

基于 STD-ST-Former 的现货电价长步时空预测

张鹏飞¹, 胡博², 胡展硕³, 罗桓桓⁴, 刘桁宇⁵, 邢作霞¹

(1. 沈阳工业大学电气工程学院, 辽宁省 沈阳市 110870; 2. 国网大连供电公司, 辽宁省 大连市 116001;
3. 沈阳工程学院国际教育学院, 辽宁省 沈阳市 110136; 4. 辽宁电力交易中心有限公司,
辽宁省 沈阳市 110055; 5. 国网辽宁省电力有限公司电力科学研究院, 辽宁省 沈阳市 110000)

Long Step Spatial-temporal Prediction of Spot Electricity Price Based on STD-ST-Former

ZHANG Pengfei¹, HU Bo², HU Zhanshuo³, LUO Huanhuan⁴, LIU Hengyu⁵, XING Zuoxia¹

(1. School of Electrical Engineering, Shenyang University of Technology, Shenyang 110870, Liaoning Province, China; 2. State Grid Dalian Electric Power Supply Company, Dalian 116001, Liaoning Province, China; 3. College of International Education, Shenyang Institute of Engineering, Shenyang 110136, Liaoning Province, China; 4. Liaoning Power Exchange Center Co., Ltd., Shenyang 110055, Liaoning Province, China; 5. Electric Power Research Institute of State Grid Liaoning Electric Power Co., Ltd., Shenyang 110000, Liaoning Province, China)

ABSTRACT: Accurate prediction of spot price can provide beneficial guidance for market participants in trading strategies. In order to improve the long step spatial-temporal prediction accuracy of regional spot electricity price, a prediction framework using dual channel spatial-temporal transformer (ST-Former) based on seasonal trend decomposition is proposed. This prediction framework adapts the idea of decomposition, and obtain more predictable results including a stable trend part and a fluctuating seasonal part by untangling the entangled time patterns in regional spot electricity prices. On this basis, a ST-Former model based on the combination of variable Patch temporal attention and spatial self-attention is proposed, to explore the long step spatial-temporal features of regional spot electricity prices. Then, the ST-Former model is used to model the trend part and seasonal part, forming a dual channel model framework. Finally, the output results of the dual channel model framework are aggregated to obtain the final prediction of regional spot electricity prices. Taking the spot electricity price data of the Australian electricity market as an example, the effectiveness and superiority of the proposed prediction framework are verified by comparing with twelve prediction methods.

KEY WORDS: electricity market; prediction of regional spot electricity price; seasonal trend decomposition; attention mechanism; spatial-temporal correlation; long step spatial-

temporal prediction

摘要: 精准的现货电价预测能够为市场参与者的交易策略提供有利的指导, 为提高区域现货电价的长步时空预测精度, 提出一种基于季节趋势分解-双通道时空 transformer(spatial-temporal transformer, ST-Former)的区域现货电价预测框架。该预测框架采用分解的思想, 通过解开区域现货电价中纠缠的时间模式, 从而获得更可预测的稳定的趋势部分和波动的季节部分, 在此基础上, 提出一种基于可变 Patch 时间注意力和空间自注意力相结合的 ST-Former 模型, 以挖掘区域现货电价的长步时空特征; 然后, 通过 ST-Former 模型分别对趋势部分和季节部分进行建模构成双通道模型框架; 最后, 将双通道模型框架的输出结果进行聚合得到最终预测结果, 实现区域现货电价的长步时空预测。以澳大利亚电力市场的区域现货电价数据为例, 与 12 种预测方法进行对比, 验证所提预测框架的有效性和优越性。

关键词: 电力市场; 区域现货电价预测; 季节趋势分解; 注意力机制; 时空相关性; 长步时空预测

0 引言

在全球低碳化转型的背景下, 电力市场在推动电能高效利用、优化电力资源配置等方面具有重要作用, 目前已有 30 多个国家和地区拥有电力市场, 不少发展中国家开始探索和建立电力市场制度^[1]。精准的电价预测对电力市场稳定运行具有重要意义, 也为市场参与者的交易策略提供有利的指导。从现有运营的电力市场来看, 现货市场电价波动大, 容易受到供需关系、地理位置、突发事件、环

基金项目: 国家自然科学基金项目(U22B20115)。

Project Supported by National Natural Science Foundation of China (U22B20115).

境变化、投标策略、新能源波动等方面的影响^[2], 相比于中长期电价预测难度更大, 因此现货电价预测得到了学者们的广泛关注^[3-5]。

目前, 针对现货电价预测的相关研究, 大多数是将现货电价预测作为时间序列预测问题来处理, 使用统计模型、机器学习、深度学习等方法进行预测。经典的统计模型具有较强的通用性和可解释性, 文献[6]采用自回归移动平均模型(auto-regressive integrated moving average model, ARIMA)预测短期电价, 并以此为依据优化市场竞价策略; 文献[7]提出一种基于季节函数的自回归移动平均模型(auto-regressive moving average with exogenous input, ARMAX)来预测边际电价。统计模型在平稳的电价预测中效果良好, 但在随机性大、非平稳的现货电价预测中能力有限。

随着机器学习的快速发展, 许多机器学习模型被应用在现货电价预测中, 文献[8]将随机森林(random forest, RF)算法应用到现货电价预测中, 与基线算法对比提升了 25%; 文献[9]提出一种基于在线顺序极限学习机(online sequential extreme learning machine, OS-ELM)的电价预测方法; 文献[10]研究不同机器学习模型在电价预测方面的效果, 并考虑多种特征因素来提高电价预测效果。由于影响电价波动所涉及的因素较多, 如果没有悉心设计的特征工程, 传统的机器学习方法并不能较好地预测。

深度学习作为近年来的热点研究, 在现货电价预测方面得到了广泛的应用。文献[11]提出一种基于快速集成经验模态分解(fast ensemble empirical mode decomposition, FEEMD)、变分模态分解(variational mode decomposition, VMD)和基于萤火虫算法(firefly algorithm, FA)优化的反向传播(back propagation, BP)神经网络的混合模型, 利用 VMD 将 FEEMD 产生的高频固有模态函数(intrinsic mode functions, IMFs)进一步分解为多个模态, 以提高预测精度; 文献[12]提出一种基于最大信息系数与改进多层级门控长短期记忆网络(long short-term memory, LSTM)的短期电价预测方法; 文献[13]提出一种基于奇异谱分析和纵横交叉算法优化长短期记忆网络的日前电价预测模型; 文献[14]提出一种基于深度学习的多步电价预测模型, 通过捕捉陡峭的价格变化来提高电价预测效果; 文献[15]提出一种基于双向长短期记忆网络(bidirectional long

short-term memory, Bi-LSTM)的分布式节点边际价格的概率预测方法, 通过捕捉电价复杂的时间动态特征来提高预测精度; 文献[16]提出一种基于随机森林、改进马氏距离(improved mahalanobis distance, IMD)、改进集合经验模态分解(improve complete ensemble empirical mode decomposition, ICEEMD)、变分模态分解和双向长短期记忆网络的电价预测模型, 该混合模型对分解后的多个子序列进行预测, 综合处理后得到最终的电价预测结果; 文献[17]提出一种基于 Transformer 的电价预测模型来辅助市场参与者的短期交易策略, 该预测模型可以准确地对实时电价进行概率预测, 同时能够为市场参与者提供有利的决策辅助; 文献[18]提出一种基于最大信息系数、集成经验模态分解和改进 Informer 的短期电价多步预测模型, 实现未来一天内 24 点的多步预测。

上述文献将现货电价预测做为一元时间序列预测问题来处理, 未考虑不同现货电价之间的相关性, 忽略了现货电价数据中蕴含的时空特性。近些年, 学者们通过加入空间相关性来提高现货电价的预测精度。文献[19]提出一个具有时间和空间滞后的电价预测模型, 通过统计检验显示, 跨区域的电价具有显著的空间相关性, 电价的空间相关性与时空相关性一样重要; 文献[20]提出一种基于图卷积网络(graph convolutional network, GCN)和长短期记忆网络的预测模型, 对日前市场边际电价进行预测; 文献[21]提出一种基于图卷积网络和时间卷积网络(temporal convolutional network, TCN)结合的预测模型来预测节点边际电价, 并设计了三支结构来匹配节点边际电价的构成, 以提高预测模型的可解释性。

虽然以上研究工作极大地推动了现货电价预测领域的发展, 但目前的现货电价预测方法大多是在短步预测下设计的, 比如预测未来现货电价的 1~24 步等。然而, 电力市场中现货电价的间隔点可能为 5 或 15 min, 当预测未来 1 天现货电价时, 这意味着预测步长为 96 步或 288 步。目前国内外针对现货电价长步预测的研究仍处于起步阶段, 亟需有效的技术方法填补该研究的空白。最近 2 年, 自注意力机制、改进 Transformer 等方法的出现使得长步预测这一研究领域得以快速发展。文献[22]提出一种改进的 Transformer 长步预测模型, 其有效降低了模型计算的复杂度, 并提高了处理长输入序

列的能力；文献[23]提出一种基于自相关机制与分解架构的 Transformer 长步预测模型；文献[24]提出一种将模型输入序列分割为 Patch 子序列的改进 Transformer 长步预测模型。虽然上述模型能够较好地提高长步预测精度，但都是仅挖掘时间特征的方法，如果直接应用在区域现货电价的长步时空预测中，会忽略区域现货电价之间的空间相关性，从而限制了预测精度的进一步提升。此外解开时间序列中纠缠的时间模式有利于提高长步预测精度，但传统分解是对整个数据集进行分解，这将会导致未来信息的引入，在训练模型时，滑动窗口中输入数据会不可避免地获取到滑动窗口中输出数据，从而造成未来信息泄露问题。

针对上述问题，本文提出一种基于季节趋势分解-双通道 ST-Former 的区域现货电价预测框架，该预测框架旨在提高区域现货电价的长步时空预测精度。首先，利用季节趋势分解将滑动窗口中输入数据分解为稳定的趋势部分和波动的季节部分，使分解后的区域现货电价具有更加清晰的模式特征，从而增强预测模型对数据分布漂移的鲁棒性，通过对滑动窗口中输入数据进行分解的方式来解决传统分解思路存在未来信息泄露的问题；其次，提出一种基于可变 Patch 时间注意力和空间自注意力相结合的 ST-Former 模型，以挖掘各个区域现货电价的长步时空特征，该模型在挖掘区域现货电价时间特征的基础上，考虑不同区域现货电价之间的空间相关性，来解决过去长步预测模型中忽略变量之间空间特征的问题；然后，通过提出的 ST-Former 模型分别对趋势部分和季节部分进行建模构成双通道模型框架；最后，将双通道模型框架的输出结果进行聚合得到最终预测结果，从而实现区域现货电价准确的长步时空预测。以澳大利亚电力市场的区域现货电价数据为例，与 12 种预测方法进行对比，验证了所提预测框架的有效性和优越性。

1 季节趋势分解与自注意力机制

1.1 季节趋势分解

季节趋势分解是一种经典的时间序列处理方法，其可以将时间序列分解为趋势部分和季节部分等^[25]，从而使时间序列变得容易预测和可解释。经典的时间序列预测模型，通常使用分解手段对历史时间序列进行预处理，如 Prophet^[26]采用趋势季节性分解，本文采用的季节趋势分解如下^[23]：

$$\begin{cases} \mathbf{x}_T = \text{AvgPool}(\text{Padding}(\mathbf{x})) \\ \mathbf{x}_S = \mathbf{x} - \mathbf{x}_T \end{cases} \quad (1)$$

式中： $\mathbf{x}$ 为滑动窗口中输入数据； $\mathbf{x}_T$ 、 $\mathbf{x}_S$ 分别为输入的趋势部分和季节部分； $\text{AvgPool}(\cdot)$ 为移动平均操作； $\text{Padding}(\cdot)$ 表示填充操作。

1.2 自注意力机制

自注意力是 Transformer 类模型的基础，对于时间序列预测问题，自注意力的输入 $\mathbf{x}_{t-h+1:t}^i \in \mathbf{R}^{h \times D}$ 表示第 i 个时间序列中输入长度为 h 的 D 个特征值，首先将 $\mathbf{x}_{t-h+1:t}^i$ 通过线性变化转换为查询矩阵 $\mathbf{Q} = \mathbf{x}_{t-h+1:t}^i \mathbf{W}_Q$ 、密钥矩阵 $\mathbf{K} = \mathbf{x}_{t-h+1:t}^i \mathbf{W}_K$ 和值矩阵 $\mathbf{V} = \mathbf{x}_{t-h+1:t}^i \mathbf{W}_V$ 。 $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V \in \mathbf{R}^{D \times d}$ 为投影矩阵，是自注意力的可学习参数，自注意力的输出 $\mathbf{h}_{t-h+1:t}^i \in \mathbf{R}^{h \times d}$ 为值矩阵中各个值与对应的权重运算后得到的加权和，其中对应的权重为注意力分数，表示任意 2 个时刻的值的两两相关性，注意力分数由查询矩阵和密钥矩阵计算得到， $\mathbf{h}_{t-h+1:t}^i$ 计算如下：

$$\mathbf{h}_{t-h+1:t}^i = \text{att}(\mathbf{x}_{t-h+1:t}^i | \mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{att}(\mathbf{x}_{t-h+1:t}^i, \mathbf{x}_{t-h+1:t}^i) = \varphi\left(\frac{(\mathbf{x}_{t-h+1:t}^i \mathbf{W}_Q)(\mathbf{x}_{t-h+1:t}^i \mathbf{W}_K)^T}{\sqrt{d}} (\mathbf{x}_{t-h+1:t}^i \mathbf{W}_V)\right) \quad (2)$$

式中： $\text{att}(\cdot)$ 为注意力操作； φ 为 Softmax 激活函数； d 为隐藏维度。自注意力的输入中每个时刻的值与所有时刻的值进行交互得到相关性，因此能够挖掘整个输入的时间特征。

2 基于季节趋势分解-双通道 ST-Former 的区域现货电价预测框架

2.1 区域现货电价的长步时空预测

跨区电力市场交易的日渐频繁、远距离电力传输等会影响不同区域现货电价的波动，电力市场中不同电价的区域之间的电网潮流流动存在着空间相关性，这不但是一个时间演变过程，而是一个时空演变过程。区域现货电价的波动与区域空间负荷大小、区域间输电阻塞程度有着密切的关系，仅挖掘时间特征的传统预测方法会忽略区域现货电价数据中蕴含的空间特征，难以对其空间特征进行描述。对电力市场中第 i 个区域的现货电价进行预测时，不仅要考虑第 i 个区域现货电价的历史数据，也要考虑与第 i 个区域相邻的其他区域的现货电价的历史数据。因此，区域现货电价的长步时空预测可以看作是一个考虑空间相关性的多元时间序列长步预测问题。

构造区域现货电价的时空信息数据集如图 1 所示。首先，将电力市场中不同现货电价的各个区域视为节点；其次，将区域现货电价作为该节点的特

征信息；然后，将各个区域之间的输电线路拓扑结构作为节点之间的边；最后，各个区域现货电价之间的相关性作为边的特征信息。

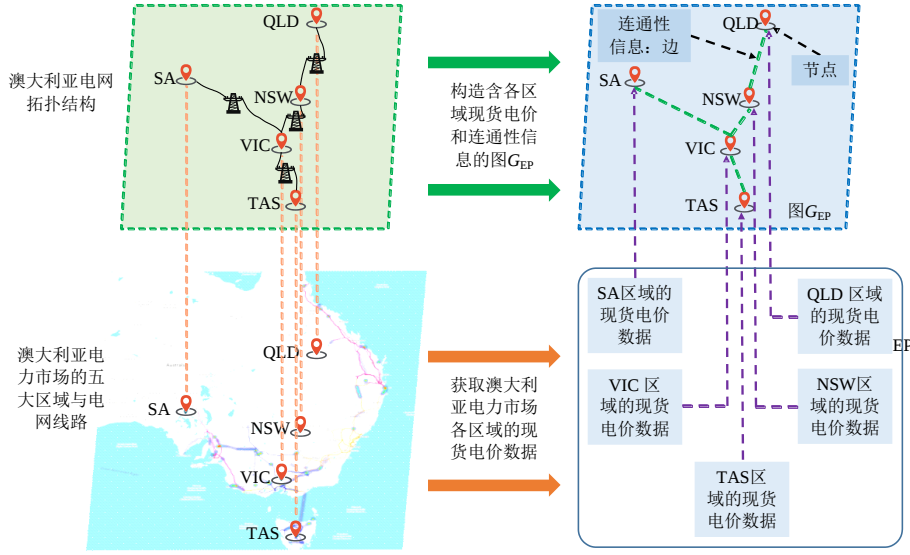


图 1 区域现货电价的时空信息数据集

Fig. 1 Spatial-temporal information dataset of regional spot electricity prices

本文使用图 $G_{EP}=(V_{EP}, E_{EP})$ 来描述区域现货电价的时空信息数据集，其中： V_{EP} 为所有节点的集合； E_{EP} 为所有边的集合； $v_i \in V_{EP}$ 为第 i 个节点及节点的特征信息； $e_{ij} \in E_{EP}$ 为节点 v_i 到节点 v_j 之间的边的特征信息。邻接矩阵 $A \in \mathbf{R}^{N \times N}$ 表示各个节点之间的边的连通性， N 为节点的总数。

根据上述定义，本文建立了区域现货电价的长步时空预测框架的函数 ξ ，通过 H 步的历史数据，来预测 F 步的未来数据，通过改变 H 和 F 的设定值来实现长步时空预测，区域现货电价的长步时空预测如下：

$$\xi_{\psi}(\mathbf{x}_{t-H+1}, \mathbf{x}_{t-H+2}, \dots, \mathbf{x}_{t-1}, \mathbf{x}_t) = (\tilde{\mathbf{x}}_{t+1}, \tilde{\mathbf{x}}_{t+2}, \dots, \tilde{\mathbf{x}}_{t+F-1}, \tilde{\mathbf{x}}_{t+F}) \quad (3)$$

式中： ψ 为预测框架的可学习参数； $\mathbf{x}_t \in \mathbf{R}^{N \times D}$ 为所有 N 个节点在 t 时刻的 D 个特征值； $\tilde{\mathbf{x}}_{t+F} \in \mathbf{R}^{N \times D}$ 为所有 N 个节点在 $t+F$ 时刻的 D 个预测特征值。

2.2 区域现货电价预测框架

2.2.1 区域现货电价的季节趋势分解

利用季节趋势分解将滑动窗口中输入数据分解为稳定的趋势部分和波动的季节部分并进行单独建模，这将有助于挖掘该时间序列的时间特征。在理论上基于循环神经网络、时间卷积网络和 Transformer 等深度学习预测方法可能不会自动解开不正交的稳定的趋势部分和波动的季节部分^[27]，因此，在深度学习预测方法中，需要考虑对输入数

据进行分解操作。但文献[11,13,16,18]中的分解方式，对整个数据集分解或对训练集、验证集、测试集单独分解的操作均会引入未来信息，在训练模型时，滑动窗口中输入数据会不可避免地获取到滑动窗口中输出数据，从而获得较好的提升效果。对数据集做分解时，应当保证同一滑动窗口中输入数据不能接触到输出数据，否则会造成未来信息泄露的问题，不当分解造成未来信息泄露如附图 A1 所示。

针对上述问题，本文对预测框架的滑动窗口中输入数据进行分解操作，从而解决传统分解存在未来信息泄露的问题。将滑动窗口中输入数据通过季节趋势分解得到稳定的趋势部分和波动的季节部分，来分别反映区域现货电价的稳定特征和波动特征。对于滑动窗口中长度为 H 的输入数据 $\mathbf{x}_{t-H+1:t} \in \mathbf{R}^{N \times H \times D}$ ，分解过程如下：

$$\mathbf{x}_{t-H+1:t,T} = \text{AvgPool}(\text{Padding}(\mathbf{x}_{t-H+1:t})) \quad (4)$$

$$\mathbf{x}_{t-H+1:t,S} = \mathbf{x}_{t-H+1:t} - \mathbf{x}_{t-H+1:t,T} \quad (5)$$

式中 $\mathbf{x}_{t-H+1:t,T}$ 、 $\mathbf{x}_{t-H+1:t,S}$ 分别为 $t-H+1 \sim t$ 时间段内滑动窗口中输入数据的趋势部分和季节部分。通过移动平均操作来提取滑动窗口中输入数据的趋势部分，其中包含填充操作，以保证分解操作后的趋势部分和季节部分长度不变。

2.2.2 ST-Former 模型

空间负荷分布不均、区域间输电阻塞等原因，导致区域现货电价具有明显的时空特性，如果直接

采用仅挖掘时间特征的长步预测模型会忽略区域现货电价蕴含的空间特征,从而限制了预测精度的进一步提升。

针对上述问题,本文提出一种基于可变 Patch 时间注意力和空间自注意力相结合的 ST-Former 模型,用于挖掘区域现货电价的长步时空特征。该模型在挖掘区域现货电价时间特征的基础上,考虑了不同区域现货电价之间的空间相关性,来解决过去长步预测模型中忽略变量之间空间特征的问题。ST-Former 模型结构如附图 A2 所示。该模型的设计主要考虑了 2 方面:首先,设计了一种可变 Patch 时间注意力结构,来挖掘 Patch 中输入值的时间特征,并且可以根据输入长度的不同来改变网络结构,例如,当输入较短时,变为较小的 Patch,来全面挖掘输入中每个时间步的时间特征,以提高预测精度;当输入较长时,变为较大的 Patch,使模型中每层的输入长度呈指数级缩小,以解决在堆叠多层注意力时因计算效率无法引入较长输入的问题。其次,设计了一种空间自注意力结构,与可变 Patch 时间注意力结构相互叠加,以挖掘区域现货电价之间的相关性。通过这 2 方面的设计,使 ST-Former 模型实现了对区域现货电价的长步时空特征挖掘。

ST-Former 模型中每层注意力的所有 N 个变量的输入长度 H 相同,将第 l 层的 H_l 个输入沿时间维度划分为 $P_l=H_l/S_l$ 个 Patch,式中: S_l 为第 l 层的 Patch 大小(H_l 、 S_l 是可调节的参数), l 为模型中堆叠的注意力层数, $1 \leq l \leq L$, $H_{l+1}=P_l$ 。使用 $\mathbf{x}_p^{i,l} = (\mathbf{x}_{(p-1) \times S_l + 1}^{i,l}, \dots, \mathbf{x}_{p \times S_l}^{i,l})$ 为第 l 层的第 i 个变量的第 p 个 Patch 的输入值,式中, $\mathbf{x}_{(p-1) \times S_l + 1}^{i,l} \in \mathbf{R}^D$ 为 $\mathbf{x}_p^{i,l}$ 中第 $(p-1) \times S_l + 1$ 个输入, $1 \leq i \leq N$, $1 \leq p \leq P_l$ 。 $\mathbf{x}_{t-H+1:t,\varepsilon} \in \mathbf{R}^{N \times F \times D}$ 为第 ε 个通道的 ST-Former 模型的初始输入。

可变 Patch 时间注意力结构灵感来自文献[24],为了增加模型输入序列的长度,对于每个 Patch 时间注意力,本文引入了一个可学习的伪时间戳 $\mathbf{T}_p^{i,l} \in \mathbf{R}^{1 \times d}$, $\mathbf{T}_p^{i,l}$ 为第 l 层的第 i 个变量的第 p 个伪时间戳,其中伪时间戳作为可变 Patch 时间注意力中的查询矩阵来查询 Patch 中的所有输入值,因此只需要计算所有输入值的单个注意力得分,从而降低了计算的复杂度,实现模型输入长度的增加。可变 Patch 时间注意力输出如下:

$$\rho(\mathbf{T}_p^{i,l}) = \text{patt}(\mathbf{T}_p^{i,l}, \mathbf{x}_p^{i,l}) = \phi \left[\frac{\mathbf{T}_p^{i,l} (\mathbf{x}_p^{i,l} \mathbf{W}_K^{i,l})^T}{\sqrt{d}} \right] (\mathbf{x}_p^{i,l} \mathbf{W}_V^{i,l}) \quad (6)$$

式中: $\rho(\mathbf{T}_p^{i,l}) \in \mathbf{R}^{1 \times d}$ 为第 l 层的第 i 个变量的第 p 个可变 Patch 时间注意力输出; $\text{patt}(\cdot)$ 为 Patch 时间注意力操作; $\mathbf{x}_p^{i,l} \in \mathbf{R}^{S_l \times D}$; $\mathbf{W}_K^{i,l} \in \mathbf{R}^{D \times d}$ 及 $\mathbf{W}_V^{i,l} \in \mathbf{R}^{D \times d}$ 为第 l 层的第 i 个变量的可变 Patch 时间注意力的可学习的时间投影矩阵。

为了挖掘区域现货电价之间的相关性,接下来,将通过空间自注意力结构来挖掘每 1 层的所有 N 个变量的第 p 个 Patch 之间的相关性,并经过邻接矩阵 $\mathbf{A}$ 剪枝空间自注意力的冗余相关性后,得到各个区域现货电价之间的相关性,空间自注意力输出如下:

$$\zeta(\rho[\mathbf{T}_p^l]) = \text{satt}(\rho[\mathbf{T}_p^l], \rho[\mathbf{T}_p^l]) = \phi \left(\frac{(\rho[\mathbf{T}_p^l] \mathbf{U}_Q^{p,l})(\rho[\mathbf{T}_p^l] \mathbf{U}_K^{p,l})^T}{\sqrt{d}} \odot \mathbf{A} \right) (\rho[\mathbf{T}_p^l] \mathbf{U}_V^{p,l}) \quad (7)$$

式中: $\zeta(\rho[\mathbf{T}_p^l]) \in \mathbf{R}^{N \times d}$ 为第 l 层的拓扑结构变量的第 p 个 Patch 之间的空间自注意力输出; $\text{satt}(\cdot)$ 表示空间注意力操作; $\rho[\mathbf{T}_p^l] = [\rho(\mathbf{T}_p^{1,l}), \dots, \rho(\mathbf{T}_p^{N,l})] \in \mathbf{R}^{N \times d}$ 为第 l 层的所有 N 个变量的第 p 个可变 Patch 时间注意力输出; $\mathbf{T}_p^l = [\mathbf{T}_p^{1,l}, \dots, \mathbf{T}_p^{N,l}] \in \mathbf{R}^{N \times d}$ 为第 l 层的所有 N 个变量的第 p 个伪时间戳; $\mathbf{U}_Q^{p,l} \in \mathbf{R}^{d \times d}$; $\mathbf{U}_K^{p,l} \in \mathbf{R}^{d \times d}$ 及 $\mathbf{U}_V^{p,l} \in \mathbf{R}^{d \times d}$ 为第 l 层的第 p 个空间自注意力的可学习的空间投影矩阵。

为了提高可变 Patch 时间注意力和空间自注意力的效果,本文将多层注意力叠加在一起形成多层时空注意力,与只使用最后一层的汇总输出 $\mathbf{O}_L$ 相比,采用所有层的汇总输出 $\mathbf{O}$ 会有 2 个好处:首先是所有层的汇总输出 $\mathbf{O}$ 代表了不同的时间尺度,从而形成了不同尺度的时间特征;其次是提供了多个梯度反馈路径,从而简化了学习过程。第 l 层的汇总输出 $\mathbf{O}_l$ 是通过连接每一个 $\zeta(\rho[\mathbf{T}_p^l])$ 导出的,具体如下:

$$\mathbf{O}_l = [\zeta(\rho[\mathbf{T}_1^l]), \zeta(\rho[\mathbf{T}_2^l]), \dots, \zeta(\rho[\mathbf{T}_{P_l}^l])] \quad (8)$$

式中 $\mathbf{O}_l \in \mathbf{R}^{N \times P_l \times d}$ 。

由于输入的维数在每 1 层都呈指数下降,因此本文采用了含有 1 层神经网络的跳跃连接来确保每 1 层之间的输出形状一致,所有层的汇总输出 $\mathbf{O}$ 如下:

$$\mathbf{O} = \sum_{l=1}^L \mathbf{W}_l \mathbf{O}_l \quad (9)$$

式中： $W_l \in \mathbf{R}^{d \times P_l}$ 为第 l 层跳跃连接的可学习参数； $O \in \mathbf{R}^{N \times d \times d}$ 。

然后将所有层的汇总输出 O 连接到预测器。最后，本文采用了含有 2 层神经网络的预测器，该预测器将隐藏表示的 O 映射到预测输出 $\tilde{x}_{t+1:t+F,\varepsilon}$ ：

$$\tilde{x}_{t+1:t+F,\varepsilon} = (\text{ReLU}(W_1 O)) W_2 \quad (10)$$

式中： $\tilde{x}_{t+1:t+F,\varepsilon} \in \mathbf{R}^{N \times F \times D}$ 为第 ε 个通道的 ST-Former 模型的预测输出；ReLU 为整流线性单元激活函数； $W_1 \in \mathbf{R}^{F \times d}$ 和 $W_2 \in \mathbf{R}^{d \times D}$ 为可学习参数。

2.2.3 基于季节趋势分解-双通道 ST-Former 的预测框架及流程步骤

本文提出的基于季节趋势分解-双通道 ST-Former 的预测框架如附图 A3 所示，该预测框架主要由滑动窗口中输入数据的时间趋势分解、ST-Former 模型、双通道模型框架 3 部分组成。

本文提出的预测框架的流程如附图 A4 所示，对于训练过程，首先，通过滑动窗口提取到训练集中 $t-H+1 \sim t$ 时间段内区域现货电价数据，将该数据通过季节趋势分解处理后，得到 $t-H+1 \sim t$ 时间段内区域现货电价的趋势部分 $x_{t-H+1:t,T}$ 和季节部分 $x_{t-H+1:t,S}$ ；然后，将 $x_{t-H+1:t,T}$ 和 $x_{t-H+1:t,S}$ 分别输入到 2 个 ST-Former 模型中进行建模构成双通道模型框架；最后，双通道 ST-Former 模型分别输出 $t+1 \sim t+F$ 时间段内区域现货电价的趋势部分的预测值 $\tilde{x}_{t+1:t+F,T}$ 和季节部分的预测值 $\tilde{x}_{t+1:t+F,S}$ ，将 $\tilde{x}_{t+1:t+F,T}$ 和 $\tilde{x}_{t+1:t+F,S}$ 进行聚合得到区域现货电价的预测值 $\tilde{x}_{t+1:t+F}$ ，将 $\tilde{x}_{t+1:t+F}$ 和区域现货电价的真实值 $x_{t+1:t+F}$ 输入到损失函数中，并以损失函数最小化为目标进行训练。之后，通过滑动窗口沿时间维度移动，重复上述步骤，完成预测框架的训练，再经过验证集对预测框架进行调整优化。

对于预测过程，首先，通过滑动窗口提取到测试集中 $t-H+1 \sim t$ 时间段内区域现货电价数据，将该数据通过季节趋势分解处理后，得到 $t-H+1 \sim t$ 时间段内区域现货电价的趋势部分 $x_{t-H+1:t,T}$ 和季节部分 $x_{t-H+1:t,S}$ ；然后，将 $x_{t-H+1:t,T}$ 和 $x_{t-H+1:t,S}$ 分别输入到 2 个训练完成的 ST-Former 模型中；最后，双通道 ST-Former 模型分别输出 $t+1 \sim t+F$ 时间段内区域现货电价的趋势部分的预测值 $\tilde{x}_{t+1:t+F,T}$ 和季节部分的预测值 $\tilde{x}_{t+1:t+F,S}$ ，将 $\tilde{x}_{t+1:t+F,T}$ 和 $\tilde{x}_{t+1:t+F,S}$ 进行聚合得到区域现货电价的预测值 $\tilde{x}_{t+1:t+F}$ ，将 $\tilde{x}_{t+1:t+F}$ 和区域现货电价的真实值 $x_{t+1:t+F}$ 输入到评价指标中计算。

之后，通过滑动窗口沿时间维度移动，重复上述步骤，完成预测框架的预测。

2.3 预测模型的损失函数

本文以平均绝对误差为训练目标，损失函数如下：

$$L(\{\tilde{x}_{t+1:t+F}, x_{t+1:t+F}\}; \Phi) = \frac{1}{NFD} \sum_{i=1}^N \sum_{j=1}^F \sum_{k=1}^D |x_{t+j,k}^i - \tilde{x}_{t+j,k}^i| \quad (11)$$

式中： Φ 为模型参数； $x_{t+j,k}^i$ 为第 i 个区域现货电价在 $t+j$ 时刻的第 k 个特征值； $\tilde{x}_{t+j,k}^i$ 为第 i 个区域现货电价在 $t+j$ 时刻的第 k 个预测特征值。

3 算例分析

3.1 数据集分析与处理

为了对比不同预测模型对区域现货电价的预测效果，本文使用现货价格波动性很强的澳大利亚电力市场数据，使用 2022 年 1 月 1 日—2022 年 4 月 30 日之间，4 个月的区域现货电价为数据集，包含新南威尔士州(New South Wales, NSW)、昆士兰州(Queensland, QLD)、南澳大利亚州(South Australia, SA)、塔斯马尼亚(Tasmania, TAS)、维多利亚州(Victoria, VIC)这 5 个区域，每个区域的数据包时间间隔为 5 min 的现货电价，总共有 2 880 h，34 559 个数据点。该数据集存在非常多的尖峰电价，在澳大利亚电力市场中每兆瓦时平均现货价格在几十澳元到几百澳元之间，但出现尖峰价格时可以达到上千澳元。

澳大利亚电力市场覆盖了其东部、南部范围广阔的几个州，绵延几千 km。各区域电网之间联络线及区域电网内部分线路存在输电功率限制的情况，其输电线路拓扑结构如图 1 所示，QLD 区域与 NSW 区域相连接，NSW 区域与 VIC 区域相连接，VIC 区域分别与 SA 和 TAS 区域相连接。澳大利亚电力市场中 SA 和 VIC 区域中 1 天 288 个点的现货电价情况如附图 B1 所示，SA 区域的现货电价相比于 VIC 区域的现货电价波动变化大，这反映了空间因素对区域现货电价的影响，使得不同区域的现货电价存在着明显差异。在黑线左侧部分，SA 和 VIC 区域的现货电价非常相似，但在黑线右侧部分，SA 和 VIC 区域的现货电价非常不同，如果采用 ARIMA、LSTM 等仅挖掘时间特征的预测模型，通过历史区域现货电价(黑线左侧部分)预测未来区域现货电价(黑线右侧部分)，则很难对这 2 个来自同一时间但不同空间的区域现货电价产生不同的预

测结果,为此,提出区域现货电价在空间上的不可区分性这一概念,用于描述传统时序模型在区域现货电价时空预测中的局限性。

将区域现货电价数据集按 7:1:2 分为训练集、验证集和测试集。采用 Z 分数法对上述 3 个数据集进行归一化:

$$\hat{x} = \frac{x - \text{mean}(\mathbf{x}_{\text{train}})}{\text{std}(\mathbf{x}_{\text{train}})} \quad (12)$$

式中: x 为数据集中的真实值; $\hat{x}$ 为 x 的归一化值; $\text{mean}(\mathbf{x}_{\text{train}})$ 为训练集的平均值; $\text{std}(\mathbf{x}_{\text{train}})$ 为训练集的标准差。

3.2 模型参数与评价指标

本文预测框架的参数设置:采用历史 288 步区域现货电价来预测未来 288 步区域现货电价,因此, $H=288$, $F=288$ 。区域现货电价数据集只包含电价数据,因此, $D=1$ (考虑电价单一特征,未加入其他特征)。澳大利亚电力市场中具有 5 个区域现货电价,因此, $N=5$ 。模型中堆叠了 5 层时空注意力,每 1 层 Patch 的大小依次为 6、4、3、2、2,因此, $S, L=6, 4, 3, 2, 2$ 。模型其他参数设置:隐藏维度 $d=128$, 移动平均 $m_a=25$, 丢弃率 $d_0=0.3$, 多头数量 $h=8$, 学习率 $l_r=0.001$, 批量大小 $b_s=64$, 迭代次数 $e=100$, 优化器为 Adam。

本文通过区域现货电价预测的平均结果来评价不同预测模型的性能,该结果是重复了 5 次实验的平均值。本文采用平均绝对误差 I_{MAE} 、均方根误差 I_{RMSE} 作为评价指标:

$$I_{\text{MAE}} = \frac{1}{NF_T} \sum_{i=1}^N \sum_{j=1}^{F_T} |x_{t+j}^i - \tilde{x}_{t+j}^i| \quad (13)$$

$$I_{\text{RMSE}} = \sqrt{\frac{1}{NF_T} \sum_{i=1}^N \sum_{j=1}^{F_T} (x_{t+j}^i - \tilde{x}_{t+j}^i)^2} \quad (14)$$

式中: x_{t+j}^i 为第 i 个区域现货电价在 $t+j$ 时刻的值; $\tilde{x}_{t+j}^i$ 为第 i 个区域现货电价在 $t+j$ 时刻的预测值; F_T 为测试集中的总用例数。

3.3 算例结果与分析

本文比较了所提预测框架的 3 种不同形式: T-Former, 由多层可变 Patch 时间注意力构成的预测模型,可以挖掘区域现货电价的长步时间特征; ST-Former, 由多层可变 Patch 时间注意力与空间自注意力相互叠加构成的预测模型,可以挖掘区域现货电价的长步时空特征; STD-ST-Former, 对滑动窗口进行季节趋势分解的基础上,通过 ST-Former

模型分别对趋势部分和季节部分进行建模构成的双通道模型框架,该预测框架将使区域现货电价变得更容易预测和可解释。上述 3 种不同形式的模型中其他模块、结构均相同。此外,本文还与 ARIMA、SVR、LSTM、STGCN、GWN、Informer、Autoformer、Dlinear、MLP 和 STID 这 10 种预测模型进行对比,10 种预测模型的超参数设置详见附录 B。

通过上述 13 种预测模型对澳大利亚 5 个区域的现货电价进行训练、预测,13 种模型的 I_{MAE} 指标对比如图 2 所示,结果表明: STD-ST-Former 的 I_{MAE} 在 288 步预测中均优于 ST-Former 的 I_{MAE} , 平均降低了 1.67 澳元,由此可知 STD-ST-Former 中的基于季节趋势分解的双通道建模预测框架具有重要作用,该预测框架可以缓解区域现货电价数据存在分布漂移的问题,使分解后的区域现货电价具有更加清晰的模式特征,从而增强预测模型对数据分布漂移的鲁棒性,当波动的季节部分发生分布漂移时,稳定的趋势部分仍然可以准确预测,从而保证预测性能; ST-Former 的 I_{MAE} 在 288 步预测中均优于 T-Former 的 I_{MAE} , 平均降低了 2.32 澳元,由此可知 ST-Former 中的空间自注意力可以在可变 Patch 时间注意力挖掘长步时间特征的基础上,有效挖掘区域现货电价之间的空间相关性,从而提升区域现货电价长步时空预测的精度;在 288 步预测中,13 种预测模型的平均 I_{MAE} 由低到高依次为 STD-ST-Former、ST-Former、T-Former、SVR、Autoformer、Dlinear、Informer、ARIMA、MLP、STID、STGCN、LSTM、GWN, STD-ST-Former 的 I_{MAE} 与其他 10 种预测模型的 I_{MAE} 相比,始终具有明显优势,附表 B1 展示了 13 种预测模型的 I_{MAE} 评价结果。

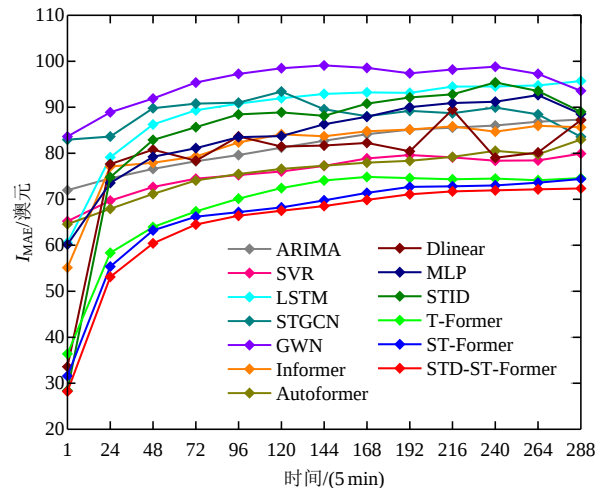


图2 13种预测模型的 I_{MAE} 对比

Fig. 2 I_{MAE} comparison of thirteen predictive models

13 种预测模型的 I_{RMSE} 指标对比如图 3 所示, 结果表明: 与 I_{MAE} 指标的趋势相同, STD-ST-Former 的 I_{RMSE} 在 288 步预测中均优于 ST-Former 的 I_{RMSE} , 平均降低了 0.80 澳元, ST-Former 的 I_{RMSE} 在 288 步预测中均优于 T-Former 的 I_{RMSE} , 平均降低了 3.36 澳元; 在 288 步预测中, 13 种预测模型的平均 I_{RMSE} 由低到高依次为 STD-ST-Former、ST-Former、T-Former、SVR、Autoformer、Informer、ARIMA、STID、Dlinear、STGCN、LSTM、MLP、GWN, STD-ST-Former 的 I_{RMSE} 与其他 10 种预测模型的 I_{RMSE} 相比, 始终具有明显优势。附表 B2 展示了 13 种预测模型的 I_{RMSE} 评价结果。

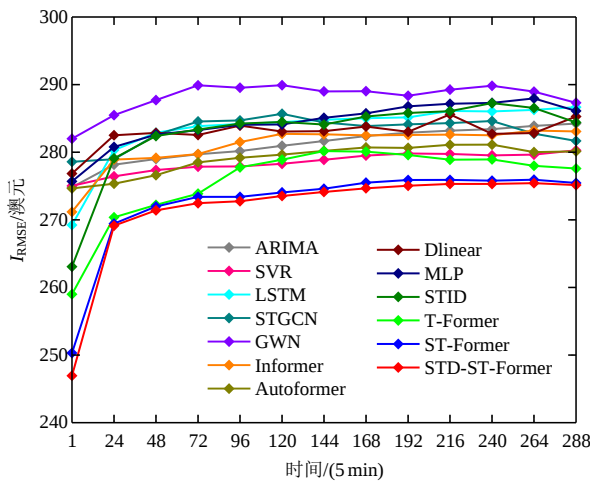


图 3 13 种预测模型的 I_{RMSE} 对比

Fig. 3 I_{RMSE} comparison of thirteen predictive models

图 2、3 的结果表明: 本文提出的 STD-ST-Former 预测框架在区域现货电价的长步时空预测中具有优越的预测性能, 以澳大利亚电力市场为算例, 在 13 种预测模型中达到最优评价结果; 此外, 结果还表明 13 种预测模型在 24 步之后的 I_{MAE} 和 I_{RMSE} 的趋势变化逐渐稳定, 能够清楚地反映 13 种预测模型之间长步时空预测性能的差距。其中, GWN、STGCN、LSTM 等方法预测性能欠佳的主要原因是基于循环神经网络和时间卷积网络的预测模型很难挖掘远距离的时间特征, 当直接应用于长步时空预测任务时, 不仅长步时空预测性能不佳, 而且还会导致短步时空预测性能的显著下降。

除了上文 3 种形式的对比, 本文还分析了预测框架中各部分设计的作用, 附表 B3 和 B4 展示了 STD-ST-Former 对伪时间戳 $T_p^{i,j}$ 和邻接矩阵 A 的消融实验结果, 其中, STD-ST*-Former 表示 STD-ST-Former 没有伪时间戳 $T_p^{i,j}$ 部分, STD-S*T-Former 表示 STD-ST-Former 没有邻接矩阵 A 部分。结果表

明: 增加伪时间戳 $T_p^{i,j}$ 部分可以提高模型的预测性能, 伪时间戳 $T_p^{i,j}$ 通过降低计算复杂度, 使得模型能够挖掘更长的时间特征, 从而达到更好的预测效果; 此外, 增加邻接矩阵 A 部分也可以提高模型的预测性能, 由输电线路拓扑结构构成的邻接矩阵 A 通过剪枝空间自注意力中冗余的相关性, 可以更加准确地描述各个区域现货电价之间的相关性, 从而达到更好的预测效果。

本文还分析了预测框架中重要超参数的敏感性, Patch 大小对 STD-ST-Former 的 I_{MAE} 和 I_{RMSE} 的影响分别如附图 B2 和 B3 所示, 结果表明: 在不同的设置下对第 1 步的 I_{MAE} 有较大的影响, 对其他步的 I_{MAE} 和 I_{RMSE} 有略微的影响, 从平均结果看, (6,4,3,2,2) 为最优的 Patch 大小。注意力头数大小对 STD-ST-Former 的 I_{MAE} 和 I_{RMSE} 的影响分别如附图 B4 和 B5 所示, 结果表明: 与 Patch 大小的影响类似, 在不同的设置下对第 1 步的 I_{RMSE} 有较大的影响, 对其他步的 I_{MAE} 和 I_{RMSE} 有略微的影响; 此外, 虽然注意力头数设置为 2 时, 第 1 步的 I_{RMSE} 明显降低, 但之后步长的 I_{RMSE} 全部高于其他设置下的 I_{RMSE} , 从平均结果看, 8 为最优的注意力头数大小。隐藏维度大小对 STD-ST-Former 的 I_{MAE} 和 I_{RMSE} 的影响如附图 B6 和 B7 所示, 结果表明: 隐藏维度设置为 64 时, 各个步长的 I_{MAE} 明显升高, 在其他设置时对 I_{MAE} 有略微的影响; 此外, 在不同的设置下对 I_{RMSE} 有略微的影响, 从平均结果看, 128 为最优的隐藏维度大小。丢弃率大小对 STD-ST-Former 的 I_{MAE} 和 I_{RMSE} 的影响如附图 B8 和 B9 所示, 结果表明: 丢弃率设置为 0.1 和 0.2 时, 在第 48 步之后的 I_{MAE} 明显升高, 在其他设置时对 I_{MAE} 有略微的影响; 此外, 在不同的设置下对 I_{RMSE} 有略微的影响, 从平均结果看, 0.3 为最优的丢弃率大小。总体上看, Patch 大小和注意力头数大小对 I_{RMSE} 的影响较大, 隐藏维度大小和丢弃率大小对 I_{MAE} 的影响较大。

澳大利亚电力市场中 NSW 的现货电价预测如图 4 所示, 展示了从测试集中提取的 288 个输入数据(历史区域现货电价)来预测 288 个输出数据(未来区域现货电价)。如图 4 可知: STD-ST-Former 预测效果最好, 可以很好地预测未来区域现货电价的趋势变化, 其次预测效果较好的是 ST-Former 和 T-Former, 也可以预测未来区域现货电价的趋势变

化,但 ST-Former 和 T-Former 预测效果均差于 STD-ST-Former,结果表明,解开区域现货电价中纠缠的时间模式和挖掘区域现货电价之间的空间相关性均有利于提高区域现货电价的预测精度;SVR 和 ARIMA 在预测评价结果中尚可,但在实际预测中表现欠佳,不能预测出区域现货电价的趋势变化;Autoformer 和 Informer 在预测评价结果和实际预测中表现优于 LSTM、STGCN 和 GWN,这也验证了 Transformer 类预测模型在长步时空预测中更具优势,优于以往基于循环神经网络和时间卷积网络的预测模型;LSTM、STGCN 和 GWN 在预测评价结果和实际预测中表现不佳,能够预测区域现货电价

的部分趋势变化,但预测值偏离真实值较多;Dlinear、MLP 和 STID 在预测评价结果和实际预测中表现优于 LSTM、STGCN 和 GWN,这是因为澳大利亚电力市场中区域现货电价数据的间隔为 5 min,具有随机性大、波动性强等特性,Dlinear、MLP 和 STID 属于线性预测模型,线性模型结构简单,其归纳偏差相对较弱,相较于 LSTM、STGCN 和 GWN 不容易出现过拟合现象,因此,线性模型在分布漂移明显和模式不明确区域现货电价长步时空预测中优于大多数基于循环神经网络和时间卷积网络的预测模型。QLD、SA、TAS 和 VIC 的结果类似,详见附图 B10—B13。

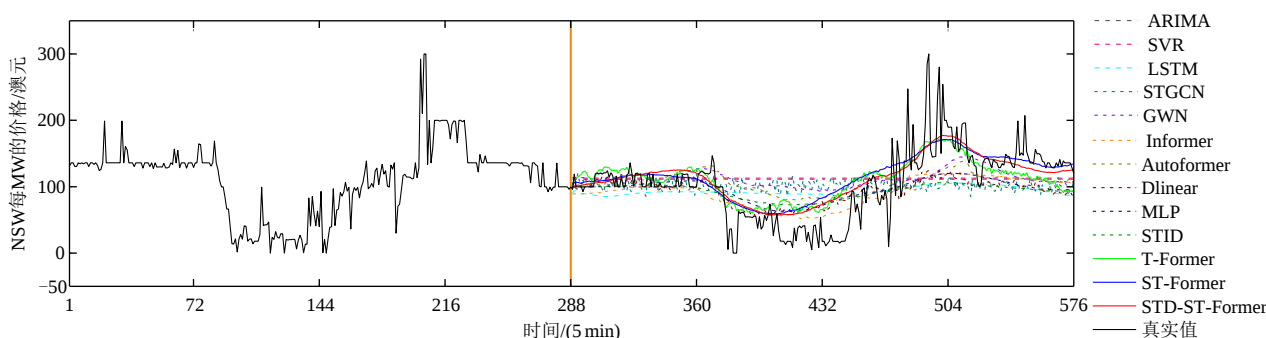


图4 13种预测模型在NSW的现货电价预测对比

Fig. 4 Comparison of thirteen prediction models for spot electricity price prediction in NSW

图5展示了STD-ST-Former在NSW的现货电价预测结果的散点图,结果表明:STD-ST-Former能够对正常区域现货电价进行准确的长步时空预测,但对极端区域现货电价的长步时空预测效果不佳,例如负电价和尖峰电价等。可见,STD-ST-Former能够有效地挖掘到区域现货电价的全局时空特征,但并未很好地挖掘到区域现货电价的局部

时空特征,因而对极端区域现货电价的长步时空预测效果不佳。QLD、SA、TAS和VIC的结果类似,详见附图B14—B17。

为了进一步呈现13种预测模型的性能,附图B18展示了13种预测模型在5个区域的预测误差分布,结果表明:STD-ST-Former具有最好的预测性能,预测误差主要集中在-40~40澳元之间,其次是ST-Former和T-Former。进一步分析附图B18可发现,除了本文所提预测框架的3种形式之外,其余10种预测模型的预测误差分布都存在向左侧倾斜的情况,这意味着10种预测模型倾向于低估实际电价,这也表明本文所提预测框架的优越性能。

4 结论

针对电力市场中区域现货电价长步时空预测精度低问题,本文提出一种基于季节趋势分解-双通道ST-Former的预测框架,并以澳大利亚电力市场的区域现货电价数据为算例,验证了所提预测框架的有效性和优越性。主要结论如下:

1) 本文提出的基于可变Patch时间注意力和空间自注意力相结合的ST-Former模型,用于挖掘区

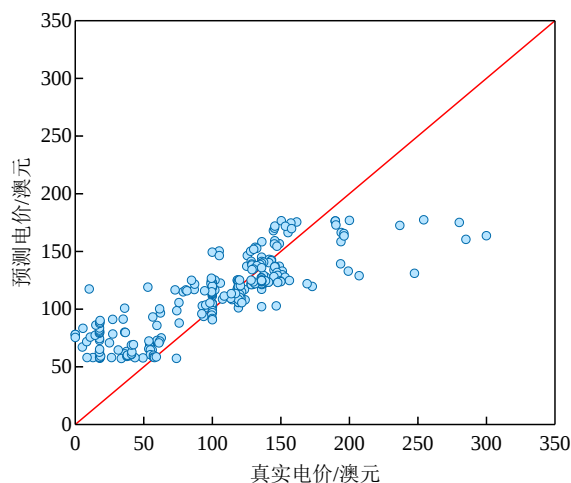


图5 NSW现货电价预测结果的散点图

Fig. 5 Scatter plot of spot electricity price prediction results in NSW

域现货电价的长步时空特征, 在可变 Patch 时间注意力挖掘区域现货电价时间特征的基础上, 考虑了其他区域现货电价的影响, 采用空间自注意力挖掘区域现货电价之间的空间相关性, 从而缓解区域现货电价在空间上的不可区分性, 算例结果表明挖掘区域现货电价之间的空间相关性可以有助于提高长步时空预测精度。

2) 本文提出的基于季节趋势分解-双通道 ST-Former 的预测框架, 用于提高区域现货电价的可预测性和可解释性, 利用季节趋势分解将滑动窗口中输入数据分解为稳定的趋势部分和波动的季节部分, 从而避免未来信息泄露的问题; 然后通过 ST-Former 模型分别对趋势部分和季节部分进行建模构成双通道模型框架。算例结果表明, 解开区域现货电价中纠缠的时间模式可以有助于提高长步时空预测精度。

3) 本文提出的区域现货电价预测框架以澳大利亚电力市场的电价数据为算例, 与 12 种预测方法进行对比, 算例结果表明, 本文所提的预测框架在 288 步预测中预测评价结果和实际预测效果均为最优。

STD-ST-Former 在区域现货电价长步时空预测方面拥有广阔的潜力, 它可以挖掘不同区域现货电价之间经由电气连接的空间相关性, 以及通过对滑动窗口中输入数据的季节趋势分解, 从而使区域现货电价变得更容易预测和可解释。虽然 STD-ST-Former 可以有效地预测未来区域现货电价的趋势变化, 但区域现货电价中波动的局部时空特征并未很好地挖掘到, 后续工作将考虑挖掘区域现货电价中波动的局部时空特征, 来进一步提升长步时空预测的精度。

参考文献

- [1] 尹积军, 吴文传. 浙江构建新型电力系统的技术路径与实践[J]. 中国电机工程学报, 2023, 43(14): 5404-5415.
YIN Jijun, WU Wenchuan. Technical path and practice for constructing the new type power system in Zhejiang Province[J]. Proceedings of the CSEE, 2023, 43(14): 5404-5415(in Chinese).
- [2] 张粒子, 陈皓轩, 黄弦超, 等. 欧洲电力平衡市场机制设计逻辑及对我国碳中和目标下辅助服务市场发展规划的启示[J]. 中国电机工程学报, 2023, 43(S1): 14-30.
ZHANG Lizi, CHEN Haoxuan, HUANG Xianchao, et al. The logic of the design of the European electricity equilibrium market and its implications for the development planning of China's power ancillary services market under the carbon-neutral target[J]. Proceedings of the CSEE, 2023, 43(S1): 14-30(in Chinese).
- [3] 王进, 张粒子, 丛野, 等. 应用图论建模输电网的电力现货市场出清模型[J]. 中国电机工程学报, 2024, 44(9): 3429-3440.
WANG Jin, ZHANG Lizi, CONG Ye, et al. A clearing model of electricity spot market based on graph theory to modeling transmission network [J]. Proceedings of the CSEE, 2024, 44(9): 3429-3440(in Chinese).
- [4] 潘虹锦, 高红均, 杨艳红, 等. 基于主从博弈的售电商多元零售套餐设计与多级市场购电策略[J]. 中国电机工程学报, 2022, 42(13): 4785-4799.
PAN Hongjin, GAO Hongjun, YANG Yanhong, et al. Multi-type retail packages design and multi-level market power purchase strategy for electricity retailers based on master-slave game[J]. Proceedings of the CSEE, 2022, 42(13): 4785-4799(in Chinese).
- [5] 王小君, 窦嘉铭, 刘翌, 等. 可解释人工智能在电力系统中的应用综述与展望[J]. 电力系统自动化, 2024, 48(4): 169-191.
WANG Xiaojun, DOU Jiaming, LIU Zhao, et al. Review and prospect of explainable artificial intelligence and its application in power systems[J]. Automation of Electric Power Systems, 2024, 48(4): 169-191(in Chinese).
- [6] 艾欣, 周志宇, 魏妍萍, 等. 基于自回归积分滑动平均模型的可转移负荷竞价策略[J]. 电力系统自动化, 2017, 41(20): 26-31, 104.
AI Xin, ZHOU Zhiyu, WEI yanping, et al. Bidding strategy for time-shiftable loads based on autoregressive integrated moving average model[J]. Automation of Electric Power Systems, 2017, 41(20): 26-31, 104(in Chinese).
- [7] GONZÁLEZ J P, ROQUE A M S M S, PÉREZ E A. Forecasting functional time series with a new hilbertian ARMAX model: application to electricity price forecasting[J]. IEEE Transactions on Power Systems, 2018, 33(1): 545-556.
- [8] 魏勤, 陈仕军, 黄炜斌, 等. 利用随机森林回归的现货市场出清价格预测方法[J]. 中国电机工程学报, 2021, 41(4): 1360-1367.
WEI Qin, CHENG Shijun, HUANG Weibin, et al. Forecasting method of clearing price in spot market by random forest regression[J]. Proceedings of the CSEE, 2021, 41(4): 1360-1367(in Chinese).
- [9] XIAO Chixin, SUTANTO D, MUTTAQI K, et al. Online

- sequential extreme learning machine algorithm for better predispatch electricity price forecasting grids[J]. IEEE Transactions on Industry Applications, 2021, 57(2): 1860-1871.
- [10] TSCHORA L, PIERRE E, PLANTEVIT M, et al. Electricity price forecasting on the day-ahead market using machine learning[J]. Applied Energy, 2022, 313: 118752.
- [11] WANG Deyun, LUO Hongyuan, GRUNDER O, et al. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm [J]. Applied Energy, 2017, 190: 390-407.
- [12] 赵雅雪, 王旭, 蒋传文, 等. 基于最大信息系数相关性分析和改进多层级门控 LSTM 的短期电价预测方法[J]. 中国电机工程学报, 2021, 41(1): 135-146.
- ZHAO Yaxue, WANG Xu, JIANG Chuanwen, et al. A novel short-term electricity price forecasting method based on correlation analysis with the maximal information coefficient and modified multi-hierarchy gated LSTM[J]. Proceedings of the CSEE, 2021, 41(1): 135-146(in Chinese).
- [13] 殷豪, 丁伟锋, 陈顺, 等. 基于长短时记忆网络-纵横交叉算法的含高比例新能源电力市场日前电价预测[J]. 电网技术, 2022, 46(2): 472-480.
- YIN Hao, DING Weifeng, CHEN shun, et al. Day-ahead electricity price forecasting of electricity market with high proportion of new energy based on LSTM-CSO model [J]. Power System Technology, 2022, 46(2): 472-480(in Chinese).
- [14] DENG Zhoufu, LIU Chenxu, ZHU Zhiliang. Inter-hours rolling scheduling of behind-the-meter storage operating systems using electricity price forecasting based on deep convolutional neural network[J]. International Journal of Electrical Power & Energy Systems, 2021, 125: 106499.
- [15] TOUBEAU J F, MORSTYN T, BOTTIEAU J, et al. Capturing spatio-temporal dependencies in the probabilistic forecasting of distribution locational marginal prices[J]. IEEE Transactions on Smart Grid, 2021, 12(3): 2663-2674.
- [16] WANG Keke, YU Min, NIU Dongxiao, et al. Short-term electricity price forecasting based on similarity day screening , two-layer decomposition technique and Bi-LSTM neural network[J]. Applied Soft Computing, 2023, 136: 110018.
- [17] BOTTIEAU J, WANG Yi, DE GRÈVE Z, et al. Interpretable transformer model for capturing regime switching effects of real-time electricity prices[J]. IEEE Transactions on Power Systems, 2023, 38(3): 2162-2176.
- [18] 许越, 李强, 崔晖. 基于 MIC-EEMD-改进 Informer 的含高比例清洁能源与储能的电力市场短期电价多步预测[J]. 电网技术, 2024, 48(3): 949-957.
- XU Yue, LI Qiang, CUI Hui. Short-term multi-step price prediction for the electricity market with a high proportion of clean energy and energy storage based on MIC-EEMD-improved informer[J]. Power System Technology, 2024, 48(3): 949-957(in Chinese).
- [19] ABATE G D, HALDRUP N. Space-time modeling of electricity spot prices[J]. The Energy Journal, 2017, 38(5): 175-196.
- [20] 韩升科, 胡飞虎, 陈之腾, 等. 基于 GCN-LSTM 的日前市场边际电价预测[J]. 中国电机工程学报, 2022, 42(9): 3276-3285.
- HAN Shengke, HU Feihu, CHEN Zhiteng, et al. Day ahead market marginal price forecasting based on GCN-LSTM[J]. Proceedings of the CSEE, 2022, 42(9): 3276-3285(in Chinese).
- [21] YANG Yuyun, TAN Zhenfei, YANG Haitao, et al. Short-term electricity price forecasting based on graph convolution network and attention mechanism[J]. IET Renewable Power Generation, 2022, 16(12): 2481-2492.
- [22] ZHOU Haoyi, ZHANG Shanghang, PENG Jieqi, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, America: AAAI, 2021: 11106-11115.
- [23] WU Haixu, XU Jiehui, WANG Jianmin, et al. Autoformer: decomposition transformers with auto-correlation for long-term series forecasting[C]//Proceedings of the 35th International Conference on Neural Information Processing Systems. New York, America: NIPS, 2021: 22419-22430.
- [24] NIE Yuqi, NGUYEN N H, SINTHONG P, et al. A time series is worth 64 words: long-term forecasting with transformers[C]//Proceedings of the International Conference on Learning Representations. Kigali, Rwanda: ICLR, 2023: 1-24.
- [25] CLEVELAND R B, CLEVELAND W S, MCRAE J E, et al. STL: a seasonal-trend decomposition procedure based on loess[J]. Journal of Official Statistics, 1990, 6(1): 3-73.
- [26] TAYLOR S J, LETHAM B. Forecasting at scale[J]. The American Statistician, 2018, 72(1): 37-45.
- [27] CAO D, JIA F, ARIK S, et al. TEMPO: prompt-based

- generative pre-trained transformer for time series forecasting[C]//Proceedings of the International Conference on Learning Representations. Vienna, Austria: ICLR, 2024: 1-33.
- [28] DRUCKER H, BURGESS C J C, KAUFMAN L, et al. Support vector regression machines[C]//Proceedings of the 10th International Conference on Neural Information Processing Systems. Denver, America: NIPS, 1996: 155-161.
- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [30] YU Bing, YIN Haoteng, ZHU Zhanxing. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden: IJCAI, 2018: 3634-3640.
- [31] WU Zonghan, PAN Shirui, LONG Guodong, et al. Graph wavenet for deep spatial-temporal graph modeling[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. Macao, China: IJCAI, 2019: 1907-1913.
- [32] ZENG Ailing, CHEN Muxi, ZHANG Lei, et al. Are transformers effective for time series forecasting?[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC, America: AAAI, 2023: 11121-11128.
- [33] 胡珊珊, 张乐平, 王吉. 基于多层感知机的多维条件下电能表准确度预测模型[J]. 电测与仪表, 2018, 55(16): 125-131.
- HU Shanshan, ZHANG Leping, WANG Ji. Power meter accuracy forecast model based on multidimensional conditions by multi-layer perceptions[J]. Electrical Measurement & Instrumentation, 2018, 55(16): 125-131(in Chinese).
- [34] SHAO Zezhi, ZHANG Zhao, WANG Fei, et al. Spatial-temporal identity: a simple yet effective baseline for multivariate time series forecasting [C]// Proceedings of the 31st ACM International Conference on Information & Knowledge Management. Atlanta, America: CIKM, 2022: 4454-4458.

附录

<http://ntps.epri.sgcc.com.cn/djgxcb/CN/10.13334/j.0258-8013.pcsee.240616>



张鹏飞

在线出版日期: 2024-07-19。

收稿日期: 2024-03-17。

作者简介:

张鹏飞(1995), 男, 博士研究生, 研究方向为电力市场电价预测、电力负荷预测和人工智能在电力系统中的应用, dianlipf@smail.sut.edu.cn;

胡博(1972), 男, 博士, 教授级高级工程师, 博士生导师, 研究方向为电力系统自动化、电力市场、需求侧管理, dianlihobo@sina.com。

(责任编辑 邱丽萍)