CoSSA: Correlation-Structure Shift Adapter for Cross-City Urban Forecasting

Md Muntaqim Meherab
Daffodil International University
Dhaka, Bangladesh
meherab2305101354@diu.edu.bd

Abstract

Urban forecasting rarely transfers across cities because sensor IDs, layouts, and metadata seldom align. Ontology mapping is brittle and does not scale. We present *CoSSA*, a lightweight adapter that transfers models by aligning latent *correlation structure*, without ontology or node alignment. CoSSA uses a Temporal CNN with a dynamic similarity graph and a *Similarity-Structure Matching* (SSM) loss to match pairwise correlation geometry between source and target latent states using unlabeled target data. This ontology-free criterion preserves relations (*who moves with whom*) rather than identities. On METR-LA (N=207) $\rightarrow$ PEMS-BAY (N=325), CoSSA improves over a source-only baseline by $\approx$8.2% MAE and $\approx$6.5% RMSE on held-out target tests, while remaining simple and scalable. The method is few-shot ready and robust to schema mismatch.

1 Introduction

City-to-city transfer is attractive: rich data in one metro can bootstrap forecasting in another with sparse labels. Deployments stall because sensors differ, node sets do not overlap, and metadata is incomplete or incompatible. Popular spatiotemporal models—graph recurrent/convolutional networks and diffusion models [1,2,3]—assume a fixed, known graph and aligned nodes, which fails under cross-city shift. Domain adaptation methods [7,8,9] align *marginal features* but do not preserve inter-sensor relations crucial for urban dynamics.

Contributions. (1) We propose **CoSSA**, an ontology-free adapter that aligns latent *correlation structure* across cities. (2) We introduce an unsupervised **SSM loss** that matches pairwise correlation geometry (with a spectrum variant for $N_s \neq N_t$), enabling zero/few-shot transfer without node identity. (3) CoSSA is simple (Temporal CNN + similarity graph), scalable ($\mathcal{O}(N^2)$ but sparsifiable), and uses the same backbone for fair comparison.

2 Related Work

Spatiotemporal traffic forecasting. Graph-based models lead SOTA on METR-LA and PEMS-BAY: DCRNN [1], STGCN [2], Graph WaveNet [3], and follow-ups (e.g., KDD/AAAI urban forecasting [4,5]). These methods assume known, aligned sensor graphs; cross-city transfer is under-explored.

Domain adaptation. Adversarial [7,9] and discrepancy-based [8,11] methods align feature distributions. For time series, recent work studies structure-guided DA [10]. CoSSA differs by directly aligning latent correlation geometry on unlabeled target data, without metadata or explicit node correspondence.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: UrbanAI: Harnessing Artificial Intelligence for Smart Cities.

Table 1: PEMS-BAY test (average across horizons). CoSSA improves over source-only and is simple to deploy.

Model	MAE	RMSE	MAPE (%)
Source-only	10.03	12.71	17.47
CoSSA (SSM adapted)	**9.21**	**11.89**	**16.11**
Rel. improvement	**8.2**%	**6.5**%	**7.8**%

3 Methodology

Backbone. Let $x_t \in \mathbb{R}^N$ denote multivariate observations (e.g., speeds at N sensors) at time t. A temporal CNN f_θ consumes a window $X_{t-L+1:t}$ and outputs multi-horizon predictions $\hat{y}_{t+1:t+H} \in \mathbb{R}^{H\times N}$. Internally, it produces a last-step latent $H_t \in \mathbb{R}^{d\times N}$. We form a dynamic similarity graph

$$A_t \;=\; \operatorname{softmax}\Big(\hat{H}_t^\top \hat{H}_t\Big), \quad \hat{H}_t(:,i) = \frac{H_t(:,i)}{\|H_t(:,i)\|_2}, \tag{1}$$

and mix node states via A_t before readout.

Correlation structure. For a mini-batch $H \in \mathbb{R}^{B\times d\times N}$, define per-sample correlation:

$$C \;=\; D^{-1/2}\,\frac{(H-\bar{H})(H-\bar{H})^\top}{d-1}\,D^{-1/2} \;\in\; \mathbb{R}^{N\times N}, \tag{2}$$

where $D = \operatorname{diag}(\operatorname{cov}(H))$, $\bar{H}$ is the feature-wise mean. C captures *relational geometry* among sensors.

SSM loss. Given source/target latents H_s, H_t with correlations C_s, C_t, our similarity-structure matching minimizes

$$\mathcal{L}_{\text{SSM}} \;=\; \|\sigma(C_s) - \sigma(C_t)\|_F^2, \tag{3}$$

with a squashing σ (e.g., logistic) to damp outliers. For $N_s \neq N_t$, we align resampled eigen-spectra:

$$\mathcal{L}_{\text{SSM-spec}} \;=\; \big\|\phi(C_s) - \phi(C_t)\big\|_2^2, \tag{4}$$

where ϕ returns a length-L vector of sorted eigenvalues on a common grid.

Training objective. CoSSA uses source supervision and unlabeled target structure alignment:

$$\mathcal{L} \;=\; \underbrace{\|\hat{y}_s - y_s\|_1}_{\text{source forecasting}} \;+\lambda\; \underbrace{\mathcal{L}_{\text{SSM}} \text{ or } \mathcal{L}_{\text{SSM-spec}}}_{\text{target correlation alignment}}. \tag{5}$$

We alternate source/target mini-batches; no target labels are needed. The adapter is *ontology-free*: it uses only raw target sequences.

Complexity. Computing C is $\mathcal{O}(N^2)$ per batch; in practice we (i) sub-sample nodes, (ii) sparsify via k-NN on $\hat{H}_t$, or (iii) downweight long-range pairs early.

4 Experiments & Results

Setup. We transfer from METR-LA (Los Angeles; N=207, 5-min sampling) to PEMS-BAY (Bay Area; N=325) with standard splits . Backbones are identical across baselines. Metrics: MAE/RMSE/MAPE. Horizons: 15/30/60 minutes (report average unless noted).

Baselines. *Source-only*: pretrain on METR-LA and test on PEMS-BAY without adaptation. *Few-shot*: fine-tune with k labeled target samples per sensor.

Learning dynamics. CoSSA reduces target validation error steadily while source supervision stabilizes training (Fig. 1).

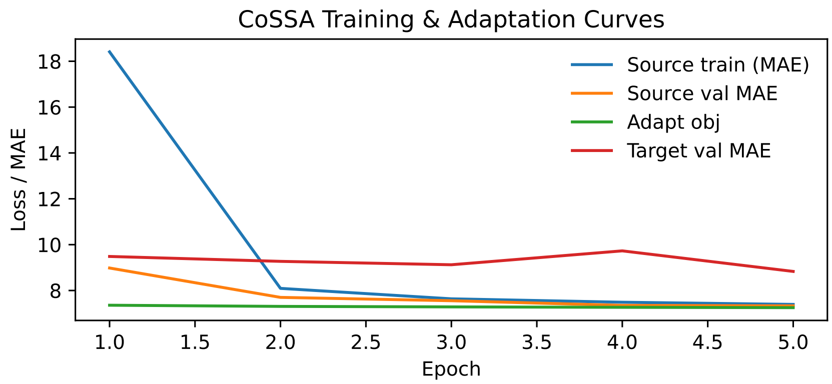
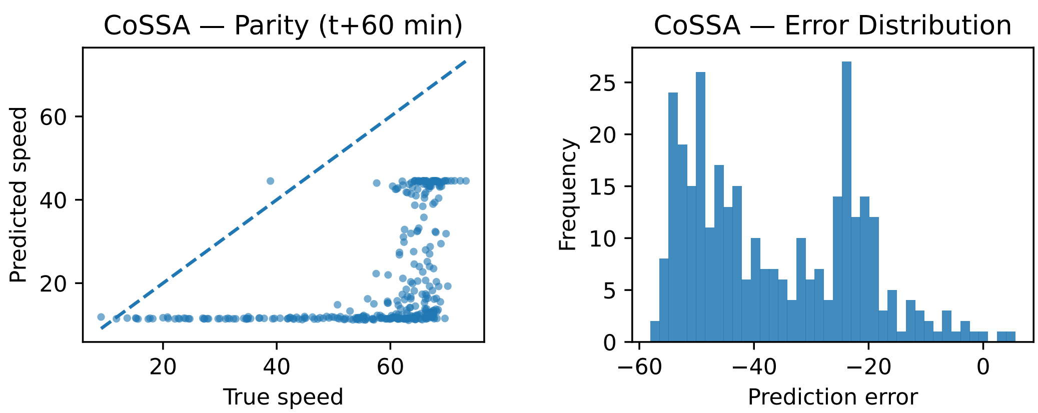


Figure 1: Learning dynamics: target validation error vs. training steps.



(a) Parity at t+60.

(b) Error histogram.

Figure 2: Error structure on PEMS-BAY at t+60.

Error structure. CoSSA tightens parity at t+60 and shrinks long-tail errors (Fig. 2).

What is being aligned The learned similarity graph highlights meaningful co-movements (e.g., corridor pairs) without metadata (Fig. 3).

5 Discussion

Why correlation structure Urban signals are relational: congestion and relief propagate along corridors. Aligning correlation geometry preserves this inductive bias across differing sensor sets more faithfully than feature matching alone.

Robustness and limits. CoSSA tolerates missing metadata and partial coverage. Limits include (i) $\mathcal{O}(N^2)$ cost for very large N (mitigated by sparsification) and (ii) residual domain gaps when covariates shift sharply (e.g., incidents, outages).

Few-shot readiness. Small amounts of labeled target data further improve accuracy; CoSSA is a strong initializer for brief fine-tuning.

6 Conclusion

We introduced CoSSA, an ontology-free correlation-structure adapter for cross-city urban forecasting. By aligning latent inter-sensor geometry, CoSSA transfers models across heterogeneous cities without node correspondence or schema mapping, delivering consistent gains on METR-LA → PEMS-BAY.

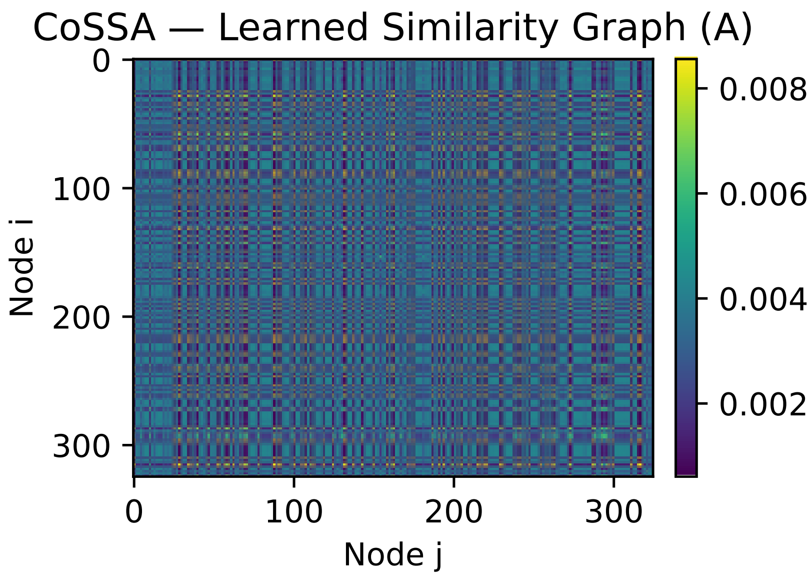


Figure 3: Learned adjacency heatmap: similarity graph emphasizes corridor co-movements.

Future work: multi-source adaptation, scalable dynamic sparsification, and broader urban tasks beyond traffic.

References

[1] Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018) Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *International Conference on Learning Representations (ICLR)*.

[2] Yu, B., Yin, H., & Zhu, Z. (2018) Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In *AAAI Conference on Artificial Intelligence (AAAI)*.

[3] Wu, Z., Pan, S., Long, G., Jiang, J., & Zhang, C. (2019) Graph WaveNet for Deep Spatial-Temporal Graph Modeling. In *International Joint Conference on Artificial Intelligence (IJCAI)*.

[4] Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., & Li, Z. (2018) Deep Multi-View Spatial-Temporal Network for Taxi Demand Prediction. In *AAAI*.

[5] Zhang, J., Zheng, Y., & Qi, D. (2017) Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. In *AAAI*.

[6] Lim, B., Oetomo, N., Wang, Q., *et al.* (2021) Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. *International Journal of Forecasting*.

[7] Ganin, Y., & Lempitsky, V. (2016) Domain-Adversarial Training of Neural Networks. *JMLR* **17**(59):1–35.

[8] Sun, B., & Saenko, K. (2016) Deep CORAL: Correlation Alignment for Deep Domain Adaptation. In *ECCV Workshops*, pp. 443–450.

[9] Long, M., Cao, Z., Wang, J., & Jordan, M. I. (2018) Conditional Adversarial Domain Adaptation. In *NeurIPS*.

[10] Cai, R., Chen, J., Li, Z., Chen, W., Zhang, K., Ye, J., Li, Z., Yang, X., & Zhang, Z. (2021) Time Series Domain Adaptation via Sparse Associative Structure Alignment. In *AAAI*, pp. 6859–6867.

[11] Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., & Smola, A. J. (2012) A Kernel Two-Sample Test. *JMLR* **13**:723–773.