AEM JOURNAL .ORG

PA-ResTCN: A Phase-Aware Residual Calibration Model for Radar Tide Gauges

Yang Han[1,2,*], Xinglin Zhang[3], Chenglong Sun[1,2], Yuchao Wei[1,2]
[1]National Ocean Technology Center, Tianjin 300112, China
[2]Key Laboratory of Ocean Observation Technology, Ministry of Natural Resources, Tianjin 300112, China
[3]Wenzhou Ocean Center, Ministry of Natural Resources, Wenzhou 318000, China

Corresponding author: Yang Han (e-mail: shirlleyhan@126.com)

ABSTRACT Non-contact radar water level gauges are widely used in marine and coastal observation because of their convenient deployment and low maintenance cost. However, in practical operation, radar observations are often affected by installation zero drift, weak temporal lag, wave disturbance, and local nonlinear fluctuations, which leads to systematic deviation from reference gauges. To address this problem, this paper proposes a phase-aware residual temporal convolutional network (PA-ResTCN) for tide level calibration. The proposed framework first performs deterministic lag correction and constant-bias correction according to the physical consistency between radar and buoy-gauge observations, and then learns only the dynamic residual component by a multi-branch temporal convolutional architecture. Specifically, the model contains a trend branch for long-range tidal evolution, a derivative branch for short-term local fluctuation modeling, and a phase feature branch for representing tidal-state priors. A weighted residual-learning objective is further introduced to emphasize turning-point intervals. Experiments were conducted on minute-level paired radar-buoy tide observations collected at a tide station established by our team. The results show that the proposed method reduces the RMSE from 8.526 of raw radar observations to 1.585 on the test set, and further outperforms lag-bias correction, linear residual regression, and vanilla TCN-based calibration models. Ablation experiments verify that residual learning, derivative-aware modeling, and phase-aware auxiliary features are the main sources of performance gain. The proposed method provides an effective and interpretable calibration solution for practical radar water-level monitoring.

INDEX TERMS radar tide gauge; tide level calibration; temporal convolutional network; residual learning; phase-aware modeling; sensor fusion

I. INTRODUCTION

Marine and coastal water-level observation is a fundamental support for storm-surge warning, tidal analysis, coastal engineering, and marine disaster prevention. Among different observation devices, non-contact radar water level gauges have become increasingly important because they avoid direct immersion, are convenient to install, and can operate continuously in harsh environments. However, although radar gauges have clear engineering advantages, their measurements are still influenced by multiple factors, including installation datum mismatch, local environmental interference, and signal instability under complex sea states. As a result, the radar sequence may present systematic bias, mild temporal misalignment, and local fluctuation inconsistency with respect to high-reliability reference gauges.

Traditional calibration and correction methods are mainly based on linear regression, statistical filtering, and state-space estimation. These methods are effective when the discrepancy between sensors is dominated by linear offset or smooth noise, but they are less suitable when the residual error is small in magnitude yet highly structured in time. In the present dataset, preliminary analysis shows that the radar and buoy sequences are already highly correlated in the stable observation interval, while the remaining error mainly consists of a constant offset, a weak lag, and localized nonlinear residuals near rapid transition intervals. This characteristic implies that directly fitting buoy values from scratch using a black-box neural network is neither efficient nor sufficiently interpretable.

Recent deep learning methods have demonstrated strong ability in time-series representation learning and nonlinear mapping. In particular, temporal convolutional networks (TCNs) have shown competitive performance for sequence modeling with stable training behavior and long effective memory. However, directly applying standard TCN, LSTM, or GRU models to radar-gauge calibration still suffers from two limitations. First, the model has to simultaneously learn deterministic sensor bias and dynamic residuals, which increases learning difficulty. Second, conventional architectures do not explicitly emphasize tidal phase information and turning-point-sensitive local fluctuations, which are the key intervals where residual mismatch is more likely to occur.

To address these issues, this paper proposes a phase-aware residual temporal convolutional network, denoted as PA-ResTCN. Instead of directly mapping radar observations to buoy values, the method first uses lag-bias pre-correction to remove the deterministic discrepancy, and then uses a multi-branch temporal convolutional model to learn the remaining residual. The main contributions of this paper are as follows:

1 A two-stage intelligent calibration framework is proposed for radar water level gauges, in which deterministic lag and bias are corrected first, and the deep network learns only the dynamic residual component.

2 A phase-aware residual temporal convolutional architecture is designed, which combines a trend branch, a derivative branch, and a phase feature branch to improve residual modeling under tidal transition conditions.

3 A weighted residual-learning strategy is introduced to emphasize turning-point intervals and improve local dynamic consistency. Ablation experiments verify the effectiveness of each component on real paired radar-buoy observations.

II. RELATED WORK

Current water-level quality control and sensor calibration methods are still dominated by traditional statistical approaches. Kalman-filter-based methods have been used in sea-level monitoring and smoothing tasks and have shown good performance in real-time estimation scenarios [1]. However, these methods usually rely on explicit assumptions about process noise and observation noise, and their performance may degrade when the discrepancy between sensors contains time-varying nonlinear residuals.

For non-contact water-level observation, radar- and vision-based measurement techniques have received extensive attention in recent years. A recent review pointed out that non-contact measurement methods have clear advantages in safety and maintenance, but robustness under complex environmental disturbances remains a critical challenge [2]. Recent deep-learning-based water-level measurement studies have also shown that visual estimation can reach high accuracy when the observation scene is favorable, although its performance may still degrade under complex backgrounds, reflections, and illumination changes [3], [4].

Beyond direct level measurement, data-driven water-level prediction and multi-station modeling have developed rapidly. Graph convolutional recurrent networks have been used for multi-site tide-level prediction [5], two-stage modeling strategies have been explored for multi-station water-level forecasting [6], and attention-enhanced ConvLSTM architectures have been introduced for nearshore water-level prediction [7]. These studies indicate that modern sequence models are effective at learning temporal dependence and cross-scale fluctuation patterns in hydrological and marine processes.

More recently, general-purpose time-series foundation architectures have provided stronger sequence representation ability. TimesNet models temporal variations in a two-dimensional manner [8], PatchTST reformulates long-horizon prediction with patch-based transformers [9], and TimeMixer performs decomposable multiscale mixing for complex time-series forecasting [10]. In parallel, diffusion-based methods have been introduced for probabilistic time-series imputation and forecasting, including CSDI and autoregressive denoising diffusion models [11], [12], and recent surveys have summarized their broader potential for time-series and spatio-temporal data analysis [13], [14].

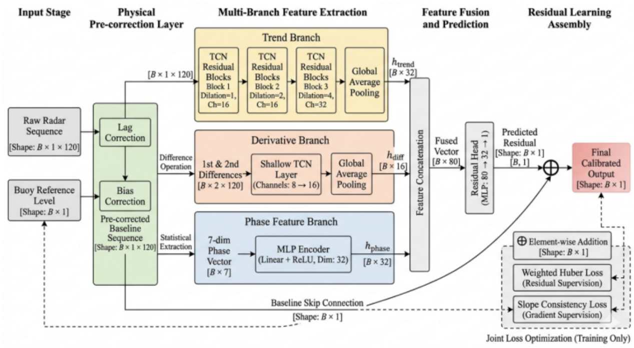

Fig. 1. Overall architecture of PA-ResTCN.

However, most existing neural-network-based sequence models are designed for measurement, forecasting, or imputation tasks in a general setting. For radar water-level calibration, the discrepancy structure is more specific: the dominant part is deterministic and interpretable, while the remaining part is localized and phase-dependent. Therefore, a specialized model should preserve physical interpretability at the front end and reserve nonlinear representation learning only for the residual component. This motivates the proposed PA-ResTCN framework.

III. THE PROPOSED METHOD

A. OVERALL FRAMEWORK OF PA-RESTCN

This paper proposes a physics-informed phase-aware residual calibration framework for paired radar and buoy water-level sequences. The model architecture is shown conceptually in Fig. 1 and consists of two stages:

1 Physics-informed pre-correction stage: estimate the optimal lag and constant bias from paired observations and obtain a pre-corrected radar series.

2 Residual learning stage: use a multi-branch temporal convolutional network to predict the remaining nonlinear residual between the pre-corrected radar signal and the buoy reference.

Let the original radar observation at time step t be x_t^r, and the buoy observation be y_t. The pre-corrected radar value is defined as

$$x_t^{(0)} = x_{t+\tau^*}^r - b \tag{1}$$

where τ^* is the optimal lag estimated by cross-correlation, and b is the constant bias estimated from the training set. The residual to be learned is then

$$r_t = y_t - x_t^{(0)} \tag{2}$$

The final prediction is

$$\widehat{y}_t = x_t^{(0)} + \widehat{r}_t \tag{3}$$

where $\widehat{r}_t$ is produced by the proposed residual network.

This design has two advantages. First, it reduces the burden on the neural network by removing deterministic error sources. Second, it preserves physical interpretability because lag and offset are estimated explicitly rather than hidden inside a black-box mapping.

B. LAG-BIAS PRE-CORRECTION MODULE

The first stage estimates the temporal alignment and systematic offset between radar and buoy signals. For a candidate lag τ within a predefined search interval, the correlation coefficient is computed on the overlapping sequence segments. The optimal lag is given by

$$\tau^* = arg \max_{\tau \in [-L, L]} \rho\left(x_{t+\tau}^r, y_t\right) \tag{4}$$

where $\rho(\cdot)$ denotes the Pearson correlation coefficient and L is the maximum lag search range.

After lag correction, the constant bias is estimated by

$$b = \frac{1}{N} \sum_{t=1}^{N} \left(x_{t+\tau^*}^r - y_t\right) \tag{5}$$

In the present experiment, the optimal lag estimated on the training set is 4 min and the mean bias is -7.992 in the original engineering unit. This confirms that the dominant discrepancy between the two sensors is a stable datum mismatch plus weak temporal misalignment.

C. MULTI-BRANCH RESIDUAL TEMPORAL CONVOLUTIONAL NETWORK

The residual learning stage is designed to model the remaining nonlinear discrepancy after pre-correction. The architecture contains three branches.

1) Trend Branch

The trend branch receives the pre-corrected radar sequence within a sliding time window:

$$s_t = [x_{t-W+1}^{(0)}, x_{t-W+2}^{(0)}, \ldots, x_t^{(0)}] \tag{6}$$

where W is the input window length. This branch uses dilated causal convolutions and residual blocks to capture long-range tidal evolution and slow variation patterns. Compared with recurrent models, TCN provides stable optimization and can model long effective memory using fewer parameters [3].

2) Derivative Branch

Because the main residual error is concentrated in short-term local fluctuations, this paper introduces a dedicated derivative branch. The first-order and second-order differences are defined as

$$\Delta x_t = x_t^{(0)} - x_{t-1}^{(0)} \tag{7}$$

$$\Delta^2 x_t = \Delta x_t - \Delta x_{t-1} \tag{8}$$

The derivative branch takes Δx and $\Delta^2 x$ as inputs and focuses on modeling local rate-of-change and curvature information. This branch is particularly useful for capturing the mismatch near rapid tidal transition intervals, where direct value-level similarity remains high but local slope-level consistency may deteriorate.

3) Phase Feature Branch

To encode the physical state of the water-level process, a phase feature branch is introduced. It contains auxiliary descriptors that are not represented explicitly by the raw level sequence, including:

- hour-of-day sine and cosine embeddings;
- rolling standard deviation within the recent 15 min;
- rolling range within the recent 30 min;
- turning-intensity indicator;
- instantaneous first-order and second-order derivatives.

The turning-intensity indicator is defined as

$$p_t = \frac{|\Delta^2 x_t|}{|\Delta x_t| + 1} \tag{9}$$

These features are passed through a lightweight multilayer perceptron to generate a compact phase representation.

4) Residual Fusion and Output

The trend and derivative branches are first fused in the feature space, after which the phase representation is concatenated before the output head. The network predicts only the residual $\widehat{r}_t$, rather than the absolute water level $\widehat{y}_t$. This residual-learning formulation makes the network focus on the most difficult part of the calibration task and avoids wasting model capacity on deterministic components that have already been removed.

In the implementation used in this paper, the best-performing model on the current dataset does not rely on an explicit multiplicative gate. A gated variant was also evaluated and is reported in the ablation section.

D. LOSS FUNCTION

The calibration objective should not only reduce pointwise value error but also preserve local dynamic consistency. Therefore, this paper adopts a weighted residual-learning objective based on the Huber loss:

$$\mathcal{L}_{\text{res}} = \frac{1}{N} \sum_{t=1}^{N} w_t \, \text{Huber}(\widehat{r}_t, r_t) \tag{10}$$

where w_t is the turning-point weight. Specifically,

$$w_t = 1 + min\left(\frac{p_t}{6}, 1\right) \tag{11}$$

This design increases the contribution of intervals with stronger local curvature, thereby encouraging the model to focus on physically difficult segments. In addition, an auxiliary slope head is used to predict the one-step buoy increment

$$\Delta y_t = y_t - y_{t-1} \tag{12}$$

and the corresponding auxiliary loss is

$$\mathcal{L}_{\text{slope}} = \frac{1}{N}\sum_{t=1}^{N}(\widehat{\Delta y}_t - \Delta y_t)^2 \tag{13}$$

The total loss is

$$\mathcal{L} = \mathcal{L}_{\text{res}} + \lambda\mathcal{L}_{\text{slope}} \tag{14}$$

where λ is set to 0.1 in this study.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL SETUP

The dataset used in this paper consists of paired minute-level radar and buoy water-level observations from a coastal station. According to prior quality-control analysis, the stable overlapping interval is from 2026-03-21 00:00 to 2026-04-12 23:59, containing 33,120 valid paired minute points. After constructing sliding-window samples with a window length of 120 min, 32,997 supervised samples are obtained.

To avoid temporal leakage, the data are split by complete days rather than randomly by points. The first 16 days are used for training, the next 3 days for validation, and the final 4 days for testing. All models are implemented in PyTorch and trained on CPU in the present experiment environment. The Adam optimizer is used with a learning rate of 10^{-3}, batch size 512, and a maximum of 8 epochs with early stopping.

B. BASELINE MODELS AND EVALUATION METRICS

To evaluate the effectiveness of the proposed calibration strategy, the following baselines are considered:

1. Raw Radar: direct radar observation without correction.
2. Lag-Bias: deterministic lag and constant-bias correction only.
3. Linear Residual: linear regression on residual features.
4. Vanilla TCN Direct: standard TCN that directly predicts buoy water level.
5. Res-TCN: TCN residual model without derivative branch or phase branch.
6. Res-TCN-Diff: residual TCN with derivative branch only.
7. PA-ResTCN: the proposed full calibration model without explicit gating.

TABLE I. Performance comparison of different calibration models

Model	MAE	RMSE	R^2	P95	Slope-MAE
Raw Radar	8.115	8.526	0.9849	12.000	1.410
Lag-Bias	1.238	1.724	0.9994	3.008	1.420
Linear Residual	1.280	1.719	0.9994	3.399	1.256
Vanilla TCN Direct	4.913	5.895	0.9928	10.597	0.688
Res-TCN	1.277	1.751	0.9994	3.567	1.436
Res-TCN-Diff	1.211	1.591	0.9995	3.201	0.842
PA-ResTCN (Ours)	1.204	1.585	0.9995	3.199	0.835

The evaluation metrics include mean absolute error (MAE), root mean square error (RMSE), coefficient of determination (R^2), 95th percentile absolute error (P95), and slope MAE. The slope MAE is defined on the first-order difference sequence and measures local dynamic consistency:

$$\text{Slope-MAE} = \frac{1}{N-1}\sum_{t=2}^{N}\left|(\widehat{y}_t - \widehat{y}_{t-1}) - (y_t - y_{t-1})\right| \tag{15}$$

C. QUANTITATIVE RESULTS

Table I summarizes the experimental results on the test set.

The results indicate that deterministic pre-correction already removes most of the discrepancy between the two sensors, reducing the RMSE from 8.526 to 1.724. This confirms that the radar-buoy difference is strongly dominated by stable offset and weak lag. However, deterministic correction alone still leaves a structured residual pattern.

Compared with Lag-Bias correction, the proposed PA-ResTCN further reduces RMSE from 1.724 to 1.585, corresponding to an additional reduction of about 8.1%. Compared with raw radar observations, the total RMSE reduction reaches 81.4%. More importantly, the proposed model improves slope consistency substantially, reducing Slope-MAE from 1.420 to 0.835. This suggests that the model not only calibrates the absolute level more accurately but also better preserves local dynamic evolution.

The comparison between Vanilla TCN Direct and residual-learning models is particularly instructive. Directly predicting buoy values with a generic TCN yields much worse accuracy than residual formulations, showing that for this task the key challenge is not to reconstruct the entire tidal process but to model the small structured discrepancy left after physics-informed pre-correction.

To further visualize the calibration behavior of different methods on the test set, Fig. 2 presents the full-test-set tide-level comparison, and Fig. 3 shows a local zoom view in a challenging interval. These two figures provide intuitive evidence of how the proposed model improves both global fitting quality and local dynamic consistency.

From Fig. 2, it can be seen that the raw radar sequence follows the overall tidal trend but exhibits a clear systematic negative bias relative to the buoy reference. After explicit lag-bias correction, the radar sequence becomes much closer to the buoy reference. Once residual learning is introduced, all deep models further improve the fit to different extents,

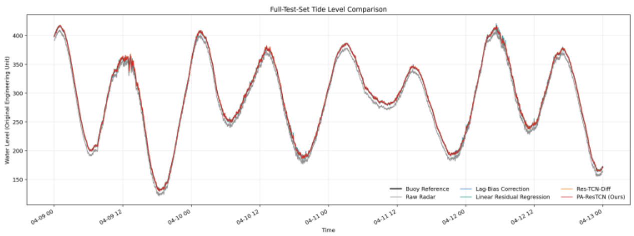

Fig. 2. Full-test-set tide level comparison.

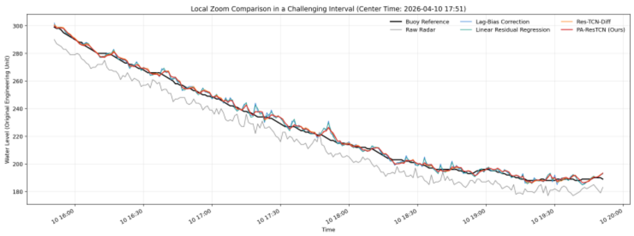

Fig. 3. Local zoom comparison in a challenging interval.

among which the proposed PA-ResTCN stays closest to the buoy reference during most time periods. This indicates that the proposed model can preserve the main tidal trend while further correcting local structured residual errors.

Fig. 3 provides a clearer view of model behavior in a difficult interval characterized by a continuously falling tide level with local slope fluctuations. The raw radar sequence deviates the most from the buoy reference, while lag-bias correction restores the overall level but still leaves visible local discrepancies near several transition points. In contrast, the proposed PA-ResTCN tracks the local fluctuations of the buoy reference more accurately, and its curve remains closer to the reference sequence than the other methods at most critical turning points. This result further demonstrates the advantage of the proposed model in local dynamic consistency and complex residual compensation.

D. ABLATION STUDY

To verify the contribution of each component, this paper designs the following ablation configurations:

1. Vanilla TCN Direct: no pre-correction-based residual learning.
2. Res-TCN: residual learning only.
3. Res-TCN-Diff: residual learning with derivative branch.
4. PA-ResTCN w/o Gate: residual learning with derivative branch, phase branch, and weighted loss.
5. PA-ResTCN Full: adds an explicit multiplicative gate for trend-difference fusion.

The ablation results are shown in Table II.

Several conclusions can be drawn from Table II.

First, residual learning is essential. Once the model is reformulated from direct buoy prediction to residual prediction, RMSE drops from 5.895 to 1.751. This strongly supports the physics-informed decomposition adopted in this paper.

TABLE II. Ablation results of model components

Configuration	MAE	RMSE	R^2	P95	Slope-MAE
Vanilla TCN Direct	4.913	5.895	0.9928	10.597	0.688
Res-TCN	1.277	1.751	0.9994	3.567	1.436
Res-TCN-Diff	1.211	1.591	0.9995	3.201	0.842
PA-ResTCN w/o Gate	1.204	1.585	0.9995	3.199	0.835
PA-ResTCN Full	1.204	1.586	0.9995	3.195	0.785

Second, the derivative branch contributes most of the nonlinear modeling gain. Adding derivative features reduces RMSE from 1.751 to 1.591 and reduces Slope-MAE from 1.436 to 0.842. This confirms that local mismatch between radar and buoy is mainly concentrated in short-term rate-of-change behavior rather than in the smooth background tide.

Third, phase-aware auxiliary features and weighted learning still provide additional improvement. The PA-ResTCN variant without explicit gating achieves the best overall RMSE and MAE, indicating that recent statistical fluctuation and tidal-state descriptors are useful for residual correction.

Finally, the explicit gate does not produce stable improvement in the current single-station, small-sample setting. Although the full gated model obtains the best P95 and Slope-MAE, its RMSE is slightly higher than the non-gated variant. Therefore, the non-gated PA-ResTCN is adopted as the main model in this paper, while the gated version is regarded as a promising extension for larger multi-station datasets.

E. DISCUSSION

The present results reveal an important engineering implication: for radar water-level calibration, a purely data-driven end-to-end model is not always the best choice. When the discrepancy structure is dominated by explicit sensor lag and datum mismatch, combining physical pre-correction with compact residual learning is both more interpretable and more effective.

Another notable observation is that local dynamic consistency benefits more from derivative-aware modeling than absolute-value accuracy alone would suggest. In practical marine monitoring, this is important because anomaly warning and short-term process analysis are often more sensitive to local variation than to small constant shifts.

At the same time, this study still has limitations. The experiment is conducted on a single station with a relatively short continuous high-quality interval. The model has not yet been tested under cross-station transfer, seasonal changes, or extreme weather disturbances. Therefore, the present work should be interpreted as a station-specific intelligent calibration study rather than a universal radar correction model.

V. CONCLUSION

This paper proposes a physics-informed phase-aware residual temporal convolutional network for intelligent calibration of radar water level gauges. The key idea is to decouple deterministic discrepancy and nonlinear residual discrepancy: lag correction and constant-bias correction are handled explicitly, while a compact multi-branch TCN is used to learn only the dynamic residual component. By combining a trend branch, a derivative branch, and a phase feature branch, the proposed model effectively captures both long-period tidal evolution and local mismatch near rapid transition intervals.

Experiments on real paired radar-buoy observations show that the proposed method substantially improves calibration accuracy over raw radar data and achieves better performance than deterministic correction, linear residual regression, and standard TCN calibration baselines. Ablation experiments further demonstrate that residual learning and derivative-aware modeling are the principal sources of improvement, while the contribution of explicit gating remains limited under the current sample scale.

The proposed framework offers an interpretable and practically deployable solution for intelligent calibration of radar water-level gauges. Future work will focus on three directions:
1) extending the model to multi-station and multi-season datasets;
2) introducing joint calibration and anomaly-detection learning; and
3) incorporating external environmental variables such as wind, wave, and meteorological observations to further improve robustness under extreme conditions.

FUNDING

This research was financially supported by the Research Fund of the Key Laboratory of Ocean Observation Technology MNR(No. 2024KLOOTB03).

REFERENCES

[1] J. Strandberg, T. Hobiger, and R. Haas, "Real-time sea-level monitoring using Kalman filtering of GNSS-R data," GPS Solutions, vol. 23, Art. no. 61, 2019.
[2] Z. Wu, Y. Huang, K. Huang, K. Yan, and H. Chen, "A review of non-contact water level measurement based on computer vision and radar technology," Water, vol. 15, no. 18, Art. no. 3233, 2023.
[3] Z. Xie, J. Jin, J. Wang, R. Zhang, and S. Li, "Application of deep learning techniques in water level measurement: Combining improved SegFormer-UNet model with virtual water gauge," Applied Sciences, vol. 13, no. 9, Art. no. 5614, 2023.
[4] W. C. Liu and W. C. Huang, "Evaluation of deep learning computer vision for water level measurements in rivers," Heliyon, vol. 10, no. 4, Art. no. e25989, 2024.
[5] X. Zhang, T. Wang, W. Wang, P. Shen, Z. Cai, and H. Cai, "A multi-site tide level prediction model based on graph convolutional recurrent networks," Ocean Engineering, vol. 269, Art. no. 113579, 2023.
[6] Z. Yuan, J. Liu, Y. Liu, Q. Zhang, Y. Li, and Z. Li, "A two-stage modelling method for multi-station daily water level prediction," Environmental Modelling & Software, vol. 156, Art. no. 105468, 2022.
[7] J. Yang, T. Zhang, J. Zhang, X. Lin, H. Wang, and T. Feng, "A ConvLSTM nearshore water level prediction model with integrated attention mechanism," Frontiers in Marine Science, vol. 11, Art. no. 1470320, 2024.
[8] H. Wu, T. Hu, Y. Liu, H. Zhou, J. Wang, and M. Long, "TimesNet: Temporal 2D-variation modeling for general time series analysis," in Proc. ICLR, 2023.
[9] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with transformers," in Proc. ICLR, 2023.
[10] S. Wang, H. Wu, X. Shi, et al., "TimeMixer: Decomposable multiscale mixing for time series forecasting," in Proc. ICLR, 2024.
[11] Y. Tashiro, J. Song, Y. Song, and S. Ermon, "CSDI: Conditional score-based diffusion models for probabilistic time series imputation," in Advances in Neural Information Processing Systems, 2021.
[12] K. Rasul, C. Seward, I. Schuster, and R. Vollgraf, "Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting," in Proc. ICML, 2021.
[13] L. Lin, Z. Li, R. Li, X. Li, and J. Gao, "Diffusion models for time series applications: A survey," Frontiers of Information Technology & Electronic Engineering, 2024.
[14] Y. Yang, M. Jin, H. Wen, et al., "A survey on diffusion models for time series and spatio-temporal data," arXiv:2404.18886, 2024.