Meta Guidance: Incorporating Inductive Biases into Deep Time Series Imputers

Jiacheng You, Xinyang Chen✉, Yu Sun†, Weili Guan, Liqiang Nie

Harbin Institute of Technology, Shenzhen, †Nankai University

NEURAL INFORMATION PROCESSING SYSTEMS

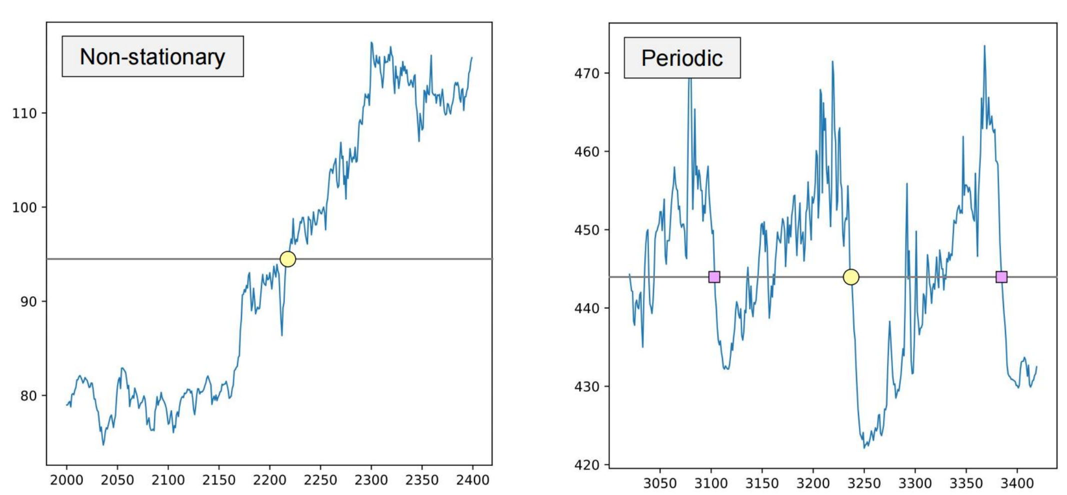
Summary

- Meta Guidance (MG) is a lightweight, plug-and-play module that can be integrated into most deep imputation networks, consistently enhancing performance across diverse architectures.
- MG introduces two inductive biases — Non-Stationary Guidance and Periodic Guidance — to explicitly encode non-stationarity and periodic dependencies in time series imputation.

Background

- Time series data suffer from missing valuesdue to sensor faults, communication interruptions, or irregular sampling, which destroy temporal dependencies and deteriorate downstream model.
- The non-stationarity and periodic of time series are critical characteristics that can substantially impact the performance of imputation models.
- Current approaches always adopt end-to-end learning to implicitly infer temporal patterns, largely overlooking the strategic integration of domain-specific inductive biases known to enhance imputation.

Motivation

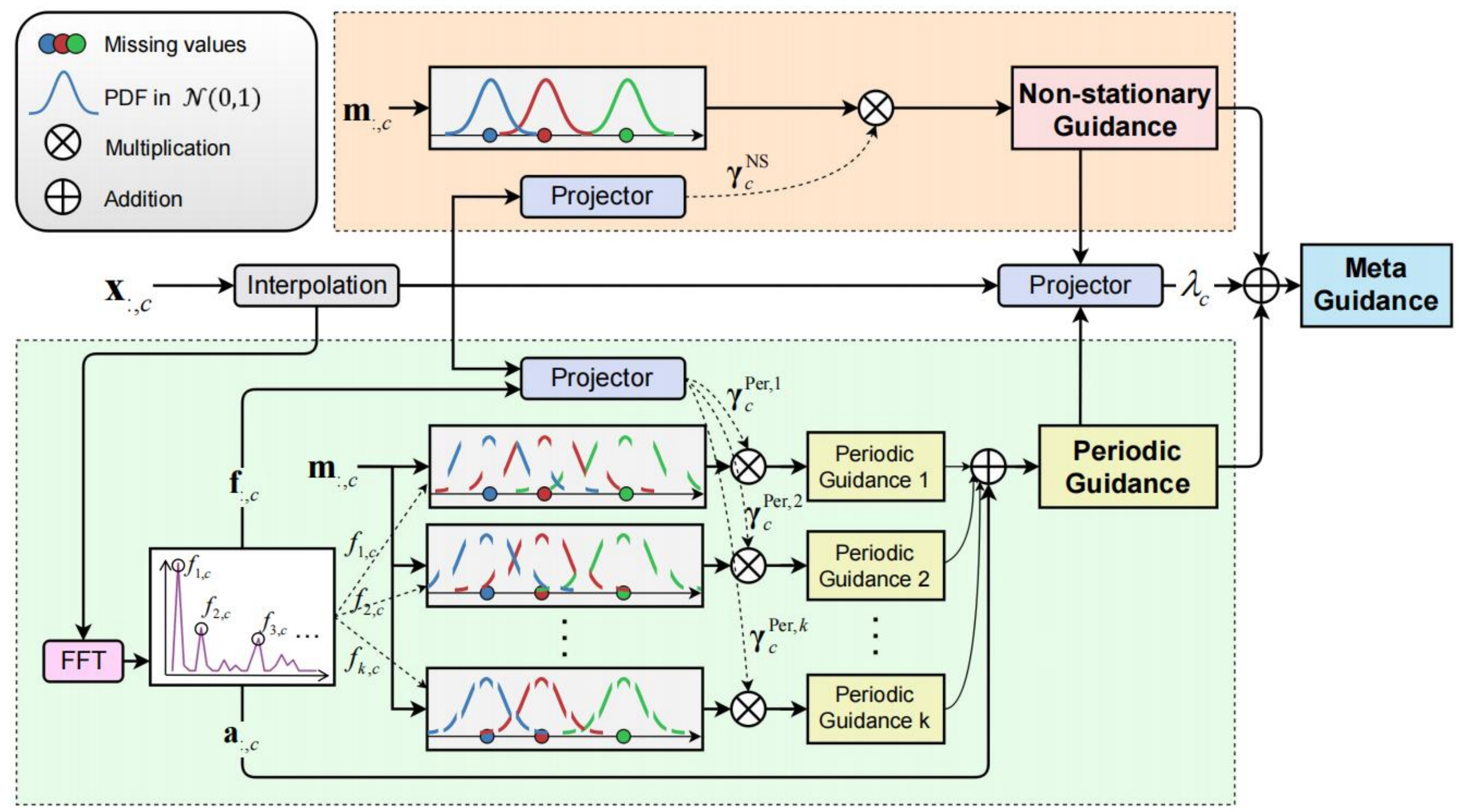


- Missing values in non-stationary time series should be inferred mainly from temporally local contexts.
- Periodic dependencies imply that observations at regular intervals are more likely to share similar patterns.

Contributions

- We introduce two inductive biases for time series imputation — Non-Stationary Guidance (NSG) and Periodic Guidance (PG) — which explicitly encode local continuity and periodic dependencies.
- We design Meta Guidance (MG), a lightweight, model-agnostic module that adaptively fuses NSG and PG according to input characteristics, enabling dynamic adaptation to diverse temporal dynamics.
- MG achieves up to 27.39% error reduction across nine real-world datasets.

Method



Three steps to learn Meta Guidance:

1. NSG assigns stronger weights to nearby timestamps following the proximity principle, where timestamps closer to the target moment are prioritized for weight enhancement:

$$g_{t,c}^{\mathrm{NS}} = \sum_{i=-r}^{i=r} \psi(i) \cdot \gamma_c^{\mathrm{NS}} \cdot \mathbb{1}(1 \le i+t \le T) \cdot \mathbb{1}(m_{i+t,c} = 0)$$

2. PG assigns higher weights to timestamps at regular intervals that are identified through FFT analysis, leveraging frequency-domain characteristics to capture periodic patterns:

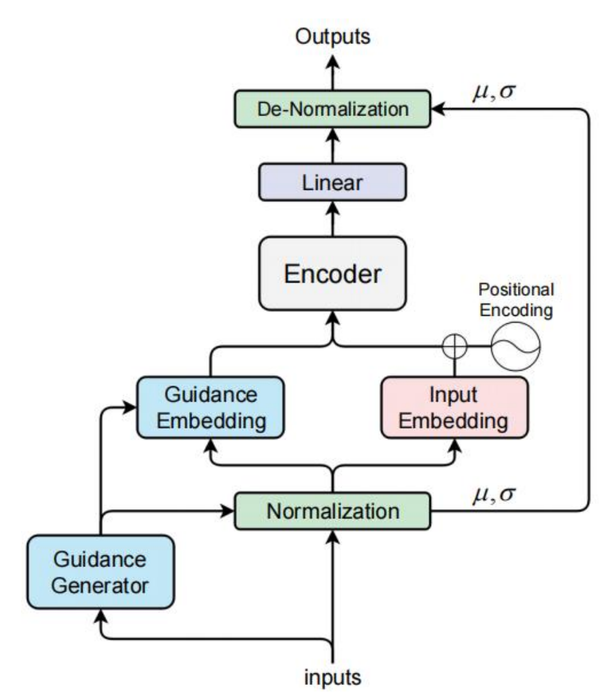
$$g_{t,c}^{\mathrm{Per},l} = \sum_{i=-r}^{i=r} \psi(i) \cdot \gamma_c^{\mathrm{Per},l} \cdot \mathbb{1}(1 \le i \cdot p_{l,c} + t \le T) \quad \cdot \mathbb{1}(m_{i \cdot p_{l,c}+t,c} = 0)$$

3. Learning to weight two guidances (i.e., NSG and PG), enabling dynamic adjustment of their contributions based on data characteristics.:

$$\mathbf{g}_{:,c}^{\mathrm{Meta}} = \lambda_c \cdot \mathbf{g}_{:,c}^{\mathrm{NS}} + (1-\lambda_c) \cdot \mathbf{g}_{:,c}^{\mathrm{Per}}$$

Three steps to incorporate MG into Transformer:

1. Add MG-based Normalization and De-Normalization layers prior to the input layer of the model and subsequent to its output layer, respectively.
2. Multiply the MG sequence by the raw input data element-wise and subsequently feed the resulting product into an Embedding layer for feature mapping.
3. Concatenate the Embedded MG with the original input data along the feature dimension to form the new input data for subsequent network modules.



Experiments

- We evaluate our method on six widely used time series datasets — HD, Electricity, Traffic, Weather, ETT, and TCPC, covering diverse application domains such as healthcare, energy, and transportation.
- We integrate Meta Guidance (MG) into five strong deep imputation baselines: Transformer, CSDI, TimesNet, SAITS, and iTransformer.
- Across all datasets, these models show a consistent performance boost after incorporating MG, achieving an average error reduction of 27.39%.

dataset		HD		Weather		Electricity		Traffic		ETTm1		TCPC	
metric		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Transformer	10%	0.112	0.188	0.088	0.182	0.276	0.379	0.233	0.441	0.177	0.257	0.177	0.248
	25%	0.119	0.234	0.095	0.196	0.284	0.394	0.231	0.453	0.213	0.308	0.233	0.328
	40%	0.134	0.247	0.101	0.204	0.291	0.407	0.233	0.466	0.238	0.344	0.276	0.382
Transformer +MG	10%	**0.046**	**0.082**	**0.049**	**0.159**	**0.180**	**0.261**	**0.198**	**0.388**	**0.108**	**0.169**	**0.069**	**0.118**
	25%	**0.050**	**0.095**	**0.052**	**0.168**	**0.191**	**0.276**	**0.203**	**0.397**	**0.134**	**0.207**	**0.086**	**0.145**
	40%	**0.057**	**0.105**	**0.058**	**0.176**	**0.200**	**0.289**	**0.217**	**0.422**	**0.153**	**0.235**	**0.103**	**0.169**
⋮													
iTransformer	10%	0.107	0.153	0.081	0.193	0.184	0.262	0.198	0.390	0.155	0.236	0.087	0.136
	25%	0.114	0.190	0.103	0.214	0.213	0.298	0.230	0.433	0.181	0.269	0.111	0.164
	40%	0.131	0.226	0.122	0.234	0.237	0.329	0.257	0.474	0.207	0.303	0.122	0.181
iTransformer +MG	10%	**0.075**	**0.109**	**0.055**	**0.176**	**0.166**	**0.241**	**0.185**	**0.375**	**0.128**	**0.205**	**0.052**	**0.103**
	25%	**0.071**	**0.110**	**0.055**	**0.185**	**0.186**	**0.267**	**0.208**	**0.411**	**0.142**	**0.226**	**0.064**	**0.119**
	40%	**0.073**	**0.118**	**0.060**	**0.192**	**0.205**	**0.293**	**0.231**	**0.448**	**0.156**	**0.244**	**0.076**	**0.138**
Promotion		**↑43.76%**		**↑26.49%**		**↑17.91%**		**↑9.88%**		**↑25.82%**		**↑34.61%**	

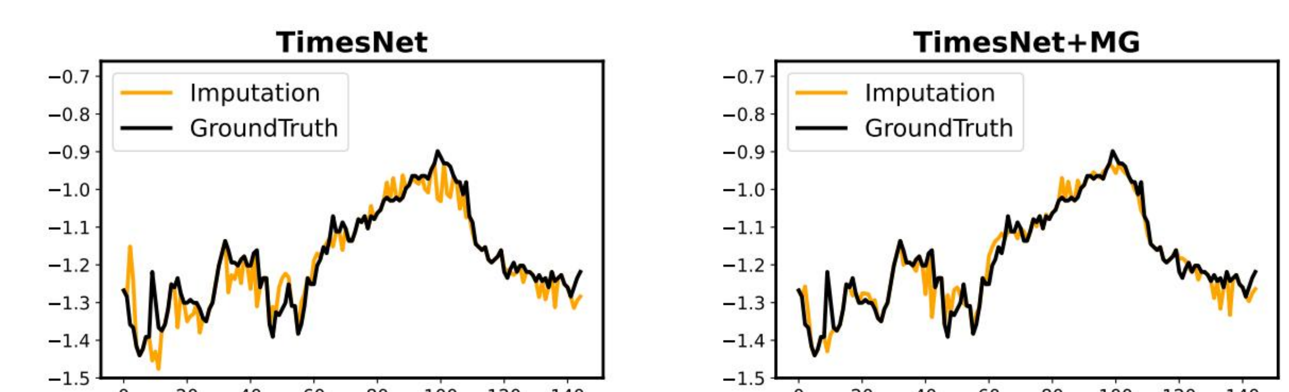
Analysis

The two guidance modules are both effective and complementary, jointly leading to optimal performance.

				HD			Electricity			Weather			TCPC		
$\mathbf{G}^{\mathrm{NS}}$	$\mathbf{G}^{\mathrm{Per}}$	$\mathbf{G}^{\mathrm{Meta}}$		10%	25%	40%	10%	25%	40%	10%	25%	40%	10%	25%	40%
-	-	-	MAE	0.112	0.119	0.134	0.276	0.284	0.291	0.088	0.095	0.101	0.177	0.233	0.276
			RMSE	0.188	0.234	0.247	0.379	0.394	0.407	0.182	0.196	0.204	0.248	0.328	0.382
✓	-	-	MAE	0.047	0.052	0.058	0.182	0.194	0.202	0.053	0.056	0.064	0.076	0.112	0.115
			RMSE	0.083	0.096	0.107	0.262	0.279	0.292	0.161	0.172	0.182	0.131	0.176	0.178
-	✓	-	MAE	0.051	0.056	0.062	0.186	0.208	0.208	0.051	0.055	0.061	0.072	0.093	0.104
			RMSE	0.088	0.099	0.112	0.267	0.288	0.298	0.160	0.174	0.179	0.127	0.159	0.175
✓	✓	✓	MAE	**0.046**	**0.050**	**0.057**	**0.180**	**0.191**	**0.200**	**0.049**	**0.052**	**0.058**	**0.069**	**0.086**	**0.103**
			RMSE	**0.082**	**0.095**	**0.105**	**0.261**	**0.276**	**0.289**	**0.159**	**0.168**	**0.176**	**0.118**	**0.148**	**0.169**

Showcase

The example below shows that incorporating Meta Guidance significantly improves the imputation performance of TimesNet.



School of Computer Science and Technology - Harbin Institute of Technology, Shenzhen - China

Mail: jiachengyou2hit@gmail.com