

9-10-2024

Short-term prediction method of distributed PV output power based on K-I-ELM multi-model integration

Zhuohan JIANG

Economic & Technical Research Institute, State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Hunan Key Laboratory of Energy Internet Supply-demand and Operation, Changsha 410004, China

Shengyu ZHOU

Economic & Technical Research Institute, State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Hunan Key Laboratory of Energy Internet Supply-demand and Operation, Changsha 410004, China

Yuqing HE

Economic & Technical Research Institute, State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; Hunan Key Laboratory of Energy Internet Supply-demand and Operation, Changsha 410004, China

Renjun ZHOU

School of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China

Chenhao SUN

School of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China, chenhaosun@csust.edu.cn

Follow this and additional works at: <https://jepst.researchcommons.org/journal>

Recommended Citation

JIANG, Zhuohan; ZHOU, Shengyu; HE, Yuqing; ZHOU, Renjun; and SUN, Chenhao (2024) "Short-term prediction method of distributed PV output power based on K-I-ELM multi-model integration," *Journal of Electric Power Science and Technology*. Vol. 39: Iss. 4, Article 17.

DOI: 10.19781/j.issn.1673-9140.2024.04.017

Available at: <https://jepst.researchcommons.org/journal/vol39/iss4/17>

This Clean energy and energy storage is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology. For more information, please contact dlxb04@163.com.

引用格式:江卓翰,周胜瑜,何禹清,等.基于K-I-ELM多模型集成的分布式光伏出力短期预测方法[J].电力科学与技术学报,2024,39(4):146-152.

Citation: JIANG Zhuohan, ZHOU Shengyu, HE Yuqing, et al. Short-term prediction method of distributed PV output power based on K-I-ELM multi-model integration[J]. Journal of Electric Power Science and Technology, 2024, 39(4): 146-152.

基于 K-I-ELM 多模型集成的分布式光伏 出力短期预测方法

江卓翰^{1,2}, 周胜瑜^{1,2}, 何禹清^{1,2}, 周任军³, 孙辰昊³

(1. 国网湖南省电力有限公司经济技术研究院, 湖南 长沙 410004; 2. 能源互联网供需运营湖南省重点实验室, 湖南 长沙 410004;

3. 长沙理工大学电气与信息工程学院, 湖南 长沙 410114)

摘要:为响应“双碳”目标,高比例新能源接入的新型电力系统已成为下一个发展目标。光伏作为当前电力系统能源发电主体形式之一,其出力特性数据尚存在多源、异构及高维等分布特点,导致不同特征作用机理、机制较为复杂,继而加大分布式光伏系统出力的预测难度。为此,首先构建核主成分分析(kernel principle component analysis, KPCA)模型,通过核函数在特征空间中依据不同特征的有效信息蕴含度提取主成分;然后采用信息熵(information entropy, IE)模型,根据各主成分信息负载度量加权系数,综合求解相应作用权重;最后依据特征评估结果,针对性设置极限学习机(extreme learning machine, ELM)网络参数,降低预测不确定度。最终整合多类别数据挖掘模型,构建 K-I-ELM 预测方法,在复杂数据环境下实施光伏出力短期预测。基于某实际台区光伏发电数据进行案例分析,论证所提方法针对不同数据环境的适应性及较高的预测精度。

关键词:信息熵;核主成分分析;极限学习机;短期预测;光伏出力

DOI: 10.19781/j.issn.1673-9140.2024.04.017

中图分类号: TM615

文章编号: 1673-9140(2024)04-0146-07

Short-term prediction method of distributed PV output power based on K-I-ELM multi-model integration

JIANG Zhuohan^{1,2}, ZHOU Shengyu^{1,2}, HE Yuqing^{1,2}, ZHOU Renjun³, SUN Chenhao³

(1. Economic & Technical Research Institute, State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China; 2. Hunan Key Laboratory of

Energy Internet Supply-demand and Operation, Changsha 410004, China; 3. School of Electrical & Information Engineering,

Changsha University of Science & Technology, Changsha 410114, China)

Abstract: In response to the "dual carbon" strategy, a new type of power system with a high proportion of renewable energy access has become the next development goal. As one of the main forms of current energy generation, photovoltaic (PV) power generation has characteristics such as multi-source, heterogeneous, and high-dimensional data distribution, which makes the mechanisms and effects of different features relatively complex and subsequently increases the difficulty of predicting the output of distributed PV systems. To address this, multiple categories of data mining models are integrated to construct an K-I-ELM prediction method for short-term PV output prediction in complex data environments. First, a kernel principal component analysis (KPCA) model is constructed to extract principal components based on the effective information contained in different features in the feature space through a kernel function. An information entropy (IE) model is employed to measure the weighting coefficients based on the information load of each principal component and comprehensively solve the corresponding effect weights. Finally,

收稿日期: 2022-09-21; 修回日期: 2023-03-24

基金项目: 国家自然科学基金(52207074); 国网湖南省电力有限公司科技项目(5216A22001J); 湖南省科技创新平台与人才计划(2019TP1053)

通信作者: 孙辰昊(1991—), 男, 博士, 讲师, 主要从事电力数据挖掘及应用和人工智能的研究; E-mail: chenhaosun@csust.edu.cn

based on the feature evaluation results, the network parameters of the extreme learning machine (ELM) are set specifically to reduce prediction uncertainty. A case study based on actual PV power generation data from a certain substation demonstrates the adaptability and high prediction accuracy of the proposed method in different data environments.

Key words: information entropy; kernel principle component analysis; extreme learning machine; short-term prediction; photovoltaic output

近年来,世界人口正在快速增长,全球能源消耗和需求也在增长。减少二氧化碳排放量的问题则引起了全世界对可再生能源的关注。由于太阳能资源丰富、安全可靠且能够减少土地资源浪费等特点,光伏发电逐渐成为解决化石能源短缺的重要方案^[1-4]。但光伏发电的间歇性和随机波动性无疑给电网带来了诸多难题,因此,准确的光伏发电预测能更合理地帮助电力调度部门及时调整计划,促进电网对新能源的消纳能力,这对电网的优化运行和控制具有重要意义^[5]。

随着光伏发电厂容量的不断增长及其在电力系统中的渗透性不断提高,如何处理准确的短期预测工具显得尤为重要。近几年,光伏发电的短期预测方法大致可分为两大类:通过曲线拟合、数学统计等方法直接预测光伏系统输出功率的直接预测法;通过对太阳辐射的先验预测和对电力的后验变换来预测光伏电站发电量的间接预测法。

直接预测方法主要使用分类回归法、时间序列法、灰色理论法和模糊理论法等来预测光伏发电量。文献[6]针对目前油纸绝缘状态在单一特征量评估中容易存在评估结果的片面性、不准确性等问题,提出一种基于回复电压多参量回归的油纸绝缘状态诊断方法;文献[7]为了减少光伏系统接入电网产生的不利影响,对预测光伏功率输出进行研究,提出了一种基于数据中潜在季节类别的混合深度学习模型;文献[8]运用灰色模型对光伏发电量进行总体趋势的预测,再加入加权马尔科夫链预测理论,建立了灰色—加权马尔科夫链预测模型;文献[9]为了对变压器状态做出快速准确评价,提出一种基于优劣解距离(technique for order preference by similarity to ideal solution, TOPSIS)和灰色关联分析的状态评价方法;文献[10]使用一种自适应神经复杂模糊推理系统来预测光伏发电量。以上方法需要确定国内电力生产总值中的消耗量、负荷、时间以及行业总产值之间的关系和变化,再利用数学模型进行预测,这增加了数据处理与获取的难度,且整个过程将持续进行数学模型校准和调整,从而需要更长的时间来完成。

而间接预测法近些年已成为许多学者的重点研究方向之一,其中应用最为广泛的是人工智能算法,包括神经网络、支持向量机、粒子群算法、遗传算法(genetic algorithm, GA)和聚类算法等。文献[11]基于地基云图分析和处理,将图像亮度和云量作为云对辐射影响的重要特征量,参与基于径向基函数(radical basis function, RBF)神经网络的地表辐射预测建模,然后结合光电转化模型实现光伏功率超短期预测。由于其方法简单、学习过程收敛速度快,但是预测精度不高,不适用于所有天气类型,因此,学者们利用BP、深度前馈及长短期记忆神经网络进行光伏预测^[12-14]。文献[15]针对传统相似度函数在聚类过程中所存在的问题,提出改进灰色—欧氏距离相似度函数,再采用二次自适应支持向量机模型分别进行光伏功率预测;文献[16]提出了一种包含天气预报信息的粒子群—稀疏贝叶斯混合算法的发电预测方法;文献[17]针对光伏发电系统出力波动问题,提出GA-模糊RBF神经网络的光伏发电功率预测模型,将功率预测值应用于光伏发电的蓄电池储能功率调节系统,以降低对电网的冲击;文献[18]针对聚类优化处理导致预测耗时较长、短期负荷预测精度偏低等问题,提出一种基于改进K-means算法的电力短期负荷预测方法。这类方法不需要获得相关的物理模型,而是通过对历史数据的不断训练得到光伏发电出力的预测模型,广泛应用于分析众多不确定性与电力负荷预测的相关性。但是,如何准确描述人工不确定性时所采用的标准相对困难。

由于光伏出力数据具有多源、异构及高维等分布特性,致使各数据特征作用机理、机制较为复杂,进而增大分布式光伏系统出力预测难度。为此,首先基于日平均温度、最大温度、最小温度风向、平均风速以及地形等多类别数据特征,构建核主成分分析(kernel principle component analysis, KPCA)模型,采用核函数在特征空间综合评估各特征所涵盖的有效信息,从而筛选主成分;然后采用信息熵(information entropy, IE)模型,根据信息负载度量各主成分加权系数以求解作用权重,为后续针对性

挖掘提供条件;最后,依据特征评估结果,针对性设置极限学习机(extreme learning machine, ELM)预测模型相关参数,并最终整合为K-I-ELM多模型集成预测方法,降低预测过程中的潜在不确定度,基于某区域配网的光伏发电历史数据,实施实例仿真分析。仿真结果论证所提方法在复杂数据环境下仍可保持较高的预测精度。

1 特征评估途径

1.1 核主成分分析

KPCA是一种特征提取技术,它是一种基于核的主元成分分析(principle component analysis, PCA)方法,利用核函数将数据映射到非线性空间,可以使用核模型将数据减少到 n (样本数)个非线性主分量,从而减少高维空间下的数据计算负担。

在标准PCA中,特征空间中增加的数据维度有助于绘制变量之间的线性关系。主要通过对数据协方差矩阵进行对角化实现降维,以便解耦数据中存在的非线性相关度,假设特征空间 $X = (x_1, x_2, \dots, x_M)$, M 为其数据集的样本个数,数据在特征空间中的协方差矩阵为

$$C^X = \frac{1}{M} \sum_{j=1}^M \phi(x_j) \phi(x_j)^T \quad (1)$$

式中, $\phi(\cdot)$ 为将数据从输入空间映射到特征空间的核函数。

求解特征空间中的特征值:

$$v'kv = \lambda \quad (2)$$

式中, $k = V\Lambda V'$, Λ 是一个包含 k 的特征值 λ_i 的对角矩阵, V 是特征向量矩阵; v 为特征向量, 是特征空间中样本点的线性组合, $v = \sum_{i=1}^M \alpha_i \phi(x_j)$, 其中 α_i 是线性组合的系数。

在特征空间 X 中, 样本数据的特征向量 $V \in X$ 满足以下关系:

$$\lambda v = C^X v \quad (3)$$

整合式(1)~(3)可得:

$$C^X v = \left(\frac{1}{M} \sum_{j=1}^M \phi(x_j) \phi(x_j)^T \right) v = \frac{1}{M} \sum_{j=1}^M \langle \phi(x_j), v \rangle \phi(x_j) \quad (4)$$

式中, $\langle \phi(x_j), v \rangle$ 为 $\phi(x_j)$ 和 v 之间的点积。

由式(4)可得, 非零特征向量的解位于集合 $\{\phi(x_1), \phi(x_2), \dots, \phi(x_M)\}$ 中。因此, 结合式(3)、(4)可得:

$$\lambda \sum_{i=1}^M \alpha_i \langle \phi(x_k), \phi(x_i) \rangle = \frac{1}{M} \sum_{i=1}^M \alpha_i \left\langle \phi(x_k), \sum_{j=1}^M \phi(x_j) \right\rangle \langle \phi(x_j), \phi(x_i) \rangle \quad (5)$$

当 $k = 1, 2, \dots, M$ 时, 针对所有样本数据集, 式(5)所表示的特征值问题只需在特征空间中映射函数的点积即可解决。然而, 即使确定特征值存在, 映射 $\phi(\cdot)$ 的计算过程也存在一定的困难。为此, 基于特征空间中 2 个向量的点积, 通过 $K_{ij} = \langle \phi(x_i), \phi(x_j) \rangle$ 定义核矩阵 K , 式(5)可简化为

$$\lambda \sum_{i=1}^M \alpha_i \langle \phi(x_k), \phi(x_i) \rangle = \lambda \sum_{i=1}^M \alpha_i K_{ki} \quad (6)$$

则需解决的特征值问题可转化为

$$\lambda \alpha = K \alpha \quad (7)$$

式中, α 为 K 的特征向量; 核矩阵 K 是实现 KPCA 降维的关键, 主要用来执行从输入空间到特征空间的转换。

因数据在高斯核函数传递的过程中不容易发散, 所以本文采用高斯核函数, 其计算公式如下:

$$k(x, y) = \exp(-\gamma \|x - y\|^2) \quad (8)$$

式中, x, y 为输入空间中的 2 个样本点; γ 为预先制定的参数, 其值取决于训练数据集; $\|x - y\|$ 表示样本点之间的欧氏距离。

与标准 PCA 一样, 需求解最大特征值对应的特征向量, 即 KPCA 需要投影的方向, 则主成分的计算由核函数计算得到:

$$\Phi^m = \sum_{i=1}^n \alpha_i^m k(x^i, x^j) \quad (9)$$

在通常情况下, 数据集的中心不再是特征空间的原点。为此, 为方便数据计算, 映射数据集中化可得核矩阵公式如下:

$$K_c = K - KE_N - E_N K + E_N KE_N \quad (10)$$

式中, K_c 为经过集中化替换后的核矩阵; E_N 为 N 阶方阵。

1.2 信息熵

信息量度量的是一个具体事件发生所带来的信息, 而熵则是在结果出来之前对可能产生的信息量的期望即考虑该随机变量的所有可能取值, 即所有可能发生事件所带来的信息量的期望。熵的概念最初来源于热力学, 用于描述不可逆现象的运动过程, 后来在信息论中用于表示事物的不确定性。熵是一种客观赋权法, 其通过随机抽样实验结果的抽样不确定度来定义信息熵大小, 可以用来解决信

息冗余和不确定多属性问题。由于熵与有序度之间存在一定的关系,熵越大其有序度越低、不确定性越大;相反,熵越小其有序度越高、不确定性越小。因此,可以根据建立的指标体系,利用熵来定量确定各指标的权重,从而进行多指标综合评价,最终确定对整个系统影响较大的因素。

熵值法是根据信息负载指标来确定指标权重的大小。基于评估矩阵如下:

$$D = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (11)$$

其中, x_{ij} 是第 j 个指标属性的第 i 个指标值。

1) 将评估矩阵 D 归一化得到矩阵 $R = (r_{ij})_{n \times m}$, 其根据不同的需求选择不同的形式, 若较大值为更优选择, 则 $r_{ij} = (x_{ij} - \min x_{ij}) / (\max x_{ij} - \min x_{ij})$; 若较小值为更优选择, 则 $r_{ij} = (\min x_{ij} - x_{ij}) / (\max x_{ij} - \min x_{ij})$ 。

2) 求指标在各方案下的比值 l_{ij} , 即第 j 项指标在第 i 个方案中占该指标的比重, 其数学表达式为

$$l_{ij} = r_{ij} / \sum_{i=1}^n r_{ij} \quad (12)$$

3) 通过 l_{ij} 可计算第 j 项指标的熵, h_j 表示所有方案对第 j 项指标的熵贡献总量, 其数学表达式为

$$h_j = -k \sum_{i=1}^n l_{ij} \ln l_{ij} \quad (13)$$

其中, 常数 $k = 1 / \ln n$, 当每个方案指标的贡献率趋于一致时, 即可保证 h_j 趋于 1。

4) 计算指标差异系数 g_j 。对于给定的 j , r_{ij} 的差异性越小则 h_j 越大; r_{ij} 全部相等时 $h_j = h_{\max} = 1$, 此时对于方案的比较将毫无作用, 而各方案的指标值相差越大时 h_j 越小, 该项指标对于方案的比较作用越大。 g_j 数学表达式为

$$g_j = 1 - h_j \quad (14)$$

5) 确定权重系数。加权后的系数 w_j 为权重系数的归一化权重, 其数学表达式为

$$w_j = g_j / \sum_{j=1}^m g_j \quad (15)$$

6) 得到评估分数为

$$S = \lambda D w^T \quad (16)$$

其中, w 为权重向量, 是通过熵值法计算出来的各指标的权重。评估组权重 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$ 计算方式为

$$\sum_{j=1}^m \lambda_m = 1 \quad (17)$$

2 数据预测模型

极限学习机是一种新型的前馈神经网络学习方法, 具有训练简单、抗干扰性强的特点。传统神经网络在训练前需要人为输入训练次数、训练精度、隐含层节点以及输入层节点等信息。如果隐含层节点太少就可能学习不够充分, 而隐含层节点太多, 则可能存在过拟合无法预测的情况。ELM 可以克服基于梯度的学习神经网络速度慢、输入参数多的传统缺点, 其只需输入隐含层的节点数就可以随机产生输入权值和隐含层偏移值, 这种方法基于简单的线性回归, 但可以得到较为满意的结果。

神经网络利用数学方法模拟人脑网络的特性, 根据不同网络拓扑结构, 神经网络大致可分为 3 类: 前向、反馈和自组织网络。ELM 属于单隐层前向网络, 其具体结构如图 1 所示。

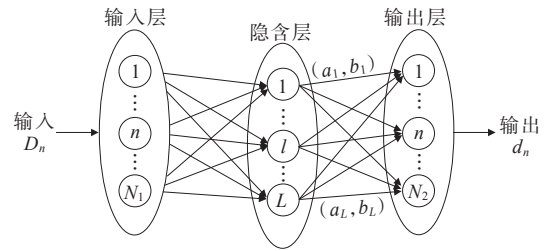


图1 ELM 网络结构

Figure 1 Illustration of ELM structure

根据图 1 中的信息, ELM 由输入层、隐藏层和输出层组成。假设输入数据集为 $D_n (n = 1, 2, \dots, N)$, 输出数据集为 $d_n (n = 1, 2, \dots, N)$, N 为训练数据集总数, 且输入层包含 N_1 个神经元, 输出层包含 N_2 个神经元, 隐藏层包含 L 个神经元, 其神经元个数代表各数据集的维数。ELM 的目标是找出输入与输出数据的关系, 输出项可由以下公式计算:

$$\sum_{l=1}^L \omega^T \cdot f(a_l, b_l, D_l) = d_n \quad (18)$$

式中, a_l, b_l 均为隐藏节点的参数; ω^T 为连接隐含和输出节点的输出权重参数; f 为激励函数, 本文选取 sigmoid 函数。

式(18)可简写为

$$H\beta = d \quad (19)$$

$$H = \begin{bmatrix} f(a_1, b_1, D_1) & \cdots & f(a_L, b_L, D_1) \\ \vdots & \ddots & \vdots \\ f(a_1, b_1, D_N) & \cdots & f(a_L, b_L, D_N) \end{bmatrix}_{N \times L}$$

式中, H 为隐含层输出矩阵; β 为输出权值矩阵, $\beta = [\beta_1, \beta_2, \dots, \beta_L]^T$; d 为期望的输出矩阵, $d = [d_1, d_2, \dots, d_N]^T$ 。

若 H 为可逆矩阵, 则可容易求解, 而 H 可逆的前提是其矩阵为一个方阵。当隐层神经元数与训练样本数相同时, ELM 的神经网络可以有效地逼近训练样本。然而, 当隐层神经元数超过一定值时, 极端学习机神经网络可能会出现过拟合现象; 同时, 当隐层神经元较大时, 需要计算的变量数量将呈指数增长, 增加了计算负担。因此, 隐层神经元数量通常小于训练样本数量, 求解矩阵 β 转化为以下优化问题。

若前馈神经网络隐藏层的激活函数是无穷可微的, 则可以得到线性系统的最小二乘解 $\hat{\beta}$, 即输出权重的 β 范数最小, 由此, 可以通过求解 $\hat{\beta}$ 获得最小的训练误差, 即

$$\hat{\beta} = H^T d \quad (20)$$

$$\min \|\beta^T H - d\| \quad (21)$$

其中, H' 为 H 的广义逆矩阵, $H' = H^T (1/c_r + H^T H)^{-1}$, c_r 为正则化参数。

3 预测流程

本文所构建 K-I-ELM 模型运行流程如图 2 所示。

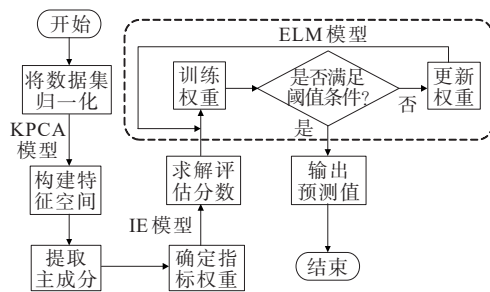


图 2 K-I-ELM 预测流程

Figure 2 Prediction procedure of K-I-ELM

4 算例分析

本文基于东北某地区分布式光伏接入配网数据实施算例仿真, 样本数据以天为采集单位周期, 每日出力数据的采集以小时为单位时限, 算例中主要采用该配网出力浮动最为显著的 7 月出力数据作为对象。除去出力数据、日期, 样本数据中的环境因素主要由气象、地理特征构成。气象特征包括日平均温度、最大温度、最小温度, 相对湿度、日降水量, 平均气压、平均风速, 风向、天气类型以及云量

等; 地理特征包括地形与海拔。仿真采用十重交叉验证, 以 70% 数据样本作为训练集、15% 为验证集、15% 为测试集。

将标准化的数据样本集进行主成分计算, 以衡量各特征所包含信息的有效程度。本文选用累计百分比代表主成分方差量即数据有效信息, 故得到的累计方差百分比为该特征的累计贡献率, 可用于判定主成分 (principle component, PC)。总计 12 个特征的累计贡献率具体如图 3 所示, 可知以 95% 的累计方差贡献率为限, 共计 4 个 PC 即可满足后续分析需求。在此基础上, 基于 IE 模型求解各特征相应主成分对应的加权系数, 并整合计算作用权重, 结果如表 1 所示, 加权系数 $\omega_1 \sim \omega_4$ 分别对应 4 个 PC。

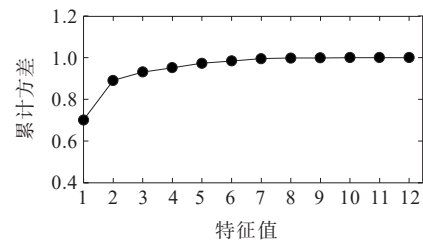


图 3 各特征累计方差百分比

Figure 3 Percentage of cumulative variance for each feature

表 1 各特征相应主成分对应加权系数

Table 1 Principal components of each feature correspond to weighted coefficients

特征	ω_1	ω_2	ω_3	ω_4	作用权重
平均温度	-0.028	0.034	0.095	-0.047	0.089
最大温度	-0.009	-0.022	0.057	-0.017	0.081
最小温度	-0.041	0.027	0.031	-0.068	0.092
相对湿度	0.025	0.010	-0.027	0.005	0.038
日降水量	-0.005	-0.056	0.045	0.012	0.152
平均气压	0.002	0.031	-0.055	0.028	0.041
风向	0.005	-0.012	-0.019	0.027	0.013
平均风速	-0.010	-0.062	-0.029	-0.033	0.057
天气类型	0.057	0.076	-0.034	0.088	0.186
云量	0.048	-0.092	-0.037	0.085	0.168
地形	-0.027	0.006	-0.012	0.020	0.073
海拔	-0.024	-0.017	0.054	-0.033	0.010

最后, 基于上述结果针对性设置 ELM 模型参数, 并将基于所提 K-I-ELM 模型的预测结果分别与广义回归神经网络 (generalized regression neural network, GRNN) 和 RBF 神经网络模型预测结果进行对比分析, 如图 4 所示 (选择 7 月 15 日数据作为典型日)。

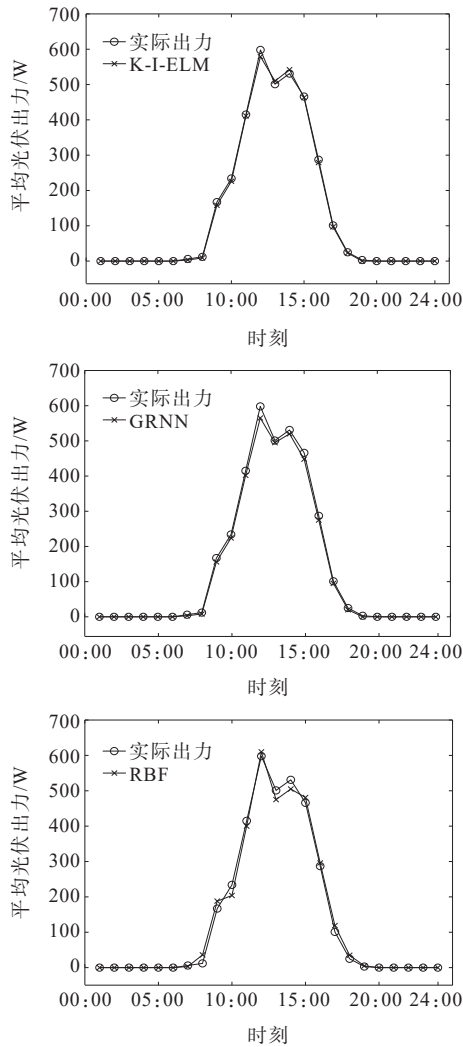


图4 各预测模型效果对比

Figure 4 Effect comparison of each prediction model

为量化预测效果,本文采用均方根误差(root mean squared error, RMSE)和决定系数作为评价指标,具体计算方法如下:

$$\Delta e_{\text{RMSE}} = \frac{1}{N} \sum_{i=1}^N (p_i - x_i)^2 \quad (22)$$

$$R^2 = \frac{\sum_{i=1}^N (p_i - p_{av})^2}{\sum_{i=1}^N (x_i - p_{av})^2} \quad (23)$$

式中, p_i 为任意一个预测值, x_i 为对应真实值, p_{av} 是预测值均值, N 是样本总数。 Δe_{RMSE} 越小, 则预测精度越高; R^2 取值为 0~1, 其值越接近 1 则预测结果越接近真实值。

预测效果对比如表 2 所示, 相较其他 2 种预测模型, 所提 K-I-ELM 模型的决定系数更接近 1, RMSE 也相对更低。综合来看, K-I-ELM 模型在预测精度上具有一定的优势。

表2 预测模型对比

Table 2 Comparison of prediction models

预测模型	Δe_{RMSE}	R^2
K-I-ELM	0.203	0.977
GRNN	0.328	0.957
RBF	0.559	0.913

5 结语

针对分布式光伏出力数据分布特性, 本文提出一种基于多模型集成的 K-I-ELM 预测模型, 以实现多源、异构及高维数据环境下的光伏出力短期预测。

1) 构建 KPCA 模型, 选择核函数并建立特征空间从而提取主成分, 以评估数据集中各特征有效性;

2) 应用 IE 模型, 依据信息负载程度, 度量各主成分相应的作用权重;

3) 依据特征评估结果设置 ELM 预测模型参数, 并整合为 K-I-ELM 预测方法, 降低预测过程中的不确定度;

4) 仿真实验验证了 K-I-ELM 预测模型在数据环境适应度以及预测精度上具有一定优势, 可满足实际应用需求。

参考文献:

- [1] 丁明, 王伟胜, 王秀丽, 等. 大规模光伏发电对电力系统影响综述[J]. 中国电机工程学报, 2014, 34(1): 1-14.
DING Ming, WANG Weisheng, WANG Xiuli, et al. A review on the effect of large-scale PV generation on power systems[J]. Proceedings of the CSEE, 2014, 34(1): 1-14.
- [2] 贾科, 宣振文, 林瑶琦, 等. 基于 Adaboost 算法的并网光伏发电系统的孤岛检测法[J]. 电工技术学报, 2018, 33(5): 1106-1113.
JIA Ke, XUAN Zhenwen, LIN Yaoyi, et al. An islanding detection method for grid-connected photovoltaic power system based on Adaboost algorithm[J]. Transactions of China Electrotechnical Society, 2018, 33(5): 1106-1113.
- [3] 杜磊, 赵涛, 冯之健, 等. 单相短路故障条件下级联模块中压光伏发电系统的有功功率回流抑制[J]. 电工技术学报, 2022, 37(20): 5201-5213.
DU Lei, ZHAO Tao, FENG Zhijian, et al. Active power backflow suppression of cascaded module medium-voltage PV power generation system during single-phase short-circuit fault[J]. Transactions of China Electrotechnical Society, 2022, 37(20): 5201-5213.
- [4] 田雨果, 王彤, 邢其鹏, 等. 计及虚拟惯量控制与低电压穿越的光伏发电系统暂态稳定分析[J]. 电力系统保护与控制, 2022, 50(2): 52-59.

- TIAN Yuguo, WANG Tong, XING Qipeng, et al. Transient stability analysis of a photovoltaic generation system considering virtual inertia control and low voltage ride-through[J]. Power System Protection and Control, 2022, 50 (2):52-59.
- [5] 虎智峰,陈静,张婧菲,等.考虑新能源不确定性边界的主动配电网优化调度[J].智慧电力,2022,50 (11):48-55.
HU Zhifeng, CHEN Jing, ZHANG Jingfei, et al. Optimal dispatch of active distribution network considering uncertainty boundary of renewable power generation[J]. Smart Power, 2022, 50 (11):48-55.
- [6] 蔡金锭,叶荣,陈汉城.回复电压多元参数回归分析的油纸绝缘老化诊断方法[J].电工技术学报,2018,33(21):5080-5089.
CAI Jinding, YE Rong, CHEN Hancheng. Aging diagnosis method of oil-paper insulation based on multiple parameter regression analysis of recovery voltage[J]. Transactions of China Electrotechnical Society, 2018, 33(21):5080-5089.
- [7] 李琰,吕南君,刘雪涛,等.考虑新能源消纳和网损的分布式光伏集群出力评估方法[J].电力建设,2022,43 (10):136-146.
LI Yan, LÜ Nanjun, LIU Xuetao, et al. Output evaluation method of distributed photovoltaic cluster considering renewable energy accommodation and power loss of network[J]. Electric Power Construction, 2022, 43(10): 136-146.
- [8] 李民,杨暑森,李科锋,等.覆雪状态下光伏发电功率预测方法研究[J].高压电器,2023,59 (9):250-257.
LI Min, YANG Shusen, LI Kefeng, et al. Research on power prediction method of photovoltaic power generation under snow coating conditions[J]. High Voltage Apparatus, 2023, 59 (9):250-257.
- [9] 王文森,贺馨仪,杨晓西,等.基于多参量数据回归分析的电力变压器状态监测方法[J].电网与清洁能源,2023, 39 (4):83-90.
WANG Wensen, HE Xinyi, YANG Xiaoxi, et al. Power transformer condition monitoring method based on multivariate statistical analysis[J]. Power System and Clean Energy, 2023, 39 (4):83-90.
- [10] YAZDANBAKSH O, KRAHN A, DICK S. Predicting solar power output using complex fuzzy logic[C]//Joint IFSA World Congress and NAFIPS Annual Meeting Edmonton: IFSA World Congress and NAFIPS Annual Meeting, Edmonton, AB, Canada, 2013.
- [11] 张惠娟,刘琪,岑泽尧,等.基于GWO-MLP的光伏系统输出功率短期预测模型[J].电测与仪表,2022,59 (7):72-77+113.
ZHANG Huijuan, LIU Qi, CEN Zeyao, et al. Short-term prediction model of output power of photovoltaic system based on GWO-MLP[J]. Electrical Measurement & Instrumentation, 2022, 59 (7):72-77+113.
- [12] 冯昌森,张瑜,文福拴,等.基于深度期望Q网络算法的微电网能量管理策略[J].电力系统自动化,2022,46 (3): 14-22.
FENG Changsen, ZHANG Yu, WEN Fushuan, et al. Energy management strategy for microgrid based on deep expected Q network algorithm[J]. Automation of Electric Power Systems, 2022, 46 (3):14-22.
- [13] 赵亮,刘友波,余莉娜,等.基于深度信念网络的光伏电站短期发电量预测[J].电力系统保护与控制,2019,47 (18):11-19.
ZHAO Liang, LIU Youbo, YU Lina, et al. Short-term power generation forecast of PV power station based on deep belief network[J]. Power System Protection and Control, 2019, 47(18):11-19.
- [14] 杨晶显,张帅,刘继春,等.基于VMD和双重注意力机制LSTM的短期光伏功率预测[J].电力系统自动化,2021, 45(3):174-182.
YANG Jingxian, ZHANG Shuai, LIU Jichun, et al. Short-term photovoltaic power prediction based on variational mode decomposition and long short-term memory with dual-stage attention mechanism[J]. Automation of Electric Power Systems, 2021, 45(3):174-182.
- [15] 殷豪,陈云龙,孟安波,等.基于二次自适应支持向量机的光伏输出功率预测[J].太阳能学报,2019,48(7):1866-1873.
YIN Hao, CHEN Yunlong, MENG Anbo, et al. Forecasting photovoltaic power based on quadric self-adaptive SVM model[J]. Acta Energetica Solaris Sinica, 2019, 48(7):1866-1873.
- [16] 李元诚,白恺,曲洪达,等.基于粒子群-稀疏贝叶斯混合算法的光伏功率预测方法[J].太阳能学报,2016,37 (5):1153-1159.
LI Yuancheng, BAI Kai, QU Hongda, et al. A photovoltaic power forecasting model using sparse bayesian regression optimized by particle swarm algorithm[J]. Acta Energetica Solaris Sinica, 2016, 37(5):1153-1159.
- [17] 叶林,陈政,赵永宁,等.基于遗传算法-模糊径向基神经网络的光伏发电功率预测模型[J].电力系统自动化, 2015, 39(16):16-22.
YE Lin, CHEN Zheng, ZHAO Yongning, et al. Photovoltaic power forecasting model based on genetic algorithm and fuzzy radial basis function neural network [J]. Automation of Electric Power Systems, 2015, 39(16): 16-22.
- [18] 苟超,陈伯建,吴翔宇,等.基于改进K-means算法的电力短期负荷预测方法研究[J].电力科学与技术学报,2022, 37(1):90-95.
XUN Chao, CHEN Bojian, WU Xiangyu, et al. Research on short-term power load forecasting method based on improved K-means algorithm[J]. Journal of Electric Power Science and Technology, 2022, 37(1):90-95.