文章编号:1674-7607(2026)01-0077-08 DOI:10.19805/j.cnki.jcspe.2026.240722

基于BSimilar优化PTransformer的光伏功率短期预测

张文广[1], 蔡 浩[1,2], 刘 科[1,2], 孙盼荣[1,2]
(1. 华北电力大学 新能源电力系统全国重点实验室,北京 102206;
2. 华北电力大学 控制与计算机工程学院,北京 102206)

摘 要:为提高光伏功率短期预测的精度,提出了考虑光伏设备性能退化因素的相似日算法优化的分时段多通道独立光伏功率短期预测方法。首先,在PTransformer模型中用分时段与通道独立的方法来处理光伏输入数据,以降低空间复杂度及提高长时间数据序列的关注度。其次,运用Transformer的编码器模型,通过自身注意力机制捕捉光伏序列特征之间的依赖关系,进行光伏功率的短期预测。最后,运用夹角余弦距离计算相似度并考虑光伏设备性能退化因素确定相似日,利用其功率数据优化PTransformer模型,以改善功率数据的滞后性。结果表明:相比典型的光伏功率短期预测方法,所提方法训练速度更快,预测精准度更高,并且对复杂天气状况下的光伏功率也有较好的预测结果。

关键词:光伏功率;短期预测;性能退化;贝叶斯分析;Transformer;相似日
中图分类号:TM615 **文献标志码**:A **学科分类号**:470.40

Short-term Prediction of PV Power Based on BSimilar Optimized PTransformer

ZHANG Wenguang[1], *CAI Hao*[1,2], *LIU Ke*[1,2], *SUN Panrong*[1,2]
(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China; 2. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: To enhance the accuracy of short-term photovoltaic power forecasting, a segmented multi-channel independent short-term photovoltaic power prediction method optimized by a similar day algorithm with consideration of photovoltaic equipment performance degradation factors was proposed. Firstly, the photovoltaic input data were processed in PTransformer model with time period and channel independent methods to reduce the spatial complexity and improve the attention of long-term data series. Secondly, the encoder model of Transformer was applied to capture the dependencies between photovoltaic sequence features through its own attention mechanism for short-term photovoltaic power prediction. Finally, the PTransformer model was optimized using its power data to improve the lag of the power data by applying the angle cosine distance to calculate the similarity and considering the performance degradation factors of the photovoltaic equipment to determine the similar day. Results show that compared with typical short-term photovoltaic power prediction method, the proposed method has faster training speed, higher predic-

收稿日期:2024-11-19 **修订日期**:2024-12-30
基金项目:华能集团总部科技项目(HNKJ20-H88-05)
作者简介:张文广(1975—),男,山东烟台人,教授,博士生导师,博士,研究方向为新能源发电功率预测及机组智能故障诊断与运维等。
蔡 浩(通信作者),男,硕士研究生,E-mail:caihao0212@163.com。

tion accuracy and better prediction results for photovoltaic power under complex weather conditions.

Key words: photovoltaic power; short-term forecast; performance degradation; Bayesian analysis; Transformer; similar day

光伏发电作为一种可再生能源发电形式，具有清洁环保、资源丰富的优点，蕴含了巨大的发展前景[1]。随着“双碳”目标的实施，近年来，光伏发电装机规模增长迅速，数据表明光伏发电已经成为电力能源的重要组成部分[2]。光伏发电受天气影响较大，具有较强的波动性与间歇性[3]。因此，对光伏出力进行精准预测对电力系统的安全稳定和调度运营极为重要。

国内外学者提出了较多光伏功率预测的方法，这些方法根据预测对象的不同可分为直接预测和间接预测[4]。间接预测方法是对地表辐照度进行预测，再以辐照度预测值为基础预测光伏功率[5-6]，但间接预测需要较为精准的天气预报数据。直接预测方法是基于历史数据建立模型进行预测，预测的精度主要由数据的相关性和模型的泛化能力决定。文献[7]通过选取待测日前 50 天数据逐时刻计算历史日和待预测日的加权欧氏距离，确定相似日优化预测模型，提高了光伏预测精度。有学者将 Transformer 模型应用于长期时间序列的预测[8-9]。文献[10]提出了基于 Transformer 编码器的功率预测模型，利用多头自注意力机制挖掘气象要素与光伏发电功率的关联信息，提升了短期功率预测效果。文献[11]采用基于 LogSparse 设计的卷积自注意力层捕获局部信息，降低了预测空间的复杂度。文献[12]利用改进的相似序列注意力模块，提出了嵌入在预测框架中的细粒度时间和云覆盖空间注意力，提高了光伏功率数据相似度的计算精度。

光伏电站的运行状态是时刻变化的，设备性能随着时间推移也会出现衰减，但现有光伏预测算法并未考虑光伏设备性能退化的因素。目前，为了提升模型预测精度，基于深度学习的光伏功率预测模型对长时间光伏序列的特征进行提取，增加了网络层数和时空复杂度，使得训练时间增加。

基于以上研究，本文提出了基于 BSimilar 优化的 PTransformer 光伏功率短期预测方法(BSPT)。在 PTransformer 模型中用分时段(Patch)与通道独立的方法来降低空间复杂度以及提高长时间数据序列的关注度，并用解码器(Encoder)预测短期的光伏功率。采用夹角余弦算法选取待测日前一年中与其相似度最高的一天，再根据贝叶斯退化模型得到性能退化的相似日功率，用该功率数据优化模型的输出结果。最后对所提出的模型方法进行了对比分析验证。

1 PTransformer 模型

PTransformer 模型具有计算复杂度低且预测精度高[13]的特点，主要由分时段、通道独立以及 Encoder 组成，前两部分主要作用是处理数据输入模型的方式，而 Encoder 则是利用注意力机制进行时空的特征提取和预测输出。

1.1 数据预处理

本文所用的数据为国内某光伏场站的历史数据，其中气象数据包含总辐照度(total solar irradiance，TSI)、直接辐照度(direct normal irradiance，DNI)、水平辐照度(global horizontal irradiance，GHI)、湿度(relative humidity，Air_H)以及水平 1.5 m 处温度(air temperature，Air_T)和压强(atmosphere，Air_P)。原始数据在采集和传输过程中由于各种因素的影响，出现了异常的数据。采用 3σ 原则[14]剔除了原始数据中的异常值，其箱线图如图 1 所示。

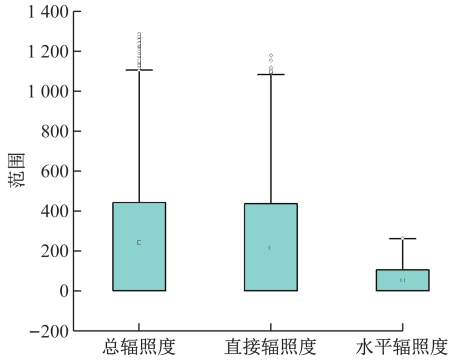


图 1 数据缺失值的箱线图

Fig. 1 Boxplot of missing data values

根据箱线图数据的缺失值，对于小步长缺失的数据部分采用线性插值法补全，对于大步长缺失的数据部分采用滑动平均法补全。

1.2 分时段

PTransformer 模型的分时段是将每个输入的单变量序列 $X(i)$ 划分为可以重叠的时间段，该过程

将生成一个时间段序列，时间段的数量 N 为

$$N=\frac{L-P}{S}+2 \tag{1}$$

式中：P 为时间段长度；S 为非重叠区域的步长；L 为输入数据数量。

分时段的引入使输入的数量从 L 减少至 L/S。与用时间段做输入方式相比，用时间点做输入方式的内存使用量和计算复杂度均降低。因此，在相同的训练时间和内存条件下，可以让模型关注更长的历史序列，显著提高其预测性能。

1.3　通道独立

通道独立指每个通道的特征表示是单个变量，每个通道与其他通道的特征表示是独立的，而 Transformer 的通道混合主要在于不同通道之间的相关性和交互性。与通道混合相比，其主要作用是扩展模型的输入数据，提升模型的泛化能力。

长度为 L 的第 i 个单变量序列为 $\boldsymbol{x}_{1:L}^{(i)}=(x_1^{(i)},x_2^{(i)},\cdots,x_L^{(i)})$，其中 $i=1,2,\cdots,M$。输入 $(x_1,x_2,\cdots,x_L)$ 被分割成 M 个单变量序列 $\boldsymbol{x}^{(i)}\in\mathbf{R}^{1\times L}$，其中每个序列都根据通道独立设置并单独输入 Encoder，然后 Encoder 输出预测结果 $\hat{\boldsymbol{x}}^{(i)}=(\hat{x}_{L+1}^{(i)},\hat{x}_{L+2}^{(i)},\cdots,\hat{x}_{L+T}^{(i)})\in\mathbf{R}^{1\times T}$，如图 2 所示。

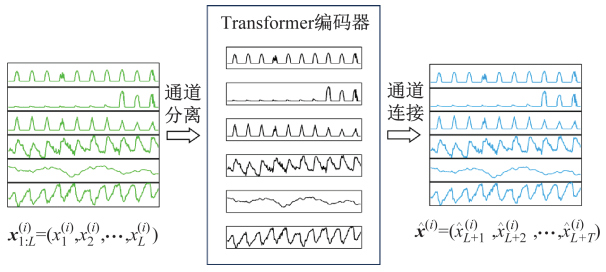


图 2　PTransformer 通道独立的示意图

Fig. 2　Schematic diagram of PTransformer channel independence

1.4　Transformer 编码器

PTransformer 模型中的主体部分用到了编码器，由多头注意力模块、残差连接和层归一化、全连接前馈网络组成，如图 3 所示。

图 3(a)为编码层：将 Patch 映射到 D 维度的 Transformer 潜在空间 $\boldsymbol{W}_p\in\mathbf{R}^{D\times P}$，并应用可学习的加权位置编码 $\boldsymbol{W}_{\text{pos}}\in\mathbf{R}^{D\times N}$ 来监控 Patch 的时间顺序 $\boldsymbol{x}_d^{(i)}=\boldsymbol{W}_p\boldsymbol{x}_p^{(i)}+\boldsymbol{W}_{\text{pos}}$，其中：$\boldsymbol{W}_p$ 为映射矩阵；$\boldsymbol{x}_p^{(i)}$ 为编码层的输入；$\boldsymbol{W}_{\text{pos}}$ 为编码矩阵；$\boldsymbol{x}_d^{(i)}$ 为 Encoder 部分的输入；p 和 d 为其各自的维度；θ 为映射变换。

图 3(b)为多头注意力层：模型中将多头注意力中的每个 head 即 $h=1,2,\cdots,H$ 转化为 $\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 矩阵，其矩阵分别为：

$$\begin{cases}\boldsymbol{Q}_h^{(i)}=(\boldsymbol{x}_d^{(i)})^{\mathrm{T}}\boldsymbol{W}_h^Q\\ \boldsymbol{K}_h^{(i)}=(\boldsymbol{x}_d^{(i)})^{\mathrm{T}}\boldsymbol{W}_h^K\\ \boldsymbol{V}_h^{(i)}=(\boldsymbol{x}_d^{(i)})^{\mathrm{T}}\boldsymbol{W}_h^V\end{cases} \tag{2}$$

式中：$\boldsymbol{Q}_h^{(i)}$、$\boldsymbol{K}_h^{(i)}$ 和 $\boldsymbol{V}_h^{(i)}$ 分别为查询矩阵、键矩阵和值矩阵；i 为第 i 个注意力头；h 为头部索引；$\boldsymbol{W}_h^Q$、$\boldsymbol{W}_h^K$ 和 $\boldsymbol{W}_h^V$ 为上述 3 个矩阵的权重矩阵；$\boldsymbol{W}_h^Q$、$\boldsymbol{W}_h^K\in\mathbf{R}^{D\times d_k}$，$\boldsymbol{W}_h^V\in\mathbf{R}^{D\times D}$。

利用缩放来获得注意力输出 $\boldsymbol{O}_h^{(i)}\in\mathbf{R}^{D\times N}$。

$$(\boldsymbol{O}_h^{(i)})^{\mathrm{T}}=\text{Attention}(\boldsymbol{Q}_h^{(i)},\boldsymbol{K}_h^{(i)},\boldsymbol{V}_h^{(i)})=$$

$$\text{Softmax}\left[\frac{\boldsymbol{Q}_h^{(i)}(\boldsymbol{K}_h^{(i)})^{\mathrm{T}}}{\sqrt{d_k}}\right]\boldsymbol{V}_h^{(i)} \tag{3}$$

式中：$\sqrt{d_k}$ 为缩放因子；Softmax 为归一化函数。

图 3(c)为残差连接和层归一化：引入残差连接来解决模型的网格退化现象和梯度爆炸。图 3 中：x 为输入特征；μ 为该层特征的均值；σ 为该层特征的标准差；$\hat{x}$ 为归一化后的输出；F 为非线性变换包

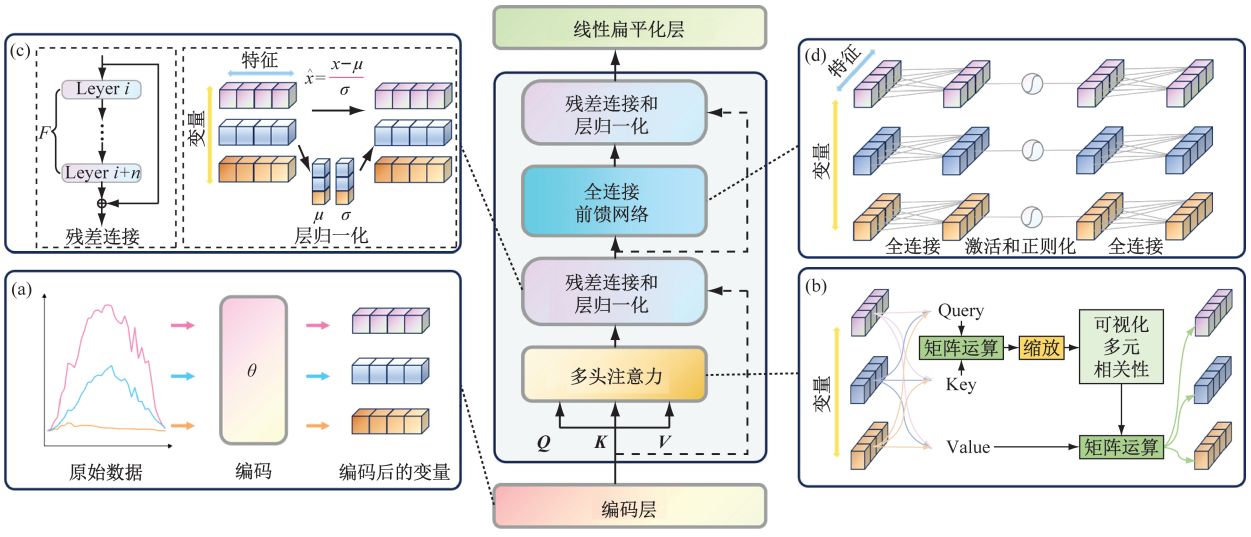


图 3　Transformer 编码器

Fig. 3　Transformer encoder

含卷积和激活等操作；i 为起始层；n 为跳跃的层数。图 3(d)为全连接前馈网络层：主要作用是对输入数据进行特征提取和非线性变换。将上一层的权重矩阵和偏置向量进行线性变换，再通过激活函数进行非线性变换输出至下一层。最后使用线性的扁平化层得到预测结果。

2 考虑设备性能退化的 BSimilar 优化模型

BSimilar 是基于相似日算法并考虑性能退化的优化模型，通过相似日算法得到相似日功率数据，利用贝叶斯退化模型计算得到性能退化的相似日功率数据。

2.1 皮尔逊相关系数分析

影响光伏发电功率的因素众多，在光伏数据中相关性较小的变量会增加相似日计算的复杂度，采用皮尔逊相关系数[15]来表示变量之间线性相关性的程度，其越接近于−1 或 1 时，变量之间的相关性就越强，而越接近 0 时，相关性就越弱。皮尔逊相关系数 r 为

$$r=\frac{\sum_{i=1}^{n_1}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n_1}(x_i-\bar{x})^2}\cdot\sqrt{\sum_{i=1}^{n_1}(y_i-\bar{y})^2}} \tag{4}$$

式中：x_i 和 y_i 分别为 2 个变量的观测值；$\bar{x}$ 和 $\bar{y}$ 分别为 2 个变量的平均值；n_1 为观测值的数量。

采用皮尔逊相关系数对光伏数据进行分析，得到图 4 所示光伏数据变量相关性热图，图中 Power 为功率。

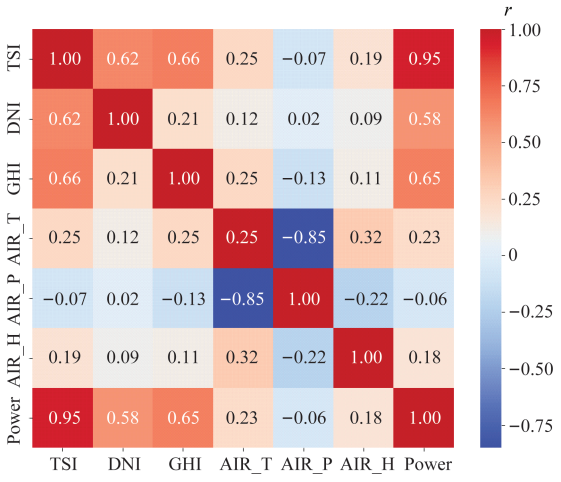


图 4 光伏数据变量相关性热图

Fig. 4 Heatmap of correlation of photovoltaic data variables

由图 4 可知，TSI、DNI、GHI 这 3 个天气因素与功率的相关性系数较高，因此选取这三项数据计算相似日。

2.2 相似日算法

采用相似日算法优化 PTransformer 模型的结果，该方法不会增加时空复杂度，并且能够达到提高预测精度的目的。利用 TSI、DNI、GHI 天气因素的相似日算法如图 5 所示。

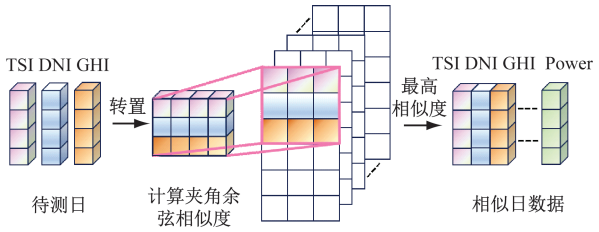


图 5 相似日算法

Fig. 5 Similar day algorithm

相似日算法采用夹角余弦[16]相似度来计算相似日和待测日的相似度。余弦相似度 θ_s 为

$$\theta_s=\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\|\boldsymbol{A}\|\times\|\boldsymbol{B}\|} \tag{5}$$

式中：$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 2 个向量；$\|\boldsymbol{A}\|$ 和 $\|\boldsymbol{B}\|$ 分别表示 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的范数，即向量长度。

2.3 贝叶斯退化模型

光伏设备性能随着时间的推移会出现衰减状况，引起光伏电站的发电效率降低。采用贝叶斯线性回归模型引入先验信息，通过不确定性估计量化退化率，计算光伏电站的性能退化程度。

首先，定义先验分布，即在没有数据时模型参数的先验概率。其似然函数 Y_{obs} 为

$$Y_{obs}\sim N(\mu,\sigma) \tag{6}$$

式中：$\mu=\alpha+\beta\cdot X_s$；$\alpha$ 为截距；β 为斜率；$N(\mu,\sigma)$ 表示均值为 μ 和标准差为 σ 的正态分布；X_s 为时间索引。

其次，采样后验分布，使用 No-U-Turn Sampler (NUTS)采样器采样得到后验样本。

最后，采用贝叶斯推断结合先验分布和观测数据来更新对参数的信念，再使用后验分布的样本取平均来估计后验均值，得到 α_{mean} 和 β_{mean}，并计算性能退化率 D_L：

$$D_L=\alpha_{mean}+\beta_{mean}\times X_m \tag{7}$$

式中：α_{mean} 为截距 α 的后验均值；β_{mean} 为斜率 β 的后验均值；X_m 为 24 个月的索引，取值范围为(0,23)。

贝叶斯退化模型引入了先验信息和后验分布，能够更全面和准确地估计性能退化率。

2.4 Squeeze-and-Excitation 层

Squeeze-and-Excitation(SE)层是一种有效的

通道注意力机制，它通过自适应调整通道上的特征响应，增强对重要特征的捕捉能力。本文将含性能退化因素的相似日功率数据与 PTransformer 模型预测的功率数据输入至 SE 层，生成最优权重向量 $\boldsymbol{W}_{s1}$ 和 $\boldsymbol{W}_{s2}$，从而利用相似日功率数据优化 PTransformer 模型预测的功率数据，得到最终优化后的功率数据。

由此，上述相似日算法、贝叶斯退化模型和 SE 层构成了 BSimilar 优化模型。运用夹角余弦方法计算相似度，得到相似度最高的相似日数据；进而利用贝叶斯退化模型计算得到性能退化后的相似日数据；最后，将其与 PTransformer 预测的功率数据输入至 SE 层，实现了对上述功率数据的优化。优化后的功率数据结合了光伏设备的性能退化因素，更接近真实情况。

3　模型实现

首先，要对原始数据进行预处理，使用第 1.1 节方法将数据中的异常值剔除，并用数据补全方法填充空缺值。本文的预测方法由 PTransformer 模型和 BSimilar 优化模型 2 个部分组成。

PTransformer 模型的训练步骤如下：

(1) 初始化模型，设置超参数和随机种群粒子数，将数据集划分为训练集和测试集，并进行正则化和归一化。

(2) 将训练集的数据进行分时段和通道独立处理，用 Encoder 进行训练，得到 PTransformer 模型。

BSimilar 优化模型的步骤如下：

(1) 根据对光伏功率影响程度筛选主要的相似度计算变量，构建相似日数据集。

(2) 利用夹角余弦相似度方法计算相似日与待测日的相似度，得到相似日功率数据。

(3) 运用贝叶斯性能退化模型计算性能退化的功率数据。

(4) 将考虑了性能退化的相似日功率数据和 PTransformer 模型预测的功率数据输入至 SE 层，得到优化后的最终功率数据。

最后使用均方误差(mean squared error，MSE)、平均绝对误差(mean absolute error，MAE)、相对平方误差(relative squared error，RSE)作为评价指标对模型进行优化调整。

BSimilar 优化 PTransformer 的光伏功率短期预测方法的流程图如图 6 所示。

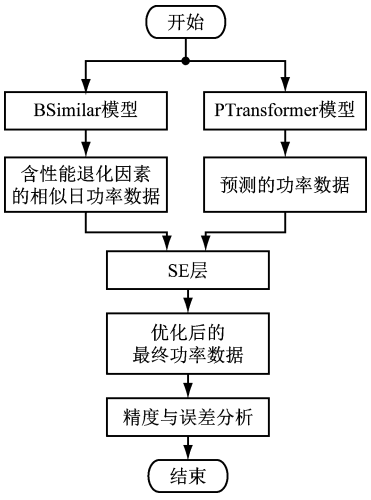


图 6　光伏功率短期预测的流程图

Fig. 6　Flowchart for short-term photovoltaic power prediction

4　算例分析

以国内某 50 MW 光伏电站作为研究对象，选取 2019 年和 2020 年全年的数据，由数据采集与监视控制(SCADA)系统间隔 15 min 采集。用线性插值法和滑动平均法对剔除异常值的数据进行补充，共计 70 177 个数据。利用皮尔逊方法得到各个变量与发电功率之间的相关度，如表 1 所示。

表 1　气象因素与功率相关度

Tab. 1　Correlation of meteorological factors and power

气象因素	TSI	GNI	GHI	AIR_T	AIR_P	AIR_H
与功率相关度	0.95	0.58	0.65	0.23	−0.06	0.18

4.1　PTransformer 的参数设置

在 PTransformer 模型中神经元数量、迭代次数、学习率等超参数的选择对模型的预测效果起着重要作用。当迭代次数超过 100 次后，测试误差有明显上升趋势，若再进行迭代则会出现过拟合的现象。所以本模型训练的迭代次数设置为 100 次。经过多次预测结果的对比，调整后的 PTransformer 参数如表 2 所示。

表 2　PTransformer 参数设置

Tab. 2　Parameter setting of PTransformer

参数	取值
Encoder 层数	2
迭代次数	100
学习率	0.000 1
Patch_len	16
Batch_size	128

4.2　BSimilar 优化

将待测日的日期设置为 2020-10-15，运用夹角余弦方法得到相似日数据集和待测日的余弦相似度 θ_s，如表 3 所示。

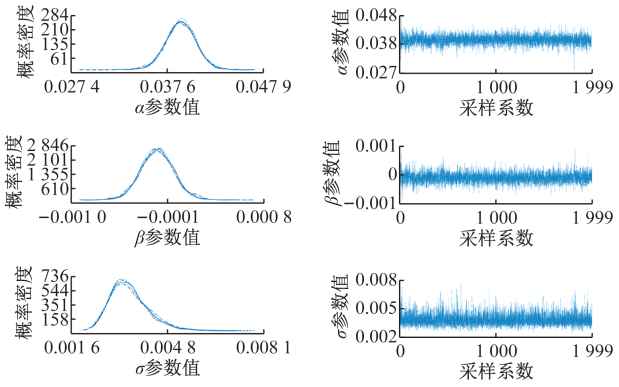
表 3 相似日与待测日相似度

Tab. 3 Similarity of similar day and target day

日期	2019-09-29	2019-10-03	2019-09-30	2020-10-14	2019-10-01
θ_s	0.988	0.996	0.994	0.930	0.994

从表 3 可以看出，2019-10-03 相似度最高，根据这天的数据可得到相似日功率数据。

将相似日数据集进行归一化处理，利用 NUTS 采样器采样 α、β、σ 的先验分布和后验分布，采样结果图 7 所示。



(a) 后验分布密度图 (b) 后验分布跟踪图

图 7 后验分布密度图和跟踪图

Fig. 7 Posteriori distribution density maps and trace maps

图 7(a)为后验分布密度图，展示每个参数在采样过程中的值，图中显示参数在采样过程中没有周期性波动，也不存在多峰分布，表明采样已收敛。图 7(b)为后验分布跟踪图，展示了参数的概率密度，根据后验分布参数计算光伏设备性能退化曲线，如图 8 所示。

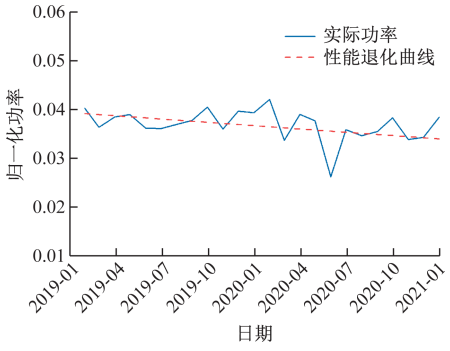


图 8 光伏设备性能退化曲线

Fig. 8 Performance degradation curves of photovoltaic devices

由图 8 可以看出，两年的功率存在退化趋势，通过采样器进行采样，生成包含 α 和 β 的后验样本，使用后验样本取平均来估计后验均值，得到 $\alpha_{mean}=0.039\ 2$，$\beta_{mean}=-0.000\ 2$，相似日和待测日的 X_m 分别为 9 和 21，根据退化曲线 $D_L=\alpha_{mean}+\beta_{mean}\times X_m$，分别得出相似日和待测日的 D_L 值，则相似日与待测日相比，退化率相差 6.41%。

通过 SE 层得到最优权重向量部分值，如表 4 所示。

表 4 SE 层的最优权重向量

Tab. 4 Optimal weight vector for SE layer

W_{s1}	W_{s2}
0.497 745 3	0.502 254 7
0.495 194 3	0.504 805 7
0.490 244 2	0.509 755 8
⋮	⋮
0.481 154 1	0.518 845 9
0.465 857 7	0.534 142 3
0.483 234 5	0.516 765 5

基于 BSimilar 优化模型，利用性能退化率和相似日功率数据可计算得到性能退化的相似日功率数据。待测日、相似日功率数据以及考虑性能退化的相似日功率数据如图 9 所示。

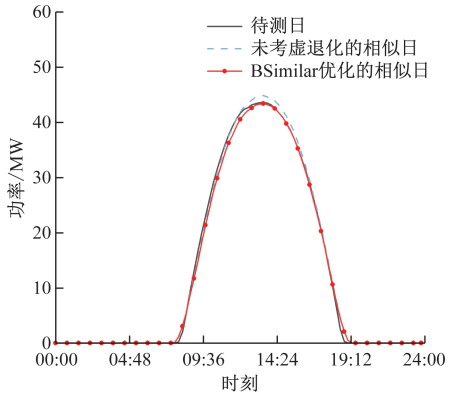


图 9 相似日与待测日功率对比

Fig. 9 Comparison of power between similar day and target day

从图 9 可以看出，未考虑性能退化因素的相似日在中午时刻比待测日功率数值高，而考虑了性能退化的相似日数据更接近待测日，精度更高。

图 9 所示天气情况良好时的功率变化小且较为稳定，而光伏发电在多云和阴雨天往往波动较大，对于相似日选取的要求更高，因此又选择了波动较大的天气作为待测日，以进一步验证本文所提方法的有效性。

图 10 为阴雨天气待测日的功率数据，由于辐照度的变化较大，导致功率波动非常明显。图 10 中展示的结果表明本文所提的 BSimilar 优化模型能够预测突变时刻的变化趋势，且在时间上不存在滞后，表明此模型也适应复杂天气情况。

4.3 模型预测结果分析

采用测试集中 2020-08-10 的功率数据，对 PTransformer 预测结果与 BSPT 预测结果进行对比。

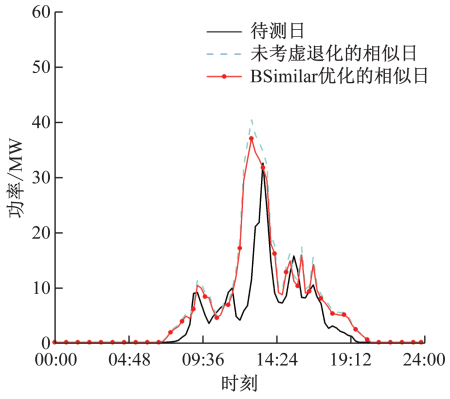


图 10　阴雨天气相似日与待测日功率的对比

Fig. 10　Comparison of power between similar day and target day in cloudy weather

从图 11 可以看出，PTransformer 的预测值比真实值有一定的滞后，而 BSPT 的预测值更接近真实值，表明 BSimilar 优化模型改善了 PTransformer 的预测滞后问题，提高了模型的预测精度。

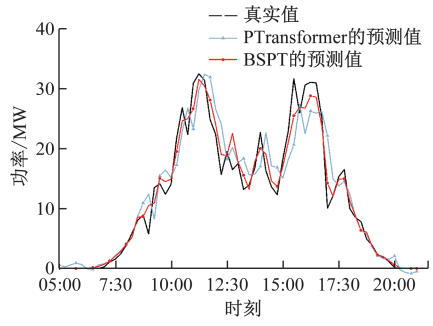


图 11　模型优化效果对比

Fig. 11　Comparison of model optimization effect

使用测试集中 2020-08-08 的功率数据进行本文方法和各典型方法预测精度的对比实验。典型方法有 Autoformer、Informer、Transformer 和 PatchTST。为能更好地显示出各模型的预测性能，在设置其迭代次数以及其他超参数时，经过多次训练选取了各个模型的最优迭代次数以及超参数。本文所提方法与其他典型方法的预测结果如图 12 所示。

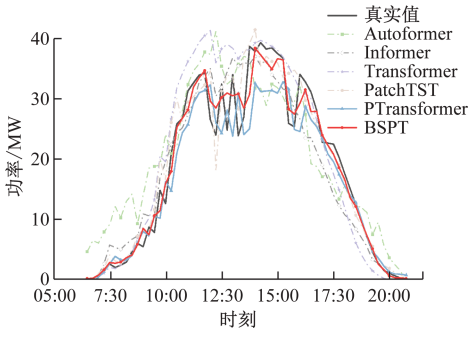


图 12　不同方法预测效果的对比

Fig. 12　Comparison of predictive effectiveness of different methods

由图 12 可以看出，这些典型方法的预测结果较真实值均有不同程度的滞后，本文 BSPT 的预测结果最接近真实值。本文模型与其他模型的误差对比见表 5。

表 5　本文模型与其他模型的误差对比

Tab. 5　Comparison of error between model in this paper and other models

模型	MSE	MAE	RSE
Autoformer	0.381	0.395	0.618
Informer	0.165	0.244	0.407
Transformer	0.155	0.217	0.394
PatchTST	0.154	0.210	0.393
PTtransformer	0.157	0.203	0.392
BSPT	0.141	0.181	0.359

从表 5 可以看出，BSPT 模型误差最小。与 Transformer 模型相比，本文 BSPT 模型的 MSE、MAE 和 RSE 分别降低了 10.19%、10.83% 和 8.41%。与以上典型模型中误差最小的 PatchTST 模型相比，本文 BSPT 模型的 MSE、MAE 和 RSE 分别降低了 8.44%、13.81% 和 8.65%，这表明 BSPT 模型的预测精度最高。

本文所提方法并不会提高模型复杂程度，表 6 分别从时间和空间的角度对比各个模型的复杂程度。由于各模型的迭代次数不同，表 6 给出了每个模型迭代的平均时间以及每次迭代所占用的服务器内存和 GPU 峰值内存。

表 6　模型训练的迭代时间和占用内存的对比

Tab. 6　Comparison of iteration time and memory usage for model training

模型	训练时间/s	所占服务器内存/MB	所占 GPU 内存/MB
Autoformer	47.98	59.27	1 171.87
Informer	45.03	59.32	809.45
Transformer	24.37	59.31	976.15
BSPT	20.62	59.00	367.14

从表 6 可以看出，基于 Transformer 所改进的 Autoformer 模型的训练时间和占用内存均比 Transformer 模型增加，Informer 模型虽然训练占用内存比 Transformer 模型低，但其迭代时间远高于 Transformer 模型，表明这 2 个模型的占用资源有所增加，而本文所提模型每次迭代的时间和占用内存最少，表明本文所提模型计算时间短，占用系统资源少。

综上所述，BSPT 预测方法对光伏功率具有很

好的预测能力，并且较之其他模型，该模型精度更高、误差更小、速度更快。

5 结论

（1）PTransformer 使用分时段减少了光伏数据输入的复杂度，提升了光伏短期预测的速度。采用通道独立的输入方式扩展了模型的输入数据，提升了模型的泛化能力，提高了模型预测精度。

（2）采用贝叶斯性能退化模型，考虑光伏设备性能退化对光伏功率预测的影响，进一步提高了 PTransformer 模型的预测精度。

（3）与其他方法相比，BSPT 的光伏功率短期预测方法训练速度更快、误差更小、精度更高。

参考文献：

[1] 陆王琳，李政，姜山. 分体式太阳能 PVT 热泵系统试验运行性能分析[J]. 动力工程学报，2024，44(10)：1574-1581.
LU Wanglin，LI Zheng，JIANG Shan. Experimental operation performance analysis on a solar PVT heat pump system with combined modules[J]. Journal of Chinese Society of Power Engineering，2024，44(10)：1574-1581.

[2] 国家统计局. 中华人民共和国 2023 年国民经济和社会发展统计公报[R]. 北京：中国统计局，2024.

[3] 董坤，冉鹏，刘旭，等. 基于权参数优化的并行深度学习光伏功率预测[J]. 动力工程学报，2024，44(1)：91-98.
DONG Kun，RAN Peng，LIU Xu，et al. Photovoltaic power prediction by weight parameter optimization-based parallel deep learning[J]. Journal of Chinese Society of Power Engineering，2024，44(1)：91-98.

[4] 董存，王铮，白捷予，等. 光伏发电功率超短期预测方法综述[J]. 高电压技术，2023，49(7)：2938-2951.
DONG Cun，WANG Zheng，BAI Jieyu，et al. Review of ultra-short-term forecasting methods for photovoltaic power generation[J]. High Voltage Engineering，2023，49(7)：2938-2951.

[5] 张姗，冬雷，纪德洋，等. 基于 NWP 相似性分析的超短期光伏发电功率预测[J]. 太阳能学报，2022，43(4)：142-147.
ZHANG Shan，DONG Lei，JI Deyang，et al. Power forecasting of ultra-short-term photovoltaic station based on NWP similarity analysis[J]. Acta Energiae Solaris Sinica，2022，43(4)：142-147.

[6] 陈志宝，丁杰，周海，等. 地基云图结合径向基函数人工神经网络的光伏功率超短期预测模型[J]. 中国电机工程学报，2015，35(3)：561-567.
CHEN Zhibao，DING Jie，ZHOU Hai，et al. A model of very short-term photovoltaic power forecasting based on ground-based cloud images and RBF neural network[J]. Proceedings of the CSEE，2015，35(3)：561-567.

[7] 李超然，潘鹏程，杨伟荣，等. 基于改进相似日优化 HBA-BiLSTM-KELM 的光伏发电功率预测[J]. 太阳能学报，2024，45(5)：508-516.
LI Chaoran，PAN Pengcheng，YANG Weirong，et al. Research on PV system power prediction based on improved similar day and HBA-BiLSTM-KELM neural network[J]. Acta Energiae Solaris Sinica，2024，45(5)：508-516.

[8] ZHOU Haoyi，ZHANG Shanghang，PENG Jieqi，et al. Informer：beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the **35**th AAAI Conference on Artificial Intelligence. Honolulu，USA：AAAI，2021：11106-11115.

[9] 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. Red Hook，USA：Curran Associates Inc.，2021：22419-22430.

[10] 黄莉，甘恒玉，刘兴举，等. 基于 Transformer 编码器的超短期光伏发电功率预测[J]. 智慧电力，2024，52(5)：16-22，59.
HUANG Li，GAN Hengyu，LIU Xingju，et al. Ultra-short-term photovoltaic power generation prediction based on transformer encoder[J]. Smart Power，2024，52(5)：16-22，59.

[11] LI Shiyang，JIN Xiaoyong，XUAN Yao，et al. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems . Red Hook，USA：Curran Associates Inc.，2019：471.

[12] YANG Jing，HE Hui，ZHAO Xuemeng，et al. Day-ahead PV power forecasting model based on fine-grained temporal attention and cloud-coverage spatial attention[J]. IEEE Transactions on Sustainable Energy，2024，15(2)：1062-1073.

[13] NIE Yuqi，NGUYEN N H，SINTHONG P，et al. A time series is worth 64 words：long-term forecasting with transformers[C]//International Conference on Learning Representations. Kigali：ICLR，2023.

（下转第 102 页）

Generation, 2023, 52(9): 155-161.

[39] 李旭东. 电热熔盐单罐设计方法及释热过程中罐内温度场变化规律研究[D]. 石家庄: 河北科技大学, 2023.

[40] 张腾飞, 田书, 郭成威. 考虑可控负荷的光热电站和风电系统调度策略[J]. 电力系统及其自动化学报, 2022, 34(9): 32-39.
ZHANG Tengfei, TIAN Shu, GUO Chengwei. Scheduling strategy for concentrating solar power plant and wind power system considering controllable load[J]. Proceedings of the CSU-EPSA, 2022, 34(9): 32-39.

[41] XU Ti, ZHANG Ning. Coordinated operation of concentrated solar power and wind resources for the provision of energy and reserve services[J]. IEEE Transactions on Power Systems, 2017, 32(2): 1260-1271.

[42] 吴玉庭, 张晓明, 王慧富, 等. 基于弃风弃光或低谷电加热的熔盐蓄热供热技术及其评价[J]. 中外能源, 2017, 22(2): 93-99.
WU Yuting, ZHANG Xiaoming, WANG Huifu, et al. Molten salt heat storage and supply technology based on heating using abandoned wind power, PV power or off-peak power[J]. Sino-Global Energy, 2017, 22(2): 93-99.

[43] BLOESS A, SCHILL W P, ZERRAHN A. Power-to-heat for renewable energy integration: a review of technologies, modeling approaches, and flexibility potentials[J]. Applied Energy, 2018, 212: 1611-1626.

[44] 张学镭, 梁建雄, 王普, 等. 基于 S-CO_2 布雷顿循环的太阳能发电储能一体化系统性能分析[J]. 动力工程学报, 2024, 44(8): 1205-1215.
ZHANG Xuelei, LIANG Jianxiong, WANG Pu, et al. Performance analysis of an integrated solar power generation and storage system based on the S-CO_2 brayton cycle[J]. Journal of Chinese Society of Power Engineering, 2024, 44(8): 1205-1215.

(上接第 84 页)

[14] 杨茂, 孟玲建, 李大勇, 等. 基于类 3σ 准则的光伏功率异常数据识别[J]. 可再生能源, 2018, 36(10): 1443-1448.
YANG Mao, MENG Lingjian, LI Dayong, et al. Identification of abnormal data of photovoltaic power based on class 3σ[J]. Renewable Energy Resources, 2018, 36(10): 1443-1448.

[15] MISHRA M, BYOMAKESHA DASH P, NAYAK J, et al. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction[J]. Measurement, 2020, 166: 108250.

[16] 王言文, 邱启荣, 王宝坤. 基于 Lasso 和 SVR 的向量夹角余弦变权重组合预测模型[J]. 统计与决策, 2020, 36(18): 22-26.
WANG Yanwen, QIU Qirong, WANG Baokun. Forecasting model of vector angle cosine variable weight combination based on lasso and support vector regression[J]. Statistics & Decision, 2020, 36(18): 22-26.