

DOI: 10.19783/j.cnki.pspc.250781

# 基于时序金字塔双层集成学习架构的短期风速区间值预测

冯涛<sup>1</sup>, 艾学轶<sup>1</sup>, 韦善阳<sup>2</sup>, 甘伟<sup>3</sup>, 艾小猛<sup>2</sup>

(1. 武汉科技大学管理学院, 湖北 武汉 430065; 2. 强电磁技术全国重点实验室(华中科技大学电气与电子工程学院), 湖北 武汉 430074; 3. 卡迪夫大学工程学院, 威尔士 卡迪夫 CF24 3AA)

**摘要:** 风速区间值预测通过捕捉实际风速波动范围, 能够有效反映风速的随机性和不确定性。然而, 由于风速区间值序列的多尺度波动特性, 单一预测模型往往难以全面表征其复杂波动趋势, 预测性能受限。基于此, 提出一种基于时序金字塔双层集成学习架构的短期风速区间值预测方法。该方法主要包括数据预处理优化、多模型融合集成预测机制构建、集成输出优化三部分。在数据预处理优化中, 采用红嘴蓝鹊(red-billed blue magpie optimizer, RBMO)算法优化变分模态分解参数, 实现信号在频域上的有效分离。在多模型融合集成预测机制中, 结合时序金字塔注意力机制融合多种模型优势, 对子序列的多尺度特征模式进行差异化建模。集成输出优化则利用灰狼优化(grey wolf optimizer, GWO)算法优化子序列预测结果的输出权重, 以捕捉各子序列与真实风速之间的不同特征相关性。算例分析表明: 所提预测方法能够显著提升短期风速预测精度, 具有较强的泛化能力和稳定性。

**关键词:** 区间值预测; 集成学习; Pyraformer; 金字塔注意力机制

## Short-term wind speed interval prediction based on a temporal pyramid dual-layer ensemble learning architecture

FENG Tao<sup>1</sup>, AI Xueyi<sup>1</sup>, WEI Shanyang<sup>2</sup>, GAN Wei<sup>3</sup>, AI Xiaomeng<sup>2</sup>

(1. School of Management, Wuhan University of Science and Technology, Wuhan 430065, China; 2. State Key Laboratory of Advanced Electromagnetic Technology (School of Electrical and Electronic Engineering, Huazhong University of Science and Technology), Wuhan 430074, China; 3. School of Engineering, Cardiff University, Cardiff CF24 3AA, Wales, UK)

**Abstract:** Wind speed interval prediction captures the fluctuation range of actual wind speed, effectively reflecting its randomness and uncertainty. However, due to the multi-scale fluctuation characteristics of wind speed interval sequences, a single prediction model often fails to fully represent their complex dynamics, leading to limited prediction performance. To address this issue, this paper proposes a short-term wind speed interval prediction method based on a temporal pyramid dual-layer ensemble learning architecture. The method mainly comprises three parts: data preprocessing optimization, multi-model ensemble prediction mechanism construction, and ensemble output optimization. The data preprocessing optimization employs the red-billed blue magpie optimizer to optimize the parameters of variational mode decomposition, enabling effective signal decomposition in the frequency domain. In the multi-model ensemble prediction stage, a temporal pyramid attention mechanism is introduced to integrate the advantages of multiple models and perform differentiated modeling of multi-scale feature patterns in sub-sequences. In the ensemble output optimization stage, the grey wolf optimizer is used to optimize the output weights of sub-sequence predictions, capturing the varying feature correlations between each sub-sequence and actual wind speed. Case studies demonstrate that the proposed method can significantly improve the accuracy of short-term wind speed prediction, and has strong generalization ability and stability.

This work is supported by the National Natural Science Foundation of China (No. 52407103).

**Key words:** interval prediction; ensemble learning; Pyraformer; pyramidal attention mechanism

## 0 引言

风速固有的随机性和间歇性易导致风电出力波

动, 影响风电安全并网<sup>[1-4]</sup>。此外, 风电装机规模的持续攀升加剧了电力系统运行的不确定性<sup>[5-7]</sup>, 电网安全运行和新能源高效消纳之间的矛盾日益凸显<sup>[8]</sup>。在此背景下, 高精度的短期风速预测能够为电网实时调度和调频提供有力支撑, 促进大规模风

基金项目: 国家自然科学基金项目资助(52407103)

电消纳<sup>[9]</sup>, 对优化能源输出, 保障电力系统稳定运行至关重要<sup>[10]</sup>。

目前风速预测研究大多集中在基于单值风速的点预测或概率区间预测上<sup>[11]</sup>。其中, 点预测仅提供未来时刻风速的单一具体值, 难以有效捕捉实际风速的波动范围<sup>[12]</sup>; 概率区间预测则通常基于点预测结果, 在不同显著性水平下得到未来时刻风速的概率区间, 相比于点预测虽然能够进一步描述风速的潜在波动<sup>[13]</sup>, 但其本质上仍局限于反映点值附近的潜在变化, 导致在表征风速的不确定性和波动性方面仍存在不足<sup>[14]</sup>。相比之下, 基于区间值风速的预测方法因其独特的优势, 克服了上述基于单值风速预测方法的局限。与单值序列不同, 区间值序列通过涵盖每个时间点风速的上下界来刻画变量的动态变化范围, 蕴含更丰富的预测信息<sup>[15]</sup>。区间值序列能够更加直接地描述复杂因素引起的风速波动特征和不确定信息<sup>[16]</sup>, 尤其在极端天气等风速波动较剧烈的场景下, 有望为能源规划和调度决策提供更全面、更可靠的预测信息。然而, 由于风速固有的随机性和波动性, 区间值序列数据信息增加的同时, 也引入了更高的不确定性和建模复杂性, 使得预测难度进一步增加。因此, 如何开发能够深度捕捉风速区间值序列不确定信息和高维复杂波动特征的预测方法, 是充分发挥区间值预测优势的关键问题, 也是当前研究中亟待解决的核心难题。

同时, 随着研究的不断深入, 风速预测方法已从早期的单一预测模型发展为更复杂的集成预测模型。集成预测建模时通常会考虑多种模型, 以减轻单个模型的预测偏差, 从而提高模型预测精度和泛化能力<sup>[17]</sup>。例如, 文献[18]构建了包含多种预测模型的模型库, 分别为各风速子序列选取最优模型以

执行预测任务; 文献[19]提出了基于加权信息准则的模型选择策略, 为各风速子序列选择最佳预测模型。尽管上述集成方法在整体预测任务建模中考虑了多种模型, 但在风速子序列预测任务中仍采用单一建模策略, 未能充分发挥多模型融合、协同预测优势, 难以充分捕捉风速区间值序列包含的大量不确定信息以及高维复杂波动特征, 制约了模型整体的预测精度与泛化能力。

基于此, 本文提出了基于时序金字塔双层集成学习架构的短期风速区间值预测方法, 在数据驱动的基础上融合多模型性能优势。首先, 提出基于红嘴蓝鹊优化(red-billed blue magpie optimizer, RBMO)算法优化变分模态分解参数的数据预处理优化方法, 实现风速信号在频域上的有效分解和各模态分量的精准分离; 其次, 构建基于 Pyraformer 时序金字塔注意力机制的多模型融合预测机制, 对风速区间值序列的多尺度波动模式进行差异化建模; 最后, 提出基于灰狼优化(grey wolf optimizer, GWO)算法的集成输出优化策略, 捕捉各子序列与真实风速之间不同的特征相关性, 进一步提升预测精度。

## 1 考虑模型多样性的区间值预测方法

本文所提预测方法的整体架构如图 1 所示。首先, 模块 1 通过 RBMO 优化变分模态分解(variational mode decomposition, VMD)的参数取值, 得到分解后的若干风速子序列。其次, 模块 2 建立了基于 Pyraformer 的多模型融合预测机制, 在多时间尺度上对风速不同波动模式以及子模型不同预测优势之间的复杂特征关系进行建模, 输出风速子序列预测结果。最后, 模块 3 通过 GWO 算法优化模块 2 风速子序列预测结果的输出权重, 得到最终的风速预测结果。

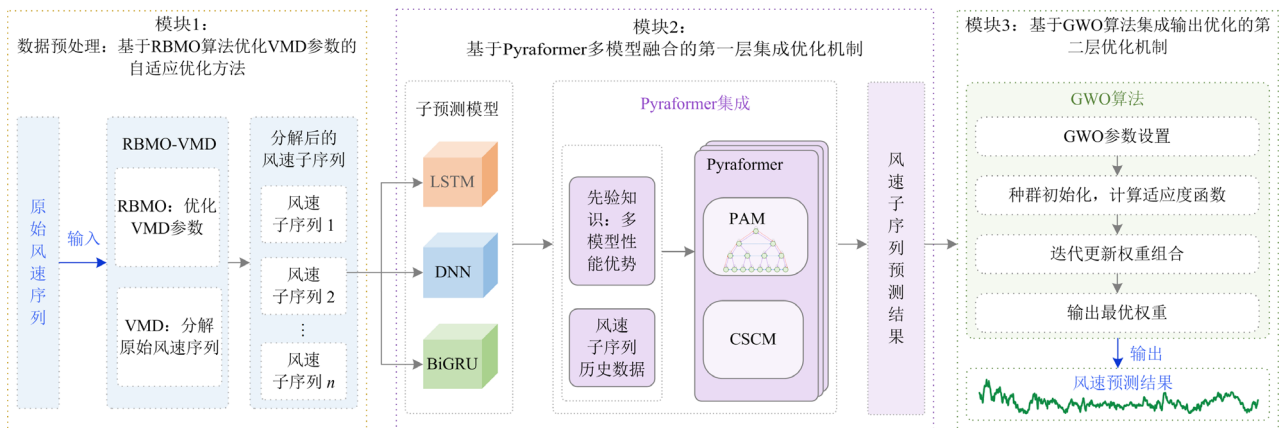


图 1 所提短期风速区间值预测方法框架图

Fig. 1 Architecture of the proposed short-term wind speed interval-valued prediction method

### 1.1 数据预处理: 基于 RBMO 算法优化 VMD 参数的自适应分解优化方法

风速区间值序列包含大量不确定信息和高维复杂波动特征, VMD 作为一种信号处理方法, 在分解中不仅可以抑制噪声影响, 而且能够有效提取不同尺度的频率特征<sup>[20]</sup>。但是, 其分解效果在很大程度上依赖于模态数  $K$  和惩罚因子  $\alpha$  的取值。

首先, VMD 通过构造并求解变分问题, 将复杂信号分解为多个具有不同中心频率的本征模态函数<sup>[21]</sup>。

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ \text{s.t.} \sum_{k=1}^K u_k(t) = x(t) \end{cases} \quad (1)$$

式中:  $x(t)$  表示输入信号;  $u_k(t)$  是分解后的第  $k$  个分量;  $\delta(t)$  是狄拉克  $\delta$  函数;  $\omega_k$  是中心频率; “\*” 代表卷积运算;  $\partial_t$  表示对时间  $t$  求偏导数。

然后, 构建增广拉格朗日函数, 将约束变分问题转化为无约束问题, 如式(2)所示。

$$\begin{cases} L(\{u_k\}, \{\omega_k\}, \{\lambda\}) = A + B + \langle \lambda(t), x(t) - C \rangle \\ A = \alpha \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ B = \left\| x(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 \\ C = \sum_{k=1}^K u_k(t) \end{cases} \quad (2)$$

式中:  $\lambda(t)$  是拉格朗日乘子。

最后, 使用交替方向乘子法迭代求解, 直至满足收敛阈值。

$$\sum_{k=1}^K \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (3)$$

式中:  $\hat{u}_k^n$  为第  $k$  个模态在第  $n$  次迭代时的频域值;  $\varepsilon$  为预设的收敛阈值。

由上述步骤可知, 模态数  $K$  设定过小会导致信息丢失, 无法充分捕捉信号特征; 过大又会导致过度分解, 纳入噪声信号。惩罚因子  $\alpha$  控制各分量带宽,  $\alpha$  过小易产生模态混叠, 降低分解精度;  $\alpha$  过大会使信号分割细碎。

因此, 本文提出了基于 RBMO 算法的参数寻优策略, 自适应确定 VMD 分解的最佳  $K$ 、 $\alpha$  组合, 以更准确地将风速信号分解为不同的模态分量, 降低区间值序列数据复杂度。RBMO<sup>[22]</sup> 参数自适应调整机制能够动态优化搜索策略, 即使在高维、复杂的风速频域分解场景中也能保持稳定的性能。本文

所提 RBMO-VMD 自适应分解方法的详细优化过程如图 2 所示。

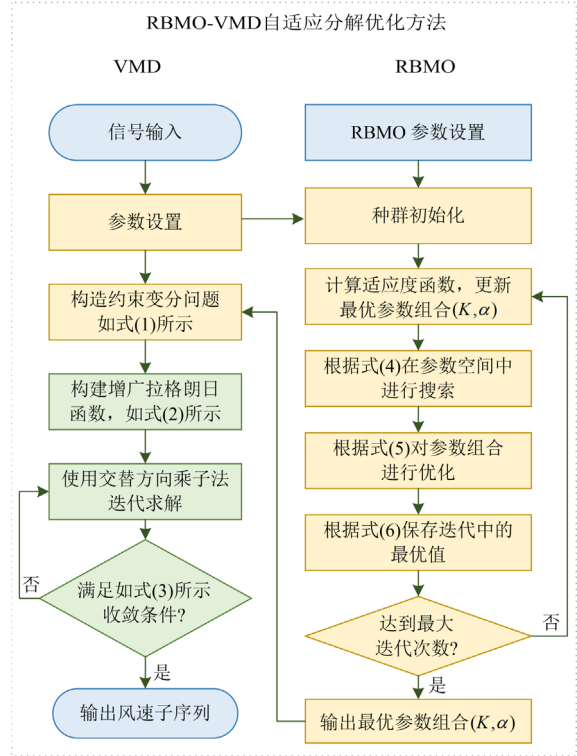


图 2 RBMO-VMD 自适应分解优化方法流程图

Fig. 2 Flow chart of RBMO-VMD algorithm

步骤 1: 随机生成 VMD 分解的  $K$  和  $\alpha$  初始值。

步骤 2: 确定适应度函数  $f$ 。本文的适应度函数为风速序列各模态分量样本熵的最小值。

步骤 3: 根据式(4)在 VMD 参数取值空间中进行全面搜索, 找到更优的 VMD 参数组合。

$$\begin{cases} X_i(t+1) = X_i(t) + (X_p^{\text{mean}}(t) - X_r(t)) \cdot r_1 \\ X_i(t+1) = X_i(t) + (X_q^{\text{mean}}(t) - X_r(t)) \cdot r_2 \end{cases} \quad (4)$$

式中:  $X_i(t+1)$  表示第  $t+1$  次迭代中生成的  $(K, \alpha)$  取值组合;  $r_1$ 、 $r_2$  是  $(0, 1)$  之间的随机数;  $X_p^{\text{mean}}(t)$  是随机选择的  $p$  个个体的平均值;  $X_q^{\text{mean}}(t)$  是随机选择的  $q$  个个体的平均值;  $X_r$  是随机选择的一个个体; 当个体数量  $m \in [2, 5]$  时,  $p = m$ , 当  $m \geq 10$  时,  $q = m$ 。

步骤 4: 根据式(5), 利用已有的搜索信息, 对当前发现的较为优秀的 VMD 参数取值组合进行进一步优化。

$$\begin{cases} X_i(t+1) = X_{\text{food}}(t) + C_F \cdot (X_p^{\text{mean}}(t) - X_i(t)) \cdot r_3 \\ X_i(t+1) = X_{\text{food}}(t) + C_F \cdot (X_q^{\text{mean}}(t) - X_i(t)) \cdot r_4 \\ C_F = \left( 1 - \left( \frac{t}{T} \right)^{\left( \frac{2t}{T} \right)} \right) \end{cases} \quad (5)$$

式中： $X_{\text{food}}(t)$  为第  $t$  次迭代得到的最优参数组合； $T$  是最大迭代次数； $r_3$ 、 $r_4$  是正态随机扰动； $C_F$  是协作因子。

步骤 5：根据式(6)，在优化过程中保存并更新 VMD 的最优参数组合。在完成全部迭代后，输出全局最优的  $(K, \alpha)$  组合。

$$X_i(t+1) = \begin{cases} X_{\text{old}}, & f(X_i(t+1)) > f_{\text{old}} \\ X_i(t+1), & f(X_i(t+1)) \leq f_{\text{old}} \end{cases} \quad (6)$$

式中： $X_{\text{old}}$  是之前迭代所产生的最优参数组合； $f_{\text{old}}$  是  $X_{\text{old}}$  所对应的适应度函数值。

### 1.2 基于 Pyraformer 多模型融合的第一层集成优化机制

所提 Pyraformer 多模型融合集成预测机制的具体优化流程如图 3 所示。

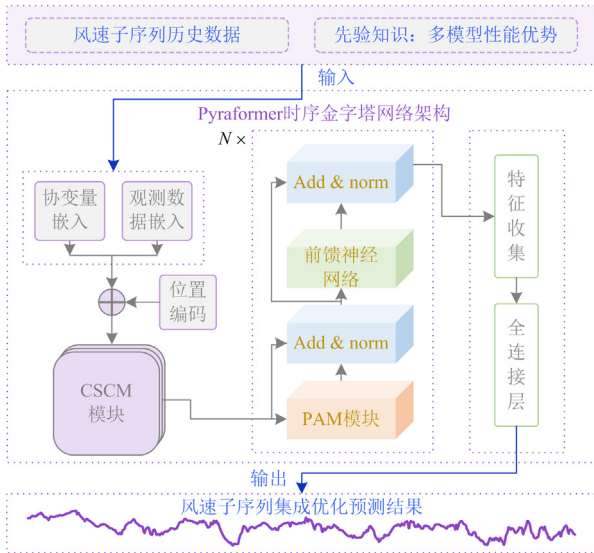


图 3 Pyraformer 多模型融合集成预测机制  
Fig. 3 Multi-model fusion ensemble prediction mechanism based on Pyraformer

Pyraformer 是一种基于多分辨率金字塔注意力机制的神经网络<sup>[23]</sup>，能够在不同时间分辨率上提取输入数据的多尺度特征<sup>[24]</sup>。因此，本文结合 Pyraformer 的多尺度特征建模能力，获取各子模型在不同时间尺度下的相对预测优势，以自适应构建关于模型性能偏好的先验知识。同时，捕捉风速序列在不同时间尺度下的波动模式，并进一步建模这些波动特征与各子模型预测优势之间的复杂特征关系，构建深度融合风速多尺度特征与多模型性能偏好的集成预测机制。首先，量化不同子模型性能优势，并与风速子序列历史数据耦合；然后，耦合数据共同作为观测数据嵌入到所提基于 Pyraformer 的多模型融合预测机制中，对风速不同波动模式以及子模型不同

预测优势之间的复杂特征关系进行建模。此外，历史风速数据对应的时间戳信息作为协变量嵌入，以捕捉风速的时间周期规律。

本文在开展正式算例验证前进行了大量的子模型选择实验，选择出 3 种具有不同性能优势且在不同数据集上性能较稳健的子预测模型执行后续实验。即：长短时记忆网络(long short-term memory, LSTM)、深度神经网络(deep neural network, DNN)和双向门控循环单元(bidirectional gated recurrent unit, BiGRU)。其中，LSTM 通过细胞状态和门控机制可以有效捕捉风速序列的长期依赖关系<sup>[25]</sup>；DNN 模型利用多层非线性变换能够有效提取风速子序列的高维隐层特征<sup>[26]</sup>；BiGRU 模型通过双向递归结构融合历史与未来上下文信息<sup>[27]</sup>，擅长捕捉风速瞬时波动特征。

具体而言，图 3 中的嵌入数据首先被传输到粗粒度构造模块(coarse-scale construction module, CSCM)中，该模块通过逐层卷积下采样捕获风速子序列在不同时间尺度上的特征。设  $L$  为输入序列长度，CSCM 将核大小为  $C$ 、步长为  $C$  的  $S$  个卷积层依次应用于嵌入序列，生成由细到粗共  $S$  个尺度的序列，每个尺度下的序列长度为： $L/C^{s-1}, s=1, 2, \dots, S$ 。第一个尺度(最细尺度)下的序列长度为： $L^{(1)} = L/C^0 = L$ ，对应原始输入序列，用于保留细节信息。第  $S$  尺度(最粗尺度)下的序列长度为  $L^{(S)} = L/C^{S-1}$ ，用于提取全局信息。通过上述操作，CSCM 在多时间尺度下构造了从精细到粗糙的信息层次，以促进后续操作中各节点间信息的高效交互<sup>[28]</sup>。

随后，如图 4 所示，Pyraformer 中的金字塔注意力模块(pyramidal attention module, PAM)接收 CSCM 输出的多尺度拼接序列。PAM 利用金字塔注意力机制，通过在不同尺度节点间交互传递信息，实现不同

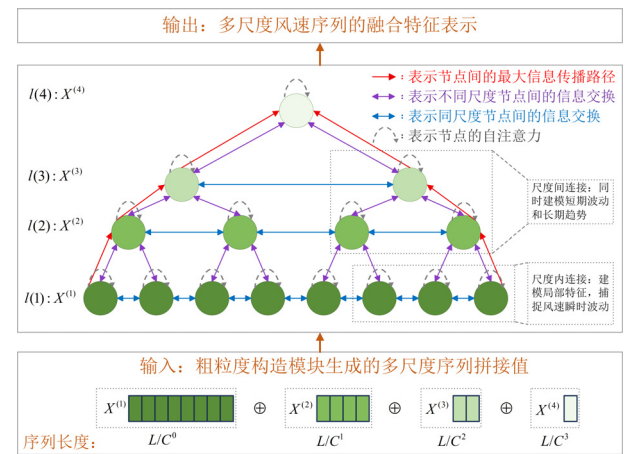


图 4 金字塔注意力模块  
Fig. 4 Pyramidal attention module

层级间的信息交换<sup>[29]</sup>, 从而捕捉风速子序列与子模型预测优势在不同时间尺度上更丰富、细致的依赖关系。

图4展示了4个层级的节点 $l(1)$ 、 $l(2)$ 、 $l(3)$ 和 $l(4)$ , 其中 $l(1)$ 是金字塔底层最精细尺度的节点, 对应原始输入序列, 用于保留细节信息, 捕捉局部特征。而上层 $l(2)$ 、 $l(3)$ 和 $l(4)$ 的较粗尺度节点则表征更长时间尺度的特征, 捕捉整体波动趋势。图中的边可分为两类: 跨尺度连接与尺度内连接<sup>[30]</sup>。跨尺度连接捕捉风速子序列以及子预测模型在不同时间尺度下的特征模式, 帮助模型同时建模短期波动和长期趋势。尺度内连接则在金字塔结构每一层级的内部, 通过连接相邻节点捕捉风速子序列以及子预测模型的局部特征模式。具体而言, 中间层级 $l(2)$ 的每个粗尺度节点通过 $C$ 叉树( $C=2$ ), 与其在 $l(1)$ 层级中的 $C$ 个子节点相连, 形成尺度内连接。如此, $l(1)$ 层级中的子节点信息可以通过注意力汇聚到 $l(2)$ 层级的父节点中; 同时, 父节点也能够将全局上下文信息传递给子节点, 以构建不同层级间跨尺度的注意力连接。此外, 在同一层级中, 每个节点只需关注包括自身在内的同尺度的 $A$ 个相邻节点, 形成尺度内连接。例如当 $A=3$ 时, 底层 $l(1)$ 处的节点只需关注其自身以及左右相邻的2个节点。

如式(7)所示, 设 $n_\ell^{(s)}$ 表示尺度 $s$ 上的第 $\ell$ 个节点, 其中 $s=1, 2, \dots, S$ 分别表示金字塔从底层到顶层的尺度。 $n_\ell^{(s)}$ 以及图中每个节点仅需关注3个尺度的相邻节点集合 $N_\ell^{(s)}$ 。

$$\begin{cases} N_\ell^{(s)} = A_\ell^{(s)} \cup C_\ell^{(s)} \cup P_\ell^{(s)} \\ A_\ell^{(s)} = \left\{ n_j^{(s)} : |j - \ell| \leq \frac{A-1}{2}, 1 \leq j \leq \frac{L}{C^{s-1}} \right\} \\ C_\ell^{(s)} = \{ n_j^{(s-1)} : |\ell-1|C < j \leq \ell C \} \text{ if } s \geq 2 \text{ else } \emptyset \\ P_\ell^{(s)} = \left\{ n_j^{(s+1)} : j = \left\lceil \frac{\ell}{C} \right\rceil \right\} \text{ if } s \leq S-1 \text{ else } \emptyset \end{cases} \quad (7)$$

式中:  $s$ 是尺度索引;  $\ell$ 是节点索引;  $L$ 是输入序列的长度;  $S$ 是金字塔结构的总尺度层数;  $N_\ell^{(s)}$ 由同尺度下包含自身在内的相邻 $A$ 个节点的邻域集(记为 $A_\ell^{(s)}$ )、在 $C$ 叉树中的 $C$ 个子节点(记为 $C_\ell^{(s)}$ )和其在 $C$ 叉树中的唯一父节点(记为 $P_\ell^{(s)}$ )组成。

由式(7)可知, 节点 $n_\ell^{(s)}$ 仅需对3个尺度的相邻节点集合 $N_\ell^{(s)}$ 计算注意力, 而非全局节点。如此, 在显著降低计算复杂度的同时, 实现了对信息的跨尺度关注。因此, 节点 $n_\ell^{(s)}$ 的注意力可表示为

$$y_i = \sum_{\ell \in N_\ell^{(s)}} \frac{\exp(\mathbf{q}_i \mathbf{k}_\ell^T / \sqrt{d_k}) \mathbf{v}_\ell}{\sum_{\ell \in N_\ell^{(s)}} \exp(\mathbf{q}_i \mathbf{k}_\ell^T / \sqrt{d_k})} \quad (8)$$

式中:  $\mathbf{q}_i$ 是当前节点的查询向量;  $\mathbf{k}_\ell^T$ 是邻居节点键向量的转置;  $d_k$ 是键向量维度;  $\mathbf{v}_\ell$ 是邻居节点的值向量。

经过 $N$ 层注意力堆叠后, 最粗尺度节点可获得全局感受野。

$$\frac{L}{C^{s-1}} - 1 \leq \frac{(A-1)N}{2} \quad (9)$$

式中:  $N$ 是金字塔注意力层数。

通过上述操作, PAM构建了跨尺度的树状连接和同尺度的局部连接, 实现了对序列多时间尺度依赖关系的高效建模。随后, PAM输出多尺度序列的融合特征表示。通过特征提取操作, 从每个尺度中提取最后一个节点的特征表示并进行拼接, 再传递给全连接层。全连接层将PAM计算得到的多尺度时间依赖特征映射为具体的目标预测值, 最终输出风速子序列预测结果。

### 1.3 基于GWO算法集成输出优化的第二层优化机制

不同风速子序列预测结果包含不同的预测信息, 与实际风速的相关性可能呈现不平衡分布。因此对各子序列的影响权重占比进行区分, 能够进一步提升整体预测精度。GWO<sup>[31]</sup>能够在广阔且复杂的权重求解空间中快速适应风速的非线性特征, 在保证优化质量的同时能够有效降低计算资源消耗和时间成本。基于GWO算法的集成输出优化流程如图5所示。

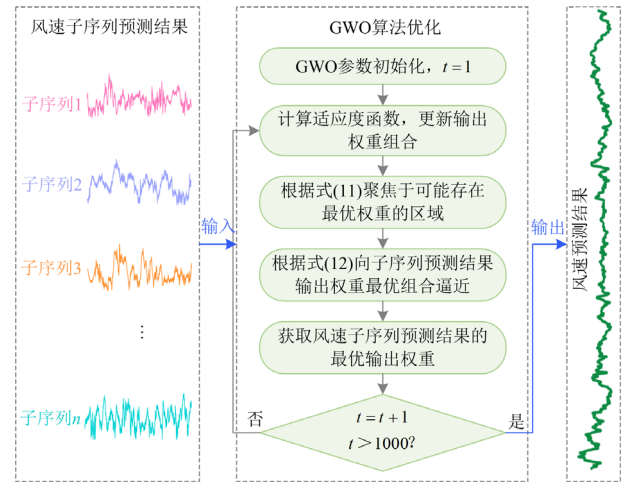


图5 基于GWO的集成输出优化策略

Fig. 5 Ensemble output optimization strategy based on GWO

步骤1: 参数初始化。生成风速子序列预测结果的初始输出权重。

步骤2: 构造如式(10)所示的适应度函数 $f_{\min}$ , 其值越小, 预测优化结果越好。

$$f_{\min} = \sqrt{\frac{1}{M} \sum_{k=1}^M (y(k) - \hat{y}(k))^2} \quad (10)$$

式中:  $y(k)$  是真实风速;  $\hat{y}(k)$  是预测风速;  $M$  是数据点数量。

步骤 3: 权重初步寻优。根据式(11)在权重输出空间中不断缩小搜索范围。

$$\begin{cases} \bar{X}(t+1) = \bar{X}_p(t) - \bar{A} \cdot \bar{D} \\ \bar{D} = |\bar{C} \cdot \bar{X}_p(t) - \bar{X}(t)| \\ \bar{A} = 2\bar{a} \cdot \bar{r}_1 - \bar{a} \\ \bar{C} = 2 \cdot \bar{r}_2 \\ \bar{a} = 2 - t \cdot \frac{2}{T} \end{cases} \quad (11)$$

式中:  $\bar{X}_p$  是当前风速子序列预测结果的最优输出权重;  $\bar{X}(t+1)$  是第  $t+1$  次迭代时的风速子序列预测结果输出权重;  $\bar{a}$  是收敛因子, 线性递减控制搜索范围;  $\bar{A}$  是控制全局探索与局部开发的系数向量;  $\bar{C}$  是随机系数向量;  $\bar{D}$  是距离向量;  $\bar{r}_1$ 、 $\bar{r}_2$  是  $[0, 1]$  之间的随机向量。

步骤 4: 聚焦最优输出权重。按式(12)进一步更新预测结果的输出权重, 并向最优解逼近。

$$\begin{cases} \bar{D}_\alpha = |\bar{C}_1 \cdot \bar{X}_\alpha - \bar{X}|, \bar{X}_1 = 2 \cdot \bar{X}_\alpha - \bar{A}_1 \cdot \bar{D}_\alpha \\ \bar{D}_\beta = |\bar{C}_2 \cdot \bar{X}_\beta - \bar{X}|, \bar{X}_2 = 2 \cdot \bar{X}_\beta - \bar{A}_2 \cdot \bar{D}_\beta \\ \bar{D}_\gamma = |\bar{C}_3 \cdot \bar{X}_\gamma - \bar{X}|, \bar{X}_3 = 2 \cdot \bar{X}_\gamma - \bar{A}_3 \cdot \bar{D}_\gamma \\ \bar{X}(t+1) = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{3} \end{cases} \quad (12)$$

式中:  $\bar{X}_\alpha$ 、 $\bar{X}_\beta$ 、 $\bar{X}_\gamma$  是不同位置向量;  $\bar{X}$  是优化对象的当前位置向量。

步骤 5: 输出最优权重。当  $|\bar{A}| < 1$  时, 输出当前迭代中的最优权重组合, 直至满足最大迭代次数, 结束优化。

## 2 数据集与评价指标构建

### 2.1 实验数据

数据集 1 和数据集 2 分别来自美国国家海洋和大气管理局全球监测实验室的 Barrow(BRW) 和 Samoa(SMO)气象站。分别在 BRW 和 SMO 气象站上搜集每 10 min 间隔下的最小风速、最大风速以及平均风速。数据集 1 和数据集 2 的搜集时间分别是: 2024 年 11 月 1 日 00:00 至 2024 年 11 月 30 日 23:59、2023 年 8 月 8 日 16:19 至 2024 年 9 月 30 日 23:59。本文进行单步预测, 根据前 20 个时间点的已知风速区间值数据估算第 21 个时间点的数值。如图 6 和图 7 所示: 前 60% 为训练集, 中间 25% 为验证集, 最后 15% 为测试集。

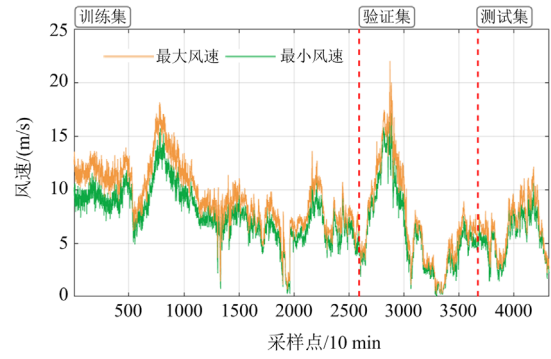


图 6 BRW 气象站每 10 min 间隔下的风速区间值

Fig. 6 Wind speed interval-valued with 10 min span in BRW

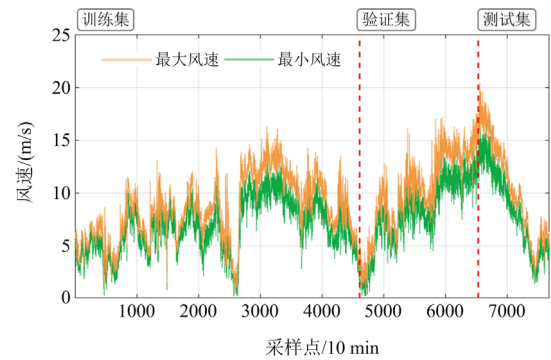


图 7 SMO 气象站每 10 min 间隔下的风速区间值

Fig. 7 Wind speed interval-valued with 10 min span in SMO meteorological station

### 2.2 评价指标

本文引入 4 个区间预测评价指标: 区间平均绝对误差(interval mean absolute error, IMAE)、区间均方根误差(interval root mean square error, IRMSE)、区间平均绝对百分比误差(interval mean absolute percentage error, IMAPE)<sup>[17]</sup> 和预测区间覆盖率(prediction interval coverage probability, PICP)<sup>[15]</sup>, 以及 3 个确定性点预测评价指标: 平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)和平均绝对百分比误差(mean absolute percentage error, MAPE)<sup>[32]</sup>, 评估所提预测方法的性能。

## 3 算例分析

### 3.1 模型参数设置

本研究模型参数设置首先参考相关文献取值, 然后进行大量参数取值试验, 选用试验中的最优参数执行后续实验。其中, RBMO 和 GWO 的种群数量分别为 20、100, 最大迭代次数分别为 50、1000; Pyraformer 模型的训练批次大小为 32, 训练迭代次

数为 500, 注意力头数为 8, 学习率为 0.0001, 激活函数为 ReLU; LSTM、DNN 和 BiGRU 模型的训练迭代次数为 500, 学习率为 0.001, 优化器为 Adam。本研究相关实验在 Matlab 和 Pytorch 中进行, 硬件配置为: 2.50 GHz CPU, 16 GB RAM; GPU, RTX 3080 Ti。

### 3.2 所提预测方法有效性分析

本节将在具有差异化时空分布和波动特性的两个风速数据集上验证所提短期风速区间值预测方法的有效性。

**方案 1:** 不考虑模型多样性, 不进行集成输出优化。仅采用单一模型对风速子序列进行预测, 再将子预测结果相加得到最终风速预测结果。

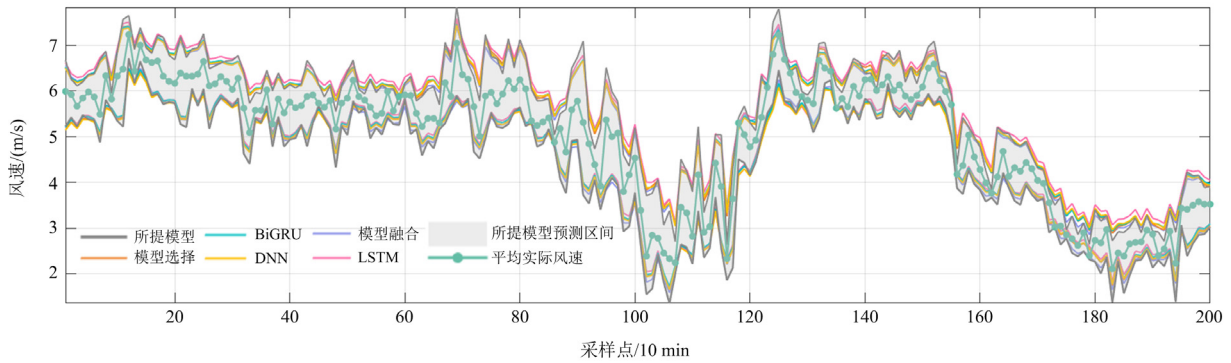
**方案 2:** 不采用所提 Pyraformer 多模型融合预测方法, 不进行集成输出优化。根据文献[18-19]中所提模型选择集成预测方法, 分别从方案 1 的 3 种单一预测模型中为各风速子序列选择最佳预测模型, 再将子预测结果相加得到最终的预测结果。

**方案 3:** 考虑模型多样性, 采用所提 Pyraformer 多模型融合集成预测方法, 但不进行集成输出优化。

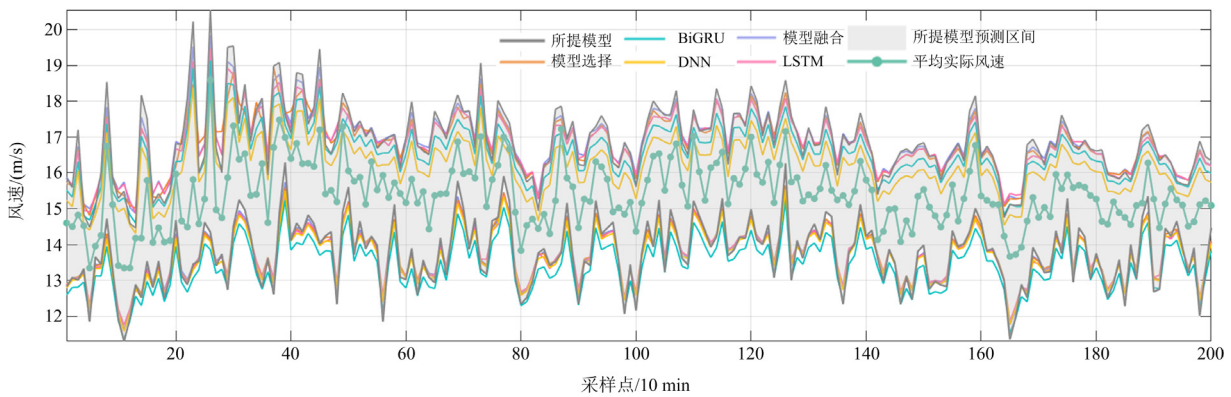
**方案 4:** 本文所提预测方法。算例评价结果如表 1 所示, 其中最佳指标用粗体表示。此外, 图 8 展示了风速区间值预测的可视化结果, 选取了测试集中前 200 个时间间隔下的风速拟合数据, 以更加清晰地展示各预测方法之间的预测差异。

表 1 不同模型风速区间值预测性能评价结果

Table 1 Evaluation results on wind speed interval-valued prediction for different models						
数据集	方案	模型	评价指标			
			IMAE	IMAPE	IRMSE	PICP
数据集 1	方案 1	LSTM	0.1827	0.0366	0.227	0.9244
		DNN	0.1682	0.0324	0.2153	0.9506
		BiGRU	0.164	0.0317	0.2092	0.9429
	方案 2	模型选择	0.1588	0.03	0.2039	0.9537
	方案 3	模型融合	0.1454	0.0273	0.1866	0.9707
	<b>方案 4</b>	<b>Proposed</b>	<b>0.1156</b>	<b>0.0214</b>	<b>0.1459</b>	<b>0.9707</b>
数据集 2	方案 1	LSTM	0.1934	0.0241	0.2624	0.9991
		DNN	0.2202	0.0254	0.316	0.9922
		BiGRU	0.2035	0.0241	0.2859	0.9983
	方案 2	模型选择	0.1749	0.0217	0.2451	0.9991
	方案 3	模型融合	0.169	0.0211	0.2298	0.9991
	<b>方案 4</b>	<b>Proposed</b>	<b>0.136</b>	<b>0.0166</b>	<b>0.1824</b>	<b>0.9991</b>



(a) 数据集 1 不同预测方法的风速区间值拟合曲线对比



(b) 数据集 2 不同预测方法的风速区间值拟合曲线对比

图 8 数据集 1 和数据集 2 中不同预测方法的风速区间值拟合曲线

Fig. 8 Fitting curves of wind speed interval-valued prediction for different methods in Dataset 1 and Dataset 2

根据表 1 的评价结果可以看出, 在两个数据集上, 本文所提预测模型的所有评价指标都取得了最优值, 预测区间覆盖率(prediction interval coverage probability, PICP)达到了 97%以上, 预测性能相比于对比方法提升了 19.51%以上。同时, 从图 8 可以看出, 在风速数据波动较为剧烈的时段, 所提模型的拟合曲线更贴近实际风速, 预测区间宽度相比于其他方法更窄。这表明所提区间值预测方法能够在较窄的预测区间宽度上更精准地捕捉未来风速的潜在波动趋势。此外, 从方案 2 和方案 3 的实验对比结果可以看出, 本文所提基于 Pyraformer 时序金字塔的多模型融合预测方法能够在有效提升风速预测精度的同时, 保持较强的稳定性, 表明其能够稳定、有效地表征风速区间值序列中所包含的大量不确定信息和高维复杂波动特征, 提升预测精度。方案 3 与方案 4 的实验对比结果进一步表明, 基于 GWO 算法的第二层集成输出优化机制, 能够在多模型融合的基础上, 进一步捕捉各子序列与真实风速之间不同的特征相关性, 提升整体预测性能。

因此, 上述算例结果表明, 本文所提方法进一步提升了风速预测的准确性和稳定性。一方面, 所提基于双层集成优化的区间值预测方法包含了更丰富的预测信息, 能够更直接地描述风速的潜在波动趋势。另一方面, 所提方法在数据驱动的基础上有效融合了多模型预测性能优势。即使在风速波动较为剧烈的场景下, 也能够稳定表征复杂因素引起的风速波动特征和不确定信息, 为机组功率调节、电网调频、电力调度等策略的制定提供更加可靠、全面的信息依据。

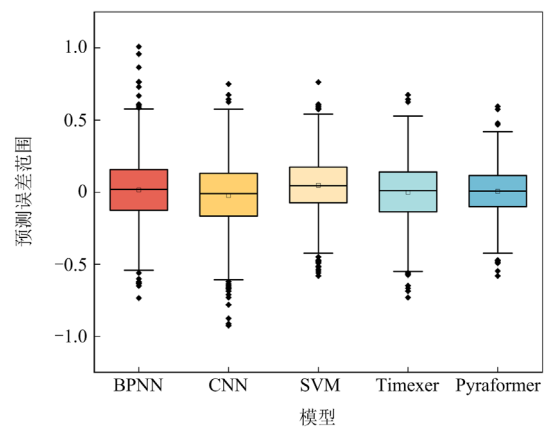
### 3.3 基于 Pyraformer 的多模型融合集成优化机制必要性分析

本节分别在数据集 1 的最大风速集和最小风速集上进行确定性点预测实验, 以验证所提基于 Pyraformer 多模型融合机制的必要性。实验对比模型分别是: 卷积神经网络(convolutional neural network, CNN)、反向传播神经网络(back propagation neural network, BPNN)、支持向量机(support vector machine, SVM)和 Timexer<sup>[33]</sup>。算例结果如表 2 和图 9 所示。

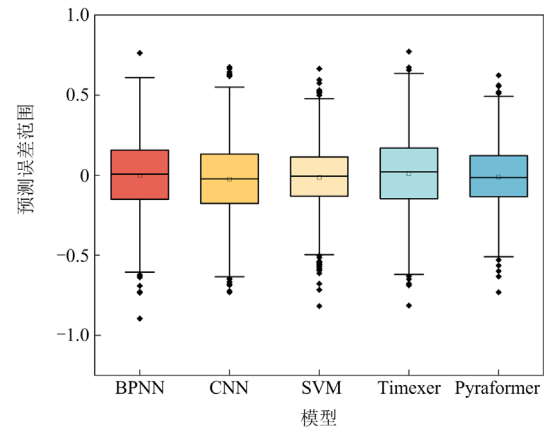
从表 2 可以看出, 不同模型架构的集成学习表现存在显著差异。部分模型如 CNN、SVM 未能有效融合多模型预测优势, 在各数据集上的性能表现均弱于其他模型。Timexer 模型在两个风速集上的预测性能差异较大, 稳定性不强。相比之下, 本文所提基于 Pyraformer 时序金字塔的多模型融合集成优化机制, 在不同数据集中始终保持最佳预测效果, 预测性能相比于对比模型至少提升了 16.92%, 具有

表 2 数据集 1 不同模型性能评价结果

数据集 1	集成模型	评价指标		
		MAE	MAPE	RMSE
最小风速集	BPNN	0.1669	0.0359	0.2201
	CNN	0.1885	0.0407	0.2427
	SVM	0.1761	0.0415	0.2182
	Timexer	0.1649	0.0346	0.2098
	<b>Pyraformer</b>	<b>0.1309</b>	<b>0.0275</b>	<b>0.1687</b>
最大风速集	BPNN	0.1890	0.0307	0.2415
	CNN	0.2057	0.0367	0.2550
	SVM	0.1747	0.0293	0.2247
	Timexer	0.1924	0.0324	0.2460
	<b>Pyraformer</b>	<b>0.1599</b>	<b>0.0270</b>	<b>0.2044</b>



(a) 数据集 1 最小风速集下不同模型预测误差箱线图对比



(b) 数据集 1 最大风速集下不同模型预测误差箱线图对比

图 9 数据集 1 不同模型预测误差箱线图

Fig. 9 Boxplot of prediction errors for different models on Dataset 1

较高的预测精度和较强的稳定性。此外, 从图 9 可以看出, 其预测误差更小且范围更集中, 充分验证了所提基于 Pyraformer 时序金字塔的多模型融合集成优化机制的必要性和有效性。

### 3.4 GWO 集成输出优化策略的必要性分析

本节将在数据集 1 的最大风速集和最小风速集上进行确定性点预测实验, 以验证 GWO 集成输出优化策略的必要性。对比方法分别是: 对比方法分别是: 麻雀优化算法(sparrow search algorithm, SSA)、正余弦优化算法(sine cosine algorithm, SCA)、牛顿-拉夫逊优化算法(Newton-Raphson-based optimizer, NRBO)以及直接求和法(summation method, SUM)。算例结果如表 3 所示。此外, 图 10 分别在上述数据集上截取了连续 100 个时间间隔下的风速预测拟合数据, 以更加清晰地展示各方法之间的预测差异。

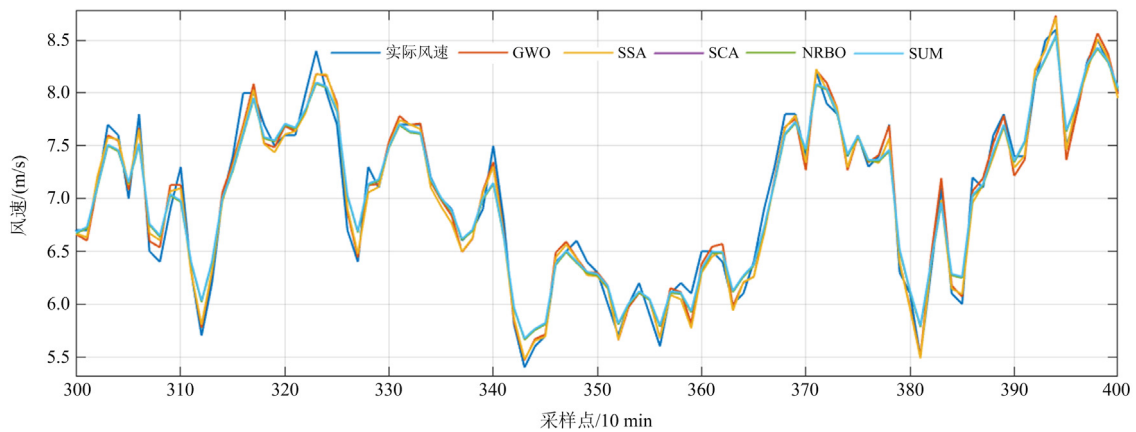
从表 3 和图 10 可以看出, 与 SUM 相比, SCA 以及 NRBO 算法没有表现出显著的集成优化效果, SCA 算法在最小风速集上甚至逊色于 SUM 法, 未能有效捕捉子序列与真实风速之间的特征相关性。相比之下, 本文所提 GWO 集成输出优化策略能够有效拟合风速在波峰、波谷的波动趋势, 在两个数据集上都取得了最佳的预测表现, 预测性能提升了

17.52%, 证明其能够在复杂的权重求解空间中快速适应风速数据的非线性特征, 高效捕捉子序列与真实风速之间的特征相关性, 提升整体预测精度, 验证了 GWO 集成输出优化策略的必要性和有效性。

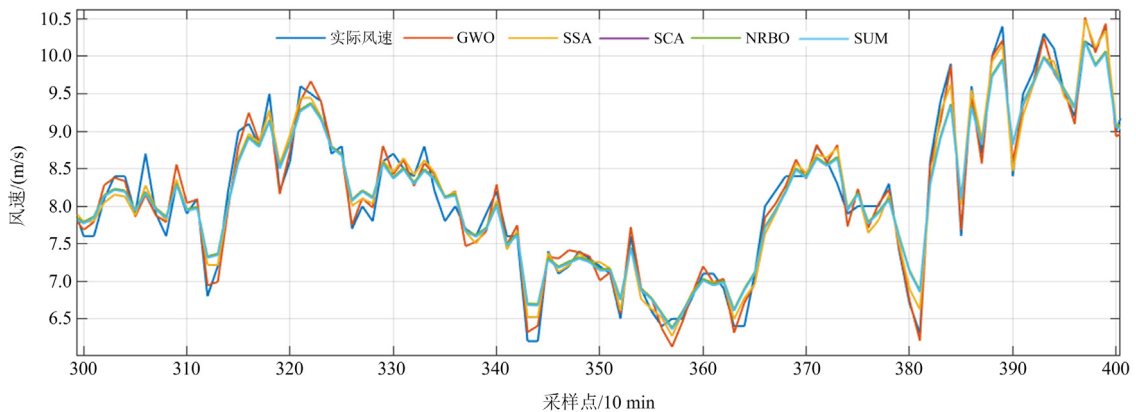
表 3 数据集 1 不同集成输出优化策略评价结果

Table 3 Evaluation results for different ensemble output strategies on Dataset 1

数据集 1	集成输出策略	评价指标		
		MAE	MAPE	RMSE
最小 风速集	SUM	0.1309	0.0275	0.1687
	SSA	0.1233	0.0258	0.1543
	SCA	0.1313	0.0276	0.1692
	NRBO	0.1304	0.0272	0.1685
	<b>GWO</b>	<b>0.1076</b>	<b>0.0222</b>	<b>0.1346</b>
最大 风速集	SUM	0.1599	0.0270	0.2044
	SSA	0.1442	0.0246	0.1822
	SCA	0.1597	0.0270	0.2041
	NRBO	0.1585	0.0268	0.2025
	<b>GWO</b>	<b>0.1236</b>	<b>0.0206</b>	<b>0.1572</b>



(a) 数据集 1 最小风速数据集下不同集成输出优化策略拟合曲线对比



(b) 数据集 1 最大风速数据集下不同集成输出优化策略拟合曲线对比

图 10 数据集 1 不同集成输出优化策略拟合曲线

Fig. 10 Fitting curves of different ensemble output optimization strategies on Dataset 1

## 4 结论

本文提出了基于时序金字塔双层集成学习架构的短期风速区间值预测方法。该方法能够显著提升短期风速预测精度,具有较强的泛化能力和稳定性,主要结论如下。

1) 所提模型预测性能相比于对比模型至少提升了 19.51%,能够有效捕捉风速区间值序列包含的不确定信息和高维复杂波动特征,提升短期风速预测的准确性和稳定性。

2) 所提基于 Pyraformer 时序金字塔的多模型融合集成预测机制能够有效融合多模型性能优势,突破子序列对单一预测模型的依赖,进一步提升集成预测模型性能。

3) 所提基于 GWO 的集成输出优化策略能够有效捕捉各子序列与真实风速之间不同的特征相关性,提升整体预测精度。

### 参考文献

- [1] 胡宇晗, 朱利鹏, 李佳勇, 等. 融合深度误差反馈学习和注意力机制的短期风电功率预测[J]. 电力系统保护与控制, 2024, 52(4): 100-108.  
HU Yuhan, ZHU Lipeng, LI Jiayong, et al. Short-term wind power forecasting with the integration of a deep error feedback learning and attention mechanism[J]. Power System Protection and Control, 2024, 52(4): 100-108.
- [2] AI X, LI S, XU H. Short-term wind speed forecasting based on two-stage preprocessing method, sparrow search algorithm and long short-term memory neural network[J]. Energy Reports, 2022, 8: 14997-15010.
- [3] 田晓煜, 赵海宇, 翟文辉, 等. 适用于混合风电场频率电压支撑的分布式自适应解耦控制[J]. 电力系统保护与控制, 2025, 53(6): 126-139.  
TIAN Xiaoyu, ZHAO Haiyu, ZHAI Wenhui, et al. Distributed adaptive decoupling control for frequency-voltage support in hybrid wind farms[J]. Power System Protection and Control, 2025, 53(6): 126-139.
- [4] 李东东, 张先明, 姚寅, 等. 考虑风速时空分布及机组运行状态差异的风电场有效惯量估计[J]. 电力建设, 2024, 45(1): 112-124.  
LI Dongdong, ZHANG Xianming, YAO Yin, et al. Estimation of effective inertia of wind farms considering temporal and spatial distribution of wind speed and differences in unit operating states[J]. Electric Power Construction, 2024, 45(1): 112-124.
- [5] 杨茂, 郭镇鹏, 王达, 等. 基于图神经网络的短期风电功率群体预测方法[J]. 电力系统保护与控制, 2025, 53(19): 79-88.  
YANG Mao, GUO Zhenpeng, WANG Da, et al. Short-term wind power group forecasting method based on graph neural networks[J]. Power System Protection and Control, 2025, 53(19): 79-88.
- [6] ZHAO S, SHAO C, DING J, et al. Unreliability tracing of power systems for identifying the most critical risk factors considering mixed uncertainties in wind power output[J]. Protection and Control of Modern Power Systems, 2024, 9(5): 96-111.
- [7] 余光正, 信德政, 吕辉清, 等. 考虑有限信息背景下的风电场短期功率预测方法[J]. 电力建设, 2025, 46(9): 130-143.  
YU Guangzheng, XIN Dezheng, LÜ Huiqing, et al. Short-term power prediction method for wind farms considering limited information background[J]. Electric Power Construction, 2025, 46(9): 130-143.
- [8] 张志艳, 兰龙, 白景升, 等. 小样本条件下风电功率预测方法的研究[J]. 电力系统保护与控制, 2025, 53(18): 120-130.  
ZHANG Zhiyan, LAN Long, BAI Jingsheng, et al. Research on wind power forecasting methods under small-sample conditions[J]. Power System Protection and Control, 2025, 53(18): 120-130.
- [9] 张野, 李凤婷, 张高航, 等. 考虑风电爬坡备用需求的风电高渗透电力系统优化调度方法[J]. 电力系统保护与控制, 2024, 52(23): 95-106.  
ZHANG Ye, LI Fengting, ZHANG Gaohang, et al. Optimization and scheduling methods for wind power high-penetration power systems considering wind power ramping reserve requirements[J]. Power System Protection and Control, 2024, 52(23): 95-106.
- [10] ZHOU F, PANG J, WANG B, et al. Analysis and suppression on tower vibration of permanent magnet synchronous generator-based wind energy conversion system[J]. Protection and Control of Modern Power Systems, 2025, 10(3): 65-82.
- [11] ZHU J, ZHENG P, NIU L, et al. An enhanced interval-valued PM2.5 concentration forecasting model with attention-based feature extraction and self-adaptive combination technology[J]. Expert Systems with Applications, 2025, 264.
- [12] ZOU R, SONG M, WANG Y, et al. Deep non-crossing probabilistic wind speed forecasting with multi-scale features[J]. Energy Conversion and Management, 2024, 44(11): 88-93, 133.
- [13] LV M, WANG J, WANG S, et al. A multi-input and dual-output wind speed interval forecasting system based

- on constrained multi-objective optimization problem and model averaging[J]. *Energy Conversion and Management*, 2024, 319.
- [14] WANG J, ZHANG H, LI Q, et al. Design and research of hybrid forecasting system for wind speed point forecasting and fuzzy interval forecasting[J]. *Expert Systems with Applications*, 2022, 209.
- [15] YANG W, ZANG X, WU C, et al. A new multi-objective ensemble wind speed forecasting system: mixed-frequency interval-valued modeling paradigm[J]. *Energy*, 2024, 304.
- [16] WANG J, LU Y, LI Q. An interpretable interval-valued wind power prediction system based on multi-objective feature extraction and base model selection with dynamic ensemble[J]. *Swarm and Evolutionary Computation*, 2025, 96.
- [17] HAO Y, WANG X, WANG J, et al. A new perspective of wind speed forecasting: multi-objective and model selection-based ensemble interval-valued wind speed forecasting system[J]. *Energy Conversion and Management*, 2024, 299.
- [18] ZHANG L, WANG J, NIU X, et al. Ensemble wind speed forecasting with multi-objective Archimedes optimization algorithm and sub-model selection[J]. *Applied Energy*, 2021, 301.
- [19] WANG C, ZHANG S, LIAO P, et al. Wind speed forecasting based on hybrid model with model selection and wind energy conversion[J]. *Renewable Energy*, 2022, 196: 763-781.
- [20] ZHANG Y, CHEN B, PAN G, et al. A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting[J]. *Energy Conversion and Management*, 2019, 195: 180-197.
- [21] DRAGOMIRETSKIY K, ZOZZO D. Variational mode decomposition[J]. *IEEE Transactions on Signal Processing*, 2013, 62(3): 531-544.
- [22] FU S, LI K, HUANG H, et al. Red-billed blue magpie optimizer: a novel metaheuristic algorithm for 2D/3D UAV path planning and engineering design problems[J]. *Artificial Intelligence Review*, 2024, 57(6): 134-143.
- [23] 蔡文豪. 基于改进 Pyraformer 的光伏发电功率预测[D]. 北京: 华北电力大学, 2024.  
CAI Wenhao. Photovoltaic power prediction based on improved Pyraformer[D]. Beijing: North China Electric Power University, 2024.
- [24] LIU S, YU H, LIAO C, et al. Pyraformer: low-complexity pyramidal attention for long-range time series modeling and forecasting[C]// *International Conference on Learning Representations (ICLR)*, April 25-29, 2022.
- [25] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [26] SHAH D, CAMPBELL W, ZULKERNINE F H. A comparative study of LSTM and DNN for stock market forecasting[C]// *2018 IEEE International Conference on Big Data*, December 10-13, 2018, WA, USA: 4148-4155.
- [27] WANG S, SHI J, YANG W, et al. High and low frequency wind power prediction based on Transformer and BiGRU-Attention[J]. *Energy*, 2024, 288.
- [28] 张翺, 李红月. 基于优化 VMD 与 TCN-ISE-Pyraformer 的短期电力负荷预测[J]. *安徽工业大学学报(自然科学版)*, 2025, 42(3): 263-275.  
ZHANG Xuan, LI Hongyue. Short-term electric load forecasting based on optimized VMD and TCN-ISE-Pyraformer[J]. *Journal of Anhui University of Technology (Natural Science)*, 2025, 42(3): 263-275.
- [29] LIU X, YANG K, ZOU B, et al. Accurate estimation of state of health for lithium-ion batteries based on Pyraformer and TimeGAN data augmentation[J]. *Journal of Power Sources*, 2025, 640.
- [30] 蔡华锋, 唐艺豪. 基于 Pyraformer 网络的短期电力负荷预测[J]. *武汉大学学报(工学版)*, 2023, 56(9): 1105-1113.  
CAI Huafeng, TANG Yihao. Short-term power load forecasting based on Pyraformer network[J]. *Engineering Journal of Wuhan University*, 2023, 56(9): 1105-1113.
- [31] MIRJALILI S, MIRJALILI SM, LEWIS A. Grey wolf optimizer[J]. *Advances in Engineering Software*, 2014, 69: 46-61.
- [32] SHANG Z, CHEN Y, LAI D, et al. A novel interpretability machine learning model for wind speed forecasting based on feature and sub-model selection[J]. *Expert Systems with Applications*, 2024, 255.
- [33] WANG Y, WU H, DONG J, et al. Timexer: empowering transformers for time series forecasting with exogenous variables[J]. *Advances in Neural Information Processing Systems*, 2024, 37: 469-498.

收稿日期: 2025-07-18; 修回日期: 2025-11-24

作者简介:

冯涛(1999—), 女, 硕士研究生, 研究方向为风电功率预测; E-mail: fengtao20230324@wust.edu.cn

艾学轶(1983—), 女, 通信作者, 博士, 副教授, 研究方向为大规模新能源功率预测技术; E-mail: aixueyi1030@wust.edu.cn

韦善阳(1993—), 男, 博士(后), 研究方向为电力系统负荷预测。E-mail: shanyangwei@hust.edu.cn

(编辑 魏小丽)