(\rho) : \mathbb{R} \to \mathbb{R}$, i.e., $v(x,t) = V(\rho(x,t))$. The evolution of traffic density is then given as:

$$\partial_t \rho(x,t) + \partial_x \left(\rho(x,t) V(\rho(x,t)) \right) = 0. \tag{2}$$

Considering the physics constraints, we have Assumption 1 on the fundamental diagram $V(\rho)$.

Assumption 1 *The fundamental diagram $V(\rho)$ is a non-negative, non-increasing function.*

To describe the dependence between speed and density, some closed-form functions with parameters have been adopted as the fundamental diagram. (Greenshields et al., 1935) takes the first step to calibrate the fundamental diagram as a linear function:

$$V(\rho) = v_f \left(1 - \rho/\rho_m \right), \tag{3}$$

3

with $v_f > 0$ and $\rho_m > 0$ being parameters representing free speed and maximum density. In (Underwood, 1961), the fundamental diagram is:

$$V(\rho) = v_f \exp\left(-\rho/\rho_c\right), \tag{4}$$

with the two parameters $v_f > 0$ and $\rho_c > 0$ being free speed and critical density respectively. (Drake et al., 1966) describes the fundamental diagram via:

$$V(\rho) = v_f \exp\left(-(\rho/\rho_c)^2/2\right). \tag{5}$$

Besides using closed-form functions with parameters that have explicit physical meaning, researchers have also focused on getting the fundamental diagram via data-driven methods, such as generalized least square method in (Qu et al., 2015), machine learning in (Shi et al., 2021), and Gaussian process in (Cheng et al., 2022). However, all these methods focus on the calibration of speed and local density. It is unclear how these models fit in with nonlocal dynamics.

Unlike the local LWR model, the "look-ahead" nonlocal model assumes that the speed at (x, t) depends not on the local density $\rho(x, t)$, but instead the weighted average of downstream traffic density within a length of $\eta > 0$. The speed becomes $v(x, t) = V_\eta(\rho_\eta(x, t))$, where the nonlocal traffic density takes the form of

$$\rho_\eta(x, t) = \int_0^\eta \rho(x + y, t)\omega(y)\mathrm{d}y, \tag{6}$$

with $\omega : [0, \eta] \to \mathbb{R}$ being the integral nonlocal kernel function. Similarly, for "look-behind" nonlocal LWR, the nonlocal traffic density is a weighted average of upstream traffic. The nonlocal PDE traffic flow model is

$$\partial_t \rho(x, t) + \partial_x\left(\rho(x, t)V_\eta(\rho_\eta(x, t))\right) = 0. \tag{7}$$

For well-posedness of the nonlocal PDE (7), besides the same constraints on the fundamental diagram V_η as in Assumption 1, we also have constraints on the kernel function $\omega(x)$ as follows.

Assumption 2 *The function $\omega(x)$ is a non-negative, non-increasing function with $\int_0^\eta \omega(x)\mathrm{d}x = 1$.*

Theorem 1 *(Karafyllis et al., 2022) Under Assumption 1 on V_η and Assumption 2 on ω, for every initial condition $\rho(t, 0) \in W^{2,\infty}(\mathbb{R}) \cap \mathrm{Per}(\mathbb{R})$, the initial value problem for the nonlocal LWR (7) has a unique solution $\rho(x, t) \in C^1(\mathbb{R}^+ \times \mathbb{R})$ with $\rho(t, \cdot) \in W^{2,\infty}(\mathbb{R}) \cap \mathrm{Per}(\mathbb{R})$ for all $t \geq 0$, where $W^{2,\infty}(\mathbb{R})$ is the Sobolev space of C^1 functions on $\mathbb{R}$ with Lipschitz derivative, and $\mathrm{Per}(\mathbb{R})$ is the set of continuous, positive mappings $\rho : \mathbb{R} \to (0, \infty)$ with a period of L.*

Two commonly used kernel functions take the constant or linear decreasing function form as:

$$\omega(x) = 1/\eta, \tag{8}$$
$$\omega(x) = 2(\eta - x)/\eta^2. \tag{9}$$

Although these two kernel functions satisfy the conditions in Assumption 2, it is unclear how well they fit real traffic flow. In this paper, we learn the kernel function that best fits real traffic data.

3. Learning fundamental diagram, look-ahead kernel, & spatial-temporal dynamics

We give in Fig. 2 the constructed NN. There are three parameters to be trained: θ to learn the spatial-temporal dynamics of density $\rho(x, t)$; θ_ω to learn the nonlocal look-ahead kernel ω, and θ_v to learn
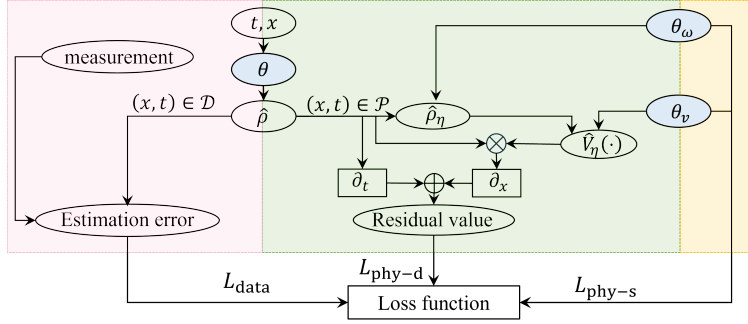
Figure 2: Diagram of the proposed NN. We use $\mathcal{D}$ to evaluate the data discrepancy between learned density and ground truth, and $\mathcal{P}$ to evaluate the model discrepancy between the learned dynamics and the nonlocal LWR.

the fundamental diagram V_η. We design three loss functions: data loss $L_{\text{data}}(\theta)$ in the red box to describe the discrepancy between the learned density and measurement, physics dynamics loss $L_{\text{phy}-\text{d}}$ in the green box to describe the discrepancy between the learned dynamics and nonlocal traffic flow model, and physics static loss $L_{\text{phy}-\text{s}}$ in the yellow box to constrain the learned kernel and fundamental diagram to ensure well-posedeness of the nonlocal PDE. The overall loss function is:

$$L(\theta, \theta_v, \theta_\omega) = L_{\text{data}}(\theta, \theta_v, \theta_\omega) + L_{\text{phy}-\text{d}}(\theta, \theta_v, \theta_\omega) + L_{\text{phy}-\text{s}}(\theta_v, \theta_\omega). \tag{10}$$

We specify the loss functions as follows.

3.1. Data loss

We use an NN with parameter θ to learn the density at location x and time t as $\hat{\rho}(x, t; \theta)$. We assume that the initial condition $\rho(0, x)$ is given, and there are N_l loop detectors located at $x = l_i$ to measure the density $\rho(t, l_i)$. So we have measured traffic density at $\mathcal{D} = \{(0, x_i)|i = 0, \cdots N_x - 1\} \cup \{(t_i, l_j)|i = 0, \cdots, N_t - 1; j = 0, \cdots N_l - 1\}$. To minimize the difference between learned density $\rho(x, t; \theta)$ and the measurement density $\rho(x, t)$, the data loss is designed as

$$L_{\text{data}}(\theta) = \alpha_{-1} L_{\text{initial}}(\theta) + \sum_{i=0}^{N_l-1} \alpha_i L_{\text{detector}}(l_i; \theta), \tag{11}$$

where $\alpha_{-1} > 0$ and $\alpha_i > 0$ are coefficients, $L_{\text{initial}}(\theta)$ and $L_{\text{detector}}(\theta)$ are the estimation error on the initial data and loop detector data respectively:

$$L_{\text{initial}}(\theta) = \frac{1}{N_x} \sum_{i=0}^{N_x} \left(\rho(i\Delta x, 0) - \hat{\rho}(i\Delta x, 0; \theta) \right)^2. \tag{12}$$

$$L_{\text{detector}}(l_i; \theta) = \frac{1}{N_t} \sum_{k=0}^{N_t} \left(\rho(t_k, l_i) - \hat{\rho}(t_k, l_i; \theta) \right)^2. \tag{13}$$

3.2. Physics loss

Physics dynamics loss: Since NN cannot directly calculate integral in the nonlocal density, we adopt the numerical scheme with convergence proved in (Karafyllis et al., 2022). A numerically

approximated nonlocal density is:

$$\bar{\rho}_\eta(t_i, x_j; \theta, \theta_\omega) = \sum_{k=0}^{N_\eta - 1} \rho(t_i, x_j + k\Delta x)\bar{\omega}_k, \tag{14}$$

where $N_\eta = \eta/\Delta x$, and $\bar{\omega}_k = \int_{k\Delta x}^{(k+1)\Delta x} \omega(s)ds$. We use a trainable vector $\theta_\omega \in \mathbb{R}^{N_\eta}$ to learn the kernel $\bar{\omega}$, and the learned kernel $\hat{\bar{\omega}}(\theta_\omega)$ is:

$$\hat{\bar{\omega}}_i(\theta_\omega) = \frac{\theta_{\omega,i}}{\mathbf{1} \cdot \theta_\omega}. \tag{15}$$

The learned nonlocal density is

$$\hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega) = \sum_{k=0}^{N_\eta - 1} \hat{\rho}(t_i, x_j + k\Delta x; \theta)\hat{\bar{\omega}}_k(\theta_\omega). \tag{16}$$

We use a neural network with parameter θ_v to learn the fundamental diagram $V_\eta(\cdot)$. The NN takes density as input and outputs the estimated speed $\hat{V}_\eta(\cdot; \theta_v)$. To evaluate the discrepancy between the learned dynamics and the nonlocal LWR model, we define a residual value as

$$\begin{aligned} f(t_i, x_j; \theta, \theta_v, \theta_\omega) =& \partial_t \hat{\rho}(t_i, x_j; \theta) + \partial_x \left(\hat{\rho}(t_i, x_j; \theta)\hat{V}_\eta(\hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega); \theta_v) \right) \\ =& \partial_t \hat{\rho}(t_i, x_j; \theta) + \partial_x \hat{\rho}(t_i, x_j; \theta)\hat{V}_\eta(\hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega); \theta_v) \\ &+ \hat{\rho}(t_i, x_j; \theta)\partial_\rho \hat{V}_\eta(\hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega); \theta_v)\partial_x \hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega), \end{aligned} \tag{17}$$

where $\partial_x \hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega)$ is given by the same numerical discretization in (16) as

$$\partial_x \hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega) = \sum_{k=0}^{N_\eta - 1} \partial_x \hat{\rho}(x, t + k\Delta x; \theta)\hat{\bar{\omega}}_k(\theta_\omega); \tag{18}$$

the partial differentiation $\partial_t \hat{\rho}(t_i, x_j; \theta)$, $\partial_x \hat{\rho}(t_i, x_j; \theta)$, and $\partial_\rho \hat{V}_\eta(\hat{\rho}_\eta(t_i, x_j; \theta, \theta_\omega); \theta_v)$ are calculated via automatic differentiation in Tensorflow. We design a physics loss to minimize the residual value:

$$L_{\mathrm{phy}}(\theta, \theta_v, \theta_\omega) = \frac{1}{N_p} \sum_{(i,j) \in P} \left(f(t_i, x_j; \theta, \theta_v, \theta_\omega) \right)^2, \tag{19}$$

where $\mathcal{P} \in \mathcal{G}$ are temporal-spatial points selected from the whole domain to evaluate the learned dynamics.

Physics static loss: Given the conditions on $\omega(x)$ in Assumption 2, we have three conditions on $\bar{\omega}_k$:

$$\bar{\omega}_k \geq 0, \forall k = 0, 1, \cdots, N_\eta - 1, \tag{20}$$

$$\bar{\omega}_{k+1} \leq \bar{\omega}_k, \forall k = 0, 1, \cdots, N_\eta - 2, \tag{21}$$

$$\sum_{k=0}^{N_\eta - 1} \bar{\omega}_k = 1. \tag{22}$$

By the definition of the learned kernel $\hat{\bar{\omega}}$ (15), the constraint (22) is automatically satisfied. To meet the constraint (20), we define a penalty as:

$$L_{p,\omega,1}(\theta_\omega) = \sum_{i=0}^{N_\eta - 1} \left(\min\{\hat{\bar{\omega}}_i(\theta_\omega), 0\} \right)^2. \tag{23}$$

To meet the constraint (21), we define a penalty as:

$$L_{p,\omega,2}(\theta_\omega) = \sum_{i=0}^{N_\eta-2} \left(\max\{\hat{\bar{\omega}}_{i+1}(\theta_\omega) - \hat{\bar{\omega}}_i(\theta_\omega), 0\} \right)^2. \tag{24}$$

For the fundamental diagram, given the constraints in Assumption 1, to satisfy $\hat{V}_\eta(\rho; \theta_v) \geq 0$, we design a penalty as

$$L_{p,v,1} = \sum_{i=0}^{N_\rho-1} \left(\min\left\{ \hat{V}_\eta(i\Delta\rho; \theta_v), 0 \right\} \right)^2, \tag{25}$$

where $N_\rho = \rho_m/\Delta\rho$ with ρ_m being the maximum density. To make $\hat{V}_\eta(\rho; \theta_v)$ be non-increasing with respect to the density ρ, we design a penalty as

$$L_{p,v,2} = \sum_{i=0}^{N_\rho-1} \left(\max\left\{ \partial_\rho \hat{V}_\eta(i\Delta\rho; \theta_v), 0 \right\} \right)^2, \tag{26}$$

where $\partial_\rho \hat{V}_\eta$ is calculated using automatic differentiation provided in Tensorflow. The physics static loss is a weighted sum of the four penalties:

$$L_{\mathrm{phy-s}} = p_{\omega,1} L_{p,\omega,1}(\theta_\omega) + p_{\omega,1} L_{p,\omega,2}(\theta_\omega) + p_{v,1} L_{p,v,1}(\theta_v) + p_{v,2} L_{p,v,2}(\theta_v), \tag{27}$$

where $p_{v,1} > 0$, $p_{v,2} > 0$, $p_{v,3} > 0$, $p_{\omega,1} > 0$, and $p_{\omega,2} > 0$ are coefficients.

3.3. Training settings

For the variable θ, we use a fully connected feedforward neural network with six hidden layers and 64 neurons in each hidden layer. For the fundamental diagram θ_v, we use a fully connected feedforward neural network with two hidden layers and 64 neurons in each hidden layer.

We use $N_l = 5$ loop detectors evenly distributed among the ring road including one at the boundary $x = 0$. We train NN parameters by first running 50,000 iterations of ADAM for rough training and then using L-BFGS to refine the NN. We randomly choose $N_P = 512$ points from $\mathcal{G}$ to evaluate the physics dynamic loss $L_{\mathrm{phy-d}}$. We decide coefficients α_i in data loss L_{data} by grid search. We set the penalty coefficients as $p_{v,1} = p_{v,2} p_{\omega,1} = p_{\omega,2} = 10^4$.

4. Analyze nonlocal effect on ring-road traffic

In this section, we analyze the learned dynamics with local and nonlocal LWR PDEs from two aspects: traffic dynamics in Section 4.1 and static speed-density fundamental diagram in Section 4.2.

4.1. Nonlocal effect on traffic dynamics

To evaluate the error between learned density $\hat{\rho}$ and actual density, we use relative RMSE:

$$\mathrm{E}_\rho = \frac{\sqrt{\sum_{i,j} \left(\hat{\rho}(t_i, x_j) - \rho(t_i, x_j) \right)^2}}{\sqrt{\sum_{i,j} \left(\rho(t_i, x_j) \right)^2}} \times 100\%. \tag{28}$$

In Table 1, we give the estimation error with different fundamental diagrams and look-ahead kernels. From each row, we see that given a fundamental diagram, such as Underwood and Drake, using

There are three significant traffic features:
A: upstream propagation of congested waves (dark red)
B: downstream propagation of free waves (dark blue)
C: stop-and-go oscillation (light red, light blue, and yellow)

Greenshields (3)

(a) Local   (b) Constant kernel (8)   (c) Linear kernel (9)   (d) Learned kernel

Underwood (4)

(e) Local   (f) Constant kernel (8)   (g) Linear kernel (9)   (h) Learned kernel

Drake (5)

(i) Local   (j) Constant kernel (8)   (k) Linear kernel (9)   (l) Learned kernel

Learned

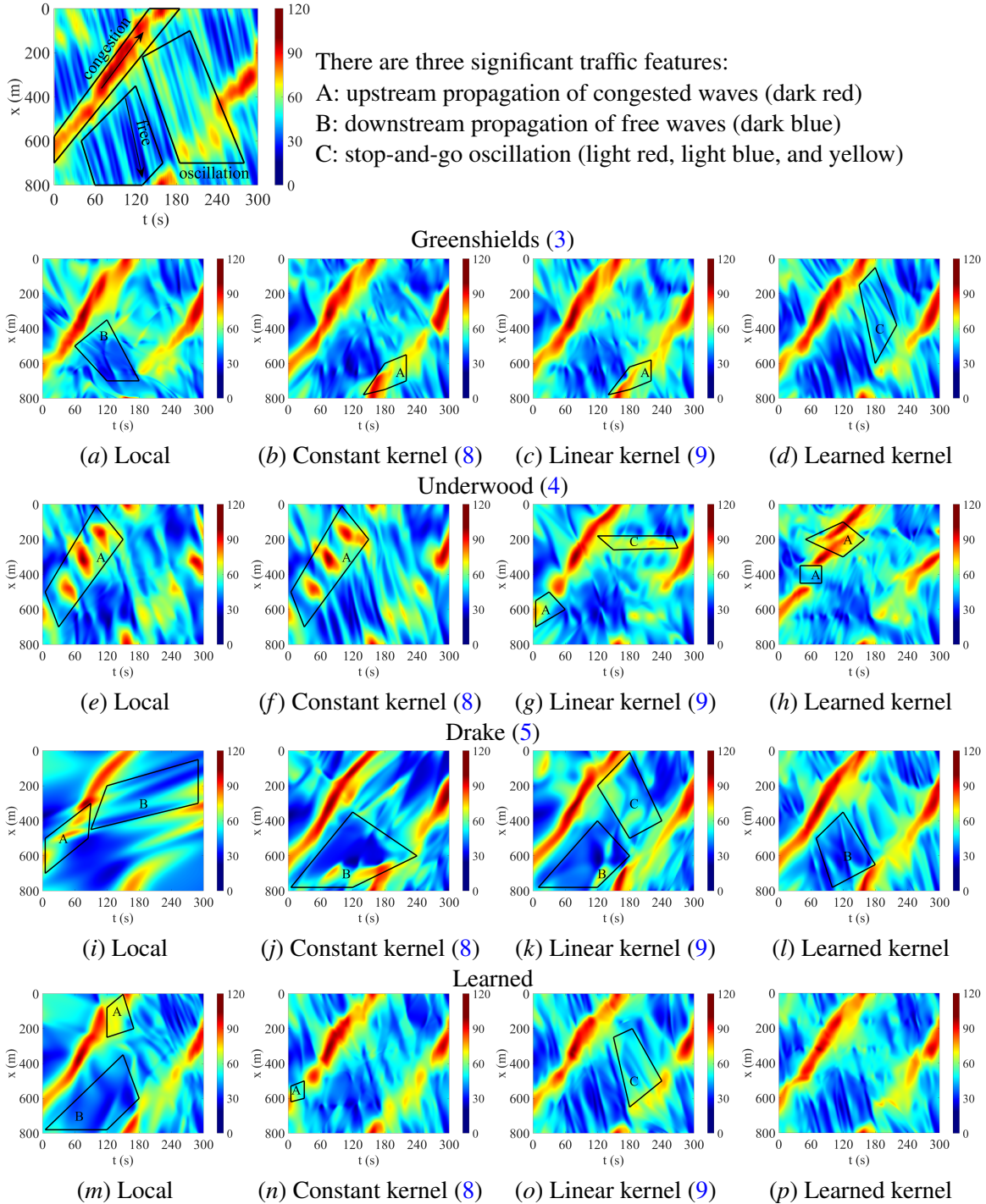(m) Local   (n) Constant kernel (8)   (o) Linear kernel (9)   (p) Learned kernel

Figure 3: The ground truth (first row) and learned density dynamics (subfigs (a) to (p)). The text specifies the fundamental diagram, and the subtitle shows the nonlocal kernel. We identify three types of traffic dynamics from the ground truth, and label inaccurate estimation in the learned dynamics.

Table 1: Estimation error (%) of density

FD \ Kernel	Local	Constant (8)	Linear (9)	Learned
Greenshields (3)	16.80	16.07	14.08	12.70
Underwood (4)	19.30	20.36	17.50	17.27
Drake (5)	18.71	20.62	17.82	14.66
Learned	14.92	12.70	12.20	11.90

$(a)\ \eta = 35$ m $(b)\ \eta = 40$ m $(c)\ \eta = 45$ m $(d)\ \eta = 50$ m
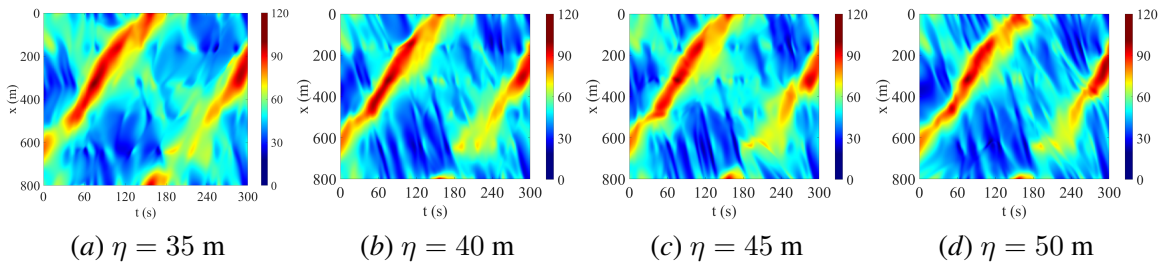
Figure 4: The learned density dynamics with varying kernel length η.

constant look-head kernel (8) may instead increase the estimation error. Using NN to learn a look-ahead kernel reduces the estimation error for all four cases of fundamental diagrams. Comparing the estimation error of the same look-ahead kernel and different fundamental diagrams in each column, we find that learning a fundamental diagram via NN consistently has lower estimation errors.

In Fig. 3, we plot the learned density dynamics. In each row, we see that with the learned look-ahead kernel, there is a more accurate traffic pattern than both the local LWR and nonlocal LWR with linear and constant kernels. Taking the last row as an example, we see that local LWR fails to learn the propagation of free traffic wave as Fig. 3(m) shows, and nonlocal LWR with a constant look-ahead kernel causes false propagation of free flow as Fig. 3(n) shows. This not only validates the existence of nonlocal effect in real traffic but also underscores the importance of choosing physics prior for PIDL.

Comparing the estimation obtained by different fundamental diagrams, we see that the learned fundamental diagram yields more accurate dynamics. For example, Fig. 3(d) shows that the Green-shields fundamental diagram estimates a distorted propagation of the stop-and-go oscillation. For the Underwood fundamental diagram, as the black box in Fig 3(h) shows, it fails to learn the congestion wave and gets false free traffic flow.

We run simulations with the kernel length η ranging from 5 m to 60 m, and find that 35 m to 50 m yields more accurate estimated dynamics. In Fig. 4, we give the learned density with varying kernel length. We see that all of them have an accurate estimation of traffic waves. In Fig. 5, we plot the learned kernel with different kernel length η. We see that the optimal kernel length has a similar pattern: a majority of the nonlocal effect is within 5 to 10 meters, which is approximately the length of one to two vehicles. For the kernel shown in the subfigures in Fig. 5, the look-ahead kernel within 4 meters accounts for $68\% - 80\%$ of the whole look-ahead effect. This indicates that the speed is mainly dictated by the vehicle and its leader vehicle.

4.2. Nonlocal effect on fundamental diagram

In Fig. 6, we give the scatter plot of the speed and nonlocal density under different look-ahead kernels with learned fundamental diagram. We also plot the calibrated fundamental diagram. As
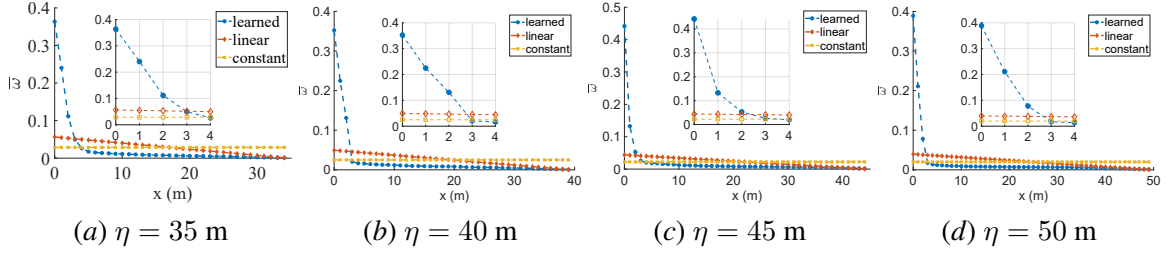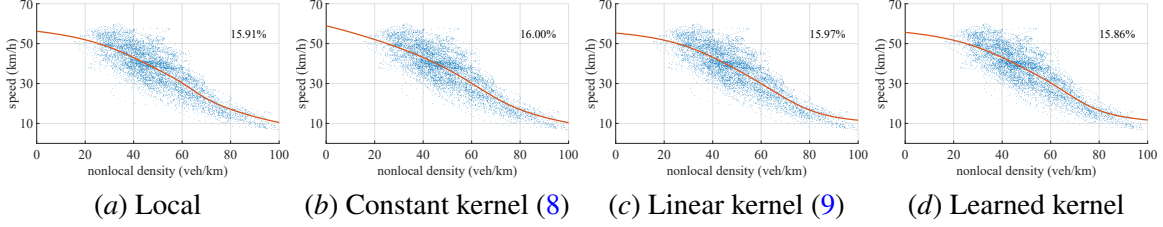
(*a*) $\eta = 35$ m (*b*) $\eta = 40$ m (*c*) $\eta = 45$ m (*d*) $\eta = 50$ m

Figure 5: The kernel function learned by NN with varying kernel length η.

(*a*) Local (*b*) Constant kernel (8) (*c*) Linear kernel (9) (*d*) Learned kernel

Figure 6: Learned fundamental diagram with local and nonlocal models.

shown in Fig. 6, in nonlocal LWR, there still exists a static relationship between speed and nonlocal density. To evaluate the calibration accuracy of fundamental diagram, we use relative RMSE as:

$$\mathrm{E}_v = \frac{\sqrt{\sum_{i,j} \left(\hat{V}_\eta(\hat{\rho}_\eta(t_i, x_j)) - v(t_i, x_j) \right)^2}}{\sqrt{\sum_{i,j} \left(v(t_i, x_j) \right)^2}} \times 100\%. \tag{29}$$

We give the calibration error in each subfig in Fig. 6. Comparing the fundamental diagram of local model in Fig. 6(a) and nonlocal models in Fig. 6(b)-(d), we see that the nonlocal term only has a marginal effect on the fundamental diagram. For other fundamental diagrams, Greenshields (3), Un-derWood (4), and Drake (5), we also find similar results, i.e., the fundamental diagram of nonlocal and local density is approximately the same.

5. Conclusion

In this paper, we learn the look-ahead kernel and fundamental diagram of nonlocal LWR model that best fit ring-road experimental traffic data using physics-informed deep learning. We found out that the traffic data indeed demonstrates look-ahead nonlocal phenomena, and the learned nonlocal LWR model yields a more accurate estimation of traffic wave propagation. Our result also underscored the importance of choosing physics prior in PIDL. For the optimal look-ahead kernel, we found that one to two leader vehicles account for a majority of the anticipation effect in the ring road setting. For the fundamental diagram, we found that look-ahead has only a marginal effect on the calibrated fundamental diagram. The future extension of this paper includes analyzing nonlocal effects in straight roads and control design for nonlocal LWR models.

10

References

Alexandre Bayen, Jean-Michel Coron, Nicola De Nitti, Alexander Keimer, and Lukas Pflug. Boundary controllability and asymptotic stabilization of a nonlocal traffic flow model. *Vietnam Journal of Mathematics*, 49(3):957–985, 2021.

Sebastien Blandin and Paola Goatin. Well-posedness of a conservation law with non-local flux arising in traffic flow modeling. *Numerische Mathematik*, 132(2):217–241, 2016.

Zhanhong Cheng, Xudong Wang, Xinyuan Chen, Martin Trépanier, and Lijun Sun. Bayesian calibration of traffic flow fundamental diagrams using gaussian processes. *IEEE Open Journal of Intelligent Transportation Systems*, 3:763–771, 2022.

Maria Colombo, Gianluca Crippa, and Laura V Spinolo. On the singular local limit for conservation laws with nonlocal fluxes. *Archive for Rational Mechanics and Analysis*, 233:1131–1167, 2019.

Gunes Dervisoglu, Gabriel Gomes, Jaimyoung Kwon, Roberto Horowitz, and Pravin Varaiya. Automatic calibration of the fundamental diagram and empirical observations on capacity. In *Transportation Research Board 88th Annual Meeting*, volume 15, pages 31–59, 2009.

Joseph S Drake, Joseph L Schofer, and Adolf D May Jr. A statistical analysis of speed-density hypotheses. *Highway Research Record*, 154:53–87, 1966.

Shimao Fan and Benjamin Seibold. Data-fitted first-order traffic models and their second-order generalizations: Comparison by trajectory and sensor data. *Transportation research record*, 2391 (1):32–43, 2013.

FHWA. Next Generation Simulation (NGSIM), 2007. URL https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm.

Derek Gloudemans, Yanbing Wang, Junyi Ji, Gergely Zachar, William Barbour, Eric Hall, Meredith Cebelak, Lee Smith, and Daniel B Work. I-24 motion: An instrument for freeway traffic science. *Transportation Research Part C: Emerging Technologies*, 155:104311, 2023.

Paola Goatin and Sheila Scialanga. Well-posedness and finite volume approximations of the lwr traffic flow model with non-local velocity. *Networks and Heterogeneous Media*, 11(1):107–121, 2016.

Bruce D Greenshields, JR Bibbins, WS Channing, and HH Miller. A study of traffic capacity. In *Highway research board proceedings*, volume 14, pages 448–477. Washington, DC, 1935.

Kuang Huang and Qiang Du. Stability of a nonlocal traffic flow model for connected vehicles. *SIAM Journal on Applied Mathematics*, 82(1):221–243, 2022.

Iasson Karafyllis, Dionysios Theodosis, and Markos Papageorgiou. Analysis and control of a nonlocal pde traffic flow model. *International Journal of Control*, 95(3):660–678, 2022.

Alexander Keimer and Lukas Pflug. On approximation of local conservation laws by nonlocal conservation laws. *Journal of Mathematical Analysis and Applications*, 475(2):1927–1955, 2019.

Maria Kontorinaki, Anastasia Spiliopoulou, Claudio Roncoli, and Markos Papageorgiou. First-order traffic flow models incorporating capacity drop: Overview and real-data validation. *Transportation Research Part B: Methodological*, 106:52–75, 2017.

Robert Krajewski, Julian Bock, Laurent Kloeker, and Lutz Eckstein. The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2118–2125, 2018. doi: 10.1109/ITSC.2018.8569552.

Michael James Lighthill and Gerald Beresford Whitham. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229(1178):317–345, 1955.

Saeed Mohammadian, Zuduo Zheng, Md Mazharul Haque, and Ashish Bhaskar. Performance of continuum models for realworld traffic flows: Comprehensive benchmarking. *Transportation Research Part B: Methodological*, 147:132–167, 2021.

John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.

Emanuel Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.

Xiaobo Qu, Shuaian Wang, and Jin Zhang. On the fundamental diagram for freeway traffic: A novel calibration approach for single-regime models. *Transportation Research Part B: Methodological*, 73:91–102, 2015.

Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Paul I Richards. Shock waves on the highway. *Operations research*, 4(1):42–51, 1956.

Daniel Rodriguez-Rey, Marc Guevara, Ma Paz Linares, Josep Casanovas, Juan Salmerón, Albert Soret, Oriol Jorba, Carles Tena, and Carlos Pérez García-Pando. A coupled macroscopic traffic and pollutant emission modelling system for barcelona. *Transportation Research Part D: Transport and Environment*, 92:102725, 2021.

Rongye Shi, Zhaobin Mo, Kuang Huang, Xuan Di, and Qiang Du. A physics-informed deep learning paradigm for traffic state and fundamental diagram estimation. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):11688–11698, 2021.

Michael J Smith, Takamasa Iryo, Richard Mounce, Marco Rinaldi, and Francesco Viti. Traffic control which maximises network throughput: Some simple examples. *Transportation Research Part C: Emerging Technologies*, 107:211–228, 2019.

Raphael E Stern, Shumo Cui, Maria Laura Delle Monache, Rahul Bhadani, Matt Bunting, Miles Churchill, Nathaniel Hamilton, Hannah Pohlmann, Fangyu Wu, Benedetto Piccoli, et al. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. *Transportation Research Part C: Emerging Technologies*, 89:205–221, 2018.

Yuki Sugiyama, Minoru Fukui, Macoto Kikuchi, Katsuya Hasebe, Akihiro Nakayama, Katsuhiro Nishinari, Shin-ichi Tadaki, and Satoshi Yukawa. Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam. *New journal of physics*, 10(3): 033001, 2008.

Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2446–2454, 2020.

Robin T Underwood. Speed, volume, and density relationship: Quality and theory of traffic flow. *Yale Bureau of Highway Traffic*, pages 141–188, 1961.

Huan Yu and Miroslav Krstic. *Traffic Congestion Control by PDE Backstepping*. Springer, 2022.

Chenguang Zhao and Huan Yu. Observer-informed deep learning for traffic state estimation with boundary sensing. *IEEE Transactions on Intelligent Transportation Systems*, 2023.

Shi-Teng Zheng, Rui Jiang, Jun-Fang Tian, HM Zhang, Zhen-Hua Li, Lan-Da Gao, and Bin Jia. Experimental study on properties of lightly congested flow. *Transportation Research Part B: Methodological*, 149:1–19, 2021.