电能质量参数集成预测

Max Domagk Peter
Feistel Jan Meyer 德累斯顿工业大学，德国德累斯顿
m
ax.do
magk@tu-dresden.de

Marco Lindner
TransnetBW 德国斯图加特

Jako Kilter 塔林理工大学
爱沙尼亚塔林

摘要——**随着电力电子技术的应用日益广泛，输电系统对电能质量（PQ）监测的需求也随之增加。尽管运营商收集了大量数据集，但这些数据通常仅用于合规性评估。中长期预测能够通过支持主动式资产管理和趋势检测，提升这些数据集的价值，尽管数据异质性和季节性仍带来挑战。本文系统评估了输电系统电能质量参数的单个模型和集成预测方法。研究分析了来自德国和爱沙尼亚测量活动的 700 多条周时间序列，以便在结构化集成框架内评估各种模型和聚合策略。结果表明，就准确度和稳健性而言，集成预测始终优于单个模型，其效果显著胜过季节性朴素基准及表现最佳的单一模型。因此，集成预测被证实是输电系统电能质量长期预测的一种稳健且可扩展的方法。**

关键词——电能质量、输电系统、时间序列分析、长期预测、集成预测

I. 引言

随着风能和太阳能发电、电池储能系统、高压直流输电系统以及柔性交流输电系统装置等基于逆变器的技术日益普及，电能质量在输电网络中的重要性不断提升。因此，输电系统运营商加大了对高压及特高压电网的电能质量监测力度。迄今为止，对这些数据的分析主要局限于针对规划水平开展合规性评估。然而，对未来数月的电能质量参数进行准确的中长期预测，有助于及早发现不利趋势和潜在的阈值超限。这类预测有助于制定主动维护规划并支持明智的投资决策，

尽管中长期预测电能质量参数具有重要意义，但针对输电等级的相关研究仍然有限。有效预测电能质量时间序列需要具备若干关键特性。稳健性对于应对各种电能质量参数和异质数据源至关重要，而处理大量时间序列时，计算效率也不可或缺。此外，除了点预测准确性之外，还必须考虑预测可靠性，因为单个模型不可避免地存在结构性局限。在此背景下，集成多个预测模型有望提升稳健性并降低模型不确定性[1]。

本文通过对历史电能质量数据进行统计分析来应对这些挑战。首先介绍适用于电能质量参数中长期预测的预测模型，随后构建一个系统的集成预测框架。然后对单个模型和集成模型的性能进行评估和比较，以验证集成方法用于电能质量预测的稳健性和实际应用价值。

II. 数据特征与预处理

A. 电能质量参数变化

电能质量参数会随时间变化，这是受到许多不同因素影响所致，且这些因素主要与电气环境有关。通常可以区分三种不同类型的变化：短期、中期和长期变化 [2]。短期变化通常以日和周为周期，主要受用户和发电行为的影响。某些用户配置呈现出明显不同的日模式（例如，工作日与周末）。

持续数月的中期变化与用户和发电行为的整体变化有关，而这些变化大多与季节性影响相关。冬季夜幕降临较早，可能导致照明等设备的使用增加，进而造成谐波排放升高。光伏系统在夏季产生的能量显著更多（白昼更长，日照角度更高）。

本文聚焦于中期至长期变化，其中长期指持续数月至一年的时间范围，这对于可靠的系统性能预测至关重要。

B. 数据集

该数据集来自德国和爱沙尼亚两家输电系统运营商在其输电系统内开展的两次测量活动。数据集包含德国的14个测量站点和爱沙尼亚的13个测量站点，覆盖不同电压等级：110 kV（17个站点）、220 kV（2个站点）、330 kV（2个站点）和380 kV（6个站点）。所有站点的测量时长均至少为3年。数据采集采用符合IEC 61000-4-30 A级标准的固定式电能质量仪器。

监测的电能质量参数包括多种电压质量参数：长时间闪变严重度（Uplt）、电压不平衡（UNB）、电压总谐波畸变（Uthd）以及各次谐波电压（U03、U05、U07、...、U15），记录间隔为10分钟，并按此进行聚合。

C. 数据预处理

首先，对原始的10分钟测量值进行聚合，按每个完整日历周（周一至周日）计算 95^{th} 百分位数，得到周值。如果一周内至少有95 %的10分钟测量数据可用，则该周视为有效；为提高数据可用性，每周最多允许缺失8.5小时的数据。

生成的周时间序列中经常会出现缺失周或间隔，例如由于维护

停机所致。为最大限度地增加可用时间序列的数量，允许缺失周数最多占 20%。使用最近 10 周内可用的最新每周 95th 百分位数来填补间隔。

随后，将每周的数值与各自的规划水平进行比较，以此进行归一化，从而得到相应的利用率。所采用的规划水平遵循国家电网规范：在德国，高压电网采用 VDE-AR-N 4120，超高压电网采用 VDE-AR-N 4130；在爱沙尼亚，则采用输电系统运营商规定的规划水平，这些水平与 IEC TR 61000-3-6 保持一致。

完成全部预处理后，共有 716 条涵盖 3 年的时间序列可供分析。图 1 展示了一条示例时间序列，说明电压不平衡的利用率。

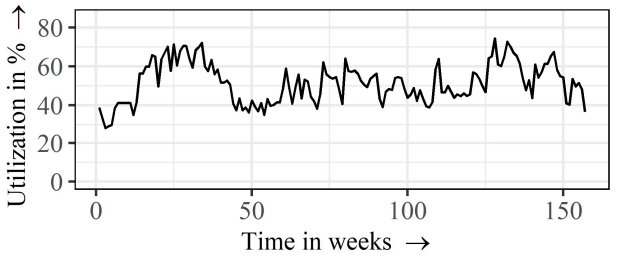


图 1. 德国输电系统 220 kV 电压等级下某测量站点电压不平衡规划等级的利用情况

III. 预测模型

预测时域从超短期（秒至分钟）到长期（数年至数十年）不等。尽管机器学习和深度学习在短期预测方面表现出色，但由于依赖特定的趋势和周期性结构，它们在长期预测中的效能往往受到限制 [3]。此外，这类“黑箱”方法通常缺乏可解释性，且需要大量计算资源 [4]。因此，本研究采用可解释的模型，将时间序列分解为趋势、季节性和余项成分。研究使用指数平滑、ARIMA/SARIMA 和 Prophet，因为这些方法计算效率高、经过充分验证，并且在各种应用中具有稳健性。

A. 模型概述预测方法的一类

包括简单的基准方法。朴素法将最后一个观测数据点外推至所有未来值，而漂移法则外推历史平均变化量。对于季节性时间序列，季节性朴素法（SNaive）将未来值设为上一周期对应时段的值，并作为基准准确度评估所必需的持续性基线。

另一类重要的预测技术是指数平滑法。指数平滑（ES）将过去的观测值进行指数加权平均，以生成预测 [5], [6]。霍尔特-温特斯法（HW）等扩展方法会纳入趋势和季节性成分。

更复杂的统计方法包括ARIMA（自回归积分滑动平均）模型。Box-Jenkins方法使用非季节性（ARIMA）和季节性（SARIMA）对自相关结构进行建模差分 [7]。这些模型对具有强自相关性和明确周期的数据十分有效。

近年来广受欢迎的一种现代预测框架是 Prophet，它是一种由趋势、季节性和节假日成分构成的加性回归框架 [8]。该方法使用分段线性或逻辑斯蒂增长曲线建模趋势，并使用傅里叶级数处理多种季节性，能够灵活应对周期性和非周期性变化。

最后，基于分解的方法为时间序列预测提供了一种直观的途径。STL分解（基于局部加权回归的季节性和趋势分解）可将序列拆分为独立成分，并对各成分分别应用不同的预测方法 [9]。例如，可对季节性成分应用季节性朴素法，同时对季节调整序列使用 ARIMA [10]。

B. 选定模型

总共有八种个体模型适用于电能质量参数预测，列于表 I 中。

表 I。 个体预测模型概述

模型	描述
季节性朴素法	季节性朴素法（基准）
HW	霍尔特-温特斯法
SARIMA	季节性ARIMA
Prophet	含趋势和季节性的加法模型
STL-Drift	基于漂移法的 STL 分解
STL-ES	基于指数平滑的 STL 分解
STL-Holt	基于 Holt 法的 STL 分解
STL-ARIMA	基于 ARIMA 法的 STL 分解

对于图 1 所示的电压不平衡利用率，图 2 示例性地展示了三种具有代表性的模型预测。该三年数据集被划分为 105 周的训练集和 52 周的测试集，以进行评估。在此示例中，HW 模型的预测值略高，而 Prophet 和 SNaive 模型的预测值略低。

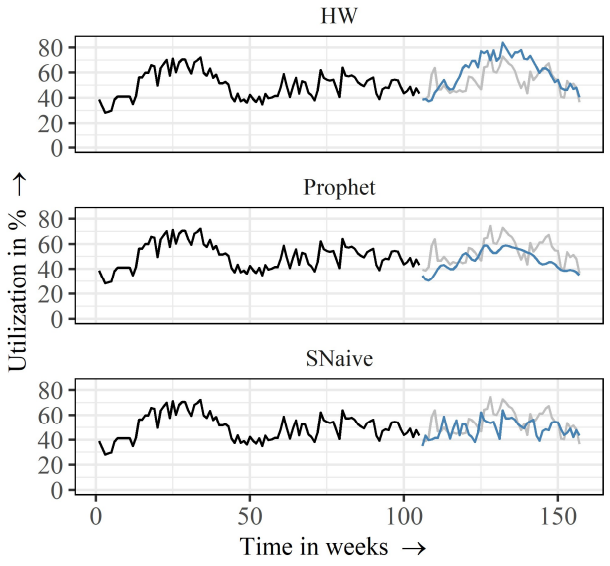


图 2. 图 1 中电压不平衡利用情况的示例预测，采用三种不同模型；黑色表示训练集（105 周）；测试集以灰色表示（52周预测期），模型预测以蓝色表示

IV. 集成预测

A. 集成预测框架

集成预测基于生成各个（基准）预测，然后将它们组合起来。该过程通常包括两个阶段：预测阶段，即根据给定的时间序列生成多个单独的预测；组合阶段，即使用指定的组合函数汇总这些预测，从而形成预测集成。生成单独预测的方法以及组合这些预测的策略可能存在很大差异。可以使用自助法、层次预测或相互独立、互不干扰的模型等方法来生成预测集成 [1]。

本文仅关注第三类，即仅考虑由相互独立、互不干扰的预测模型构成的集成。

B. 集成模型

使用表 I 中列出的八种个体预测模型，可以构建不同规模的集成模型，组合数量从两个模型到全部八个模型不等。本文考虑了所有 247 种可能的配置，并为每个集成指定一个简称：字母表示集成规模（B = 2 个模型、C = 3 个模型，……，H = 8 个模型），数字表示其在该组中的序号。例如，B01-B28 对应由 2 个模型组成的集成，C01-C56 对应由 3 个模型组成的集成，而 H01 则表示由全部 8 个模型组成的集成。

C. 预测组合方法

预测组合在提升预测准确度和降低模型选择风险方面有着悠久的历史，大量实证证据表明，集成通常优于其个体组成部分 [11]。等权均值被广泛视为稳健的基准，能够有效降低方差，并在各种时间序列上表现良好。中位数组合提供了另一种选择，对异常值不那么敏感；不过，均值与中位数结果的比较并无定论，且很大程度上取决于数据 [1]。

加权集成旨在让表现更好的模型发挥更大作用，例如根据近期模型准确度或相对表现分配权重 [1]。本研究考虑了四种组合方法，并将其应用于全部 247 种集成配置：

- **均值**（对个体预测进行等权平均）• **中位数**（对个体预测逐点取中位数）• **对称平均绝对百分比误差**（基于模型平均表现进行准确度加权的集成）• **排名**（基于模型平均表现排序进行排名加权的集成）

采用对称平均绝对百分比误差和排名构建非等权集成预测 $\hat{y}_{\text{ens},\,T+h|T}$ 遵循对 M 个体模型预测 $\hat{y}_{i,\,T+h|T}$进行线性组合的方式：

$$\hat{y}_{\text{ens},T+h|T} = \sum_{i=1}^{M} w_i \cdot \hat{y}_{i,T+h|T} \qquad (1)$$

为确保准确度更高（误差更低或排名更低）的模型获得更大权重，权重w_i 定义为由某一性能指标 ϕ_i 在所有模型间归一化后的倒数：

$$w_i = \phi_i^{-1} / \sum_{j=1}^{M} \phi_j^{-1} \qquad (2)$$

输入 ϕ_i 表示模型 i 在性能评估程序中得出的平均对称平均绝对百分比误差或平均排名。

图 3 展示了集成 C01 的一个预测，该预测是通过对图 2 中三个单个模型的点预测取均值得出的。

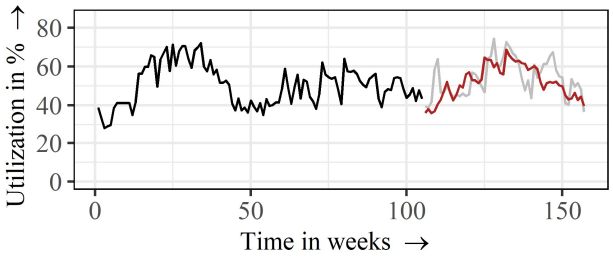


图 3。　利用图 2 中三种不同模型对电压不平衡进行集成预测；训练集（105 周）以黑色表示；测试集以灰色表示（52周预测期），集成预测以红色

V. 性能评估

A. 性能指标

点预测准确度采用平均绝对误差 (MAE) 和对称平均绝对百分比误差 (sMAPE) 进行评估。对于预测值 $\hat{y}_{T+h|T}$ 、实际观测值 y_{T+h} 以及预测时域 $H = 52$，sMAPE 定义如下：

$$\text{sMAPE} = \frac{200}{H} \sum_{h=1}^{H} \frac{|y_{T+h} - \hat{y}_{T+h|T}|}{|y_{T+h}| + |\hat{y}_{T+h|T}|} \qquad (3)$$

该指标的取值范围限定为 0 % - 200 %，可为评估预测性能提供均衡的度量 [12]。在本分析中，低于 10 % 视为良好，最高 25 % 视为可接受，超过此范围则视为较差或很差。

B. 性能评估程序

通过对所有时间序列和模型使用105周训练期和52周预测时域来保持一致性。按照与[13], 类似的方法，准确度分三步评估：

1. 计算每个模型和每条时间序列在整个预测时域上的准确度（平均绝对误差和对称平均绝对百分比误差）。2. 确定每条时间序列中每个模型的排名R(对称平均绝对百分比误差)。3. 计算每个模型在全部716个序列上的平均准确度（平均绝对误差$\overline{\text{MAE}}$ 和对称平均绝对百分比误差$\overline{\text{sMAPE}}$）以及平均排名（R(对称平均绝对百分比误差)$\overline{\text{R(sMAPE)}}$）。

这些汇总指标（对称平均绝对百分比误差$\overline{\text{sMAPE}}$ 和R(对称平均绝对百分比误差)$\overline{\text{R(sMAPE)}}$ ）随后被用作构建加权集成的输入，如第 IV.C 节所述。为评估相对于基准模型的性能，对任意模型i 定义基准比率（基准比率）如下：

$$\text{BR}_i = \overline{\text{sMAPE}}_i / \overline{\text{sMAPE}}_{\text{SNaive}} \qquad (4)$$

C. 个体模型性能

表 II 概述了全部 716 条时间序列上八个个体预测模型的性能。

TABLE II. 个体预测MO的性能 DELS

#	模型	$\overline{\text{MAE}}$	sMAPE	R(sMAPE)	BR
1	STL-ARIMA	3.9%	18.22 %	3.34	0.873
2	STL-ES	3.93 %	18.67 %	3.28	0.894
3	SARIMA	4.08 %	19.38 %	4.08	0.928
4	季节性朴素法	4.62 %	20.88 %	5.15	1
5	Prophet	4.51 %	21.41 %	4.3	1.025
6	STL-Holt	4.49 %	21.53 %	4.56	1.031
7	STL-Drift	5.34 %	24.2 %	5.39	1.159
8	HW	5.15 %	25.03 %	5.9	1.199

结果表明，基于分解的模型表现出明显优势，其中 STL-ARIMA 和基于 STL 的指数平滑法表现最佳，取得了最低的平均 sMAPE（18.22 % 和 18.67 %）以及最佳平均排名。只有这些 STL 变体和 SARIMA 成功超越了季节性朴素基准（BR ＜1），准确率提升最高达 12.7 %；而 Prophet 和 HW 等更复杂的模型则未能胜过基准。这些结果验证了平均排名的稳健性以及各项指标的互补性。例如，尽管 Prophet 的 sMAPE（21.41 %）高于季节性朴素法（20.88 %），但它们的平均排名顺序并未反映出这一差异（4.56 对 5.15）。这表明，即使误差幅度有所不同，序数可靠性仍具有竞争力，从而证明了基于排名的集成方法是合理的。

D. 集成模型表现对 988 个集成的分析

种组合模型（采用四种组合方法形成的 247 种配置）表现显著优于单个模型。如表 III 所示，集成模型始终具有更高的准确度，排名前 100 的模型均优于表现最佳的个体模型（STL-ARIMA，对称平均绝对百分比误差为 = 18.22 %）。

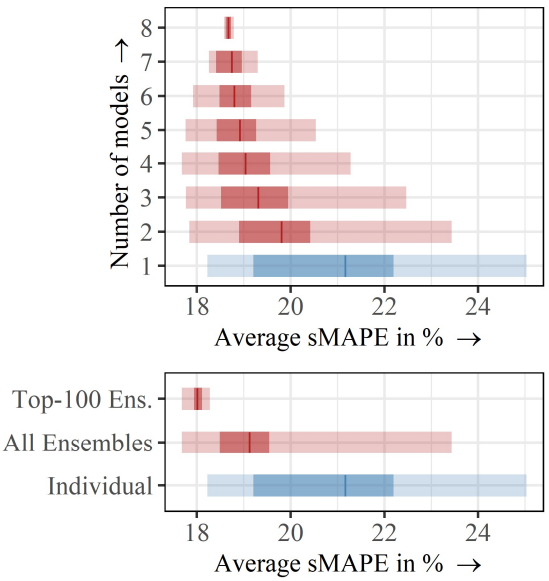


图4. 按集成规模（上）和模型集合（下）划分的平均对称平均绝对百分比误差，包含均值（实线）、25th-75th 百分位范围（深色条）和最小值至最大值范围（浅色条）。

值得注意的是，表现最佳的个体模型未能跻身集成模型前 100 名，这凸显了预测组合在稳健降低误差方面的能力。D28集成配置（由 STL-ARIMA、STL-ES、SARIMA 和季节性朴素法组成）脱颖而出，成为总体表现最佳的配置；其中，中位数组合取得了最低的对称平均绝对百分比误差 17.68 %，基准比率为 0.847。与表现最佳的个体模型相比，该配置提升了 2.6 %；与季节性朴素法基准相比，则提升了 15.3 %。

表 III. 排名前 100 的集成模型性能

#	模型	$\overline{\text{MAE}}$	$\overline{\text{sMAPE}}$	R sMAPE ( )	BR
1	D28 中位数	3.79 %	17.68 %	387.56	0.847
2	D16 中位数	3.76 %	17.74 %	388.91	0.85
3	D28 排名	3.81 %	17.74 %	372.36	0.85
4	E23 中位数	3.81 %	17.76 %	359.49	0.851
5	D28 对称平均绝对百分比误差	3.83 %	17.77 %	386.79	0.851
…	…	…	…	…	…
95	D26 对称平均绝对百分比误差	3.9 %	18.18 %	410.57	0.871
96	F18 对称平均绝对百分比误差	3.91 %	18.18 %	372.94	0.871
97	E33 中位数	3.89 %	18.19 %	427.86	0.871
98	E25 均值	3.9 %	18.19 %	393.94	0.871
99	E01 排名	3.87 %	18.19 %	368.09	0.871
100	D67 排名	3.86 %	18.19 %	370.8	0.871

如图4所示，增加集成中单个模型的数量，可持续降低平均对称平均绝对百分比误差，从而提高准确度，同时降低性能方差，增强稳健性；与单个模型相比，排名前100的集成能够实现显著更高的准确度。此外，图5显示，最成功的集成通常会将五个最佳单个模型中的四个或五个组合在一起。虽然中位数是排名前100的集成中最常用的方法，但具体的聚合方式最终并不如模型池的构成和规模重要。因此，集成的成功似乎取决于纳入多样化且表现良好的单个模型。

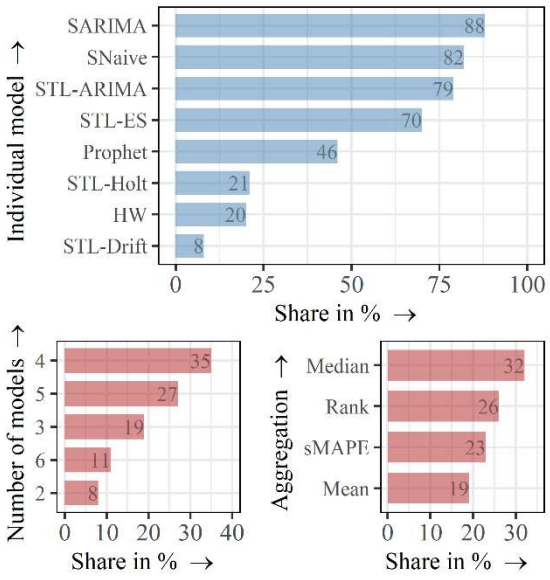


图5. 前100个集成模型的构成：单个模型占比（上）、集成规模分布（左下）以及各种聚合方法的利用（右下）

E. 最佳单体模型与集成模型的比较

图6展示了最佳个体模型和最佳集成模型的不同预测示例。

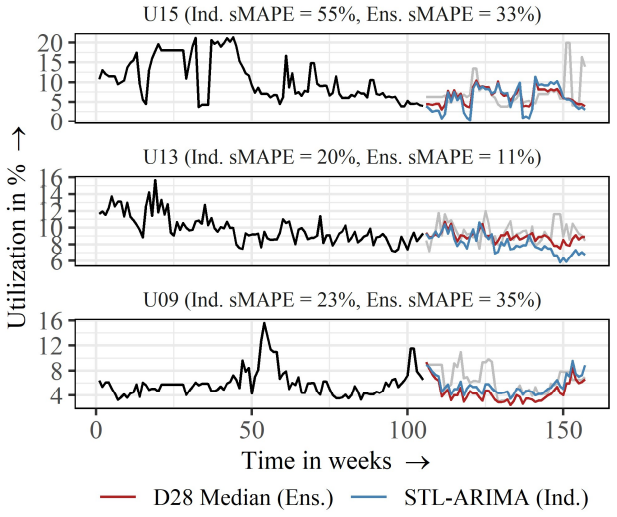


图6。 最佳个体模型（蓝色）与最佳集成模型（红色）的预测示例：训练集（105 周）以黑色表示；测试集以灰色表示（52周预测期）；从上到下：集成模型的表现明显优于、略优于和逊于最佳个体模型

为评估预测组合是否能在审慎的模型选择之外提供附加价值，本文对其性能进行了详细比较。相对对称平均绝对百分比误差改进（参见图7左）表明，最佳集成在大多数情况下优于最佳个体模型，在全部比较中有62%取得了更低的对称平均绝对百分比误差。当集成更优时，收益具有系统性且具有实际意义，中位数相对改进约为4%。尽管个体模型在少数情况下占优，集成的主要价值在于其稳健性。通过比较个体模型与集成的对称平均绝对百分比误差（图7右）可以看出，集成能够稳定性能，避免因选择单一次优模型而常出现的极端预测失败。

VI. 结论

本研究利用输电系统数据评估了用于中长期电能质量预测的集成预测。尽管 STL-ARIMA 和 STL-ES 等基于分解的模型表现出较强的单独预测能力，但没有任何单个模型能在所有时间序列中始终表现最优。集成方法能够有效降低模型选择风险，持续提高预测准确度并减少误差变异性。结果还表明，集成效果主要取决于单个模型的多样性，而非聚合方法的复杂程度，因为简单的均值或中位数组合已能带来稳健且可靠的改进。

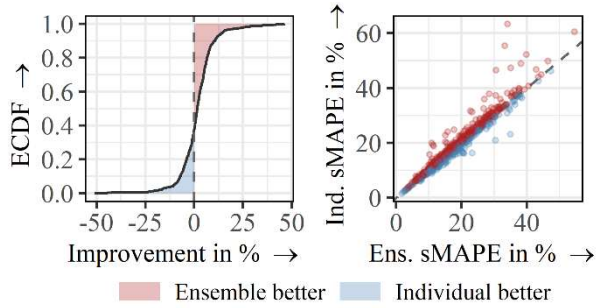


图7。 最佳个体模型与集成模型的比较：相对对称平均绝对百分比误差（sMAPE）改进的经验累积分布函数（ECDF）（左侧）以及个体模型与集成模型 sMAPE 的散点图（右侧）

因此，集成预测为输电系统运营商提供了一种可扩展且计算高效的框架，以加强趋势识别并支持基础设施规划。未来研究将聚焦于配电系统应用，将其扩展至概率预测，并开发非对称准确度指标，以优先保障电能质量限值附近的预测准确度，因为该区域的预测误差具有最大的运行影响。

致谢

本研究得到德国输电系统运营商 TransnetBW、爱沙尼亚输电系统运营商 Elering AS 和欧盟的支持，并由爱沙尼亚教育与科学部共同资助（项目编号：TemTa-134）。

参考文献

[1] X. Wang、R. J. Hyndman、F. Li 和 Y. Kang，《预测组合：五十余年的回顾》，《国际预测学杂志》，第39卷，第4期，第1518–1547页，2023年10月，doi：10.1016/j.ijforecast.2022.11.005。[2] M. Domagk 等，《基于长期测量活动的电能质量参数趋势分析》，载于《2022年第20届谐波与电能质量国际会议》（ICHQP），意大利那不勒斯，2022年5月，第1–6页。doi：10.1109/ICHQP53011.2022.9808645。[3] A. Zeng、M. Chen、L. Zhang 和 Q. Xu，《Transformer 对时间序列预测是否有效？》，《AAAI人工智能会议论文集》，第37卷，第9期，第11121–11128页，2023年6月，doi：10.1609/aaai.v37i9.26317。[4] P. F. Austnes、C. García-Pareja、F. Nobile 和 M. Paolone，《基于经验Copula的配电系统概率负荷预测》，《可持续能源、电网与网络》，第42卷，文章编号101708，2025年6月，doi：10.1016/j.segan.2025.101708。[5] C. C. Holt，《通过指数加权移动平均法预测季节性和趋势》，《国际预测学杂志》，第20卷，第5–10页，2004年，doi：10.1016/j.ijforecast.2003.09.015。[6] P. R. Winters，《通过指数加权移动平均法预测销售额》，《管理科学》，第6卷，第3期，第324–342页，1960年。[7] G. Box，《Box与Jenkins：时间序列分析、预测与控制》，载于《一桩非常英伦的往事：六位英国人与20世纪时间序列分析的发展》，T. C. Mills 编，伦敦：Palgrave Macmillan UK，2013年，第161–215页。doi：10.1057/9781137291264_6。[8] S. J. Taylor 和 B. Letham，《大规模预测》，《美国统计学家》，第72卷，第1期，第37–45页，2018年1月，doi：10.1080/00031305.2017.1380080。[9] R. Cleveland 和 W. Cleveland，《STL：一种基于局部加权回归的季节性-趋势分解方法》，《官方统计学杂志》，第6卷，第3–73页，1990年。[10] K. H. Lem，《用于季节性时间序列预测的STL-ARIMA方法：初步研究》，《ITM网络会议》，第67卷，文章编号01008，2024年，doi：10.1051/itmconf/20246701008。[11] M. Hibon 和 T. Evgeniou，《是否进行组合：预测及其组合的选择》，《国际预测学杂志》，第21卷，第1期，第15–24页，2005年1月，doi：10.1016/j.ijforecast.2004.05.002。[12] R. J. Hyndman 和 G. Athanasopoulos，《预测：原理与实践》，第2版。OTexts，2018年。[13] S. Makridakis、E. Spiliotis 和 V. Assimakopoulos，《M4竞赛：10万条时间序列与61种预测方法》，《国际预测学杂志》，第36卷，第1期，第54–74页，2020年1月，doi：10.1016/j.ijforecast.2019.04.014。