

DOI:10.12158/j.2096-3203.2020.02.016

# 基于小波去噪-KPCA神经网络的光伏功率预测方法

孙新程<sup>1</sup>, 万玥<sup>1</sup>, 丁宏<sup>1</sup>, 葛晨阳<sup>2,3</sup>, 史文斌<sup>4</sup>

(1. 国网江苏省电力有限公司高邮供电分公司, 江苏 高邮 225600; 2. 国网江苏省电力有限公司连云港供电分公司, 江苏 连云港 222100; 3. 南京理工大学自动化学院, 江苏 南京 210094; 4. 上海绿色环保能源有限公司, 上海 202156)

**摘要:** 为了提高光伏发电功率的预测精度, 提出了一种基于小波去噪/聚类/核主成分分析(KPCA)神经网络的光伏发电功率预测方法。首先, 应用二维小波阈值去噪法预处理光伏出力数据; 然后, 应用  $k$ -means 算法将预测模型分为 4 种不同模式下的子预测模型; 引入 KPCA 对输入空间降维重构, 利用粒子群优化(PSO)神经网络算法建立基于聚类/KPCA/神经网络的光伏发电功率预测模型。采用某光伏电站的实例数据进行预测分析, 结果表明该模型实现了不同模式下的光伏出力较为精准的预测, 显示出良好的预测性能, 验证了预测模型的可行性和有效性。

**关键词:** 小波去噪; 核主成分分析; 粒子群算法; 神经网络; 光伏发电预测

**中图分类号:** TM715

**文献标志码:** A

**文章编号:** 2096-3203(2020)02-0110-07

## 0 引言

光伏发电受气象、环境等因素影响, 出力具有较强的波动性和随机性, 因此对光伏发电功率进行精准预测具有重要的现实意义。

光伏发电预测方法主要分为两类: 物理方法<sup>[1-2]</sup>和统计方法<sup>[3-7]</sup>。常用的统计方法包括神经网络<sup>[8-13]</sup>, 支持向量机(support vector machine, SVM)<sup>[14-18]</sup>, 马尔科夫链<sup>[19-20]</sup>等。光伏输出功率与天气条件密切相关, 文献[21]根据不同季节建立预测子模型, 但没有给出科学的分类方法。信道噪声等因素会降低数据质量, 影响预测拟合精度。文献[22-24]采用非线性小波去噪方法对数据进行去噪预处理, 然而一维小波去噪的方法具有局限性。光伏发电输出功率受复杂因素的影响, 输入空间维数高, 需要将输入空间降维解耦。主成分分析(principal component analysis, PCA)<sup>[25-29]</sup>是常见的降维算法, 然而这种线性降维方法是基于嵌入高维数据空间中的子空间是线性或近似线性的前提。另外传统的反向传播(back propagation, BP)神经网络也存在权重阈值的初始值是随机的, 容易陷入局部最小值的缺点, 需要对预测模型进行改进。

为解决上述问题, 提出了一种基于二维小波去噪方法,  $k$ -means/核主成分分析(kernel principal component analysis, KPCA)/粒子群优化神经网络

(particle swarm optimization back propagation neural network, PSOBP)的光伏输出功率预测方法。首先, 采用二维小波阈值去噪方法用于预处理历史输出数据以减少数据噪声; 采用  $k$ -means 算法对预测模型进行科学分类, 分为 4 种模式; 然后, 引入 KPCA 来减少输入的维数并将其转换为几个主成分, 简化预测模型; PSO 用于提高神经网络算法的性能。最后, 利用某光伏电站的实例数据验证了预测模型的可行性和有效性。

## 1 光伏发电功率的影响因素分析

光伏发电输出功率的工程模型如式(1)所示:

$$P = N\eta_1\eta_2\eta_3AR_\beta[1 - a(T_c - 25)] \quad (1)$$

式中:  $N$  为光伏组件个数;  $\eta_1$  为光电转换效率, %;  $\eta_2$  为最大功率跟踪的工作效率, %;  $\eta_3$  为逆变器效率, %;  $A$  为光伏组件面积,  $\text{m}^2$ ;  $\beta$  为倾角;  $R_\beta$  为斜面总辐射,  $\text{W}/\text{m}^2$ ;  $a$  为温度系数;  $T_c$  为光伏组件的板温度,  $^\circ\text{C}$ 。

从式(1)可以看出, 光伏输出功率主要与电气效率、电站设计、气象等因素有关。电气效率、电站设计在预测周期内都是固定不变的, 因此在讨论光伏发电功率的影响因素时, 主要是考虑气象因素。这些气象因子(辐照度、温度、湿度、风速等)对光伏发电功率的影响程度各不相同, 而且各因子之间还存在着多重的关联和耦合。为了确定预测模型的输入向量, 需要对光伏发电功率与气象影响因子之间的关系进行定性和定量分析。文中采用 Pearson 相关系数法对关联度进行定量分析。

Pearson 相关系数法计算公式如式(2)所示:

收稿日期:2019-09-06; 修回日期:2019-10-20

基金项目:江苏省重点研发计划资助项目“源网荷储一体化级联多端口变换器多目标解耦与优化控制关键技术”(BE2-019018)

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

式中:  $R$  为变量  $x$  与  $y$  的相关系数;  $\bar{x}$ ,  $\bar{y}$  分别为变量  $x$ ,  $y$  的平均数。

Pearson 相关系数是用来反映 2 个变量线性相关程度的统计量,  $R$  的绝对值大小与变量相关性成正比, 且正数为正相关, 负数为负相关。

光伏发电功率与辐照度、温度、湿度和风速的实测值归一化后的曲线如图 1 所示。

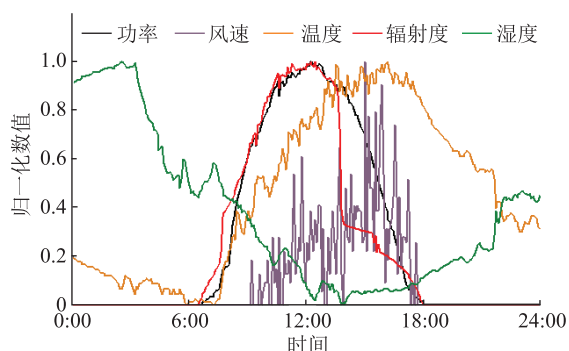


图 1 发电功率与影响因子归一化曲线

Fig.1 Curves of power generation and impact factor normalization

通过曲线图可以直观地看出, 发电功率与辐照度、温度等因素之间存在很强的相关性, 利用 Pearson 相关系数法计算发电功率与辐照度的相关系数, 具体如表 1 所示。

表 1 光伏发电功率与影响因子的相关系数

Table 1 Correlation coefficient between photovoltaic power generation and impact factor

参数	相关系数	参数	相关系数
辐照度	0.940	湿度	-0.598
温度	0.549	风速	0.311

通过图 1 和表 1 对影响因素的定性与定量分析, 可以得出辐照度是影响光伏发电的最主要因素, 相关性最强, 呈正相关; 温度与湿度也是光伏发电功率的重要影响因素, 其中温度呈正相关, 湿度呈负相关; 风速的相关性程度低、波动过大。因此在确定预测模型的输入向量时, 主要考虑辐照度、温度和湿度这 3 种影响因子。

## 2 光伏功率预测模型

### 2.1 小波去噪数据预处理

光伏发电曲线呈现很强的随机性和波动性, 同时由于信道传输等影响, 历史数据曲线会产生突变

或者部分缺失。将此类数据作为历史数据进行光伏发电功率预测工作, 会影响光伏发电预测的精准度, 因此需要对历史数据的噪声进行去噪预处理, 提高数据质量。常用的去噪算法是在一维度中进行去噪处理, 具有一定的局限性; 不同天气条件下的光伏发电数据特性也完全不一样。因此, 在不同天气条件下, 采用二维小波阈值对数据进行去噪处理。

传统的阈值函数包括硬阈值和软阈值, 都存在不同程度的存在不连续或者导数不连续等缺陷。因此文中的阈值函数选取阈值函数如下:

$$\tilde{\omega}_{j,k} = \begin{cases} (1 - \mu) \omega_{j,k} + \mu \text{sgn}(\omega_{j,k}) (|\omega_{j,k}| - T) & |\omega_{j,k}| \geq T \\ 0 & |\omega_{j,k}| < T \end{cases} \quad (3)$$

式中:  $\omega_{j,k}$  为小波系数;  $\tilde{\omega}_{j,k}$  为阈值处理后的小波系数;  $\mu$  为加权因子。

$$\mu = \frac{T}{|\omega_{j,k}|} e\left(-\frac{|\omega_{j,k}| - T}{|\omega_{j,k}|}\right) \quad (4)$$

二维小波阈值去噪的具体步骤如下:

步骤一: 在不同天气条件下, 选取  $n$  天  $m$  个时间点的数据样本, 形成  $n \times m$  矩阵的二维数据集;

步骤二: 对二维数据矩阵归一化处理;

步骤三: 对原始含噪数据进行小波分解, 得到小波系数  $\omega_{j,k}$ ;

步骤四: 对小波系数进行阈值处理得到处理后的小波系数  $\tilde{\omega}_{j,k}$ , 并利用  $\tilde{\omega}_{j,k}$  进行小波重构;

步骤五: 对二维数据矩阵反归一化处理, 输出去噪后的数据集。

### 2.2 k-means 算法

光伏发电功率预测精度与天气状态密切相关。不同天气状态下发电功率的变化规律有着明显的区别。根据“相似原则”仅采用单一天气条件下的预测模型对不同天气类型进行预测, 降低了整体的预测效果。

直接采用气象部门提供的天气类型进行分类, 分类方法简单易行。但是由于发电功率不完全受天气条件影响, 相同天气下可能出现发电曲线不相似, 或者不同天气下存在曲线相似的情况。

因此利用  $k$ -means 算法的统计学方法对光伏发电的历史数据集分类, 将原有的预测模型分成 4 种广义天气条件(晴, 阴, 雨, 大雨)下的 4 个预测子模型, 即 Model I, Model II, Model III 和 Model IV。

采用每天光伏出力数据的特征参数信息作为输入量, 可以简化分类模型的复杂度, 输入向量由 5

个因子组成:

$$I = [P, T_{OD}, L_s, T, n] \quad (5)$$

式中:  $P$  为光伏出力平均值, kW;  $T_{OD}$  为光伏出力差三阶导数最大值;  $L_s$  为光伏出力归一化离散差;  $T$  为环境温度平均值, °C;  $n$  为积日。

### 2.3 KPCA-神经网络预测模型

#### 2.3.1 KPCA 降维重构

光伏发电功率受辐照度、温度、湿度等因素的影响,影响因子复杂耦合。同时根据“邻近原则”,与待预测日日期邻近的历史数据被选择作为模型的输入。这些复杂数据若不经处理直接作为模型输入向量,会使输入空间的维数过高、模型结构趋于复杂。目前常用的降维算法是 PCA,算法从原始输入空间中提取线性独立的输入变量,有效地简化了模型的结构。然而,它是基于嵌入高维数据空间中的子空间是线性或近似线性的前提,因此文中将非线性降维算法 KPCA 引入到模型中,实现输入数据的降维重构,达到简化模型的目的。

基于核函数的 KPCA 特征提取算法的原理是先通过非线性映射  $\varphi$  将原始低维空间  $d$  样本映射到高维线性空间  $D$  中,然后再在该特征空间中进行 PCA。

通常  $\varphi$  是隐形的,直接求解  $D$  维空间协方差矩阵  $C$  的特征值与特征向量很困难,可设  $\nu$  是  $C$  的特征向量,  $\lambda$  是对应的特征值,则有:

$$\nu = \frac{1}{N} \sum_{i=1}^N (\psi(X_i)^T \nu) \psi(X_i) \quad (6)$$

$$\psi(X_i) = \varphi(X_i) - \frac{1}{N} \sum_{j=1}^N \varphi(X_j) \quad (7)$$

由于  $C\nu = \lambda\nu$ ,进一步化简可得:

$$\lambda Na = \tilde{K}Na \quad (8)$$

$$\tilde{K} = \psi(X_i)\psi(X_j)^T \quad (9)$$

$$a = \frac{1}{N\lambda} (\psi(X_i)^T \nu) \quad (10)$$

$$\tilde{K} = K - I_N K - K I_N - I_N K I_N \quad (11)$$

式中:  $I_N = 1/N$ ,  $K$  即为核矩阵。求出  $\tilde{K}$  后,特征值  $\lambda$  与特征向量也可相继算出。主成分贡献率即为:

$$\lambda_i / \sum_{i=1}^D \lambda_i, \text{ 累计主成分贡献率为 } \sum_{i=1}^k \lambda_i / \sum_{i=1}^D \lambda_i。$$

当累计主成分贡献率达到某一限定值时,即可用  $k$  维主成分代替原有的样本集,在保证有效信息的基础上模型得到了简化,提高了模型的学习能力和泛化能力。

#### 2.3.2 PSO 算法优化神经网络

通常 BP 神经网络的权值和阈值的初始值是随

机选取的,预测模型易陷入局部极值。PSO 算法是依据适应度不断更新粒子的位置和速度以达到最优的一种具有全局搜索能力的算法。

PSO 算法优化神经网络的步骤如下:

步骤一:将神经网络的权值和阈值作为粒子的初始位置;

步骤二:将神经网络的训练误差作为 PSO 算法的适应度;

步骤三:更新粒子的速度和位置直至迭代结束,寻找粒子的个体最优值和全局最优值;

步骤四:最后将经过 PSO 算法计算后得到的全局最优值作为神经网络算法的初始权值和阈值。

引入 PSO 算法可以优化初始值的选取,降低传统神经网络陷入局部最优的可能性,提高收敛速度和精度。

### 2.4 光伏发电功率预测流程

光伏发电功率预测方法是首先在 SCADA 系统中完成光伏发电量、气象等数据的收集;然后采用二维小波去噪方法对收集的数据进行预处理;利用  $k$ -means 方法将原有的预测模型分成 4 个预测子模型;利用 KPCA 对预测子模型的输入量降维重构;最后输入 PSOBP 的预测算法,实现光伏发电功率预测。光伏发电功率预测的流程如图 2 所示。

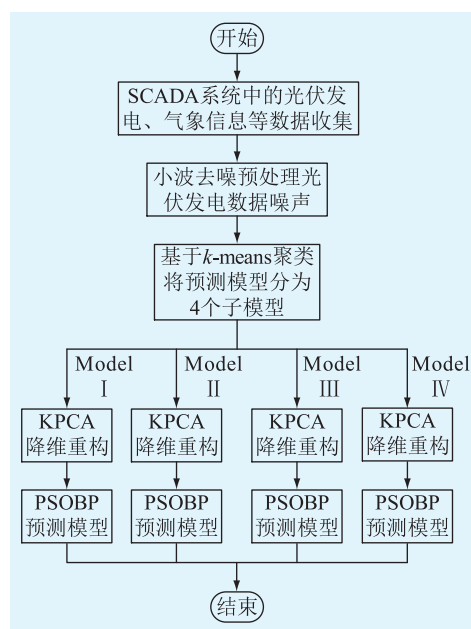


图 2 光伏发电功率预测的流程

Fig.2 Flow chart of photovoltaic power forecasting

## 3 光伏预测实例

### 3.1 实验描述

文中选择某光伏电站 2016 年 8 月 1 日至 2017 年 8 月 1 日每天 5 min 的电气量数据和气象数据作

为数据样本,共计 365 天。预测的评价指标主要有:相关系数  $R$ , 平均绝对百分误差  $e_{\text{mape}}$  和均方根误差  $e_{\text{rmse}}$ 。

$R$  由式(2)确定,用于反映预测值与实际值之间的相关性。 $e_{\text{mape}}$  由式(12)确定,用于反映模型的全局预测能力。

$$e_{\text{mape}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{P_i - \tilde{P}_i}{P_i} \right| \times 100\% \quad (12)$$

$e_{\text{rmse}}$  由式(13)确定,用于反映预测值和实际值之间的离散性。

$$e_{\text{rmse}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - \tilde{P}_i)^2} \quad (13)$$

式中:  $P_i$  为实际值;  $\tilde{P}_i$  为预测值;  $n=156$ 。

### 3.2 实验流程

步骤一:首先利用二维小波去噪方法对 SCADA 系统中的历史数据进行预处理,去除数据噪声。图 3 为某日的光伏发电功率曲线和去噪后的功率曲线。由图 3 可以看出,原曲线中含有曲线异常点,经过二维小波去噪处理后异常点去除,曲线数据更加科学。

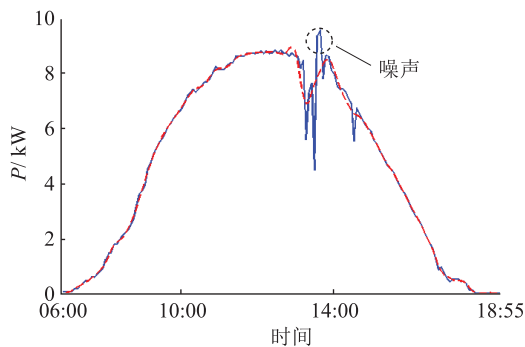


图 3 去噪后的光伏发电功率序列

Fig.3 Power diagram of Photovoltaic after denoising

步骤二:利用  $k$ -means 算法对光伏发电功率的历史数据分类。

步骤三:确定模型的输入量,并利用 KPCA 实现输入空间的降维重构。

输入量包括前一天 6:00 到 18:55 每 5 min 的光伏出力,预测日的温度、湿度、日期和天气类型量化值。引入 KPCA 对输入空间降维,本预测实例选用的核函数为高斯核函数。经过 KPCA 得出的主成分贡献率以及累计贡献率如表 2 所示。

可以看出当提取输入变量的维数为 8 时,累计贡献率已经达到 95%,可以代替原始的高维输入空间,在保证有效信息的前提下输入空间实现了降维重构,网络模型结构得到了简化。

步骤四:将重构后的输入空间作为 PSO 算法优

表 2 主成分贡献率以及累计贡献率

Table 2 Principal component contribution rate and accumulated contribution rate

编号	特征值	贡献率/%	累计贡献率/%
1	3.210 4	53.502 4	53.502 4
2	1.854 5	22.564 6	76.067 0
...	...	...	...
8	0.436 8	4.933 9	95.056 2

化神经网络模型的输入并进行训练,得到光伏发电预测模型。

### 3.3 结果分析

为了分析和比较预测效果,文中对提出的预测方法与 BP、PSOBP、KPCA-PSOBP 以及是否采用去噪处理进行了比较。4 种不同模型的结果和 Model I 模型下不同方法的预测结果见表 3、表 4 和图 4。

表 3 文中预测模型精度分析

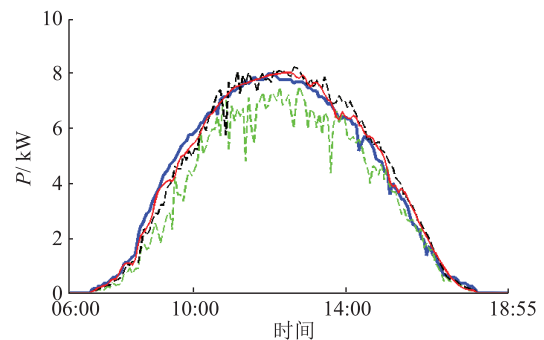
Table 3 Precision analysis of prediction models

参数	Model I	Model II	Model III	Model IV
$R$	0.979 0	0.961 2	0.898 9	0.764 5
$e_{\text{rmse}}/\text{kW}$	0.866 0	0.880 3	0.796 1	0.647 7
$e_{\text{mape}}/\%$	0.067 5	0.646 3	0.857 9	0.377 7

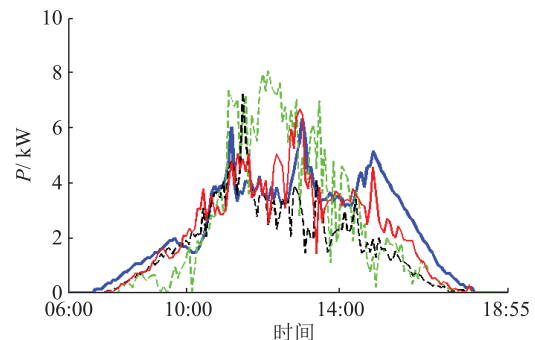
表 4 Model I 模型下不同方法的精度与时间分析

Table 4 Accuracy and time analysis of different methods under Model I

参数	BP	PSOBP	KPCA-PSOBP	去噪-KPCA-PSOBP
$e_{\text{mape}}/\%$	0.677 4	0.342 6	0.168 0	0.067 5
$t/\text{s}$	8.231 0	8.412 0	2.423 0	3.121 0



(a) Model I



(b) Model II

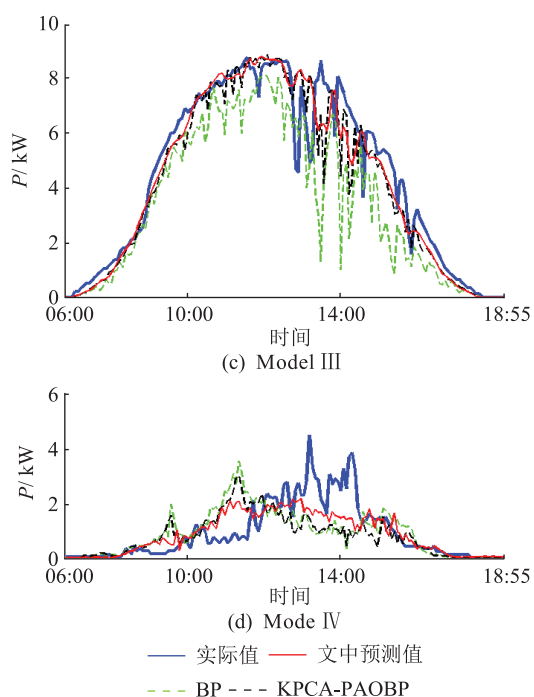


图4 预测结果

Fig.4 Prediction results

通过对表3和图4中结果的分析,文中的预测方法可以反映实际的PV输出,证明了该模型的可行性。其中Model I的预测效果最好, $e_{\text{mapc}}$ 指标达到了0.0675,同时Model IV的 $e_{\text{mapc}}$ 指标也达到了0.3777。证明在复杂气象条件下,预测精度仍能达到较好的标准。

通过对图4中的不同方法的预测曲线与实际值的拟合度,与表4中选取Model I模型下不同方法的精度与时间,BP与PSOBP的预测结果分析得出PSO优化BP提高了算法的预测精度;PSOBP与KPCA-PSOBP的预测结果分析KPCA消除了冗余并简化了预测模型,使得算法的运算时间得到了显著提高,同时精度也得到了改善。此外,基于小波去噪方法,数据预处理提高了数据质量,预测模型的运算速度有所下降,但预测精度 $e_{\text{mapc}}$ 从0.168%下降到0.0675%,证明了文中所提模型的有效性。

#### 4 结语

文中建立了基于小波去噪/KPCA/PSOBP的光伏输出功率预测方法。 $k$ -means算法提高了模型的泛化能力;二维小波阈值去噪方法降低了数据噪声;KPCA减少了输入的维数并简化了模型;PSO改善了神经网络的性能。该预测模型可行、高效,对光伏输出功率预测具有一定的实用价值。

#### 参考文献:

[1] 陈昌松,段善旭,殷进军. 基于神经网络的光伏阵列发电预

测模型的设计[J]. 电工技术学报,2009,24(9):153-158.

CHEN Changsong,DUAN Shanxu,YIN Jinjun. Design of photovoltaic array power forecasting model based on neural network [J]. Transactions of China Electrotechnical Society, 2009, 24 (9): 153-158.

[2] 王守相,张娜. 基于灰色神经网络组合模型的光伏短期出力预测[J]. 电力系统自动化,2012,36(19):37-41.

WANG Shouxiang,ZHANG Na. Short-term output power forecast of photovoltaic based on a grey and neural network hybrid model[J]. Automation of Electric Power Systems, 2012, 36 (19): 37-41.

[3] 李乐,刘天琪. 基于近邻传播聚类 and 回声状态网络的光伏预测[J]. 电力自动化设备,2016,36(7):41-46.

LI Le,LIU Tianqi. PV power forecasting based on AP-ESN[J]. Electric Power Automation Equipment,2016,36(7):41-46.

[4] 朱想,居蓉蓉,程序,等. 组合数值天气预报与地基云图的光伏超短期功率预测模型[J]. 电力系统自动化,2015,39(6):4-10.

ZHU Xiang,JU Rongrong,CHENG Xu. A very short-term prediction model for photovoltaic power based on numerical weather prediction and ground-based cloud images[J]. Automation of Electric Power Systems,2015,39(6):4-10.

[5] LARSON D P,NONNENMACHER L,COINBRA C F M. Day-ahead forecasting of solar power output from photovoltaic plants in the American southwest[J]. Renewable Energy, 2016, 91: 11-20.

[6] 王昕,黄柯,郑益慧,等. 基于PNN/PCA/SS-SVR的光伏发电功率短期预测方法[J]. 电力系统自动化,2017,41(2):455-461.

WANG Xin,HUANG Ke,ZHENG Yihui, et al. Combined PV power forecast based on firefly algorithm generalized regression neural network [J]. Power System Technology, 2017, 41 (2): 455-461.

[7] 赖昌伟,黎静华,陈博,等. 光伏发电出力预测技术研究综述[J]. 电工技术学报,2019,34(6):1201-1217.

LAI Changwei,LI Jinghua,CHEN Bo, et al. Review of photovoltaic power output prediction technology [J]. Transactions of China Electrotechnical Society,2019,34(6):1201-1217.

[8] 朱红路,李旭,姚建曦,等. 基于小波分析与神经网络的光伏电站功率预测方法[J]. 太阳能学报,2015,36(11):2725-2730.

ZHU Honglu,LI Xu,YAO Jianxi, et al. The power prediction method for photovoltaic power station based on wavelet neural networks[J]. Acta Energeiae Solaris Sinica, 2015, 36 (11): 2725-2730.

[9] 郝思鹏,吴清,李佳伟,等. 基于人工蜂群算法优化RBF神经网络光伏组件故障诊断定位方法[J]. 供用电,2019,36(10):87-92.

HAO Sipeng,WU Qing,LI Jiawei, et al. Optimizing RBF neural network fault diagnosis and location method for PV module based on ABC algorithm[J]. Distribution & Utilization,2019, 36(10):87-92.

- [10] 李德智, 张文文, 嵇建飞. 基于果蝇优化模糊神经网络模型的售电市场评估[J]. 电力需求侧管理, 2019, 21(1): 27-31.  
LI Dezhi, ZHANG Wenwen, JI Jianfei. Electricity retail market evaluation based on fruit fly optimized fuzzy neural network model[J]. Power Demand Side Management, 2019, 21(1): 27-31.
- [11] TAO Y, CHEN Y. Distributed PV power forecasting using genetic algorithm based neural network approach[C]// 2014 International Conference on Advanced Mechatronic Systems, 2014, Tokai, Japan.
- [12] 王晓兰, 葛鹏江. 基于相似日和径向基函数神经网络的光伏阵列输出功率预测[J]. 电力自动化设备, 2013, 33(1): 100-103.  
WANG Xiaolan, GE Pengjiang. PV array output power forecasting based on similar day and RBFNN[J]. Electric Power Automation Equipment, 2013, 33(1): 100-103.
- [13] 丁明, 刘志, 毕锐, 等. 基于灰色系统校正-小波神经网络的光伏功率预测[J]. 电网技术, 2015, 39(9): 2438-2443.  
DING Ming, LIU Zhi, BI Rui, et al. Photovoltaic output prediction based on grey system correction-wavelet neural network [J]. Power System Technology, 2015, 39(9): 2438-2443.
- [14] 蒋浩然, 刘小聪, 李扬. 基于多变量灰色预测模型的省级电力公司中长期电量预测[J]. 电力需求侧管理, 2019, 21(1): 37-41.  
JIANG Haoran, LIU Xiaocong, LI Yang. Mid and long term electricity consumption forecasting of provincial power companies based on multi-variable gray [J]. Power Demand Side Management, 2019, 21(1): 37-41.
- [15] 朱永强, 田军. 最小二乘支持向量机在光伏功率预测中的应用[J]. 电网技术, 2011, 35(7): 54-59.  
ZHU Yongqiang, TIAN Jun. Application of least square support vector machine in photovoltaic power forecasting [J]. Power System Technology, 2011, 35(7): 54-59.
- [16] 黄磊, 舒杰, 姜桂秀, 等. 基于多维时间序列局部支持向量回归的微网光伏发电预测[J]. 电力系统自动化, 2014, 38(5): 19-24.  
HUANG Lei, SHU Jie, JIANG Guixiu, et al. Photovoltaic generation forecast based on multidimensional time-series and local support vector regression in microgrids [J]. Automation of Electric Power Systems, 2014, 38(5): 19-24.
- [17] 陈道君, 龚庆武, 金朝意, 等. 基于自适应扰动量子粒子群算法参数优化的支持向量回归机短期风电功率预测[J]. 电网技术, 2013, 37(4): 974-980.  
CHEN Daojun, GONG Qingwu, JIN Zhaoyi, et al. Short-term wind power prediction based on support vector regression machine optimized by adaptive disturbance quantum-behaved particle swarm optimization [J]. Power System Technology, 2013, 37(4): 974-980.
- [18] 茆美琴, 龚文剑, 张榴晨, 等. 基于 EEMD-SVM 方法的光伏电站短期出力预测[J]. 中国电机工程学报, 2013, 33(34): 17-24.  
MAN Meiqin, GONG Wenjian, ZHANG Liuchen, et al. Short-term photovoltaic generation forecasting based on EEMD-SVM combined method [J]. Proceedings of the CSEE, 2013, 33(34): 17-24.
- [19] 丁明, 徐宁舟. 基于马尔可夫链的光伏发电系统输出功率短期预测方法[J]. 电网技术, 2011, 35(1): 152-157.  
DING Ming, XU Ningzhou. A method to forecast short-term output power of photovoltaic generation system based on Markov chain [J]. Power System Technology, 2011, 35(1): 152-157.
- [20] 李英姿, 何琳, 牛进苍. 基于马尔可夫链的光伏并网发电量预测[J]. 太阳能学报, 2014, 35(4): 611-616.  
LI Yingzi, HE Lin, NIU Jincang. Forecasting power generation of grid-connected solar PV system based on Markov chain [J]. Acta Energetica Solaris Sinica, 2014, 35(4): 611-616.
- [21] 袁晓玲, 施俊华, 徐杰彦. 计及天气类型指数的光伏发电短期出力预测[J]. 中国电机工程学报, 2013, 33(34): 57-63.  
YUAN Xiaoling, SHI Junhua, XU Jieyan. Short-term power forecasting for photovoltaic generation considering weather type index [J]. Proceedings of the CSEE, 2013, 33(34): 57-63.
- [22] 王育飞, 付玉超, 孙路, 等. 基于混沌-RBF神经网络的光伏发电功率超短期预测模型[J]. 电网技术, 2018, 42(4): 1110-1116.  
WANG Yufei, FU Yuchao, SUN Lu. Ultra-short prediction model of photovoltaic output power based on Chaos-RBF neural network [J]. Power System Technology, 2018, 42(4): 1110-1116.
- [23] 徐中一, 刘远龙. 基于 RBF 神经网络的配电变压器不良数据辨识[J]. 供用电, 2018, 35(6): 41-46.  
XU Zhongyi, LIU Yuanlong. The bad data identification of distribution transformer based on RBF neural network [J]. Distribution & Utilization, 2018, 35(6): 41-46.
- [24] 史静, 李琥, 李冰洁, 等. 基于长短期记忆神经网络的配电网负荷预测方法研究[J]. 供用电, 2019, 36(7): 71-77, 90.  
SHI Jing, LI Hu, LI Bingjie, et al. Distribution network load forecasting based on LSTM neural network [J]. Distribution & Utilization, 2019, 36(7): 71-77, 90.
- [25] 郑凌蔚, 刘士荣, 毛军科, 等. 一种综合 ANFIS 和 PCA 的光伏发电功率预测新方法[J]. 太阳能学报, 2012, 33(6): 993-1001.  
ZHENG Lingwei, LIU Shirong, MAO Junke, et al. A new PV generation power forecasting method integrating ANFIS and PCA [J]. Acta Energetica Solaris Sinica, 2012, 33(6): 993-1001.
- [26] 廖卫强, 张认成, 俞万能, 等. 基于相似样本及 PCA 的光伏输出功率预测[J]. 太阳能学报, 2016, 37(9): 2377-2385.  
LIAO Weiqing, ZHANG Rencheng, YU Wanneng. Prediction of output of photovoltaic based on similar samples and principal component analysis [J]. Acta Energetica Solaris Sinica, 2016, 37(9): 2377-2385.
- [27] 李志海, 张春平, 王子壬, 等. 结合 PCA 的 K-means 算法在专变用户用电行为分析中的应用[J]. 电力信息与通信技术, 2018, 16(12): 62-67.

LI Zhihai,ZHANG Chunping,WANG Ziren, et al. Application of K-means algorithm combined with PCA in behavior analysis of private transformer customer[J]. Electric Power Information and Communication Technology,2018,16(12):62-67.

[28] 谢晓帆,刘秋林,李斌,等. 基于主成分分析法与对应分析法的县域配电网状况评估[J]. 智慧电力,2018,46(6):68-73.

XIE Xiaofan,LIU Qiulin,LI Bin, et al. Evaluation of country distribution network status based on principal component analysis and correspondence analysis[J]. Smart Power,2018,46(6):68-73.

[29] 陈勇,李鹏,张忠军,等. 基于 PCA-GA-LSSVM 的输电线路覆冰负荷在线预测模型[J]. 电力系统保护与控制,2019,47(10):110-119.

CHEN Yong,LI Peng,ZHANG Zhongjun,et al. Online prediction model for power transmission line icing load based on PCA-GA-LSSVM[J]. Power System Protection and Control,2019,47(10):110-119.

作者简介:



孙新程

孙新程(1993),男,硕士,工程师,从事智能电网与人工智能算法相关工作(E-mail:sunxc24@163.com);

万玥(1992),女,硕士,工程师,从事新能源发电与配网自动化相关工作;

丁宏(1983),男,学士,工程师,从事变电运维与智能变电站相关工作。

### Forecasting method of photovoltaic output power based on wavelet denoising/KPCA/PSO-BP

SUN Xincheng<sup>1</sup>, WAN Yue<sup>1</sup>, DING Hong<sup>1</sup>, GE Chenyang<sup>2,3</sup>, SHI Wenbin<sup>4</sup>

(1. State Grid Gaoyou Power Supply Company of Jiangsu Electric Power Co.,Ltd., Gaoyou 225600, China;

2. State Grid Lianyungang Power Supply Company of Jiangsu Electric Power Co.,Ltd., Lianyungang 222100, China;

3. School of Automation, Nanjing University of Technology, Nanjing 210094, China;

4. Shanghai Green Environmental Energy Co., Ltd., Shanghai 202156, China)

**Abstract:** In order to improve the forecasting accuracy of photovoltaic (PV) output power, a forecasting method is proposed based on k-means, KPCA and PSO-BP. Firstly, wavelet threshold de-noising algorithm is used to pretreat PV output data. Then, the k-means clustering algorithm is applied to divide the forecasting model into four sub models under different modes. The kernel principal component analysis (KPCA) method is used to reduce the dimensionality of the input space. Neural network algorithm is optimized based on particle swarm optimization (PSO). Finally, the PV output power forecasting model based on k-means /KPCA/PSO-BP is established. The example data is used to verify the forecasting model, the results show that it can forecast the PV output power accurately in different modes and have good forecasting performance.

**Keywords:** wavelet denoising; kernel principal component analysis; particle swarm optimization; neural network; photovoltaic forecast

(编辑 钱悦)