Supplementary Material

Table1 Fine-grained classification based on different vehicle types

Type	Sized	Classification Criteria
Bus (passenger vehicle)		
k_{B1}	Small-sized bus	Number of seats & length of vehicle
k_{B2}	Mid-sized bus	
k_{B3}	Large-sized bus	
Truck (freight vehicle)		
k_{T1}	Small-sized Truck	Total weight & length of vehicle
k_{T2}	Mid-sized Truck	
k_{T3}	Large-sized Truck	

Table2 Results of diversion coefficients for fine-grained vehicles at each ramp

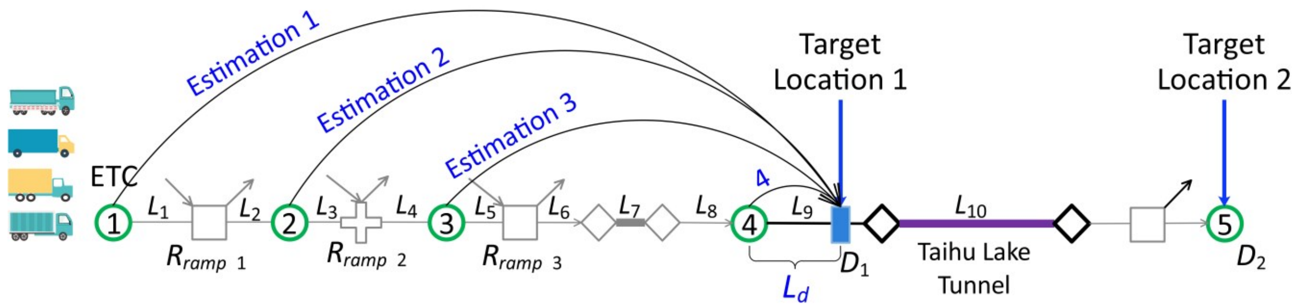
Type	Ramp 1			Ramp 2			Ramp 3		
	$R_{off-ramp1}$	$R_{on-ramp1}$	R_{ramp1}	$R_{off-ramp2}$	$R_{on-ramp2}$	R_{ramp2}	$R_{off-ramp3}$	$R_{on-ramp3}$	R_{ramp3}
Total	0.074	0.288	1.301	0.438	0.405	0.945	0.205	0.095	0.879
K_{B1}	0.071	0.305	1.337	0.457	0.414	0.926	0.208	0.084	0.865
K_{B2}	0.077	0.271	1.266	0.350	0.396	1.075	0.139	0.079	0.935
K_{B3}	0.051	0.313	1.381	0.463	0.364	0.844	0.295	0.081	0.767
K_{T1}	0.093	0.259	1.224	0.357	0.368	1.017	0.173	0.154	0.977
K_{T2}	0.083	0.180	1.118	0.321	0.361	1.063	0.171	0.130	0.952
K_{T3}	0.095	0.391	1.485	0.462	0.389	0.879	0.200	0.122	0.911

In the **"Estimation of Fine-Grained Traffic Flow"** section of the experiment, given the challenge of traffic state estimation on highway segments with sparse sensor coverage (i.e., sections without detectors), direct access to ground truth data for validation is not possible. Therefore, in addition to the validation methods designed in the manuscript, we introduce the **Pearson Correlation Coefficient (PCC)** as an additional evaluation metric to further assess the model's performance.

$$PCCs = \frac{\mathrm{cov}(y^1, y^2)}{\delta_{y^1}\delta_{y^2}} = \frac{\sum_{i=1}^{n} (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}\sqrt{\sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}}_i)^2}}$$

In this analysis, we designate the target location as D_1 (Fig. 11 in this paper), positioned 500 meters upstream of the Taihu Lake Tunnel, with a segment length of $L_d = L_9 - 0.5 = 1.13\,km$. To examine the impact of ramp diversion coefficients on estimation accuracy, we design four estimation schemes, *Estimation 1 – Estimation 4*, each utilizing upstream traffic data from

$Seg_{ETC1} - Seg_{ETC4}$ to estimate traffic flow at the target location. Since direct traffic flow measurements at $\boldsymbol{D}_1$ are unavailable, PCCs are employed to quantify the similarity of estimation results across different schemes. **Fig. S1** presents the validation results based on PCCs for each estimation scheme, leading to the following key observations:



a) **Impact of traffic demand duration**: Estimation accuracy varies significantly under different traffic demand conditions, with generally higher PCC values observed in Condition 1 compared to Condition 2. In Condition 1, the shorter demand period reduces temporal discrepancies, leading to a more accurate representation of spatial traffic flow propagation. Conversely, in Condition 2, the extended demand duration introduces greater variability and uncertainty in traffic evolution, highlighting the need to account for the relationship between T_{demand} and $T_{traveling}$ in traffic state estimation—a factor often overlooked in short-term prediction models.

b) **Effect of vehicle classification**: PCC values vary significantly across vehicle categories. Small-sized buses (K_{B1}) consistently exhibit higher correlation coefficients, indicating more stable traffic flow patterns. In contrast, larger vehicles such as K_{T2} are subject to greater fluctuations due to factors like weight restrictions, speed limits, and lane-changing behaviors. These nonlinear dynamics make their traffic patterns more challenging to predict, reinforcing the importance of fine-grained vehicle classification in traffic estimation models.

c) **Spatial correlation and gantry distance**: When two ETC gantries are closer together, the corresponding estimation results show higher correlation coefficients, aligning with the common practice of using distance as a proxy for spatial correlation in traffic flow prediction tasks. However, as the distance between gantries increases, the estimation uncertainty grows due to factors such as traffic dispersion, interactions with ramps, and the influence of special road structures.

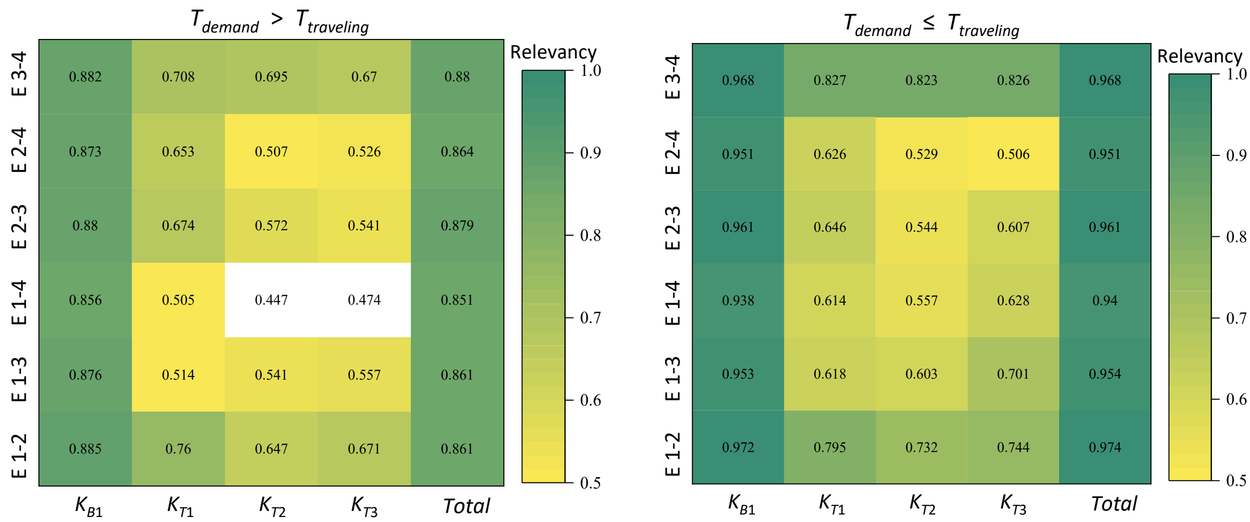


Fig. S1. Validation results based on Pearson Correlation Coefficient (PCC) for each estimation scheme: Validation results for **Condition 1** (July 29, 2022, $T_{demand} = 10min \leq T_{traveling}$); (b) Validation results for **Condition 2** (July 29, 2022, $T_{demand} = 30min > T_{traveling}$).