

低成本微电网轻量化在线超短期光伏功率预测算法设计

陈晓科¹, 陈奇芳², 何 婷^{3†}, 黄锦成⁴

(1. 广东电网有限责任公司电力科学研究院; 广东省智能电网新技术企业重点实验室, 广东 广州 510080;

2. 华北电力大学 新能源电力系统国家重点实验室, 北京 102206;

3. 华南理工大学 电力学院 广东省绿色能源技术重点实验室, 广东 广州 510640;

4. 广东智造能源科技研究有限公司, 广东 广州 511458)

摘要: 本文结合用户侧光伏微电网的实际工程需求, 研究了满足嵌入式应用需求的轻量化在线超短期光伏功率预测算法. 采用了核函数极限学习机算法作为在线超短期光伏功率预测算法的核心, 通过使用特征序列代替传统时间序列作为训练样本集, 实现了数据量的轻量化设计; 通过基于趋势加权相似度的训练样本精选, 提高了算法精度, 简化了计算量, 降低了算法计算时间. 通过嵌入式系统实验测试结果表明, 本文提出的轻量化超短期光伏预测算法在精度、计算时间和数据存储等方面都能满足嵌入式系统低成本应用的需求.

关键词: 光伏功率预测; 超短期; 轻量化; 核函数极限学习机

中图分类号: TP273 **文献标识码:** A

Design of simplified online ultra-short term photovoltaic output forecasting algorithm for low cost microgrid

CHEN Xiao-ke¹, CHEN Qi-fang², HE Ting^{3†}, HUANG Jin-cheng⁴

(1. Electric Power Research Institute of Guangdong Power Grid Co., Ltd.; Guangdong Province Key Laboratory of Smart Grid Technology, Guangzhou Guangdong 510080, China;

2. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China;

3. Guangdong Key Laboratory of Clean Energy Technology, School of Electric Power, South China University of Technology, Guangzhou Guangdong 510640, China;

4. Guangdong Intetwork Energy Technology Co., Ltd., Guangzhou Guangdong 511458, China)

Abstract: Based on the engineering requirement of user side photovoltaic (PV) microgrid, a simplified online short term PV output forecasting algorithm is studied for embedded system application. Extreme learning machine with kernel (ELM.K) algorithm is adopted as the main part. The traditional time sequence of training dataset is replaced with characteristic sequence of history data, therefore, the amount of storage place of training data is reduced. Because of the optimal training dataset is selected from original training dataset by trend weighted similarity, the accuracy is improved, also the amount of calculation and runtime of forecasting algorithm is reduced. The test results of embedded system show that the performance of proposed online short term PV output forecasting algorithm on accuracy, runtime and storage occupation can satisfy the requirement of low cost embedded system application.

Key words: photovoltaic output forecasting; ultra-short term; simplified; extreme learning machine with kernel

1 引言(Introduction)

随着经济的快速发展, 能源危机和环境污染问题日益严重, 发展清洁能源成为社会共识, 并受到了社会的广泛关注^[1]. 据英国石油公司预测, 到2030年, 全球可再生能源将占全球能源供应比例的30%^[2]. 光伏发电作为主要的新能源发电方式之一, 可以应用于智能楼宇型光伏微电网系统中、家庭型户用型光伏微电网系统中及偏远地区的大规模光伏发电厂等. 但是,

由于光伏功率具有间歇性与波动性, 给光伏微电网或光伏发电厂的能量管理、经济调度和运行控制带来了巨大的挑战. 因此, 如何准确预测未来时刻的光伏功率成为解决光伏间歇性与波动性问题的关键^[3-4], 并得到国内外学者的广泛关注.

从技术层面上可以将光伏预测算法分为统计方法和人工智能方法两类. 文献[5]提出了一种基于马尔科夫链理论的光伏电站功率预测方法, 主要从转移矩

阵、不同阶段间的衔接和状态转移的修正等方面建立统计预测模型. 该方法需要通过对大量历史数据的统计得到状态转移概率, 并且需要较多存储空间存放每个预测时刻的转移矩阵, 因而, 算法性能会受到嵌入式系统有限资源的约束. 文献[6]提出了一种基于时间序列的自回归滑动平均模型(auto-regressive and moving average model, ARMA). 首先证明了晴天的光伏功率序列是平稳的; 其次, 进行了模型定阶、自相关函数和偏相关函数的求取; 最后对模型参数进行了整定, 得到ARMA预测模型. 该模型需要用到多个月份的历史数据进行训练, 数据存储量和计算量对嵌入式应用偏大. 文献[7]提出了一种基于机器学习的自适应光伏超短期功率预测方法. 该方法采用小波分析结果训练支持向量机(support vector machine, SVM)分类器, 结合前30 min的历史数据对未来15 min的出力曲线类型进行预测, 最后从ARMA算法和人工神经网络算法(artificial neuro network, ANN)中挑选一种构建预测模型. 文献[8]提出了一种基于神经网络的光伏功率预测算法, 并研究了不同时间尺度的历史数据对预测精度的影响. 基于神经网络的预测算法需要较多的历史数据和较大计算量完成网络权重的训练, 因而不适合资源受限的嵌入式系统在线预测应用. 文献[9]提出了一种基于核函数极限学习机的微电网短期负荷预测方法. 该方法采用了大量的时间序列样本数据对模型进行训练, 由于数据量较大, 计算量也较大, 因此会在嵌入式系统中的应用受限. 此外, 由于光伏特性与负荷特性差异巨大, 适用于负荷预测的模型结构并不能直接用于光伏预测, 需根据光伏特性进行改进.

从预测方式上可将光伏预测算法分为直接预测和间接预测. 直接预测是通过历史数据信息直接预测光伏出力^[5-8]; 间接预测是先预测影响光伏出力的关键因素值(如光照强度、温度、云量等), 再根据光伏电池的光电转换模型计算出光伏功率. 文献[10]基于光伏发电量、地表太阳能辐射量和气温序列, 分别按统一建模和时间序列建模两种方法建立了最小二乘支持向量机预测模型. 文献[11]提出了一种间接预测方法. 该方法将大气层外辐射、大气质量、图像亮度和云量作为输入因子, 基于径向基函数神经网络算法建立了地表辐射的预测模型, 然后根据光电转换模型计算得到最终的超短期光伏功率. 文献[10-11]提出的间接预测算法需要采集云层信息, 需要进行多种数据分析和处理, 数据计算量大、消耗内存资源多, 并且对数据采集、处理要求较高, 投入成本较大, 不适合低成本微电网的应用.

目前, 基于数据统计规律或基于人工智能的光伏预测算法都是基于充足的数据信息资源和强大的设备计算能力, 通过对大量历史数据和实时数据的特征挖掘, 结合昂贵的数值天气预报, 采用复杂的预测模

型, 实现较高精度的预测结果. 这类预测算法对数据资源和设备计算能力的要求较高, 且实现成本高, 使其仅适用于规模较大且具有高性能监控服务器的光伏电站系统的应用.

随着国家分布式光伏的大力推广政策, 用户侧光伏型微电网进入了公众视野, 并迅速成为了集成分布式光伏发电系统的主要形式之一. 对于商用型光伏微电网和民用型光伏微电网, 其建设规模小, 为了便于大范围建设和推广, 经济性是首要投资目标. 因此, 这类光伏微电网都不再具备独立高性能监控服务器, 取而代之的是嵌入式中央控制器, 其计算资源和计算能力受到严重的制约. 此外, 这类光伏微电网也不再具备完善的信息资源获取渠道, 从而在很大程度上限制了现有算法所依赖的数据资源.

本文结合用户侧光伏微电网的实际工程需求, 研究了满足嵌入式应用需求的轻量化光伏超短期预测算法. 采用了计算速度快、精度高的核函数极限学习机(extreme learning machine with kernel, ELM_K)算法作为超短期光伏功率预测算法的核心, 在满足工程精度要求的情况下, 主要从计算量和数据量两个方面对算法进行轻量化设计. 通过嵌入式系统实验测试结果表明, 本文提出的轻量化超短期光伏预测算法在精度、计算时间和数据存储等方面都能满足实际工程需求.

2 极限学习机理论(Extreme learning machine theory)

基本极限学习机(extreme learning machine, ELM)是一种前馈神经网络算法, 其网络层次结构如图1所示, 可表示为

$$f(x) = \sum_{i=1}^L \beta_i G(a_i, b_i, x_j), \quad j = 1, \dots, N, \quad (1)$$

式中: a_i 为链接输入层与隐藏层节点的输入权重, b_i 为隐藏层节点的阈值, 且 a_i, b_i 为随机生成. $G(a_i, b_i, x_j)$ 为隐藏层的激励函数; β_i 为连接隐藏层和输出层的输出权重; N 为输入层输入数据 x_j 的维数.

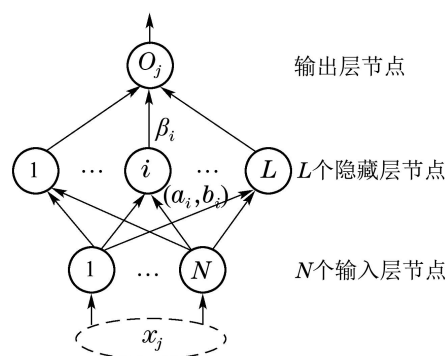


图 1 极限学习机算法的神经网络结构图

Fig. 1 Structure of extreme learning machine algorithm

与传统基于回归的神经网络预测算法相比(如BP神经网络算法), ELM算法在具有两个目标: 1) 获得最小的预测误差; 2) 最小化输出权重向量 β , 如式(2)所示:

$$\min \|H\beta - O\|^2 \text{ 和 } \min \|\beta\|, \quad (2)$$

式中 H 为ELM算法的隐藏层激励函数构成的矩阵, 如式(3)所示:

$$H = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix}. \quad (3)$$

ELM算法的输出权重向量 β 和目标值 O , 如式(4)所示:

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix}, \quad O = \begin{bmatrix} O_1 \\ \vdots \\ O_N \end{bmatrix}. \quad (4)$$

由Bartlett的相关理论可知, 在前馈神经网络的预测误差最小的时候, 其网络输出权重越小, 算法对应的泛化能力越强. 在确定ELM算法的网络结构时, 通过最小二乘法求解输出权重 β 是常用的方法, 如式(5)所示:

$$\beta = H^\dagger O = H^T (HH^T)^{-1} O = H^T \left(\frac{1}{C} + HH^T \right)^{-1} O, \quad (5)$$

式中 H^\dagger 为 H 的广义逆矩阵^[15]. 为了提高结果的泛化能力和稳定性, 通常基于岭回归理论适当增加正常数 $1/C$.

ELM_K算法是在基本ELM算法基础上, 利用Mercer条件构造得到. 主要用于激励函数 $h_i(x)$ 未知的情况. 在极限学习机基本网络特性的基础上引入核函数, 可以是ELM_K获得更好的预测精度. ELM_K网络特性方程可描述为式(6)所示:

$$f(x) = h(x)H^T \left(\frac{1}{C} + HH^T \right)^{-1} O = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{1}{C} + \Omega_{ELM} \right)^{-1} O, \quad (6)$$

式中 $K(x_i, x_j)$ 为核函数, 在实际应用中一般将高斯函数作为核函数, 其表示形式如式(7)所示:

$$\Omega_{ELM}(x_i, x_j) = K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2). \quad (7)$$

通过引入高斯核函数后, 在实际应用中无需知道带求解问题准确的激励函数, 从而简化了算法的难度.

3 轻量化超短期光伏功率预测模型 (Simplified ultra-short term photovoltaic output forecasting model)

3.1 需求分析与建模 (Demand analysis and modeling)

相对于集中式光伏电厂的运行条件, 实际工程中用户侧微电网光伏功率预测具有以下特点: 1) 无高性能监控服务器, 光伏功率预测算法需运行在计算能力与计算资源受限的嵌入式系统中; 2) 信息资源获取渠道单一, 无低成本数值天气预报, 甚至没有任何天气信息, 只能通过逆变器获取光伏功率信息; 3) 数据存储空间和数据存、取速度受限, 不适合进行在线大数据特征分析与挖掘.

考虑到实际工程中用户侧微电网光伏功率预测的特点, 本文将用户侧微电网光伏功率预测算法分解为离线参数优化整定与在线功率预测两部分, 分别如图2和图3所示.

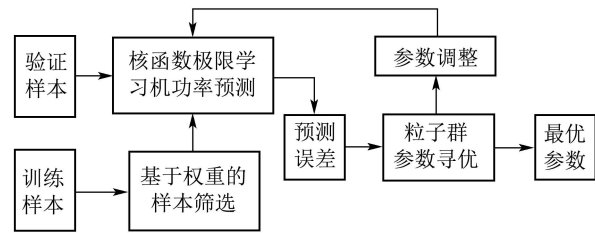


图2 光伏功率预测算法离线参数寻优

Fig. 2 Off-line parameters optimization of PV output forecasting algorithm

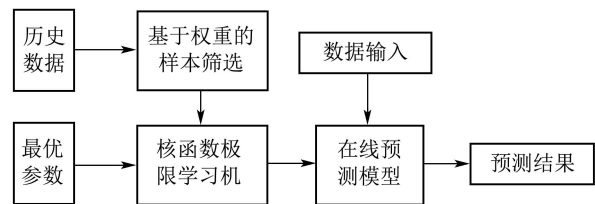


图3 光伏功率在线预测算法结构

Fig. 3 On-line PV output forecasting algorithm

其中: 离线参数优化整定在PC机上运行, 通过粒子群算法和大量历史数据, 以整体预测误差最小为目标, 对核函数极限学习机中的两个参数 C 和 γ 进行一次性离线最优化整定, 形成配置文件即可. 在线功率预测部分设计在嵌入式系统中运行, 通过有限数据信息在规定时间内得到预测结果.

3.2 轻量化预测模型设计 (Simplified forecasting model design)

3.2.1 基于嵌入式系统的系统架构设计 (System architecture design based on embedded system)

本文选用的嵌入式系统实验测试平台为基于

STM32F407 的嵌入式系统开发板, 主控芯片主频最高为 168 MHz, 片内 RAM 存储 192 kB, 片外 RAM 存储 2 MB, 用于存储程序的 NOR Flash 存储空间 16 MB, NAND Flash 存储空间 128 MB, 操作系统为 uCOS-III, 采用的主要系统功能架构如图 4 所示。

由图 4 可知, 公共基础层主要分为信息通信、状态监控和数据存取 3 个功能任务; 应用功能层分为功率预测、能量优化调度和保护与控制 3 个功能任务; 人机交互层

交互层任务主要实现运行状态展示、用户信息录入等功能。在分时多任务的操作系统中, 每个功能任务需由系统按设定的时间周期进行调用并为其分配独立的 RAM 空间。因而当系统任务较多且较复杂时, 再加上系统自身的多个任务, 嵌入式系统主频低和 RAM 存储空间不足的短板问题非常突出。因此, 必须对每个任务的计算量和存储空间占有量进行轻量化设计, 特别是复杂度较高、数据量较大的功率预测算法。



图 4 基于嵌入式的主要系统功能架构

Fig. 4 Main system architecture based on embedded system

3.2.2 数据样本属性结构设计(Attribute structure design of data samples)

超短期光伏预测算法中, 输入数据的样本属性结构至关重要, 对预测精度、计算量、计算时间、数据存储量和开发成本等有很大影响。1) 数据属性所包含的信息全面程度对预测精度有决定性作用, 如数值天气预报可大幅度提高预测精度, 但是价格昂贵, 不利于推广应用。2) 数据属性的维数对计算量、计算时间和数据存储量有决定性作用, 随着数据维数增长, 计算量和计算时间会成指数, 同时数据存储量也会增长。3) 工程实际中为了降低成本, 往往只留有最基础的数据来源的渠道, 如只有光伏功率信息。因此, 预测算法输入数据样本属性选择, 需要紧密结合实际工程中的数据来源和设备性能, 在保证预测精度的同时降低数据属性维度。本文针对应用于无天气信息的用户侧微电网光伏功率预测算法, 选择相邻 4 个时刻的历史功率信息作为数据属性, 设计样本属性结构如式(8)所示:

$$S_{t+1} = \{P_t, P_{t-1}, P_{t-2}, P_{t-3}\}, \quad (8)$$

式中 S_{t+1} 为预测 $t+1$ 时刻光伏功率的输入数据样

本, 分别由 $t, t-1, t-2, t-3$ 时刻的光伏功率构成。

3.2.3 训练数据轻量化设计(Simplified design of training data)

基于机器学习类型的预测算法, 训练样本数据所包含特征信息的多少是决定算法效果的重要影响因素。一般情况下, 训练样本数量大, 则其所包含的特征信息多, 反之, 包含的特征信息少。为了得到较好的预测结果, 预测算法都会选择数十天或几个月的历史数据作为训练样本。以 60 天的超短期光伏预测算法训练样本数据为例, 每天 96 个点, 算法中的数据为浮点型, 需占用 4 个字节。仅计算时的第一步将历史数据读入内存共需要占用 22.5 kB 空间, 加上算法的数据处理、模型训练等过程中的数据量, 需要的内存空间按 22.5 kB 的倍数增长。这可能导致包括能量管理算法、负荷预测算法、保护控制算法等多个算法的嵌入式系统中央控制器出现内存资源短缺的问题。因此, 在保证预测精度的前提下, 减少超短期光伏功率预测算法的训练样本数据量至关重要。

为了实现减少训练样本而不影响样本特征信息

的多样性,本文采用基于特征序列的训练方法代替基于传统的时间序列的训练法.对式(8)中的样本属性 S_{t+1} 按式(9)进行特征提取,得到特征训练样本如式(10)所示:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}, \quad (9)$$

$$\hat{S}_{t+1} = \{r_t, r_{t-1}, r_{t-2}\}, \quad (10)$$

式中: r_t 为 t 时刻光伏功率相对于 $t-1$ 时刻光伏功率的特征信息, \hat{S}_{t+1} 为特征序列.

设 X 为历史数据特征序列 \hat{S} 构成的训练样本集,如式(11)所示:

$$X = \begin{bmatrix} \hat{S}_{1,1} & \hat{S}_{1,2} & \cdots & \hat{S}_{1,m} \\ \vdots & \vdots & & \vdots \\ \hat{S}_{n,1} & \hat{S}_{n,2} & \cdots & \hat{S}_{n,m} \end{bmatrix}, \quad (11)$$

式中: n 为历史数据天数, m 为每天的数据点数; $\hat{S}_{n,m}$ 为第 n 天第 m 个特征序列元素.采用时间序列的样本集训练核函数极限学习机,在任意时刻 t 只能从训练样本集中获取同一时刻的样本数据进行训练,因此,需要有足够多天数的历史数据才能保证 t 时刻训练样本的多样性,导致算法的数据量大幅度增加,系统内存资源消耗大.然而,采用特征序列作为训练样本集, X 中的每一个特征序列 \hat{S} 都可以在任意时刻 t 用于训练核函数极限学习机,因此,用于训练的特征序列的数量为 $n \times m$ 个,能够用少量天数的历史数据充分保证训练数据特征的多样性,从而大幅度减少算法的数据量,节约系统内存资源.

3.2.4 计算量轻量化设计(Simplified design of calculation data)

对机器学习算法的计算量和计算时间起决定性作用的因素是算法中浮点运算矩阵的维度.计算两个 $n \times n$ 矩阵的乘法,需要执行 $n^2(n-1)$ 次加法运算和 n^3 次乘法运算,因此矩阵计算的时间复杂度为 $O(n^3)$.对于核函数极限学习机算法,关键的两步为隐藏层矩阵 H 的计算和输出权重 β 的计算.执行光伏功率预测算法的过程中,对核函数极限学习机的训练和采用训练好的学习机模型进行预测时,都需要进行多次浮点矩阵乘法和浮点矩阵求逆运算.随着隐藏层矩阵 H 维度的增大,算法的逼近精度会有相应的提高,但是计算量将程指数增长.因此,为了保证预测算法的实时性,在满足精度要求的情况下,需要合理的降低隐藏层矩阵 H 的维度.采用高斯核函数极限学习机,隐藏成矩阵的维度由 $N \times L$ 转化为 $N \times N$, N 为输入层节点数.为了确保算法的预测精度、降低核函数矩阵维度,本文采用了趋势加权

相似度指标对训练样本集进行优选,用优选后的训练样本对核函数极限学习机进行训练,同时优选的数据样本数量即为选用的高斯核函数矩阵的维度.

在当前时刻 t ,需要对 $t+1$ 时刻的光伏功率进行预测.当前时刻的输入数据属性特征序列为 $\hat{S}_{t+1} = \{r_t^*, r_{t-1}^*, r_{t-2}^*\}$,根据式(12)对训练样本集按趋势加权相似度由小到大进行优选,从 $n \times m$ 的特征序列训练样本集中挑选出 k 个精选训练样本作为 $t+1$ 时刻核函数极限学习机的训练样本. k 为精选样本数量,根据预测精度和计算量综合选取.

$$R = \frac{1}{v} \sum_{i=0}^{v-1} |r_{t-i} - r_{t-i}^*| \cdot a_i, \quad (12)$$

式中: R 为趋势加权相似度值; v 为特征序列样本属性的维度; a_i 为 $t-i$ 时刻的趋势权重,未来 $t+1$ 时刻的光伏功率与各历史时刻的光伏功率关联程度随着时间的推移而依次减弱,因此取 $a_0 > a_1 > \cdots > a_{v-1}$,具体取值需根据光伏功率历史数据进行整定.

采用精选后的特征序列样本训练核函数极限学习机,能够得到适应于当前时刻变化趋势的输出权重 β ,从而在较少训练样本数据量的情况下得到较好的预测精度,并且能有效降低计算量,但是,算法在适应突变方面稍有欠缺.

采用传统的时间序列样本对核函数极限学习机进行训练时,一般用几十天同一时刻的光伏历史数据属性作为输入训练样本.为了保证算法整体的预测精度,需要有足够多天数的历史数据样本进行训练,得到具有较好泛化能力的输出权重 β ,达到较好的回归预测效果,对突变的适应性较好.但是训练样本天数过多,会导致核函数极限学习机矩阵维度的增加,从而使单次矩阵计算量成指数增长,使整个预测算法的计算量成指数的倍数增长.

4 算例分析(Analysis of examples)

4.1 实验环境与实验参数(Experimental environment and experimental parameters)

选择北京地区某10kW装机容量1月1日开始连续38天光伏功率历史数据进行验证实验,采样时段为5:00~19:00,采样时间间隔15min,每天共56个数据点.第1天到第8天连续8天的光伏功率历史数据作为训练样本数据库 A ,如式(13)所示.后30天的连续光伏功率历史数据作为验证样本.

$$A = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,m} \\ \vdots & & & \vdots \\ P_{n,1} & P_{n,2} & \cdots & P_{n,m} \end{bmatrix}, \quad (13)$$

式中 $P_{n,m}$ 为第 n 天第 m 个时刻的光伏功率值.

经过参数寻优之后, 参数 C 与 γ 的 2 次幂指数值分别取 17.02 和 16.34; 趋势加权相似度权重为 [1.8, 1.3, 1]; 按照趋势加权相似度精选训练样本数量 k 为 15. 采用平均绝对百分误差 (mean absolute percentage error, MAPE) 作为离线参数寻优阶段的预测精度评价函数, 如式 (14) 所示:

$$\begin{cases} \text{MAPE} = \frac{100\%}{m} \sum_{i=1}^m \left\| \frac{x_i - \hat{x}_i}{x_i} \right\|, \\ x_i > 10\% \times \text{Cap}, \end{cases} \quad (14)$$

式中: x_i 为光伏功率实际值, \hat{x}_i 为光伏功率预测值, Cap 为预测光伏的安装容量, 单位 kW.

首先, 按式 (15) 对历史光伏功率训练样本数据库 \mathbf{A} 进行归一化处理, 缩放至 [0, 1] 的范围. 然后, 按式 (8)–(11) 形成特征序列训练样本集. 最后, 逐日逐点循环输入验证样本数据对轻量化超短期光伏功率在线预测算法进行测试.

$$y_{i,j} = \frac{\max(\mathbf{A}) - p_{i,j}}{\max(\mathbf{A}) - \min(\mathbf{A})}, \quad i \in [1, n], \quad j \in [1, m].$$

(15)

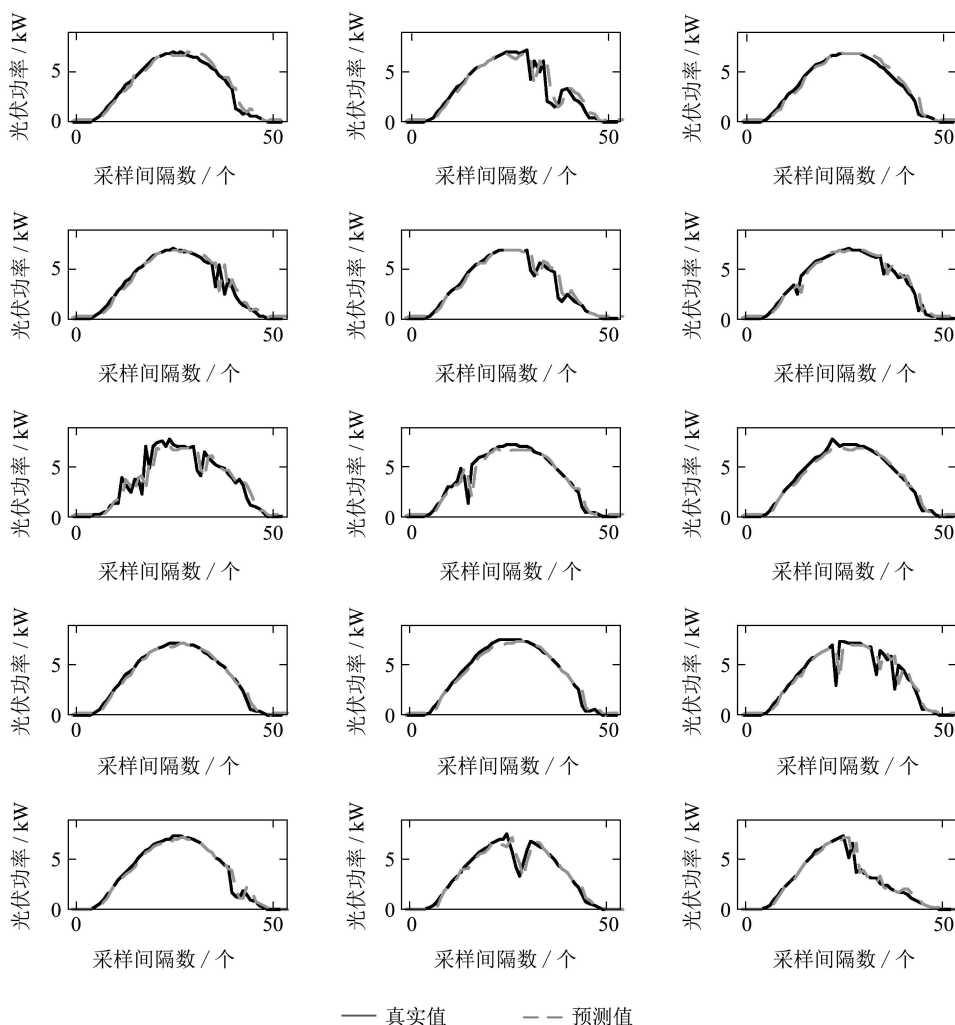


图 6(a) 前 15 天预测值与实际值对比

Fig. 6(a) The comparison between the forecasted results and actual values for the first 15 days in a month

4.2 测试结果及误差分析 (Test results and error analysis)

图 5 为第 25, 26, 27 连续 3 天的真实值和预测值的对比图, 图 6 为连续 30 天的光伏功率预测值与实际值的对比图, 图 7 为连续 30 天的光伏功率预测平均绝对百分误差图.

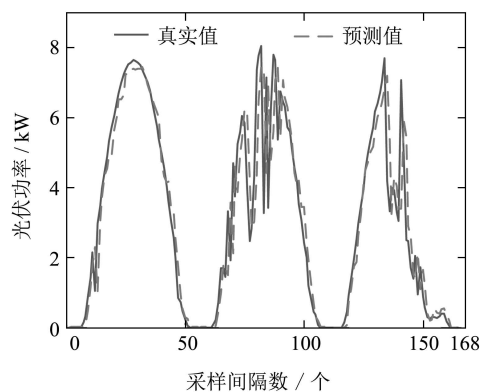


图 5 预测算法连续 3 天预测值与真实值对比

Fig. 5 The comparison between the forecasted results and actual values of the forecasting algorithm for three days

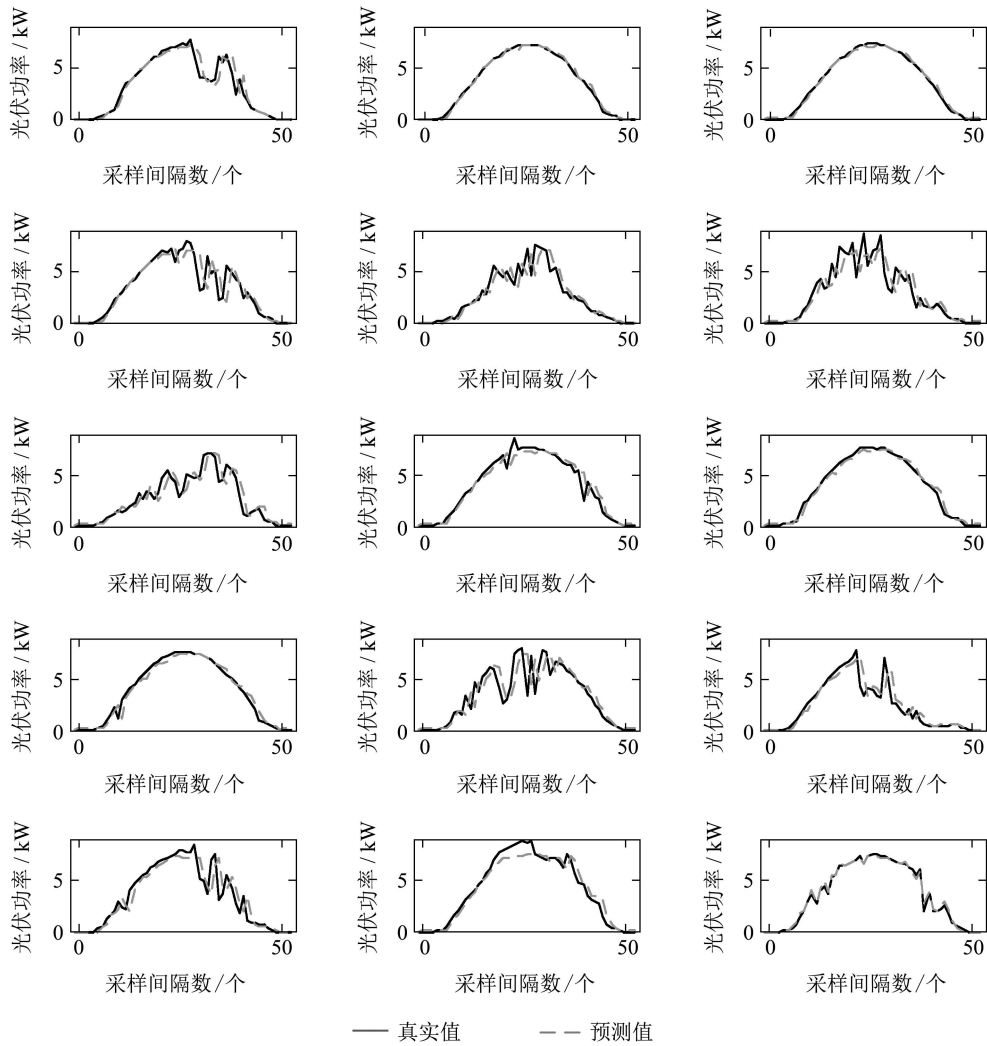


图 6(b) 后15天预测值与实际值对比

Fig. 6(b) The comparison between the forecasted results and actual values for the last 15 days in a month

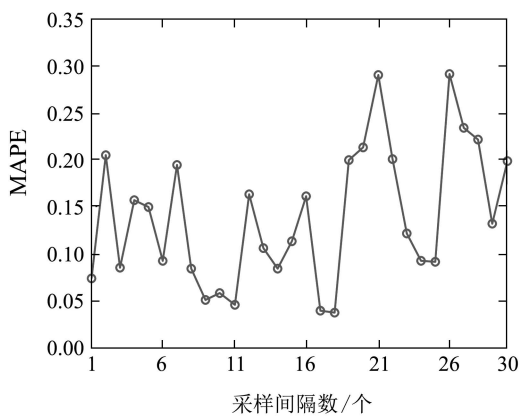


图 7 连续30天光伏功率预测平均绝对百分误差

Fig. 7 MAPE of the forecasted photovoltaic output power in 30 days

图5中,第25天为晴天气天气波动非常小,光伏出力非常平滑.第26天,初始一段时间为晴天,但在中间一段较长的时间出现了多云,之后又为晴天;天气的剧烈变化导致了当天光伏功率的剧烈波动.

第27天为晴转多云转阴,前半天光伏出力正常,后半天光伏功率迅速衰落.由图5和图6中对比情况可知,本文提出的算法既能够在晴天对光伏功率进行预测,又能适应各种天气的剧烈波动.

由图7可知,本文提出的超短期光伏功率预测算法在各种天气剧烈波动的情况下,日平均绝对百分误差不超过30%;在天气晴朗情况下,日平均绝对百分误差不超过10%;连续一个月的超短期光伏功率预测月平均绝对百分误差为13.94%.

4.3 算法的性能测试结果分析(Performance test results of algorithm analysis)

在嵌入式开发板上对轻量化在线超短期光伏功率预测算法进行4次重复运行,采用秒表计时方法得到的算法运行时间如表1所示.

由于采用人工秒表计时方法,记录的运行时间存在一定的误差(± 5 s),但与算法的长期运行时间相比可以忽略.对表1中的时间求平均值,得到每次

执行连续 30 天光伏功率预测的平均时间为 1401.25 s, 从而可以计算出每天进行 56 点光伏功率预测耗时为 46.71 s, 每天进行一次单点预测耗时为 0.834 s, 能充分满足在线超短期光伏功率预测实时性的需求。

表 1 算法 4 次重复运行时间

Table 1 The algorithm running time that repeated four times

序号	预测天数	运行时间
1	30	23' 23"
2	30	23' 19"
3	30	23' 21"
4	30	23' 22"

在算法测试过程中, 为算法任务开辟了 4 kB 的 RAM 堆栈空间, 算法任务能够流畅运行, 加上 8 天共 448 个浮点型历史数据所占用的额外 1.75 kB RAM 存储空间, 算法所占用的总的内存空间不超过 6 kB。

4.4 对比算例分析 (Analysis of contrasting examples)

为了验证本文提出的预测模型的精度, 在 PC 机上分别对本文提出的预测模型与基于支持向量机的预测模型用不同装机容量的微电网数据进行了对比实验分析, 两个方案都从历史数据中选出相同的 8 天历史数据作为训练样本集, 不同之处为:

方案 1 基于支持向量机的预测模型. 采用基于时间序列、无优选的样本集对模型进行训练, 每个时刻只能用历史数据中相同时刻的数据进行模型训练。

方案 2 本文提出的预测模型. 首先, 从历史数据中提取特征序列, 然后采用趋势加权相似度指标对样本集进行优选, 最后用优选的数据对模型进行训练。

用两个方案分别对装机容量为 10 kW 的光伏系统和 40 kW 的光伏系统出力进行连续一个月的预测, 得到的月平均绝对预测误差结果如表 2 所示。

表 2 两个方案月平均绝对预测误差对比情况

Table 2 Comparison of monthly average absolute prediction error between two schemes

光伏类型	本文算法/%	支持向量机算法/%
PV1(8 kW)	12.42	13.89
PV2(15 kW)	11.76	12.57

从表 2 中可以看出, 本文提出的预测算法的预测

精度高于支持向量机算法. 这是因为方案 1 中采用时间序列的训练方法在某个时刻只有 8 个可用的训练样本, 因而, 用于对模型进行训练的样本数据较少, 不能将支持向量机算法训练到合适的状态. 而本文提出的预测模型采用了基于趋势加权相似度指标进行优选的方法, 可以从 8 天历史数据的任意时刻筛选出合适的特征训练样本集对模型进行训练, 从而提高了预测精度。

综上所述, 本文提出的轻量化在线超短期光伏功率预测算法, 采用特征序列代替训练样本中的传统时间序列, 从而大幅度减少了训练样本对历史数据量的依赖性, 可以用较少天数的数据量得到较好的效果; 采用基于趋势加权相似度的训练样本精选方法, 在确保算法预测精度的前提下, 简化了算法计算量, 大大降低了算法计算时间. 因此, 本文提出的轻量化在线超短期光伏功率预测算法非常适用于低成本微电网的应用。

5 结论(Conclusions)

针对低成本用户侧微电网中的如下问题: 1) 无高性能监控服务器, 系统计算能力与计算资源受限; 2) 信息资源获取渠道单一, 无低成本数值天气预报, 甚至没有任何天气信息; 3) 数据存储空间和数据存取速度受限, 不适合进行在线大数据特征分析与挖掘等, 本文提出了一种轻量化设计的在线超短期光伏功率预测算法. 为了保证预测精度, 预测算法中选用了具备快速学习能力的核函数极限学习机算法, 同时结合特征序列训练样本集和基于趋势加权相似度精选训练样本的方法, 实现了预测算法的轻量化设计. 经过实验测试验证, 本文提出的轻量化在线超短期光伏功率预测算法能够适应各种天气变化, 在预测精度、计算时间和占用内存等方面具有较好的效果, 能够满足低成本用户侧微电网的应用需求。

参考文献(References):

- [1] QIAN Zhen, CAI Shibo, GU Yuqing, et al. Review of PV power generation prediction [J]. *Journal of Mechanical & Electrical Engineering*, 2015, 32(5): 651 – 659.
(钱振, 蔡世波, 顾宇庆, 等. 光伏发电功率预测方法研究综述 [J]. *机电工程*, 2015, 32(5): 651 – 659.)
- [2] BOB D. BP statistical review of world energy 2014 [EB/OL]. [2014-6-18]. <http://www.bp.com/en/global/corporate/about-bp/energy-economics/statistical-review-of-world-energy.html>.
- [3] CHINE W, MELLIT A, PAVAN A M, et al. Fault detection method for grid-connected photovoltaic plants [J]. *Renewable Energy*, 2014, 66(10): 99 – 110.
- [4] WU Hongbin, CAI Liang. Economic dispatching of renewable energy distribution generation systems [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2010, 26(12): 287 – 292.

- (吴红斌, 蔡亮. 可再生能源分布式发电系统的经济调度 [J]. 农业工程学报, 2010, 26(12): 287 – 292.)
- [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.
(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.)
- [6] LAN Hua, LIAO Zhimin, ZHAO Yang. ARMA model of the solar power station based on output prediction [J]. *Electrical Measurement & Instrumentation*, 2011, 48 (542): 31 – 35.
(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(542):31 – 35.)
- [7] GAO Yang, ZHANG Biling, MAO Jingli, et al. Machine learning-based adaptive very-short-term forecast model for photovoltaic power [J]. *Power System Technology*, 2015, 39(2): 307 – 311.
(高阳, 张碧玲, 毛京丽, 等. 基于机器学习的自适应光伏超短期出力预测模型 [J]. 电网技术, 2015, 39(2): 307 – 311.)
- [8] IGI E, OZTOPAL A, YERLIB B, et al. Short-mid-term solar power prediction by using artificial neural networks [J]. *Solar Energy*, 2012, 86(2): 725 – 733.
- [9] LIU Nian, ZHANG Qingxin, LIU Haitao. Online short-term load forecasting based on ELM with kernel algorithm in micro-grid environment [J]. *Transactions of China Electrotechnical Society*, 2015, 30(8): 218 – 224.
(刘念, 张清鑫, 刘海涛. 基于核函数极限学习机的微电网短期负荷预测方法 [J]. 电工技术学报, 2015, 30(8): 218 – 224.)
- [10] ZHU Yongqiang, TIAN Jun. Application of least square support vector machine in photovoltaic power forecasting [J]. *Power System Technology*, 2011, 35(7): 54 – 59.
(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)
- [11] 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.
(陈志宝, 丁杰, 周海, 等. 地基云图结合径向基函数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.)
- [12] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications [J]. *Neurocomputing*, 2006, 70(1): 489 – 501.
- [13] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classification [J]. *Systems, Man, and Cybernetics, Part B: IEEE Transactions on Cybernetics*, 2012, 42(2): 513 – 529.
- [14] HUANG G B, WANG D H, LAN Y. Extreme learning machines: a survey [J]. *International Journal of Machine Learning and Cybernetics*, 2011, 2(2): 107 – 122.
- [15] SERRE D. *Matrices: Theory and Applications* [M]. Berlin: Springer, 2010.
- [16] TOH K A. Deterministic neural classification [J]. *Neural Computation*, 2008, 20(6): 1565 – 1595.

作者简介:

陈晓科 (1978–), 男, 高级工程师, 主要研究方向为电力系统分析与电网规划, E-mail: 1594496644@qq.com;

陈奇芳 (1986–), 男, 博士研究生, 主要研究方向为光伏功率预测, E-mail: amiqicqf@163.com;

何婷 (1993–), 女, 硕士研究生, 主要研究方向为主动配电网优化运行, E-mail: 136487166@qq.com;

黄锦成 (1984–), 男, 高级工程师, 主要研究方向为微电网规划设计, E-mail: 546512173@163.com.