

DOI:10.19651/j.cnki.emt.2518489

# Bi-EMamba: 基于 Mamba 的高压 电缆接地电流预测模型\*

周景董晖刘心唐振洋

(华北电力大学控制与计算机工程学院 北京 102206)

**摘要:** 高压电缆接地电流是保障电缆运行安全与稳定的关键指标,其精准预测对于故障预防和提升电网可靠性具有重要意义。针对传统时序预测模型在预测精度和计算效率上的局限性,本文提出了一种基于 Mamba 架构的接地电流预测模型——Bi-EMamba 模型。模型通过构建时空依赖关系编码器,显著增强了对多变量时间序列中长期依赖关系和空间关联性的建模能力,同时提升了模型的记忆性能。为应对非平稳数据问题,模型引入可逆归一化进行数据归一化处理,并通过超参数优化进一步提升了模型预测精度和泛化能力。基于北京某高压电缆线路数据集的实验结果表明,Bi-EMamba 在多种预测步长下均显著优于现有基准模型,尤其在长期预测场景中展现出更强的泛化性和计算效率。与当前 SOTA 模型 iTransformer 相比,Bi-EMamba 的均方误差降低了 6.52%,平均绝对误差降低了 3.21%,内存使用量降低了 29.49%。

**关键词:** 接地电流;多变量时序预测;状态空间模型;时空特征;深度学习

**中图分类号:** TP391.4;TM615;TN929.5 **文献标识码:** A **国家标准学科分类代码:** 470.4030

## Bi-EMamba: A Mamba-based prediction model for ground currents in high-voltage cable

Zhou Jing Dong Hui Liu Xin Tang Zhenyang

(School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

**Abstract:** Ground current in high-voltage cable systems serves as a critical indicator for ensuring operational safety and stability. Accurate ground current prediction is crucial for fault prevention and enhancing grid reliability. To address the limitations of traditional time-series prediction models in terms of prediction accuracy and computational efficiency, this paper proposes a ground current prediction model based on the Mamba architecture, referred to as the Bi-EMamba model. Through a spatiotemporal dependency encoder, the model effectively captures long-term dependencies and spatial correlations in multivariate time series while maintaining high memory efficiency. To address the challenge of non-stationary data, the model incorporates Reversible Instance Normalization for data normalization and employs hyperparameter optimization to further improve prediction accuracy and generalization capability. Experimental results based on a dataset from a high-voltage cable line in Beijing demonstrate that Bi-EMamba outperforms existing benchmark models across various prediction horizons. Notably, in long-term forecasting scenarios, it exhibits superior generalization and computational efficiency. Compared to the current state-of-the-art model, iTransformer, Bi-EMamba achieves a 6.52% reduction in Mean Squared Error, a 3.21% reduction in Mean Absolute Error, and a 29.49% reduction in memory usage.

**Keywords:** ground current; multivariate time series prediction; state space model; spatio-temporal features; deep learning

## 0 引言

高压电缆是城市电力传输的重要基础设施,其稳定性

直接关系到电网的安全运行。高压电缆通常由金属护层、绝缘层和线芯导体组成,在运行过程中,交流电流会在外部金属护层上产生感应电压。为消除这一电压,电缆的金属

收稿日期:2025-04-01

\* 基金项目:国家自然科学基金(62301220)项目资助

护层需定点接地,从而在护层与大地之间形成闭合回路,产生的接地电流<sup>[1]</sup>。接地电流是高压电缆状态监测的关键指标,异常的接地电流可能导致能量损耗、漏电,甚至短路或火灾。因此,国家电网《高压电缆状态检测技术规范》对接地电流设定了严格的限值。

然而,现有监测方法多依赖固定阈值判断,仅适用于静态场景,难以适应复杂运行环境的动态变化,导致提前预警能力有限。随着城市电网规模的扩大,对接地电流的智能预测提出了更高的要求,促使学术界和工业界探索高效、实时的时序预测方法。

传统接地电流预测主要依赖电力信号仿真模型<sup>[2-4]</sup>,通过建立电缆运行状态的数学描述来模拟电流变化规律。然而,这类方法难以捕捉电力系统的非线性特征,对环境噪声敏感,且缺乏对突发事件的适应性。近年来,随着在线监测设备的广泛部署,电网系统积累了海量高频监测数据,其特性包括时空相关性强、非平稳性显著,这为基于深度学习的时序预测提供了数据支持,同时也带来了新的建模挑战。

深度学习模型已广泛应用于电力系统时间序列预测<sup>[5-7]</sup>。循环神经网络(recurrent neural network, RNN)和长短期记忆网络(long short-term memory network, LSTM)在短期预测任务中表现良好,但受限于梯度消失和计算效率,难以建模长时间依赖关系。Transformer 通过自注意力机制增强了全局建模能力,并具有良好的并行计算特性,然而,其计算复杂度为  $O(N^2)$ ,导致长序列预测的计算开销过高,难以满足电力系统对实时性的需求。

为了解决 Transformer 计算复杂度高的问题,近年来研究者们提出了一系列优化方案。例如,金字塔注意力机制(Pyraformer<sup>[8]</sup>)和时间序列分解(Autoformer<sup>[9]</sup>)减少了计算开销;指数平滑注意力(ETSformer<sup>[10]</sup>)和多尺度 Transformer(DRFormer<sup>[11]</sup>)优化了不同时间尺度的特征提取;分段策略(PatchTST<sup>[12]</sup>、Crossformer<sup>[13]</sup>)降低了输入序列长度的影响。然而,这些方法但仍未能根本性突破长序列建模的计算瓶颈。为此,iTransformer<sup>[14]</sup>通过反转注意力层建立序列间直接依赖关系,但其基于多层感知机(multilayer perceptron, MLP)进行序列映射的方式忽略了时间序列的内部动态演化特性。综上,Transformer 及其变体在计算效率与预测性能之间仍然存在权衡,促使研究者探索计算更高效的替代方案。

为降低计算开销,研究者们提出一系列线性时间序列预测模型,如 DLinear<sup>[15]</sup>和 TiDE<sup>[16]</sup>,它们通过直接建模时间序列的线性趋势,显著降低了计算复杂度。然而,在电力系统的高波动性、非周期性和非平稳时间序列中,这些模型仍面临显著挑战,主要体现在:局部模式捕捉能力不足,难以准确预测短期突变;对长期趋势的刻画较弱,难以有效建模复杂时间依赖;MLP 结构在捕捉全局模式方面表现有限。此外,相较 Transformer,MLP 需要更长的历史窗口输入,这对数据完整性提出了更高的要求<sup>[17]</sup>。

近期提出的 Mamba 模型<sup>[18]</sup>基于选择性状态空间模型

(selective structured space model, selective SSM),采用状态空间更新机制代替 Transformer 的自注意力,从而以  $O(N)$  的计算复杂度建模长距离依赖关系。研究者们针对预测任务围绕其结构也进行了多项改进,例如,C-Mamba<sup>[19]</sup>通过通道注意力机制增强不同变量间的交互;DTMamba<sup>[20]</sup>采用通道独立策略并引入双 Mamba 块,以提升长序列预测能力;UMamba<sup>[21]</sup>结合 U 形结构,在不同尺度上提取时间序列特征,提升泛化能力。尽管 Mamba 及其变体已在金融市场、加密货币以及电力系统等领域展现了良好的性能,但在高压电缆接地电流预测任务中仍面临以下挑战:接地电流数据受环境因素影响显著,存在漂移和突变,大多数模型未针对非平稳时间序列设计归一化策略,可能导致预测不稳定;多个接地电流监测点之间存在复杂空间依赖,但现有模型未能充分建模这一特性,影响预测精度;部分变体增加了计算复杂度,处理大规模监测数据的高计算成本限制了实时应用的可行性。

针对上述挑战,本研究提出了一种基于 Mamba 的高压电缆接地电流预测模型——Bi-EMamba,旨在提升预测精度、计算效率及泛化能力,以满足高压电缆接地电流预测的实时性要求。首先,针对非平稳数据,引入可逆归一化层对输入数据进行通道独立标准化,提高模型对分布变化的适应能力。其次,为提升对短期动态特征的捕捉能力,Bi-EMamba 采用分段机制,使模型能够更精准地建模短时间内的突变模式,提高预测精度。此外,为同时建模多个监测点之间的空间关联性和长时间依赖关系,Bi-EMamba 设计了时空依赖关系编码器(spatiotemporal dependency encoding)。该编码器使用了动态排列层,以增强监测点之间的时空依赖建模能力,并通过堆叠双向 EMamba 模块提高长时序建模的精度,编码器还引入遗忘门结构以调节历史信息的保留比例,从而减少信息丢失,优化长期趋势建模。最后,使用独立预测层对各通道数据进行解码,从而减少变量间的耦合效应,提高预测的准确性和稳定性。

## 1 方 法

### 1.1 接地电流监测数据特性分析

高压电缆接地电流数据具有显著的非平稳性,即其统计特性(均值、方差、周期性等)随时间变化<sup>[22]</sup>。这种非平稳性主要源于外部环境因素(如温度、湿度、电网负荷波动)和设备运行状态的动态变化。对于深度学习预测模型而言,非平稳性可能导致训练过程中数据分布的漂移,影响模型的泛化能力和预测精度。

图 1 展示了北京某高压电缆线路 6 个监测点的全年接地电流数据。从图中可以观察到,不同监测点的接地电流波动幅度存在较大差异,并在某些时间段(如夏季 6 月~7 月)出现显著峰值。这些峰值的幅度可达平稳电流值的 2~3 倍,与外部环境变化密切相关。此外,某些监测点的电流波动模式较为一致,表明数据具有空间相关性,即某些

监测点的电流变化可能受到周围设备状态的影响<sup>[23]</sup>。

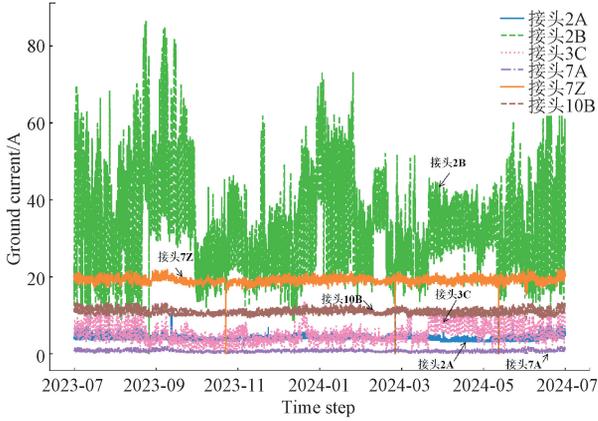


图 1 年度高压电缆接地电流波动特性图

Fig. 1 Annual ground current fluctuation characteristics of high-voltage cable

进一步分析局部时间窗口内的接地电流特性(如图 2),可以发现不同相位(A、B、C、Z 相)之间的电流数据存在明显的周期性波动模式。此外,不同接头处的接地电流数据呈现出较强的时间相关性,即当前时刻的电流值与过去的趋势密切相关,进一步增加了预测的难度。同时,不同接头(如接头 1 与接头 5)之间的接地电流数据也表现出一定的相关性,表明接地电流不仅具有时间相关性,还存在明显的空间相关性,这种关联性主要源于线路结构特性及电缆系统内各单元的相互作用。

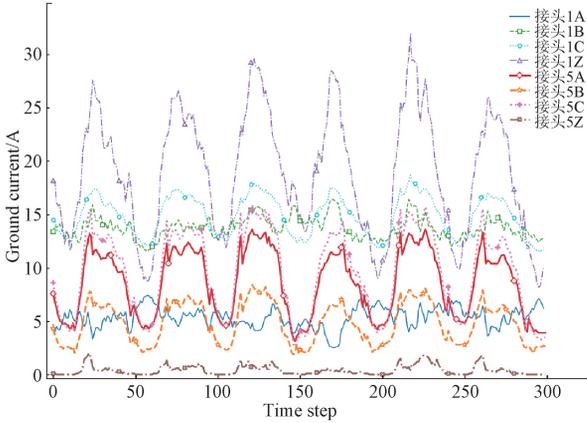


图 2 300 个时间步内两个接头处接地电流的监测数据

Fig. 2 The monitoring data of grounding current at two joints within 300 time steps

综上,接地电流数据的非平稳性、时间相关性和空间相关性相互交织,对预测模型的建模能力提出了更高要求。因此,构建能够同时考虑这些特性的预测模型成为研究的关键问题。

### 1.2 模型基础

#### 1) 多变量时间序列预测问题描述

给定一个多变量时间序列  $\mathbf{X}_{in} = \{\mathbf{x}_1, \dots, \mathbf{x}_L\} \in$

$\mathbb{R}^{L \times M}$ , 其中  $\mathbf{x}_i = \{p_{i1}, \dots, p_{iM}\}$  表示第  $i$  个时间步的  $M$  个变量值。多变量时间序列数据预测的目标是基于长度为  $L$  的历史观测,预测未来数据  $\mathbf{X}_{out} = \{\mathbf{x}_{L+1}, \mathbf{x}_{L+2}, \dots, \mathbf{x}_{L+H}\} \in \mathbb{R}^{H \times M}$ 。理想情况下,预测结果应尽可能接近真实值。其中,  $L$  表示历史回溯窗口长度,  $p$  为单变量,  $M$  为变量总数,  $H$  为预测视野,即未来的预测步数。

#### 2) 状态空间模型

状态空间模型(structured space model, SSM)的灵感来自连续系统,使用一阶微分方程通过隐藏状态  $h(t)$  将输入函数  $x(t)$  映射到输出函数  $y(t)$ , 定义如下<sup>[18]</sup>:

$$h'(t) = \mathbf{A}h(t) + \mathbf{B}x(t) \quad (1)$$

$$y(t) = \mathbf{C}h(t) \quad (2)$$

其中,参数  $\mathbf{A} \in \mathbb{R}^{N \times N}$ ,  $\mathbf{B} \in \mathbb{R}^{D \times N}$ ,  $\mathbf{C} \in \mathbb{R}^{N \times D}$ 。

连续参数  $\mathbf{A}$ 、 $\mathbf{B}$  可以通过零阶保持和以  $\Delta$  为间隔的时间采样离散为  $\bar{\mathbf{A}}$ 、 $\bar{\mathbf{B}}$ , 定义如下:

$$\bar{\mathbf{A}} = \exp(\Delta \mathbf{A}) \quad (3)$$

$$\bar{\mathbf{B}} = (\Delta \mathbf{A})^{-1} (\exp(\Delta \mathbf{A}) - \mathbf{I}) \cdot \Delta \mathbf{B} \quad (4)$$

离散化后 SSM 的计算公式可以写成:

$$h_k = \bar{\mathbf{A}}h_{k-1} + \bar{\mathbf{B}}x_k \quad (5)$$

$$y_k = \mathbf{C}h_k \quad (6)$$

离散化的 SSM 以卷积运算的方式进行并行训练,并以递归神经网络的方式进行有效推理。并通过将 HIPPO 矩阵<sup>[24]</sup>引入参数矩阵  $\mathbf{A}$  的初始化,得到了 SSM 的一种变体,即结构化状态空间模型(structured state space sequence model, S4)<sup>[25]</sup>,提高了对长期依赖关系的建模能力,但它要求可学习矩阵  $\mathbf{A}$ 、 $\mathbf{B}$ 、 $\mathbf{C}$  以及  $\Delta$  在所有时间步长中保持不变。

#### 3) Mamba 模型

为解决 S4 模型中缺乏选择性和序列建模中的效率瓶颈,近期研究提出了 Mamba 模型。Mamba 在 S4 模型中引入了一种选择机制,允许可学习参数  $\mathbf{A}$ 、 $\mathbf{B}$ 、 $\mathbf{C}$  以及  $\Delta$  动态地与输入序列动态交互,并在其循环模式中采用了硬件感知并行算法。因此,Mamba 可以根据输入选择性地记忆或忽略信息,同时保持对序列长度的线性扩展能力。

Mamba 的核心是 Selective SSM,见图 3<sup>[26]</sup>。Mamba 模型输入序列形状为  $B \times M \times D$ , 其中  $B$  表示批量大小、 $M$  表示变量的数量,  $D$  表示隐藏状态维度。首先,模型通过线性投影将输入映射到更高维度  $E \times D$ , 生成中间表示  $\mathbf{x}$  和  $\mathbf{z}$ 。随后,模型使用卷积操作和激活函数对  $\mathbf{x}$  进行处理,生成新的投影表示  $\mathbf{x}'$ 。通过结合  $\mathbf{x}'$  和 Selective SSM,生成状态表示  $\mathbf{y}$ 。最后,  $\mathbf{y}$  与激活后的残差连接  $\mathbf{z}$  结合,通过线性变换获得时间步  $t$  的最终输出  $y_t$ 。残差连接有助于保留输入中的关键信息,从而增强序列特征的完整性。Mamba 的主要参数包括 Selective SSM 的状态扩展因子  $N$ , 卷积核大小  $k$ , 以及用于输入输出线性投影的块扩展因子  $E$ 。  $N$  和  $E$  的值越大则计算成本越大, Mamba 模型的最终输出为  $\mathbf{Y} \in \mathbb{R}^{B \times M \times D}$ 。

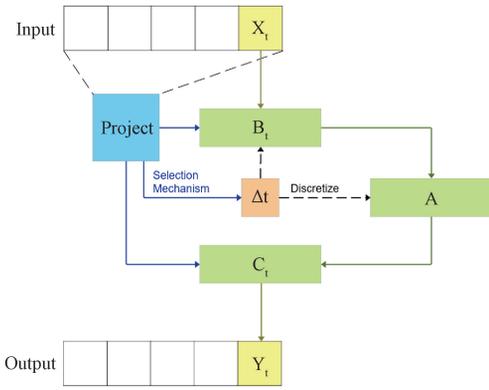


图 3 Selective SSM 模型结构

Fig. 3 The structure of Selective SSM

## 2 模型架构

本节将详细介绍本文提出的 Bi-EMamba 模型。图 4 展示了整体结构,主要由可逆归一化层、分段层、时空依赖关系编码器和预测层组成。

### 2.1 可逆归一化层

为应对接地电流数据集的非平稳性问题,使用可逆归一化方法(reversible instance normalization, RevIN)<sup>[27]</sup>,将每个输入通道标准化为零均值和单位标准差,并在输出时进行反归一化,确保归一化前后的特征数量一致。常规的归一化操作会影响模型获取输入数据的原始分布,从而降低模型最终的预测效果,但可逆归一化方法将归一化操作删除的非平稳信息返回给模型,这样就有效处理了非平稳

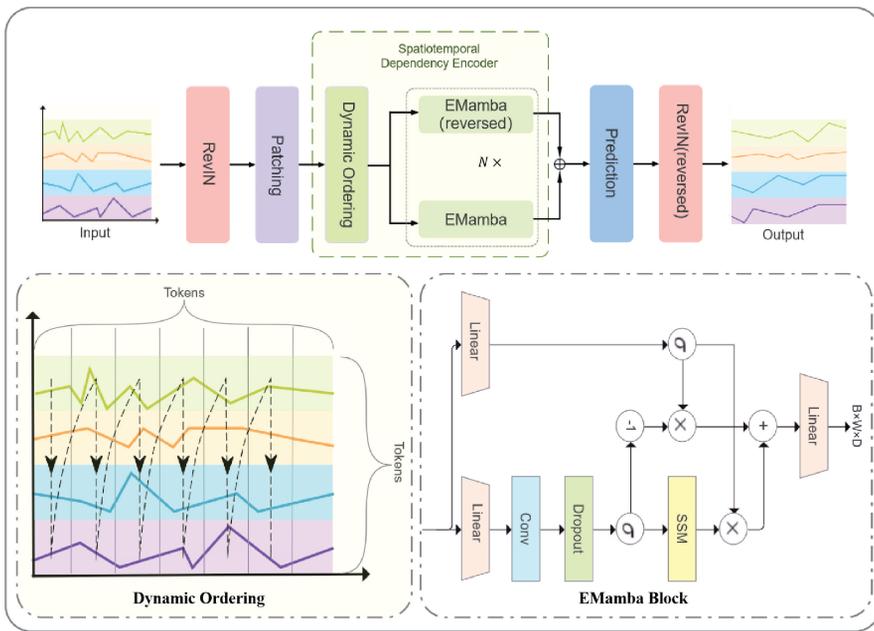


图 4 Bi-EMamba 的整体架构

Fig. 4 The overall architecture of Bi-EMamba

时间序列的数据分布漂移。

### 2.2 分段层

由于单纯建立时间点之间的关系难以充分提取序列中的语义信息,采用了一种分段机制(patching),以捕捉接地电流序列中的局部演化模式。受到 PatchTST<sup>[12]</sup>启发,对每个输入的单变量时间序列  $X_t$  进行分段,时间步长为  $S$ ,将每个变量分割成大约  $J = L/S$  个切片  $p$  (patch),  $p^i \in \mathbb{R}^{J \times P}$ ,其中  $J$  是切片总数, $P$  为每个切片的长度, $J = \lfloor \frac{L-P}{S} + 1 \rfloor$ 。这种分段策略使输入切片的数量从  $L$  减少到大约  $L/S$ ,从而减少了内存使用和计算复杂度。分段机制使得模型可以观察到更长的历史序列,提升了预测性能。此外,每个切片聚合来自相邻时间点的信息,模型可以更好地捕捉到接地电流序列的局部演化模式,同时减少

内存使用并去除冗余信息。

将每个切片通过式(7)所示的线性变换映射到  $D$  维空间。

$$p_{out} = Linear( Batch(p) ) \quad (7)$$

### 2.3 时空依赖关系编码器

为了能够充分利用高压电缆各监测点间接地电流的时空相关性,设计一种时空依赖关系编码器(spatiotemporal dependency encoder),包括动态排列层(dynamic ordering)和堆叠的双向 EMamba 模块,具体实现如算法 1 和算法 2 所示。

动态排列层通过使各变量的切片在每个时间步上交替排列,构建了一个全面且紧密联系的整体回溯窗口。这种结构化的组织使得模型能够更准确地捕获接地电流数据中的长期依赖关系和监测点间的空间相关性。经过动

算法 1: EMamba Block 的实现

Algorithm1: Implementation of EMamba block

Input: Patch-wise token sequence  $\mathbf{E}_x : (J, M, D)$

Output:  $\mathbf{E}_y : (J, M, D)$

```

1  $\mathbf{x} : (J, M, D) \leftarrow \text{Linear}^x(\mathbf{E}_x)$ 
2  $\mathbf{z} : (J, M, D) \leftarrow \text{Linear}^z(\mathbf{E}_x)$ 
3  $\mathbf{x}' : (J, M, D) \leftarrow \text{SiLU}(\text{Conv1D}(\mathbf{x}))$ 
4  $\mathbf{x}' : (J, M, D) \leftarrow \text{Dropout}(\mathbf{x}')$ 
5  $\mathbf{A} : (D, N) \leftarrow \text{Parameter}^A$ 
6  $\mathbf{B} : (J, M, N) \leftarrow \text{Linear}^B(\mathbf{x}')$ 
7  $\mathbf{C} : (J, M, N) \leftarrow \text{Linear}^C(\mathbf{x}')$ 
8  $\mathbf{\Delta} : (J, M, D) \leftarrow \tau_A(\text{Parameter}^A + \text{Linear}^A(\mathbf{x}'))$ 
9  $\bar{\mathbf{A}}, \bar{\mathbf{B}} : (J, M, D, N) \leftarrow \text{discretize}(\mathbf{\Delta}, \mathbf{A}, \mathbf{B})$ 
10  $\mathbf{y} : (J, M, D) \leftarrow \text{SSM}(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \mathbf{C})(\mathbf{x}')$ 
11  $\mathbf{y}' : (J, M, D) \leftarrow \mathbf{y} \otimes \text{SiLU}(\mathbf{z}) + \mathbf{x}' \otimes (1 - \sigma(\mathbf{z}))$ 
12  $\mathbf{E}_y : (J, M, D) \leftarrow \text{Linear}^{y'}(\mathbf{y}')$ 
13 return  $\mathbf{E}_y$ 

```

算法 2: 时空依赖关系编码器的实现

Algorithm2: Implementation of spatiotemporal dependency encoder

Input: Patch-wise token sequence  $\mathbf{E}_x^{(t)} : (J, M, D)$

Output:  $\mathbf{E}_x^{(t+1)} : (J, M, D)$

```

1  $\mathbf{E}_x^{(t)} : (J, M, D) \leftarrow \text{Ordering}$ 
2  $\overrightarrow{\mathbf{E}_x^{(t)}} : (J, M, D) \leftarrow \overrightarrow{\text{EMamba}}(\mathbf{E}_x^{(t)})$ 
3  $\overleftarrow{\mathbf{E}_x^{(t)}} : (J, M, D) \leftarrow \overleftarrow{\text{EMamba}}(\mathbf{E}_x^{(t)})$ 
4  $\mathbf{E}_x^{(t+1)} : (J, M, D) \leftarrow \overrightarrow{\mathbf{E}_x^{(t)}} + \overleftarrow{\mathbf{E}_x^{(t)}}$ 
5 return

```

态排列层后的切片输出为  $\mathbf{E}_{mix} \in \mathbb{R}^{J \times M \times D}$ 。

原始的 Mamba 模型通过两个分支处理输入特征。第 1 个分支  $F_1$  将输入特征传递给 1 维卷积模块和选择性 SSM 块,以捕捉长时间依赖关系;第 2 个分支  $F_2$  传递输入特征至 SiLU 激活函数,用于生成门控信息。然而,由于选择性 SSM 块的输出直接受到门控层的影响,模型倾向于优先处理较近的输入信息,从而限制了历史信息的保留能力<sup>[15]</sup>。为此,本文设计了一个新的 EMamba 模块。

具体而言,引入遗忘门(forget gate)结构,其计算公式为:

$$\text{gate}_f = 1 - \text{gate}_{F_2} \quad (8)$$

遗忘门通过与  $F_1$  中 1 维卷积模块的输出相乘,将结果与选择性 SSM 的输出相加,从而实现新特征和历史特征的选择性组合。相比直接过滤新特征的方式,这样可以使模型有效地保留更长时间范围内的历史信息。此外,在 SSM 块前增加 dropout 层,以降低过拟合风险并提高模型的泛化性能。

原始 Mamba 模型仅能单向处理一维序列,而接地电

流序列数据具有双向模式及空间依赖关系。因此,采用两个 EMamba 模块分别处理正向和反向信息,从而同时建模序列的双向关联性。编码器包含多个堆叠的双向 EMamba 模块,增强模型对时空特征的捕捉能力。

## 2.4 预测层

在时空依赖关系编码器之后,使用独立的预测层对每个通道进行解码。为了提高解码效率,采用类似于 PatchTST 的方法,使用单通道独立解码策略,即每个通道通过一个简单的线性预测头单独解码。最后,经过逆归一化操作,将预测结果恢复到与输入数据一致的分布范围内。

## 3 实验准备

### 3.1 实验参数设置

本实验在 Python 3.10 环境下实现,操作系统为 Ubuntu 20.04, GPU 为 NVIDIA RTX 4060。实验过程中采用了同类深度学习模型的超参数设置作为参考。表 1 列出了各参数的取值和调优范围,其中主要超参数包括激活函数、学习率、损失函数、丢弃概率、窗口大小和预测步长等。这些参数的选择旨在提升模型的预测性能,同时确保模型的训练稳定性和收敛速度。

表 1 实验超参数设置

Table 1 Experimental hyperparameter settings

参数名	参数取值
激活函数(activation)	gelu
优化器(optimizer)	Adam
学习率(learning rate)	0.001, 0.0005
损失函数(loss)	MSE
参数丢弃系数	0.2
窗口大小	[48, 96, 192, 336, 720, 1440]
预测步长	[48, 96, 192, 336, 720]
维度(dimension)	[16, 32, 64, 128, 256, 512]
encoder layers	3
decoder layers	2
dimension of FCN	2048
n_head	32
切片长度	[24, 48, 60, 72, 120, 240]

### 3.2 数据集

实验数据来源于北京某高压电缆线路的接地电流监测系统,采集时间范围覆盖 2023 年 7 月 1 日~2024 年 6 月 30 日,为期一年,并以每 30 min 一次的频率进行了连续的数据采集。线路采用交叉互连接地方式,并在沿线设置了 12 个接地监测点位,每个点位均配备了能够记录 A、B、C 以及 Z 相接地电流的传感器。由此形成了由 48 个独立通道构成的多变量时间序列数据集。

在数据处理过程中,对于监测设备故障或人为操作导

致的数据缺失以及较大的突变点,采用线性插值法进行填充,以确保数据的完整性。数据集划分采用留出法,分为训练集、验证集和测试集,比例为 7:2:1。训练集用于模型参数优化,验证集用于模型性能评估并调整参数,最终在测试集上验证模型的泛化能力。训练过程持续 10 轮,并引入提前停止机制,即在验证集上若三轮未见误差下降则停止训练,避免过拟合。

### 3.3 损失函数

为了衡量预测值与真实值之间的差距,使用均方误差(mean squared error, MSE)损失。MSE 是一种常用的回归损失,能够有效地惩罚较大的预测误差,有助于提高预测的整体精度。损失函数定义如下:

$$\mathcal{L}(Y, \hat{Y}) = \frac{1}{|Y|} \sum_{i=1}^{|Y|} (y_{(i)} - \hat{y}_{(i)})^2 \quad (9)$$

### 3.4 基准方法

为了验证本文提出的 Bi-EMamba 模型的有效性,将

其与时序预测领域的 7 种代表性模型进行对比,包括 iTransformer、Autoformer、Crossformer、FourierGNN<sup>[28]</sup>、MICN<sup>[29]</sup>、PatchTST 和 S-Mamba<sup>[26]</sup>。这些方法在长序列预测、多变量时序建模等方面具有较强的代表性。

## 4 实验结果与分析

### 4.1 对比实验

本节展示了不同模型在高压电缆线路接地电流预测任务中的性能表现。为了评估各模型在中长期预测任务中的适应能力,采用了多种预测步长( $Pre = H = 48, 96, 192, 336, 720$ )。所有模型的回溯窗口长度设为  $L = 720$ ,即利用过去 15 天的接地电流数据预测未来不同天数(1 天、2 天、4 天、7 天、15 天)的接地电流。表 2 展示了各模型在不同预测步长下的 MSE 和平均绝对误差(mean absolute error, MAE)对比,其中最佳结果以粗体标注,次佳结果以下划线表示。

表 2 不同模型的接地电流预测性能对比  
Table 2 Classification metrics for each label

$Pre$	Bi-EMamba		S-Mamba		iTransformer		Autoformer	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
48	<b>0.510 26</b>	<b>0.447 36</b>	0.534 64	0.469 20	0.530 78	0.463 44	1.095 73	0.768 25
96	<b>0.626 87</b>	<b>0.517 62</b>	0.668 34	0.536 42	<u>0.656 82</u>	<u>0.531 83</u>	1.259 65	0.830 61
192	<b>0.831 17</b>	<b>0.610 29</b>	<u>0.843 20</u>	<u>0.617 29</u>	0.858 13	0.619 17	1.434 42	0.897 46
336	<b>0.961 50</b>	<b>0.663 81</b>	1.052 32	0.682 74	1.114 22	0.691 50	1.451 31	0.909 49
720	<b>1.130 52</b>	<b>0.732 25</b>	<u>1.202 18</u>	0.740 21	1.220 17	0.766 42	1.595 10	0.948 29
$Pre$	Crossformer		FourierGNN		MICN		PatchTST	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
48	0.534 81	0.463 68	0.604 97	0.524 05	<u>0.527 63</u>	<u>0.458 93</u>	0.558 08	0.493 10
96	0.682 49	0.543 18	0.711 11	0.580 60	0.667 23	0.535 74	0.680 13	0.556 54
192	1.013 87	0.651 54	0.848 06	0.648 58	0.855 44	0.626 90	0.855 43	0.626 89
336	1.233 31	0.790 87	1.362 28	0.695 35	<u>1.033 84</u>	0.705 19	1.082 87	<u>0.680 45</u>
720	1.404 36	0.839 27	1.501 89	0.873 26	1.227 52	0.784 59	1.211 12	<u>0.734 81</u>

表 2 中的实验结果显示,Bi-EMamba 模型在所有预测步长下的 MSE 和 MAE 均优于现有基线模型,尤其在长期预测(如  $Pre = 336$  和  $Pre = 720$ )中展现出更强的泛化能力和稳定性。与当前 SOTA 的 iTransformer 相比,Bi-EMamba 的 MSE 和 MAE 在各个预测步长上平均降低了 6.52% 和 3.21%,与 S-Mamba 相比,Bi-EMamba 的 MSE 和 MAE 在各个预测步长上平均降低了 5.36% 和 2.63%,表明其捕捉时序特征和提升预测精度方面具有明显优势。而与 Autoformer 相比,Bi-EMamba 的 MSE 和 MAE 降幅更为显著,分别达到了 41.72% 和 32.25%。

总体而言,Bi-EMamba 模型能够有效利用长时间序列的历史数据,表现出卓越的精度和稳定性,是接地电流长期预测领域的理想选择。

为更直观地展示模型的预测性能,图 5 提供了多步预测( $Pre = 96$ )条件下,某监测点接地电流预测结果的可视化对比。整体而言,各模型均能较好地拟合接地电流的变化趋势。然而,相较于其他模型,Bi-EMamba 在捕捉实际曲线的局部波动方面表现更优。

此外,图 6 展示了接地点 2 的四相接地电流,使用 Bi-EMamba 模型在相同预测步长下的预测结果。从整体趋势来看,模型能够有效拟合电流的主要波动模式。然而,在部分电流剧烈变化的区间,预测曲线仍存在一定的滞后或振幅缩减现象。

同时,从预测精度、内存使用和训练速度 3 个方面对各模型进行全面评估,实验配置中输入长度  $L$  和预测长度  $H$  均设为 720。如图 7 所示,Bi-EMamba 模型在各项指标

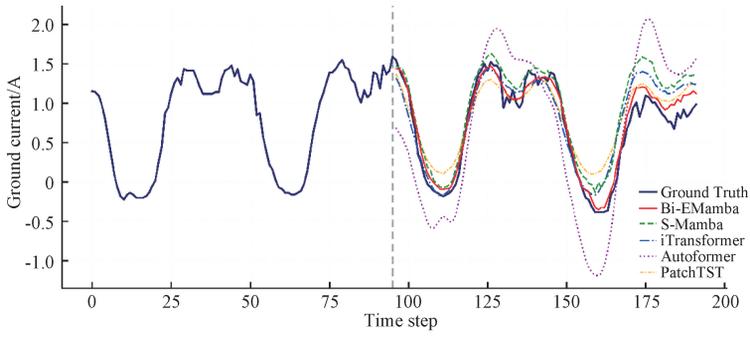


图 5 接地电流多步预测 ( $Pre=96$ ) 结果

Fig. 5 Multi-step ground current prediction results ( $Pre=96$ )

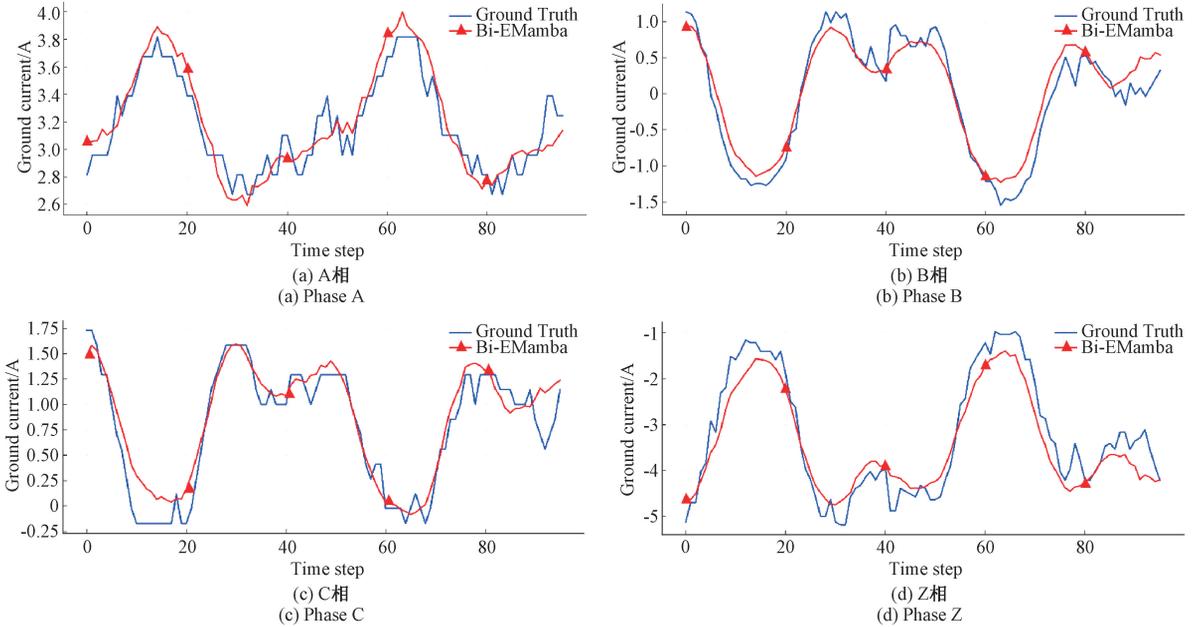


图 6 接地点 2 接地电流多步预测 ( $Pre=96$ ) 结果

Fig. 6 Multi-step prediction results of ground current at monitoring point 2 ( $Pre=96$ )

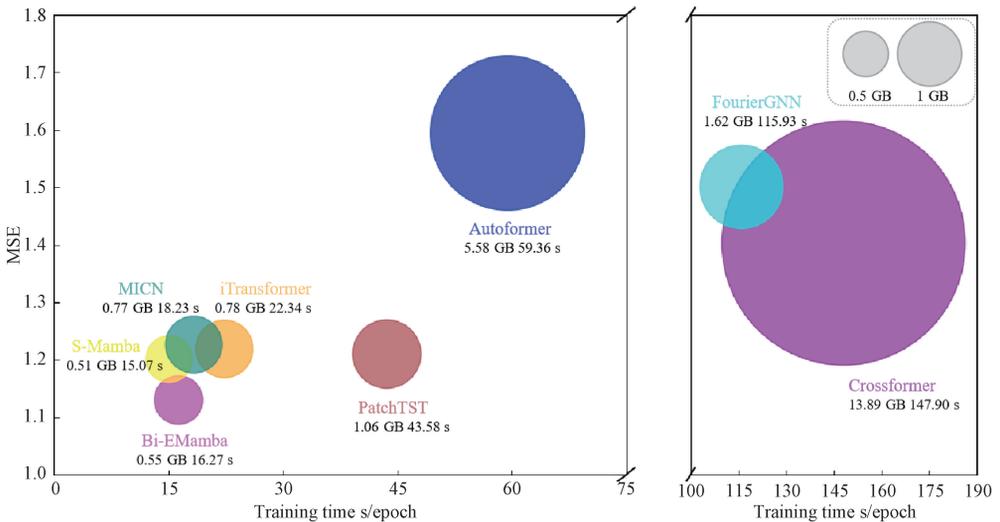


图 7 各模型效率比较

Fig. 7 Comparison of model efficiencies

中均表现优异,达到最佳预测精度,MSE 为 1.130 52。在内存使用方面,Bi-EMamba 展示了显著的资源优势,仅消耗 0.55 GB 内存,得益于其选择性状态空间模型所实现的线性复杂度设计。相比之下,Crossformer 和 Autoformer 的内存消耗分别为 13.89 GB 和 5.58 GB,远高于 Bi-EMamba,这使得 Bi-EMamba 在资源受限的环境中更具扩展性。在训练速度方面,Bi-EMamba 也展现出明显优势,平均每轮训练时间为 12.032 s。尽管 S-Mamba 的训练速度和内存效率有所优势,但其预测精度仍不及 Bi-EMamba。

综上所述,Bi-EMamba 模型凭借其低内存消耗、快速训练时间和高预测精度,在综合性能上表现突出。其基于 SSM 的线性复杂度设计尤其适用于大规模和实时预测任务,使其在计算和内存效率方面具有显著优势,完全满足接地电流预测的应用需求。

#### 4.2 消融实验

为验证本文提出的各个组件(双向结构、遗忘门、动态排列层)的合理性和有效性,进行多种消融实验,结果如表 3 所示。所有模型均采用  $L=720$ ,  $Pre=192$ 。

当模型同时包含双向结构、遗忘门和动态排列层时,MSE 和 MAE 分别达到 0.831 17 和 0.610 29,表现最佳。这表明双向结构与遗忘门的组合在捕捉接地电流长期依赖关系方面具有显著优势,而动态排列层进一步增强了模

表 3 消融实验结果

Table 3 Results of ablation experiments

模型	双向结构	遗忘门	动态排列层	MSE	MAE
本文模型	✓	✓	✓	0.831 17	0.610 29
模型 1	×	✓	✓	0.862 33	0.619 54
模型 2	✓	×	✓	0.852 51	0.626 14
模型 3	✓	✓	×	0.840 89	0.623 32
模型 4	×	×	×	0.863 02	0.619 51

型对监测点间空间关系的建模能力。该配置的最低误差验证了各组件协同作用下捕捉时序复杂模式的有效性。

另外,为验证 RevIN 层在缓解分布差异问题上的有效性,实验对所提方法在各个步骤中的训练和测试数据分布进行了分析,如图 8 所示。从图 8(a)~(b)可以观察到,相较于原始输入数据(图 8(a)),RevIN 归一化后的数据(图 8(b))显著降低了训练和测试数据的分布差异。这表明,实验中设置的归一化层能够有效削弱数据的非平稳性,从而减少由于数据分布不匹配导致的模型泛化误差。在标准化数据作为输入的情况下,模型能够较好地对齐训练数据和测试数据的分布(图 8(c)),从而提升预测稳定性。随后,通过 RevIN 反归一化操作,模型预测结果可以返回到原始数据分布(图 8(d)),保证最终输出数据仍然符合实际物理特性。

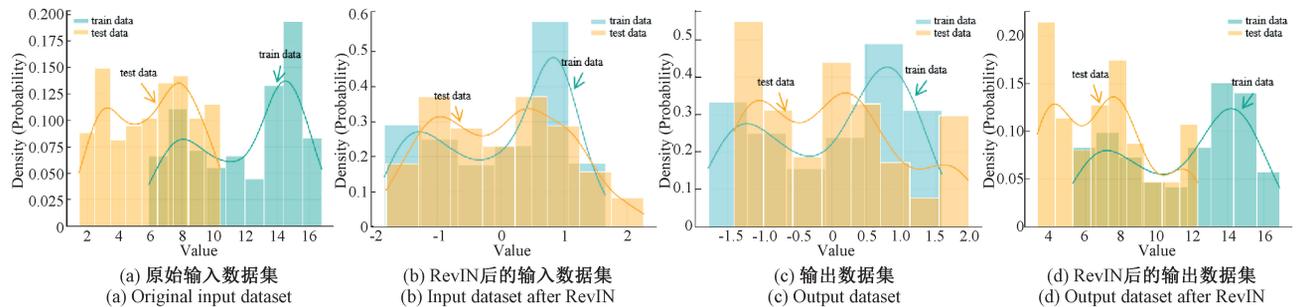


图 8 RevIN 对训练数据和测试数据分布差异的影响

Fig. 8 The impact of RevIN on the distribution between training and test data

#### 4.3 超参数分析

##### 1) 回溯窗口大小

先前的研究表明,基于 Transformer 的方法在回溯窗口不断增长的情况下,不一定能够充分受益<sup>[15]</sup>。这可能是因为在自注意力机制下,模型对时间序列的时序关系可能存在忽视。因此,针对回溯窗口大小对 Bi-EMamba 模型性能的影响进行评估,并将其与其他基线模型进行了对比。

实验结果如图 9 所示,随着输入序列的加长,Bi-EMamba 的预测精度不断提高,表现出良好的适应能力和对长时间序列的利用效果。iTransformer、MICN 和 PatchTST 也表现出一定的窗口增益能力,但整体效果不

如 Bi-EMamba,并且 iTransformer 随窗口增加展示出的增益是不连续的。综上所述,相比其他模型,本文提出的模型能够更好地利用长时间历史数据,适应较长的回溯窗口,适用于接地电流长期监测的预测任务。

##### 2) Dropout 层比率

图 10 展示了 Dropout 层在不同比率(0~0.5)下,模型预测性能的变化情况。可以观察到,与不使用 dropout 层相比,随着 dropout 率的增加,模型的 MSE 先下降后上升,在 dropout 率为 0.2 时取得了最佳性能。这表明适当的 dropout 率可以有效提高模型的泛化能力,降低预测误差。

图 11 展示了 dropout 率为 0.2 时的损失曲线。可以观察到,训练过程中,dropout 层有助于防止过早收敛和过

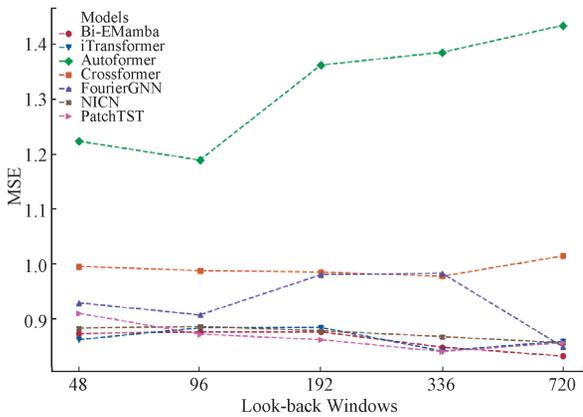


图 9 不同回溯窗口对性能的影响

Fig. 9 Impact of different look-back windows on performance

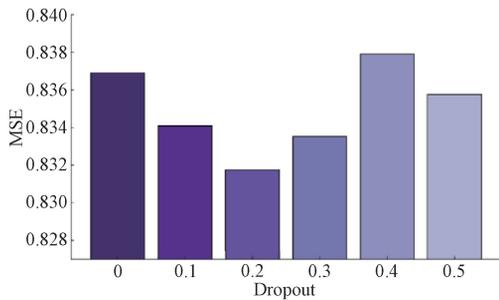


图 10 Dropout 层比率对性能的影响

Fig. 10 Impact of Dropout layer rates on performance

拟合;还有助于降低验证损失。该实验验证了 dropout 层在 Bi-EMamba 模型中的有效性。

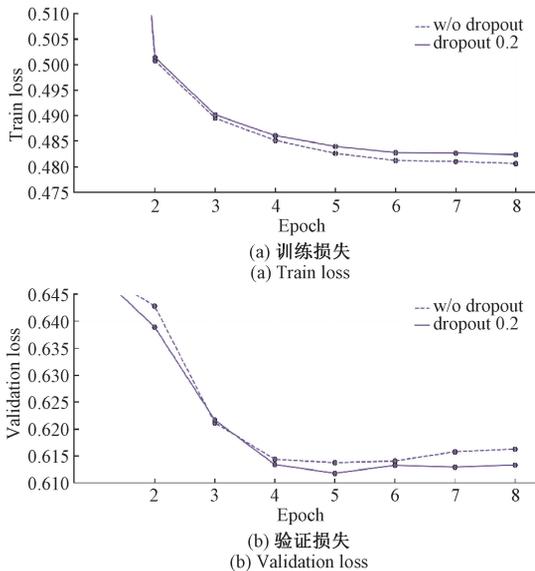


图 11 损失曲线

Fig. 11 Loss curve

### 3) Patch 长度

Patch 长度  $P$  对模型的性能有显著影响。较小的 Patch

长度可以捕获更细粒度的时间特征,但会产生更多的切片,更适合用于复杂的模型进行精细建模,但同时也显著增加了计算成本。较大的 Patch 长度则生成较少的切片,更适合长序列建模,计算效率更高,但可能无法充分捕捉细节信息。

实验结果(图 12)表明,随着  $P$  的增大,模型的 MSE 逐渐降低,预测精度提高,同时内存消耗也随之下降。这说明增大 Patch 长度在减少内存占用和计算负担方面有积极作用。然而,当 Patch 长度过大时,模型在接地电流序列中的细节捕捉能力下降,导致局部特征信息的损失。

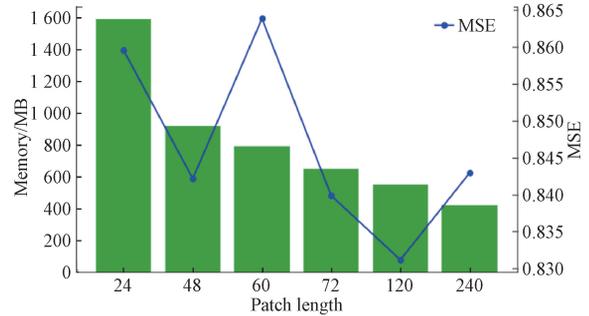


图 12 Patch 长度对性能的影响

Fig. 12 Effect of Patch length on performance

综合考虑预测精度和内存效率,本实验选择  $P = 120$  作为最佳 Patch 长度,以在捕捉细节信息和降低计算成本之间取得平衡。

### 4) 隐藏状态维度

隐藏状态维度  $D$  的大小直接影响模型的复杂性和预测性能。随着  $D$  增大,模型能够捕捉到更加复杂的系统演化模式。然而,过大的  $D$  会显著增加模型的参数量,导致训练效率降低,并可能引入过拟合风险。

在本实验中,测试了不同  $D$  值(16, 32, 64, 128, 256, 512)对模型性能的影响(图 13)。结果显示,当  $D$  大于 32 时,模型预测精度的提升幅度逐渐减小,且在某些点不再具有连续的增益。同时,参数量随着  $D$  值呈指数级增长,导致模型的计算负担加重。考虑到接地电流监测系统对实时性和计算效率的要求,本模型最终选择  $D = 32$  作为隐藏状态的最佳维度,以平衡模型的复杂性和实时性能。

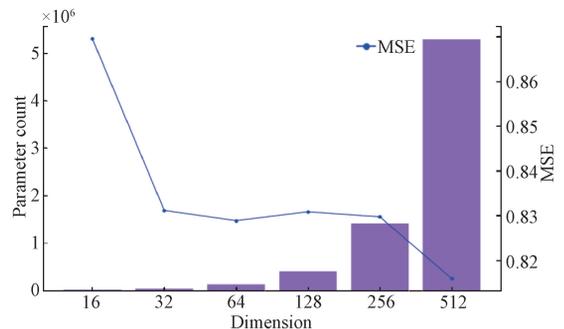


图 13 隐藏状态维度对性能的影响

Fig. 13 Effect of hidden state dimension on performance

#### 4.4 模型泛化性研究实验

为了全面评估本文模型的适应性和泛化能力,实验选取了 5 个公共数据集进行测试,包括 ETTh2、ETTM2、Weather、Electricity 和 Traffic。这些数据集涵盖了多个实际应用领域,如天气预测、交通流量分析和能源管理,能够有效验证模型在不同时间序列场景下的性能。表 4 给出了各数据集的基本统计信息。

在实验设置上,所有数据集的历史回溯窗口大小均设定为  $L=96$ ,并在 4 种不同的预测步长  $Pre \in \{96, 192, 336, 720\}$  下进行评估。实验结果如表 5 所示,其中最佳结果以粗体标注,次佳结果以下划线表示。从表 5 可观察到,在各个数据集的不同预测步长上,Bi-EMamba 模型在大多数情况下均表现出最佳的预测精度,显著优于其他对

表 4 泛化性实验数据集的统计数据

Table 4 Statistical data of the generalization experiment dataset

数据集	变量	频率	长度
ETTh2	7	1 hour	17 420
ETTM2	7	15 min	17 420
Weather	21	10 min	52 696
Electricity	321	1 hour	26 304
Traffic	862	1 hour	17 544

比模型。这一结果表明 Bi-EMamba 具备较强的泛化性和鲁棒性,能够在不同数据分布和预测任务中保持稳定的预测性能。

表 5 公共数据集下模型泛化性验证实验结果

Table 5 Results of model generalization verification experiments on public datasets

Models	Bi-EMamba		S-Mamba		iTransformer		Autoformer		Crossformer		MICN		PatchTST		
	Pre	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh2	96	0.295	0.344	<u>0.296</u>	<u>0.348</u>	0.297	0.349	0.346	0.388	0.745	0.584	0.299	0.364	0.302	<u>0.348</u>
	192	<b>0.369</b>	<u>0.398</u>	<u>0.376</u>	<b>0.396</b>	0.380	0.400	0.456	0.452	0.877	0.656	0.441	0.454	0.388	0.400
	336	<b>0.417</b>	<b>0.428</b>	<u>0.424</u>	<u>0.431</u>	0.428	0.432	0.482	0.486	1.043	0.731	0.654	0.567	0.426	0.433
	720	<b>0.424</b>	<b>0.442</b>	<u>0.426</u>	<u>0.444</u>	0.427	0.445	0.515	0.511	1.104	0.763	0.956	0.716	0.431	0.446
ETTM2	96	<u>0.177</u>	0.268	0.179	<u>0.263</u>	0.180	0.264	0.255	0.339	0.287	0.366	0.179	0.275	<b>0.175</b>	<b>0.259</b>
	192	<b>0.227</b>	<b>0.296</b>	0.250	0.309	0.250	0.309	0.281	0.340	0.414	0.492	0.307	0.376	<u>0.241</u>	<u>0.302</u>
	336	<b>0.289</b>	<b>0.341</b>	0.312	0.349	0.311	0.348	0.339	0.372	0.597	0.542	0.325	0.388	<u>0.305</u>	<u>0.343</u>
	720	<b>0.387</b>	<b>0.395</b>	0.411	0.406	0.412	0.407	0.433	0.432	1.730	1.042	0.502	0.490	<u>0.402</u>	<u>0.400</u>
Weather	96	<u>0.161</u>	0.235	0.165	<b>0.210</b>	0.174	0.214	0.266	0.336	<b>0.158</b>	0.230	<u>0.161</u>	0.229	0.178	0.219
	192	<b>0.203</b>	0.264	0.214	<b>0.252</b>	0.221	<u>0.254</u>	0.307	0.367	<u>0.206</u>	0.277	0.220	0.281	0.224	0.259
	336	<b>0.268</b>	0.321	0.274	<u>0.297</u>	0.278	<b>0.296</b>	0.359	0.395	<u>0.272</u>	0.335	0.278	0.331	0.292	0.306
	720	<u>0.350</u>	0.371	<u>0.350</u>	<b>0.345</b>	0.358	<u>0.347</u>	0.419	0.428	0.398	0.418	<b>0.311</b>	0.356	0.354	0.348
Electricity	96	<b>0.139</b>	0.243	<b>0.139</b>	<u>0.235</u>	<u>0.148</u>	<b>0.240</b>	0.201	0.317	0.219	0.314	0.164	0.269	0.174	0.259
	192	<u>0.160</u>	0.257	<b>0.159</b>	<u>0.255</u>	0.162	<b>0.253</b>	0.222	0.334	0.231	0.322	0.177	0.285	0.178	0.265
	336	<b>0.172</b>	<b>0.266</b>	<u>0.176</u>	0.272	0.178	<u>0.269</u>	0.231	0.338	0.246	0.337	0.193	0.304	0.196	0.282
	720	<u>0.207</u>	<u>0.305</u>	<b>0.204</b>	<b>0.298</b>	0.225	0.317	0.254	0.361	<b>0.280</b>	0.363	0.212	0.321	0.237	0.316
Traffic	96	0.406	0.289	<b>0.382</b>	<b>0.261</b>	<u>0.395</u>	<u>0.268</u>	0.613	0.388	0.522	0.290	0.519	0.309	0.457	0.295
	192	0.429	0.298	<b>0.396</b>	<b>0.267</b>	<u>0.417</u>	<u>0.276</u>	0.616	0.382	0.530	0.293	0.537	0.315	0.471	0.299
	336	0.456	0.310	<b>0.417</b>	<b>0.276</b>	<u>0.433</u>	<u>0.283</u>	0.622	0.337	0.558	0.305	0.534	0.313	0.482	0.304
	720	0.483	0.319	<b>0.460</b>	<b>0.300</b>	<u>0.467</u>	<u>0.302</u>	0.660	0.408	0.589	0.328	0.577	0.325	0.514	0.322

## 5 结 论

本研究针对现有时序预测模型在高压电缆接地电流预测中的计算复杂度高、预测精度不足等问题,提出了 Bi-EMamba 模型。该模型专为高压电缆大数据环境中的实时预测设计,旨在满足监测系统对精准性和实时性的需求,填补了接地电流智能预测领域的空白。通过实验分析,得出以下结论:1)Bi-EMamba 模型结合选择性状态空

间模型与时空依赖关系编码器,针对接地电流数据的非平稳性和时空相关性问题提出了有效解决方案。实验结果表明,Bi-EMamba 在多种预测步长下均表现出优异的性能,显著优于现有基线模型。与 SOTA 模型 iTransformer 相比,其 MSE 平均降低 6.52%,内存消耗减少 29.49%,特别是在长期预测任务中展现出较强的泛化能力和实时预测优势。2)Bi-EMamba 模型的线性复杂度设计显著降低了计算成本,能够高效适应大规模接地电流监测任务,在

资源受限环境下具有良好的扩展性。通过超参数和关键组件的实验验证,进一步证明了模型的合理性和高效性。本研究为高压电缆接地电流的智能预测提供了创新思路,在提升电网运行的安全性和稳定性方面具有重要的工程应用价值。

### 参考文献

- [1] 吴春芳,陈云,黄韬,等. 电力电缆状态感知与检测技术研究综述[J]. 高压电器, 2024, 60(10): 86-103.  
WU CH F, CHEN Y, HUANG T, et al. A review of condition sensing and detection technologies for power cables[J]. High Voltage Apparatus, 2024, 60(10): 86-103.
- [2] 陈晓儒,黄龙毅,赖胜杰,等. 高压电缆护套环流泄漏电流分量的分析与计算[J]. 电气传动, 2024, 54(3): 23-29.  
CHEN X R, HUANG L Y, LAI SH J, et al. Analysis and calculation of leakage current components in sheath circulating current of high-voltage cables[J]. Electric Drive, 2024, 54(3): 23-29.
- [3] LI ZH L, DU B X, LI W. Evaluation of high-voltage AC cable grounding systems based on the real-time monitoring and theoretical calculation of grounding currents[J]. High Voltage, 2018, 3(1): 38-43.
- [4] 李根,周文俊,喻莹,等. 邻近共地高压电缆线路护层电流特征与故障诊断[J]. 高电压技术, 2025, 51(2): 698-707.  
LI G, ZHOU W J, YU Y, et al. Current characteristics and fault diagnosis of sheath in adjacent common-ground high-voltage cable lines [J]. High Voltage Engineering. 2025, 51(2): 698-707.
- [5] AHN H K, PARK N. Deep RNN-based photovoltaic power short-term forecast using power IoT sensors[J]. Energies, 2021, 14(2): 436-452.
- [6] 张安安,何明光,于璐,等. 基于有限元法的海底电缆埋设深度实时预测研究[J]. 电子测量与仪器学报, 2023, 37(12): 18-28.  
ZHANG AN AN, HE M G, YU L, et al. Real-time prediction of submarine cable burial depth based on finite element method [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(12): 18-28.
- [7] 樊浩,宁博扬,何森. 基于SVM改进 PSO算法的电缆接头温度预测分析[J]. 电子测量技术, 2019, 42(21): 53-56.  
FAN H, NING B Y, HE S. Temperature prediction analysis of cable joints based on SVM improved PSO algorithm[J]. Electronic Measurement Technology, 2019, 42(21): 53-56.
- [8] LIU SH ZH, YU H, LIAO C, et al. Pyraformer: Low-complexity pyramidal attention for long-range time series modeling and forecasting[C]. The Tenth International Conference on Learning Representations, 2022.
- [9] WU H X, XU J H, WANG J M, et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting[J]. Advances in Neural Information Processing Systems, 2021, 34: 22419-22430.
- [10] WOO G, LIU CH H, SAHOO D, et al. ETSformer: Exponential smoothing transformers for time-series forecasting[J]. ArXiv preprint arXiv: 2202.01381, 2022.
- [11] DING R X, CHEN Y Q, LAN Y T, et al. DRFormer: Multi-scale transformer utilizing diverse receptive fields for long time-series forecasting [J]. ArXiv preprint arXiv: 2408.02279, 2024.
- [12] NIE Y Q, NGUYEN N H, SINTHONG P, et al. A time series is worth 64 words: Long-term forecasting with transformers [J]. ArXiv preprint arXiv: 2211.14730, 2022.
- [13] ZHANG Y H, YAN J CH. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting [C]. The Tenth International Conference on Learning Representations, 2023.
- [14] LIU Y, HU T G, ZHANG H R, et al. iTransformer: Inverted transformers are effective for time series forecasting [J]. ArXiv preprint arXiv: 2310.06625, 2023.
- [15] ZENG A L, CHEN M H, ZHANG L, et al. Are Transformers Effective for Time Series Forecasting? [C]. AAAI Conference on Artificial Intelligence, 2023, 37(9): 11121-11128.
- [16] DAS A, KONG W H, LEACH A B, et al. Long-term forecasting with TiDE: Time-series dense encoder [J]. ArXiv preprint arXiv: 2304.08424, 2023.
- [17] YI K, ZHANG Q, FAN W, et al. Frequency-domain MLPs are more effective learners in time series forecasting [J]. Advances in Neural Information Processing Systems, 2023, 36: 76656-76679.
- [18] GU A, DAO T. Mamba: Linear-time sequence modeling with selective state spaces [J]. ArXiv preprint arXiv: 2312.00752, 2023.
- [19] ZENG CH L, LIU ZH Y, ZHENG G J, et al. C-mamba: Channel correlation enhanced state space models for multivariate time series forecasting [J]. ArXiv preprint arXiv: 2406.05316, 2024.
- [20] WU Z X, GONG Y F, ZHANG A Q. DTMamba:

- Dual twin mamba for time series forecasting [J]. ArXiv preprint arXiv: 2405.07022, 2024.
- [21] WU L, PEI W B, JIAO J L, et al. UmambaTsf: A u-shaped multi-scale long-term time series forecasting method using mamba [J]. ArXiv preprint arXiv: 2410.11278, 2024.
- [22] 冯强,赵建光,杨茸,等.时间序列中非平稳性和波动性的建模及预测[J].计算机科学与探索,2025,19(5):1313-1321.
- FENG Q, ZHAO J G, YANG R, et al. Modeling and forecasting of non-stationarity and volatility in time series [J]. Journal of Computer Science and Exploration, 2025, 19(5): 1313-1321.
- [23] 周念成,兰雪珂,莫复雪,等.计及线路接头温升约束的受端电网转供优化模型[J].电工技术学报,2021,36(15):3237-3249.
- ZHOU N CH, LAN X K, MO F X, et al. Optimization model of power transfer for the receiving-end grid considering the junction temperature rise constraint [J]. Transactions of China Electrotechnical Society, 2021, 36(15): 3237-3249.
- [24] GU A, DAO T, ERMON S, et al. HiPPO: Recurrent memory with optimal polynomial projections[J]. ArXiv preprint arXiv: 2008.07669, 2020.
- [25] GU A, GOEL K, RE C. Efficiently modeling long sequences with structured state spaces [J]. ArXiv preprint arXiv: 2111.00396, 2021.
- [26] WANG Z H, KONG F H, FENG SH, et al. Is mamba effective for time series forecasting? [J]. ArXiv preprint arXiv: 2403.11144, 2024.
- [27] KIM T, KIM J, TAE Y, et al. Reversible instance normalization for accurate time-series forecasting against distribution shift[C]. The Tenth International Conference on Learning Representations, 2022.
- [28] YI K, ZHANG Q, FAN W, et al. FourierGNN: Rethinking multivariate time series forecasting from a pure graph perspective [J]. ArXiv preprint arXiv: 2311.06190, 2023.
- [29] WANG H Q, PENG J, HUANG F H, et al. MICN: Multi-scale local and global context modeling for long-term series forecasting[C]. The Tenth International Conference on Learning Representations, 2023.

### 作者简介

周景(通信作者),博士,副教授,硕士生导师,主要研究方向为人工智能、机器学习和电力大数据分析。

E-mail:zhoujing108@ncepu.edu.cn

董晖,硕士研究生,主要研究方向为电力大数据分析、人工智能。

E-mail:donghui99@ncepu.edu.cn

刘心,硕士研究生,主要研究方向为计算机视觉与电力故障检测方面的研究工作。

E-mail:liuxin0710@ncepu.edu.cn

唐振洋,硕士研究生,主要研究方向为计算机视觉。

E-mail:tangzhenyang@ncepu.edu.cn