

基于样本双重筛选的光伏发电功率预测

冬 雷¹, 周 晓¹, 郝 颖¹, 廖晓钟¹, 高 阳²

(1. 北京理工大学自动化学院, 北京 100081; 2. 沈阳工程学院电力学院, 沈阳 110136)

摘 要: 首先通过建立辐照度和功率之间的关系,剔除功率奇异值来完成初步筛选,其次通过计算辐照度、温度、云量等气象因素与发电功率的相关系数,选取相关系数较大的气象因素来评价历史天与预测天的相似度,提取与预测天最相似的历史天作为训练样本来完成二次筛选,最后利用BP神经网络和遗传算法进行光伏发电的功率预测,结果表明该方法具有较高的预测精度。

关键词: 光伏发电预测; BP神经网络; 双重筛选; 相似日; 遗传算法

中图分类号: TK514

文献标识码: A

0 引 言

太阳能作为一种清洁、环保、可再生的新能源,已成为对应能源短缺、气候变化与节能减排的重要选择之一^[1]。但太阳能光伏发电受到太阳辐照度、温度等气象因素的影响,具有较强的随机性与间歇性^[2],因此光伏发电系统相对于大电网是一个不可控源,进而对电网的调峰、调频、母线电压等造成不利影响^[3],这正是限制光伏发电大规模应用的难点之一。随着光伏产业的迅猛发展和装机容量的增长,为保证电力系统经济、安全和可靠运行,光伏发电的预测显得越来越重要^[4]。目前光伏发电功率的预测方法可分为两类——直接预测法和间接预测法^[5]。直接预测法是根据历史光伏发电功率数据及历史天气信息直接预测功率;间接预测法是利用历史天气信息预测辐照度,再利用辐照度和光伏发电功率之间的关系得到功率值。现在对光伏发电功率的预测方法有很多种,如神经网络法^[6-10]、自回归滑动平均模型法^[11]、马尔科夫链法^[12,13]、卡尔曼预测法^[14]、支持向量机法^[15-17]和混合智能算法^[18-20]等。

本研究以新疆维吾尔自治区奇台县石城子光伏电站为研究对象,首先将数据分为春、夏、秋、冬4个季节,每个季节根据天气情况分为晴、多云和雨雪3种天气类型,通过剔除奇异值和选取相似

日双重筛选的方法选出与预测天最相似的20个历史天作为BP神经网络的训练数据,采用遗传算法(genetic algorithms, GA)寻优神经网络的权值,最后利用GA-BP神经网络得到预测日的功率。

1 神经网络和遗传算法

1.1 神经网络的基本原理

BP(Back Propagation)神经网络又称误差反向传播神经网络,是一种按误差逆向传播算法训练的多层前馈网络,是目前应用最广泛和成功的神经网络模型之一。BP神经网络结构包括输入层、隐含层和输出层,其基本思想包括信号的正向传播和误差的反向传播两个过程。正向传播时输入样本从输入层传入,经隐含层处理后传向输出层,若输出层的实际输出与期望输出不同,则转向误差的反向传播阶段。误差的反向传播是将输出误差通过隐含层向输入层反传,反传过程中依据误差信号修正各单元的权值。这两个过程周而复始的进行,权值也不断调整,一直到输出的误差减少到可以接受的程度或进行到预先设定的学习次数为止^[21]。

1.2 遗传算法的基本原理

遗传算法(genetic algorithms, GA)是一类借鉴生物界自然选择和自然遗传机制的随机化搜索算

收稿日期: 2016-02-16

通信作者: 冬 雷(1967—),男,博士、副教授,主要从事电力电子与电力传动、风力发电与太阳能发电系统、微电网技术等方面的研究。
correspondent_dong@163.com

法。遗传算法以编码空间代替问题的参数空间,以适应度函数为评价依据,以编码群体为进化为基础,以群体中个体位串的遗传操作实现选择和遗传机制,建立一个迭代过程。在这一过程中,通过随机重组编码位串中重要的基因,使新一代的位串集合优于老一代的位串集合,群体的个体不断进化,逐渐接近最优解,最终达到求解问题的目的^[22]。

遗传算法运行的基本步骤如下:

- 1) 选择编码策略,把参数集合 X 和域转换为位串空间 S 。
- 2) 定义适应度函数 $f(x)$ 。
- 3) 确定遗传策略,包括选择群体大小 n ,选择、杂交、变异方法,以及确定杂交概率 P_c 、变异概率 P_m 等遗传参数。
- 4) 随机初始化生成群体 P 。
- 5) 计算个体适应度 $f(x)$ 。
- 6) 按照遗传策略,运用选择、杂交和变异算子作用于群体得到下一代群体。
- 7) 判断终止条件,若满足则进化过程中具有最大适应度个体作为最优解,若不满足终止条件则转到 5)。

1.3 GA-BP 神经网络

遗传算法具有良好的全局搜索能力,可以快速地将解空间中的全体解搜索出来,而不会陷入局部最优解的快速下降陷阱,并且利用它的内在并行性,可以方便地进行分布式计算,加快求解速度。BP 神经网络在训练权值时可能陷入局部最优,因此本文采用遗传算法训练神经网络的权值。

2 光伏发电功率的预测模型

光伏发电功率受很多因素影响,如太阳辐照度、太阳入射角度、光伏阵列的安装角度、转换效率、气压、云量、温度等。对于一个既定的光伏发电系统,输出功率的时间序列具有自相关性,光伏发电输出功率的历史数据来自于同一套发电系统,数据本身就包含了光伏阵列的地理位置、安装角度等对功率的影响^[7],因此选择历史数据训练神经网络来预测光伏发电功率。

图 1 是相同天气类型、不同季节类型条件下石城子 20 MW 光伏发电的功率曲线,由于季节性的太

阳辐照度差异导致光伏发电功率的季节性变化,在相同的天气类型条件下,功率曲线具有相同的趋势,但是不同季节类型也会影响功率大小。图 2 是相同季节、不同天气类型下的光伏发电的功率曲线,可以看出受辐照度、温度、云量等气象因素的影响,天气类型不同时光伏发电功率差异很大。因此,可以先按照季节对历史数据进行分类,然后每个季节再按照天气类型分为晴、多云和雨雪。

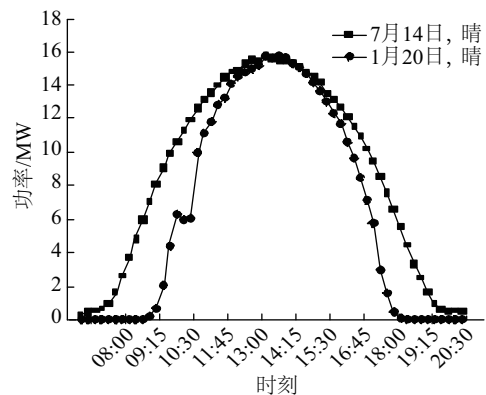


图1 冬季和夏季条件下的光伏发电功率对比图

Fig. 1 Comparison of photovoltaic power generation under winter and summer

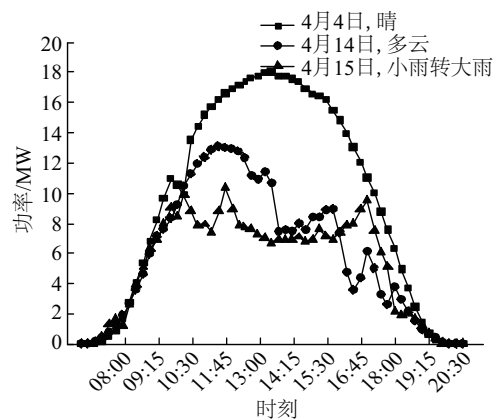


图2 晴天、多云和雨天条件下的光伏发电功率对比图

Fig. 2 Comparison of photovoltaic power generation under clear weather, cloudy weather and rainy days

由于采集误差,光伏电站采集的功率值可能与实际值相差很大,然而神经网络模型的准确建立很大程度上依赖于训练样本,因此为了得到最佳神经网络模型,本文采用双重筛选训练样本的方法。初步筛选——通过建立功率和辐照度的关系式来验证功率是否为奇异值并对其进行删除。二次筛选——根据各种气象因素选出与预测天最相似的

历史天做训练样本。

由于 BP 神经网络在预测时易陷入局部最优,而遗传算法具有全局搜索最优解的特点,因此本文采用遗传算法来优化神经网络的权值,从而得到基于样本双重筛选的遗传算法-神经网络光伏发电功率预测模型。预测模型如图 3 所示,先按照季节对历史数据进行分类,再将其按照天气分为晴、多云和雨雪 3 种类型,训练样本时在对应季节、对应天气类型的历史数据中经过双重筛选选出与预测天最相近的 20 个历史天作为训练数据,最后采用 GA-BP 神经网络预测得到预测日的功率数据。

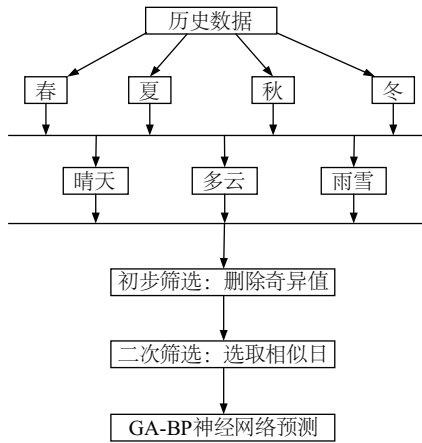


图3 光伏发电功率预测流程图

Fig. 3 Flow chart of photovoltaic power generation forecast

3 预测步骤

3.1 数据来源及处理

本研究以新疆维吾尔自治区奇台县石城子光伏电站为研究对象,功率值由发电站直接提供。相应的各种气象数据来源于欧洲中期天气预报中心(ECMWF),包括地面辐照度、温度、全云量、低云量、风速和湿度,其中表面辐照度是累积值,其他因素为瞬时值,时间间隔为 3 h。

若无大气层和云层的阻挡,地面接收到的辐射将接近于地外辐射,地外辐射可通过地球与太阳的相对运动关系得到。因此假定在短时间内地外辐射和地面辐射成比例且比例不变,根据得到的地外辐射趋势和下载的地面辐照度的累积值可得到地面辐射的瞬时值。任何地点、任何时刻的地外辐射度可通过式(1)~式(7)得到:

$$I = I_0 \cos \theta \quad (1)$$

$$I_0 = I_{sc} (1 + 0.033 \cos \frac{2\pi N}{365}) \quad (2)$$

$$\cos \theta = \sin \varphi \sin \delta + \cos \varphi \cos \delta \cos \omega \quad (3)$$

$$\delta = 23.4^\circ \times \sin \frac{2\pi(284 + N)}{365} \quad (4)$$

$$\omega = (e - 12) \times 15^\circ \quad (5)$$

$$e = \text{time} - \frac{4 \times (120 - c)}{60} + \frac{E}{60} \quad (6)$$

$$E = 9.87 \sin \frac{4\pi(N-81)}{365} - 7.53 \cos \frac{2\pi(N-81)}{365} - 1.5 \sin \frac{2\pi(N-81)}{365} \quad (7)$$

式中, I ——大气层外切面的太阳辐照度, W/m^2 ; I_0 ——大气层外水平面辐照度, W/m^2 ; I_{sc} ——太阳常数,当日地距离为平均值时,地球大气层外垂直于太阳光线的每平方米面积上每秒获得的太阳辐射能量称为太阳常数, $I_{sc} = 1367 \text{ W}/\text{m}^2$; N ——年序日,即一年中的第 N 天; θ ——太阳天顶角; φ ——纬度; c ——经度; δ ——赤纬角, $(^\circ)$,地球上某点相对于赤道面的方向和夹角,向北为正,向南为负,量值大小为 $0^\circ \sim \pm 90^\circ$; ω ——时角, $(^\circ)$,太阳午时 $\omega = 0^\circ$,上午取负值,下午取正值; e ——真太阳时, h ; time ——当前时间, h ; E ——地球绕太阳公转时运动和转速变化而产生的时差, min 。

通过插值方法可以得到温度、云量等其他气象因素的时间间隔为 15 min 的数据。由于夜晚太阳辐照度为零,输出功率也为零,所以只对预测日 07:00~20:00 时间段的光伏发电功率进行预测,故只需要截取在 07:00~20:00 时间段的历史数据。

3.2 双重筛选训练样本

为提高预测结果的精确度,本次研究采取双重筛选训练样本的方法,初步筛选保证了功率数据的合理性,二次筛选保证了神经网络训练样本与预测天的相近性。

3.2.1 初步筛选——删除奇异值

光伏发电功率主要受太阳辐照度的影响,当太阳辐照度较大时,光伏发电阵列的输出功率也较大,因此光伏阵列的发电功率与太阳辐照度成高度

正相关。因此可根据它们的正相关建立功率和辐照度的关系式。

筛选步骤:

第1步:根据历史数据和统计方法建立功率和辐照度的关系式(8)。其中不同的天气类型对应的表达式系数不同。

$$P = a + b \times H \quad (8)$$

式中, P ——功率, W; H ——辐照度, W/m²; a 、 b ——系数。

第2步:由于光伏发电的功率不只与辐照度有关,还与温度、云量等其他多种因素有关,因此对表达式的输出设定一个阈值,如果实际功率不在此阈值范围内,则认为该数据是奇异值将其删除。如果一天的数据超过60%都被判断为奇异值,则这一天的数据无代表性,不将其作为样本。

$$S = \sqrt{k_f(Q_f - q_f)^2 + k_t(Q_t - q_t)^2 + k_{tc}(Q_{tc} - q_{tc})^2 + k_{lc}(Q_{lc} - q_{lc})^2 + k_w(Q_w - q_w)^2 + k_h(Q_h - q_h)^2} \quad (10)$$

式中, $k_f = \frac{r_f}{R}$; $k_t = \frac{r_t}{R}$; $k_{tc} = \frac{r_{tc}}{R}$; $k_w = \frac{r_w}{R}$; $k_h = \frac{r_h}{R}$; $R = r_f + r_t + r_{tc} + r_{lc} + r_w + r_h$; Q ——历史日; q ——预测日。下标 f ——辐照度; t ——温度; tc ——全云量; lc ——低云量; w ——风速; h ——湿度。

3)相似度 S 越小则历史天与预测天越相似,将相似度按照从小到大排序,选取前20天作为神经网络的训练数据。

3.3 GA-BP神经网络预测模型

GA-BP神经网络预测模型训练及预测步骤:

1)数据归一化。神经网络的输入变量单位不同,数量级差别也很大,为了使数据处理方便、程序运行时收敛加快要进行归一化,归一化公式如式(11)所示。

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (11)$$

式中, x' 、 x ——归一化后的值和归一化前的值; x_{\min} 、 x_{\max} —— x 的最小值和最大值。

2)确定神经网络的输入和输出。选择不同的输入进行预测,根据式(12)分别计算平均绝对百分比误差(MAPE),选取误差最小时的输入量作为最终输入。最后确定神经网络的输入包括4个变量,分别为辐照度、温度、全云量、低云量。神经网络的输出为对应时刻的功率。

3.2.2 二次筛选——选取相似日

根据辐照度、温度、云量等气象因素选取与预测天最相似的20d作为神经网络的训练样本。

1)根据历史数据和式(9)计算各种气象因素与功率的相关系数,包括辐照度、温度、全云量、低云量、风速和湿度,对应的相关系数分别为 r_f , r_t , r_{tc} , r_{lc} , r_w , r_h 。式(9)为相似度计算公式。

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (9)$$

式中, $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$; $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$, n ——数据的数量。

2)根据式(10)计算历史天气与预测天的相似度 S ,其中 k 表示各种因素对功率影响的大小。

$$e_{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|p_i - p'_i|}{p_i} \times 100\% \quad (12)$$

式中, e_{MAPE} ——平均绝对百分比误差; p_i ——功率真实值; p'_i ——神经网络输出后反归一化的值; n ——数据数量。

3)确定神经网络的隐含层神经元的个数。采用反复实验法确定隐含层神经元个数,即改变隐含层个数分别预测,观察预测误差的大小,选择误差最小时的隐含层个数为最终个数。

4)神经网络隐含层函数选择 Matlab 的 tansig 函数,输出层选择线性函数。

5)利用选取的样本训练 GA-BP 神经网络,得到预测天的功率数据,将预测结果反归一化,反归一化公式如式(13)所示。

$$y = y' \cdot (y_{\max} - y_{\min}) + y_{\min} \quad (13)$$

式中, y ——归一化后的值; y' ——神经网络输出值; y_{\max} ——神经网络输出值的最大值, y_{\min} ——神经网络输出值的最小值。

4 仿真结果

以新疆维吾尔自治区奇台县石城子 20 MW 光伏电站为研究对象,分别对直接进行预测的 GA-BP 神经网络预测直接预测模型和经过双重筛选训练样本的 GA-BP 神经网络预测模型进行仿真分析,

对预测日 07:00~20:00 的功率进行预测,预测时间间隔为 15 min。

仿真数据描述:石城子 1~6 月份的功率为原始训练数据,7 月份的功率为预测对比数据。

首先,对原始训练数据进行初步筛选:截取 07:00~20:00 的历史数据,建立功率与辐照度的关系式,式(14)为晴天(P 、 H 下标为 s)、式(15)为多云(P 、 H 下标为 c)、式(16)为雨天(P 、 H 下标为 r),将历史数据带入相应的表达式中,删除奇异值。

$$P_s = 0.86043 + 0.03893 \times H_s \quad (14)$$

$$P_c = 1.03043 + 0.03152 \times H_c \quad (15)$$

$$P_r = 0.81969 + 0.03111 \times H_r \quad (16)$$

其次,对原始训练数据进行二次筛选:计算各种气象因素与功率的相关系数,得到 $r_t = 0.873$ 、 $r_{tc} = -0.08$ 、 $r_{lc} = -0.244$ 、 $r_w = 0.053$ 、 $r_h = -0.113$ 。根据相似度公式(10)计算历史天与预测天的相似度,按照从小到大排序,选择前 20 天的数据作为神将网络的训练样本。

然后,根据预测模型训练及预测步骤利用 MATLAB 进行 GA-BP 神经网络仿真,得到不同天气类型下的预测结果。图 4 分别为晴天、多云和晴转雨条件下的真实值和预测值曲线。

为了进行预测效果对比,同时给出了未对原始训练数据进行双重筛选而直接进行预测的 GA-BP 神经网络预测直接预测模型的预测功率曲线。根据式(12)和式(17)计算均方根误差(RMSE)和平均绝对百分比误差(MAPE)对预测结果进行评估,如表 1。

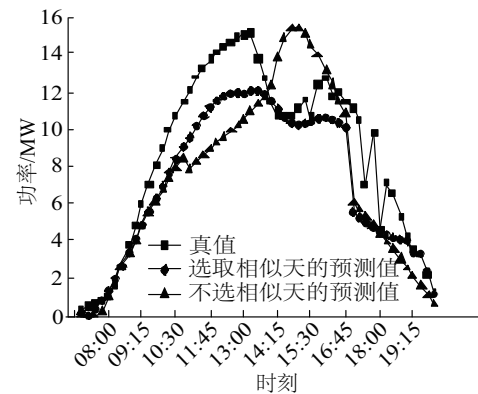
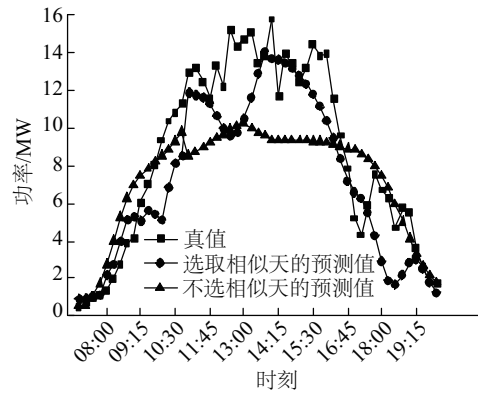
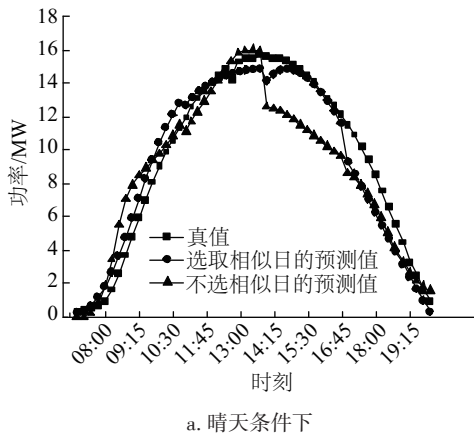


图 4 不同条件下的真实值和预测值曲线

Fig. 4 Real and predictive curves under different conditions

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{p_i - p_i'}{P} \right)^2} \quad (17)$$

式中, P ——功率的额定值, MW; p_i ——功率真实值, MW; p_i' ——神经网络输出后反归一化的值, MW; n ——数据数量。

由图 4 和表 1 可看出:1)3 种天气类型下,基于双重筛选训练样本的 GA-BP 神经网络预测模型的预测效果均优于 GA-BP 神经网络直接预测模型的结果;2)多云和晴转雨时,天气变化较频繁和剧烈,而欧洲中期天气预报中心(european centre for medium-range weather forecasts, ECMWF)的气象数据时间间隔为 3 h,不能准确体现天气短时间内的波动性,因此预测功率的精度不高;3)相比较而言,晴天时的预测精度较高。

表1 预测误差的统计指标

Table 1 Statistical index of forecast error

误差指标	GA-BP神经网络直接预测			双重筛选样本后 GA-BP神经网络预测		
	晴天	多云	多云转雨	晴天	多云	多云转雨
RMSE/%	9.78	14.88	14.88	5.70	11.51	10.71
MAPE/%	25.33	30.07	28.81	15.48	28.07	20.68

5 结 论

本文以新疆维吾尔自治区奇台县石城子 20 MW 光伏电站为研究对象,考虑到天气不同光伏发电规律不同、天气分类不精确、光伏发电功率采集误差等因素,建立了基于双重筛选训练样本的 GA-BP 神经网络预测模型。首先对原始训练数据进行初步筛选,即根据辐照度和功率的关系式剔除功率数据的奇异值,然后进行二次筛选,即分析气象因素对天气的影响权重,定义天气相似度函数,选择与预测天天气最相似的 20 d 作为训练样本,最后利用 GA-BP 神经网络进行预测。采用遗传算法可有效避免 BP 神经网络训练权值时陷入局部最优。与未对原始训练数据进行双重筛选而直接进行预测的 GA-BP 神经网络直接预测模型相比较,仿真结果表明基于双重筛选训练样本的 GA-BP 神经网络预测模型的预测效果更优,而且预测结果的满足《光伏发电功率预测系统功能规范》中“光伏发电短期功率中次日 00:00~24:00 预测月均方根误差小于 20%”的要求。因此该预测模型的提出对提高光伏发电功率的预测效果有很高的参考价值。

【参考文献】

[1] 闫云飞,张智恩,张 力,等.太阳能利用技术及其应用[J].太阳能学报,2012,(Sup1): 47—56.

[1] Yan Yunfei, Zhang Zhien, Zhang Li, et al. Solar energy utilization technology and its application [J]. Acta Energiae Solaris Sinica, 2012, (Sup1): 47—56.

[2] 张 曦,康重庆,张 宁,等.太阳能光伏发电的中长期随机特性分析[J].电力系统自动化,2014, 38 (6): 6—13.

[2] Zhang Xi, Kang Chongqing, Zhang Ning, et al. Analysis of mid/long term random characteristics of photovoltaic power generation[J]. Automation of Electric Power Systems, 2014, 38(6): 6—13.

[3] 邓 雅,胡书举,孟岩峰,等.光伏发电系统功率预

测研究方法综述[J].电气制造,2013,(6): 18.

[3] Deng Ya, Hu Shuju, Meng Yanfeng, et al. A review of research methods on power prediction of photovoltaic power generation system [J]. Electrical Manufacturing, 2013, (6): 18.

[4] 杨金焕,于华丛,葛 亮.太阳能光伏发电应用技术[M].北京:电子工业出版社,2009.

[4] Yang Jinhuan, Yu Huacong, Ge Liang. Application technology of solar photovoltaic power generation [M]. Beijing: Publishing House of Electronics Industry, 2009.

[5] 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.

[5] 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.

[6] 王 雨,苏 适,严玉廷.基于Kalman滤波和BP神经网络的光伏超短期功率预测模型[J].电气技术,2014,(1): 42—46.

[6] Wang Yu, Su Shi, Yan Yuting. Very short-term PV power forecasting model based on Kalman filter algorithm and BP neural network [J]. Electrical Engineering, 2014, (1): 42—46.

[7] 王 飞,米增强,杨奇逊,等.基于神经网络与关联数据的光伏电站发电功率预测方法[J].太阳能学报,2012, 33(7): 1171—1177.

[7] Wang Fei, Mi Zengqiang, Yang Qixun, et al. Power forecasting approach of PV plant based on ANN and relevant data [J]. Acta Energiae Solaris Sinica, 2012, 33 (7): 1171—1177.

[8] 丁 明,王 磊,毕 锐.基于改进BP神经网络的光伏发电系统输出功率短期预测模型[J].电力系统保护与控制,2012, 40(11): 93—99.

[8] Ding Ming, Wang Lei, Bi Rui. A short-term prediction model to forecast output power of photovoltaic system based on improved BP neural network [J]. Power System Protection and Control, 2012, 40(11): 93—99.

- [9] 张 岚, 张艳霞, 郭嫦敏, 等. 基于神经网络的光伏系统发电功率预测[J]. 中国电力, 2010, (9): 75—78.
- [9] Zhang Lan, Zhang Yanxia, Guo Changmin, et al. Photovoltaic system power forecasting based on neural networks[J]. Electric Power, 2010, (9): 75—78.
- [10] Singh V P, Vijay V, Bhatt M S, et al. Generalized neural network methodology for short term solar power forecasting [A]. 2013 13th International Conference on Environment and Electrical Engineering (EEEIC) [C], Wroclaw, Poland, 2013, 58—62.
- [11] 兰 华, 廖志民, 赵 阳. 基于 ARMA 模型的光伏电站出力预测[J]. 电测与仪表, 2011, 48(2): 31—35.
- [11] Lan Hua, Liao Zhimin, Zhao Yang. ARMA model of the solar power station based on output prediction[J]. Electrical Measurement & Instrumentation, 2011, 48(2): 31—35.
- [12] 李英姿, 贺 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.
- [12] Li Yingzi, He Lin, Niu Jincang. Forecasting power generation of grid-connected solar PV system based on Markov chain[J]. Acta Energaie Solaris Sinica, 2014, 35(4): 611—616.
- [13] Li Yingzi, Ru Luan, Niu Jincang. Forecast of power generation for grid-connected photovoltaic system based on grey model and Markov chain[A]. 3rd IEEE Conference on Industrial Electronics and Applications (ICIEA 2008)[C], Singapore, 2008, 1729—1733.
- [14] Tuyishimire B, McCann R, Bute J. Evaluation of a Kalman predictor approach in forecasting PV solar power generation[A]. 2013 4th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)[C], Rogers, AR, USA, 2013, 1—6.
- [15] 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电网光伏发电预测[J]. 电力系统自动化, 2014, 38(5): 19—24.
- [15] Huang Lei, Shu Jie, Jiang Guixiu, et al. Photovoltaic generation forecast based on multidimensional time-series and local support vector regression in micro grid [J]. Automation of Electric Power Systems, 2014, 38(5): 19—24.
- [16] 傅美平, 马红伟, 毛建容. 基于相似日和最小二乘支持向量机的光伏发电短期预测[J]. 电力系统保护与控制, 2012, 40(16): 65—69.
- [16] Fu Meiping, Ma Hongwei, Mao Jianrong. Short-term photovoltaic power forecasting based on similar days and least square support vector machine [J]. Power System Protection and Control, 2012, 40(16): 65—69.
- [17] Yang Xiyun, Jiang Feifei, Liu Huan. Short-term solar radiation prediction based on SVM with similar data[A]. 2nd IET Renewable Power Generation Conference (RPG 2013)[C], Beijing, China, 2013.
- [18] Hossain M R, Oo A M T, Ali A B M S. Hybrid prediction method of solar power using different computational intelligence algorithms[A]. 2012 22nd Australasian Universities Power Engineering Conference (AUPEC)[C], Bali, Indonesia, 2012, 1—6.
- [19] Haque A U, Nehrir M H, Mandal P. Solar PV power generation forecast using a hybrid intelligent approach [A]. Power and Energy Society General Meeting (PES), 2013 IEEE[C], Vancouver, BC, Canada, 2013, 1—5.
- [20] 杨锡运, 刘 欢, 张 彬, 等. 基于熵权法的光伏输出功率组合预测模型[J]. 太阳能学报, 2014, 35(5): 744—749.
- [20] Yang Xiyun, Liu Huan, Zhang Bin, et al. A combination method for photovoltaic power forecasting based on entropy weight method [J]. Acta Energaie Solaris Sinica, 2014, 35(5): 744—749.
- [21] 朱大奇, 史 慧. 人工神经网络原理及应用[M]. 北京: 科学出版社, 2006.
- [21] Zhu Daqi, Shi Hui. Principle and application of artificial neural network[M]. Beijing: Science Press, 2006.
- [22] 韩瑞锋. 遗传算法原理与应用实例[M]. 北京: 兵器工业出版社, 2009.
- [22] Han Ruifeng. Principle and application of genetic algorithm[M]. Beijing: Weapon Industry Press, 2009.

POWER PREDICTION MODEL OF PV POWER GENERATION BASED ON DOUBLE SCREENING OF SAMPLES

Dong Lei¹, Zhou Xiao¹, Hao Ying¹, Liao Xiaozhong¹, Gao Yang²

(1. College of Automation, Beijing Institute of Technology, Beijing 100081, China;

2. College of Electric Power, Shenyang Institute of Engineering, Shenyang 110136, China)

Abstract: PV power generation has characteristics such as volatility, intermittence, but the more similar the weather condition is, the more similar the generation law of PV power generation system is. Firstly, the power singular value was removed to complete the preliminary screening through setting up the relations between irradiance and power. Secondly, the correlation coefficient between power and various meteorological factors such as irradiance, temperature and cloudiness, etc. was calculated, and then the meteorological factors with bigger correlation coefficient were collected to evaluate the similarity between historical days and the forecast day. The most similar historical days from the predicted day was extracted as training samples to complete the secondary screening. At last, the power of PV power generation was predicted using BP neural network and genetic algorithm. The results show that the method has high prediction accuracy.

Keywords: PV power generation prediction; BP neural network; double screening; similar days; genetic algorithm