

Graph Neural Controlled Differential Equations for Traffic Forecasting

Jeongwhan Choi, Hwangyong Choi, Jeehyun Hwang, Noseong Park

Department of Artificial Intelligence, Yonsei University

{jeongwhan.choi, hwangyong753, hwanggh96, noseong}@yonsei.ac.kr



Overview

- We present the method of spatio-temporal graph neural controlled differential equation (**STG-NCDE**).
- We design two NCDEs for learning the temporal and spatial dependencies of traffic conditions and combine them into a single framework.
- **STG-NCDE** is robust to the irregularity of the temporal sequence.
- Our large-scale experiments with 6 datasets and 20 baselines clearly show the efficacy of the proposed method.

Spatio-temporal Data

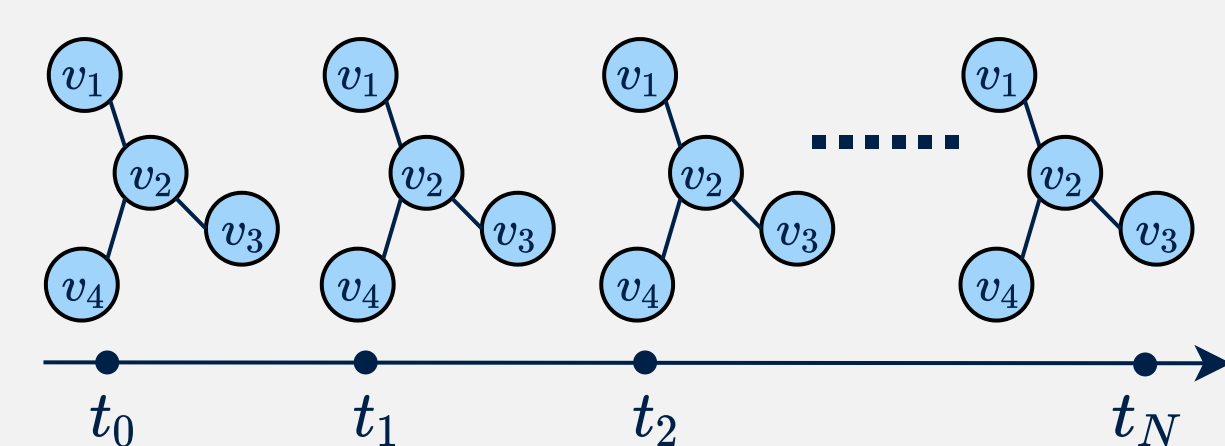


Figure 1: Example of spatio-temporal graph data

- The spatio-temporal graph data frequently happens in real-world applications.
- For instance, the traffic forecasting task is one of the most popular problems in the area of spatio-temporal processing.

Neural Controlled Differential Equations (NCDEs)

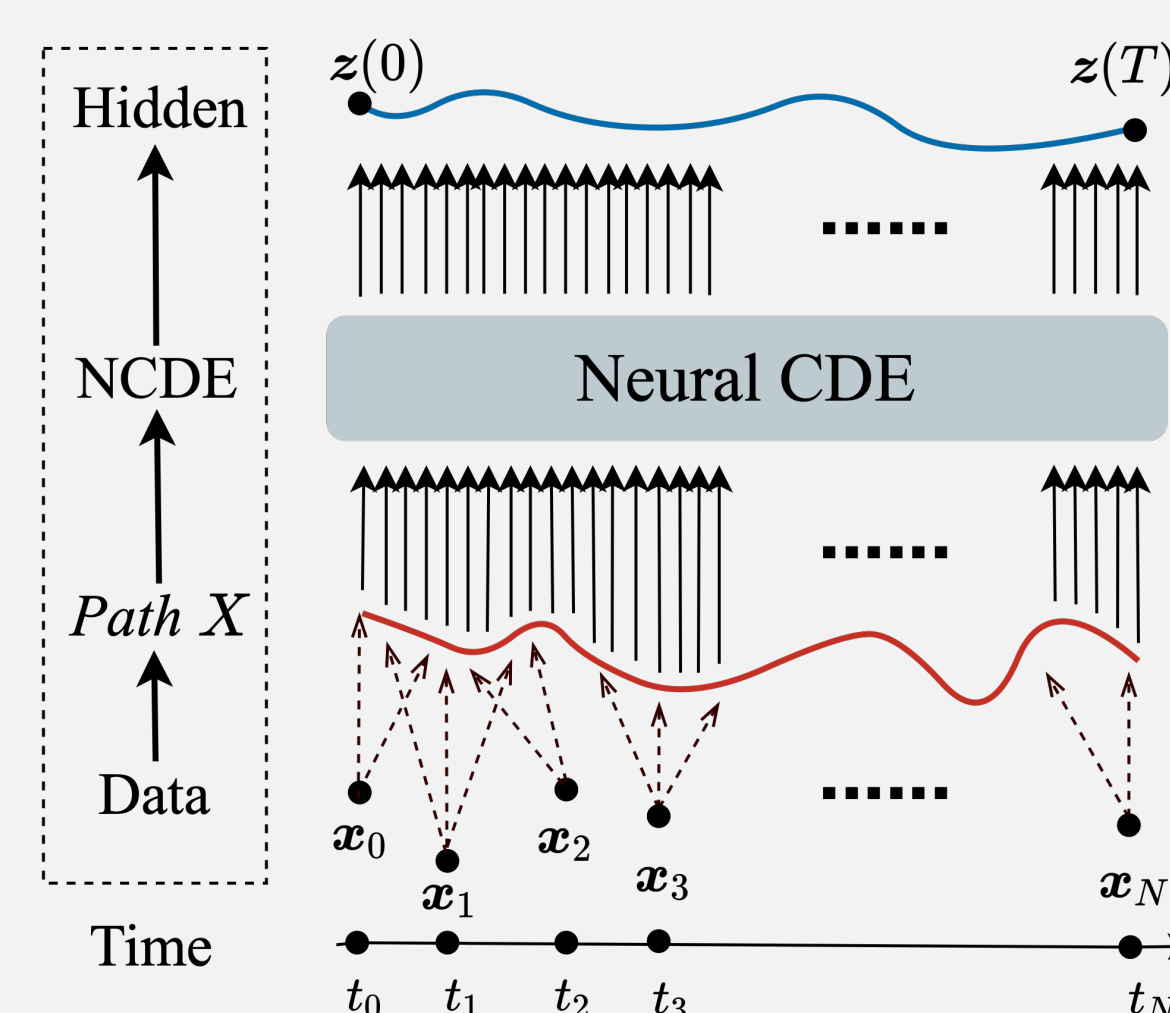


Figure 2: The overall workflow of the original NCDE for processing time-series.

- NCDEs generalize RNNs in a continuous manner and can be written as follows:

$$z(T) = z(0) + \int_0^T f(z(t); \theta_f) dX(t) \quad (1)$$

$$= z(0) + \int_0^T f(z(t); \theta_f) \frac{dX(t)}{dt} dt, \quad (2)$$

- It has not been studied yet how to combine the NCDE technology and the graph convolutional processing technology.
- We integrate them into a single framework to solve the spatio-temporal forecasting problem.

Graph Neural Controlled Differential Equations

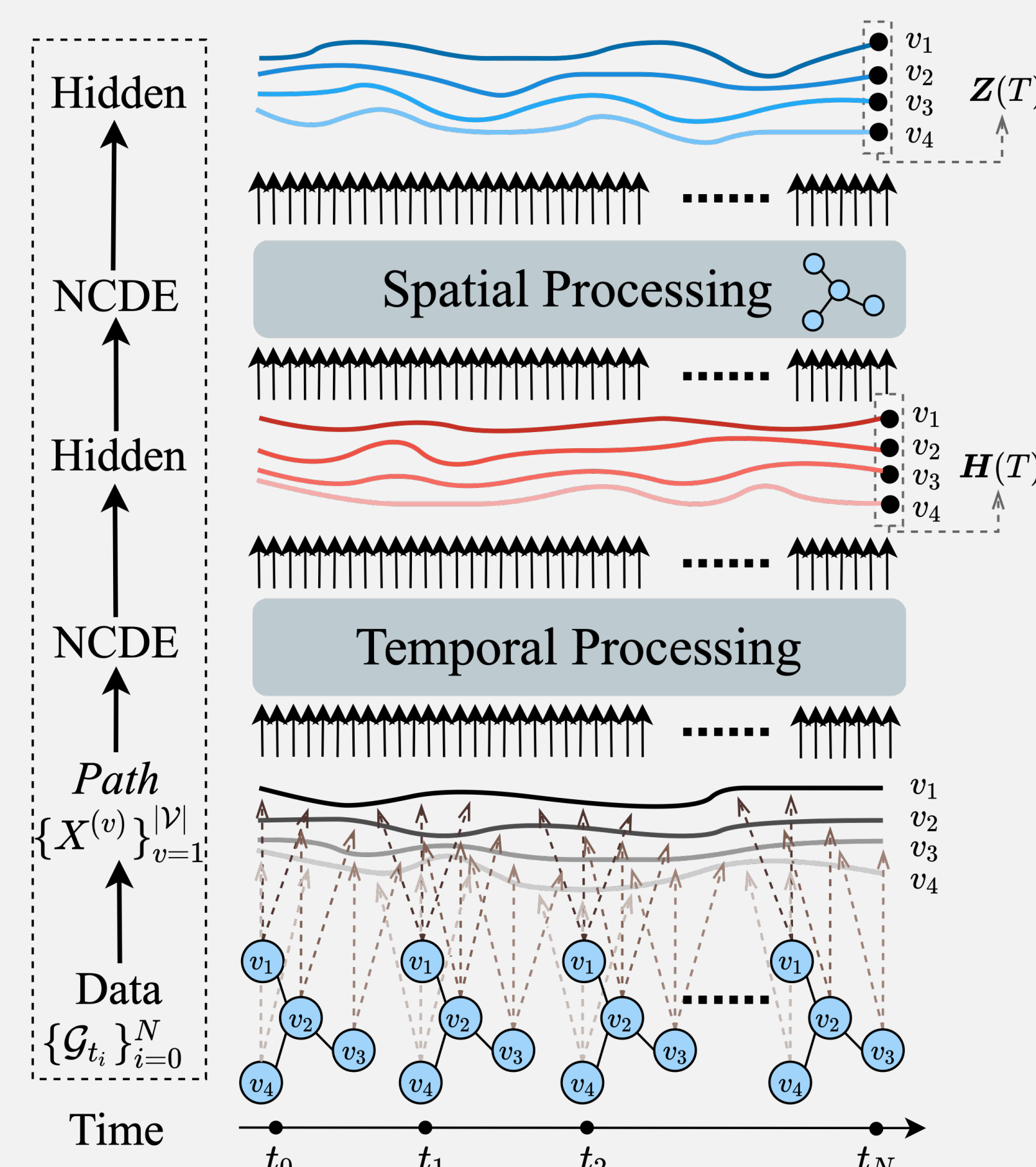


Figure 3: The overall workflow in our proposed STG-NCDE.

- Our proposed spatio-temporal graph neural controlled differential equation (**STG-NCDE**) consists of two NCDEs.

Temporal processing.

- The first NCDE for the temporal processing can be written as follows:

$$h^{(v)}(T) = h^{(v)}(0) + \int_0^T f(h^{(v)}(t); \theta_f) \frac{dX^{(v)}(t)}{dt} dt, \quad (3)$$

where $h^{(v)}(t)$ is a hidden trajectory (over time $t \in [0, T]$) of the temporal information of node v .

- Eq. (3) can be rewritten as follows using the matrix notation:

$$H(T) = H(0) + \int_0^T f(H(t); \theta_f) \frac{dX(t)}{dt} dt, \quad (4)$$

where $X(t)$ is a matrix whose v -th row is $X^{(v)}(t)$.

Spatial processing.

- The second NCDE starts for its spatial processing as follows:

$$Z(T) = Z(0) + \int_0^T g(Z(t); \theta_g) \frac{dH(t)}{dt} dt, \quad (5)$$

where the hidden trajectory $Z(t)$ is controlled.

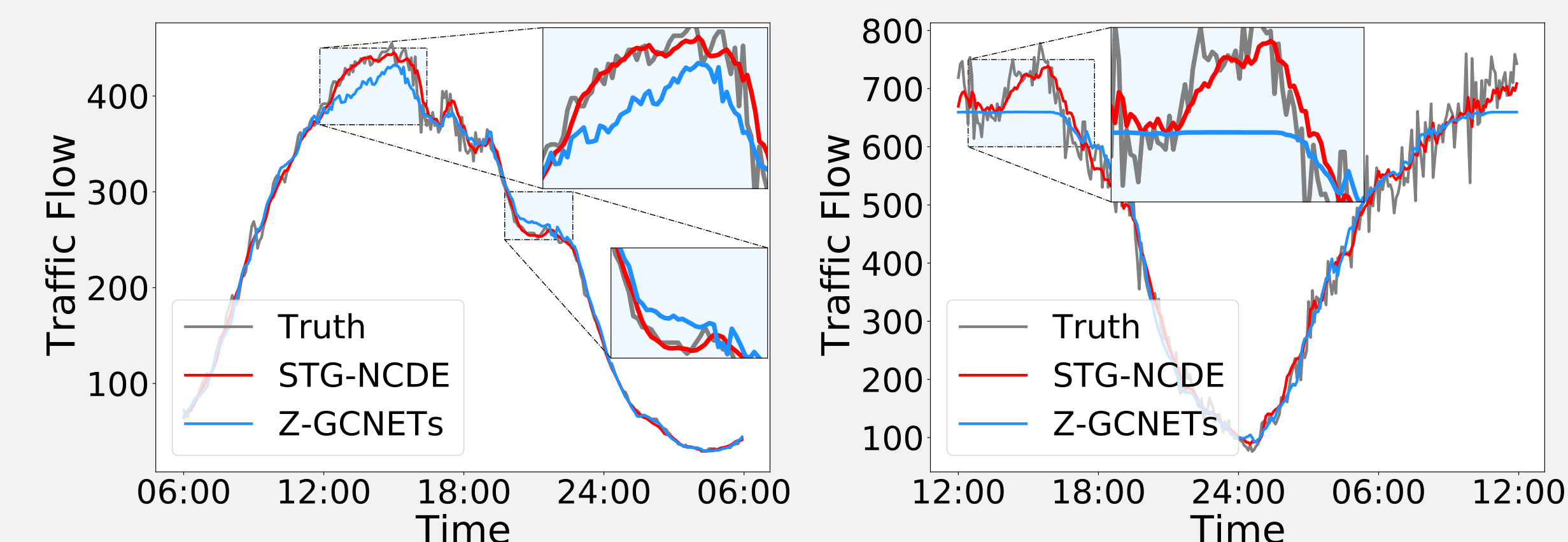
- After combining Eqs. (4) and (5), we have the following single equation which incorporates both the temporal and the spatial processing:

$$Z(T) = Z(0) + \int_0^T g(Z(t); \theta_g) f(H(t); \theta_f) \frac{dX(t)}{dt} dt. \quad (6)$$

Experiments

Table 1: The average error of some selected highly performing models across all the six datasets. Inside the parentheses, we show their performance relative to our method.

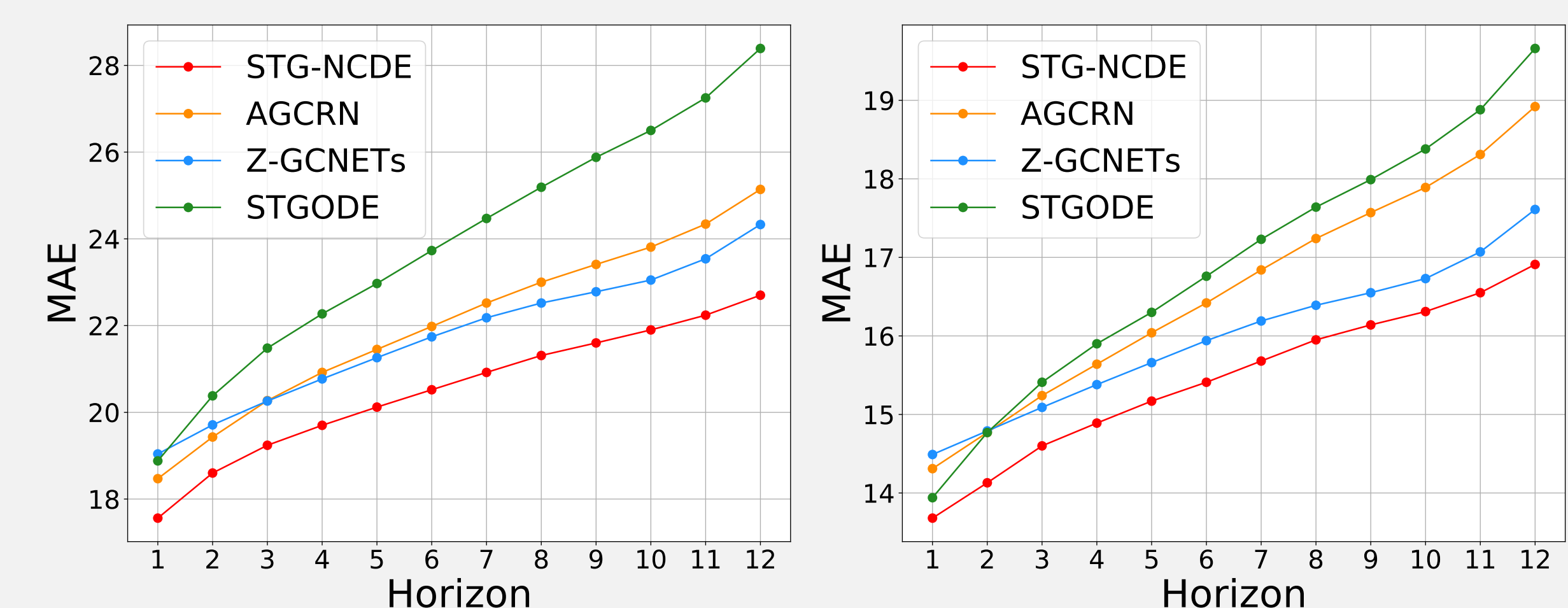
Model	MAE	RMSE	MAPE
STGCN	14.88 (117.0%)	24.22 (113.6%)	12.30 (121.8%)
DCRNN	14.90 (117.1%)	24.04 (112.7%)	12.75 (126.1%)
GraphWaveNet	15.94 (125.3%)	26.22 (122.9%)	12.96 (128.2%)
ASTGCN(r)	14.86 (116.9%)	23.95 (112.3%)	12.25 (121.3%)
STSGCN	14.45 (113.5%)	23.58 (110.5%)	11.42 (113.0%)
AGCRN	13.32 (104.7%)	22.29 (104.5%)	10.37 (102.7%)
STFGNN	13.92 (109.5%)	22.57 (105.8%)	11.30 (111.9%)
STGODE	13.56 (106.6%)	22.37 (104.8%)	10.77 (106.6%)
Z-GCNETs	13.22 (104.0%)	21.92 (102.7%)	10.44 (103.4%)
STG-NCDE	12.72 (100.0%)	21.33 (100.0%)	10.10 (100.0%)



(a) Node 261 in PeMSD4

(b) Node 9 in PeMSD8

Figure 4: Traffic forecasting visualization.



(a) MAE on PeMSD7

(b) MAE on PeMSD8

Figure 5: Prediction error at each horizon.

Table 2: Forecasting error on irregular PeMSD4

Model	Missing rate	MAE	RMSE	MAPE
	10%	19.36	31.28	12.79%
STG-NCDE	30%	19.40	31.30	13.04%
	50%	19.98	32.09	13.48%

- We, for the first time, perform irregular traffic forecasting to reflect real-world environments where sensing values can be missing.