EasyST: Modeling Spatial-Temporal Correlations and Uncertainty for Dynamic Wind Power Forecasting via PaddlePaddle

Yiji Zhao[1,2], Haomin Wen[1,2,3], Junhong Lou[3], Jinji Fu[1,2], Jianbin Zheng[3], Youfang Lin[1,2,*]

[1]School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China

[2]Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing, China. *Team advisor

[3]Artificial Intelligence Department, Cainiao Network, Hangzhou, China.

Content

Background

Problem Introduction

Wind Power Forecasting (WPF)

- Predict future wind power based on historical data
- Complex Spatial-Temporal Correlations
- Huge data uncertainties

Dataset

SDWPF from Longyuan Power Group Corp

- Spatial distribution: relative location of all wind turbines
- Dynamic context: weather situations and turbine internal contexts

Metrics

$$Score = \frac{1}{2}(MAE(\hat{y}, y) + RMSE(\hat{y}, y))$$

prediction length：288

Content

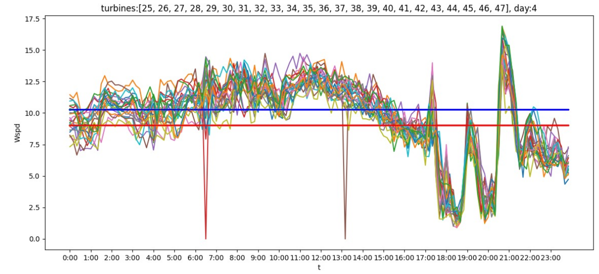
- Background
- Challenge
- Solution
- Experiment

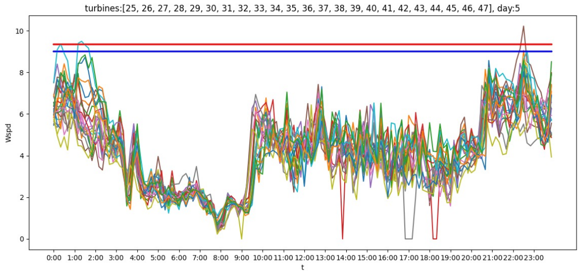
Challenge

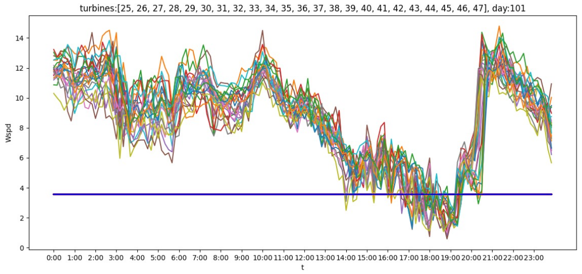
■ Complex spatial-temporal correlations

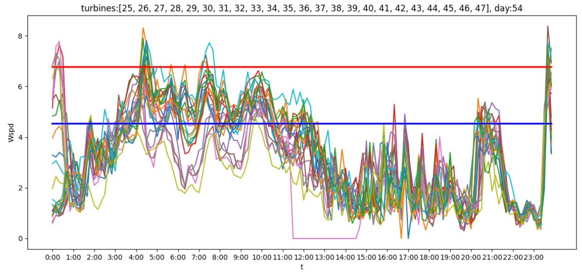
- Spatial correlation
- Temporal correlation

■ Huge data uncertainties









Content

- Background
- Challenge
- Solution
- Experiment

Spatial and Temporal Modeling

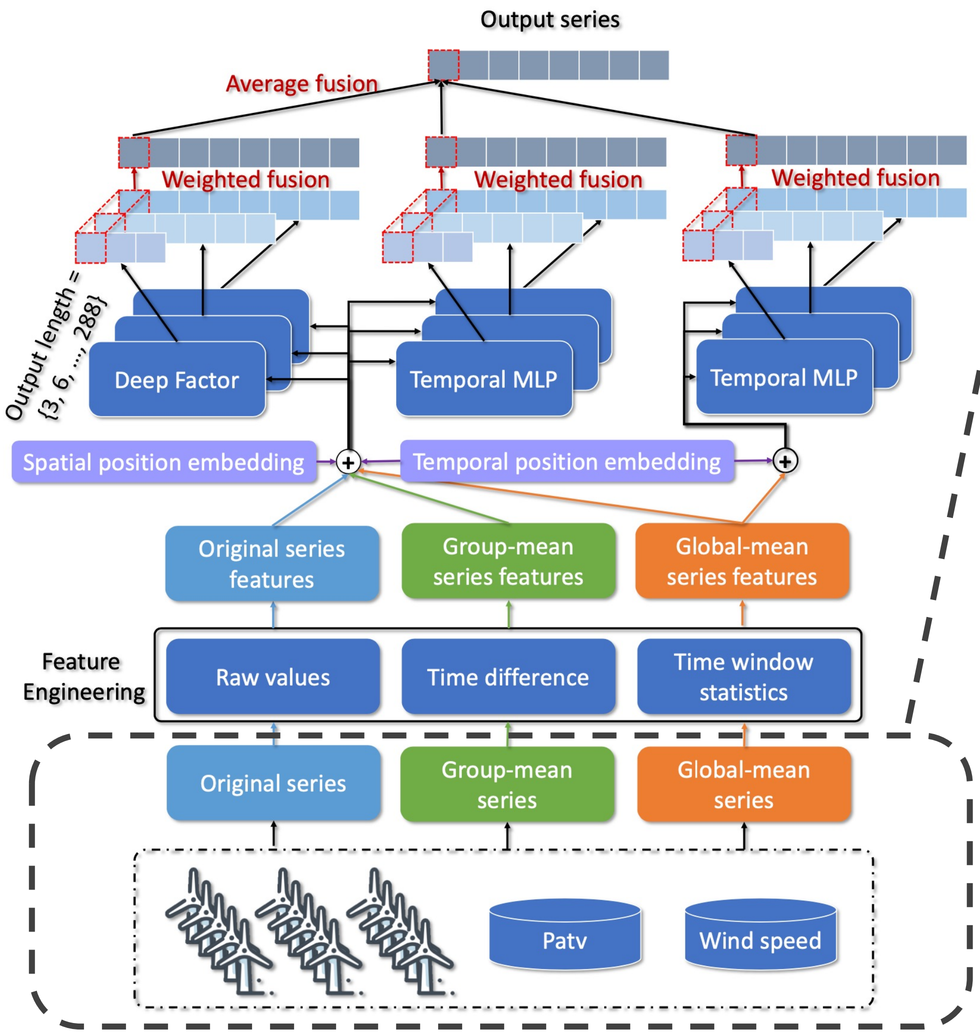


Fig 1. Overview of solution

■ Data Pre-processing

Methods

- ➢ Abnormal values clipping
- ➢ **Spatial group partitioning**
- ➢ **Group & Global mean aggregation（Fig 2.）**
- ➢ Spatial filling of missing values
- ➢ **Coarse-grained aggregation of time series（Fig 3.）**

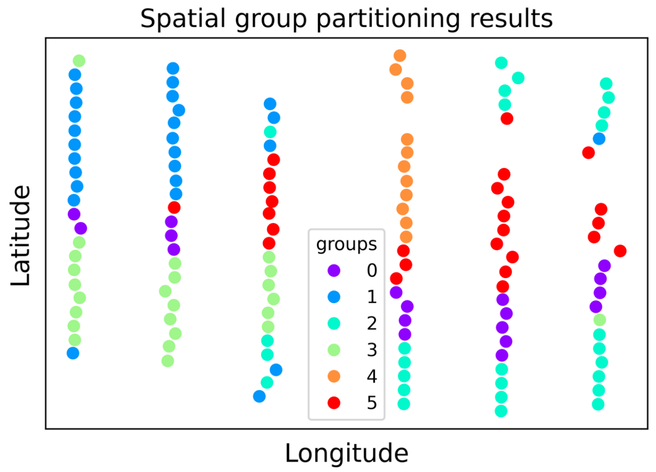


Fig 2. Spatial group partition.

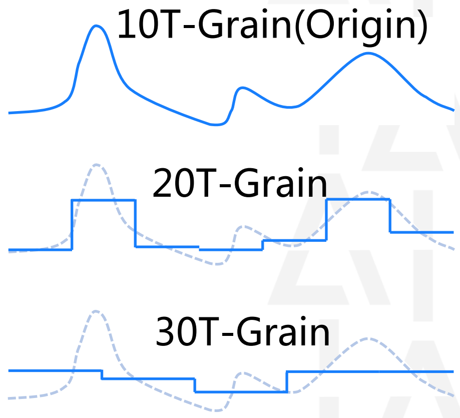


Fig 3. Coarse-grained aggregation

Spatial and Temporal Modeling

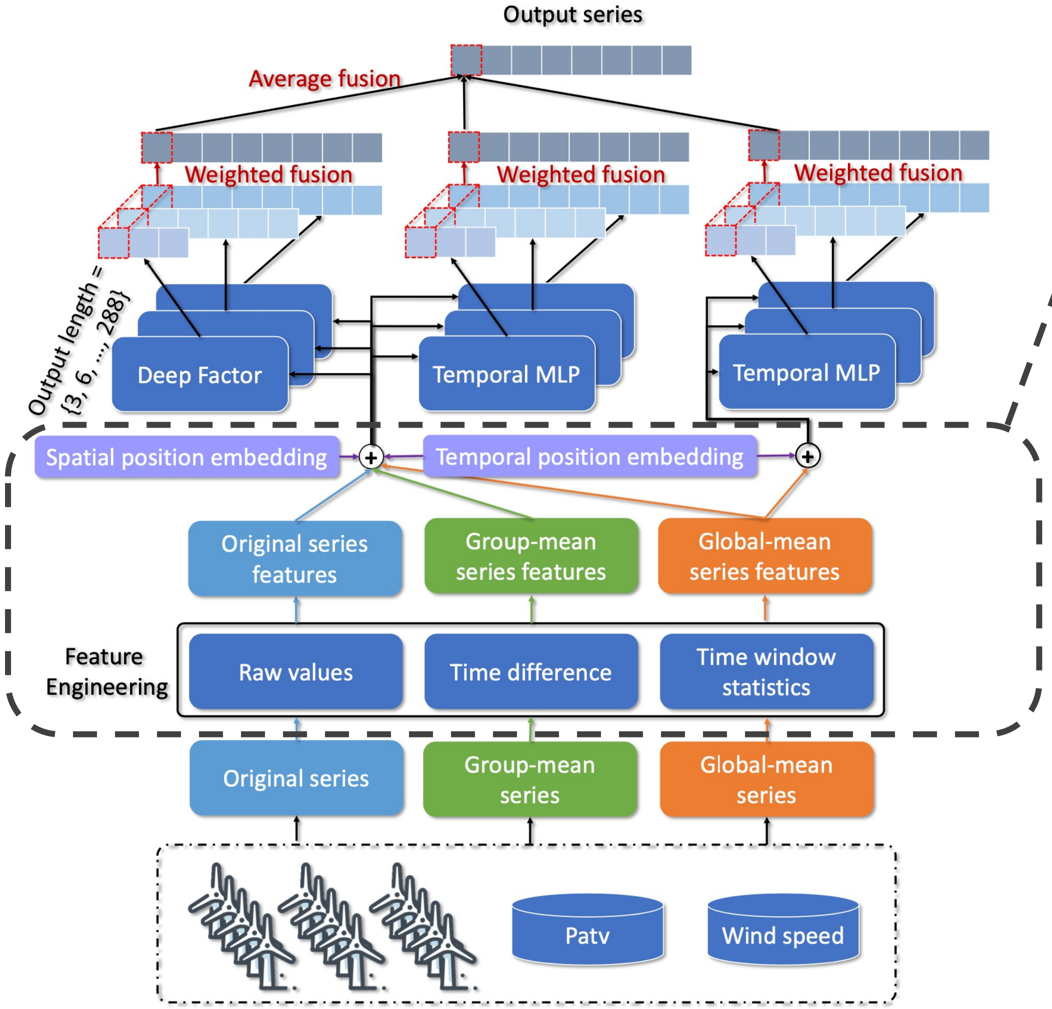


Fig 1. Overview of solution

■ Feature Engineering

feature type	features	configure
raw values	patv,wspd	granularity = 10 min
time different features	diff	Δt = 1,2,3
time window statistics	max, min, std, var, mean, median	window = 1h, 3h, 6h, 12h, 24h, 48h, 72h

➢ Spatial position embedding(SE):

A trainable spatial embedding for each turbine.

$$\mathbf{SE} = \left[z_1^s, z_2^s, \ldots, z_N^s\right] \in \mathbb{R}^{N\times d}$$

➢ Temporal position embedding(TE):

Time-of-day & Hour of day

$$\mathbf{TE} = \left[z_1^t, z_2^t, \ldots, z_{T_d}^t\right] \in \mathbb{R}^{T_d\times d}$$

Deterministic and Probabilistic Modeling

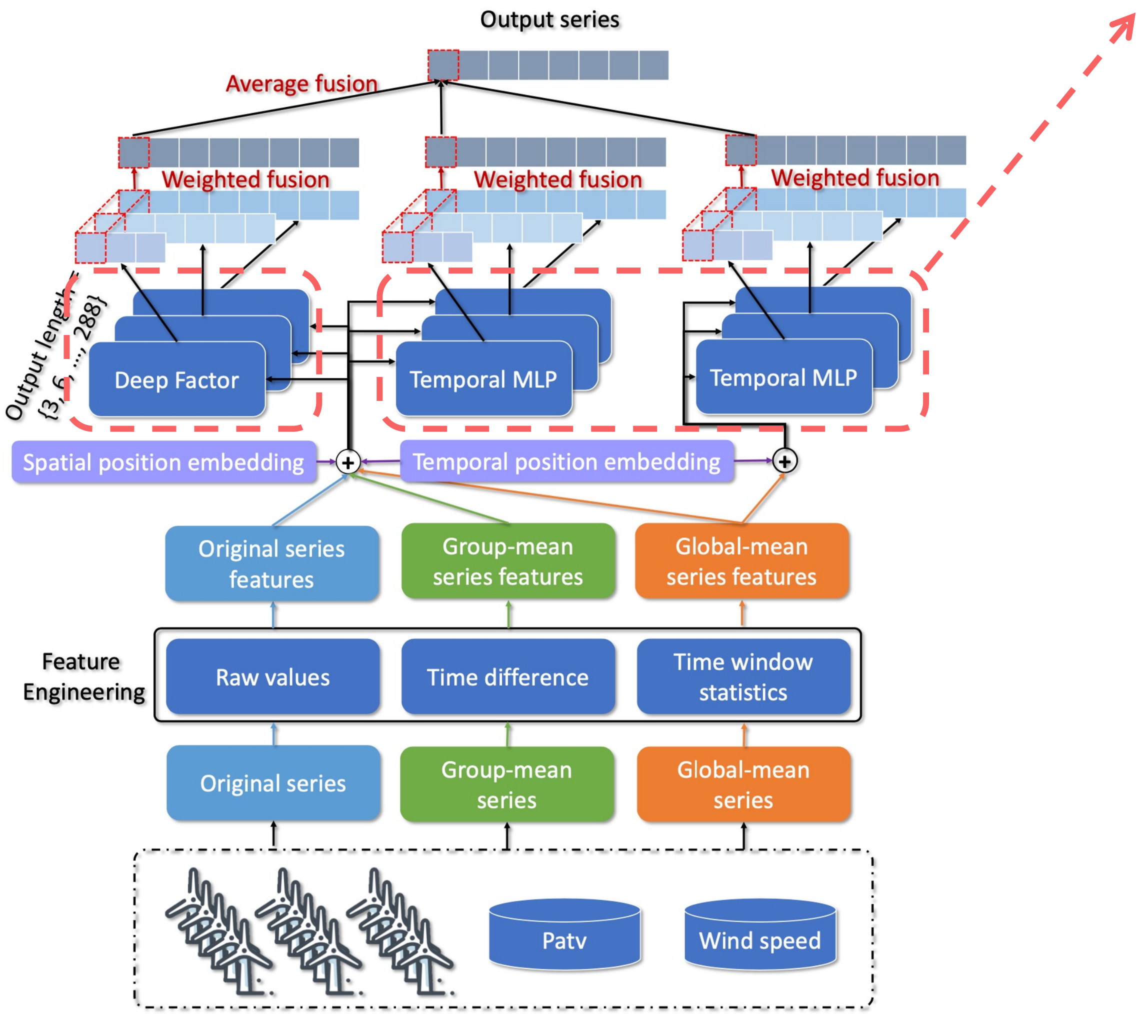


Fig 1. Overview of solution

■ Deterministic Forecasting Model

Temporal MLP: Directly forcast the value of target (Patv)

- ➢ **Temporal-Linear Layer (TLinear)**
- ➢ Gated Linear Unit (GLU)

$$\mathcal{H}^l = \mathrm{GLU}(\mathrm{Linear}_{C\to 2C}(\mathcal{H}^{l-1})).$$

$$\tilde{\mathcal{H}}^l = \mathrm{TLinear}_{TC\to C}(\mathcal{H}^l) \in \mathbb{R}^{B\times N\times C},$$

$$\tilde{\mathcal{H}}^l = \mathrm{GLU}(\mathrm{Linear}_{C\to 2C}(\tilde{\mathcal{H}}^l)) \in \mathbb{R}^{B\times N\times C}.$$

$$\mathcal{O} = \mathrm{GLU}(\mathrm{Linear}_{(L+1)C\to C'}(\tilde{\mathcal{H}}^0 \parallel \cdots \parallel \tilde{\mathcal{H}}^L)) \in \mathbb{R}^{B\times N\times C'},$$

$$\hat{\mathcal{Y}} = Reshape(\mathrm{Linear}_{C'\to P}(\mathcal{O})) \in \mathbb{R}^{B\times N\times P\times 1}.$$

$$\mathcal{L}_{\mathrm{score}} = \frac{1}{2SP}\left(\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{t=1}^{P} m_{s,i,t}\cdot\left[\left(y_{s,i,t}-\hat{y}_{s,i,t}\right)^2 + \left|y_{s,i,t}-\hat{y}_{s,i,t}\right|\right]\right)$$

Deterministic and Probabilistic Modeling

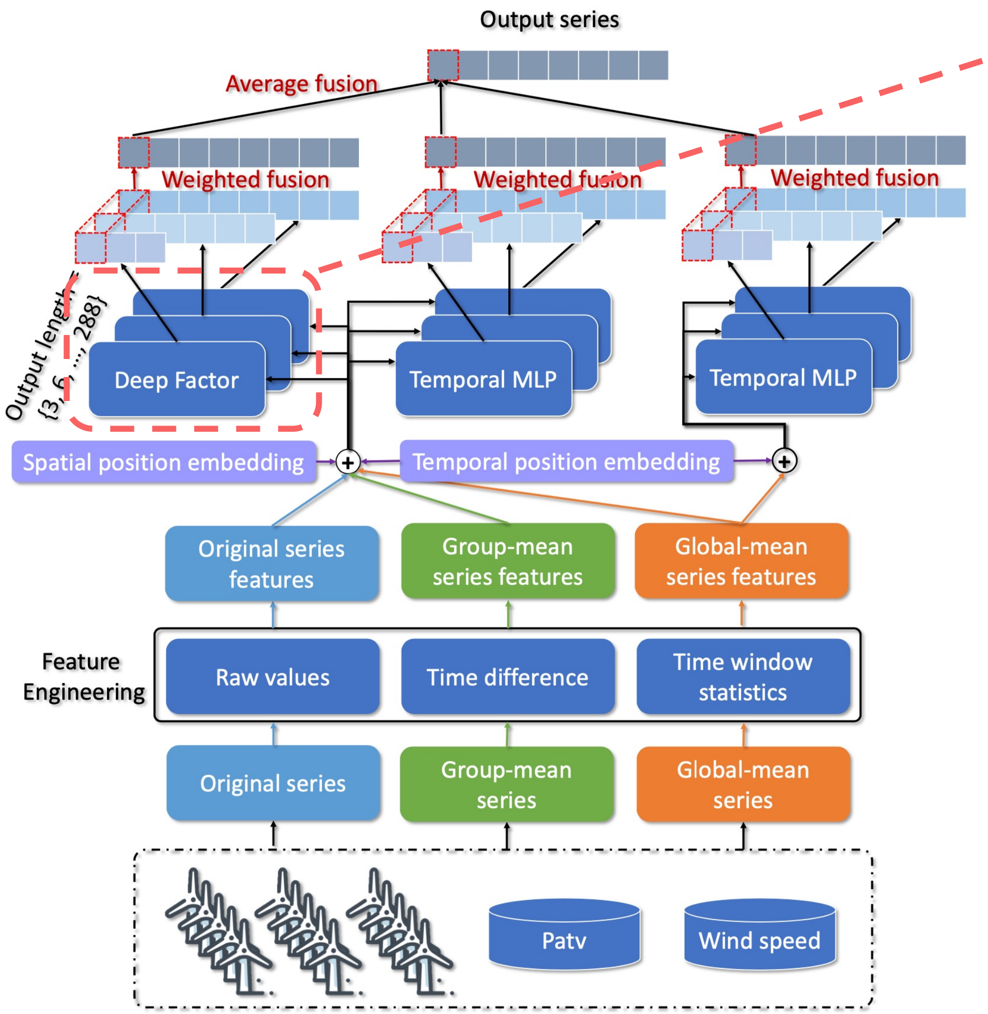


Fig 1. Overview of solution

■ Probabilistic Forecasting Model

Deep Factor: calculate the joint predictive distribution of future observations

$$\begin{aligned} \text{fixed effect}: &\quad \boldsymbol{\mu}_t = \mathrm{Linear}_{C\to 1}(\mathrm{GRU}(\mathbf{H}_{t-1})) \in \mathbb{R}^{BN}, \\ \text{random effect}: &\quad \boldsymbol{\sigma}_t = \mathrm{Linear}_{C\to 1}(\mathrm{GRU}(\mathbf{H}_{t-1})) \in \mathbb{R}^{BN}, \\ \text{emission}: &\quad \boldsymbol{y}_t \sim p(\boldsymbol{y}_t | \boldsymbol{\mu}_t, \boldsymbol{\sigma}_t), \end{aligned}$$

$$\begin{aligned} \mathcal{L}_{\mathrm{NLL}} &= -\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j=1}^{P} m_{s,i,t} \cdot \log p(y_{s,i,t} | \mu_{s,i,t}, \sigma_{s,i,t}) \\ &= \sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j=1}^{P} m^{s}_{i,j} \cdot \left(\frac{\log \sigma^2_{s,i,t}}{2} + \frac{(y_{s,i,t} - \mu_{s,i,t})^2}{2\sigma^2_{s,i,t}} + C \right) \end{aligned}$$

Hybrid Model

Table 1. Model table list

model	term	pred length	granularity
Temporal MLP (global-mean)	short	3,6,9,12,24,34	10min
	middle	72	20min
	long	144,288	10min
	long	144,288	30min
Temporal MLP (all-turbine)	short	24,34	10min
Deep Factor	short	3,6,9,12,24,34	10min
	middle	108	30min
	long	144,288	10min

Multi-model & Multi-term & Multi-grain

- **Multi-model (Table 1.)**

  Temporal MLP (global mean & all turbine) & Deep Factor
- **Multi-term (Fig 4.)**

  Short term & middle term & long term
- **Multi-grain**

  10 min (origin) & 20 min & 30 min

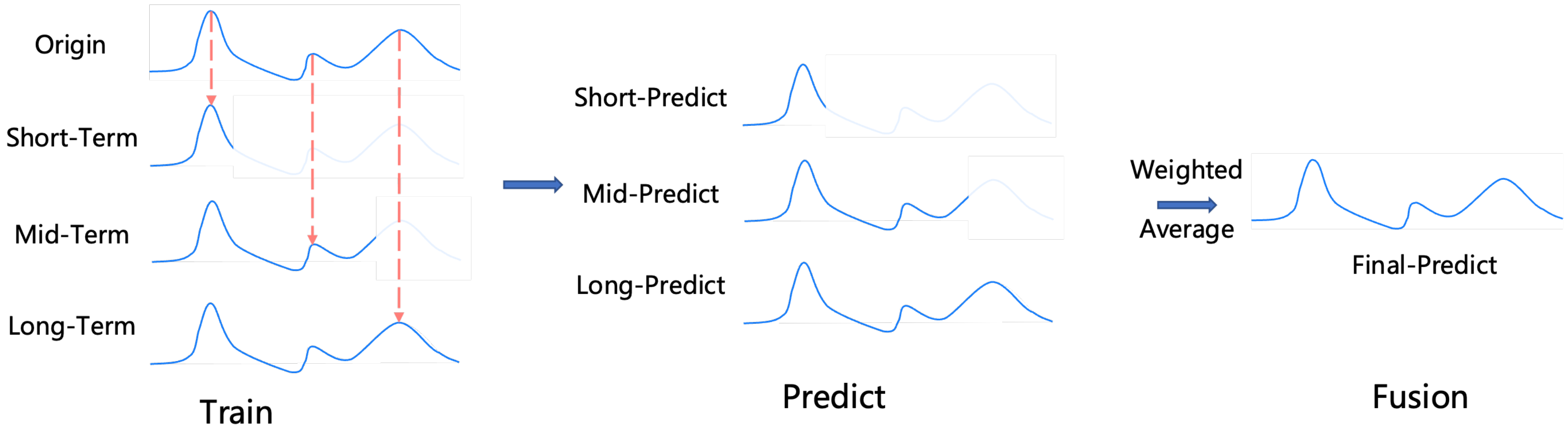


Fig 4. Multi-Term

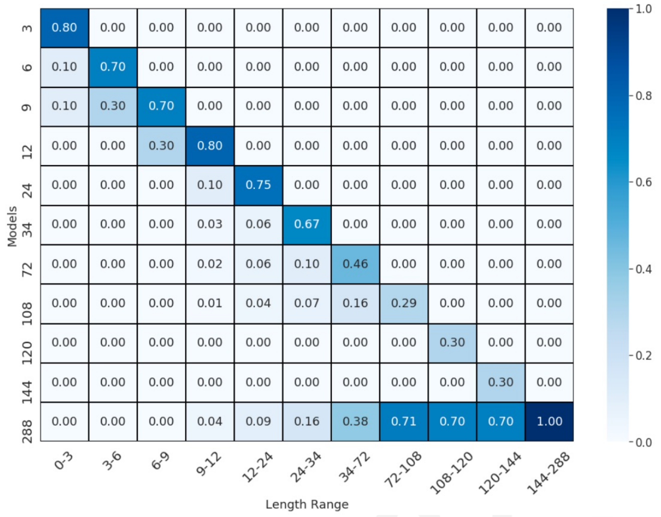
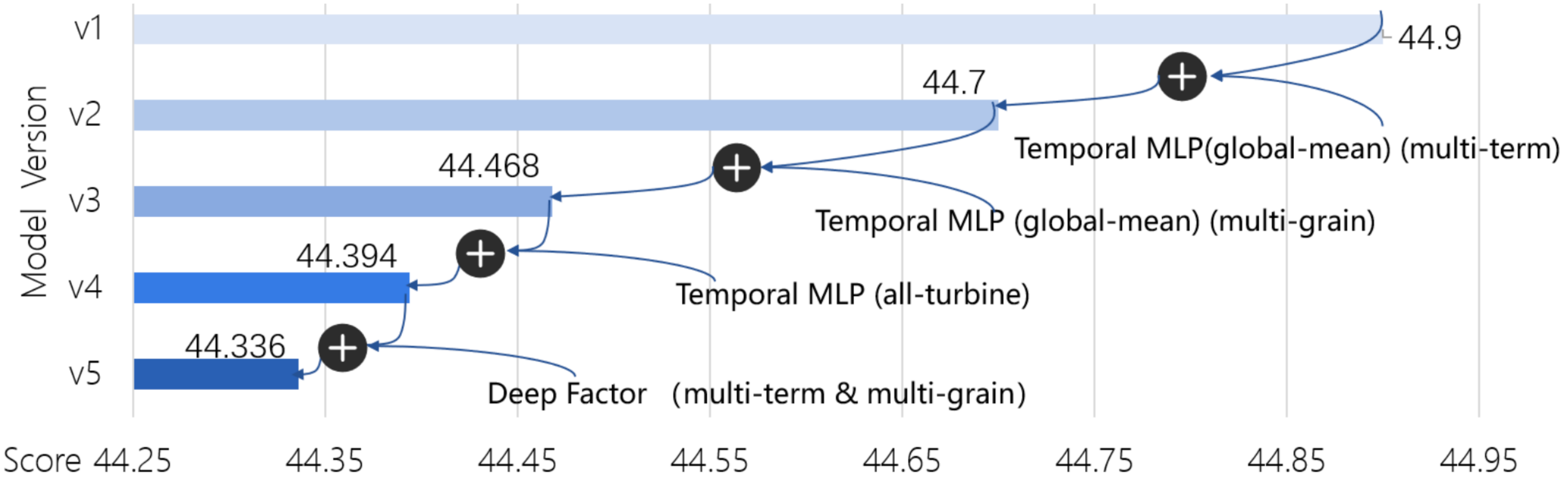


Fig 5. Fusion Weight

Content

- Background
- Challenge
- Solution
- Experiment

Experiment



- **V1** : Temporal MLP (global-mean) model
- **V2** : V1 + **Multi-term** prediction
- **V3** : V2 + **Multi-grain** prediction
- **V4** : V3 + poral MLP (**all-turbine**)
- **V5** : V4 + **Deep Factor**
- **Rank** : Regular rank : 5 ; PaddlePaddle Rank : 3

Conclusion

The proposed model : H-STWPF

- Spatial and Temporal Modeling
- Deterministic and Probabilistic Modeling
- Multi-term & Multi-grain

For more our works

Official account:

INSIS Lab@BJTU

AILab@Cainiao Network

Personal webpage:

- Yiji Zhao: https://yijizhao.github.io/
- Haomin Wen: https://wenhaomin.github.io/

THANKS