Full Bayesian Significance Testing for Neural Networks in Traffic Forecasting

Zehua Liu[1], Jingyuan Wang[1,2,3*], Zimeng Li[1], Yue He[4]

[1]School of Computer Science and Engineering, Beihang University, Beijing, China

[2]School of Economics and Management, Beihang University, Beijing, China

[3]Key Laboratory of Data Intelligence and Management (Beihang University), Ministry of Industry and Information Technology, Beijing, China

[4]Department of Computer Science and Technology, Tsinghua University, Beijing, China

Background: Traffic Data

- Intelligent Transportation Systems (ITS) is crucial for tackling the growing challenges of increasing transportation demands.

- On one hand, substantial amounts of daily traffic data, including flow, volume and speed, are collected via city sensors.

Background: Traffic Forecasting

- Intelligent Transportation Systems (ITS) is crucial for tackling the growing challenges of increasing transportation network demands.

- On the other hand, traffic forecasting, a core constituent of ITS, plays a crucial role across traffic management, planning and control functionalities.

Traffic Congestion

Urban Planning Team

We can apply traffic forecasting methods to predict future speeds.

Background: Existing Methods

- The existing traffic forecasting methods can be divided into two categories: statistical models and deep learning methods.

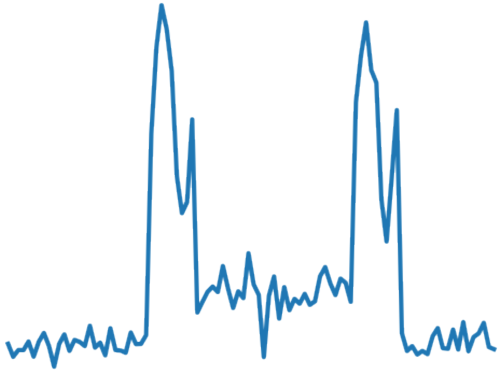
- Statistical models assume stringent data assumptions and constrain their abilities to capture intricate non-linear correlations.
  - ARIMA (Ahmed and Cook 1979; Min and Wynter 2011).

- Deep learning methods benefit from powerful approximation abilities, leveraging GNNs to extract spatial dependencies from graphs and sequence learning techniques to capture temporal dependencies.
  - DCRNN (Li et al. 2018), STGCN (Yu et al. 2018), STTN (Xu et al. 2020), ACGRN (Bai et al. 2020), CCRNN (Ye et al. 2021).

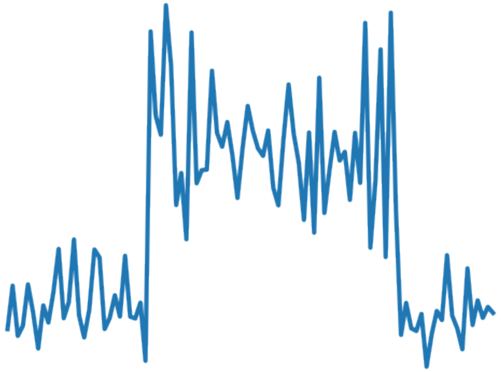
Background: Motivation

- First, most existing approaches involve inputting all factors as features into the model, lacking analysis on

① whether features are significant or not under different circumstances

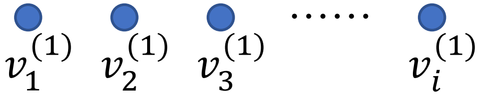


Traffic Flow on Weekdays (morning and evening peaks)

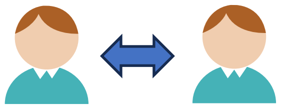


Traffic Flow on Weekends (dispersed and disordered)

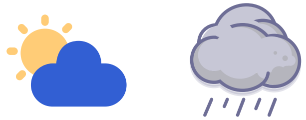
② which features are really significant with multi-features interacting



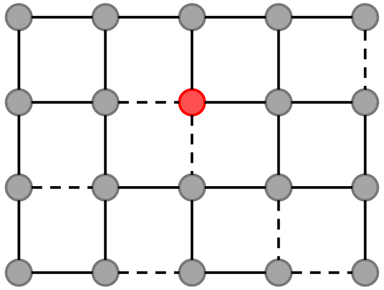
1) sensor observations



2) social connections



3) weather



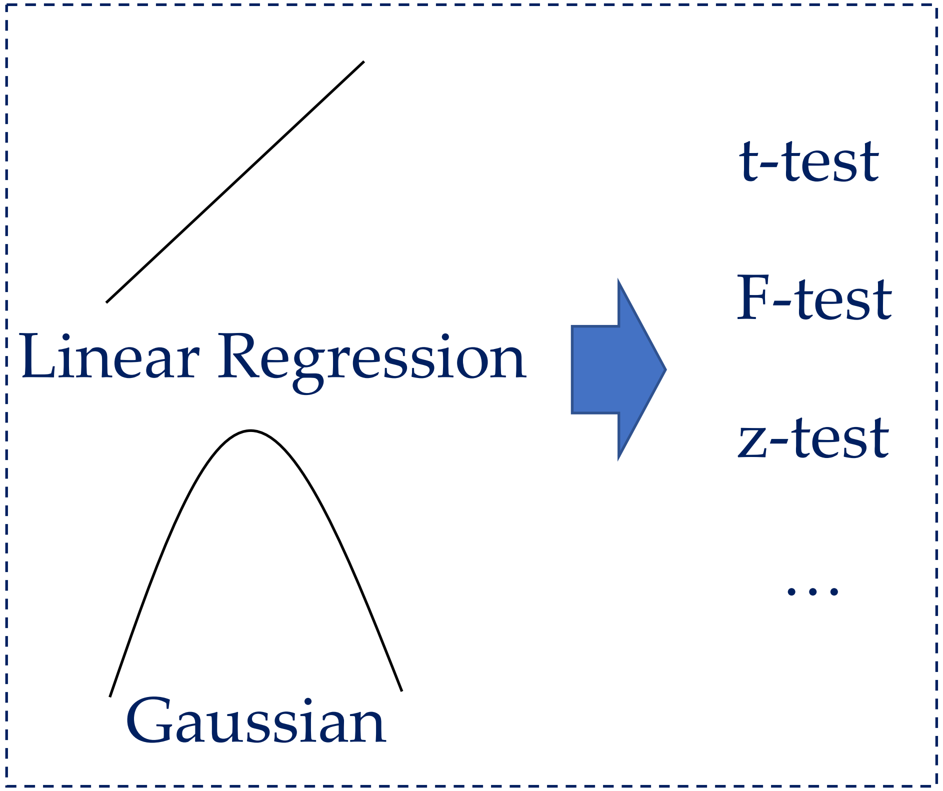
4) road network

"Stuck in a major traffic jam on the 405 freeway. #trafficjam #annoyed"

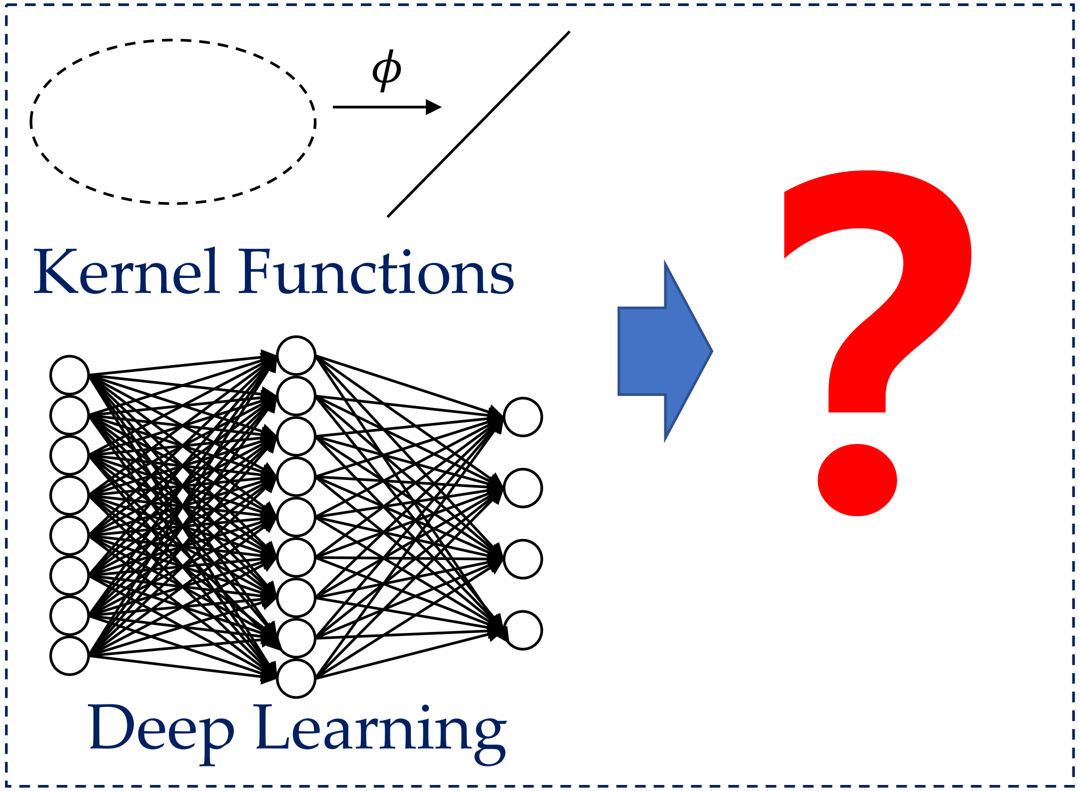
5) media text

Background: Motivation

- Second, significance testing facilitates to identify key traffic factors and eliminate noise. However, it is limited to shallow models (such as ARIMA).
  - Deep models are necessary to guarantee predictive performance due to the high degree of complexity in traffic.



parametric methods



non-parametric methods

Background: Motivation

- Third, most existing approaches only provide deterministic predictions, failing to account for uncertainty.

aleatoric uncertainty

The data acquired from sensors deployed under diverse conditions encompasses heterogeneous noise.

a new sensor with high collection precision

$\underbrace{v_1^{(1)} \quad v_2^{(1)} \quad v_3^{(1)} \quad \cdots\cdots \quad v_i^{(1)} \quad v_{n_1}^{(1)}}$

an old sensor with low collection precision

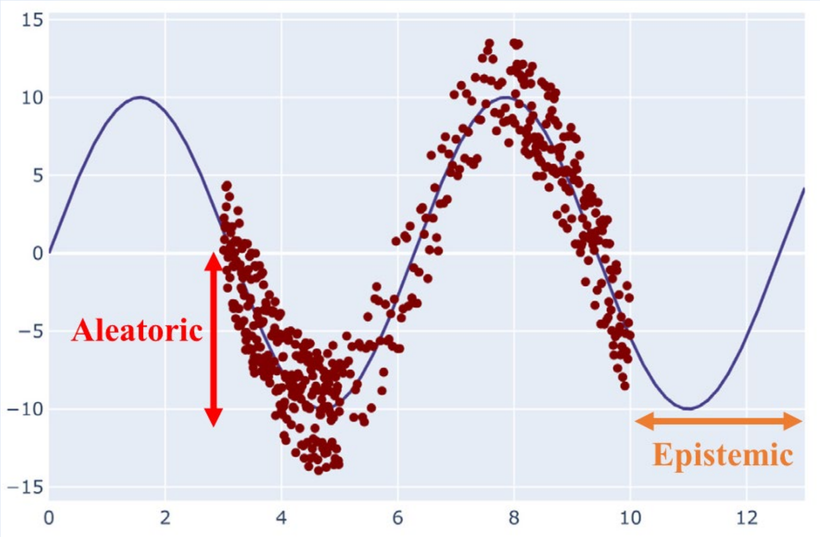
$\underbrace{v_1^{(2)} \quad v_2^{(2)} \quad v_3^{(2)} \quad \cdots\cdots \quad v_i^{(2)} \quad v_{n_2}^{(2)}}$

$v_i^{(j)}$: the observation speeds by the j-th sensor.

epistemic uncertainty

The curve represents the true relationship.

The figure reflects the epistemic uncertainty caused by the lack of data.



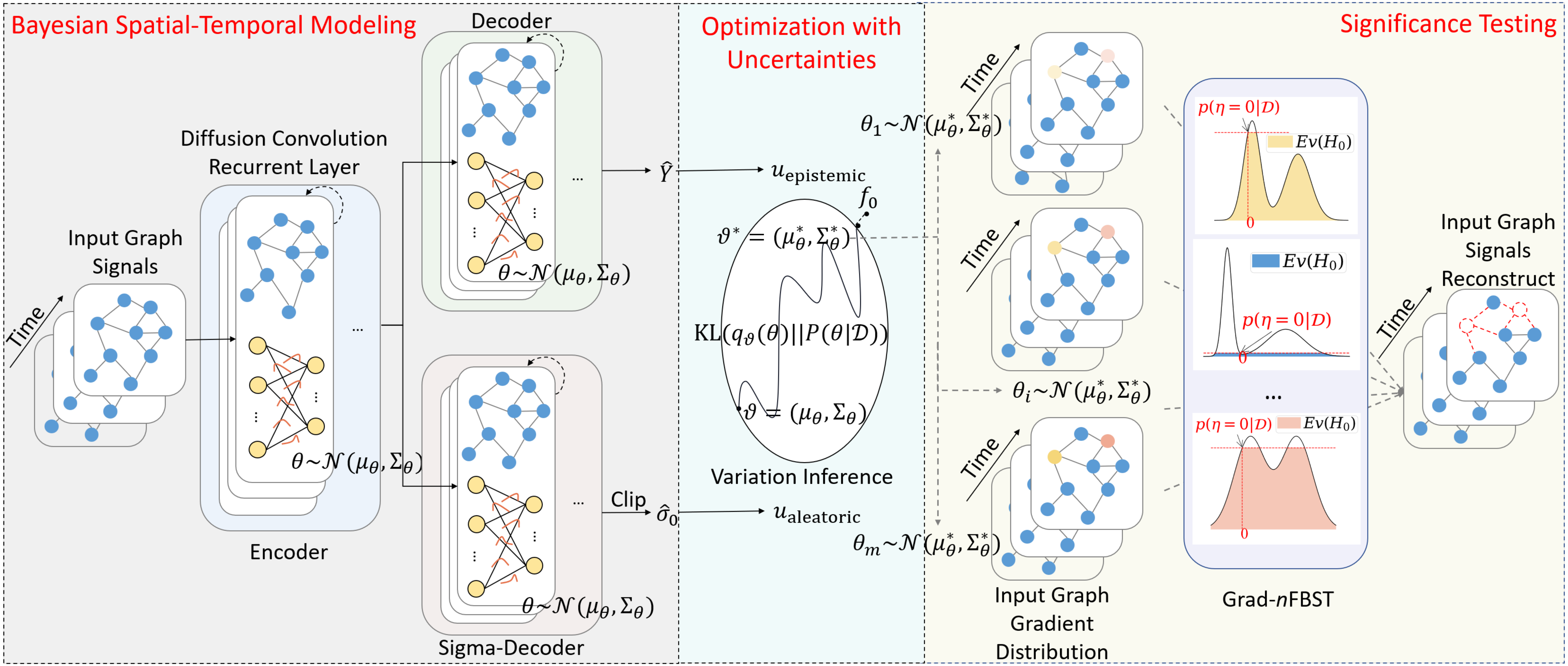
Our Solution

- Input all factors as features, lack specific analysis. → By significance testing to analyze the significance of features.

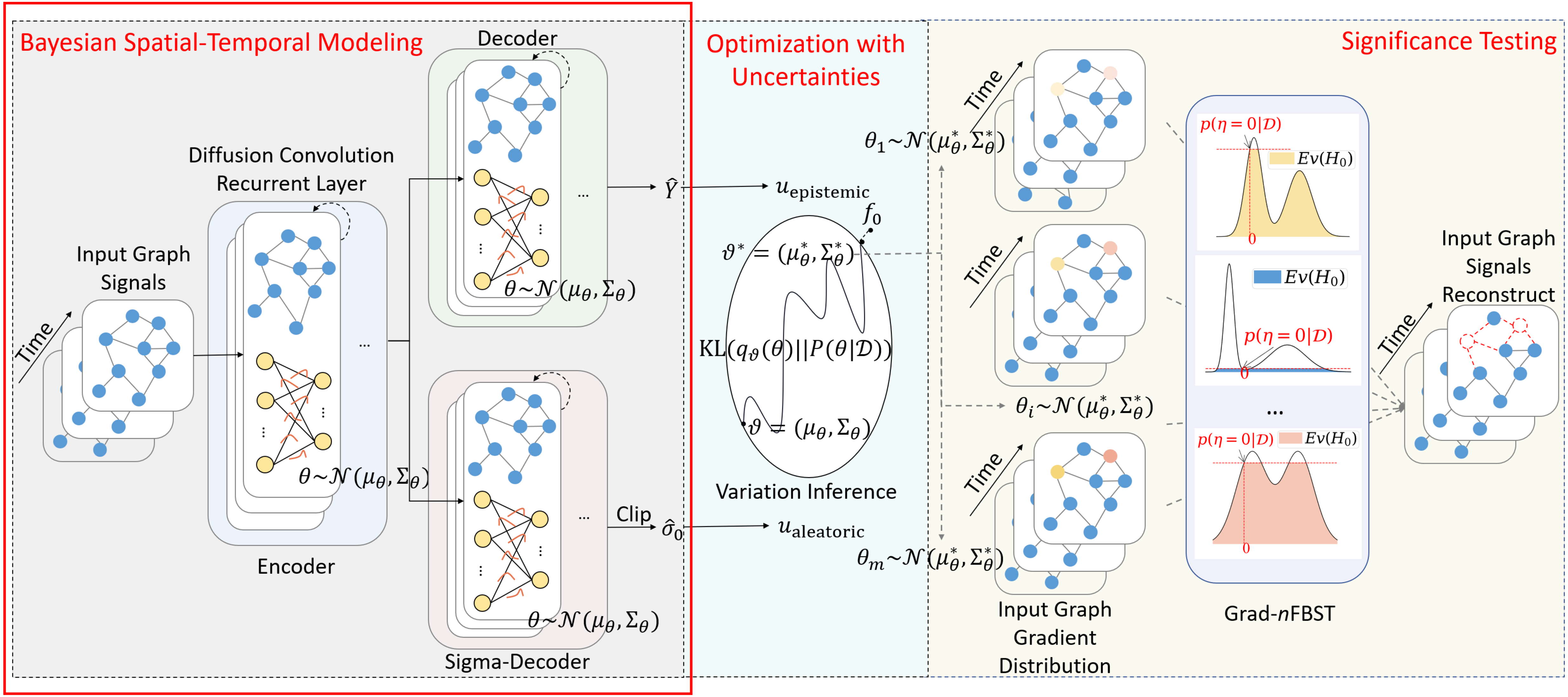
- Significance testing are restricted by model architecture, especially shallow models. → By *n*FBST (Full Bayesian Significance Testing for neural networks) flexibly and powerfully.

- Deterministic prediction ignoring uncertainty. → By combing heteroscedastic aleatoric and epistemic uncertainties into optimization.

ST-nFBST: Overall



ST-nFBST: Overall



ST-nFBST: Bayesian Spatial-Temporal Modeling

- Instead of regarding traffic forecasting as deterministic, we model it from a Bayesian perspective.

- The relation between Y_t and X_t follows a Gaussian distribution. For a pair of observations (X_t, Y_t) at time point t, the regression process can be modeled as corrupted with Gaussian random noise.

    - **Heteroscedastic:** The observation noise $\sigma_0(X_t)$ vary with inputs X_t.

$$Y_t \sim \mathcal{N}(f(X_t), \sigma_0^2(X_t))$$

    - **Outputs:** We build an encoder and two independent decoders to model spatial-temporal dependencies and observation noise respectively.

$$[\hat{Y}_t, \hat{\sigma}_0(X_t)] = f(X_t)$$

ST-nFBST: Bayesian Spatial-Temporal Modeling

- Instead of regarding traffic forecasting as deterministic, we model it from a Bayesian perspective.

- The overall architecture is a Bayesian Neural Network (BNN), whose parameters θ follow a distribution rather than deterministic values.
  - **Prediction:** The weighted average of an ensemble on the whole parameter space

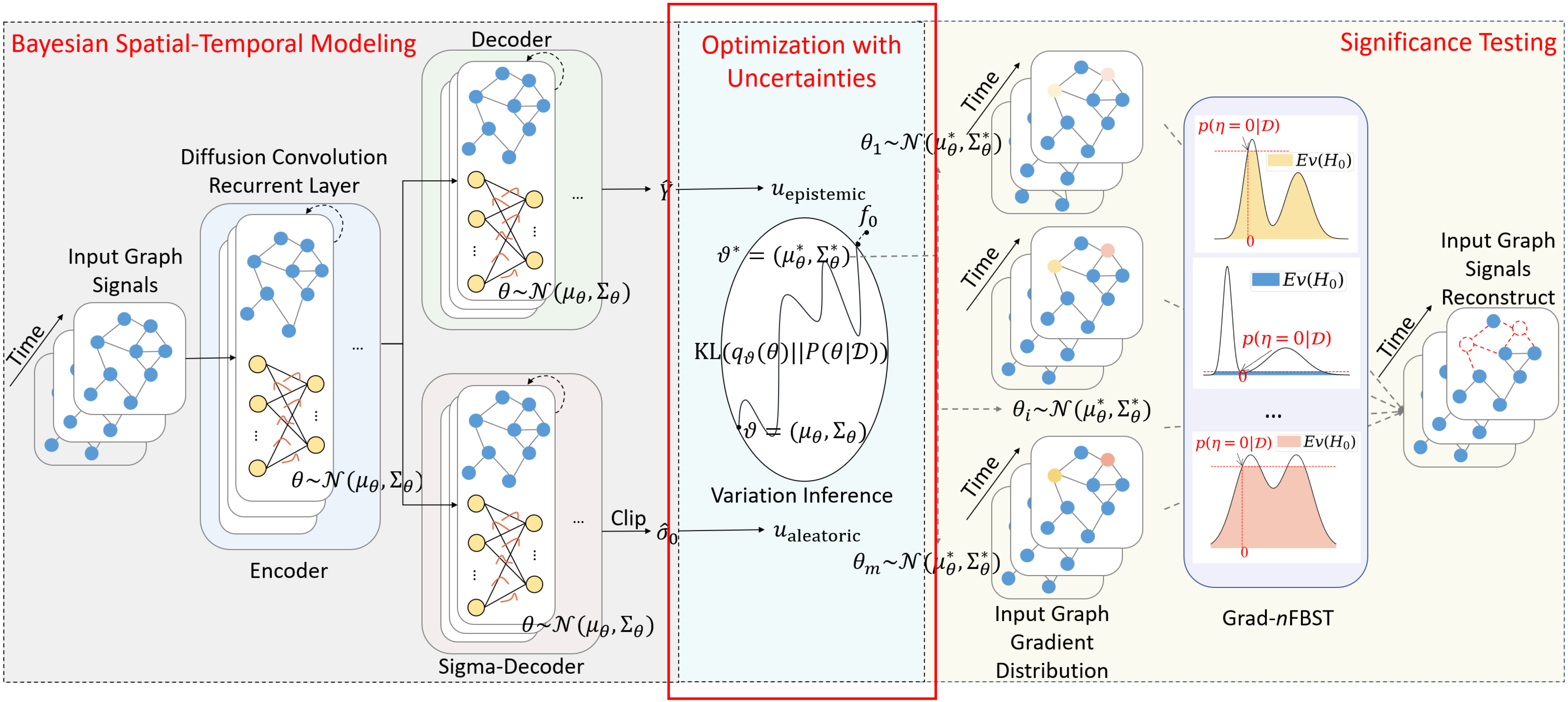
$$P(y|x,\mathcal{D}) = \int_{\Theta} P(y|x,\theta)\boldsymbol{P(\theta|\mathcal{D})}\, d\theta$$

  - **Learning:** To learn the posterior distribution of parameters θ

$$\boldsymbol{P(\theta|\mathcal{D})} = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$$

$P(\mathcal{D}) = \int_{\Theta} P(\mathcal{D}|\theta)\, d\theta$ is intractable!

ST-nFBST: Overall



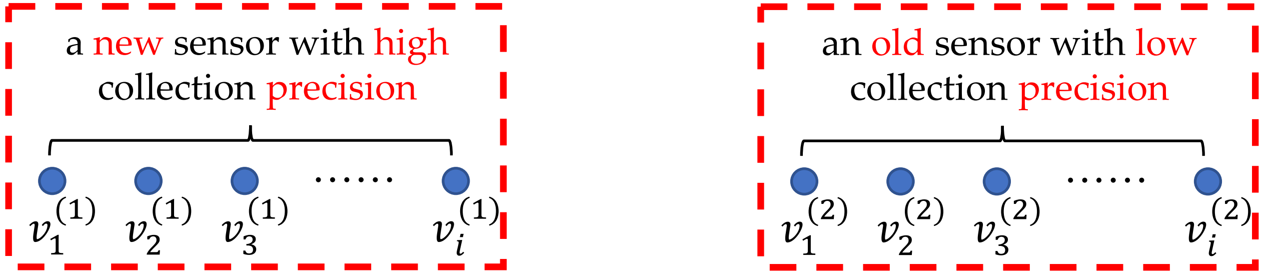
ST-*n*FBST: Optimization with Uncertainties

- In Bayesian modeling, there exist two primary categories of uncertainty, i.e., aleatoric uncertainty and epistemic uncertainty.

- ***Aleatoric uncertainty*** (also known as ***data uncertainty***) is caused by the intrinsic randomness of data, which cannot be reduced even if more data were to be collected. For example, it could be caused by sensor noise inherent in the observations.

$$u_{\text{aleatoric}} = \sigma_0^2(X_t)$$

$v_i^{(j)}$: the observation speeds by the i-th sensor.



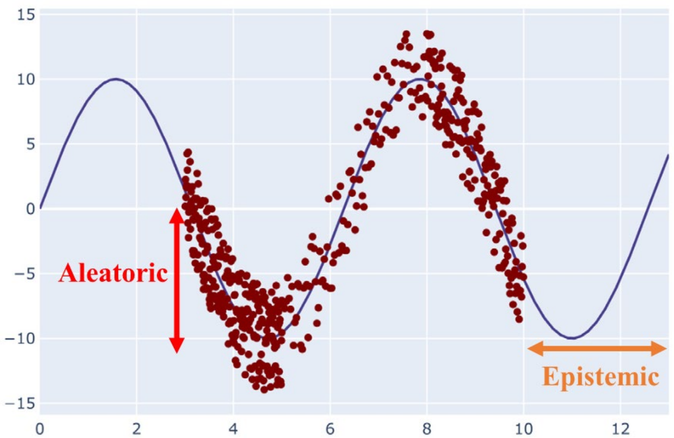
The data acquired from sensors deployed under diverse conditions encompasses heterogeneous noise.

ST-nFBST: Optimization with Uncertainties

- In Bayesian modeling, there exist two primary categories of uncertainty, i.e., aleatoric uncertainty and epistemic uncertainty.

- ***Epistemic uncertainty*** (also known as ***model uncertainty***) accounts for uncertainty in the model parameters, which arises from the lack of data or model misspecification. In traffic forecasting, the epistemic uncertainty can be captured by the predictive variance

$$u_{\text{epistemic}} = \mathbb{E}\left(\hat{Y}_t - \mathbb{E}(\hat{Y}_t)\right)^2$$



The curve represents the true relationship.

The figure reflects the epistemic uncertainty caused by the lack of data.

ST-nFBST: Optimization with Uncertainties

- Variational inference, a popular way to train BNN, entails approximating the real but intractable posterior distribution of parameters θ.

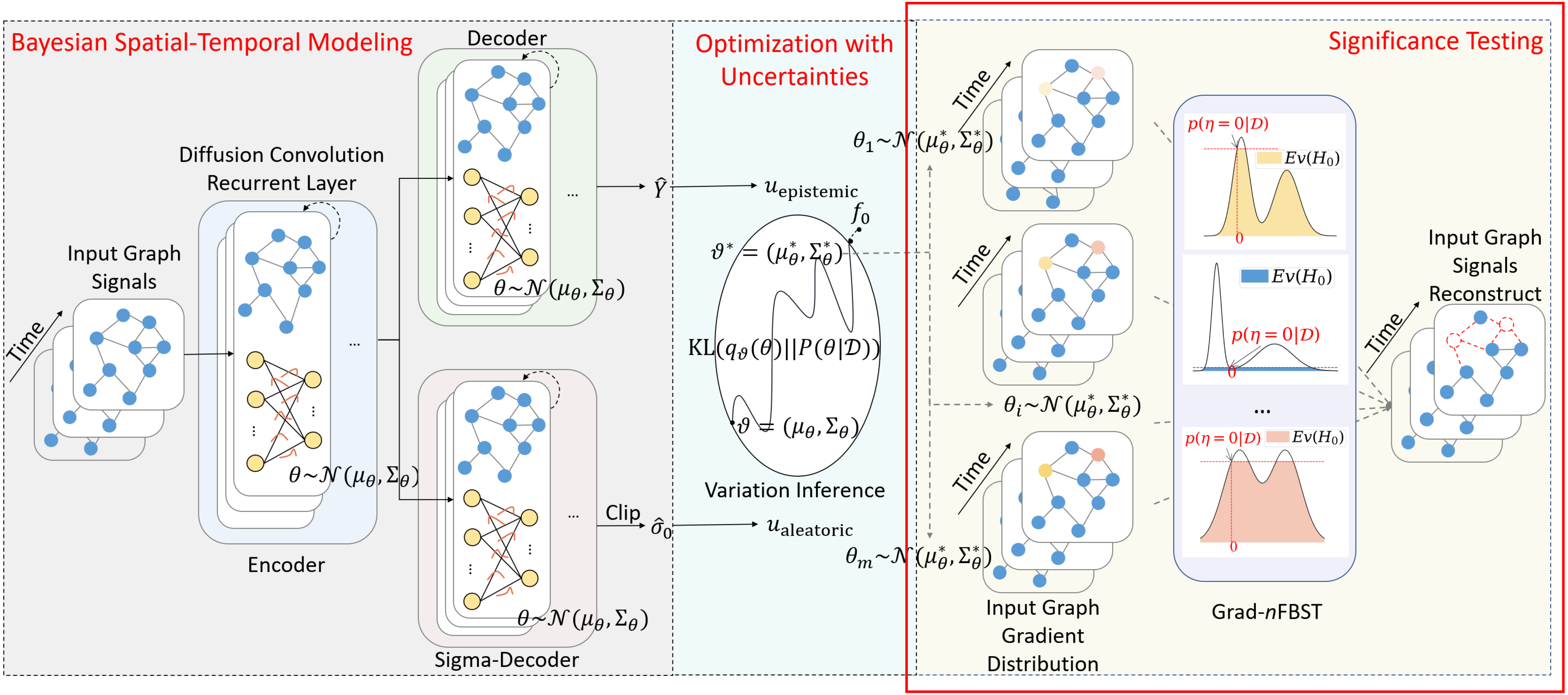
- The optimal variational distribution q_{ϑ^*} is chosen from a predefined family of tractable distributions $Q = \{q_\vartheta : \vartheta \in \Gamma\}$ by minimizing KL divergence

$$\vartheta^* = \arg\min_{\vartheta \in \Gamma} \mathrm{KL}(q_\vartheta(\theta) || P(\theta|\mathcal{D}))$$

- Using Monte Carlo integration, the aleatoric uncertainty and epistemic uncertainty are approximated as

$$u_t \approx \underbrace{\frac{1}{m}\sum_{i=1}^{m} \hat{\sigma}_{0i}^2(X_t)}_{\text{aleatoric}} + \underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(\hat{Y}_{ti} - \frac{1}{m}\sum_{i=1}^{m}\hat{Y}_{ti}\right)^2}_{\text{epistemic}}$$

ST-nFBST: Overall



ST-nFBST: Significance Testing

- We adopt Grad-nFBST (Liu et al. 2024) as the instance-wise testing statistic

$$\eta(X_t) = \frac{\delta f_0(X_t)}{\delta X_t} \in \mathbb{R}^{\tau_2 \times |\mathcal{V}| \times \tau_1 \times |\mathcal{V}|}.$$

$\eta(X_t, i, k_1, j, k_2) \in \mathbb{R}$ represents the testing statistic for predicting the traffic state of the k_1-th sensor at future time $t + i$ based on the traffic state of the k_2-th sensor at historical time $t - j$.

- The testing problem for $\eta(X_t, i, k_1, j, k_2)$ (abbreviated as η) is formulated as

$$H_0 : \eta = 0, \qquad H_1 : \eta \neq 0.$$

It aims to test whether one dimension of the inputs X_t is significant or not.

Note the hypothesis H_0 is **sharp**!

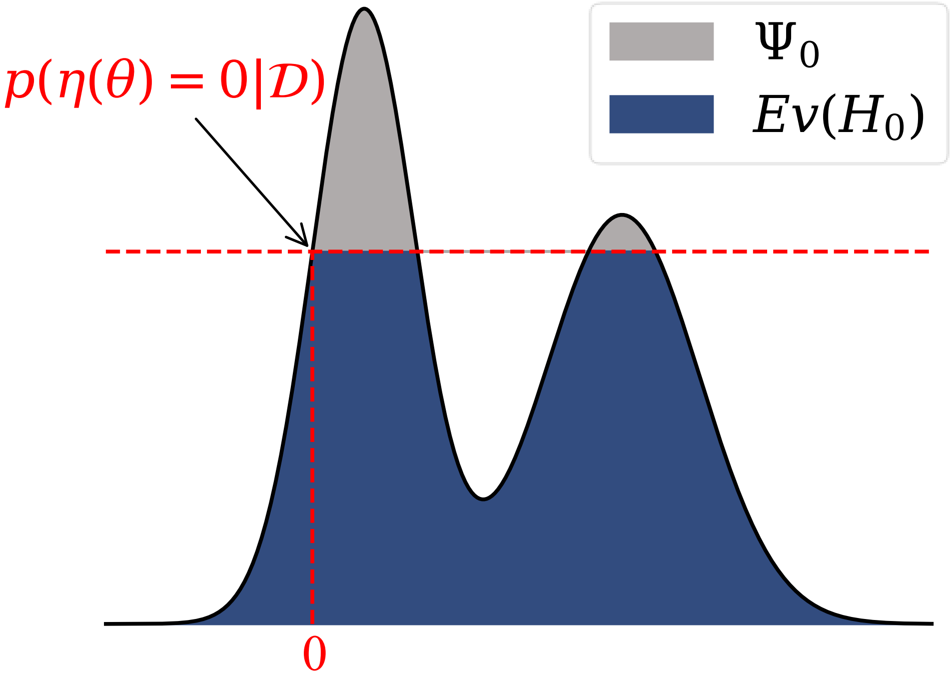
ST-nFBST: Significance Testing

- Under the assumption of H_0, the tangential set is

$$\Psi_0 = \{\eta \in \Psi : p(\eta|\mathcal{D}) > p(\eta = 0|\mathcal{D})\}.$$

- The *Bayesian evidence value supporting* H_0 is

$$Ev(H_0) = 1 - \int_{\Psi_0} p(\eta|\mathcal{D})\, d\eta.$$



Using Monte Carlo method,

$$Ev(H_0) = 1 - \int_{\Psi} \mathbb{I}(\eta \in \Psi) p(\eta|\mathcal{D})\, d\eta \approx 1 - \frac{1}{m}\sum_{i=1}^{m} \mathbb{I}(\eta_i \in \Psi_0)$$

where η_i is sampled m times based on the posterior distribution of η.

Experiment: Performance Comparison

- Performance comparison of our model and baselines on two real-world datasets is shown in the following table, where lower results are better.

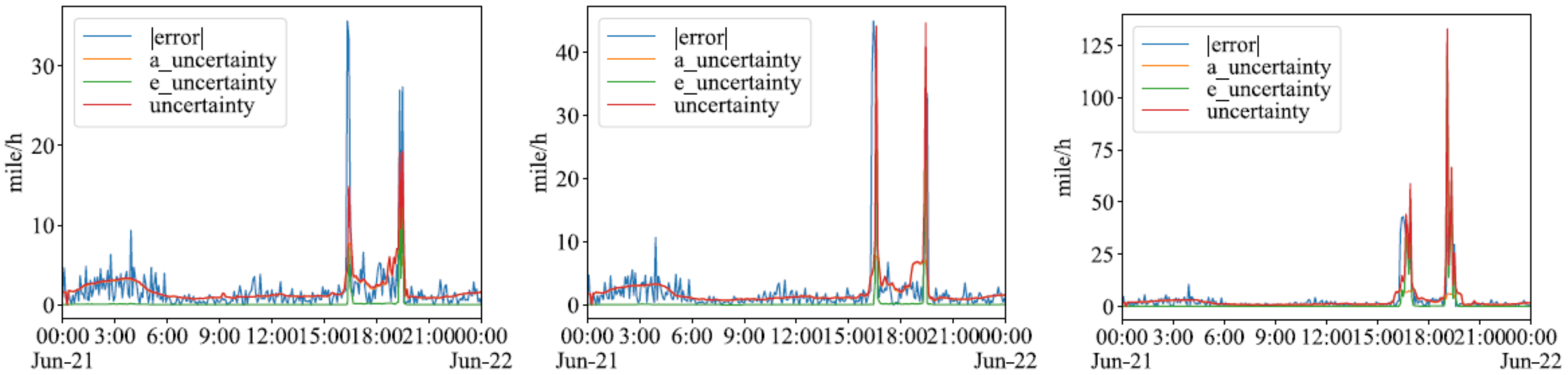
	Horizon	Metric	HA	ARIMA	VAR	SVR	LSTM	DCRNN	STGCN	STTN	AGCRN	CCRNN	DeepSTUQ	ST-*n*FBST
METR-LA	15min	MAE	4.16	3.99	4.42	3.99	3.44	2.77	2.88	2.79	2.86	2.85	2.75	**2.71**
		RMSE	7.80	8.21	7.89	8.45	6.30	5.38	5.74	5.48	5.55	5.54	5.37	**5.27**
		MAPE	13.0%	9.6%	10.2%	9.3%	9.6%	7.3%	7.6%	7.2%	7.6%	7.5%	7.2%	**7.0%**
	30min	MAE	4.16	5.15	5.41	5.05	3.77	3.15	3.47	3.16	3.25	3.24	3.14	**3.09**
		RMSE	7.80	10.45	9.13	10.87	7.23	6.45	7.24	6.50	6.57	6.54	6.38	**6.32**
		MAPE	13.0%	12.7%	12.7%	12.1%	10.9%	8.8%	9.6%	8.5%	9.0%	8.9%	8.7%	**8.5%**
	1hour	MAE	4.16	6.90	6.52	6.72	4.37	3.60	4.59	3.60	3.68	3.73	3.56	**3.52**
		RMSE	7.80	13.23	10.11	13.76	8.69	7.59	9.40	7.60	7.56	7.65	9.40	**7.47**
		MAPE	13.0%	17.4%	15.8%	16.7%	13.2%	10.5%	12.7%	10.2%	10.5%	10.6%	10.6%	**10.2%**
PSMS-BAY	15min	MAE	2.88	1.62	1.74	1.85	2.05	1.38	1.36	1.36	1.36	1.38	1.34	**1.31**
		RMSE	5.59	3.30	3.16	3.59	4.19	2.95	2.96	2.87	2.88	2.90	2.85	**2.77**
		MAPE	6.8%	3.5%	3.6%	3.8%	4.8%	2.9%	2.9%	2.9%	2.9%	3.9%	2.9%	**2.7%**
	30min	MAE	2.88	2.33	2.32	2.48	2.20	1.74	1.81	1.67	1.69	1.74	1.66	**1.64**
		RMSE	5.59	4.76	4.25	5.18	4.55	3.97	4.27	3.79	3.87	3.87	3.78	**3.75**
		MAPE	6.8%	5.4%	5.0%	5.5%	5.2%	3.9%	4.2%	3.8%	3.9%	3.9%	3.8%	**3.7%**
	1hour	MAE	2.88	3.38	2.93	3.28	2.37	2.07	2.49	1.95	1.98	2.07	1.96	**1.95**
		RMSE	5.59	6.50	5.44	7.08	4.96	4.74	5.69	**4.50**	4.59	4.65	4.56	4.54
		MAPE	6.8%	8.3%	6.5%	8.0%	5.7%	4.9%	5.8%	4.6%	4.6%	4.9%	4.6%	**4.6%**

Table 1: Performance comparison of different approaches for traffic speed forecasting. ST-*n*FBST achieves the best performance with almost all three metrics for all forecasting horizons, and the advantage becomes more evident with the increase of the forecasting horizon.

It can be concluded that our proposed ST-*n*FBST nearly outperforms all baselines across all metrics and all forecasting horizons.

Experiment: Uncertainty Analysis

- We further conduct uncertainty analysis to support that our proposed method outperforms deterministic methods.



(a) Predict 15min after for sensor 0 (b) Predict 30min after for sensor 0 (c) Predict 60min after for sensor 0

(d) Predict 15min after for sensor 26 (e) Predict 30min after for sensor 26 (f) Predict 60min after for sensor 26

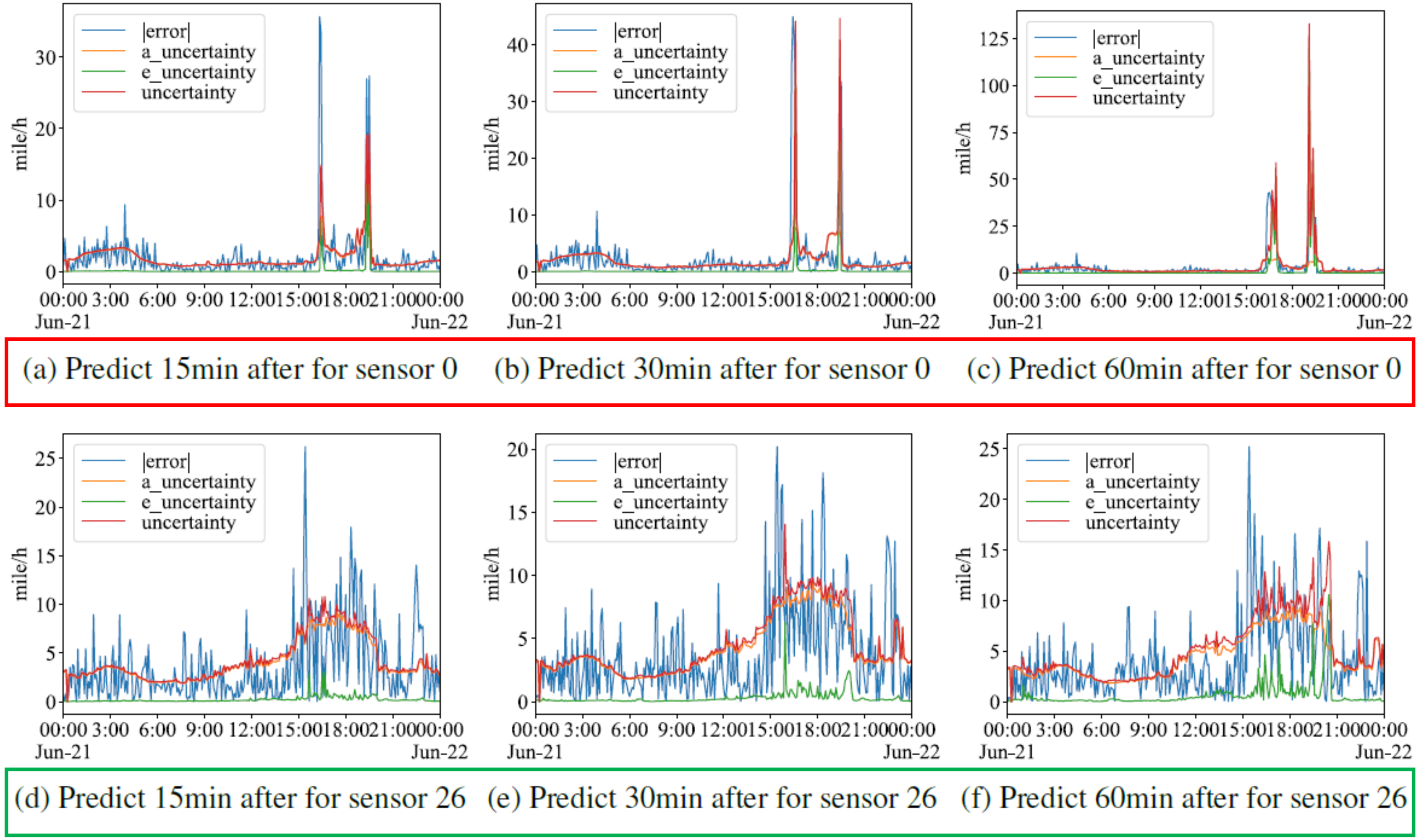
Figure 2: Comparison between uncertainties and |error| for different forecasting horizons on the METR-LA dataset.

1) MAE closely aligns with uncertainty.

2) Aleatoric uncertainty (data-related) accounts for the majority.

Experiment: Uncertainty Analysis

- We further conduct uncertainty analysis to support that our proposed method outperforms deterministic methods.



(a) Predict 15min after for sensor 0 (b) Predict 30min after for sensor 0 (c) Predict 60min after for sensor 0

(d) Predict 15min after for sensor 26 (e) Predict 30min after for sensor 26 (f) Predict 60min after for sensor 26

Figure 2: Comparison between uncertainties and |error| for different forecasting horizons on the METR-LA dataset.

3) As horizon expands, MAE and uncertainty increase for *central sensors*, while remain stable for *marginal sensors*, but predictions fluctuate.

Remark. *Central sensors* are susceptible to other sensors' influence, rendering predictions more intricate. *Marginal sensors* exhibit fewer impacts from other sensors while utilizing less information from others.

Experiment: Significance Testing

- Temporally. We aim to identify the impact of observations at different times ahead in the input sequence on the prediction by removing one's different sizes of its historical inputs.
- For short-term forecasting, the historical observations of the past 15 minutes are the most crucial, with a worsening effect as the removal window increases.

Sensor	Forecasting Horizon	Metric	Initial (-0min)	Reconstruct (-15min)	Reconstruct (-30min)	Reconstruct (-1hour)	Reconstruct (Spatial)
0 (central)	short-term (15min)	MAE	1.97	2.18	2.22	3.30	1.91
		RMSE	4.52	5.04	5.22	5.42	4.22
		MAPE	5.0%	5.3%	5.7%	5.8%	4.8%
		Uncertainty	2.20	2.48	2.42	2.43	2.05
	middle-term (30min)	MAE	2.46	2.42	2.52	2.47	2.17
		RMSE	6.11	5.83	6.28	6.00	5.24
		MAPE	6.7%	6.3%	6.6%	6.4%	5.7%
		Uncertainty	2.81	3.00	2.98	3.28	2.35
	long-term (1hour)	MAE	2.97	2.87	2.86	2.94	2.47
		RMSE	7.51	7.02	7.17	7.39	6.13
		MAPE	8.8%	8.1%	8.2%	7.6%	6.9%
		Uncertainty	3.67	4.55	4.39	4.11	2.88
181 (marginal)	short-term (15min)	MAE	2.68	3.32	3.55	3.63	1.70
		RMSE	5.57	7.02	7.49	7.46	2.60
		MAPE	6.1%	8.4%	9.0%	9.3%	2.9%
		Uncertainty	3.00	3.52	3.81	3.58	1.68
	middle-term (30min)	MAE	3.30	3.56	3.71	3.66	1.71
		RMSE	6.85	7.59	8.08	7.50	2.84
		MAPE	7.7%	9.1%	9.7%	9.2%	3.0%
		Uncertainty	3.70	3.92	4.51	4.17	1.69
	long-term (1hour)	MAE	3.77	3.76	3.78	3.77	1.71
		RMSE	7.92	8.03	8.26	7.74	2.85
		MAPE	9.0%	9.7%	9.9%	9.0%	3.1%
		Uncertainty	4.88	4.52	5.11	4.29	1.71

Table 2: Comparison of performance and uncertainty before and after reconstructing input graph signals temporally or spatially.

This indicates that short-term forecasting highly depends on historical data of its own.

Experiment: Significance Testing

- Temporally. We aim to identify the impact of observations at different times ahead in the input sequence on the prediction by removing one's different sizes of its historical inputs.
- For middle-term forecasting, the performance of marginal sensors decreases, while the performance of central sensors remains unchanged or decreases within an acceptable range.

Sensor	Forecasting Horizon	Metric	Initial (-0min)	Reconstruct (-15min)	Reconstruct (-30min)	Reconstruct (-1hour)	Reconstruct (Spatial)
0 (central)	short-term (15min)	MAE	1.97	2.18	2.22	3.30	1.91
		RMSE	4.52	5.04	5.22	5.42	4.22
		MAPE	5.0%	5.3%	5.7%	5.8%	4.8%
		Uncertainty	2.20	2.48	2.42	2.43	2.05
	middle-term (30min)	MAE	2.46	2.42	2.52	2.47	2.17
		RMSE	6.11	5.83	6.28	6.00	5.24
		MAPE	6.7%	6.3%	6.6%	6.4%	5.7%
		Uncertainty	2.81	3.00	2.98	3.28	2.35
	long-term (1hour)	MAE	2.97	2.87	2.86	2.94	2.47
		RMSE	7.51	7.02	7.17	7.39	6.13
		MAPE	8.8%	8.1%	8.2%	7.6%	6.9%
		Uncertainty	3.67	4.55	4.39	4.11	2.88
181 (marginal)	short-term (15min)	MAE	2.68	3.32	3.55	3.63	1.70
		RMSE	5.57	7.02	7.49	7.46	2.60
		MAPE	6.1%	8.4%	9.0%	9.3%	2.9%
		Uncertainty	3.00	3.52	3.81	3.58	1.68
	middle-term (30min)	MAE	3.30	3.56	3.71	3.66	1.71
		RMSE	6.85	7.59	8.08	7.50	2.84
		MAPE	7.7%	9.1%	9.7%	9.2%	3.0%
		Uncertainty	3.70	3.92	4.51	4.17	1.69
	long-term (1hour)	MAE	3.77	3.76	3.78	3.77	1.71
		RMSE	7.92	8.03	8.26	7.74	2.85
		MAPE	9.0%	9.7%	9.9%	9.0%	3.1%
		Uncertainty	4.88	4.52	5.11	4.29	1.71

Table 2: Comparison of performance and uncertainty before and after reconstructing input graph signals temporally or spatially.

This indicates that middle-term forecasting integrates information from both its own and surrounding sensors.

When excluding its own historical data, the central sensor can utilize more information from other sensors.

Experiment: Significance Testing

- Temporally. We aim to identify the impact of observations at different times ahead in the input sequence on the prediction by removing one's different sizes of its historical inputs.
- For long-term forecasting, removing a sensor's own historical data demonstrates negligible even a slight improvement.

Sensor	Forecasting Horizon	Metric	Initial (-0min)	Reconstruct (-15min)	Reconstruct (-30min)	Reconstruct (-1hour)	Reconstruct (Spatial)
0 (central)	short-term (15min)	MAE	1.97	2.18	2.22	3.30	1.91
		RMSE	4.52	5.04	5.22	5.42	4.22
		MAPE	5.0%	5.3%	5.7%	5.8%	4.8%
		Uncertainty	2.20	2.48	2.42	2.43	2.05
	middle-term (30min)	MAE	2.46	2.42	2.52	2.47	2.17
		RMSE	6.11	5.83	6.28	6.00	5.24
		MAPE	6.7%	6.3%	6.6%	6.4%	5.7%
		Uncertainty	2.81	3.00	2.98	3.28	2.35
	long-term (1hour)	MAE	2.97	2.87	2.86	2.94	2.47
		RMSE	7.51	7.02	7.17	7.39	6.13
		MAPE	8.8%	8.1%	8.2%	7.6%	6.9%
		Uncertainty	3.67	4.55	4.39	4.11	2.88
181 (marginal)	short-term (15min)	MAE	2.68	3.32	3.55	3.63	1.70
		RMSE	5.57	7.02	7.49	7.46	2.60
		MAPE	6.1%	8.4%	9.0%	9.3%	2.9%
		Uncertainty	3.00	3.52	3.81	3.58	1.68
	middle-term (30min)	MAE	3.30	3.56	3.71	3.66	1.71
		RMSE	6.85	7.59	8.08	7.50	2.84
		MAPE	7.7%	9.1%	9.7%	9.2%	3.0%
		Uncertainty	3.70	3.92	4.51	4.17	1.69
	long-term (1hour)	MAE	3.77	3.76	3.78	3.77	1.71
		RMSE	7.92	8.03	8.26	7.74	2.85
		MAPE	9.0%	9.7%	9.9%	9.0%	3.1%
		Uncertainty	4.88	4.52	5.11	4.29	1.71

Table 2: Comparison of performance and uncertainty before and after reconstructing input graph signals temporally or spatially.

This indicates that long-term forecasting does not focus on the historical data of itself but instead utilizes information about the surroundings.

Experiment: Significance Testing

- Spatially. We aim to identify the impact of sensors at different locations on the prediction by selectively removing the insignificant sensors to reconstruct the input graph signals.

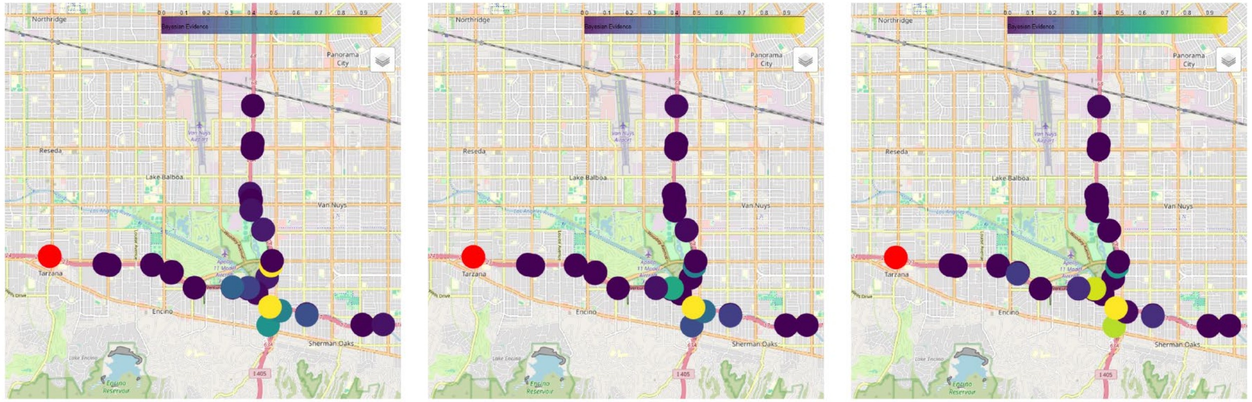
Sensor	Forecasting Horizon	Metric	Initial (-0min)	Reconstruct (-15min)	Reconstruct (-30min)	Reconstruct (-1hour)	Reconstruct (Spatial)
0 (central)	short-term (15min)	MAE	1.97	2.18	2.22	3.30	1.91
		RMSE	4.52	5.04	5.22	5.42	4.22
		MAPE	5.0%	5.3%	5.7%	5.8%	4.8%
		Uncertainty	2.20	2.48	2.42	2.43	2.05
	middle-term (30min)	MAE	2.46	2.42	2.52	2.47	2.17
		RMSE	6.11	5.83	6.28	6.00	5.24
		MAPE	6.7%	6.3%	6.6%	6.4%	5.7%
		Uncertainty	2.81	3.00	2.98	3.28	2.35
	long-term (1hour)	MAE	2.97	2.87	2.86	2.94	2.47
		RMSE	7.51	7.02	7.17	7.39	6.13
		MAPE	8.8%	8.1%	8.2%	7.6%	6.9%
		Uncertainty	3.67	4.55	4.39	4.11	2.88
181 (marginal)	short-term (15min)	MAE	2.68	3.32	3.55	3.63	1.70
		RMSE	5.57	7.02	7.49	7.46	2.60
		MAPE	6.1%	8.4%	9.0%	9.3%	2.9%
		Uncertainty	3.00	3.52	3.81	3.58	1.68
	middle-term (30min)	MAE	3.30	3.56	3.71	3.66	1.71
		RMSE	6.85	7.59	8.08	7.50	2.84
		MAPE	7.7%	9.1%	9.7%	9.2%	3.0%
		Uncertainty	3.70	3.92	4.51	4.17	1.69
	long-term (1hour)	MAE	3.77	3.76	3.78	3.77	1.71
		RMSE	7.92	8.03	8.26	7.74	2.85
		MAPE	9.0%	9.7%	9.9%	9.0%	3.1%
		Uncertainty	4.88	4.52	5.11	4.29	1.71

Table 2: Comparison of performance and uncertainty before and after reconstructing input graph signals temporally or spatially.

It can be concluded that the model's predictive performance and uncertainty are significantly improved after retraining on the reconstructed graph signals.

Experiment: Case Study

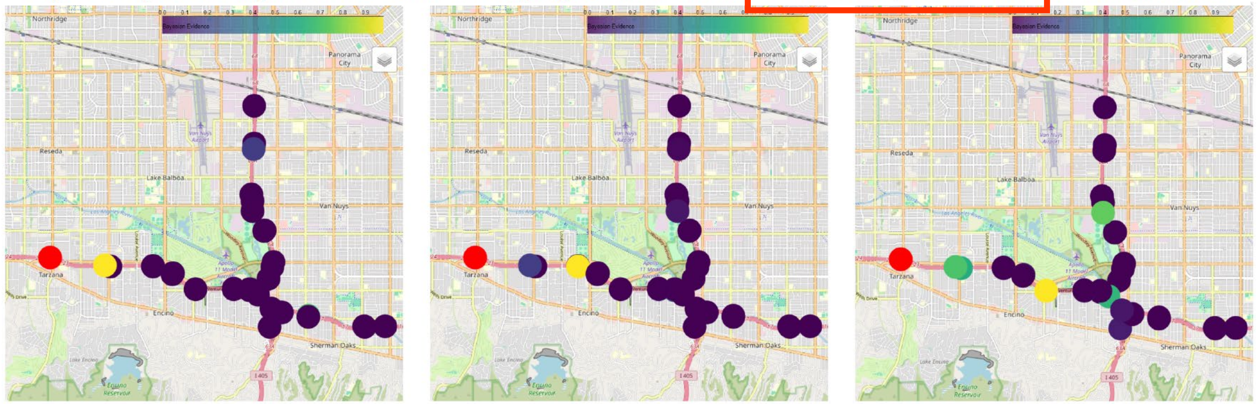
- Short-term forecasting is more concerned with the road conditions where their own sensors are located, but do not pay attention to the situation at intersections.



(a) (b) (c)

Figure 2: Spatial testing results for sensor 181 (highlighted in red) under low uncertainty and low error at different times.

Correct! surrounding sensors are significant, but sensors near intersections are insignificant.



(a) (b) (c)

Figure 3: Spatial testing results for sensor 181 (highlighted in red) under high uncertainty and high error at different times.

Wrong! surrounding sensors are insignificant, but sensors near intersections are significant.

Conclusion

- We are the first to introduce significance testing for neural networks into traffic forecasting. By modeling spatial-temporal correlations from a Bayesian perspective, we extend significance testing to accommodate GCN-based models.
- We propose a loss function in the context of epistemic uncertainty and aleatoric uncertainty to handle heteroscedastic noise in traffic data, which is more aligned with reality. The performance is improved. Moreover, a quantitative uncertainty analysis is provided, further facilitating the interpretation of the testing results.
- We propose an input reconstruction approach guided by significance testing, which improves predictive performance and reduces uncertainty by removing insignificant sensors during the reconstruction of the input graph.

Thanks for Listening!

Group Homepage: http://www.bigscity.com

Code Link: https://github.com/liuzh-buaa/ST-nFBST

Email: jywang@buaa.edu.cn liuzehua@buaa.edu.cn