DSSNet: Deep Spectral Structure Profiling Network for Traffic Flow Prediction

Tao Wang[1,2] **Zhenye Yang**[1,2] **Hua Lu**[3] **Huan Li**[4] **Senzhang Wang**[5] **Jinpeng Chen**[1,2*]

[1]School of Computer Science (National Pilot Software Engineering School),
Beijing University of Posts and Telecommunications, Beijing, China
[2]Key Laboratory of Trustworthy Distributed Computing and Service (BUPT), Ministry of Education
[3]Department of Computer Science, Aalborg University, Denmark
[4]Zhejiang University, Hangzhou, China [5]Central South University, Changsha, China
{wangtao005,yangzhenye,jpchen}@bupt.edu.cn
luhua@cs.aau.dk, lihuan.cs@zju.edu.cn, szwang@csu.edu.cn

Abstract

Traffic flow prediction is a fundamental task in intelligent transportation systems. Existing methods primarily rely on temporal dependency modeling and thus face two critical limitations: struggling to effectively capture diverse periodicity; overlooking the relationship between spectral structures in the traffic flow and spatiotemporal traffic behaviors. To address these limitations, we propose DSSNet (Deep Spectral Structure Profiling Network), an adaptive spectral profiling framework. DSSNet follows the paradigm of "spectral structure profiling" and then "dual-domain modeling". Initially, a Spectral Structure Learning module captures the dominant frequency bands of the spectrum to enhance regional representations while simultaneously serving as discriminative spatiotemporal identity between regions. Based on these identities, we construct a global Resonance Graph to uncover cross-region dependencies and introduce a Dual-Domain Knowledge Integration module to aggregate structural knowledge across both time and frequency domains. Extensive experiments on five benchmark datasets demonstrate that DSSNet outperforms all evaluated baselines on four datasets, highlighting the potential of frequency-domain modeling for understanding spatiotemporal traffic behaviors.

1 Introduction

Traffic flow prediction is a key task in intelligent transportation systems Cirstea et al. [2021], Ji et al. [2022], Dai et al. [2023]. Accurate traffic prediction improves urban efficiency while reducing carbon emissions and energy consumption Zhang et al. [2020]. Existing approaches evolved from temporal-domain models (e.g., RNNs, TCNs) Chung et al. [2014], Hochreiter and Schmidhuber [1997] to spatiotemporal graph neural networks Jiang et al. [2019], Li et al. [2018], Song et al. [2020], achieving promising performance. Frequency-domain methods emerged as a promising alternative by separating distinct frequency components Zhang et al. [2024b, 2022]. However, most treat these components uniformly, ignoring the structured knowledge in different frequency bands.

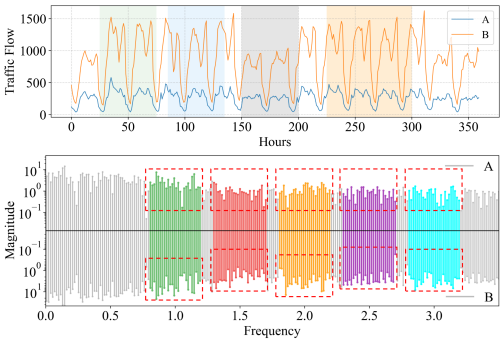
However, these methods face two major limitations. Firstly, they **struggle to effectively characterize diverse periodicity**, primarily stemming from the inherent spatiotemporal heterogeneity of traffic system Liang et al. [2022], Jiang et al. [2023], Yao et al. [2019a]. Specifically, existing approaches typically treat spatial regions uniformly or independently, failing to capture region-specific spatiotemporal correlations. In addition, existing methods **overlook the relationship between spectral structures and spatiotemporal behavior patterns**.

*Corresponding author

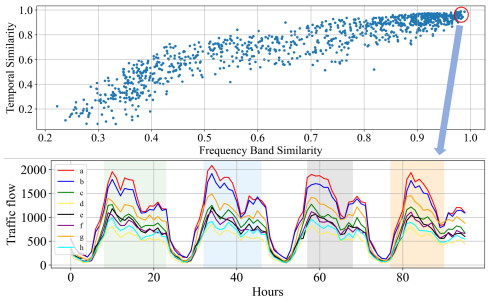
40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Although some studies have attempted frequency modeling Yi et al. [2024], Wu et al. [2025], they predominantly treat spectral components uniformly and operate in strict spatial isolation. They fail to treat regional spectral energy distributions as discriminative spatiotemporal identities. Consequently, current models remain constrained by static physical topologies, missing the critical insight to leverage spectral homophily to capture dynamic, cross-region resonance that defies spatial proximity.

To empirically validate the premise of spectral homophily—specifically, whether spectral structure similarity translates to synchronized temporal dynamics—we conduct a preliminary analysis on the T-Drive dataset Pan et al. [2019]. As illustrated in Fig. 1a, regions of different types (e.g. A.residential areas vs B.transportation hubs) exhibit heterogeneous spatiotemporal patterns, which are mirrored in the frequency domain as differences in energy (Magnitude) distribution of several *dominant frequency bands*, each corresponding to a key periodicity and its surrounding frequencies. We compute two metrics between different regions: (1) the similarity of band-wise energy distributions, and (2) the Pearson correlation of traffic flow, then visualize their relationship using a scatter plot. As shown in Fig. 1b, higher spectral similarity consistently corresponds to stronger temporal synchronization. Representative cases with the highest spectral similarity are illustrated at the bottom of Fig. 1b. More details can be found in Appendix A. These results suggest that spectral structure provides a discriminative descriptor, which can be viewed as spatiotemporal profile.



(a) Traffic Flow and spectrogram. To illustrate the spectral structure, we manually select five dominant frequency bands.



(b) Similarity correlation analysis with case studies.

Figure 1: Correlation between spectral and traffic dynamics.

In this paper, we propose DSSNet, a novel framework that explicitly capture diverse periodicity via adaptive spectral analysis. Specifically, we design a Spectral Structure Learning module to adaptively identify dominant frequency bands. The learned spectral structure serves a dual utility: it is employed to characterize intrinsic spatiotemporal patterns for enhancing region-level representations and to construct a global Resonance Graph for cross-region relationship identification. Based on these structures, we introduce a Dual-Domain Knowledge Integration (DDKI) module to perform structured knowledge aggregation jointly in both the temporal and frequency domains. Extensive experiments on multiple benchmark datasets demonstrate that DSSNet consistently achieves competitive performance compared with baseline. The main contributions are as follows:

- We propose an adaptive dominant frequency band modeling to effectively capture the relationship between frequency bands and spatiotemporal behavior patterns.
- We propose DSSNet, a deep spectral structure profiling network that effectively captures diverse periodicity by adaptive dominant frequency band modeling.
- Extensive experiments demonstrate that DSSNet achieves competitive performance on benchmark datasets, while also exhibiting strong scalability and interpretability.

2 Related Works

2.1 Temporal-Domain Methods

Traditional statistical methods (e.g., ARIMA, SVR Wu et al. [2004]) and early deep learning models (CNNs, RNNs) Hochreiter and Schmidhuber [1997], Chung et al. [2014] offered limited support for

spatial dependencies. The emergence of Graph Neural Networks (GNNs), such as STGCN Yu et al. [2018] and DCRNN Li et al. [2018], advanced the field by unifying graph convolutions with sequential modeling. Subsequent works have further refined spatial representations via adaptive graphs (Graph WaveNet Wu et al. [2019], AGCRN Bai et al. [2020]) or addressed training stability issues like over-smoothing (STGODE Fang et al. [2021]). More recent approaches leverage attention mechanisms (GMAN Zheng et al. [2020], ASTGCN Guo et al. [2019]) and signal decomposition (STWave Fang et al. [2023]) to capture intricate dynamics. Despite these advances, effectively modeling long-range periodicity and region-specific spatiotemporal heterogeneity remains a significant challenge.

2.2 Frequency-Domain Method

Frequency-domain approaches leverage spectral analysis to capture long-range dependencies and periodic patterns. Early methods, such as TDformer Zhang et al. [2022] and TimesNet Wu et al. [2023], utilize Fourier transforms or signal decomposition to model global temporal trends and periods. While effective, these methods largely overlook spatial correlations. Recent advancements focus on adaptivity: FRNet Zhang et al. [2024b] and FAN Ye et al. [2024] capture evolving periodicity through dynamic frequency selection, while FilterNet Yi et al. [2024] and FilterTS Wang et al. [2025] introduce learnable filters to handle variable-dependent patterns. Other works (e.g., TSLANet Eldele et al. [2024], FreDF Zhang et al. [2024a]) further enhance interpretability via adaptive weighting mechanisms. However, existing spectral approaches typically treat spatial regions uniformly or independently, failing to capture region-specific spectral heterogeneity and joint spatiotemporal correlations. To bridge this gap, our proposed DSSNet integrates frequency-domain modeling with spatial structures to explicitly capture diverse and heterogeneous periodic behaviors.

3 Problem Definition

Spatio-Temporal Data. The $\mathbf{X}_t \in \mathbb{R}^{C\times H\times W}$ denote the traffic flow at time step t, where C is the number of channels. The T time steps data is represented as $\mathcal{X} \in \mathbb{R}^{T\times C\times H\times W}$.

Traffic Flow Prediction. Given the history traffic data in length T_{in}, our goal is to predict the traffic flow of next T_{out} steps. At the time step t, a mapping function $\mathcal{F}_\theta$ with parameters θ is learning as:

$$\hat{\mathbf{X}}_{t+1:t+Tout} = \mathcal{F}_\theta(\mathbf{X}_{t-T_{in}+1:t}) \tag{1}$$

4 METHODOLOGY

As illustrated in Fig.2, the overall architecture of DSSNet consists of three key modules: (1) Data Preprocessing: this module enhances frequency resolution by introducing a memory bank and performs trend–seasonality decomposition on the raw traffic flow. (2) Spectral Structure Learning: this module introduces a Learnable Band Filter Bank module to adaptively identify and encode dominant frequency bands. The learned frequency bands serve a dual purpose: enhancing region-specific intrinsic representations and constructing a dynamic graph based on spectral similarity, capturing potential resonance relationships among regions. (3) Dual-Domain Knowledge Integration: based on the learned structured knowledge, this module aggregates knowledge in both the temporal and frequency domains, ultimately enabling robust spatiotemporal prediction.

4.1 Data Preprocessing

To enhance the frequency resolution of spectral analysis, we introduce a Memory Bank to store historical traffic flow in a First-In-First-Out manner. We concatenate current input X with the X_{his} from the Memory Bank to form an augmented sequence $\tilde{X} \in \mathbb{R}^{(T_{in}+T_{mem})\times C\times H\times W}$, which is then normalized with RevIN Kim et al. [2022].

We decompose $\tilde{X}$ into trend and seasonal components to separate non-stationary trends from periodic patterns. Following by Zhang et al. [2022], we employ Multi-scale 1D Average Pooling as follows:

$$X_{tre} = \sum_{k=1}^{K} w_k(x) \cdot \text{AvgPool}_k(\tilde{X}) \tag{2}$$

$$X_{sea} = \tilde{X} - X_{tre} \tag{3}$$

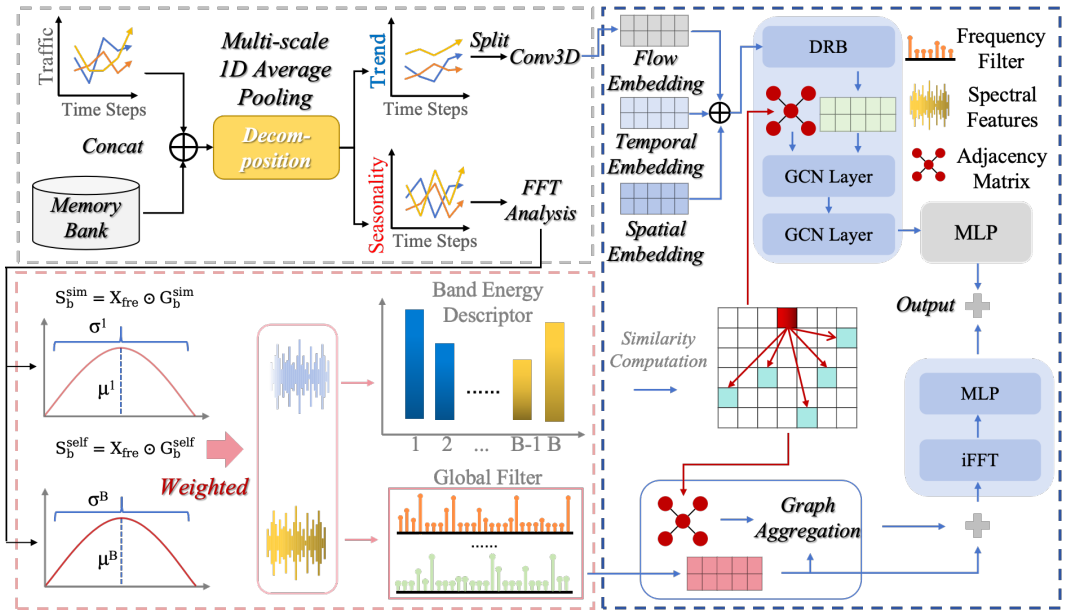


Figure 2: The overall architecture of DSSNet, consisting of data preprocessing, spectral structure learning, dual domain knowledge integration.

where w_i represent learnable weights for different kernel sizes, K is the number of kernels.

4.2 Spectral Structure Learning

This module aims to capture the spectral structure of traffic flow. We introduce Learnable Band Filter Bank to adaptively identify and encode dominant frequency bands. The identified frequency bands serve a dual purpose: capture intrinsic spatiotemporal patterns within each region to enhance region-level representations, and act as spectral semantic descriptors for constructing a Resonance Graph of cross-region relationships.

Learnable Band Filter Bank (LBFB): we construct two banks of B learnable Gaussian filters to identify dominant frequency bands. The $G_b^{self}(f)$ are used for region specific spectral feature extraction, and the $G_b^{sim}(f)$ are designed for spectral similarity computation across regions.

Each filter G_b is parameterized by a trainable center frequency μ_b and bandwidth σ_b, representing the mean and standard deviation of the Gaussian kernel respectively. This configuration allows the model to adaptively "scan" the spectrum, where μ_b defines the peak location of the frequency band and σ_b determines its coverage, effectively locking onto the most informative periodic patterns:

$$G_b(f) = \exp\left(-\frac{(f-\mu_b)^2}{2\sigma_b^2}\right) \tag{4}$$

both μ_b and σ_b are trainable during model optimization. Given seasonal component X_{sea}, we transform it into the frequency domain via fast Fourier transform (FFT), denoted as $X_{fre} \in \mathbb{C}^{F\times C\times H\times W}$, where $F = \lfloor \frac{T}{2} \rfloor + 1$. We apply $G_b(f)$ to the X_{fre} via element-wise multiplication.

$$S_b^{self} = X_{fre} \cdot G_b^{self} \quad b = 1, \ldots, B \tag{5}$$

$$S_b^{sim} = X_{fre} \cdot G_b^{sim} \quad b = 1, \ldots, B \tag{6}$$

where $S_b^{self}, S_b^{sim} \in \mathbb{R}^{F\times B\times C\times H\times W}$.

The first utility focuses on the region specific information. While S_b^{self} isolates specific dominant frequency bands, it contains redundant frequency information. To obtain a compact and robust region-level representation, we introduce a Global Frequency Filter $W_{global} \in \mathbb{R}^{B\times F}$ that is shared across all regions. We apply this global filter to the extracted sub-spectrograms to further suppress

noise and highlight universally significant frequency features:

$$x_{\text{filtered}} = S_b^{self} \odot W_{global} \tag{7}$$

the resulting $x_{filtered}$ serves as the refined spectral feature.

The second utility focuses on cross-region relationships. Regions with similar dominant bands energy distributions exhibit dependency. We quantify this by computing the **Band Energy Descriptor** (BED) $Z_n \in \mathbb{R}^B$ for region n:

$$Z_{n,b} = \frac{\sum_f |S_b^{sim}(n)(f)|^2}{\sum_{b'=1}^{B} \sum_f |S_{b'}^{sim}(n)(f)|^2}, b = 1, 2...B \tag{8}$$

Z_n captures the energy signature of the node. We then construct the Resonance Graph adjacency matrix A by measuring the similarity between Z_n using a Gaussian kernel:

$$A_{ij} = \exp\left(-\frac{\|Z_i - Z_j\|_2^2}{\tau}\right) \tag{9}$$

where τ controls the decay of affinity with spectral distance. This graph A uncovers hidden topological connections.

4.3 Dual Domain Knowledge Integration (DDKI)

The DDKI module aims to enable structure-guided knowledge aggregation in both temporal and frequency domains.

Temporal Branch: it encodes the evolving trend X_{tre} which excludes the historical traffic flow from the Memory Bank by combining intrinsic sequential patterns with extrinsic structural guidance.

We first capture comprehensive context by aggregating flow, temporal, and spatial information. The trend input is mapped to a high-dimensional embedding state E:

$$E = \text{Concat}(\text{Conv3d}(X_{tre}), E_{time}, E_{space}) \tag{10}$$

where E_{time} encodes day-of-week and hour-of-day to distinguish daily and weekly patterns, and E_{space} encodes the identity and location of each grid, $E \in \mathbb{R}^{T_{in} \times C \times H \times W \times D}$.

To efficiently capture multi-scale dependencies, we employ an Deep Residual Block (DRB) for deep representation learning, computing as:

$$Z_0 = W_2 \cdot \sigma(W_1 E + b_1) + b_2 \tag{11}$$

$$Z_{tem} = (W_4 \cdot \sigma(W_3 Z_0 + b_3) + b_4) + Z_0 \tag{12}$$

where W and b are learnable parameters.

Structure Aware GCN: Finally, we utilize the Graph A to guide trend propagation among pattern similar regions:

$$A_{norm} = \tilde{D}^{-\frac{1}{2}}(A + I)\tilde{D}^{-\frac{1}{2}} \tag{13}$$

$$Z_{tem}^{out} = \text{GCN}(Z_{tem}, A_{norm}) + Z_{tem} \tag{14}$$

This step explicitly enable the model to capture synchronized trend evolution regions.

Frequency Path: we perform Bandwise Graph Aggregation to strengthen the collaborative learning of cross region patterns and capture diverse periodicity:

$$x_{agg} = \text{Softmax}(A) \cdot x_{filtered} \tag{15}$$

To preserve node-specific characteristics while benefiting from global resonance, we apply a residual fusion with a learnable factor α:

$$Z_{fre}^{out} = (1 - \alpha) \cdot x_{filtered} + \alpha \cdot x_{agg} \tag{16}$$

We synthesize the outputs to generate the final forecast. The temporal output and the frequency output (via iFFT) are summed to obtain the final results:

$$\hat{Y} = \text{MLP}(Z_{tem}^{out}) + \text{MLP}(\text{iFFT}(Z_{fre}^{out})) \tag{17}$$

where $\hat{Y} \in \mathbb{R}^{T \times C \times H \times W}$, $iFFT$ represents the inverse FFT. We adopt MSE as the final loss:

$$Loss = MSE(Y, \hat{Y}) \tag{18}$$

5 Experiments

5.1 Experimental Setup

Datasets: We conduct experiments on five traffic datasets: **T-Drive**, **TaxiBJ**, **TaxiNYC-1**, **TaxiNYC-2** and **PEMS04**, covering both urban grids and road networks. More information are summarized in Appendix B.

Baselines: We compare DSSNet with eleven established baselines: HA, SVR Wu et al. [2004], LSTM Hochreiter and Schmidhuber [1997]; temporal domain methods: STGCN Yu et al. [2018], GWNet Wu et al. [2019], AGCRN Bai et al. [2020], STID Shao et al. [2022], STDN Cao et al. [2025]; frequency domain methods: FilterTS Wang et al. [2025], FreTS Yi et al. [2023], FilterNet Yi et al. [2024]. The more information of baseline is summarized in Appendix C. To ensure fair comparison, we equip frequency-domain methods with memory bank. In addition, We analyze the effect of Memory Bank length and compare methods under matched historical contexts in Section 5.4.

Experiment Settings: We perform multi-step prediction, where the model forecasts the next 6 steps based on 6 historical steps. We adopt MAE and RMSE as standard metrics. All experiments are conducted with NVIDIA RTX 3090 GPUs. We perform experiments five times with different random seeds and report average performance. We use the AdamW optimizer with a learning rate of 0.002.

5.2 Performance Comparison

Table 1 reports the quantitative comparison between DSSNet and eleven baseline methods. Overall, DSSNet establishes a new state-of-the-art benchmark, outperforming all baselines on four out of five datasets with substantial margins. Specifically, on the TaxiNYC-1 and TaxiNYC-2 datasets, DSSNet demonstrates exceptional dominance, achieving relative improvements of 21.11%, 23.80% in MAE and 25.63%, 27.36% in RMSE over the second-best methods. Similarly, on the T-Drive and PEMS-04 datasets, which are characterized by complex spatiotemporal heterogeneity and large-scale road networks, our model surpasses the runner-up by approximately 9.09% and 4.34% in MAE and 8.00% and 5.18% in RMSE. The results validate the effectiveness of the proposed Spectral Structure Profiling. While pure frequency-domain baselines (e.g., FilterNet) show competitive performance on TaxiBJ, they exhibit significant degradation on the other four datasets like T-Drive. In contrast, DSSNet's adaptive band filtering successfully captures diverse region-specific periodicities, ensuring consistent superiority across varying traffic scenarios.

Table 1: The performance of all selected methods on five traffic flow datasets. From top to bottom: traditional methods, temporal-domain methods, frequency-domain methods, and the present method. The best result and the second-best result are marked with **Bold** and Underline, respectively.

Methods	TaxiNYC-1		TaxiNYC-2		T-Drive		TaxiBJ		PEMS-04	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
HA	18.58	57.08	15.78	52.94	73.49	146.67	39.55	72.56	36.74	57.47
SVR	12.21	23.25	11.81	23.21	52.64	86.25	27.46	49.58	29.57	43.28
LSTM	11.24	35.82	9.65	33.11	42.74	88.12	26.24	48.11	27.43	41.44
STGCN	7.70	22.45	6.90	21.94	32.40	62.29	19.82	34.64	23.92	37.04
GWNet	8.90	27.99	7.45	23.77	32.49	68.53	26.60	51.46	24.74	38.17
AGCRN	7.20	21.22	6.80	20.95	25.05	52.64	13.90	24.54	22.39	34.84
STID	7.76	24.66	6.82	23.74	19.52	39.95	13.46	23.52	21.16	34.26
STDN	9.07	27.13	7.94	25.46	23.99	45.38	20.10	34.69	22.67	35.92
FilterTS	7.65	22.27	6.51	19.99	20.64	42.27	13.58	23.55	22.02	33.77
FreTS	8.61	25.35	7.74	24.63	21.43	42.75	13.11	22.75	19.79	32.07
FilterNet(Pai)	8.46	25.17	7.89	25.05	19.25	39.11	13.02	22.81	20.63	32.91
FilterNet(Tex)	8.37	24.82	7.89	25.21	21.57	43.36	**12.63**	**22.08**	19.78	32.01
DSSNet	**5.68**	**15.78**	**4.96**	**14.52**	**17.50**	**35.98**	12.93	22.13	**18.92**	**30.35**
Improve(%)	21.11%	25.63%	23.80%	27.36%	9.09%	8.00%	-2.37%	-0.22%	4.34%	5.18%

5.3 Ablation Study

Ablation of LBFB: To verify the effectiveness of the Learnable Band Filter Bank (LBFB), we compare it against two static strategies: **Uniform** (treating the full spectrum as equally important) and **Equal** (using fixed, equidistant frequency bands). For each strategy, we evaluate three configurations: (1) *w/o Identity* (Iden): ablating the filtering for region-level representation (Eq.7); (2) *w/o Resonance* (Reso): ablating the filtering for spectral similarity computation (Eq.8); (3) *w/o LBFB* (All): removing the module entirely. As shown in Table 2, the **Uniform** strategy causes the largest performance drop, confirming the necessity of dominant band filtering to suppress noise. The **Equal** strategy outperforms Uniform but consistently lags behind DSSNet, demonstrating the critical role of adaptive band selection. Furthermore, removing either the identity extraction (Self) or resonance construction (Sim) leads to performance degradation, validating the dual utility of the learned spectral structure.

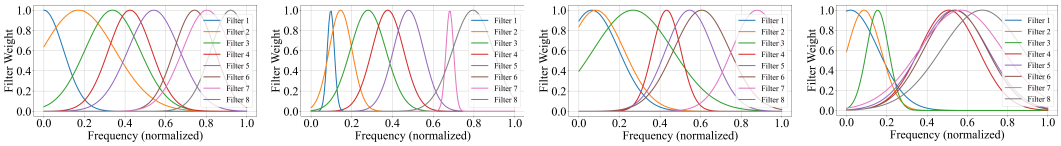
Ablation of Resonance Graph: To evaluate the validity of the Resonance Graph constructed via Eq.9, we replace it with two variants: (1) **w/ Random** (Rand): a randomly generated adjacency matrix; (2) **w/ Physical** (Phys): a spatial connectivity matrix based on physical proximity. Table 2 shows that the Random graph leads to substantial deterioration, confirming the necessity of a meaningful topology. Crucially, the **Physical** graph outperforms Random but remains inferior to our Resonance Graph. This implies that region-level traffic dependencies are more accurately characterized by "spectral functional identity" than by physical distance, supporting our hypothesis on spectral homophily.

Table 2: Results of ablation studies. The left side lists the ablation strategies and experimental configurations.

Experiments		TaxiNYC-1		T-Drive	
		MAE	RMSE	MAE	RMSE
Uniform	Iden	6.28	17.27	18.29	39.26
	Reso	5.94	16.33	17.87	36.62
	All	6.31	17.16	18.48	39.87
Equal	Iden	6.06	16.24	17.69	36.80
	Reso	5.92	15.97	17.73	36.99
	All	6.13	16.63	17.96	37.53
RG	Rand	6.12	16.94	18.13	37.24
	Phys	5.87	16.40	17.95	36.67
DDKI	Fre	6.11	16.88	17.87	36.82
	Tem	5.89	16.24	17.64	36.41
	All	6.19	17.05	17.96	37.04
DSSNet	-	**5.68**	**15.78**	**17.50**	**35.98**

Ablation of DDKI: To investigate the contribution of the Dual-Domain Knowledge Integration module, we conduct ablations by: (1) removing the Bandwise Graph Aggregation in the frequency branch (**w/o Fre**); (2) removing the Structure-Aware GCN in the temporal branch (**w/o Tem**); (3) removing the interaction entirely (**w/o All**). Results in Table 2 indicate that ablating either branch degrades performance, proving that the intrinsic spectral patterns and temporal trends are complementary. The full DDKI achieves best results, confirming the effectiveness of integrating dual-domain knowledge.

5.4 Model Analysis

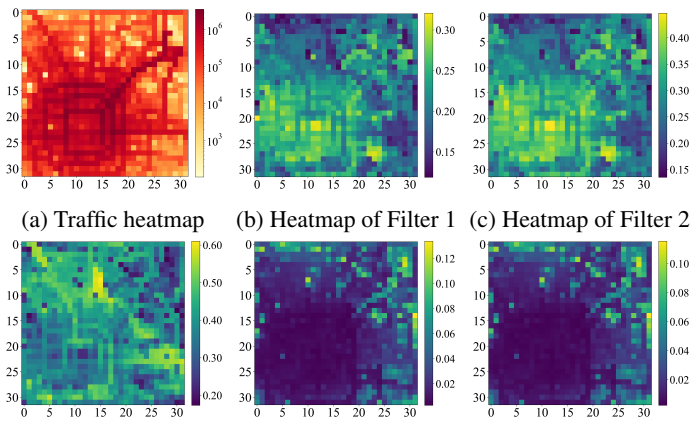


(a) TaxiNYC1-self (b) TaxiNYC1-sim (c) T-Drive-self (d) T-Drive-sim

Figure 3: The visualization of LBFB in TaxiNYC-1 and T-Drive.

Analysis of Band Filter Bank. To validate the effectiveness of the proposed Spectral Structure Profiling, we visualize the learned filters and their corresponding spatial responses in **T-Drive** dataset. We first examine the Learnable Band Filter Bank (LBFB) in Figure 3. The learned filters automatically distribute across the spectrum, ensuring multi-scale coverage. Crucially, the filter configurations for identity extraction (G^{self}) and similarity computation (G^{sim}) exhibit distinct distributions (e.g., comparing Figure 3c and 3d). This divergence implies that the spectral features required for characterizing a region's intrinsic pattern differ from those optimal for establishing cross-region resonance, validating our design of dual-view spectral filtering.

Semantics of Learned Frequency Bands. We further investigate the semantics of the learned bands (take the filters in Figure 3d as example) by visualizing the spatial attention of different learned frequency bands across diverse regions in Figure 4. By using the Ground Truth Traffic Heatmap (Figure 4a) as a reference, we observe a distinct correlation between frequency focus and traffic intensity. Specifically, Filters 1, 2, and 3, which concentrate on the low-frequency band, generate spatial heatmaps (Figure 4b-4d) that significantly highlight high-traffic regions such as arterial roads and central hubs. Conversely, Filters 4 and 5 specialize in high-frequency bands and exhibit strong attention towards low-traffic regions (Figure 4e-4f). This observation suggests a hierarchical functional mapping: low-frequency bands primarily attend to high-volume areas governed by robust periodicity; mid-frequency bands tend to capture transition zones with moderate flows; while high-frequency bands (e.g., Filter 4 in Figure 4e) focus on peripheral areas or specific local spots characterized by irregular fluctuations.

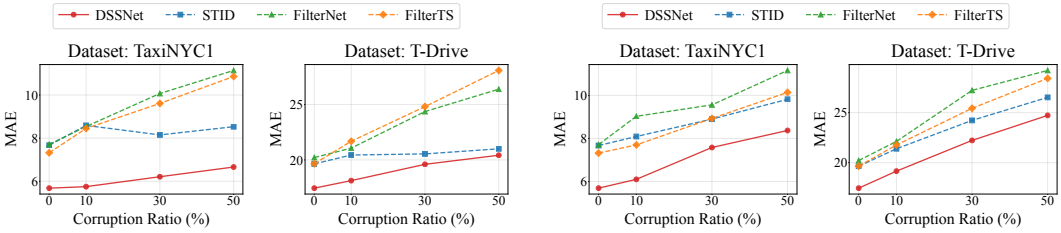


(a) Traffic heatmap (b) Heatmap of Filter 1 (c) Heatmap of Filter 2

(d) Heatmap of Filter 3 (e) Heatmap of Filter 4 (f) Heatmap of Filter 5

Figure 4: Heatmap across all regions for dominant spectral bands.

Analysis of Long-range Resonance. To verify the global modeling capability, we visualize a random Anchor (Curve a) and its top-7 resonant neighbors in Appendix D.1. As shown in Figure 11a, the learned resonant regions are widely dispersed across the city rather than being confined to the local neighborhood. This confirms that DSSNet effectively captures long-range dependencies beyond adjacent areas.Furthermore, in Figure 11b, we visualize the traffic flow of these spectrally similar regions. The consistent patterns among these spatially distant regions validate that our model successfully identifies meaningful connections.



(a) Robustness to Sensor Noise (b) Robustness to Missing Values

Figure 5: Robustness evaluation under realistic disruptions on TaxiNYC-1 and T-Drive datasets.

Robustness to Realistic Disruptions. In real-world Intelligent Transportation Systems (ITS), sensor data is frequently corrupted by hardware interference or transmission losses. Thus, we evaluate DSSNet under varying intensities of Gaussian noise (up to 50% std) and random missing values (affecting up to 50% of spatial nodes). As summarized in Figure 5 (detailed Appendix D.2), DSSNet consistently maintains the lowest performance degradation compared to temporal-domain baselines like STID. This resilience is fundamentally attributed to our *Spectral Structure Profiling* paradigm: the Learnable Band Filter Bank (LBFB) acts as an inherent frequency-domain denoiser that suppresses high-frequency sensor jitter, while the Band Energy Descriptor (BED) provides a stable, holistic regional identity that remains robust even when local temporal observations are fragmented.

Robustness against Non-stationary Surges. Real-world traffic systems are inherently non-stationary, often characterized by sudden bursts caused by accidents or extreme weather. To evaluate the resilience of DSSNet against such spectral drifts, we conduct a perturbation experiment on the TaxiNYC-1 and T-Drive datasets. Specifically, we simulate extreme traffic surges by injecting $3\times$

Table 3: Performance under matched historical buffer lengths. Each method receives $L + 6$ historical observations and predicts the next six steps. The best results are highlighted in bold.

Methods	Dataset: TaxiNYC-1											
	$L = 0$		$L = 36$		$L = 72$		$L = 168$		$L = 336$		$L = 504$	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
STID	**7.76**	24.66	7.86	24.91	7.81	24.13	8.14	24.94	7.87	24.54	8.30	24.69
FilterNet	12.78	37.83	9.33	27.62	8.76	25.93	8.40	24.86	8.37	24.82	8.21	24.30
FilterTS	11.98	34.98	9.25	27.01	9.09	26.54	8.90	25.99	7.65	22.27	8.64	25.23
DSSNet	8.13	**24.34**	**7.62**	**22.86**	**6.81**	**20.64**	**6.04**	**17.56**	**5.68**	**15.78**	**5.84**	**16.33**
Methods	Dataset: TaxiNYC-2											
	$L = 0$		$L = 36$		$L = 72$		$L = 168$		$L = 336$		$L = 504$	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
STID	6.82	23.74	6.88	23.93	6.79	23.84	6.89	23.81	6.82	23.74	6.91	23.68
FilterNet	10.80	38.17	7.80	26.79	7.04	24.34	8.35	27.72	7.89	25.21	7.31	22.96
FilterTS	34.74	104.50	8.63	29.80	8.28	29.20	8.82	29.05	6.51	19.99	9.25	31.09
DSSNet	**6.62**	**22.26**	**6.68**	**21.74**	**5.88**	**19.23**	**5.50**	**16.69**	**4.96**	**14.52**	**5.24**	**14.90**

intensity spikes into a subset of nodes (up to 50%) for 3% of the time steps. As shown in Appendix D.3, while all models experience performance degradation compared to clean data, DSSNet maintains the lowest error metrics. This robustness stems from the adaptive nature of our Learnable Band Filter Bank (LBFB), which can dynamically adjust frequency focus.

Impact of Memory Bank. To distinguish the contribution of spectral structure profiling from the benefit of longer historical context, we compare DSSNet with STID, FilterNet, and FilterTS under matched history lengths on TaxiNYC-1 and TaxiNYC-2. We vary the Memory Bank length $L \in \{0, 36, 72, 168, 336, 504\}$, where L denotes the additional historical observations preceding the most recent six input steps. Thus, each method receives the same $L+6$ causally available observations and predicts the next six steps, with identical data splits and preprocessing. As shown in Table 3, DSSNet achieves the lowest RMSE across all evaluated history lengths on both datasets, and the lowest MAE in all settings except $L = 0$ on TaxiNYC-1. At the default length of $L = 336$, DSSNet reduces MAE by 27.8% and 27.3% relative to STID on TaxiNYC-1 and TaxiNYC-2, respectively, while also outperforming both frequency-domain baselines. Longer histories substantially benefit FilterNet and FilterTS, although their improvements are non-monotonic, whereas STID shows limited benefit. DSSNet generally improves as the history grows and achieves its best results at $L = 336$ on both datasets; extending the bank to $L = 504$ slightly degrades performance. These results indicate that DSSNet's proposed spectral structure profiling method effectively exploits historical information, enabling it to maintain clear predictive advantages under matched history lengths.

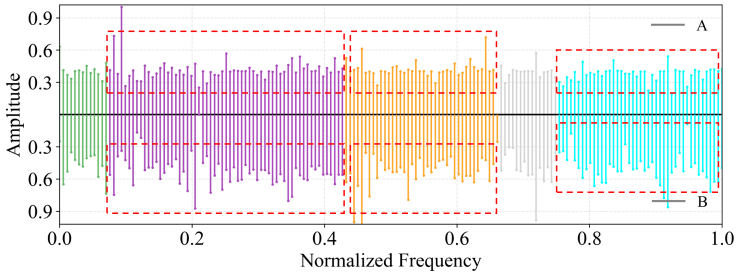


Figure 6: The visualization spectral of different types of regions.

Learned Spectral Structure. Figure 6 illustrates how our Learnable Band Filter Bank (LBFB) captures semantically meaningful spatiotemporal fingerprints. Transportation Hubs (Region B): The lower panel exhibits sharp, high-magnitude energy spikes. This reflects robust traffic regularity governed by rigid operational schedules, validating LBFB's ability to extract dominant periodicities. Residential Areas (Region A): Conversely, the upper panel shows dispersed energy with lower magnitudes. Individualized travel demands create a stochastic traffic nature, diffusing spectral

energy across broader frequencies. This divergence proves DSSNet actively encodes region-specific behavioral semantics, effectively capturing spatial-spectral heterogeneity.

5.5 Model Efficiency

As shown in Figure 7, DSSNet achieves an exceptional balance between efficiency and performance. For a fair comparison, all frequency-domain methods are equipped with a memory bank of length ($L = 336$). DSSNet maintains a modest memory usage and moderate training time, making it a highly practical alternative to resource-intensive baselines. Notably, when the memory bank scales to 336, DSSNet attains superior predictive accuracy with only marginal additional computational overhead. This sub-linear growth pattern demonstrates its strong scalability: DSSNet not only outperforms heavy models such as AGCRN in accuracy, but also remains significantly more memory-efficient than frequency-domain methods like FreTS.

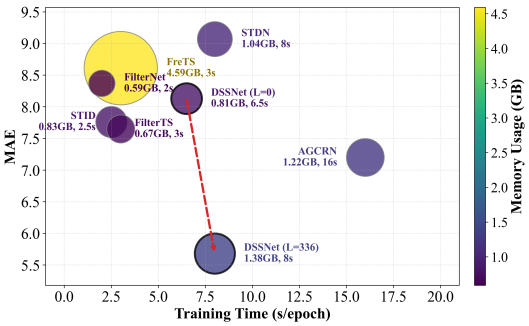


Figure 7: Assessing MAE, training time, and memory usage, evaluated on the TaxiNYC-1 Dataset.

Inference Latency. To assess edge deployment feasibility, we measured actual latency on TaxiNYC-1. As shown in Table 4, DSSNet achieves **9.30 ms** per inference, significantly faster than frequency-domain baselines like FilterNet (15.46 ms) and FilterTS (14.41 ms). Although lightweight temporal models such as STID have lower latency, DSSNet is faster than the evaluated frequency-domain baselines and provides a favorable accuracy-efficiency trade-off.

Table 4: Inference Latency (ms) comparison on TaxiNYC-1.

Method	**DSSNet**	STID	FilterNet	FreTS	FilterTS	STDN	GWNet	AGCRN	STGCN
Time	9.30	1.63	15.46	14.35	14.41	8.48	2.99	12.86	4.95

6 Conclusion

This paper proposes DSSNet, a novel frequency-aware framework that explicitly models spectral structure through its adaptive profiling paradigm. DSSNet introduces a Learnable Band Filter Bank to automatically discover dominant frequency bands. Furthermore, a dual-domain fusion strategy is designed to integrate knowledge in both temporal and frequency domains, enabling the model to jointly capture temporal dynamics and spectral correlations. Extensive experiments demonstrate that DSSNet consistently outperforms baselines across multiple prediction datasets, and revealing meaningful relationships between traffic dynamics and region spectral structures. These results highlight the effectiveness of frequency-domain modeling for spatiotemporal prediction tasks.

Acknowledgments and Disclosure of Funding

This work was supported in part by the Beijing Natural Science Foundation (No.L233034), in part by the National Natural Science Foundation of China (No.62572075, No.62402420), and in part by Fundamental Research Funds for the Beijing University of Posts and Telecommunications (No.2025TSQY01).

References

Lei Bai, Lina Yao, Can Li, Xianzhi Wang, and Can Wang. Adaptive graph convolutional recurrent network for traffic forecasting. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc.

Lingxiao Cao, Bin Wang, Guiyuan Jiang, Yanwei Yu, and Junyu Dong. Spatiotemporal-aware trend-seasonality decomposition network for traffic flow forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(11):11463–11471, Apr. 2025.

Junyoung Chung, Çaglar Gülçehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *CoRR*, abs/1412.3555, 2014.

Razvan-Gabriel Cirstea, Tung Kieu, Chenjuan Guo, Bin Yang, and Sinno Jialin Pan. Enhancenet: Plugin neural networks for enhancing correlated time series forecasting. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, pages 1739–1750, 2021.

Shaojie Dai, Jinshuai Wang, Chao Huang, Yanwei Yu, and Junyu Dong. Dynamic multi-view graph neural networks for citywide traffic inference. *ACM Trans. Knowl. Discov. Data*, 17(4), February 2023.

Emadeldeen Eldele, Mohamed Ragab, Zhenghua Chen, Min Wu, and Xiaoli Li. Tslanet: rethinking transformers for time series representation learning. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org, 2024.

Yuchen Fang, Yanjun Qin, Haiyong Luo, Fang Zhao, Bingbing Xu, Liang Zeng, and Chenxing Wang. When spatio-temporal meet wavelets: Disentangled traffic forecasting via efficient spectral graph attention networks. In *2023 IEEE 39th International Conference on Data Engineering (ICDE)*, pages 517–529, 2023.

Zheng Fang, Qingqing Long, Guojie Song, and Kunqing Xie. Spatial-temporal graph ode networks for traffic flow forecasting. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, page 364–373, New York, NY, USA, 2021. Association for Computing Machinery.

Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):922–929, Jul. 2019.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8): 1735–1780, 1997.

Jiahao Ji, Jingyuan Wang, Zhe Jiang, Jiawei Jiang, and Hu Zhang. Stden: Towards physics-guided neural networks for traffic flow prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4):4048–4056, Jun. 2022.

Bo Jiang, Ziyan Zhang, Doudou Lin, Jin Tang, and Bin Luo. Semi-supervised learning with graph learning-convolutional networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

Renhe Jiang, Zekun Cai, Zhaonan Wang, Chuang Yang, Zipei Fan, Quanjun Chen, Kota Tsubouchi, Xuan Song, and Ryosuke Shibasaki. Deepcrowd: A deep model for large-scale citywide crowd density and flow prediction. *IEEE Transactions on Knowledge and Data Engineering*, 35(1): 276–290, 2023. doi: 10.1109/TKDE.2021.3077056.

Ming Jin, Shiyu Wang, Lintao Ma, Zhixuan Chu, James Y Zhang, Xiaoming Shi, Pin-Yu Chen, Yuxuan Liang, Yuan-Fang Li, Shirui Pan, and Qingsong Wen. Time-LLM: Time series forecasting by reprogramming large language models. In *International Conference on Learning Representations (ICLR)*, 2024.

Taesung Kim, Jinhee Kim, Yunwon Tae, Cheonbok Park, Jang-Ho Choi, and Jaegul Choo. Reversible instance normalization for accurate time-series forecasting against distribution shift. In *International Conference on Learning Representations*, 2022.

Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In *International Conference on Learning Representations (ICLR '18)*, 2018.

Maohan Liang, Ryan Wen Liu, Yang Zhan, Huanhuan Li, Fenghua Zhu, and Fei-Yue Wang. Fine-grained vessel traffic flow prediction with a spatio-temporal multigraph convolutional network. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):23694–23707, 2022. doi: 10.1109/TITS.2022.3199160.

Haoxing Lin, Rufan Bai, Weijia Jia, Xinyu Yang, and Yongjian You. Preserving dynamic attention for long-term spatial-temporal prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '20, page 36–46, New York, NY, USA, 2020. Association for Computing Machinery.

Zheyi Pan, Yuxuan Liang, Weifeng Wang, Yong Yu, Yu Zheng, and Junbo Zhang. Urban traffic prediction from spatio-temporal data using deep meta learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 1720–1730, New York, NY, USA, 2019. Association for Computing Machinery.

Zezhi Shao, Zhao Zhang, Fei Wang, Wei Wei, and Yongjun Xu. Spatial-temporal identity: A simple yet effective baseline for multivariate time series forecasting. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, page 4454–4458, New York, NY, USA, 2022. Association for Computing Machinery.

Chao Song, Youfang Lin, Shengnan Guo, and Huaiyu Wan. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01):914–921, Apr. 2020.

Yulong Wang, Yushuo Liu, Xiaoyi Duan, and Kai Wang. Filterts: Comprehensive frequency filtering for multivariate time series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(20):21375–21383, Apr. 2025.

Chun-Hsin Wu, Jan-Ming Ho, and D.T. Lee. Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems*, 5(4):276–281, 2004.

Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. Timesnet: Temporal 2d-variation modeling for general time series analysis. In *International Conference on Learning Representations*, 2023.

Yuhan Wu, Xiyu Meng, Huajin Hu, Junru Zhang, Yabo Dong, and Dongming Lu. Affirm: Interactive mamba with adaptive fourier filters for long-term time series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(20):21599–21607, Apr. 2025.

Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. Graph wavenet for deep spatial-temporal graph modeling. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, IJCAI'19, page 1907–1913. AAAI Press, 2019.

Huaxiu Yao, Yiding Liu, Ying Wei, Xianfeng Tang, and Zhenhui Li. Learning from multiple cities: A meta-learning approach for spatial-temporal prediction. In *The World Wide Web Conference*, WWW '19, page 2181–2191, New York, NY, USA, 2019a. Association for Computing Machinery. ISBN 9781450366748. doi: 10.1145/3308558.3313577.

Huaxiu Yao, Xianfeng Tang, Hua Wei, Guanjie Zheng, and Zhenhui Li. Revisiting spatial-temporal similarity: a deep learning framework for traffic prediction. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI'19/IAAI'19/EAAI'19. AAAI Press, 2019b.

Weiwei Ye, Songgaojun Deng, Qiaosha Zou, and Ning Gui. Frequency adaptive normalization for non-stationary time series forecasting. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24, Red Hook, NY, USA, 2024. Curran Associates Inc.

Kun Yi, Qi Zhang, Wei Fan, Shoujin Wang, Pengyang Wang, Hui He, Defu Lian, Ning An, Longbing Cao, and Zhendong Niu. Frequency-domain mlps are more effective learners in time series forecasting. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

Kun Yi, Jingru Fei, Qi Zhang, Hui He, Shufeng Hao, Defu Lian, and Wei Fan. Filternet: harnessing frequency filters for time series forecasting. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24, Red Hook, NY, USA, 2024. Curran Associates Inc.

Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, IJCAI'18, page 3634–3640. AAAI Press, 2018.

Yuan Yuan, Jingtao Ding, Jie Feng, Depeng Jin, and Yong Li. Unist: A prompt-empowered universal model for urban spatio-temporal prediction. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, page 4095–4106, New York, NY, USA, 2024. Association for Computing Machinery.

Junbo Zhang, Yu Zheng, and Dekang Qi. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, AAAI'17, page 1655–1661. AAAI Press, 2017.

Xingyu Zhang, Siyu Zhao, Zeen Song, Huijie Guo, Jianqi Zhang, Changwen Zheng, and Wenwen Qiang. Not all frequencies are created equal: Towards a dynamic fusion of frequencies in time-series forecasting. In *Proceedings of the 32nd ACM International Conference on Multimedia*, MM '24, page 4729–4737, New York, NY, USA, 2024a. Association for Computing Machinery.

Xinyu Zhang, Shanshan Feng, Jianghong Ma, Huiwei Lin, Xutao Li, Yunming Ye, Fan Li, and Yew Soon Ong. Frnet: Frequency-based rotation network for long-term time series forecasting. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, page 3586–3597, New York, NY, USA, 2024b. Association for Computing Machinery.

Xiyuan Zhang, Xiaoyong Jin, Karthick Gopalswamy, Gaurav Gupta, Youngsuk Park, Xingjian Shi, Hao Wang, Danielle C. Maddix, and Yuyang Wang. First de-trend then attend: Rethinking attention for time-series forecasting, 2022.

Xiyue Zhang, Chao Huang, Yong Xu, and Lianghao Xia. Spatial-temporal convolutional graph attention networks for citywide traffic flow forecasting. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, CIKM '20, page 1853–1862, New York, NY, USA, 2020. Association for Computing Machinery.

Chuanpan Zheng, Xiaoliang Fan, Cheng Wang, and Jianzhong Qi. Gman: A graph multi-attention network for traffic prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34 (01):1234–1241, Apr. 2020.

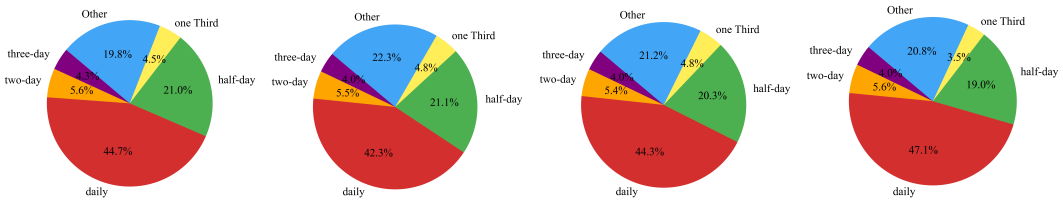
A Feasibility Analysis

To investigate whether frequency structures can effectively characterize the heterogeneity of traffic patterns, we conduct a preliminary empirical study on the T-Drive dataset Pan et al. [2019]. Our goal is to validate the following hypothesis: regions with similar dominant frequency band energy distributions tend to share more consistent traffic flow dynamics.

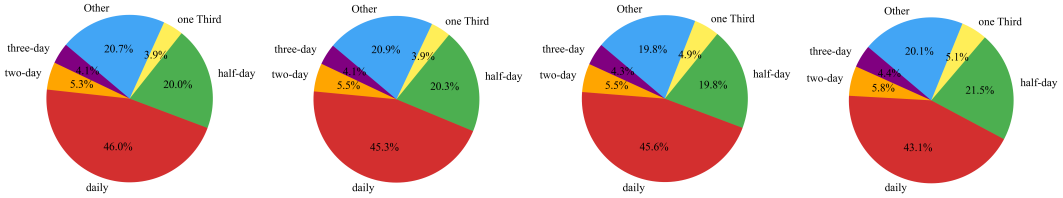
Urban traffic exhibits strong periodicity driven by daily commuting rhythms and weekly behavior cycles. To reveal these periodic structures, we transform each region's hourly traffic-flow sequence into the frequency domain using the Fast Fourier Transform (FFT). Based on typical periodic components observed in urban mobility, we manually define **five** dominant frequency bands corresponding to the following central periods: 1-day (24 hours/cycle), 0.5-day (12 hours/cycle), one-third-day (8 hours/cycle), 2-day (48 hours/cycle), and 3-day (72 hours/cycle) cycles. Each band is extended by a small neighborhood of ±0.2 in frequency center, forming the range:

$$f_{\text{band}} = f_{\text{center}} \pm 0.2 \tag{19}$$

where f_{center} corresponds to the specific frequency of selected five dominant frequency bands. In Figure 1b, the different frequency bands are shown using different colors.



(a) Region 1. (b) Region 2. (c) Region 3. (d) Region 4.



(e) Region 5. (f) Region 6. (g) Region 7. (h) Region 8.

Figure 8: Key frequency band energy ratio of similar regions: (a)–(h) represent the key frequency band energy ratios of different regions

A.1 Experiment Settings

we use all 32×32 grid regions (1024 regions in total) to ensure a comprehensive assessment. For each region, we calculate the proportion of energy contained in each frequency band relative to the total spectral energy:

$$E_i = \frac{\sum_f |F_i(f)|^2}{\sum_{i'}^{5} \sum_f |F_{i'}(f)|^2}, \quad i = 1, \ldots, 5 \tag{20}$$

where $F(f)$ is the FFT coefficient at frequency f, i represents the dominant frequency bands. This yields a compact 5-dimensional band-energy distribution for each region. We then compute the cosine similarity between these vectors to measure pairwise frequency-domain similarity.

To examine whether spectral similarity is predictive of time-domain behavioral similarity, we further normalize traffic flow for each region pair with Revin Kim et al. [2022] and compute the Pearson correlation coefficient between each region pair.

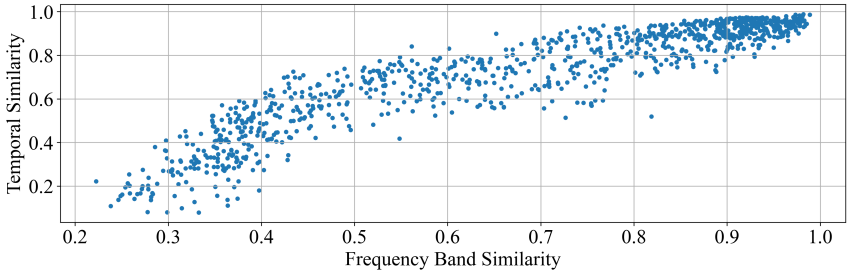


Figure 9: Correlation between temporal and frequency similarities

A.2 Result Analysis

Fig. 9 plots the scatter relationship between frequency domain similarity and temporal domain similarity. The x-axis represents Frequency Band Similarity, calculated as the cosine similarity between the Band Energy Descriptors (BED) of region pairs. The y-axis denotes Temporal Similarity, measured by the Pearson correlation coefficient of normalized traffic flow sequences. A clear positive trend emerges: as frequency domain similarity increases, the corresponding temporal domain similarity improves consistently. This strong monotonic relationship confirms that the spectral energy distribution serves as a reliable discriminator of spatiotemporal patterns.

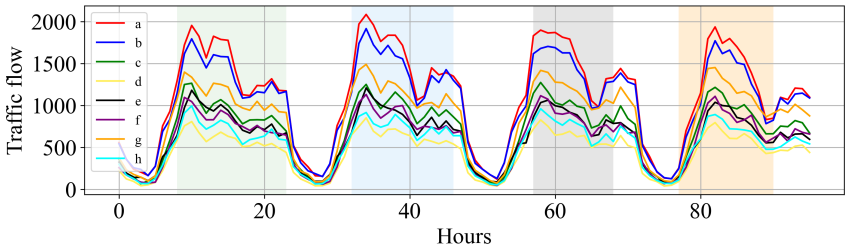


Figure 10: Traffic flow of regions with similar spectral structures

Beyond quantitative results, we visualize the traffic flow of regions with high spectral similarity (Fig. 10) and compare their band-energy distributions (Fig. 8). regions with similar frequency-band energy distributions exhibit highly synchronized temporal behaviors, including aligned peak, valley structures and consistent intra day oscillation patterns In contrast, randomly selected regions (Fig. 1a) display substantial behavioral divergence. These observations further demonstrate that the internal spectral structures of regions encode meaningful spatiotemporal characteristics.

The above empirical results validate the feasibility of exploiting frequency-domain structures to model urban traffic dynamics. In particular, they show that (1) dominant frequency bands capture intrinsic periodic patterns, (2) spectral similarity strongly correlates with time-domain similarity.

B Datasets Details

B.1 Basic Information

T-Drive. The T-Drive dataset Pan et al. [2019] contains massive taxi trajectories collected in Beijing from February 1st, 2015 to June 2nd, 2015. The city is divided into 32×32 grids, and the time interval is set to 1 hour.

TaxiBJ. The BJTaxi dataset Zhang et al. [2017] covers taxi trajectory data in Beijing from March 1st, 2014 to June 30th, 2014. The city is divided into 32×32 grids, and the time interval is set to 30 minutes.

TaxiNYC-1. The NYC-Taxi dataset Yao et al. [2019b] contains 22,349,490 taxi trip records collected in New York City from January 1st, 2015 to March 1st, 2015. The city is divided into 10×20 grids (each approximately 1 km $\times$ 1 km), and the time interval is set to 30 minutes.

TaxiNYC-2. Following Lin et al. [2020], we further use the NYC crowd flow prediction dataset, which includes both taxi and bike flow records over a 60-day period.The city is divided into 16×12 grids (each approximately 1 km $\times$ 1 km), and the time interval is set to 30 minutes.

PEMS-04. Collected from the California Transportation Performance Management System (PeMS), from January 2018 to February 2018 Song et al. [2020]. It has 307 nodes and 16992 time steps. The raw data has a fine temporal granularity of 5 minutes between consecutive timesteps.

B.2 Data Preprocessing

All datasets are normalized to the range $[-1, 1]$. To handle missing values, we apply historical average interpolation. For each dataset, we split it into three non-overlapping periods: the first 70% of the time span for training, the next 15% for validation, and the last 15% for testing. Specifically, for the PEMS04 dataset, we follow the common practice of using a 6:2:2 split ratio. To cover the full weekly cycle and seasonal variations, we maintain the Memory Bank to store recent historical flow sequences of length 168×2 time steps. Following previous practices Jin et al. [2024], Yuan et al. [2024], both the input step and prediction horizon are set as 6, i.e., $6 \rightarrow 6$.

B.3 Implementation Details and Hyperparameter Selection

The default architectural and training configurations of DSSNet are summarized in Table 5. Key hyperparameters are selected based on validation-set performance and kept fixed during testing. Our code is available at `https://github.com/xiaowang005/DSSNet`.

Table 5: Architectural and training configurations of DSSNet.

Configuration	**Value**
Hidden dimension D	128
Structure-Aware GCN layers	1
Multi-scale pooling kernels	$\{3, 7, 15, 31, 63, 127\}$
Number of frequency bands B	8
Memory Bank length L	336
Temperature τ	0.1
Dropout	0.1
Batch size	16
Maximum training epochs	80
Optimizer	AdamW
Learning rate	0.002
Weight decay	0.01
Warmup ratio	0.01

C Baselines

Traditional Methods.

- **HA**: Historical Average predicts future flow using the mean values of historical data.
- **SVR** Wu et al. [2004]: A support vector regression model widely used in time series prediction.
- **LSTM** Hochreiter and Schmidhuber [1997]: A long short-term memory network that captures temporal dependencies in sequential data.

Temporal Domain Methods.

- **STGCN** Yu et al. [2018]: Combines temporal convolution and graph convolution to capture spatial temporal correlations.
- **AGCRN** Bai et al. [2020]: Employs adaptive graph convolution with recurrent units for spatial temporal dependencies.
- **GWNet** Wu et al. [2019]: Spatial temporal method combining adaptive graph convolutions with dilated casual convolution.
- **STDN** Cao et al. [2025]: Models temporal periodicity through temporal embeddings and dynamically learns spatial relationships via a trend-seasonality decomposition mechanism.
- **STID** Shao et al. [2022]: A lightweight MLP-based spatiotemporal predictor that decomposes temporal and spatial factors for efficient modeling.

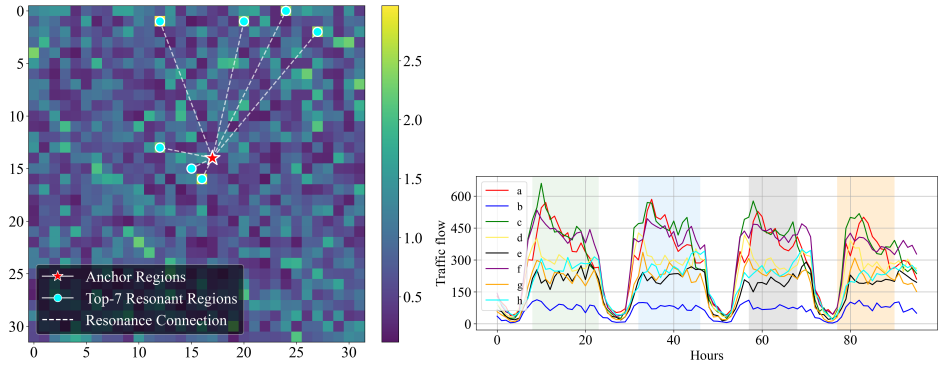
Frequency Domain Methods.

Note that since DSSNet incorporates historical traffic flow during frequency analysis, we provide historical traffic flow to the frequency domain methods to ensure a fair comparison.

- **FilterTS** Wang et al. [2025]: Applies FFT to transform data to the frequency domain, then performs dynamic and static filtering to learn shared and key frequency components.
- **FreTS** Yi et al. [2023]: Conducts discrete Fourier transform on time series, uses frequency-domain MLPs to predict real and imaginary parts separately, and employs channel and temporal learning modules for cross-variable dependencies.

- **FilterNet** Yi et al. [2024]: Employs two learnable frequency filters: a plain shaping filter (Filter Pai) and a contextual shaping filter (Filter Tex) to selectively modulate spectral components for efficient full-spectrum modeling and robust forecasting.

D Additional Analysis



(a) Spatial correlation of the anchor. (b) Traffic flow of the anchor and related nodes.

Figure 11: The visualization of long range region.

D.1 Long-range Resonance

Figure 11 visualizes the long-range connections established by our Resonance Graph to capture global spatiotemporal dependencies. Spatial Dispersion (Fig. 11a): Unlike traditional models restricted by physical adjacency, Figure 11a shows an anchor region and its top-7 resonant neighbors widely dispersed across the urban grid rather than clustered locally. Temporal Synchronization (Fig. 11b): Despite this geographical separation, Figure 11b demonstrates that these spectrally connected regions share highly consistent periodic fluctuations and evolutionary trends. This confirms DSSNet effectively overcomes physical locality constraints. By using spectral structure as a functional proxy, the Resonance Graph successfully uncovers hidden topological connections between spatially distant but behaviorally synchronized regions.

D.2 Noise Data and Missing Data Experiment

Experiment Setting: To explicitly validate our model's robustness and demonstrate its readiness for real-world deployment:

Setting A: Sensor Noise Simulation. We injected Gaussian noise directly into the raw traffic flow data to simulate hardware interference or transmission errors. To evaluate the model's tolerance threshold, we scaled the magnitude of the Gaussian noise to equal 10%, 30%, and 50% of the original signal's standard deviation. Setting B: Missing Value Simulation. To replicate scenarios where sensors temporarily go offline or drop packets, we randomly selected 10%, 30%, and 50% of the spatial nodes in the network. For each of these affected nodes, we randomly masked (dropped) 10% of their continuous time-step observations, forcing the models to predict future traffic based on fragmented historical data. Baselines: We evaluated DSSNet against STID (a robust temporal-domain baseline) and state-of-the-art frequency-domain models (FilterNet, FilterTS)

Results: thee results are shown in Table 6 and 7. across all configurations of both sensor noise and missing values, DSSNet consistently exhibits the lowest performance degradation, maintaining a clear advantage over all baseline models. This exceptional robustness is a direct consequence of our adaptive spectral structure profiling design:

Traditional temporal models are highly sensitive to point-wise local noise. In contrast, DSSNet transforms the temporal signal to isolate dominant frequency bands. The Gaussian noise we injected typically distributes uniformly across the spectrum (white noise). By employing the Learnable Band Filter Bank (LBFB) and the shared Global Frequency Filter, DSSNet selectively preserves

Table 6: Statistical Results of the Noise Experiment. Ratio denotes the noise standard deviation as a percentage of the original signal's standard deviation.

Model	Dataset-1	Ratio	MAE	RMSE	Dataset-2	Ratio	MAE	RMSE
DSSNet	TaxiNYC1	0%	**5.68**	**15.78**	T-Drive	0%	**17.50**	**35.98**
		10%	5.75	15.79		10%	18.13	36.26
		30%	6.21	16.31		30%	19.60	38.28
		50%	6.66	17.54		50%	20.42	40.08
STID	TaxiNYC1	0%	7.76	24.66	T-Drive	0%	19.52	39.95
		10%	8.59	24.47		10%	20.44	41.22
		30%	8.15	24.89		30%	20.54	43.10
		50%	8.53	25.92		50%	21.00	43.84
FilterNet	TaxiNYC1	0%	8.37	24.82	T-Drive	0%	21.57	43.36
		10%	8.54	24.59		10%	22.15	42.84
		30%	10.07	24.35		30%	24.36	45.78
		50%	11.14	30.03		50%	26.38	46.89
FilterTS	TaxiNYC1	0%	7.65	22.27	T-Drive	0%	20.64	42.27
		10%	8.48	22.43		10%	21.85	43.16
		30%	9.37	23.97		30%	24.54	44.69
		50%	10.36	24.89		50%	27.94	48.28

the low-to-mid frequency components that dictate the underlying periodic traffic patterns, while aggressively suppressing the high-frequency spectra where the artificial noise resides. When missing values occur, relying solely on local temporal sequences leads to severe performance drops.

Table 7: Statistical Results of the Missing Data Experiment. Ratio denotes the percentage of spatial nodes affected by missing data.

Model	Dataset-1	Ratio	MAE	RMSE	Dataset-2	Ratio	MAE	RMSE
DSSNet	TaxiNYC1	0%	**5.68**	**15.78**	T-Drive	0%	**17.50**	**35.98**
		10%	6.10	19.73		10%	19.15	47.01
		30%	7.58	25.98		30%	22.22	61.43
		50%	8.37	31.36		50%	24.72	69.68
STID	TaxiNYC1	0%	7.76	24.66	T-Drive	0%	19.52	39.95
		10%	8.09	26.42		10%	21.39	50.58
		30%	8.90	30.57		30%	24.22	63.63
		50%	9.82	35.01		50%	26.52	72.19
FilterNet	TaxiNYC1	0%	8.37	24.82	T-Drive	0%	21.57	43.36
		10%	9.03	28.31		10%	22.12	50.60
		30%	9.56	31.39		30%	27.22	66.27
		50%	11.16	37.43		50%	29.21	73.63
FilterTS	TaxiNYC1	0%	7.65	22.27	T-Drive	0%	20.64	42.27
		10%	7.70	23.03		10%	21.76	49.77
		30%	8.93	28.48		30%	25.43	63.97
		50%	10.14	34.05		50%	28.42	72.46

D.3 Sudden Event Experiment

Experiment Setting: To simulate the instantaneous traffic surges caused by unexpected real-world events (such as congestion propagation following an accident), we injected extreme perturbations into two benchmark datasets characterized by complex spatiotemporal heterogeneity: TaxiNYC-1 and T-Drive. The specific settings are as follows:

Table 8: Statistical Results of the Sudden Event Experiment. Ratio denotes the proportion of nodes experiencing abrupt changes.

Model	Dataset-1	Ratio	MAE	RMSE	Dataset-2	Ratio	MAE	RMSE
DSSNet	TaxiNYC1	0%	**5.68**	**15.78**	T-Drive	0%	**17.50**	**35.98**
		10%	6.59	21.30		10%	18.78	44.94
		30%	7.42	25.73		30%	20.70	57.02
		50%	7.99	30.82		50%	22.13	65.35
STID	TaxiNYC1	0%	7.76	24.66	T-Drive	0%	19.52	39.95
		10%	8.91	29.44		10%	22.21	48.46
		30%	9.67	33.92		30%	23.30	59.14
		50%	10.38	38.08		50%	24.72	67.17
FilterNet	TaxiNYC1	0%	8.37	24.82	T-Drive	0%	21.57	43.36
		10%	10.30	32.13		10%	26.66	57.81
		30%	12.14	39.04		30%	26.21	64.70
		50%	12.85	42.71		50%	30.48	76.40
FilterTS	TaxiNYC1	0%	7.65	22.27	T-Drive	0%	20.64	42.27
		10%	9.57	26.90		10%	22.32	49.41
		30%	10.36	31.57		30%	25.08	62.34
		50%	10.98	36.02		50%	27.40	71.43

We simulated extreme sudden events by artificially amplifying the ground-truth traffic flow by a multiplier of 3. Temporal Sparsity: To reflect the "instantaneous" and unpredictable nature of these events, the extreme surges were randomly injected into exactly 3% of the total time steps in the testing set. Recognizing that real-world events have varying areas of effect, we established three distinct spatial impact levels. We randomly selected 10%, 30%, and 50% of the spatial nodes to receive these extreme perturbations, allowing us to evaluate robustness against both localized incidents and wide-scale network disruptions. We evaluated our model against a strong temporal-domain baseline (STID) and state-of-the-art frequency-domain models (FilterNet, FilterTS).

Results: the results are shown in Table 8. the experimental results explicitly demonstrate that DSSNet consistently achieves the best performance and maintains remarkable stability, even when up to 50% of the spatial nodes experience extreme, sudden traffic surges. This robust performance under non-stationary conditions is fundamentally rooted in our decouple-and-filter design.

When isolated in the frequency domain, the residual effects of sudden events manifest as high-frequency broadband noise. Our Learnable Band Filter Bank (LBFB) and Global Frequency Filter act as a precise defense mechanism, effectively anchoring onto the dominant periodic bands while aggressively suppressing this event-induced broadband noise. By filtering out the chaotic noise while preserving the core periodic signals, DSSNet prevents catastrophic spectral drift. The model successfully maintains the region's stable spectral identity while relying on the dual-domain fusion to accommodate sudden temporal shifts.

Table 9: Parameter sensitivity analysis of the historical buffer length (left) and frequency bands (right). The best results are highlighted in **bold**.

Param.	**TaxiNYC-1**		**T-Drive**		**Param.**	**TaxiNYC-1**		**T-Drive**	
	MAE	RMSE	MAE	RMSE		MAE	RMSE	MAE	RMSE
$L = 0$	8.13	24.34	19.11	39.36	$B = 2$	6.11	16.24	17.89	36.65
$L = 18$	7.83	24.30	18.84	38.72	$B = 4$	5.99	16.31	17.66	36.57
$L = 72$	6.81	20.64	18.57	37.82	$B = 6$	5.94	15.99	17.59	36.18
$L = 168$	6.04	17.56	17.84	36.18	$B = 10$	5.79	16.13	17.61	36.24
$L = 504$	5.84	16.33	17.65	36.22	$B = 12$	5.80	15.99	17.72	36.43
DSSNet (336)	**5.68**	**15.78**	**17.50**	**35.98**	**DSSNet (8)**	**5.68**	**15.78**	**17.50**	**35.98**

D.4 Parameter Sensitivity Analysis

To evaluate the robustness of DSSNet to key hyperparameters, we conduct sensitivity experiments focusing on (1) the number of frequency bands $B \in \{2, 4, 6, 10, 12\}$ ($B = 8$ as default) and (2) the buffer length $L \in \{0, 18, 72, 168, 504\}$ ($L = 336$ as default). The results are shown in Table 9. Notably, with a buffer size of 0, DSSNet maintains competitive performance, outperforming most baselines. As L increases, performance improves due to enhanced FFT resolution, though excessive length (e.g., 504) introduces redundancy. This gain is primarily attributed to the enhanced FFT resolution, which allows for more precise frequency extraction. For frequency bands, ($B = 8$) provides an optimal balance between information richness and efficiency. Overall, buffer length has a greater impact, mainly due to its influence on FFT resolution.

D.5 Memory Bank Initialization

The Memory Bank is initialized as empty and progressively filled as chronologically ordered input windows are processed. Before reaching its capacity L, it retains all causally available historical observations. Once full, it follows a First-In-First-Out (FIFO) update rule, removing the oldest observations and appending newly available ones. The same causal procedure is used during training, validation, and testing: future observations and prediction targets are never inserted before the corresponding prediction is made.

D.6 Limitations

The paper acknowledges limitations regarding the current construction of the Resonance Graph. The approach may still result in high connectivity in large-scale networks, potentially increasing computational overhead. Future work will explore advanced sparsification techniques or dynamic pruning to ensure only critical long-range spectral dependencies are preserved.

D.7 Broader Impacts

DSSNet could support traffic management and transportation planning through improved traffic flow forecasting, potentially helping reduce congestion and associated energy consumption. However, inaccurate forecasts or uneven performance across regions could lead to inefficient or inequitable resource allocation. Future deployments involving fine-grained mobility data may also raise privacy concerns. Responsible use should therefore include regional performance evaluation, appropriate data aggregation and access controls, and human oversight of consequential decisions.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction clearly outline the paper's primary contributions: the DSSNet framework, the Learnable Band Filter Bank for adaptive spectral profiling, and the Dual-Domain Knowledge Integration module. These claims are consistently supported by comprehensive experimental results across five benchmark datasets.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss the limitations of our work in appendix D.6. Specifically, we reflect on the potential redundancy in the Resonance Graph construction and the computational scaling of adaptive filters for extremely large-scale urban networks. We also outline future research directions to address these constraints through more advanced graph sparsification and dynamic pruning strategies.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [N/A]

   Justification: This paper is an empirical study focused on novel architectural design and experimental validation. It does not present formal theoretical theorems, lemmas, or mathematical proofs that require a set of assumptions or formal derivation.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The paper fully discloses the DSSNet architecture through detailed mathematical formulations (Eq. 1–18) and a structural diagram. We specify all essential experimental settings, including data split ratios, preprocessing steps, and hyperparameters like the AdamW optimizer and a 0.002 learning rate. Crucially, an code repository link is provided in the appendix to ensure complete reproducibility of our results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: We provide a GitHub link containing the complete source code and implementation scripts for DSSNet. The repository includes detailed instructions for reproducing results on the five public benchmark datasets used in our study. All experimental configurations, such as data split ratios and hyperparameters, are fully disclosed.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The paper satisfies all requirements by explicitly defining the training and testing protocols within the main text and appendix. We provide all necessary information regarding data split ratios, optimization strategies, and hyperparameter selection logic to ensure the results are fully interpretable. Furthermore, the specific hardware environment and software configurations are disclosed to support a complete understanding of the experimental setup.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: We repeat the experiments five times with different random seeds and report the average MAE and RMSE to reduce the influence of randomness on the reported results.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The paper satisfies the requirements by explicitly stating the specific type of GPU hardware used for all experiments and providing a comparative analysis of memory consumption and execution time. We disclose the computational efficiency of both our proposed model and the baselines to ensure that the resources required for reproduction are clearly understood. Additionally, the sub-linear scaling of our model's overhead relative to history length is documented to demonstrate the practical compute demands for long-term forecasting tasks.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: To the best of our knowledge, the research described in this paper conforms to the NeurIPS Code of Ethics. The work uses public benchmark datasets, preserves anonymity at submission time, and does not involve human subjects or sensitive private data collection.

Guidelines:

- The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper discusses the broader societal implications in the Appendix. On the positive side, accurate traffic prediction with proposed method can significantly improve urban efficiency while reducing carbon emissions and energy consumption. We also acknowledge that while the model is intended for public management, its deployment requires careful consideration of data privacy regarding individual mobility patterns.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release high-risk generative models, scraped sensitive datasets, or other assets that would require special misuse safeguards beyond standard anonymized code release and citation of public benchmark data sources.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.

- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: We properly credit the public datasets and code framework used in this work, and we explicitly state their sources and licenses or terms of use in the appendix.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: This paper does not introduce new datasets or external assets; instead, it proposes a new model architecture and evaluates it using established, publicly available benchmark datasets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing experiments or research with human subjects. All experiments are conducted on existing public benchmark datasets.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing, user studies, or any direct research with human subjects, and therefore IRB approval or an equivalent review process was not required.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The development and implementation of the core methodology in this research, specifically the proposed framework, do not involve the use of Large Language Models (LLMs) as an original or non-standard component. LLMs does not impact the core methodology, scientific rigor, or originality of the work.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.