Confusion of Channels : Are All Channels Useful in Multivariate Time Series Forecasting?

Min Kim and [*]**Ohyun Jo**
Department of Computer Science, Chungbuk National University
{kmin010287, ohyunjo}@chungbuk.ac.kr
*Corresponding author

Abstract

Capturing all channel correlations in multivariate time series introduces spurious dependencies arising from uncorrelated channels. Weak correlated channel might introduce noise in the forecasting process, thereby hindering improvements in forecasting performance. To address performance limitation, this study proposes Channel Filtering, a correlation-based preprocessing method. The proposed method excludes weakly correlated channels to enhance performance of prediction. Experimental results demonstrate the proposed approach improves forecasting performance with appropriately adjusted channel sets.

Keywords: Multivariate Time Series, Channel Correlation, Channel Independence, Channel Dependence, Series Filtering.

1. Introduction

Multivariate time series forecasting aims to predict future values by leveraging information across multiple time series channels. Unlike univariate time series forecasting, capturing channel dependencies has garnered attention in existing studies [1, 2]. However, recent studies have reported favorable results under channel-independent strategies [3]. Previous results highlight the limitations of modeling all channel correlations. Low correlations between variables can induce noise in forecasting process, thereby limiting improvements in forecasting performance. Motivated by study observation, this study proposes ***Channel Filtering***, a channel selection scheme based on the correlation coefficient. The proposed method mitigates the effects of low correlations by adapting highly correlated channels for training.

2. Channel Filtering

Calculation of Correlation Coefficient. To identify correlations of each channel, target channel c_t and candidate channels c_n are defined. The target channel represents the channel that the model aims to predict. The candidate channels serve as auxiliary inputs that support the prediction of the target channel. After defining these two types of channels, the mutual relationship between the target channel c_t and each candidate channels c_n is measured using the Pearson Correlation Coefficient (PCC). For each pair of channels, PCC is computed based on their individual time series values, and the mean of all pairwise PCCs is adopted as the overall correlation coefficient. This step quantifies the degree of correlation between channels.

Correlation Selection. Based on computed coefficients, part of correlation coefficients is filtered to select the highly correlated channel. For correlation select, hyperparameter n is set. n is the number of channels for training, determines the balance between retaining sufficient information and suppressing noise. After setting the n, the correlation coefficients are sorted in descending order. Based on the sorted coefficients, top n correlation coefficients are selected as filtered set. Filtered set is the list of n high coefficients. Result of selection process serves as a criterion for identifying inputs with low correlations.

Rearrangement. Rearrangement is the process to create actual input to backbone models. Based on filtered set, candidate channels are selected. The selected channels and target channel are stacked into integrated input. This input excludes the noisy channels and prevents the model form being distracted by low correlations. Backbone model enables to learn effectively due to filtered input.

results indicate that an excessive reduction in the number of channels limits the amount of useful information, leading to degraded performance. When $n = 9$, the model maintained lower error rates than the baseline result with reductions of about 10.2% and 4.5%, respectively. the model recorded approximately 1.6% higher MSE and 4.0% higher MAE compared to the lowest error rate configuration. This result suggests that input includes noisy channel in excessive number of channels. Noise may hinder the learning of relevant patterns. Overall, these results demonstrate that a properly chosen number of correlated channels achieves the most effective balance between information retention and noise reduction, thereby enhancing forecasting accuracy.

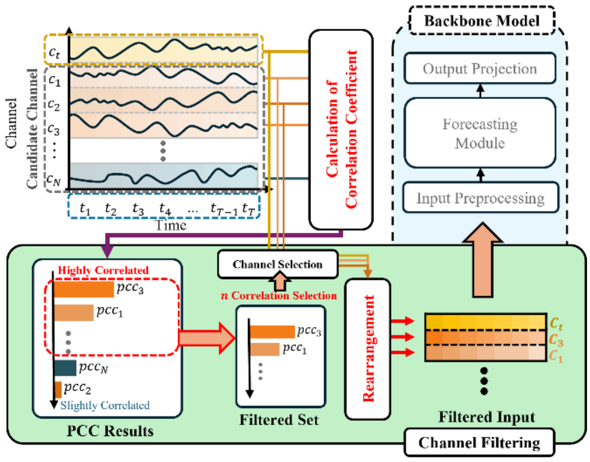

Figure 1: Structure of the proposed Channel Filtering method

3. Experiment

Number of Correlated Channels (n)							
3		5		9		Baseline	
MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
0.2797	0.3780	**0.2788**	**0.3689**	0.2832	0.3836	0.3154	0.4017

Table 1: Forecasting results measured by MSE and MAE for average prediction length.

Experimental Setup. To assess the effects of the channel correlations, experiments were conducted on the Seoul highway traffic dataset, comprising values from 105 road sections in Seoul highway. The experiments were designed to predict future traffic speeds for a specific target channel.

In the baseline configuration, TimesNet was used as the forecasting backbone model for both baseline and channel filter scheme. All 105 channels were utilized without any preprocessing, the baseline model itself. The results are following the same lookback window $L = \{24, 48, 72, 96\}$ and prediction length $H = 96$. Table 1 represents the average performance with respect to the number of correlated channels.

Experimental Results. Applying the proposed scheme led to a clear improvement over the baseline model. The baseline configuration showed an average MSE of 0.3154 and MAE of 0.4017, whereas the filtered configurations consistently reduced both errors. Among all configurations, the $n = 5$ setting produced the lowest forecasting error rates. Compared to the baseline, the error rates were reduced by approximately 11.6% and 8.2% for MSE and MAE. When $n = 3$, the model outperformed the baseline with reductions of about 11.3% in MSE and 5.9% in MAE. However, compared to case of $n = 5$, the error rates were slightly inferior, showing approximately 0.3% higher MSE and 2.5% higher MAE. Experimental

4. Conclusion

This study proposed Channel Filtering, a correlation-based channel selection preprocessing method designed to mitigate the adverse effects of weakly correlated channels in multivariate time series forecasting. Experimental results demonstrated that applying the proposed scheme in Seoul traffic dataset consistently outperformed the baseline across all lookback windows. These findings indicate that incorporating all available channels may not necessarily improve forecasting performance. Appropriately adjusting the number of correlated channels can maximize the benefits of capturing channel dependencies and enhance overall predictive performance.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT). (No.2021R1A2C2095289) In part, this work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT). (No. 2021-0-00165, Development of 5G+ Intelligent Basestation Software Modem), and this work was supported by Chungbuk National University BK21 program(2025)

References

[1] KANG, Seung Woo; JO, Ohyun. Multivariate time-series imagification with time embedding in constrained environments. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. 2024. p. 23535-23536.

[2] WU, Haixu, et al. Timesnet: Temporal 2d-variation modeling for general time series analysis. *arXiv preprint arXiv:2210.02186*, 2022.

[3] NIE, Y. A Time Series is Worth 64Words: Long-term Forecasting with Transformers. *arXiv preprint arXiv:2211.14730*, 2022.