

基于改进深度受限玻尔兹曼机算法的 光伏发电短期功率概率预测

王继东, 冉冉, 宋智林

(天津大学 智能电网教育部重点实验室, 天津 300072)

摘要:光伏发电功率受自然环境的影响具有明显的波动性、间歇性与随机性,对光伏发电进行短期功率的概率预测可以有效缓解给电网调度、能量管理等方面带来的诸多不利影响。提出一种基于改进深度受限玻尔兹曼机(RBM)算法的光伏发电短期功率概率预测模型,通过灰色关联系数法寻找待预测日的相似日,并利用遗传算法对RBM算法进行参数优化,避免模型参数寻优陷入局部最优,以提高预测模型的预测精度。仿真算例表明,所提模型可以更好地反映光伏发电功率的概率分布。

关键词:光伏发电;概率预测;受限玻尔兹曼机;灰色关联系数法;遗传算法

中图分类号:TM 615

文献标识码:A

DOI:10.16081/j.issn.1006-6047.2018.05.006

0 引言

光伏发电具有污染少、规模灵活等优点,得到了广泛应用。但由于光伏发电系统受环境因素影响明显,存在不确定性、波动性、间歇性等特点,不利于电网的安全调度和能量管理,增加了电网的运行风险^[1]。而光伏发电系统的波动性越大,对其发电功率进行确定性预测的预测精度就会越低,因此,对光伏发电的短期功率进行概率预测,可以更加全面地反映光伏发电信息,对于电网的安全调度和经济管理具有重要的意义^[2-3]。

文献[4]分析了太阳辐射强度、温度对光伏发电功率的影响,采用反传播BP(Back Propagation)神经网络算法建立短期光伏发电功率预测模型,对次日24 h的光伏输出功率进行预测。文献[5]分析了天气类型对光伏发电出力的影响,基于BP神经网络建立了计及天气类型指数的光伏发电短期出力预测模型。文献[6]利用距离分析方法分析光伏发电量与气象因素间的相关性,确定以气温和湿度作为预报输入因子,建立BP神经网络的无辐照度发电量短期预报模型;为适应天气突变,采用自组织特征映射SOM(Self-Organizing feature Map)由云量预报信息对天气类型进行聚类识别。文献[7]针对气象专业天气类型进行归纳合并,得到4种广义天气类型;提取辐照度的特征参数,建立基于支持向量机SVM(Support Vector Machine)的天气状态模式识别模型,辨识恢复部分历史数据所缺失的天气类型信

息。文献[8]建立了径向基函数神经网络模型对光伏系统发电功率进行预测。文献[9]通过人工神经网络ANN(Artificial Neural Network)的方法对光伏出力进行预测。文献[10]通过随机森林等回归算法对光伏功率进行预测。文献[11]基于天气分类,采用SVM算法预测未来短期光伏发电量的区间分布。文献[12]基于太阳辐射和环境温度等变量,采用SVM算法对光伏出力进行实时预测。文献[13]提出采用和声搜索优化高斯过程中超参数的建模方法,提高了传统高斯过程方法的预测精度。文献[14]针对天气突变的问题,通过临近传播的方法对天气进行分类,在此基础上建立回声网络状态方程对光伏出力进行预测。

文献[15]提出利用相似变量的基值点误差,采用组合权重法求取相似误差,确定相似日,然后将相似日的输出功率按不同权重系数加权得到预测功率的光伏功率预测方法。

对于光伏发电系统而言,由于其发电功率受自然环境因素的影响较为明显,对光伏发电短期功率进行预测的精确度难以准确保证,而概率预测却可以保证在满足给定精确度的要求下获取更加全面的光伏发电功率数据。

文献[16]通过分析影响光伏发电系统的各因素及其之间的关系,建立系统的动态贝叶斯网络模型,在基于当前时刻各影响因素水平的条件下,预测未来短期光伏发电量的区间分布。

本文通过分析影响光伏发电功率的环境因素数据,应用灰色关联系数法确定待预测日的相似日,采用改进深度受限玻尔兹曼机(RBM)学习算法构建短期光伏发电功率概率预测模型,以相似日的历史功率数据和待预测日的天气数据作为模型的输入数据,对待预测日的光伏发电功率进行概率预测,仿真算例验证了该模型的有效性。

收稿日期:2017-05-30;**修回日期:**2018-02-21

基金项目:国家自然科学基金资助项目(51477111);国家重点研发计划项目(2016YFB0901104)

Project supported by the National Natural Science Foundation of China(51477111) and National Key Research and Development Program of China(2016YFB0901104)

1 相似日的选取

光伏系统的发电出力受诸多因素的影响,包括地理位置、辐照角度等固定环境因素,也包括光照强度、温度、湿度、云量等可变环境因素,还有转换效率等与自身装置特性相关的因素^[17]。

本文通过分析不同环境因素对光伏发电功率的影响作用,选择对光伏发电功率影响最为明显的光照强度与温度数据作为相似日选择的环境因素判别依据。选取的每日气象特征向量如式(1)所示。

$$\mathbf{x}_i = [x_i(1), x_i(2), x_i(3), x_i(4)] = [t_{hi}, t_{li}, l_{hi}, l_{li}] \quad (1)$$

其中, t_{hi} 为第 i 日最高温度; t_{li} 为第 i 日最低温度; l_{hi} 为第 i 日最大光照强度; l_{li} 为第 i 日最小光照强度。

以 $\mathbf{x}_0$ 表示待预测日, 则待预测日 $\mathbf{x}_0$ 与第 i 个历史日 $\mathbf{x}_i$ 的第 j 个特征分量的关联系数如式(2)所示^[18]。

$$\begin{aligned} \delta_i(j) = & \left[\min_i \left(\min_j |x_0(j) - x_i(j)| \right) + \right. \\ & \left. \rho \max_i \left(\max_j |x_0(j) - x_i(j)| \right) \right] \div \\ & \left[|x_0(j) - x_i(j)| + \rho \max_i \left(\max_j |x_0(j) - \right. \right. \\ & \left. \left. x_i(j)| \right) \right] \end{aligned} \quad (2)$$

其中, ρ 一般取为 0.5。

待预测日 $\mathbf{x}_0$ 与第 i 个历史日 $\mathbf{x}_i$ 的相似度定义如式(3)所示^[19]。

$$\lambda_i = \frac{1}{4} \sum_{j=1}^4 \delta_i(j) \quad (3)$$

以相似度最大的历史日作为待预测日的相似日。

2 改进受限玻尔兹曼机

2.1 受限玻尔兹曼机

受限玻尔兹曼机包含 2 个层: 可见层(visible layer)和隐藏层(hidden layer)。其中神经元之间的连接特点为: 层内无连接、层间全连接^[20]。其结构图如图 1 所示。

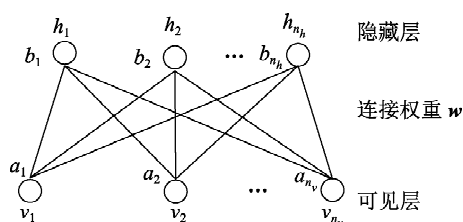


图 1 受限玻尔兹曼机结构图

Fig.1 Structure diagram of RBM

受限玻尔兹曼机的预测过程即在输入为 $\mathbf{v}$ 时,

通过条件概率 $P(\mathbf{h}|\mathbf{v})$ 来计算隐藏层的输出向量 $\mathbf{h}$, 再根据向量 $\mathbf{h}$ 以及条件概率 $P(\mathbf{v}|\mathbf{h})$ 计算可见层向量 $\mathbf{v}$, 以最新得到的可见层向量 $\mathbf{v}$ 和原始输入向量 $\mathbf{v}$ 进行比较, 不断修正参数, 最终使最新得到的可见层向量 $\mathbf{v}$ 向原始输入向量 $\mathbf{v}$ 不断靠近, 达到误差要求。

受限玻尔兹曼机的模型是一个基于能量的模型, 因给定状态 $(\mathbf{v}, \mathbf{h})$, 可定义的能量函数如式(4)所示^[21]。

$$E_{\theta}(\mathbf{v}, \mathbf{h}) = - \sum_{i=1}^{n_v} a_i v_i - \sum_{j=1}^{n_h} b_j h_j - \sum_{i=1}^{n_v} \sum_{j=1}^{n_h} h_j w_{j,i} v_i \quad (4)$$

其中, $w_{j,i}$ 为可见单元 i 与隐单元 j 之间的连接权重; a_i 为可见单元 i 的偏置; b_j 为隐单元 j 的偏置。

利用式(4)的能量函数, 得出状态 $(\mathbf{v}, \mathbf{h})$ 的联合概率分布, 其函数形式如式(5)所示。

$$P_{\theta}(\mathbf{v}, \mathbf{h}) = \frac{1}{Z_{\theta}} e^{E_{\theta}(\mathbf{v}, \mathbf{h})} \quad (5)$$

其中, Z_{θ} 为归一化因子, 其表示形式为 $Z_{\theta} = \sum_{\mathbf{v}, \mathbf{h}} e^{-E_{\theta}(\mathbf{v}, \mathbf{h})}$ 。

根据式(5)可以得出概率分布 $P_{\theta}(\mathbf{v})$ 与 $P_{\theta}(\mathbf{h})$, 其函数形式分别如式(6)、(7)所示。

$$P_{\theta}(\mathbf{v}) = \frac{1}{Z_{\theta}} \sum_{\mathbf{h}} e^{-E_{\theta}(\mathbf{v}, \mathbf{h})} \quad (6)$$

$$P_{\theta}(\mathbf{h}) = \frac{1}{Z_{\theta}} \sum_{\mathbf{v}} e^{-E_{\theta}(\mathbf{v}, \mathbf{h})} \quad (7)$$

而在给定可见层之后, 隐藏层第 j 个节点取值为 1 的概率如式(8)所示。

$$P(h_k = 1 | \mathbf{v}) = \text{sigmoid}(b_k + \sum_{i=1}^{n_v} w_{k,i} v_i) \quad (8)$$

其中, $\text{sigmoid}(x) = 1/(1 + e^{-x})$ 。

而又因为隐藏层节点之间的关系是条件独立的, 可得如式(9)所示条件概率分布。

$$P(\mathbf{h} | \mathbf{v}) = \prod_j P(h_j | \mathbf{v}) \quad (9)$$

同理, 在确定了隐藏层 $\mathbf{h}$ 之后, 可得其条件概率分布, 如式(10)、(11)所示。

$$P(v_k = 1 | \mathbf{h}) = \text{sigmoid}(a_k + \sum_{j=1}^{n_h} w_{j,k} h_j) \quad (10)$$

$$P(\mathbf{v} | \mathbf{h}) = \prod_i P(v_i | \mathbf{h}) \quad (11)$$

给定训练样本后, 训练一个受限玻尔兹曼机即是调整参数 θ , 以拟合给定的训练样本。假定训练样本集合为 $S = \{\mathbf{v}^1, \mathbf{v}^2, \dots, \mathbf{v}^{n_s}\}$, 其中, $\mathbf{v}^i = (v_1^i, v_2^i, \dots, v_{n_v}^i)^T$ ($i = 1, 2, \dots, n_s$; n_s 为训练样本的数目)。

训练受限玻尔兹曼机的目标即最大化似然函数如式(12)所示。

$$L_{\theta,S} = \prod_{i=1}^{n_v} P(\mathbf{v}^i) \quad (12)$$

对式(12)作对数处理,即训练目标可转换为如式(13)所示函数形式。

$$\ln L_{\theta,S} = \ln \prod_{i=1}^{n_v} P(\mathbf{v}^i) = \sum_{i=1}^{n_v} \ln P(\mathbf{v}^i) \quad (13)$$

此外,对于一个正在学习的受限玻尔兹曼机模型,由于似然函数的计算较为复杂,很难直接以其作为受限玻尔兹曼机的评估函数,常采用近似方法重构误差(reconstruction error)来代替似然函数作为受限玻尔兹曼机的评估函数。而重构误差是以训练样本作为初始状态,经过受限玻尔兹曼机的分布进行一次 Gibbs 采样后与原始数据的差值。

在误差满足要求时,隐藏层的输出值即为系统输出值,此时,即可以得到模型的概率分布 $P_{\theta}(\mathbf{v})$ 与 $P_{\theta}(\mathbf{h})$ [22]。

2.2 遗传算法

遗传算法 GA (Genetic Algorithm) 是一种模拟 Darwin 生物进化论过程的寻优算法 [23]。遗传算法以任一群体为起始点,通过随机性的选择、交叉和变异操作,产生具有更好适应能力的新生代个体,通过优胜劣汰的机制使得群体在约束条件的基础上向着更好的搜索空间移动,经过不断的繁衍寻找到适应度最高的个体,即优化问题的最优解。

光伏发电系统的概率预测模型,即改进受限玻尔兹曼机模型,有 3 个待优化的模型参数,分别为 $\mathbf{W}$ 、 a 与 b , 记为 $\theta = (\mathbf{W}, a, b)$, 而遗传算法具有结构简单、泛化能力强、鲁棒性强等优点,可以用作预测模型的优化选取,加快预测算法的收敛速度,从而提高模型的预测性能。

对预测模型的待优化参数 $\theta = (\mathbf{W}, a, b)$, 通过选择、交叉和变异操作进行优化。为避免基因缺失、提高算法的全局收敛性,利用个体的适应度进行选择操作,其具体操作过程为:找出当前群体中适应度最高的个体和适应度最低的个体;若当前群体中最佳个体的适应度比迄今为止最好个体的适应度还要高,就以当前群体中的最佳个体作为新的总的最好个体;用迄今为止的最好个体替换当前群体中的最差个体。交叉操作的具体操作过程为:首先在染色体中随机选择一个点作为交叉点,然后将第一个父辈交叉点前的串和第二个父辈交叉点后的串组合成一个新的染色体,将第二个父辈交叉点前的串和第一个父辈交叉点后的串组合成一个新的染色体。变异操作则是为了保持物种的多样性,以一定概率改变染色体中某个基因的字符值 [17]。

3 基于改进深度受限玻尔兹曼机的概率预测模型

相较于受限玻尔兹曼机模型的单层结构而言,深度受限玻尔兹曼机模型是一种多层结构的预测模型,其结构图如图 2 所示。

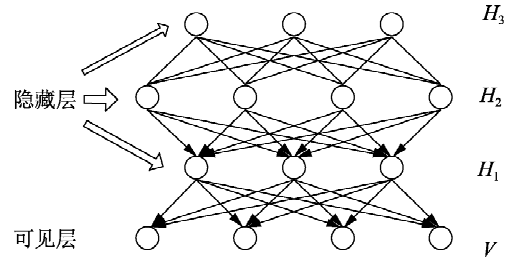


图 2 深度受限玻尔兹曼机结构图

Fig.2 Structure diagram of depth RBM

深度受限玻尔兹曼机学习算法相较于受限玻尔兹曼机学习算法具有更高的精度,通过逐层特征变换,将样本在原空间的特征表示变换到一个新特征空间,从而使预测更加容易。

样本数据,是由天气预报信息以及历史功率和气象数据,经过第 1 节所述方法选取的相似日的数据。

本文以 $\theta = (\mathbf{W}, a, b)$ 作为优化参数,通过遗传算法对预测模型进行参数寻优,建立基于深度受限玻尔兹曼机算法的概率预测模型,主要步骤如下。

a. 确定编码方式。

对改进受限玻尔兹曼机模型参数 $\theta = (\mathbf{W}, a, b)$ 的优化是一个连续复杂的非线性优化问题,采用浮点数编码方法,即个体的每个基因值用某一个范围内的一个浮点数来表示,这种编码方式使用的是决策变量的真实值,适用于决策变量表示范围较大的情况。相对于二进制的编码方式,浮点数编码改善了遗传算法的计算复杂性,提高了运算效率。

b. 生成初始群体。

如果初始种群规模较大,则能够在一定程度上保持种群基因的多样性,预测模型具有更强的全局寻优能力,但付出的代价是计算迭代的时间增加;过小的种群规模虽然具有很快的收敛速度,但可能会带来局部极小问题。

c. 确定适应度函数。

深度受限玻尔兹曼机模型在计算时的目的是使得由计算得来的可见层不断拟合可见层的真实值,其目标函数如式(14)所示。

$$L_{\theta,S} = \prod_{i=1}^{n_v} P(\mathbf{v}^i) \quad (14)$$

在确定遗传算法的适应度函数时,即根据深度受限玻尔兹曼机的目标函数确定其适应度函数的表示方式,如式(15)所示。

$$\begin{aligned}
f(\theta) &= \prod_{i=1}^{n_v} P_{\theta}(\mathbf{v}^i) \\
P_{\theta}(\mathbf{v}) &= \frac{1}{Z_{\theta}} \sum_{\mathbf{h}} e^{-E_{\theta}(\mathbf{v}, \mathbf{h})} \\
E_{\theta}(\mathbf{v}, \mathbf{h}) &= - \sum_{i=1}^{n_v} a_i v_i - \sum_{j=1}^{n_h} b_j h_j - \\
&\quad \sum_{i=1}^{n_v} \sum_{j=1}^{n_h} h_j w_{j,i} v_i \\
Z_{\theta} &= \sum_{\mathbf{v}, \mathbf{h}} e^{-E_{\theta}(\mathbf{v}, \mathbf{h})} \quad (15)
\end{aligned}$$

d. 模型参数优化。

利用遗传算法,通过选择、交叉和变异操作对待优化参数 $\theta = (\mathbf{W}, \mathbf{a}, \mathbf{b})$ 进行寻优。

e. 模型的学习过程。

预测模型反复通过自下而上的学习和自上而下的学习过程计算模型参数,即在学习得到第 $n-1$ 层后,将第 $n-1$ 层的输出作为第 n 层的输入,训练第 n 层,由此分别得到各层的参数。

f. 参数更新。

为了计算目标函数的最大值,文中采用 k 步对比散度 CD- k (Contrastive Divergence- k steps) 算法,具体过程为: $\forall \mathbf{v} \in S$, 取初始值 $\mathbf{v}^{(0)} = \mathbf{v}$, 然后执行 k 步 Gibbs 采样。其中第 t ($t = 1, 2, \dots, k$) 步先后执行如下操作: 首先,利用 $P(\mathbf{h} | \mathbf{v}^{(t-1)})$ 采样出数据 $\mathbf{h}^{(t-1)}$; 然后,利用 $P(\mathbf{v} | \mathbf{h}^{(t-1)})$ 采样出数据 $\mathbf{v}^{(t)}$; 最后,利用第 k 步采样后得到的 $\mathbf{v}^{(k)}$ 近似估计上述 3 个偏导数的计算公式,具体如式 (16)~(18) 所示。

$$\begin{aligned}
\Delta w &= \frac{\partial \ln P(\mathbf{v})}{\partial w_{i,j}} \approx P(h_i = 1 | \mathbf{v}^{(0)}) v_j^{(0)} - \\
&\quad P(h_i = 1 | \mathbf{v}^{(k)}) v_j^{(k)} \quad (16)
\end{aligned}$$

$$\Delta a = \frac{\partial \ln P(\mathbf{v})}{\partial a_i} \approx v_i^{(0)} - v_i^{(k)} \quad (17)$$

$$\begin{aligned}
\Delta b &= \frac{\partial \ln P(\mathbf{v})}{\partial b_i} \approx P(h_i = 1 | \mathbf{v}^{(0)}) - \\
&\quad P(h_i = 1 | \mathbf{v}^{(k)}) \quad (18)
\end{aligned}$$

CD- k 算法的目的是得到这些偏导数的近似值,并根据式 (19) 更新变量:

$$\begin{cases} w^{(k+1)} = w^{(k)} + \eta \Delta w \\ a^{(k+1)} = a^{(k)} + \eta \Delta a \\ b^{(k+1)} = b^{(k)} + \eta \Delta b \end{cases} \quad (19)$$

其中, η 为学习率。学习率过大,可能会导致权重异常增大,学习率过小,又会导致收敛速度变慢,因此,学习率的设置与权重 $w^{(k)}$ 的大小有关,一般地,根据 $\eta \Delta w \approx 10^{-3} w$ 来设置学习率。

最终通过设置初始群体规模、优化算法迭代次数、交叉率和变异率确定模型参数的优化结果。

图 3 为基于遗传算法优化的深度受限玻尔兹曼机光伏发电短期功率概率预测模型的流程图。

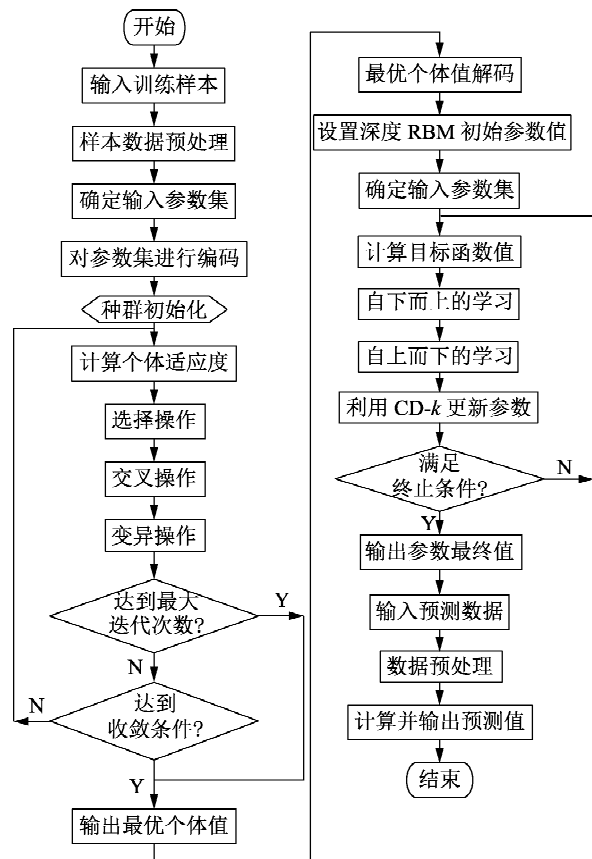


图 3 改进深度受限玻尔兹曼机模型的流程图

Fig.3 Flowchart of improved depth RBM model

预测流程主要分为 3 个部分: 首先通过遗传算法计算参数的初始值, 然后根据深度受限玻尔兹曼机算法计算参数的精确值, 最后通过输入的气象预报数据预测未来的光伏功率输出。

4 算例分析

本文采用天津某地光伏电站的实测数据与相关数据天气预报数据, 进行光伏发电系统短期功率概率预测模型的建立与有效性验证。采用 2013 年夏季数据作为模型的样本数据, 每日光伏功率发电数据的采样时间段取 06:00—19:00, 以 15 min 为采样时间间隔。采用遗传算法进行光伏发电预测模型的参数寻优, 并与 SVM 模型的预测结果进行比较。

本文建立基于遗传算法优化的深度受限玻尔兹曼机学习预测模型, 以相似日的历史发电功率数据和来自数据天气预报的预测日的天气因素信息为模型的输入量, 以光伏发电系统在预测日的输出功率信息为预测模型的输出量。以一个月的实测信息作为预测模型样本数据, 预测未来 24 h 的光伏发电输出功率, 并且得到预测点的概率密度函数。本文采用遗传算法作为优化模型, 其中遗传算法的交叉率

为0.6,变异率的取值一般很小,本文将其设定为0.2,初始群体规模设定为50,最大迭代次数为200,深度受限玻尔兹曼机的结构层数选定为4层。

图4给出了一个预测点的正态分布图。

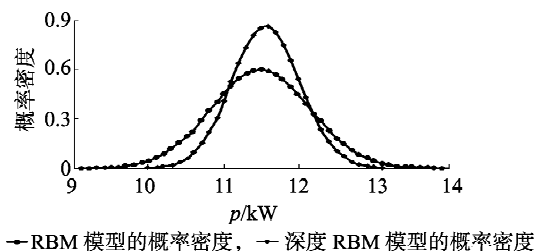


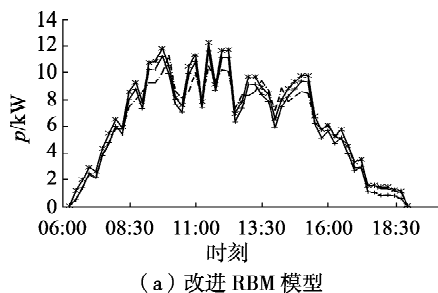
图4 单点预测值的正态分布曲线图

Fig.4 Normal distribution curve of single point forecast

在图4的正态分布曲线中,这一预测点的实际功率为10.5 kW,对于深度受限玻尔兹曼机模型,经概率预测模型的预测,其服从期望值为11.56 kW、方差值为0.469的正态分布。对于受限玻尔兹曼机预测模型,其服从期望值为11.49 kW、方差值为0.659的正态分布。可以看出,深度受限玻尔兹曼机模型要比受限玻尔兹曼机模型的预测精确度更高。

对于概率预测模型,它可以给出每一个预测点所服从的正态分布函数及其概率密度函数的期望值与方差值,而期望值是每个预测点的单点预测功率值,即模型的输出值。给定一个置信水平,即可以通过正态分布函数计算出概率上下限值预测的上限值和下限值。因此,对于概率预测模型,在给定置信水平的条件下,其概率上下限值预测精度在很大程度上取决于此模型所得到的单点预测值的预测精度。图5、6给出了不同模型在不同置信区间下的概率预测上下限。

从图5、6中可以看出,在给定的置信水平条件下,概率上下限值预测基本可以覆盖实测功率值,只有在10:00—13:00之间光照变化起伏较大的时间段,概率上下限值预测的某些预测点会与实测功率点非常接近,概率上下限值预测的区间带宽在这些预测点会变得很窄,但整体上概率上下限值预测还是可以满足要求的,这说明概率上下限值预测具有一定的可信度,并且从图中还可以看出,概率上下限值预测的上限值和下限值与实测功率值的差值并不



(a) 改进 RBM 模型

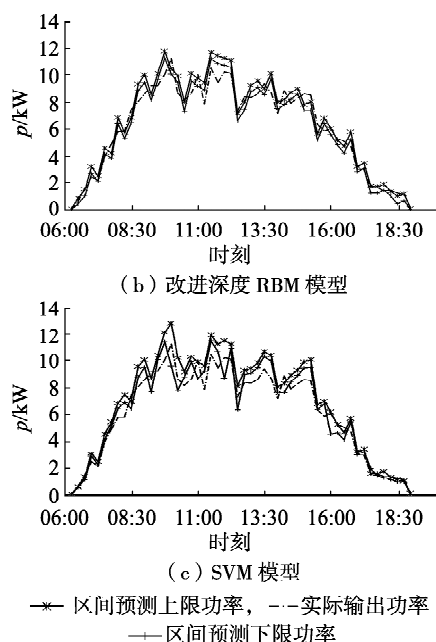


图5 不同模型在置信水平为80%时的概率上下限值预测图

Fig.5 Upper and lower limits of probability forecast diagrams for different models with 80% confidence level

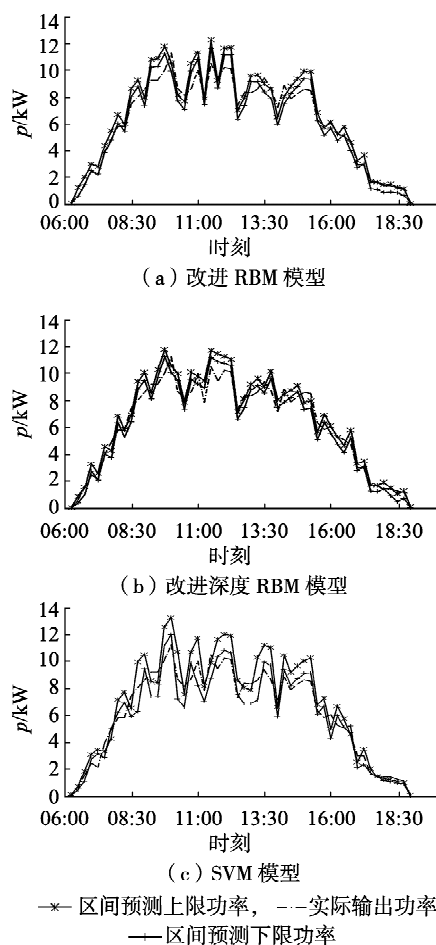


图6 不同模型在置信水平为90%时的概率上下限值预测图

Fig.6 Upper and lower limits of probability forecast diagrams for different models with 90% confidence level

是很大,整体上可以满足预测要求,这就说明概率上下限值预测还可以保证像单点预测一样的预测精度。对比 3 种预测模型可看出,深度受限玻尔兹曼机在点预测(即平均中心误差)以及区间预测(即区间平均带宽)中都有较好的表现。

从表 1 中可以看出,当置信水平较大时,预测区间的平均带宽更宽,这是因为更高的置信水平意味着更高的覆盖概率,从而确保预测结果的可靠性,相应的平均中心误差则会更大。与基于受限玻尔兹曼机算法的预测模型相比较,可以看出,在同一置信水平条件下,基于深度学习算法预测模型的区间带宽更窄,平均中心误差更小,说明其预测精度更高。即通过概率生成模型来得到光伏发电系统的置信区间时,预测模型的预测精度越高,其在同一置信水平条件下的置信区间越窄,模型的预测效果越好。

表 1 不同置信水平下的概率上下限值分析结果

Table 1 Analysis results of upper and lower limits of probability forecast with different confidence levels

置信水平/%	模型	区间平均带宽/kW	平均中心误差/kW
80	SVM	0.567 4	0.628 2
	RBM	0.526 4	0.811 9
	深度 RBM	0.461 8	0.665 2
90	SVM	0.894 9	0.628 2
	RBM	0.601 0	0.811 9
	深度 RBM	0.526 6	0.665 2

5 结论

为了有效缓解给电网规划、能量管理、稳定运行等方面带来的诸多不利影响,需要对光伏发电功率进行短期概率预测。本文建立了基于改进深度受限玻尔兹曼机学习算法的光伏系统发电短期功率概率预测模型,实现了在给定置信水平条件下的光伏发电短期功率的概率上下限值预测,并且在相同置信水平条件下改进深度受限玻尔兹曼机预测模型的预测精度要高于受限玻尔兹曼机预测模型,仿真算例表明了所建模型的有效性。

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. Review on the effect of large-scale PV generation on power systems [J]. Proceedings of the CSEE, 2014, 34(1): 1-14.

[2] KONG L, TANG X, QI Z. Study on modified EMAP model and its application in collaborative operation of hybrid distributed power generation system[C] // 2009 International Conference on Sustainable Power Generation and Supply. Nanjing, China, IEEE, 2009: 1-7.

[3] 张晓波,张保会,吴雄. 风光预测后微电网的优化运行[J]. 电

力自动化设备,2016,36(3):21-25.

ZHANG Xiaobo, ZHANG Baohui, WU Xiong. Coordinated optimal operation between micro-grid and distribution network based on dynamic award and penalty price [J]. Electric Power Automation Equipment, 2016, 36(3): 21-25.

- [4] 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.
- [5] 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):57-64.
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): 57-64.
- [6] 代倩,段善旭,蔡涛,等. 基于天气类型聚类识别的光伏系统短期无辐照度发电预测模型研究[J]. 中国电机工程学报,2011,31(34):28-35.
DAI Qian, DUAN Shanxu, CAI Tao, et al. Short-term PV generation system forecasting model without irradiation based on weather type clustering [J]. Proceedings of the CSEE, 2011, 31(34): 28-35.
- [7] 王飞,米增强,甄钊,等. 基于天气状态模式识别的光伏电站发电功率分类预测方法[J]. 中国电机工程学报,2013,33(34):75-82.
WANG Fei, MI Zengqiang, ZHEN Zhao, et al. A classified forecasting approach of power generation for photovoltaic plants based on weather condition pattern recognition [J]. Proceedings of the CSEE, 2013, 33(34): 75-82.
- [8] 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.
- [9] MELLITA, SALAM S, KALOGIROU S A. Artificial neural network-based model for estimating the produced power of a photovoltaic module [J]. Renewable Energy, 2013, 60(4): 71-78.
- [10] ZAMOM, MESTRE O, ARBOGAST P, et al. A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part II: probabilistic forecast of daily production [J]. Solar Energy, 2014, 105: 792-803.
- [11] SHI J, LEE W J, LIU Y, et al. Forecasting power output of photovoltaic systems based on weather classification and support vector machines [J]. IEEE Transactions on Industry Applications, 2012, 48(3): 1064-1069.
- [12] BAHARINK A, RAHMAN A B D, HASSAN H, et al. Hourly photovoltaics power output prediction for malaysia using support vector regression [J]. Applied Mechanics & Materials, 2015, 785: 591-595.
- [13] 李元诚,王蓓,王旭峰. 基于和声搜索-高斯过程混合算法的光伏功率预测[J]. 电力自动化设备,2014,34(8):13-18.
LI Yuancheng, WANG Bei, WANG Xufeng. Photovoltaic power forecasting based on harmony search and Gaussian process algorithms [J]. Electric Power Automation Equipment, 2014, 34(8): 13-18.
- [14] 李乐,刘天琪. 基于近邻传播聚类和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.
- [15] 杨锡运,刘欢,张彬,等. 组合权重相似日选取方法及光伏输出功率预测[J]. 电力自动化设备,2014,34(9):118-122.

- YANG Xiyun, LIU Huan, ZHANG Bin, et al. Similar day selection based on combined weight and photovoltaic power output forecasting [J]. Electric Power Automation Equipment, 2014, 34 (9): 118-122.
- [16] 董雷, 周文萍, 张沛, 等. 基于动态贝叶斯网络的光伏发电短期功率预测[J]. 中国电机工程学报, 2013, 33(增刊1): 38-45.
- DONG Lei, ZHOU Wenping, ZHANG Pei, et al. Short-term photovoltaic output forecast based on dynamic bayesian network theory [J]. Proceedings of the CSEE, 2013, 33(Supplement 1): 38-45.
- [17] ALONSO J, BATLLES F J. Solar radiation forecasting in the short- and medium-term under all sky conditions [J]. Energy, 2015, 83: 387-393.
- [18] 黎灿兵, 李晓辉, 赵瑞, 等. 电力短期负荷预测相似日选取算法[J]. 电力系统自动化, 2008, 32(9): 69-71.
- LI Canbing, LI Xiaohui, ZHAO Rui, et al. Power load forecasting similar days selecting algorithm [J]. Automation of Electric Power Systems, 2008, 32(9): 69-71.
- [19] 林辉, 刘晶, 郝志峰, 等. 基于相似日负荷修正的节假日短期负荷预测[J]. 电力系统保护与控制, 2010, 38(7): 47-51.
- LIN Hui, LIU Jing, HAO Zhifeng, et al. Short-term load forecasting for holidays based on the similar days' load modification [J]. Power System Protection and Control, 2010, 38(7): 47-51.
- [20] 陆萍, 陈志峰, 施连敏. RBM 学习方法对比[J]. 计算机时代, 2014, 11: 10-13.
- LU Ping, CHEN Zhifeng, SHI Lianmin. RBF learning method comparison [J]. Computer ERA, 2014, 11: 10-13.
- [21] 张春霞, 姬楠楠, 王冠伟. 受限玻尔兹曼机[J]. 工程数学学报, 2015(2): 159-173.
- ZHANG Chunxia, JI Nannan, WANG Guanwei. Restricted Boltzmann machines [J]. Chinese Journal of Engineering Mathematics, 2015 (2): 159-173.
- [22] 郑志蕴, 李步源, 李伦, 等. 基于云计算的受限玻尔兹曼机推荐算法研究[J]. 计算机科学, 2013, 40(12): 259-263.
- ZHENG Zhiyun, LI Buyuan, LI Lun, et al. Research on restricted Boltzmann machines recommendation algorithm based on cloud computing [J]. Computer Science, 2013, 40(12): 259-263.
- [23] 郑立平, 郝忠孝. 遗传算法理论综述[J]. 计算机工程与应用, 2003(21): 50-54.
- ZHENG Liping, HAO Zhongxiao. A review on the theory for the genetic algorithm [J]. Computer Engineering and Applications, 2003 (21): 50-54.

作者简介:



王继东

王继东(1977—),男,山东邹城人,教授,博士,主要研究方向为电力系统电能质量、分布式发电及微网、智能用电 (E-mail: jidongwang@tju.edu.cn);

冉冉(1992—),女,天津人,硕士研究生,主要研究方向为光伏发电 (E-mail: ranran1110@tju.edu.cn);

宋智林(1990—),男,河北邢台人,硕士研究生,主要研究方向为光伏发电 (E-mail: songzhilin111@163.com)。

Probability forecast of short-term photovoltaic power generation based on improved depth restricted Boltzmann machine algorithm

WANG Jidong, RAN Ran, SONG Zhilin

(Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072, China)

Abstract: Since the photovoltaic power is of obvious fluctuation, intermittence and random caused by the natural environment, the probability forecast of short-term photovoltaic power generation can effectively reduce the adverse effects on the aspects of power grid dispatching, energy management, etc. A probability forecast model based on the improved depth RBM (Restricted Boltzmann Machine) algorithm is proposed, which adopts the grey correlation method to find the day similar to the forecasted day, and uses the genetic algorithm to optimize the parameters of RBM algorithm, avoiding falling into local optimum and improving the forecast accuracy. A simulation example shows that the proposed model can better reflect the probability distribution of photovoltaic power generation.

Key words: photovoltaic power generation; probability forecast; RBM; grey correlation coefficient method; genetic algorithm