

基于 PCC-RBF 网络的风电功率短期预测方法*

游坤奇¹, 熊殷¹, 贾永青¹, 赵健², 虢优², 易灵芝²

(1. 湖南电器科学研究院有限公司, 湖南长沙 410009;

2. 湘潭大学自动化与电子信息学院 湖南省多能源协同控制技术工程研究中心, 湖南湘潭 411105)

摘要: 风电功率预测对风电场安全平稳运行、电网调度具有重要意义。针对风电功率短期预测指标选择不合理、预测精确度偏低的问题, 提出一种基于皮尔逊相关系数(PCC)和径向基函数(RBF)神经网络的风电功率短期预测方法。该方法利用 PCC 筛选出与风电功率密切相关的 3 个指标, 即电流、温度、风速, 然后以这 3 个指标作为预测模型的输入对风电功率进行 RBF 样本训练与短期预测。试验结果表明, 所提的预测模型预测误差更小, 预测精度更高, 能够满足风电功率短期预测的要求, 具有广泛的应用前景。

关键词: 风力发电; 功率预测; 皮尔逊相关系数; RBF 神经网络; 特征提取

中图分类号: TM 614 文献标志码: A 文章编号: 1673-6540(2021)01-0041-05

doi: 10.12177/emca.2020.181

Short-Term Wind Power Forecast Method Based on Pearson Correlation Coefficient and RBF Network*

YOU Kunqi¹, XIONG Yin¹, JIA Yongqing¹, ZHAO Jian², GUO You², YI Lingzhi²

(1. Hunan Electric Appliance Science Research Institute Co., Ltd., Changsha 410009, China;

2. Hunan Province Cooperative Innovation Center for Wind Power Equipment and Energy Conversion, School of Automation and Electronic Information, Xiangtan University, Xiangtan 411105, China)

Abstract: Wind power forecast is of great significance for the safe and stable operation of wind farms and power grid dispatching. At present, the selection of wind power short-term forecast indicators is unreasonable and the forecast accuracy is low. Aiming at these problems, a short-term wind power forecast model based on Pearson correlation coefficient (PCC) and radial basis function (RBF) neural network is proposed. Firstly, three indicators closely related to wind power, e. g., current, temperature and wind speed, are selected by PCC. Then, these three indicators are used as the input of the forecast model for RBF samples training and short-term forecast of wind power. The results show that the proposed forecast model has smaller forecast error and higher prediction accuracy. It can meet the requirements of short-term wind power prediction and has a wide application prospect.

Key words: wind power generation; power prediction; Pearson correlation coefficient (PCC); radial basis function (RBF) neural network; feature extraction

0 引言

随着煤炭等传统能源的日益消耗, 寻找一种

新的无污染、可再生能源显得尤其重要。风能来自大自然, 属于可再生无污染的新能源, 若能有效利用, 将大大缓解当前能源用量紧张的局面^[1]。

收稿日期: 2020-09-15; 收到修改稿日期: 2020-11-23

* 基金项目: 国家自然科学基金项目(61572416); 湖南省自科基金株洲联合基金项目(2020JJ6009); 大功率交流传动电力机车系统集成国家重点实验室开放课题项目(2021DGI3007)

作者简介: 游坤奇(1991—), 男, 助理工程师, 研究方向为电器试验设备及试验方法。

熊殷(1981—), 男, 工程师, 研究方向为低压电器测量、控制及保护相关控制器软硬件研发。

贾永青(1986—), 男, 硕士, 工程师, 研究方向为基于单片机和 PLC 的小型控制系统设计开发。

风电功率短期预测是风力发电的重要组成部分,精确的功率预测能够保证风力发电场安全运行,为电网调度提供参考^[2]。因此,研究风电功率短期预测对风电场安全运行具有重大意义。

近年来,国内外学者针对风电功率预测进行了广泛研究^[3-4],其中包括传统的物理方法^[5]、基于统计学的方法^[6]与智能预测方法。(1)物理方法。物理方法需要大量的风电机组、天气、环境等多方面信息,工作量大,预测过程复杂,难以实现风电功率准确而快速的短期预测,一般只适用于长期预测。文献[7]采用高层气象大数据进行风电场中长期功率预测,预测精度较好,但该方法需要求解大量方程组,计算量大,不适合短期预测。(2)基于统计学的方法,包括贝叶斯法^[8]、卡尔曼滤波法^[9]、概率分布法等。文献[10]根据发电调度对系统备用容量安全性和经济性的要求,建立一个能够计算任意概率水平气,且符合 Beta 分布最小概率区间的模型,但该方法无法更加准确地表达出预测误差随预测功率大小变化的关系。文献[11]使用高斯混合分布模型统计风电功率预测误差,能较好地反映出风电功率预测误差的分布与变化情况,但该模型运行时间较长,且拟合精度有待进一步提高。(3)智能预测方法。聚焦于数据挖掘、机器学习等方面,研究以时间序列为基础的预测方法,如自回归移动平均模型^[12](ARMA)、长短期记忆网络^[13](LSTM)等;以输入输出之间的映射关系为基础的预测方法,如聚类算法^[14]、支持向量机^[15]、神经网络等智能算法在功率预测方面的应用。基于 BP 神经网络^[16]的风电功率预测模型,能够有效针对风电机组的输出功率进行预测,但是 BP 神经网络容易陷入局部最优,预测精度仍有不足,预测方法有待提高。文献[17]采用最小二乘支持向量机(LS-SVM)和核密度估计法对风电功率进行预测,将统计学方法与智能算法联系在一起,一定程度上提高了预测的精度,但对数据样本要求较高,且核函数及参数的选取采取的是经验法,预测结果具有一定的随机性。基于极限学习机(ELM)^[18]的预测模型,尽管预测速度较快,但预测精度难以达到要求。

针对以上存在的问题,本文提出一种基于皮尔逊相关系数(PCC)和径向基函数(RBF)网络的风电功率短期预测方法。选取 5 项指标,分别对

其与风电功率的相关性进行分析,筛选出与风电功率相关性大的指标作为模型的输入变量。然后运用 RBF 网络对风电数据进行训练,并完成功率短期预测。试验结果表明该方法预测精度更高,误差更小,且运行时间较快。

1 基本理论

1.1 PCC

在统计学中,相关性分析方法主要是用来计算 2 个变量之间的相关程度,并进一步判断两者之间存在的关系。PCC 由英国学者卡尔·皮尔逊提出^[19],通过 2 个变量之间的协方差和标准差之间的关系进行分析,其计算式为

$$\rho_{ab} = \frac{\text{cov}(a,b)}{\sigma_a\sigma_b} = \frac{\sum_{i=1}^m (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^m (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^m (b_i - \bar{b})^2}} \quad (1)$$

式中: a, b 为 2 个变量; ρ_{ab} 为变量 a 与变量 b 之间的皮尔逊系数,其取值范围为 $[-1, 1]$; $\text{cov}(a, b)$ 为变量 a 与变量 b 之间的协方差; σ_a, σ_b 分别为变量 a 与变量 b 的标准差; m 为样本数量, $i=1, 2, \dots, m$ 。

当 $|\rho_{ab}|=0$ 时,变量 a 与变量 b 不相关;当 $0 < |\rho_{ab}| \leq 0.3$ 时,变量 a 与变量 b 弱相关;当 $0.3 < |\rho_{ab}| \leq 0.6$ 时,变量 a 与变量 b 中等相关;当 $0.6 < |\rho_{ab}| < 1$ 时,变量 a 与变量 b 强相关;当 $|\rho_{ab}|=1$ 时,变量 a 与变量 b 完全相关。

1.2 RBF 神经网络

传统 BP 神经网络进行预测时,容易陷入局部最优且泛化能力较弱。为了解决此问题,本文采用 RBF 神经网络^[20]对风电功率进行预测。RBF 神经网络采用随机逼近的方法,能够获得较好的预测结果,其结构组成为输入层、隐含层和输出层,如图 1 所示。

假设 RBF 神经网络的输入变量为 m 维,即 $\mathbf{x}=[x_1, x_2, \dots, x_m]$,输出变量为 k 维,即 $\mathbf{y}=[y_1, y_2, \dots, y_k]$,则 RBF 网络输出为

$$y_k = \sum_{j=1}^n \omega_{jk} R_j \quad (2)$$

式中: ω_{jk} 为隐含层与输出层之间的权值, $j=1, 2, \dots, n$; R_j 为隐含层第 j 个节点的输出值。

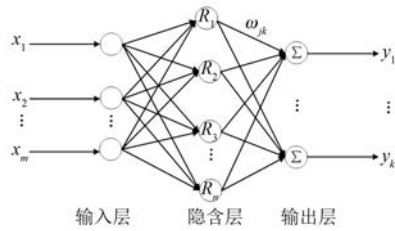


图1 RBF神经网络基本结构图

隐含层是连接输入层与输出层的桥梁,其激活函数为 RBF,本文选择高斯函数作为 RBF:

$$R_j = \exp\left(-\frac{\|x - c_j\|^2}{2\sigma_j^2}\right) \quad (3)$$

式中: c_j 为第 j 个节点的中心; $\|x - c_j\|^2$ 为欧式范数; σ_j 为第 j 个节点的基函数宽度。

1.3 极限学习机

极限学习机^[21] (ELM) 是一种针对单隐层前馈型神经网络求解的模型,与传统基于梯度下降的算法不同,ELM 可以随机初始化输入权重和偏置得到相应的输出权重,典型的极 ELM 由输入层、隐含层和输出层组成。

假设 ELM 的输入变量为 m 维,即 $\mathbf{x} = [x_1, x_2, \dots, x_m]$,输出变量为 k 维,即 $\mathbf{y} = [y_1, y_2, \dots, y_k]$,则 ELM 的输出为

$$y_k = \sum_{j=1}^l \beta_j f(\omega_j x_i + b_j) \quad (4)$$

式中: β_j 为隐含层节点与输出节点的权重值; $f(x)$ 为隐含层的激活函数; ω_j 为连接输入神经元与第 j 个隐含层节点的输入权重; b_j 为第 j 个隐含层节点的偏置; $\omega_j x_i$ 为 ω_j 与 x_i 的内积。

2 风电功率短期预测模型

2.1 特征提取

风力发电具有一定的随机性,其功率受到多种参数的影响。本文从 SCADA 监控系统^[22] 中初步选择风向、风速、温度、电流、空气密度等与风电功率密切相关的指标,并利用 PCC 法分析这 5 个指标与风电功率的相关程度,根据相关性分析结果,进一步选择与功率密切相关的指标作为风电预测模型的输入。

2.2 基于 PCC-RBF 的风电功率短期预测模型

基于 PCC-RBF 的风电功率短期预测模型实现步骤如下。

步骤 1:特征提取。将初始选择的风向、风速、温度、电流、空气密度 5 个指标分别与风电功率进行相关性分析,获取与功率相关的指标,作为预测模型的输入。

步骤 2:模型初始化。建立 RBF 网络结构,设置网络初始参数。

步骤 3:样本输入。将样本数据进行归一化,并分别将训练样本、测试样本输入至模型。

步骤 4:模型训练。采用步骤 2 中的模型结构对训练样本进行训练。

步骤 5:测试预测效果。使用测试样本对步骤 4 中训练完成的模型进行测试,并通过决定系数、多种误差函数等指标对预测效果进行评判。

2.3 评价准则

为了综合确定本文预测模型的优越性,分别选择决定系数(R^2)、平均绝对误差(MAE)、均方根误差(RMSE)、平均绝对百分比误差(MAPE)以及最大误差(ME)等评判指标对风电功率预测效果进行比较。具体计算式分别为

$$R^2 = \frac{\sum_{i=1}^m (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (5)$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (7)$$

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^m \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (8)$$

$$\text{ME} = \max |y_i - \hat{y}_i| \quad (9)$$

式中: y_i 为第 i 个样本数据的真实功率值; \hat{y}_i 为第 i 个样本数据的预测功率值; \bar{y} 为样本数据的功率平均值。

决定系数 R^2 范围为 0~1,且越接近 1,说明预测功率值越接近真实功率值。MAE、RMSE、MAPE、ME 分别为 4 种不同的误差,误差越小,说明预测越准确,预测效果越好。

3 试验验证

为了验证本文提出的 PCC-RBF 风电预测模型的有效性,选择中国某地区风力发电场 2019 年监测数据作为样本来源,每个样本记录时间间隔

30 min。其中,训练样本数量为 1 440 组,测试样本数量为 48 组,即训练一个月的风电功率值来预测短期(1 d)风电功率。

3.1 相关性分析试验

根据 PCC 法,分别计算风电、空气密度、电流、温度、风速 5 个指标与风电功率的相关性,结果如图 2 所示。图 2 为包括功率在内的 6 种指标相关性分析图,图中数值的绝对值越接近 1,两指标之间的相关性越大。由图 2 可知,电流、风速与功率强相关,温度与功率中等相关,风向、空气密度与功率弱相关。因此,本文选择电流、温度、风速作为预测模型的输入变量,功率为输出变量。

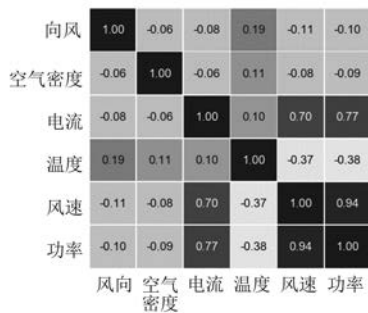


图 2 相关性分析热力图

3.2 预测结果

构建包含输入层、隐含层、输出层的 RBF 神经网络预测模型,输入层设置为 3,隐含层神经元数量为 37,输出层设置为 1,图 3~图 5 分别为 RBF 预测模型、BP 预测模型以及 ELM 预测模型的风电功率短期预测结果图。表 1 为 3 种模型风电功率短期预测效果的量化指标对比。

从图 3~图 5 和表 1 可以看出,RBF 神经网络预测模型的风电功率短期预测效果最好,其决定系数为 0.943 0,大于 BP 与 ELM 的决定指标。同时,RBF 预测的 4 种误差均低于其他 2 种模型的预测误差。综合各指标分析,RBF 预测模型能够

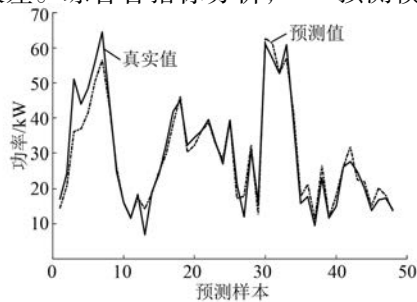


图 3 RBF 风电功率短期预测结果

有效降低预测误差,获得较高的风电功率短期预测结果。

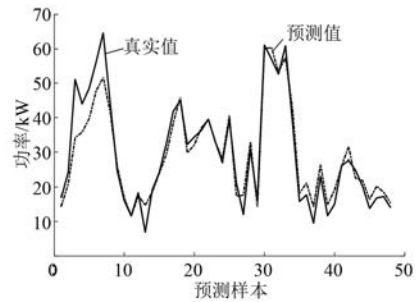


图 4 BP 风电功率短期预测结果

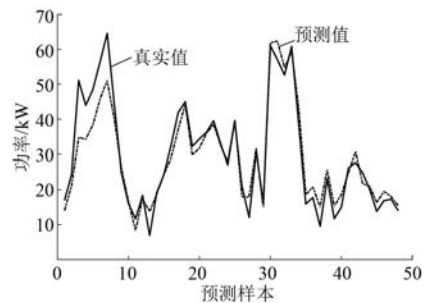


图 5 ELM 风电功率短期预测结果

表 1 3 种模型风电功率短期预测效果对比

指标	RBF	BP	ELM
R^2	0.943 0	0.903 9	0.911 8
MAE	2.705 6	3.430 2	3.399 0
RMSE	3.848 1	5.138 5	4.827 1
MAPE	10.413 5	13.229 5	13.311 8
ME/kW	14.949 9	20.036 5	16.318 8

由于算法具有随机性,为了充分体现本文提出的基于 PCC-RBF 模型的预测有效性,将该模型与 BP 网络、ELM 模型分别运行 10 次,并分别从 R^2 、MAE、RMSE、MAPE 及 ME 5 个方面分析其预测性能,结果分别如图 6~图 10 所示。

图 6~图 10 分别为 R^2 、MAE、RMSE、MAPE 及 ME 的箱型图。从图 6~图 10 的箱型图可以直观看出,RBF 网络多次运行的稳定性更高;RBF 网络的预测决定系数值(R^2)的 10 次运行平均值、最大值与最小值均比 BP 网络、ELM 模型的高;同时,RBF 网络的 MAE 值、RMSE 值、MAPE 值以及 ME 值的均值均比另外 2 种算法低,说明本文提出的 RBF 预测模型预测误差更小,预测性

能良好,能够准确对风电功率进行短期预测。

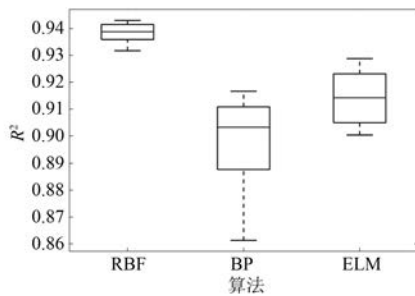


图 6 3种预测算法 R^2 值比较图

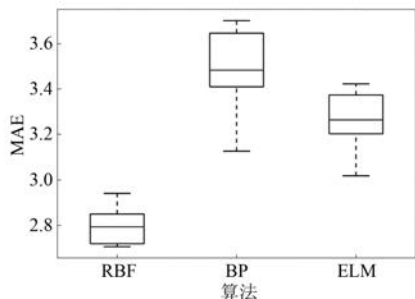


图 7 3种预测算法 MAE 值比较图

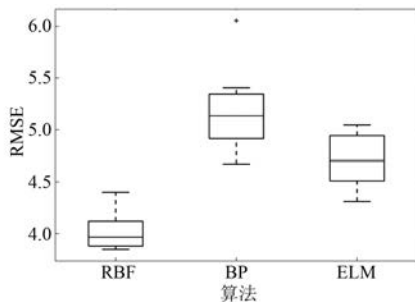


图 8 三种预测算法 RMSE 值比较图

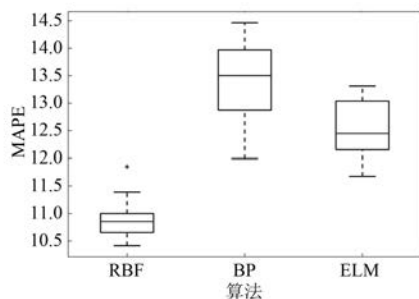


图 9 3种预测算法 MAPE 值比较图

4 结 语

风电功率短期预测对电网调度、风电场安全

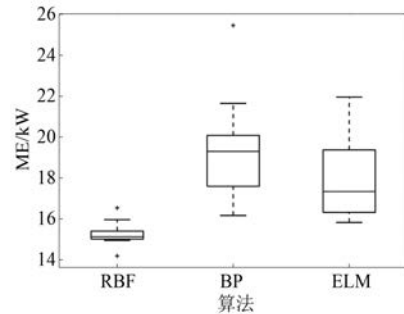


图 10 3种预测算法 ME 值比较图

运行具有重要意义。本文通过 PCC 法选择与风电功率相关程度较高的指标,并运用 RBF 网络模型进行预测,得到以下结论:

(1) 使用 PCC 法对指标进行筛选,能够有效减少风电功率短期预测的计算量;

(2) 与 BP 神经网络、ELM 预测模型相比,本文提出的 RBF 预测模型针对风电功率短期预测误差更小,决定系数更大,预测值与功率实际值更接近。

【参 考 文 献】

- [1] 朱瑛,高云波,臧海祥,等. 风电机组输出功率平滑技术综述[J]. 电力系统自动化,2018,42(18): 182.
- [2] YUAN K, ZHANG K, ZHENG Y, et al. Irregular distribution of wind power prediction[J]. Journal of Modern Power Systems and Clean Energy, 2018(6): 1172.
- [3] 钱政,裴岩,曹利宵,等. 风电功率预测方法综述[J]. 高电压技术,2016,42(4): 1047.
- [4] 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.
- [5] 牛东晓,纪会争. 风电功率物理预测模型引入误差量化分析方法[J]. 电力系统自动化,2020,44(8): 57.
- [6] 杨茂,杜刚,齐玥,等. 基于概率统计的风电功率波动规律性分析[J]. 太阳能学报,2015,36(9): 2231.
- [7] 李飞,纪元. 基于高层气象大数据的风电场中长期风功率预测研究[J]. 电力大数据,2020,23(2): 66.
- [8] 范晓旭,付果,张艳峰,等. 基于 FCM 和贝叶斯神经网络的风电功率短期预测研究[J]. 风能,2020(6): 72.

(下转第 104 页)