

# CompRestacking: Capturing Channel Dependency in Highly Correlated Multivariate Time Series Data (Student Abstract)

Min Kim, Ohyun Jo\*

Chungbuk National University, 28644, Cheongju, Republic of Korea  
kmin010287@chungbuk.ac.kr, ohyunjo@chungbuk.ac.kr

## Abstract

The consideration of channel correlation is crucial for improving the performance of multivariate time series forecasting. However, existing models fail to capture it in homogeneous and highly correlated channels. In this work, we introduce CompRestacking (Compression Restacking), a strikingly intuitive and effective method to address this problem. The approach consists of three main components: (1) PCC-Restacking for correlation-aware channel ordering, (2) Temporal embedding for time encoding, and (3) Aggregation compression for compact token generation. CompRestacking consistently outperforms in experiment results. The results demonstrate that CompRestacking leverages strong channel correlations for improved performance.

## Introduction

Multivariate time series (MTS) data are widely used in real-world domains, including electricity, energy, and transportation. In these environments, accurate forecasting based on observed data is essential for decision making, policy formulation, and strategic planning. In particular, channel correlation is crucial for improving the performance of MTS forecasting (Han et al. 2024).

To investigate how the channel independence within a dataset influences prediction performance, we conduct a systematic analysis employing well-established time series datasets and four state-of-the-art benchmark models. We conducted experiments with four models across three datasets exhibiting distinct characteristics: ETT, heterogeneous channel properties; ECL, homogeneous but weak correlations; and Seoul Traffic (Kang and Jo 2024), both homogeneous and strongly correlated.

As demonstrated in Table 1, the amount of error is markedly elevated in the Seoul Traffic dataset. This striking divergence underscores the inherent limitations of existing correlation learning methodologies. Therefore, it is evident that a methodological advancement capable of more effectively accounting for channel correlations is required to play a decisive role in enhancing forecasting performance. In this study, we introduce **CompRestacking** (Compression

Dataset	TimesNet		TimesMixer		iTransformer		FEDformer	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	0.0779	0.2099	0.0729	0.2048	0.0708	0.2038	0.0876	0.2290
ETTh2	0.1914	0.3395	0.1707	0.3228	0.1793	0.3332	0.1772	0.3288
ECL	0.0880	0.2266	0.0518	0.1609	0.0643	0.1881	0.1142	0.2684
Seoul Traffic	<b>0.3100</b>	<b>0.3872</b>	<b>0.3349</b>	<b>0.4312</b>	<b>0.4024</b>	<b>0.4893</b>	<b>0.3737</b>	<b>0.4614</b>

Table 1: The elevated amount of error in a homogeneous and highly correlated dataset.

Restacking), an intuitive and powerful framework specifically designed to structurally capture channel correlations. To this end, we particularly emphasize validating its effectiveness on the Seoul Traffic dataset, thereby demonstrating its capability in modeling strong channel dependencies.

## CompRestacking

**PCC-Restacking.** PCC-Restacking is a channel ordering methodology for MTS based on the Pearson Correlation Coefficient (PCC). As shown in Figure 1-(a), the procedure begins by positioning the target channel  $c_t$  at the center of the input series. We compute the PCC between  $c_t$  and each candidate channel  $c_n$ . In  $Restack(\cdot)$ , the channel with the highest correlation is then stacked either directly above or below the current stack according to  $pcc_n$ . At this stage, the target channel is redefined as the set of stacked channels. This procedure is applied iteratively until all  $N$  channels are arranged into restack layer  $\hat{X}$ .

$$pcc_n = PCC(c_t, c_n), \quad \hat{X} = Restack(c_t, \{pcc_n\}_{n=1}^N). \quad (1)$$

**Temporal Embedding.** Temporal information  $t_i$  is extracted from the original series. This approach enhances predictive accuracy by accounting for the periodicity of the data. The temporal granularity of the time embedding can be determined flexibly according to daily, weekly, monthly, or seasonal characteristics. As shown in Figure 1-(b), we transform time index  $i_{i=1}^T$  using sinusoidal functions with period  $P$  as follows:

$$e_i^s = \sin\left(\frac{2\pi t_i}{P}\right), \quad e_i^c = \cos\left(\frac{2\pi t_i}{P}\right). \quad (2)$$

\*Corresponding author: Ohyun Jo  
Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

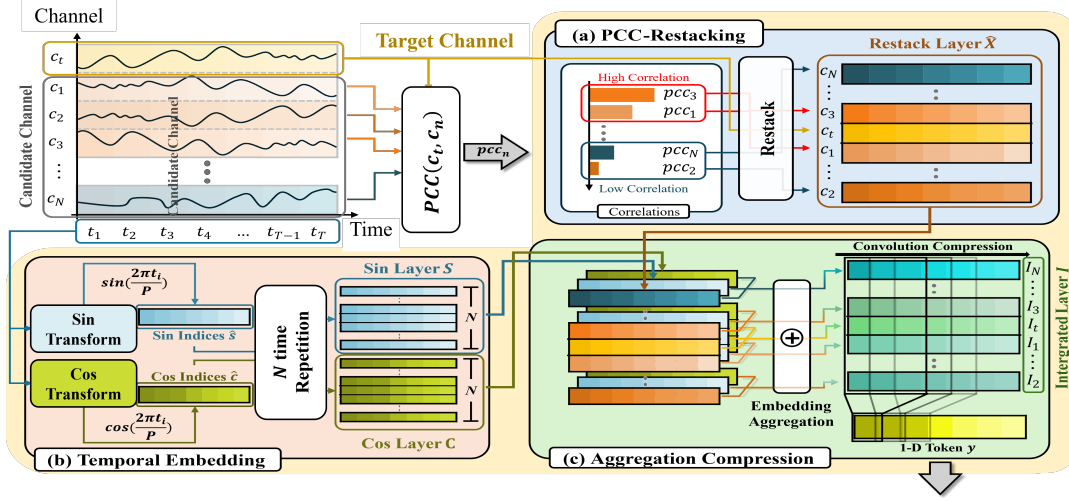


Figure 1: Overview of CompRestacking

where  $e_i^s$  and  $e_i^c$  represents the elements of sin indices  $\hat{s}$  and cos indices  $\hat{c}$  respectively.  $\hat{s}$  and  $\hat{c}$ , following the aforementioned process, are repeated as  $N$  times to transform indices into Sin layer  $S$  and Cos layer  $C$ .

**Aggregation Compression.** Aggregation compression is adopted to condense essential information into a 1-dimensional token  $y$ . As shown in Figure 1-(c), Restack layer  $\hat{X}$ , Sin layer  $S$  and Cos layer  $C$  are aggregated to make the integrated series  $I = \hat{X} + S + C$ . Thereafter, a convolutional layer is employed to compress  $I$  via convolution along the time axis into a 1-dimensional token  $y \in \mathbb{R}^{K \times 1}$ , where  $K$  denotes the length of token.

## Experiments

**Experimental Setup.** For performance evaluation of CompRestacking, experiments were conducted on the Seoul Traffic dataset, comprising measurements from 105 road sections in Seoul highway with less than 1% missing values. The experiments focus on predicting future traffic speeds for a specific target road section.

For comparison of proposed method, four benchmark models are employed in this experiment: TimesNet, CNN-based; TimeMixer, MLP-based; iTransformer and FEDformer, Transformer-based. All the benchmark models are following the same lookback window  $L = \{24, 48, 72, 96\}$  and forecasting horizon  $H = \{96, 120, 144, 168\}$ .

**Experimental Results.** Table 2 summarizes performance across horizons. The proposed method consistently achieved the best performance across all forecasting horizons. Compared to CNN-based TimesNet, the proposed method achieves a 22.07% reduction in terms of MSE and 16.14% reduction in terms of MAE. Notably, when compared to the Transformer-based iTransformer, the proposed method achieves a 39.17% reduction in terms of MSE and 31.76% reduction in terms of MAE. Compared to FEDFormer, the proposed method reduces amounts of error by 33% and

26.28% reduction in terms of MSE and MAE, respectively. Compared to TimeMixer, the reductions are 26.77% in terms of MSE and 21.68% in terms of MAE.

Forecasting	Ours		TimesNet		TimesMixer		iTransformer		FEDformer	
Horizon	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
96	<b>0.2898</b>	<b>0.3859</b>	0.3154	0.4017	0.3504	0.4290	0.4024	0.4893	0.3629	0.4510
120	<b>0.2268</b>	<b>0.3138</b>	0.3328	0.4165	0.3333	0.4294	0.4064	0.4931	0.3650	0.4521
144	<b>0.2266</b>	<b>0.3110</b>	0.3081	0.3969	0.3289	0.4261	0.4109	0.4973	0.3713	0.4582
168	<b>0.2436</b>	<b>0.3330</b>	0.3100	0.3872	0.3349	0.4312	0.4024	0.4893	0.3737	0.4614

Table 2: Forecasting results measured by MSE and MAE for different prediction lengths.

## Conclusion

The existing methodologies fail to capture channel dependencies in homogeneous and highly correlated channels. In this study, we presented CompRestacking, a strikingly intuitive and effective correlation capture method to address this problem. Experiments across multiple forecasting horizons verified that CompRestacking consistently reduced amount of error over existing models.

## Acknowledgements

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT). (No.2021R1A2C2095289)

## References

- Han, L.; Chen, X.-Y.; Ye, H.-J.; and Zhan, D.-C. 2024. Softs: Efficient multivariate time series forecasting with series-core fusion. *Advances in Neural Information Processing Systems*, 37: 64145–64175.
- Kang, S. W.; and Jo, O. 2024. Multivariate time-series imagnification with time embedding in constrained environments (student abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 23535–23536.