

DOI: 10.19783/j.cnki.pspc.180888

基于数据驱动的超短期风电功率预测综述

杨茂, 张罗宾

(东北电力大学电气工程学院, 吉林 吉林 132012)

摘要: 以数据驱动为主要特征的超短期风功率预测是大规模风电并网运行的关键基础之一。按照预测流程, 从数据挖掘、机器学习算法及风速-功率曲线等角度分析现有数据驱动方法的思想及局限性。总结离线数据驱动/深度学习算法和在线应用的预测思路, 给出风电场数据筛选的评价手段, 归纳深度学习算法的最新研究进展。最后分析超短期风功率预测的当前定位: “由模型驱动向数据驱动过渡, 由机器学习算法向深度学习算法转移”, 并指出合理的算法更迭和深层次的数据融合将是未来的研究趋势。

关键词: 数据驱动; 超短期预测; 风电功率; 多源数据融合; 深度学习

Review on ultra-short term wind power forecasting based on data-driven approach

YANG Mao, ZHANG Luobin

(School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China)

Abstract: Ultra-short-term wind power forecasting based on data-driven approach is one of the key foundations when large scale wind power integrated into the power grid. According to the forecasting process, the basic thoughts and limitations of existing methods are analyzed from the point of view of data mining approach, machine learning algorithms and wind power curve. Furthermore, the new prediction idea of offline data-driven/deep learning algorithms and online application is concluded, the evaluation methods of information screening are given, the latest research progress of deep learning algorithms in data-driven forecasting is summarized. Finally, the current position of ultra-short term wind power forecasting is summarized, that is transition from model driven to data-driven and transfer from machine learning to deep learning, and it is pointed out that the alternation of algorithms and data fusion will be the research trends in the future.

This work is supported by National Key Research and Development Program of China (No. 2018YFB0904200).

Key words: data driven; ultra-short-time forecasting; wind power; multi-source data fusion; deep learning

0 引言

截止 2018 年初, 中国风电累计并网容量达 1.68 亿 kW, 其中新增并网容量 394 万 kW, 全国风电发电量同比增长 39%^[1]。风电本身具有随机性、波动性和间歇性^[2], 难为系统所调度, 呈现出显著的反调峰性^[3], 对电力系统的稳定运行施加了无序冲击^[4-5]。准确的超短期风电功率预测(提前 15 min~4 h 的滚动多步预测^[6]) 在调度部门制定发电计划、调整运行策略等方面发挥了关键作用, 为电力市场交易提供了基础信息, 利于保持系统实时功率平衡^[7-8]。

传统风电功率预测为突出主要矛盾, 规避次要因素的影响, 多采用仅反映预测对象某一方面特征

的方法去进行功率预测, 比如以时间序列法、持续法、卡尔曼滤波法^[9]为代表的统计法, 其预测精度随着预测时间的增加而递减; 以神经网络、支持向量机、随机森林法^[4,10]为代表的机器学习法(即模型驱动法, Model to Data, M2D), 该类方法通过预测模型建立起输入与输出之间的映射关系, 是目前应用广泛的预测方法, 但预测思想的“黑箱性”难以克服, 依赖于预测模型的可靠度; 以空间相关性^[11]和数值天气预报(Numerical Weather Prediction, NWP)的降尺度处理^[12]为代表的机理驱动法, 机理预测过于聚焦物理信息, 欠缺对历史数据的利用^[13-14]。

网络的复杂度和样本规模之间始终是一对矛盾, 在样本数量保持不变的条件下, 网络规模与其泛化能力成反比。深度学习(Deep Learning, DL)算法凭借数据特征提取等优势, 被广泛应用于图像处理、

语音识别和自动驾驶等领域。DL 构造出一个多层网络结构来逼近复杂非线性函数，可以提取并学习输入样本的深层隐含特征，其中的卷积神经网络(Convolutional Neural Networks, CNN)和深度置信网络(Deep Belief Networks, DBN)等已在电力系统各领域的研究中取得了阶段性成果^[15-16]。针对机理预测法的不足，并考虑到风电场数据能够更加全面直观地反映输入参数对预测目标的影响，因此结合深度学习优势的数据驱动法成为超短期风功率预测的研究热点^[2,14,17]。

具体来看，数据驱动预测与机理预测之间存在着本质区别，如图 1 所示。首先是立足点不同：机理驱动方法建立在对大气运动刻画的基础之上，具有明确的风电转化机理^[12]；数据驱动预测基于多源数据信息(详见表 1)，核心为从已知数据集到未知预测值(Data to Data, D2D)的推导。第二是预测流程不同：机理预测法结合风电场的局地效应，利用计算流体力学(Computational Fluid Dynamics, CFD)^[12,18]或尾流模型^[18-19]对中尺度 NWP 进行降尺度处理，将其转化为轮毂处风速^[18]，通过功率曲线实现风电转换^[12]；数据驱动法先对数据进行预处理、挖掘，再将其输入至算法^[20-21]中从而获得预测值。第三是关键点不同：机理驱动重视对物理求解规则的优化以及 NWP 的质量^[4,22]；数据驱动强调多算法组合方式^[17]和信息融合手段^[4,18,23]，评估多时空场景下的数据对提升预测精度的贡献率。第四是时效性不同：机理预测法受制于数值模拟环节计算量巨大^[24-25]，一般不用于超短期预测^[18]；数据驱动法则更加适合超短期风功率预测^[14]。同时，二者也存在联系，数据驱动正逐渐渗透到机理驱动的若干环节之中，如机理预测中的功率曲线建模^[12]过程本身也具有数据驱动的特征，将在 2.1.3 节详述。

本文综述基于数据驱动的超短期风功率预测的

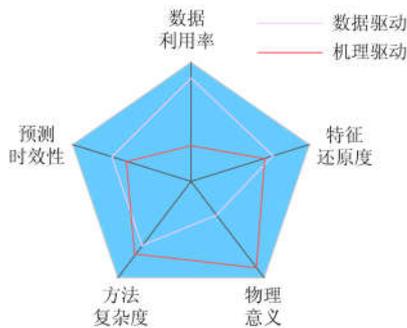


图 1 机理驱动与数据驱动的区别

Fig. 1 Contrast of mechanism-driven and data-driven

研究进展，总结 D2D 预测的框架：从输入样本的预处理方法再到预测模型，探讨了各方法的优缺点；梳理数据驱动预测研究现状，以相关系数作为信息筛选指标，以深度学习算法为预测工具，以算法组合方式的手段，归纳当前的数据驱动预测框架：“数据驱动(多源数据融合/深度学习算法组合)，在线应用”，指出预测平台已经从 Matlab 向 Python、Java 计算机语言等集群环境转移；最后对数据驱动在超短期风功率预测中的前景和局限性进行展望。

1 预测思想

1.1 基本概念

数据驱动预测就是以数据为导向的预测体系。从预测流程的角度看：分析风电场运行数据并剔除冗余数据^[14,20]，融合异质数据^[4,17]，挖掘输入条件的耦合机制及其与输出间的关联规则，从而实现输入到输出的转换，具体实现过程如图 2 所示；从预测思想的角度看：作为一种基于已知输入推导未知输出的功能结构^[13-14]，它既不是一个具体的模型，也不是某种数据分析手段，本质上是对风电场数据进行分析，进而实现 D2D 的协同工作体系；从应用条件的角度看，数据量激增为背景，深度学习算法为基础，计算机性能进步为推动力，是其在超短期风功率预测中的应用前提。

从数据集来看，数据驱动预测主要有两个特征。

1) 多源数据

外在表现为数据的体量大，来源广^[26]；内在表现为不同的数据源之间具有耦合性，如中尺度 NWP 与轮毂高度风速之间^[12,19]、风速与功率之间存在因果联系^[27]。体量大体现在：根据东北某大型风电场的长期实测结果显示，以 1 s 为两个相邻点的采样间隔，连续采集一个月的单机风速-功率数据的内存可达 25 MB 左右；来源广体现在：既有多个 NWP 数据(预报数据)，又有长期实测数据(历史数据)，即不完美的天气预报数据^[4]与相对可信的历史数据^[27]均作为信息源。

2) 数据质量

中国多数风电场位于偏远地区^[5]，加之弃风、限功率运行等问题较为严重^[3]，导致数据质量参差不齐、价值密度低^[26,28]，若直接将原始数据作为输入，可能会使模型训练失效，故需对原始数据进行预处理和挖掘，将在 2.1 节展开详细分析。

从预测框架来看，数据驱动预测主要有两个步骤。

1) 数据预处理与多源数据融合^[14,27]，将在 2.1.1 节与 4.1 节详述。

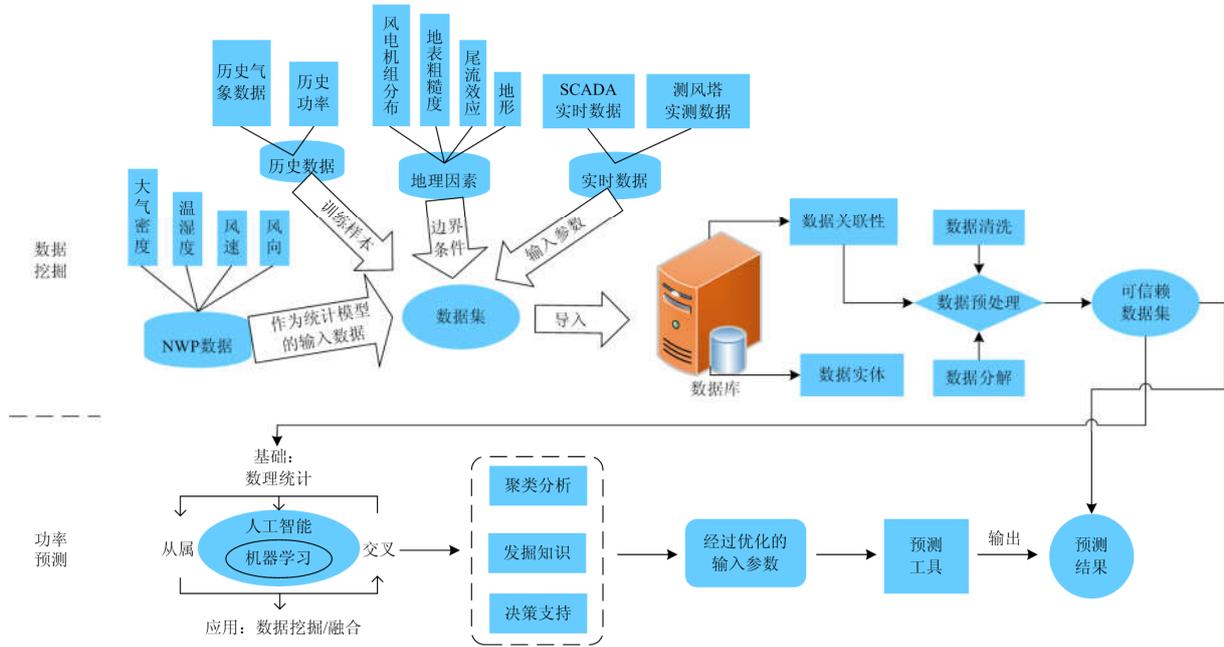


图 2 传统数据驱动预测流程

Fig. 2 Flow chart of traditional wind power forecasting based on data-driven approach

2) 形成从历史数据到未知数据的推导规则/多数据驱动算法组合以及深度学习算法的探索性应用^[13,20,29]。

表 1 风电场数据集

Table 1 Multi-source data sets of wind farms

数据分类	数据内容
地理信息	地形起伏、地表粗糙度、风机布局、流场畸变、风剪切等
历史数据	(历史上的)风速、风向、温度、湿度、大气压强等实测数据 (历史上的)风电功率
NWP 数据	风速、风向、温度、湿度、大气压强等气象预报数据
实时数据	气象站提供的实时天气信息 数字系统(SCADA)采集的实时功率数据

1.2 要素描述

风电场数据包括历史风速、功率以及 NWP 数据^[4,21,28], 本文将其统称为异质数据, 如表 1 所示。

1.2.1 风能

受大气运动的影响^[18], 风速序列在时间、空间上呈现出非稳定性和非线性的特点, 同时风向也对预测效果有一定的影响。文献[30]提出一种复数神经网络(Complex-Valued Neural Network, CVNN), 以风矢量(包括风速和风向)代替实数数据作为预测模型的输入, 仿真结果表明其预测准确率超过了实值神经网络的预测效果。

风速的骤变固然难以预测^[31], 但风速的周期变化、波动特性却有规律可循, 在固有分量的基础上对局部信号细分, 即如何对高频处的时间、低频处的频率进行划分将是下一步的研究方向。

1.2.2 风电功率

原动力间歇性决定了输出功率的非平稳性^[13,32], 现有研究集中于概率预测^[33]与波动性刻画等领域。

1) 概率预测

考虑到在风速向风功率的转化过程中存在检修、限电等人为因素或气象条件变化的干扰, 导致在相近的来流条件下可能输出不同的功率, 因此传统的确定性预测无法反映风功率的真实概率分布情况^[34-35]。风速-功率曲线建模是减小风电转化误差的途径之一, 详见 2.1.3 节, 同时区间预测^[36]亦是研究热点, 预测流程如下所述。

(1) 数据特征分析: 文献[37]考虑到影响功率预测的多样性因素, 利用 Copula 函数反映预测值与输入样本的相关性, 可使预测结果位于其真实分布区间的概率最大化。

(2) 区间预测: 基于数据间的联系与信号的非线性特征, 建立起相应的分布函数, 在满足一定的置信度下, 给定目标值所处的区间界限。贝叶斯网络(Bayesian Network, BN)模型提供了一个对不确定数据表达和推导的框架, 文献[38]由此提出了用于超短期风功率预测的自适应贝叶斯学习和高斯过程近似相结合的方法, 贝叶斯框架能够进行区间预测,

可更新计算模型权重的后验概率分布,再基于概率原理来刻画功率的非平稳性;高斯近似过程则解决了BN学习的集成问题^[39-40]。

2) 波动性刻画

文献[41]采用预处理后的风功率数据,从日、季节等时间角度分析其波动性;从空间尺度研究多个风电场出力的相关性、同时率,得出多时空尺度下的风电场出力特性:时间间隔越大,出力波动性越强;空间尺度越大,出力越平滑。文献[42]基于历史数据,给出归一化风功率波动变化率;基于Copula函数拟合风电场功率波动相关性的联合概率分布,结合蒙特卡洛模拟未来15 min内可能发生的波动组合,可提前估计全网风险趋势。

1.2.3 NWP与地理因素

NWP是在一定的初值及边界条件的约束下,通过数值方法求解大气运动方程来获得天气预报信息的技术^[43],但其解析值无法直接求取,难以精确描绘真实的大气运动^[4]。CFD作为一种将偏微分方程离散化为代数方程,利用计算机求解流场分布的数值计算方法^[18],其中涉及的动量方程的求解和边界条件参数的调整也与数据驱动密切相关。此外基于数据驱动的参数模型已被用于优化控制^[44],以最大限度地提高风电场的输出功率。

1.3 问题描述

从聚焦的问题看,数据驱动预测重点解决两个问题。

1) 剔除冗余数据^[24]后的特征提取^[14,20,26]及多时空尺度的数据融合,以加快算法训练的收敛速度,避免陷入局部最优解。

2) 预测时效性:考虑到未来中国电力市场的现货交易很可能以小时为基本交易时限^[2,7,45],因此在超短期风功率预测中,如何最大程度地发挥风电场数据集的作用取决于是否能够做到数据的快速处理、场景的快速识别以及模型的快速匹配等,相关研究表明^[14]:对于超短期风功率预测而言,数据驱动预测法比物理预测法在实现速度、预测效率上更具有优势。

2 现有预测策略

2.1 基本思想

以预测流程为主线,划分为数据预处理、预测算法与模型、风电转换三个模块分别进行介绍。

2.1.1 数据预处理/挖掘

处理方法主要包括:a)信号处理法,小波变换(Wavelet Transform, WT)^[46]、经验模态分解(Empirical Mode Decomposition, EMD)^[47]、变分模态分解

(Variation Mode Decomposition, VMD)、原子稀疏分解(Atoms Sparse Decomposition, ASD)、谱分析法(Singular Spectrum Analysis, SSA)^[48]等,均可作为对数据进行预处理并从中剥离出信息趋势^[20]的技术;b)统计分析法,一种基于统计学的历史数据处理方法^[4],常与其他方法结合使用。本节按照数据挖掘目的和效果对处理方法进行分类,这一点与其他数据驱动综述不同。

1) 剔除冗余/分解

(1) 异常数据剔除/缺失补齐

由于存在信息采集的技术性问题或人为因素干预^[28],常会产生异常数据或数据缺失。文献[20]指出冗余数据的存在或对预测效果产生不良影响,主要原因在于数据特征的重叠。文献[49]针对风电场运行过程中产生的不良数据,提出一种基于Copula理论的置信等效边界模型,进行风速分区间的异常数据清洗。文献[51]利用参数化的选择标准来甄别并剔除异常数据,该法虽简单易行,但在不同情况下需设置对应的参数,通用性有待加强。

(2) 序列分解

风速、功率序列相当于一系列不平稳信号,可采用信号分析法对其进行平稳化处理,以获得若干具有相似特征的子集,使分解后的序列有规律可循^[28]。

EMD:从本质上讲,EMD是将风速、功率等数据中各个时间尺度的波动逐级分解开来,从而产生一系列具有不同特征尺度的本征模函数(Intrinsic Mode Function, IMF)分量,进而弱化不同子序列之间的相互影响,与WT需选取小波函数相比,EMD因其自适应分解特性而更具有数据驱动的特征。

但EMD存在以下问题:a)风机输出功率的不确定性往往会形成局部的高频信号,利用EMD分析该异常数据时会产生IMF模态混叠,文献[47,50]利用集合经验模态分解(EEMD)将原始序列按照波动的幅度分解成若干个子集,并加入低频白噪声来消减模态混叠,其中文献[47]所提出的混合方法比传统的GA-BP预测方法和同时采用EMD、小波分解的GA-BP方法要更为精确,适合超短期功率预测,但该法对参数调整较为敏感。b)EEMD由于加入了白噪声,在序列重构过程中引入了误差,文献[51]采用完全集合经验模态分解(Complete EEMD, CEEMD),通过增加成对的正负白噪声来消除噪声^[52]。c)对于端点效应的抑制,文献[53]通过极值延拓法去消除端点失真现象。

文献[54]采用二次混合分解技术,在数据预处理阶段:采用EMD将原始序列分解为IMF,再通过WT把IMF继续分解为更精细的分量;训练阶段:

使用由交叉优化算法训练得到的极限学习机, 以缓解过早收敛的问题。

VMD: 文献[55]使用变分模态分解获取信号的分量, 结合样本熵算法将子序列重构为若干特征分量, 基于优化鸡群和极限学习机对分量进行单独建模并优化其参数。

ASD: 原子稀疏分解借鉴原子库的冗余性, 使信号分解系数体现出稀疏性, 可用于数据分解和自预测, 但只适用于小样本数据集。文献[56]借鉴 ASD 在处理非平稳信号的优势, 将其用于超短期风功率爬坡预测中, 实现了数据的简洁化表示。

2) 特征提取/聚类

(1) 谱分析法

谱分析是从频谱上分析风电功率时间序列的方法。文献[48]利用 SSA 提取混沌分量的趋势, 在超短期风功率预测中效果显著。

(2) 最大相关-最小冗余法

最大相关-最小冗余法(Maximum Relevance Mini-mum Redundancy, MRMR)基于评估多变量相关性的互信息理论, 核心思想为最大化信息的特征, 最小化特征之间的相关性, 保留信息特征的同时剔除冗余信息。文献[57]利用 MRMR 和互信息对高维数据进行特征提取(降维), 通过相关向量机建立特征与功率间的映射关系, 进行随机性缺失功率补齐, 由于 MRMR 获得的特征能够大致替代原始功率数据, 故补齐效果较好。

(3) 聚类

文献[58]将历史功率序列作为输入, 引入经验正交函数分解法, 提取区域内风电出力特征; 采用层次聚类法将具有相似出力特征的若干风电场整合为一个子区域, 其预测误差随着子区域数目增加而递减, 但该法对预测效果的提升存在上限, 可结合深度学习模型: 引入粗糙神经网络进行无监督的风速特征聚类学习^[59]。

2.1.2 预测算法

数据挖掘是形成具体预测算法的前提, 而算法以优化样本及融合信息为输入, 将历史数据按规则具体表现出来, 本节重点介绍广受关注的预测算法、模型(图 3)。

1) 神经网络^[23,27](Neural Network, NN): 本质是通过大量历史数据进行网络层间学习, 建立历史数据(输入样本)与预测功率(输出)间的映射关系, 主要包括 ANN、BP、RBF、Elman 等类型。文献[24]提出基于神经网络平均影响值的风电功率超短期预测方法, 引入动态神经网络, 通过考虑风速和风向的导数, 量化二者的变化趋势, 并计及各输入变量

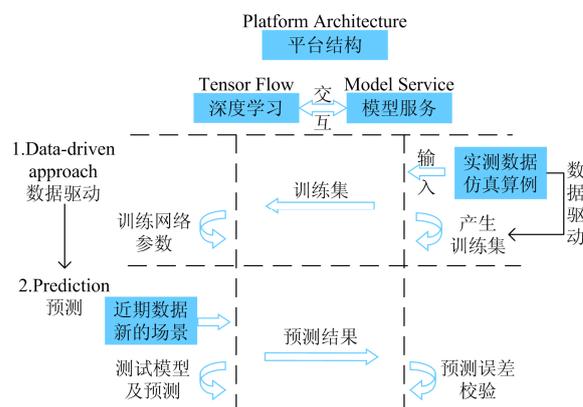


图 3 数据驱动预测模型

Fig. 3 Forecasting model based on data-driven approach

的平均影响值用来加快网络的训练速度, 预测结果表明: 该神经网络数据驱动方法的训练时间短、结构简单, 预测精度较高。

2) 支持向量机(Support Vector Machine, SVM):

将给定数据集映射至高维空间^[14,60], 与神经网络相比, 具备更好的泛化性能以及鲁棒性; 与模糊逻辑相比, 其在预测精度方面更具优势且不需经验知识^[61]; 缺点是核函数的选取困难, 对训练样本需求量^[62]。

3) 数据挖掘算法: k-means 聚类^[63], 随机森林法^[20](Random Forest, RF), 模糊逻辑^[64], 多层感知器(Multilayer Perceptron, MLP)等。a) k-means 聚类分析根据样本的内在差异进行聚类划分^[65], 利用“距离”来刻画场景间的差异, 实现场景削减, 但该算法对异常数据较为敏感; 如何确定划分标准和聚类特征^[14], 将是下一步的研究重点。b) 随机森林算法^[20]是一种集成学习算法, 作为一系列的决策树的集合, 其输出是所有回归树的预测值的平均值; 采用自助重采样技术, 克服了回归树的过拟合问题; 能够处理高维数据, 适用于数值型变量和类别型变量; 可并行化处理, 适合大数据集; 影响 RF 性能的主要因素有单棵树的预测强度和树与树之间的相关性。c) 模糊逻辑是一种评价变量不确定性的数学工具, 能够简化信息、进行经验学习和推导不确定性知识, 从而提取风速波动的关键属性^[66]。d) MLP 通过模拟人类的逻辑思维, 建立从输入到输出的映射关系, 可处理不确定性数据。

4) 基于优化算法的网络参数选取: 为获得预测算法的最优参数, 一般采用寻优算法对参数进行优化求解, 如遗传算法^[67](Genetic Algorithm, GA)、布谷鸟算法^[68](Cuckoo Search Algorithm, CS)、模拟退火算法、粒子群优化(Particle Swarm Optimization,

PSO)、萤火虫算法^[69](Firefly Algorithm, FA)、鸡群算法^[55]。上述算法借鉴仿生学原理,将生物或物理现象的行为特征移植到风功率预测算法的参数寻优、权重分配领域。文献[70]提出 MLP 与萤火虫优化算法(Firefly Optimizer Algorithm, FFA)相结合的混合模型(MLP-FFA);基于一组月度历史数据,以均方根误差(RMSE)和平均绝对误差(MAE)等统计误差为评价标准,验证了 MLP-FFA 的预测性能优于传统 MLP 算法,其原因可能是在混合智能中利用 FFA 作为附加优化器,故能在隐含层中获得最佳权重,从而使特征提取过程更具鲁棒性。

基本的 PSO 算法易陷入局部最优,可引入极值变异因子,迫使粒子在最优位置长时间停留进而变异,避免粒子的早熟现象;标准 CS 算法中的固定发现概率会降低寻优效率,可引入动态发现概率,使后期产生更多的新个体,以克服陷入局部最优解等问题^[62]。

2.1.3 风速-功率曲线建模

功率曲线的本质是通过拟合实测数据来反映风速-功率转换关系^[4,71]。

风功率曲线建模主要分为:离散方法,随机方法,参数方法以及非参数方法^[72]。离散方法将一个离散过程用连续过程近似代替;随机方法通过构造一个带参数的随机过程或者概率模型,使其参数近似等于问题的解;参数方法^[73]则是通过建立数学表达式来形成输入与输出间的映射关系。与前三种建模方法不同,非参数方法更具有数据驱动的特征,它基于大量历史数据来建模,能够较好地反映风电机组的运行状态,另外,一些人工智能法与数据挖掘方法也已用于功率曲线建模。

文献[74]考虑相邻风机的尾流效应,基于贝叶斯框架来构建多个相互作用风机的功率曲线,量化了模型参数。文献[71]对预处理的数据进行分区,选择合适的数据挖掘算法来建立动态功率曲线:数据分区用于降低模型复杂度;数据挖掘则用于构建非参数模型,建模步骤如下所述。

步骤 1:原始数据预处理;

步骤 2:将分区的数量设置为 n ,并分解风速范围至 N 分区,对应数据集的第 i 分区被标记为 S_i ;

步骤 3:将分区的质心数据点提取为新的分区数据集,基于不同核函数的 SVM 算法建立功率曲线。

2.2 典例分析

文献[46]提出以目标预测误差为阈值,利用带扰动 PSO 算法对各个频段的风速所对应的 Elman 神经网络加速训练,并基于此改进模型进行变频风

速预测,最后将不同频段的风速重构为预测风速,通过功率曲线实现超短期功率预测。该文献从小波包分解(数据预处理),到结合 PSO 寻优算法的改进 Elman 神经网络(机器学习算法),再到功率曲线(风电转换),较为完整地体现了数据驱动在预测首尾部分中的优势,即引入 WT 对风速数据进行处理,剔除了风速波动部分;功率曲线建模则实现了从风速到功率的转化。但各学者在直接预测^[12](从历史功率到预测功率)与间接预测^[12,22,27](先预测风速,再进行风电转换)的预测效果上存在分歧。

2.3 现有方法局限性

由上述分析可知,以传统信号分析学和预测算法为基础的传统数据驱动模型通用性较差,本质上是模型驱动存在天然缺陷,割裂了数据之间的耦合关系^[17],局限性总结如下所述。

1) 传统方法基于历史数据来学习风速、功率的内在规律,并执行预测机制,算法的运行速度较快,时效性更高^[14]。然而,它需要可信的历史数据集,对历史数据依赖性强,运算过程具有“黑箱性”^[75],预测精度与数据质量存在较强的因果关系,很少与空间延迟相关性等物理因素结合。

2) 不可避免地需要做出假设:传统方法大多显式或隐式地设置了某些规则和假设,如小波分解、经验模态分解假设原序列是若干子序列的叠加,原子稀疏分解也是在用表达式近似代替子序列趋势的前提之下来使用;高斯近似假设随机过程服从高斯分布^[39,76];区间预测和概率预测也通常假设预测误差服从某种概率分布^[2,35]。

3) 文献[22]提出将数据驱动与因果驱动相结合的方法,挖掘隐藏在历史数据间的相互关系,克服了机理模型在信息不足时所做出的主观假设问题,但是两种方法的融合细节以及如何从多源数据中挖掘深层特征仍是难点。

4) 尽管机器学习算法可完成信息的部分挖掘,但随着样本数量及样本类型的增加,在算法的隐含层节点数、网络层次、连接方式、参数设置等方面仍需进一步优化。文献[77]的研究结果表明:结合所有预测的平均绝对导数和精度的双目标优化步骤,可调整 SVM 超参数来进行多步预测。

3 新型预测框架

3.1 新思路

为进一步提升风功率预测的精度,扩展数据驱动的思路,本节总结新型 D2D 预测思路,即“离线数据驱动/深度学习算法/组合, NWP 融合,在线应用”的预测框架。

3.1.1 离线数据驱动/多算法组合

基于大数据的思维, 实现“离线数据驱动”: 文献[13,20]指出数据预处理以及冗余信息剔除、有效特征提取、样本相关性衡量对形成高质量的训练集具有重要意义。

波动的风功率序列会增加预测结果的不稳定性^[40], 单一预测模型难以全面提取功率波动特性^[75,78]。目前的模型选择方式有两种。1) 组合预测法^[14,20,79]: a) 基于组合方法的数据预处理; b) 模型优化组合方法; c) 基于组合方法的数据后处理; d) 基于权重系数的组合方法, 文献[21]提出基于历史数据与新观测信息的权重配比的自适应指数动态组合模型, 引入动态自适应指数能够在一定程度上反映数据的时变性, 组合模型的应用使得某些预测点的误差变小, 采用多机器学习算法的组合^[14,20]进行功率预测, 通过模型间的功能互补, 取得更具鲁棒性的预测结果, 其关键在于组合权重的选取^[4]。2) 深度学习算法, 目前主要采用深度神经网络法, 将在 3.2 节重点介绍。

3.1.2 NWP 校验

基于历史数据的离线数据驱动忽略了风速/功率序列的近期变化趋势, 加入 NWP 或有助于补充实时数据的特征信息^[81]。

主流融合方式有两种: 1) 多 NWP 耦合, 选取多家 NWP 预报数据, 分别建立预测模型, 文献[78]以气象预报部门的风速预测数据为基础, 建立基于多支持向量回归的确定性风电功率预测方法; 引入平均权值组合法确定组合权重; 借助来源不同的 NWP 数据来耦合误差特性, 提高预测精度。2) NWP 校验: 引入 NWP 实现对预测结果的预判和检验^[42], 本质是与由数据驱动形成的预测模式进行融合^[17,79], 虽然 NWP 的预报精度不高(可结合模糊聚类与关联规则进行修正^[82]), 但对风速骤变的预测效果较为可信^[4], 可利用 NWP 来校正预测结果的拐点误差^[43], 异质数据的加入计及了更多的可用信息, 能够在一定程度上克服单纯从历史数据视角进行预测的弊端。

3.1.3 在线应用

在线预测目标时刻的风速或功率时, 根据数据预处理得到的不同样本特征, 选择对应的预测算法或参数, 最后按照时序或其他规则重构并输出预测结果。

3.2 深度学习算法

卷积神经网络、深度置信网络、长短期记忆网络(Long Short Term Memory, LSTM)、GRU 网络、堆

叠降噪自动编码器(Stacked Denoising Autoencoder, SDAE)等深度学习算法已用于变压器和电气设备的故障诊断领域。深度学习算法在特征提取与重构方面展现出优越的性能, 可基于风电场的海量数据来学习特征信息, 深度学习与机器学习的区别在于:

1) 深度学习模型自身内嵌了数据特征提取模块, 输入样本为原始数据, 保留了数据的全面信息; 而机器学习则需结合信号分析法获取输入样本, 输入样本为经过预处理后的数据, 构建样本的过程中常隐含了某种假设。

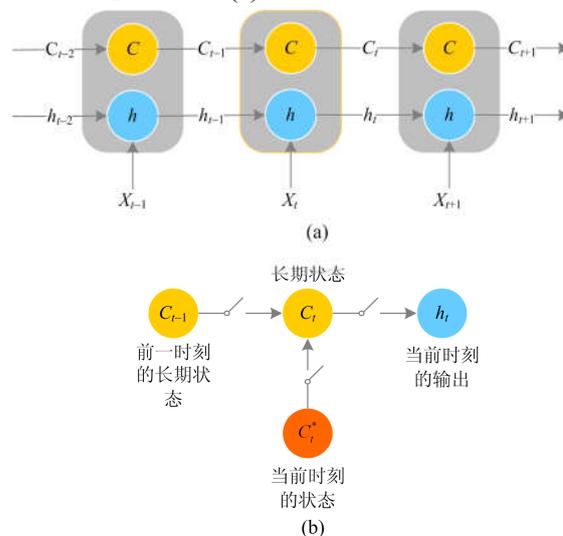
2) DL 算法的训练时间长, 测试时间较短, 这点与机器学习算法相反, 更符合超短期预测的上报时限要求; 基于 Python 的 keras 所搭建的 DL 预测模块时效性一般优于 Matlab 平台。

3) 由于具有无监督学习能力和多层非线性结构, 相比机器学习算法, DL 算法在逼近复杂函数、拟合输入样本和提取抽象特征等方面显示出了优越性能。

3.2.1 门控神经网络

1) 长短期记忆网络

LSTM 是一种特殊的循环神经网络(Recurrent Neural Networks, RNN), 它们均假设输入样本是连续的时间序列, 但当序列过长时, 中间的无用信息可能会使 RNN 出现梯度消失现象。从结构角度看: LSTM 起到网络结构更为精简; 从输入输出角度看: 在任意时刻 t , LSTM 有 3 个输入, 2 个输出, 其输入、输出结构如图 4(a)所示。LSTM 的关键在于开关状态的控制策略, 如图 4(b)所示, 开关状态的算法实现又取决于门, 具体控制对象和功能如表 2 所示, 控制过程如图 4(c)所示。



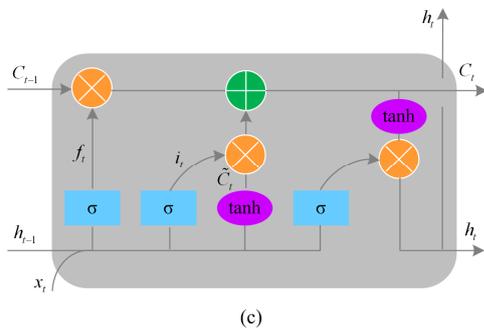


图 4 LSTM 基本结构
Fig. 4 Structure of LSTM

表 2 LSTM 门控单元

Table 2 Door control unit of LSTM

结构类型	作用对象	功能
输入门 (input gate)	细胞状态	①选择性地记录(通过)新的信息; ②用新知识来代替旧的认知
遗忘门 (forget gate)	细胞状态	①选择性地忘记过往信息; ②激活后, 将遗忘过去的细胞状态 ③弱化(减小)过往信息对当前预测体系的(所施加的影响)作用
输出门 (output gate)	隐含层	①将之前的信息保存至隐层中; ②激活后, 最新的单元输出将被传播到最终状态; ③输出与之前信息相关的内容

文献[83]基于深度学习算法的长期记忆网络训练低频子分量, 利用 Elman 递归神经网络训练高频子分量。文献[17]一方面运用 Pearson 距离分析法^[84]来评价样本与功率之间的相关程度, 实现“剔除低相关变量、保留高相关变量”功能; 另一方面基于 LSTM 对时间序列进行建模, 能够有效利用多源信息, 因此比 ANN 和 SVM 等算法具有更高的预测精度。

2) GRU 网络

GRU 也是 RNN 的一种变体, 通过添加重置门和更新门来定量刻画远期输入对近期的影响。文献[85]由此提出一种 CNN 和 GRU 相结合的深度神经网络模型, 其中 CNN 用于数据“降维”, GRU 用于提取数据的深度特征; 为克服过拟合现象, 引入 dropout 技术, 以一定的概率随机丢弃神经元, 使输入组合方式多样化。

3.2.2 卷积神经网络

CNN 最初作为一种处理视觉图像的神经网络, 其核心结构为卷积层和池化层。每个卷积层中都含有特征捕捉器, 具体结构如图 5 所示, 卷积层之间穿插可防止过拟合的池化层; 样本输入卷积层后, 特征探测器会以滑动窗口的形式进行卷积运算, 产

生相应的特征图谱, 输入样本中隐含的深层特征便会随之提取出来。

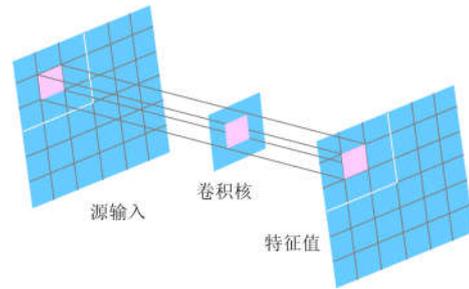


图 5 CNN 卷积层结构

Fig. 5 Convolution layer structure of CNN

文献[13]提出基于 WN、CNN 和集成技术的深度学习模型, 以量化风电不确定性; 基于该定量模型进行数据噪声筛选, 利用小波变换将原始数据分解为不同的频率, 再引入卷积神经网络对每个频段的非线性特征进行有针对性的学习。

3.2.3 堆叠降噪自动编码器

SDAE 作为一种较为成熟的 DL 算法, 其本质是为加强泛化能力, 在自编码器(AE)的基础上增加了噪声构成的。它包括基于 AE 的无监督学习模块和用于数据拟合的逻辑回归层, 主要分为“逐层训练”和“微调”两个训练步骤, 通过“多输入少输出”的维度限制, 迫使 AE 从中捕捉最重要的特征, 以压缩输入数据。但随着神经元结构的堆叠和维度的扩展, 其全局最优的搜索能力会有所下降。

文献[86]提出一种基于 SDAE 算法和多 NWP 的多对多映射网络, 建立起 NWP 与风机实测风速之间的映射关系, 但在隐含层节点数较大时, AE 的编解码过程或沦为对数据的单纯复制, 此时可对 AE 加入稀疏性限制。

3.2.4 深度置信网络

DBN 以受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)为组成元件, 可在保持样本原始特征的前提下实现数据降维。RBM 作为一种随机神经网络, 由可见层 v 和隐含层 h 构成, v 与 h 之间保持双向全连接, 同层之间相互独立, 如图 6 所示。在信息输入 v 后, 计算每个 h 被开启的概率 P , 以 $(0, 1)$ 之间的某个实数作为阈值, 当 P 大于该阈值时, 神经元被激活, 否则保持关闭。在重构阶段, 底层 h 中的输出变为反向传递中上一层的输入, 经过 v 与 h 之间的多次正反向传播, v 能够显示 h 的特征并大致还原 h , 将若干层 RBM 串联起来进行无监督训练便构成了 DBN。

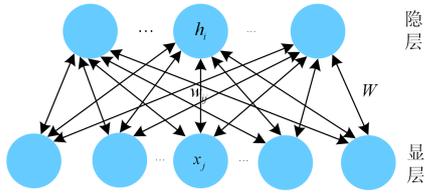


图 6 RBM 结构

Fig. 6 Structure of RBM

文献[87]基于深度 RBM 预测模型,从风速的底层特征中提取深度特征。文献[88]基于 DBN 网络作为递归预测算法,以逐层贪心算法作为预训练器,采用具有与 DBN 相同层数和隐含节点数的多层 BP 网络作为对比,进行多步滚动风速预测,验证了 DBN 具有较好的预测效果。

下一步可引入贝叶斯正则化深度置信网络 (Bayesian regularization-DBN, BR-DBN),其中 BR 可产生新的指标函数,能够削减有效权值的数量,简化了网络结构,进而提高无监督学习过程中的特征识别和泛化能力。

3.3 新特征

归纳新型数据驱动预测的特征如下所述。

1) 深度特征提取:考虑有条件的冗余数据剔除与数据特征选取,DL 算法中的特征提取用于总结数据集的一般特性。文献[14]指出预测模型构建过程中的冗余数据可能会导致所提取的“最佳特征”无法实现最佳预测效果,因此需识别具有较高相关性的不同特征。

2) D2D:既不提前对所构造的预测模型、数据分布形式做任何简化假设,也不需先验知识或人为设置参数,而是从数据本身出发去获得统计信息以及分布特征,进而去决定预测函数的映射关系、结构(Data to Model, D2M),最终实现 D2D,如自适应神经模糊推理系统^[14]以及非参数核密度估计^[43]非参数化建模并不预先假设点误差分布的表现形式:分位点回归、非参数核密度估计^[89]等直接计算出误差分布函数或分位点,该法不存在分布假设不合理问题,但缺点为需要大量已知数据,计算过程复杂^[90]。

3) 信息筛选:基于相关性系数刻画多变量间的关联度,寻找对待预测值影响最大的输入样本,表 3 给出了相关性度量方法与建议。如采用 Copula 函数模拟多个风电场的功率波动相关性^[33,42];贝叶斯网络则能够表示随机变量间的相依性^[39];欧式距离法^[91]、熵相关、秩相关等方法来衡量输入与输出间的相关性;采用聚类算法形成多个相似预测场景,建立适应不同场景的样本主成分提取模型,更好地适应运行工况的实时变化。

4) 时空数据:有两层含义, a)单一风电场的历史数据, SCADA 数据, NWP 数据等时序数据; b)具有时空相关性的毗邻风电场的空间数据集。基于此,考虑到数据集彼此之间天然地存在时空差,掩盖了数据间潜在的关联规则,因此挖掘数据间的时空相关性^[92],有利于预测下游风速的骤变^[22-23],弥补因 NWP 失效造成的预测误差。

基于数据推导的多时空尺度^[19,93]、多场景^[58]的风功率预测成为研究热点,如基于功率数据自身的变化特性及风电场运行状态,在时空尺度上分别采用“分层细化”思想,实现动态分群预测^[93];或者构建多出力场景,实现数据的特征聚类^[58,94];文献[13]验证了季节不确定性对预测精度的影响,发掘出相同类型数据在不同时序、空间位置^[25]上的演化轨迹以及不同类型数据之间的联系^[22,36,88],而后者往往无法通过统计方法或逻辑推导获得,其原因在于数据间的联系无法预知,故需进行深度融合、挖掘关联规则^[5,17,25]。

5) 逐渐开始结合风速物理特性与机组运行状态等:文献[95]为克服因不同风电场的数据特征各异而导致的预测模型失效问题,提出一种基于灰色关联分析和风速分布特征的混合预测模型,即最小二乘支持向量机(Least Square SVM, LSSVM)与径向基函数神经网络(Radial Basis Function NN, RBFNN)相结合,利用无量纲方法将风速序列转换为独立的标量,对每个风速段进行单独权重训练,根据月度风速段构建权重数据库和风速波动区段,验证了所述混合模型比单一模型更具稳健性。

6) 典例分析

(1) 基于深度特征提取的多组合模型^[21,78,93,95]

文献[20]提出了一种多模型预测方法,具备深度特征提取的功能,主要有两个贡献:①提出一种用于深度特征提取的框架来优化确定性预测方法的输入样本。质量过滤器用于数据预处理,其作用为:(a)通过纠正异常数据来实现噪声数据预处理;(b)通过将每分钟数据平均为小时数据来进行数据修整;②开发出用于风功率预测的多算法组合模型:结合了神经网络、支持向量机、梯度提升机和随机森林法等,实现各种模型间的优势互补,以期产生确定性和随机性内核。

(2) 预测典例

文献[75]提出一种预估-校正思路:基于风电场历史数据,采用 WT-NN 实现前处理环节,计及变化的环境因素,利用实时数据来修正并更新模型参数,能够在一定程度上提升下一采样点的预测准确度,该法的预测结果在风机出力变化频繁的情况下

比 BP、GABP 方法的平均相对误差更低,但其预估-校正环节也是在离线状态下进行的,同时在不同风机出力临界点处的预测效果有待提高。

文献[79]提出在选择混合预测模型的基础上,利用经验模态分解、二次分解和快速集成经验模态分解对原始数据进行去噪,以平滑风速时间序列信号中的噪声。文献[96]提出了一种基于时间序列的

k-means 聚类选择方法:一方面,采用集群选择方法对测风数据进行分类,然后通过小波变换对其进行分解,这为多层感知器神经网络提供了最佳输入;另一方面,选取最适于训练的集群,通过识别异常和不规则模式,加快神经网络的训练、预测过程,经不同风电场实测对比,验证了该法的通用性和可靠性。

表 3 输入信息筛选方法

Table 3 Methods of input information screening

文献	研究对象/目的	研究方法	方法定义	结论/建议
[17,84]	刻画风电功率数据与风电场同时期气象信息的相关性	Pearson 相关系数	Pearson 相关系数由用两个变量的协方差除以二者的标准差求得	当存在多个气象因素时,应辨识各个变量对目标点的影响程度;可度量输入样本与输出间的关系
[23]	两个及以上风电场数据集的关联性	深度学习迁移算法	DL 算法将从数据丰富的风电场获得的已知信息转移到新建风电场	若多个数据集之间存在一定的关联,那么其中一类数据就能通过其他样本进行推导 ^[55] ,利用已经掌握的数据信息弥补历史数据缺失的问题,保证遗漏数据与整体呈现一致性
[33,37]	研究相邻时段的风功率之间、功率与气象因素之间的关联关系	Copula 函数	Copula 函数通过反映多元变量的联合概率密度特性,挖掘数据间的相依关系,建立将多源信息关联起来的联合概率密度函数	采用 t-Copula 函数衡量相邻风电场出力波动相关性;采用 Copula 函数刻画多源变量的相关性,但风电功率与气象因素的联合 Copula 函数较复杂
[53]	不同历史日功率、风速信息与目标值的关联度分析	熵相关	1) 互信息:表征随机变量之间互相包含信息量的多少; 2) 熵相关:随机变量概率分布特征的广义相关性	熵相关系数能够识别多种形式的非线性相关关系,线性相关系数、秩相关系数与欧式距离法在具有严苛的相依性判据时,可能无法从历史信息中匹配到亲密样本,而熵相关由于宽松的评估指标,依旧可以捕获较多的关联样本
[66,91]	评价两个信号的相似程度	欧氏距离	两变量间的欧氏距离越小,则二者相似度越大	欧式距离的计算是基于各变量的归一化值,因此对两个单位不同的变量使用欧式距离可能失效,无法体现变量的差异度

4 展望及挑战

4.1 当前定位

经上述分析可知,数据规模扩大,深度学习算法涌现,组合策略更新,使得超短期风功率预测的重心从信号分析法结合机器学习算法向深度学习算法转移,对运行深度学习算法的软硬件要求又使得编程环境从 Matlab 向 Python、Spark 转移;进而从功率到功率的单一映射关系向多维数据对应模式转移。

4.2 关键

本节将数据驱动预测所面临的问题与研究关键点总结如下。

1) 新方法仍存在天然缺陷:如非参数核密度估计在本质上仍然属于模式匹配方法,不可避免地受到统计方法局限性(如过拟合、适应性差等问题)的

制约,若未能考虑各种出力模式或者出现风速骤变时,预测结果会出现相位误差、估计规则无法匹配等问题;多模型组合权重的选择以及组合方式的配置亦是难点,计及风的物理信息以及引入 NWP 校验或许是提升预测精度的途径之一。

2) 对待深度学习算法应持理性态度,现有的基于 DL 的风功率预测模型仍存在不足:第一,大多只使用了 DL 算法提取的最下层特征,导致了信息资源的浪费;第二,顶层设计的归类决策思路较为单一,如 sigmoid 函数的选取会对模型构建产生较大影响;第三,很少对输入样本的进行评估,可引入信息熵来确定输入之间、输入与输出之间的关联度,以突出主要影响因子。

3) 风电场数据的规模、合理性、预测对象规律性、预测方法的泛化能力及参数调整等,都会影响预测效果。目前,关于总数据集、训练集、测试集

和数据记录的采样间隔等信息规范长度的标准尚不统一,在未来的研究中需要对上述信息进行规范约束,并通过预测效果来评价该模型对数据集的敏感度。

4) 考虑到大多数的预测方案均为时效性有待提高的离线分析,而实现在线预测的关键是发挥广域数据的作用,将用户为中心的交互式设计方式和理念用于风电功率预测平台的设计,在提高硬件计算能力的同时,分享开放的多源数据支持也是关键所在。

5 结论

本文综述基于 D2D 的超短期风功率预测,剖析了数理统计-数据挖掘-大数据-人工智能的数据科学在超短期风功率预测逐步深化应用的进程中体现出来的科学思想,并将未来的研究重点总结如下。

1) 随着预测工作的逐步开展,本研究领域形成了从最初寻找合适的统计学分析和预测方法,到将数据挖掘及机器学习算法应用于超短期风功率预测,再到发现信号分析法和传统算法相结合的思路存在局限性,以及最新的深度学习算法在预测中的渗透这一完整清晰的脉络。

2) 对于 D2D 预测而言,数据挖掘手段和衡量方法、多算法组合策略及 NWP 对实时信息的补充策略,应当为研究人员所重视,一般可从算法的选择及数据融合这两个维度去提升预测精度。

3) 超短期风功率预测正处于由统计科学向人工智能转移的过渡阶段,DL 算法对编程语言和计算机环境的要求也发生了变化,如何把握统计学和算法的联系与结合点,在理性使用 DL 算法的基础上考虑更多的机理本质,以更好地统筹统计学和动力学在预测中的应用,这将是下一步的研究重点。

参考文献

- [1] 国家能源局. 2018 年一季度风电并网运行情况[EB/OL]. http://www.nea.gov.cn/2018-04/28/c_137143494.htm.
- [2] 吴问足, 乔颖, 鲁宗相, 等. 风电功率概率预测方法及展望[J]. 电力系统自动化, 2017, 41(18): 167-175.
WU Wenzu, QIAO Ying, LU Zongxiang, et al. Methods and prospects for probabilistic forecasting of wind power [J]. Automation of Electric Power Systems, 2017, 41(18): 167-175.
- [3] 舒印彪, 张智刚, 郭剑波, 等. 新能源消纳关键因素分析及解决措施研究[J]. 中国电机工程学报, 2017, 37(1): 1-9.
SHU Yinbiao, ZHANG Zhigang, GUO Jianbo, et al. Study on key factors and solution of renewable energy accommodation[J]. Proceedings of the CSEE, 2017, 37(1): 1-9.
- [4] 乔颖, 鲁宗相, 闵勇. 提高风电功率预测精度的方法[J]. 电网技术, 2017, 41(10): 3261-3269.
QIAO Ying, LU Zongxiang, MIN Yong. Research & application of raising wind power prediction accuracy[J]. Power System Technology, 2017, 41(10): 3261-3269.
- [5] 康重庆, 姚良忠. 高比例可再生能源电力系统的关键科学问题与理论研究框架[J]. 电力系统自动化, 2017, 41(9): 2-11.
KANG Chongqing, YAO Liangzhong. Key scientific issues and theoretical research framework for power systems with high proportion of renewable energy[J]. Automation of Electric Power Systems, 2017, 41(9): 2-11.
- [6] 杨茂, 陈郁林. 基于 EMD 分解和集对分析的风电功率实时预测[J]. 电工技术学报, 2016, 31(21): 86-93.
YANG Mao, CHEN Yulin. Real-time prediction for wind power based on EMD and set pair analysis [J]. Transactions of China Electrotechnical Society, 2016, 31(21): 86-93.
- [7] ZHOU Z, XIONG F, HJUANG B, et al. Game-theoretical energy management for energy internet with big data-based renewable power forecasting[J]. IEEE Access, 2017, 5: 5731-5746.
- [8] LI Zhi, YE Lin, ZHAO Yongning, et al. Short-term wind power prediction based on extreme learning machine with error correction[J]. Protection and Control of Modern Power Systems, 2016, 1(1): 9-16. DOI: 10.1186/s41601-016-0016-y.
- [9] 杨茂, 黄宾阳, 江博, 等. 基于卡尔曼滤波和支持向量机的风电功率实时预测研究[J]. 东北电力大学学报, 2017, 37(2): 45-51.
YANG Mao, HUANG Binyang, JIANG Bo, et al. Real-time prediction for wind power based on Kalman filter and support vector machines[J]. Journal of Northeast Electric Power University, 2017, 37(2): 45-51.
- [10] 杨杰, 霍志红, 何永生, 等. 基于小波与最小资源分配网络的超短期风电功率预测研究[J]. 电力系统保护与控制, 2018, 46(9): 55-61.
YANG Jie, HUO Zhihong, HE Yongsheng, et al. Ultra-short-term wind power prediction based on wavelet and minimum resource allocation network[J]. Power System Protection and Control, 2018, 46(9): 55-61.
- [11] 杨正瓴, 刘阳, 张泽, 等. 采用最近历史观测值和 PLSR 进行空间相关性超短期风速预测[J]. 电网技术, 2017, 41(6): 1815-1822.
YANG Zhengling, LIU Yang, ZHANG Ze, et al. Ultra-short-term wind speed prediction with spatial correlation

- using recent historical observations and PLSR[J]. Power System Technology, 2017, 41(6): 1815-1822.
- [12] WANG L, TAN A C, CHOLETTE M, et al. Comparison of the effectiveness of analytical wake models for wind farm with constant and variable hub heights[J]. Energy Conversion & Management, 2016, 124: 189-202.
- [13] WANG H Z, LI G Q, WANG G B, et al. Deep learning based ensemble approach for probabilistic wind power forecasting[J]. Applied Energy, 2017, 188: 56-70.
- [14] RENANI E T, ELIAS M F M, RAHIM N A. Using data-driven approach for wind power prediction: a comparative study[J]. Energy Conversion & Management, 2016, 118: 193-203.
- [15] 朱乔木, 党杰, 陈金富, 等. 基于深度置信网络的电力系统暂态稳定评估方法[J]. 中国电机工程学报, 2018, 38(3): 735-743.
- ZHU Qiaomu, DANG Jie, CHEN Jinfu, et al. A method for power system transient stability assessment based on deep belief networks[J]. Proceedings of the CSEE, 2018, 38(3): 735-743.
- [16] 杨帆, 王干军, 彭小圣, 等. 基于卷积神经网络的高压电缆局部放电模式识别[J]. 电力自动化设备, 2018, 38(5): 123-128.
- YANG Fan, WANG Ganjun, PENG Xiaosheng, et al. Partial discharge pattern recognition of high-voltage cables based on convolutional neural network[J]. Electric Power Automation Equipment, 2018, 38(5): 123-128.
- [17] 朱乔木, 李弘毅, 王子琪, 等. 基于长短期记忆网络的风电场发电功率超短期预测[J]. 电网技术, 2017, 41(12): 3797-3802.
- ZHU Qiaomu, LI Hongyi, WANG Ziqi, et al. Short-term wind power forecasting based on LSTM[J]. Power System Technology, 2017, 41(12): 3797-3802.
- [18] 李莉, 刘永前, 杨勇平, 等. 基于 CFD 流场预计算的短期风速预测方法[J]. 中国电机工程学报, 2013, 33(7): 27-32, 22.
- LI Li, LIU Yongqian, YANG Yongping, et al. Short-term wind speed forecasting based on CFD pre-calculated flow fields[J]. Proceedings of the CSEE, 2013, 33(7): 27-32, 22.
- [19] 姜文玲, 王勃, 汪宁渤, 等. 基于风速衰减因子法的大型风电场尾流效应模拟方法研究[J]. 电网技术, 2017, 41(11): 3499-3505.
- JIANG Wenling, WANG Bo, WANG Ningbo, et al. A study of the wake simulation method of the large-scale wind farms based on wind attenuation factor[J]. Power System Technology, 2017, 41(11): 3499-3505.
- [20] FENG C, CUI M, HODGE B M, et al. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting[J]. Applied Energy, 2017, 190: 1245-1257.
- [21] 叶林, 朱倩雯, 赵永宁. 超短期风电功率预测的自适应指数动态优选组合模型[J]. 电力系统自动化, 2015, 39(20): 12-18.
- YE Lin, ZHU Qianwen, ZHAO Yongning, et al. Dynamic optimal combination model considering adaptive exponential for ultra-short term wind power prediction[J]. Automation of Electric Power Systems, 2015, 39(20): 12-18.
- [22] 薛禹胜, 陈宁, 王树民, 等. 关于利用空间相关性预测风速的评述[J]. 电力系统自动化, 2017, 41(10): 161-169.
- XUE Yusheng, CHEN Ning, WANG Shumin, et al. Review on wind speed prediction based on spatial correlation[J]. Automation of Electric Power Systems, 2017, 41(10): 161-169.
- [23] HU Q, ZHANG R, ZHOU Y. Transfer learning for short-term wind speed prediction with deep neural networks[J]. Renewable Energy, 2016, 85: 83-95.
- [24] 徐龙博, 王伟, 张滔, 等. 基于神经网络平均影响值的超短期风电功率预测[J]. 电力系统自动化, 2017, 41(21): 40-45.
- XU Longbo, WANG Wei, ZHANG Tao, et al. Ultra-short-term wind power prediction based on neural network and mean impact value[J]. Automation of Electric Power Systems, 2017, 41(21): 40-45.
- [25] FILIK T. Improved spatio-temporal linear models for very short-term wind speed forecasting[J]. Energies, 2016, 9: 168.
- [26] 鲁宗相, 徐曼, 乔颖, 等. 风电功率预测的新型互联网运营模式设计[J]. 电网技术, 2016, 40(1): 125-131.
- LU Zongxiang, XU Man, QIAO Ying, et al. New internet based operation pattern design of wind power forecasting system[J]. Power System Technology, 2016, 40(1): 125-131.
- [27] KUSIAK A, ZHANG Z. Short-horizon prediction of wind power: a data-driven approach[J]. IEEE Transactions on Energy Conversion, 2010, 25(4): 1112-1122.
- [28] 彭小圣, 熊磊, 文劲宇, 等. 风电集群短期及超短期功率预测精度改进方法综述[J]. 中国电机工程学报, 2016, 36(23): 6315-6326, 6596.
- PENG Xiaosheng, XIONG Lei, WEN Jinyu, et al. A summary of the state of the art for short-term and ultra-short-term wind power prediction of regions[J]. Proceedings of the CSEE, 2016, 36(23): 6315-6326, 6596.
- [29] WAN Jianzhou, ZHAN Fanyong, LIU Feng, et al. Hybrid forecasting model based data mining and genetic

- algorithm-adaptive particle swarm optimization: a case study of wind speed time series[J]. *IET Renewable Power Generation*, 2016, 10(3): 287-298.
- [30] KITAJIMA T, YASUNO T. Output prediction of wind power generation system using complex-valued neural network[C] // *SICE Annual Conference 2010, Proceedings of IEEE*, 2010: 3610-3613.
- [31] 薛禹胜, 郁琛, 赵俊华, 等. 关于短期及超短期风电功率预测的评述[J]. *电力系统自动化*, 2015, 39(6): 141-151.
XUE Yusheng, YU Chen, ZHAO Junhua, et al. A review on short term and ultra-short-term wind power prediction[J]. *Automation of Electric Power Systems*, 2015, 39(6): 141-151.
- [32] WANG C, ZHANG H, FAN W, et al. A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction[J]. *Energy*, 2017, 138: 977-990.
- [33] WANG Z, WANG W, LIU C, et al. Probabilistic forecast for multiple wind farms based on regular vine copulas[J]. *IEEE Transactions on Power Systems*, 2017, 33(1): 578-589.
- [34] 吴晓刚, 孙荣富, 乔颖, 等. 基于风电场功率特性的风电预测误差分布估计[J]. *电网技术*, 2017, 41(6): 1801-1807.
WU Xiaogang, SUN Rongfu, QIAO Ying, et al. Estimation of error distribution for wind power prediction based on power curves of wind farms[J]. *Power System Technology*, 2017, 41(6): 1801-1807.
- [35] DOWELL J, PINSON P. Very-short-term probabilistic wind power forecasts by sparse vector auto-regression[J]. *IEEE Transactions on Smart Grid*, 2016, 7(2): 763-770.
- [36] 林优, 杨明, 韩学山, 等. 基于条件分类与证据理论的短期风电功率非参数概率预测方法[J]. *电网技术*, 2016, 40(4): 1113-1119.
LIN You, YANG Ming, HAN Xueshan, et al. Nonparametric approach for short-term probabilistic wind generation forecast based on conditional classification and evidence theory[J]. *Power System Technology*, 2016, 40(4): 1113-1119.
- [37] 兰飞, 桑川川, 梁浚杰, 等. 基于条件Copula函数的风电功率区间预测[J]. *中国电机工程学报*, 2016, 36(增刊1): 79-86.
LAN Fei, SANG Chuanchuan, LIANG Junjie, et al. Interval prediction for wind power based on conditional Copula function[J]. *Proceedings of the CSEE*, 2016, 36(S1): 79-86.
- [38] BLONBOU R. Very short-term wind power forecasting with neural networks and adaptive Bayesian learning[J]. *Renewable Energy*, 2011, 36(3): 1118-1124.
- [39] PARK J, LAW K H. Bayesian ascent: a data-driven optimization scheme for real-time control with application to wind farm power maximization[J]. *IEEE Transactions on Control Systems Technology*, 2016, 24(5): 1655-1668.
- [40] YAN J, LI K, BAI E W, et al. Hybrid probabilistic wind power forecasting using temporally local Gaussian process[J]. *IEEE Transactions on Sustainable Energy*, 2017, 7(1): 87-95.
- [41] 姜文玲, 王勃, 汪宁渤, 等. 多时空尺度下大型风电基地出力特性研究[J]. *电网技术*, 2017, 41(2): 493-499.
JIANG Wenling, WANG Bo, WANG Ningbo, et al. Research on power output characteristics of large-scale wind power base in multiple temporal and spatial scales[J]. *Power System Technology*, 2017, 41(2): 493-499.
- [42] 吕颖, 鲁广明, 谢昶, 等. 考虑大规模集中接入风电功率波动相关性的在线概率安全评估[J]. *电网技术*, 2018, 42(4): 1-8.
LÜ Ying, LU Guangming, XIE Chang, et al. Online probabilistic security assessment considering centralized integration of large scale wind power[J]. *Power System Technology*, 2018, 42(4): 1-8.
- [43] 刘晓楠, 周介圭, 贾宏杰, 等. 基于非参数核密度估计与数值天气预报的风速预测修正方法[J]. *电力自动化设备*, 2017, 37(10): 15-20.
LIU Xiaonan, ZHOU Jiegui, JIA Hongjie, et al. Correction method of wind speed prediction based on non-parametric kernel density estimation and numerical weather prediction[J]. *Electric Power Automation Equipment*, 2017, 37(10): 15-20.
- [44] GEBRAAD P M O, TEEUWISSE F W, VAN WINGERDEN J W, et al. Wind plant power optimization through yaw control using a parametric model for wake effects—a CFD simulation study[J]. *Wind Energy*, 2016, 19(1): 95-114.
- [45] 杨茂, 董骏城, 罗芑, 等. 基于近似熵的电力系统负荷预测误差分析[J]. *电力系统保护与控制*, 2016, 44(23): 24-29.
YANG Mao, DONG Juncheng, LUO Yuan, et al. Study of power system load forecasting errors based on approximate entropy[J]. *Power System Protection and Control*, 2016, 44(23): 24-29.
- [46] 叶瑞丽, 郭志忠, 刘瑞叶, 等. 基于小波包分解和改进Elman神经网络的风电场风速和风电功率预测[J]. *电工技术学报*, 2017, 32(21): 103-111.
YE Ruili, GUO Zhizhong, LIU Ruiye, et al. Wind speed

- and wind power forecasting method based on wavelet packet decomposition and improved Elman neural network[J]. Transactions of China Electrotechnical Society, 2017, 32(21): 103-111.
- [47] 李国庆, 张钰, 张明江, 等. 基于 MRMR 的集合经验模态分解和支持向量机的风电功率实时预测[J]. 东北电力大学学报, 2017, 37(2): 39-44.
LI Guoqing, ZHANG Yu, ZHANG Mingjiang, et al. The wind power real-time prediction on the EEMD and SVM of the MRAR[J]. Journal of Northeast Electric Power University, 2017, 37(2): 39-44.
- [48] SAFARI N, CHUNG C Y, PRICE G C D. A novel multi-step short-term wind power prediction framework based on chaotic time series analysis and singular spectrum analysis[J]. IEEE Transactions on Power Systems, 2018, 33(1): 590-601.
- [49] 胡阳, 乔依林. 基于置信等效边界模型的风功率数据清洗方法[J]. 电力系统自动化, 2018, 42(15): 18-23.
HU Yang, QIAO Yilin. Wind power data cleaning method based on confidence equivalent boundary model[J]. Automation of Electric Power Systems, 2018, 42(15): 18-23.
- [50] WANG S, ZHANG N, WU L, et al. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method[J]. Renewable Energy, 2016, 94: 629-636.
- [51] WANG J, DU P, NIU T, et al. A novel hybrid system based on a new proposed algorithm—multi-objective whale optimization algorithm for wind speed forecasting[J]. Applied Energy, 2017, 208: 344-360.
- [52] 王丽婕, 张利, 张岩. 基于 CEEMD 和混沌理论的超短期风电功率预测模型[J]. 系统仿真学报, 2018, 30(4): 1560-1565.
WANG Lijie, ZHANG Li, ZHANG Yan. Ultra-short-term wind power forecasting based on CEEMD and theory[J]. Journal of System Simulation, 2018, 30(4): 1560-1565.
- [53] 田波, 朴在林, 郭丹, 等. 基于改进 EEMD-SE-ARMA 的超短期风功率组合预测模型[J]. 电力系统保护与控制, 2017, 45(4): 72-79.
TIAN Bo, PIAO Zailin, GUO Dan, et al. Wind power ultra-short term model based on improved EEMD-SE-ARMA[J]. Power System Protection and Control, 2017, 45(4): 72-79.
- [54] YIN H, DONG Z, CHEN Y, et al. An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization[J]. Energy Conversion & Management, 2017, 150: 108-121.
- [55] 张亚超, 刘开培, 秦亮, 等. 基于聚类经验模态分解-样本熵和优化极限学习机的风电功率多步区间预测[J]. 电网技术, 2016, 40(7): 2045-2051.
ZHANG Yachao, LIU Kaipei, QIN Liang, et al. Wind power multi-step interval prediction based on ensemble empirical mode decomposition-sample entropy and optimized extreme learning machine[J]. Power System Technology, 2016, 40(7): 2045-2051.
- [56] 崔明建, 孙元章, 柯德平. 基于原子稀疏分解和 BP 神经网络的风电功率爬坡事件预测[J]. 电力系统自动化, 2014, 38(12): 6-11, 26.
CUI Mingjian, SUN Yuanzhang, KE Deping, et al. Wind power ramp events forecasting based on atomic sparse decomposition and BP neural networks[J]. Automation of Electric Power Systems, 2014, 38(12): 6-11, 26.
- [57] 杨茂, 张强. 基于最大相关最小冗余相关向量机的风电功率缺失数据补齐研究[J]. 太阳能学报, 2017, 38(4): 938-944.
YANG Mao, ZHANG Qiang. Polishing missing data for wind power based on MRMR of relevance vector machine[J]. Acta Energies Solaris Sinica, 2017, 38(4): 938-944.
- [58] 王尤嘉, 鲁宗相, 乔颖, 等. 基于特征聚类的区域风电短期功率统计升尺度预测[J]. 电网技术, 2017, 41(5): 1383-1389.
WANG Youjia, LU Zongxiang, QIAO Ying, et al. Short-term regional wind power statistical upscaling forecasting based on feature clustering[J]. Power System Technology, 2017, 41(5): 1383-1389.
- [59] KHODAYAR M, KAYNAK O, KHODAYAR M E. Rough deep neural architecture for short-term wind speed forecasting[J]. IEEE Transactions on Industrial Informatics, 2017, 13(6): 2770-2779.
- [60] 刘红柳, 杨茂. 基于小波变换和支持向量机的风电功率爬坡事件识别与预测[J]. 东北电力大学学报, 2016, 36(6): 30-35.
LIU Hongliu, YANG Mao. An approach for wind power climbing event identification and prediction based on wavelet and SVM[J]. Journal of Northeast Electric Power University, 2016, 36(6): 30-35.
- [61] 刘爱国, 薛云涛, 胡江鹭, 等. 基于 GA 优化 SVM 的风电功率的超短期预测[J]. 电力系统保护与控制, 2015, 43(2): 90-95.
LIU Aiguo, XUE Yuntao, HU Jianglu, et al. Ultra-short-term wind power forecasting based on SVM optimized by GA[J]. Power System Protection and Control, 2015, 43(2): 90-95.
- [62] 李燕青, 袁燕舞, 郭通. 基于 AMD-ICSA-SVM 的超短

- 期风电功率组合预测[J]. 电力系统保护与控制, 2017, 45(14): 113-120.
- LI Yanqing, YUAN Yanwu, GUO Tong. Combination ultra-short-term prediction of wind power based on AMD-ICSA-SVM[J]. Power System Protection and Control, 2017, 45(14): 113-120.
- [63] WU W, PENG M. A data mining approach combining k-means clustering with bagging neural network for short-term wind power forecasting[J]. IEEE Internet of Things Journal, 2017, 4(4): 979-986.
- [64] ZHENG D, ESEYE A T, ZHANG J, et al. Short-term wind power forecasting using a double-stage hierarchical ANFIS approach for energy management in microgrids[J]. Protection and Control of Modern Power Systems, 2017, 2(2): 136-145. DOI: 10.1186/s41601-017-0041-5.
- [65] 王钊, 潘险, 陈迎, 等. 基于实测数据的风电场风速-功率模型的研究[J]. 电力系统保护与控制, 2014, 42(2): 23-27.
- WANG Qian, PAN Xianxian, CHEN Ying, et al. Study of wind speed-active power model for wind farm based on measured data[J]. Power System Protection and Control, 2014, 42(2): 23-27.
- [66] 王贺, 胡志坚, 仇梦林. 基于模糊信息粒化和最小二乘支持向量机的风电功率波动范围组合预测模型[J]. 电工技术学报, 2014, 29(12): 218-224.
- WANG He, HU Zhijian, ZHANG Menglin. A combined forecasting model for range of wind power fluctuation based on fuzzy information granulation and least squares support vector machine[J]. Transactions of China Electrotechnical Society, 2014, 29(12): 218-224.
- [67] 肖迁, 李文华, 李志刚, 等. 基于改进的小波-BP 神经网络的风速和风电功率预测[J]. 电力系统保护与控制, 2014, 42(15): 80-86.
- XIAO Qian, LI Wenhua, LI Zhigang, et al. Wind speed and power prediction based on improved wavelet-BP neural network[J]. Power System Protection and Control, 2014, 42(15): 80-86.
- [68] ZHAO J, GUO Y, XIAO X, et al. Multi-step wind speed and power forecasts based on a WRF simulation and an optimized association method[J]. Applied Energy, 2017, 197: 183-202.
- [69] GAO Y, QU C, ZHANG K. A hybrid method based on singular spectrum analysis, firefly algorithm, and BP neural network for short-term wind speed forecasting[J]. Energies, 2016, 9(1-10): 757.
- [70] DEO R C, GHORBANI M A, SAMADIANFARD S, et al. Multi-layer perceptron hybrid model integrated with the firefly optimizer algorithm for wind speed prediction of target site using a limited set of neighboring reference station data[J]. Renewable Energy, 2018, 116: 309-323.
- [71] OUYANG T, KUSIAK A, HE Y. Modeling wind-turbine power curve: a data partitioning and mining approach[J]. Renewable Energy, 2017, 102: 1-8.
- [72] 杨茂, 杨琼琼. 风电机组风速-功率特性曲线建模研究综述[J]. 电力自动化设备, 2018, 38(2): 34-43.
- YANG Mao, YANG Qionqiong. A Review of wind speed-power curve modeling for wind turbines[J]. Electric Power Automation Equipment, 2018, 38(2): 34-43.
- [73] TASLIMI-RENANI E, MODIRI-DELSHAD M, ELIAS M F M, et al. Development of an enhanced parametric model for wind turbine power curve[J]. Applied Energy, 2016, 177: 544-552.
- [74] YU M, LIU B, BYON E, et al. Direction-dependent power curve modeling for multiple interacting wind turbines[J]. IEEE Transactions on Power Systems, 2018, 33(2): 1725-1733.
- [75] 高亚静, 刘栋, 程华新, 等. 基于数据驱动的短期风电出力预估 - 校正预测模型[J]. 中国电机工程学报, 2015, 35(11): 2645-2653.
- GAO Yajing, LIU Dong, CHENG Huaxin, et al. Predictor-corrector model of wind power forecast based on data-driven[J]. Proceedings of the CSEE, 2015, 35(11): 2645-2653.
- [76] HU Q, ZHANG S, YU M, et al. Short-term wind speed or power forecasting with heteroscedastic support vector regression[J]. IEEE Transactions on Sustainable Energy, 2017, 7(1): 241-249.
- [77] LIU J, ZIO E. SVM Hyper parameters tuning for recursive multi-step-ahead prediction[J]. Neural Computing & Applications, 2016: 1-15.
- [78] HUANG C M, KUO C J, HUANG Y C. Short-term wind power forecasting and uncertainty analysis using a hybrid intelligent method[J]. IET Renewable Power Generation, 2017, 11(5): 678-687.
- [79] YANG H F, JIANG Z P, LU H Y. A hybrid wind speed forecasting system based on a 'decomposition and ensemble' strategy and fuzzy time series[J]. Energies, 2017, 10: 1422.
- [80] WANG Jianzhou, ZHANG Fanyong, LIU Feng, et al. Hybrid forecasting model-based data mining and genetic algorithm-adaptive particle swarm optimization: a case study of wind speed time series[J]. IET Renewable Power Generation, 2016, 10(3): 287-298.
- [81] 张凯锋, 杨国强, 陈汉一, 等. 基于数据特征提取的风电功率预测误差估计方法[J]. 电力系统自动化, 2014, 38(16): 22-27, 34.

- ZHANG Kaifeng, YANG Guoqiang, CHEN Hanyi, et al. An estimation method for wind power forecast errors based on numerical feature extraction[J]. Automation of Electric Power Systems, 2014, 38(16): 22-27, 34.
- [82] ZHAO J, GUO Y, XIAO X, et al. Multi-step wind speed and power forecasts based on a WRF simulation and an optimized association method[J]. Applied Energy, 2017, 197: 183-202.
- [83] LIU H, MI X W, LI Y F. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network[J]. Energy Conversion & Management, 2018, 156: 498-514.
- [84] LIU J, WANG X, LU Y. A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system[J]. Renewable Energy, 2016, 103: 620-629.
- [85] 牛哲文, 余泽远, 李波, 等. 基于深度门控循环单元神经网络的短期风功率预测模型[J]. 电力自动化设备, 2018, 38(5): 36-42.
- NIU Zhewen, YU Zeyuan, LI Bo, et al. Short-term wind power forecasting model based on deep gated recurrent unit neural network[J]. Electric Power Automation Equipment, 2018, 38(5): 36-42.
- [86] YAN J, ZHANG H, LIU Y, et al. Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping[J]. IEEE Transactions on Power Systems, 2018, 33(3): 3276-3284.
- [87] ZHANG C Y, CHEN C L P, GAN M, et al. Predictive deep Boltzmann machine for multi-period wind speed forecasting[J]. IEEE Transactions on Sustainable Energy, 2017, 6(4): 1416-1425.
- [88] 万杰, 胡清华, 刘金福. 基于深度学习理论的短期风速多步预测方法研究[C] // 中国动力工程学会自控专业委员会, 2013: 6.
- [89] HU Bo, LI Yudun, YANG Hejun, et al. Wind speed model based on kernel density estimation and its application in reliability assessment of generating systems[J]. Journal of Modern Power Systems & Clean Energy, 2017, 5(2): 220-227.
- [90] XU Man, LU Zongxiang, QIAO Ying, et al. Modelling of wind power forecasting errors based on kernel recursive least-squares method[J]. Journal of Modern Power Systems & Clean Energy, 2017, 5(5): 1-11.
- [91] SUN G, JIANG C, CHENG P, et al. Short-term wind power forecasts by a synthetical similar time series data mining method[J]. Renewable Energy, 2017, 115: 575-584.
- [92] SIMÃO H P, POWELL W B, ARCHER C L, et al. The challenge of integrating offshore wind power in the U.S. electric grid, part II: simulation of electricity market operations[J]. Renewable Energy, 2017, 103: 418-431.
- [93] 叶林, 张慈杭, 汤涌, 等. 多时空尺度协调的风电集群有功分层预测控制方法[J]. 中国电机工程学报, 2018, 38(13): 3767-3780.
- YE Lin, ZHANG Cihang, TANG Yong, et al. Active power stratification predictive control approach for wind power cluster with multiple temporal and spatial scale coordination[J]. Proceedings of the CSEE, 2018, 38(13): 3767-3780.
- [94] ZHANG J, FLORITA A, HODGE B M, et al. Ramp forecasting performance from improved short-term wind power forecasting[C] // ASME Design Engineering Technical Conference, 2014.
- [95] SHI J, DING Z, LEE W J, et al. Hybrid forecasting model for very-short term wind power forecasting based on grey relational analysis and wind speed distribution features[J]. IEEE Transactions on Smart Grid, 2014, 5(1): 521-526.
- [96] AZIMI R, GHOFRANI M, GHAYEKHLOO M. A hybrid wind power forecasting model based on data mining and wavelets analysis[J]. Energy Conversion & Management, 2016, 127: 208-225.

收稿日期: 2018-07-17; 修回日期: 2018-08-24

作者简介:

杨茂(1982—), 男, 博士, 教授, 硕士研究生导师, 研究方向为电力系统稳定与控制、风电功率预测; E-mail: yangmao820@163.com

张罗宾(1994—), 男, 硕士研究生, 研究方向为电力系统分析及新能源发电技术。E-mail: 1447303581@qq.com

(编辑 姜新丽)