DEMamba: Decoupled Enhanced State Space Models with Selective Mechanisms for Multivariate Time Series Forecasting

Extended Abstract

Junluo Zheng
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, China
zhengjl688@163.com

Yang Liu
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, China
ly267212826@163.com

Jianyong Chen*
College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, China
jychen@szu.edu.cn

ABSTRACT

Multivariate time series forecasting (MTSF) is an important and forefront task in many real-world applications. Recently, Mamba has emerged as a powerful alternative to Transformer-based models in sequential modeling, leveraging its selective mechanism to achieve high accuracy and linear complexity. However, existing efforts to apply Mamba to MTSF face limitations, particularly in effectively and efficiently capturing temporal and cross-variate dependencies. In this work, we propose DEMamba, an enhanced model tailored for MTSF. Specifically, we redefine the roles of 1D convolution and selective SSMs (S6) and design the Time-Variable Decoupled Scanning (TVDS) Mamba block. It decouples the learning of different dimensions within a single Mamba block, with 1D convolution capturing temporal dependencies and S6 capturing cross-variate dependencies. Furthermore, we capture intra-patch feature dependencies with a feed-forward network. These designs significantly enhance the ability of DEMamba to handle complex dependencies. Extensive experimental results on eight real-world datasets demonstrate the effectiveness of DEMamba against previous state-of-the-arts. Code is available at this repository: https://github.com/LiuYang826/DEMamba.

KEYWORDS

Multivariate Time Series Forecasting, State Space Models, Selective Mechanism

ACM Reference Format:
Junluo Zheng, Yang Liu, and Jianyong Chen. 2026. DEMamba: Decoupled Enhanced State Space Models with Selective Mechanisms for Multivariate Time Series Forecasting: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 2 pages. https://doi.org/10.65109/UIXP4140

1 INTRODUCTION

Mamba's linear complexity makes it promising for multivariate time series forecasting (MTSF), but adapting it for both temporal and cross-variate dimensions remains challenging. While some approaches treat dimensions separately [1] or flatten them [2], they compromise efficiency or effectiveness. We observe that 1D convolution alone suffices for temporal dependencies with proper patching, while selective SSMs (S6) better handle cross-variate dependencies. This motivates DEMamba with a Time-Variable Decoupled Scanning (TVDS) block that allocates 1D convolution to temporal and S6 to cross-variate modeling, inspired by PatchTST [3].

Contributions: (1) DEMamba framework for MTSF; (2) TVDS block decoupling temporal and cross-variate dependencies; (3) State-of-the-art performance on eight datasets.

*Corresponding author.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/UIXP4140

2 DEMAMBA FRAMEWORK

DEMamba employs a patch-based architecture with Instance Normalization. The input is embedded into tokens $\boldsymbol{H} \in \mathbb{R}^{V \times P \times D}$ (V: variables, P: patches, D: dimension) and processed by an n-layer encoder. Each layer features a TVDS (Time-Variable Decoupled Scanning) Mamba block that explicitly separates temporal modeling (via 1D convolution) from cross-variate interactions (via bidirectional selective SSMs), avoiding the flattening strategy of vanilla Mamba (Figure 1).

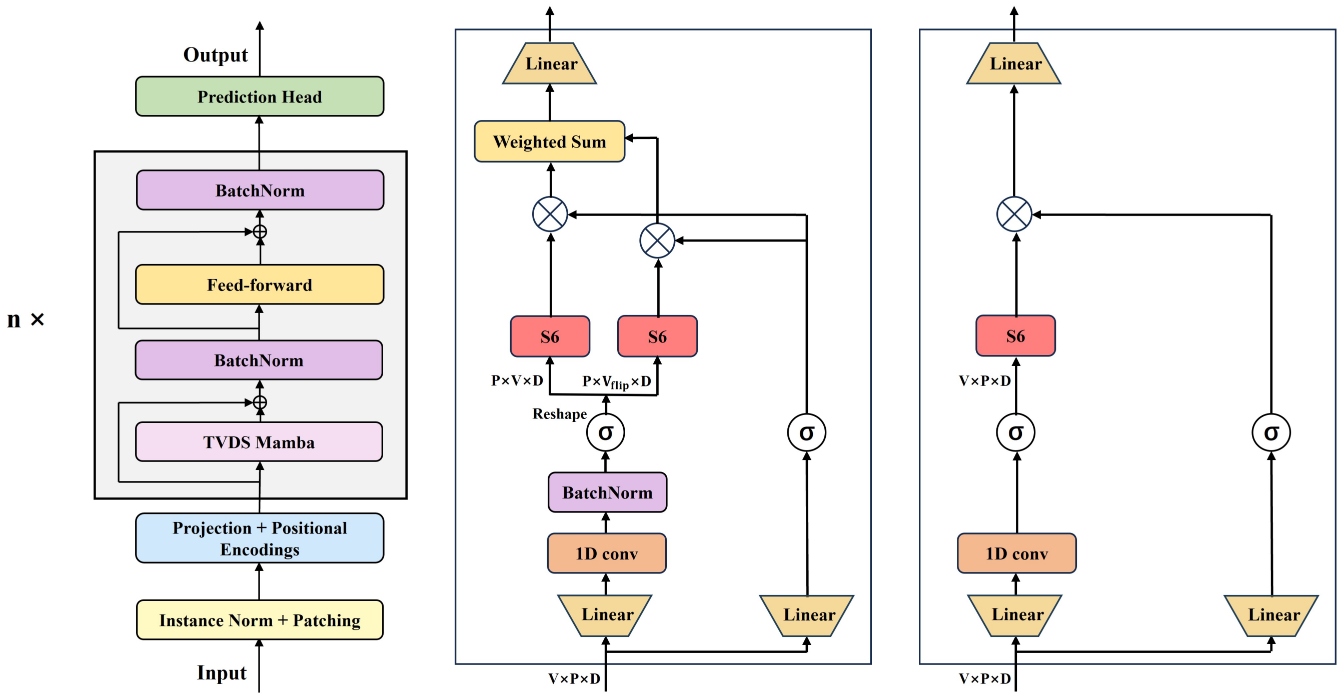
3 EXPERIMENTS AND RESULTS

3.1 Experimental Setup

We evaluate DEMamba on eight standard time series forecasting benchmarks: ETTh1, ETTh2, ETTm1, ETTm2, Weather, ECL, Traffic, and Exchange. We benchmark DEMamba against state-of-the-art methods including MambaTS, iTransformer. We use both MSE and MAE metrics for evaluation. We set the lookback window fixed at $T = 96$ with prediction horizons $K \in \{12, 24, 48, 96\}$ for Exchange dataset and $K \in \{96, 192, 336, 720\}$ for the others.

3.2 Main Results

As shown in Table 1, DEMamba demonstrates superior forecasting performance across diverse datasets. Our model achieves the best average MSE and MAE scores in 7 out of 8 datasets, obtaining the highest 1^{st} Count of 7 for both metrics. The competitive baselines exhibit distinct limitations: MambaTS, while effectively capturing multivariate correlations, shows compromised performance on several datasets due to insufficient temporal modeling. iTransformer, though designed for multivariate interactions, consistently underperforms compared to both DEMamba and MambaTS across all datasets. Notably, on the high-dimensional Traffic dataset,



(a) Overall Structure (b) TVDS Mamba Block (c) Mamba Block

Figure 1: The architecture of DEMamba. (a) The multivariate time series data is segmented into patches, which are then embedded into patch tokens. These patch tokens are fed into n layers of encoder to capture temporal, cross-variate, and feature dependencies. (b) The TVDS Mamba block captures temporal dependencies with 1D convolution. The data is then reshaped and two S6 modules are utilized to capture cross-variate dependencies in both forward and backward directions, due to the non-causal nature of the variable dimension. (c) In the original Mamba block, 1D convolution and S6 module capture dependencies along the same dimension of the data.

Table 1: MTSF results with prediction lengths $K \in \{12, 24, 48, 96\}$ for Exchange and $K \in \{96, 192, 336, 720\}$ for others. The best results are highlighted in bold. 1st *Count* indicates the number of times each method achieves the best results.

Models		DEMamba (Ours)		MambaTS (2024)		iTransformer (2024)	
Metric		MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	Avg.	**0.376**	**0.391**	0.396	0.404	0.408	0.412
ETTm2	Avg.	**0.270**	**0.318**	0.285	0.331	0.291	0.335
ETTh1	Avg.	**0.437**	**0.433**	0.447	0.444	0.458	0.450
ETTh2	Avg.	**0.376**	**0.399**	0.381	0.406	0.383	0.406
ECL	Avg.	**0.164**	**0.262**	0.184	0.273	0.175	0.266
Exchange	Avg.	**0.042**	**0.132**	0.043	0.135	0.044	0.138
Traffic	Avg.	0.477	0.295	**0.422**	**0.276**	0.423	0.283
Weather	Avg.	**0.241**	**0.273**	0.258	0.280	0.259	0.279
1st Count		**7**	**7**	1	1	0	0

DEMamba shows slightly lower performance than MambaTS, indicating potential areas for further improvement in extremely high-dimensional scenarios. These results collectively demonstrate that DEMamba's decoupled temporal and cross-variate modeling approach achieves more robust performance across diverse dataset characteristics compared to existing state-of-the-art methods.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant 62376163.

REFERENCES

[1] Ali Behrouz, Michele Santacatterina, and Ramin Zabih. 2024. Mambamixer: Efficient selective state space models with dual token and channel selection. arXiv:2403.19888 [cs.CV] arXiv preprint.
[2] Xiuding Cai, Yaoyao Zhu, Xueyao Wang, and Yu Yao. 2025. MambaTS: Improved Selective State Space Models for Long-term Time Series Forecasting. https://openreview.net/forum?id=vEtDApqkNR Under review at ICLR 2025.
[3] Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. 2023. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. In *The Eleventh International Conference on Learning Representations (ICLR)*. International Conference on Learning Representations (ICLR), Virtual, –. https://openreview.net/forum?id=Jbdc0vTOcol