

基于支持向量回归的风电场短期功率预测

林涛^{1,2}, 董栅^{1,*}, 秦冬阳¹, 马同宽¹

(1 河北工业大学 控制科学与工程学院, 天津 300130; 2 河北工业大学 计算机科学与软件学院, 天津 300130)

摘要 针对风电场的短期功率预测, 提出了一种考虑风电机组运行条件的用于风电场短期功率预测的新方法。首先, 利用风力发电机的监控和数据采集(SCADA)系统数据计算输出功率和运行条件之间的皮尔逊相关系数, 验证了 SCADA 监测项目对风力发电机输出功率的具有相关性; 其次, 建立支持向量回归(SVR)模型来预测单个风力发电机的风力与气象、运行状态的关系, 发现了考虑运行条件的模型的预测结果优于仅考虑气象信息的模型的预测结果; 最后, 考虑到不同空间位置的风力发电机组对风电场输出功率的贡献不同, 建立了各风力发电机预测功率和风电场预测功率输出之间的回归模型。试验结果表明: 所提出的风场回归模型的预测误差小于风力涡轮机所有预测功率的模型的预测误差, 从而验证了该方法的有效性。

关键词 短期预测; 监控与数据采集系统; 支持向量回归; 风力发电机

中图分类号 TM614 **文献标识码** A **文章编号** 1672-4321(2017)04-0095-05

Short-Term Forecast of Wind Farm Power Based on Support Vector Regression Model

Lin Tao^{1,2}, Dong Shan¹, Qin Dongyang¹, Ma Tongkuan¹

(1 Control Science and Engineering College, Hebei University of Technology, Tianjin 300130, China;

2 Computer Science and Engineering College, Hebei University of Technology, Tianjin 300130, China)

Abstract Aiming at the short-term power forecasting of wind farm, the paper proposed a new method which takes the operating conditions of wind turbines into account. Firstly, according to the analysis of the Pearson correlation coefficient between the output power and operating conditions, the specific relevance between SCADA monitoring project and output wind of wind turbines can be revealed. Then a Support Vector Regression model was built to predict the relationship among the wind power of a single wind turbine, meteorological information and the operation state of the wind turbine. The prediction results of the model which considered the operating conditions are better than those of the model considered only meteorological information. Finally, considering the difference contribution of wind turbines which lies in the different spatial positions, the regression model of each wind power generation and the wind farm's output power were established. The prediction error of the wind field regression model proposed in this paper is less than all the predicted power models of the wind turbine, which verified the validity of the algorithm.

Keywords wind power; short-term prediction; SCADA; support vector regression; wind turbine

风能是目前世界上最有前景的可再生能源, 被广泛用来代替化石燃料^[1]。然而, 由于风能的随机性造成风力发电具有波动性、随机性和抗峰值等问题, 并网发电会对电网的稳定性造成不利影响。为了合理安排发电计划, 提高具有风电场群的大型电力系统的可靠性, 有必要提高风电机组功率预测水平。

近年来, 混合动力系统(尤其是风力发电场和诸如电池、氢存储系统等的其它能量设备的组合)被广泛地用于提高风电场对系统频率和电压的响应性调节^[2]。混合动力系统是通过调节能量存储来平滑输出功率, 这需要及时跟踪风力的变化。因此, 开展风电场的短期功率预测对混合动力系统有着重要的

收稿日期 2017-04-30 * 通讯作者 董栅 研究方向: 计算机网络控制、人工智能 E-mail: isdongshan@foxmail.com

作者简介 林涛(1970-) 男 教授 博士 研究方向: 计算机网络控制理论、网络管理与安全、嵌入式系统及网络控制, E-mail: lintaohebut@163.com

基金项目 河北省科技支撑计划项目(17214304D)

作用^[3].

目前,风电场短期功率预测方法包括间接法和直接法.间接法首先建立一个模型来预测短期风速,通过风速和风功率曲线转换获得风电场的功率曲线^[4].由于风电场包含大量风电机组,总风力功率输出不能直接参考某个单元的功率曲线.因此,功率曲线通常是通过统计方法从风速、风向和输出功率的历史数据得到的二维矩阵,不能满足超短期预测的要求.朱建红等^[5]人通过地理位置、风电机组布局 and 气象特征将风电场分为几个区域,并选择最具代表性的风电机组功率曲线作为给定区域的功率曲线.与风电场的总功率曲线相比,该方法的精度得到提高.然而,由于风电机组的功率曲线通过拟合多组曲线获得,它将增加功率曲线变换中的附加因子的影响.为了提高预测精度,就产生了第二种预测方法,即直接法.直接法将风电场的总输出功率作为研究对象,通过找出隐藏规则或大量累积数据之间的关联,可以预测下一个时间间隔的功率.常用的分析方法包括时间序列分析方法^[6]、人工智能方法^[7].虽然该方法可以避免由功率曲线变换引起的误差,但它忽略了诸如切割机、大修、调节措施以及未达到最大输出的单元等每个风电机组的操作条件的差异.由于以上方法的局限性,提出了第3种方法,即建立单台风电机组的短期功率预测模型,然后将所有风电机组的预测功率作为整个风电场的输出功率^[8].该方法通过对每个风电机组的输出功率建模,可以有效地反映每个风电机组的操作条件的影响.此外,单台风电机组的预测误差小于风电场的总预测误差,提高了风电场输出功率的精度.

研究风电机组的短期功率预测,需要考虑的因素有很多,如风速、风向、空气温度、空气压力等气象信息^[9].风电机组将风能转化为电能的过程可以分为两步:第一步是风能通过风电机组叶片转化为动能,第二步是通过风电机组齿轮箱、发电机等将动能转化为电能.当前研究中考虑的影响因素大多是外部因素,主要集中在第一步,忽略掉第二步风电机组对功率输出的影响,导致风电机组功率预测精度降低.为此,本文提出了一种用于风电场短期风力预测的新型模型,可以考虑到风电机组运行条件对输出功率的影响,提高功率预测精度.

1 皮尔逊相关系数分析

SCADA 系统被用来监控和记录大型风电机组

的运行状态.许多运行参数包括风速、风力转子的转速和发电机功率可以从 SCADA 系统中获得^[10].SCADA 系统中的数据直接反映风力发电机的运行状况,其中包含离散监测项目和连续监测项目.将风能转化为动能的过程中影响风机功率的因素称为气象信息,将动能转化电能过程中影响风机功率因素称为风电机组的运行条件,相关数据可以通过 SCADA 系统得到,如图 1.

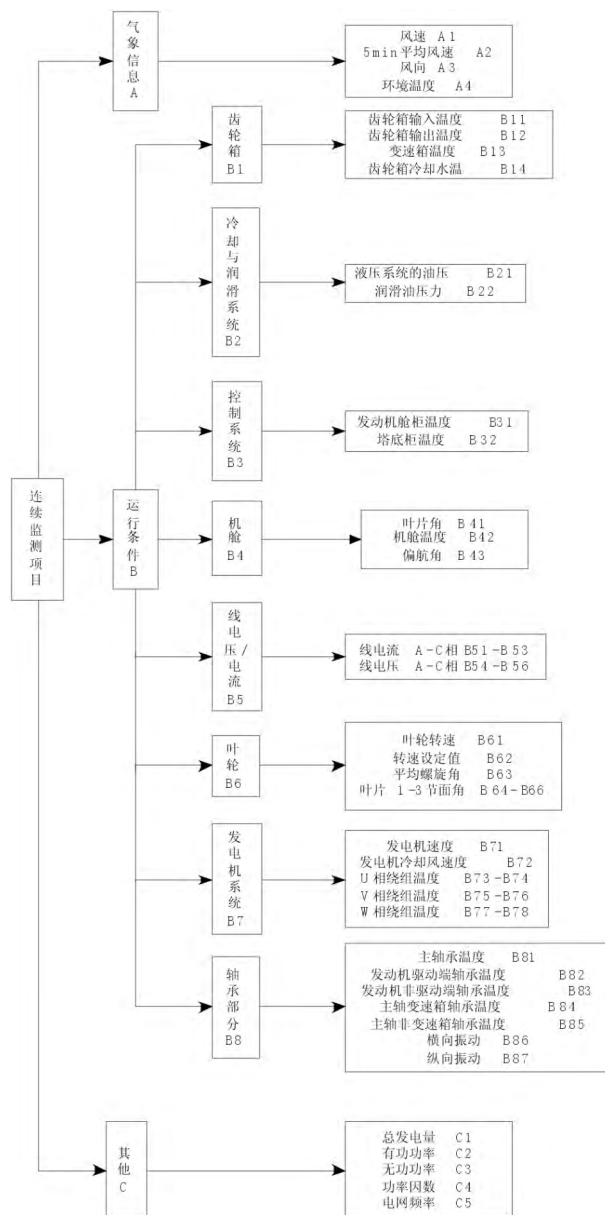


图 1 SCADA 系统的连续监控项目

Fig. 1 The continuous monitoring items of SCADA system

利用从 SCADA 系统中获得相关数据,并引入皮尔逊相关系数来定量描述操作条件、有功功率和气象信息之间的相关性,证明在进行风机功率预测时考虑风机运行状态的必要性.皮尔逊相关系数函数由式

(1) 构成:

$$r_j = \frac{\sum_{j=1}^n (X_{ji} - \bar{X}_j) (Y_i - \bar{Y})}{\sqrt{\sum_{j=1}^n (X_{ji} - \bar{X}_j)^2} \sqrt{\sum_{j=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

式中: r_j 代表皮尔逊相关系数 表示检测项目 j 与有功功率之间的关联 $r_j \in [-1, 1]$; X_{ji} 代表监测项目 j 的第 i 个监测值; Y_i 代表有功功率的第 i 个监测值; \bar{X}_j, \bar{Y} 分别代表监测项目 j 和有功功率的样本平均值.

通过随机选取 8 台风电机组, 利用 SCADA 系统中每天 3 次的历史数据, 计算出每个监测项目和有功功率的皮尔逊相关系数, 得到图 2.

由图 2 可知, 8 台风电机组的相关曲线基本相同, 说明相关曲线在风电场中具有代表性, 可以真实反映 SCADA 系统监控项目与有功功率之间的关系. 通过分析相关曲线, 可知风速 A1 和平均风速在 5min 以内的 A2 与有功功率有较强的相关性, 而风向 A3 和环境温度 A4 与有功功率的相关性较弱; 根据运行状况监测项目, 三相电流与有功功率呈现出强相关性, 网侧三相电压呈现弱关联; 另外有一些监控项目与有功功率有很强的相关性, 如齿轮箱输入/输出轴温度, 叶轮转速, 发电机转速和三相绕组温度. 同时, 叶片桨距角与有功功率之间存在负相关.

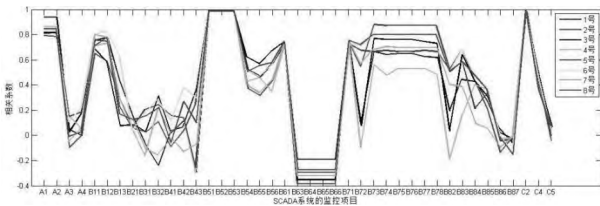


图 2 SCADA 系统的有功功率和连续监控项目的相关性

Fig.2 Correlation of active power and continuous monitoring items of SCADA system

皮尔逊相关系数表示出了 SCADA 系统监测项目与有功功率之间的相关关系, 可以明显地看出, 风电机组的运行状况直接影响着风能转化为电能的效率. 因此, 在预测风电机组功率的时候需要考虑到其运行状态.

2 风电场预测模型

风电机组是一套集合风力发电转换、机械传动、电磁耦合于一体的发电设备, 因此具有明显的非线性和强耦合特性. 为了准确提取 SCADA 系统监测项目与有功功率之间的映射函数, 本文利用支持向量机 (SVM) [11] 方法建立了回归模型.

风电机组中风向标位于轮毂的高度处, 待测量的

风向面对叶轮, 并且风速计测量的风速是风力作用在叶轮上之后的风力. 影响风电机组功率的综合数据由气象数据、叶片桨距角、叶轮转速、叶片角和偏航角等运行条件数据组成, 这意味着不同外部环境下每个风电机组的风力转换效率都是不同的. 因此, 本研究不建立尾流效应、湍流效应、表面形貌的剪切效应和风力发电机安装位置的额外模型.

SVM 是一种基于严格的统计学习理论和结构误差最小化原理开发的机器学习算法, 可以克服小样本、非线性、高维数等实际问题, 并且可以克服神经网络的缺陷, SVM 被广泛用于解决模式识别的问题. 近年来, 由于其优良的性能已经被扩展到回归领域, 命名为 SVR.

SVR 考虑给定的训练数据集 $(x_i, y_i) (i = 1, 2, \dots, m)$, 其中 m 表示训练数据对的总数, x_i 是输入向量, y_i 是对应的输出. 通过采用非线性函数 $g(x)$ 将输入空间 m 映射到特征空间 $l (l > m)$, 将相似数据聚在一起, 使不同数据分隔开, 进行数据的筛选. $f(x) = W^T g(x) + b$ 用来定义特征空间中的超平面, 称为权向量, 用来划分不同数据区域, 其中 b 称为偏差. 采样点和超平面之间的距离就可以表示为 $|y_i - f(x_i)| \leq \varepsilon (\varepsilon \geq 0)$, 即符合这个约束条件的数据属于正常数据, 可以用来计算输出有效功率. 考虑到误差, 使用惩罚系数 $C (C > 0)$ 和非负松弛变量 ξ 和 ξ^* , 通过引入这几个参数降低误差. 利用式 (2) 构建最优超平面的问题可以转换成凸二次优化问题, 即带约束的最小值求解问题:

$$\begin{aligned} \min Q(W, b, \zeta, \xi^*) &= \frac{1}{2} \|W\|^2 + C \sum_{i=1}^m (\zeta_i + \xi_i^*), \\ y_i - f(x_i) &\leq \varepsilon + \xi_i, \\ \text{s. t. } f(x_i) - y_i &\leq \varepsilon + \xi_i^*, \\ \zeta_i \geq 0, \xi_i^* &\geq 0. \end{aligned} \quad (2)$$

为了寻找最优超平面, 引入拉格朗日乘子 α_i 和 α_i^* , 可以将式 (2) 转化为相应的双重问题:

$$\begin{aligned} \max Q(\alpha, \alpha^*) &= - \sum_{i=1}^m (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) g^T(x_i) g(x_j) - \\ &\varepsilon \sum_{i=1}^m (\alpha_i - \alpha_i^*) + \sum_{j=1}^m y_j (\alpha_j - \alpha_j^*), \\ \text{s. t. } \sum_{i,j=1}^m (\alpha_i - \alpha_i^*) &= 0, \rho \leq \alpha_i, \alpha_i^* \leq C. \end{aligned} \quad (3)$$

根据公式 (3), 最佳超平面也称为回归函数可以表示为:

$$f(x) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) g^T(x_i) g(x) + b, \quad (4)$$

由于非线性方程 $g(x)$ 的确定存在难度 $g^T(x_i) g(x)$ 可以由核函数 $k(x_i, x)$ 代替, 满足 Mercer 定理的条件, 避免了计算映射函数 $g(x)$, 然后公式(4) 可以被化为:

$$f(x) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) k(x_i, x) + b, \quad (5)$$

$$y_i - f(x_i) \leq \varepsilon.$$

由式(2) ~ (4) 得到式(5) 构成风电机组功率预测的 ε -SVR 模型, 其中 y_i 表示风电机组的实际功率输出, $f(x_i)$ 表示风电机组的预测功率输出, x_i 表示影响功率因素的输入.

将操作条件和气象条件组成输入向量, 用核函数代替非线性函数, 将非线性问题转换线性问题, 通过在高维线性空间的拟合, 得到预测到的功率, 最后进行降维操作, 得到最终的风电机组输出功率.

以上通过建立 SVR 模型来解决风电机组功率预测问题. 通过推导 SVR 的基本理论, 基于 SVR 的风力预测模型也可以被简化为核函数被用于在时间 t (每个组件或子系统的操作条件风力发电机和气象信息) 通过非线性映射转换为高维空间中的线性问题. 然后, 在高维空间中进行线性拟合, 以实现在时间 $t+1$ 的风力回归预测.

3 实验与分析

3.1 SVR 模型有效性分析

为了验证 SVR 模型的有效性和可行性, 本文选择了河北某风电场的历史数据, 利用平均 5min 的数据作为样本. 随机选取 2d 的数据作为模型的训练样本, 第 3d 的数据作为测试数据.

采用 ε -SVR 模型, 可以直接通过编译的 LIBSVM 工具箱调用和训练, 操作简单. 径向基函数 (RBF) 作为核函数, 网格法用于优化参数惩罚系数 C 和核宽度参数 σ . 最后, 参数由 10 个数据集交叉验证确定.

为了验证运行条件对有功功率的影响, 建立了两个预测模型. 第一个模型的输入是基于气象信息, 第二个模型的输入是基于气象信息和操作条件. 图 3 给出了 1 号风机的实际功率和两个模型的预测结果.

在风速未达到风机的切入风速, 风电机组会切断, 存在有功功率波动段和零输出功率段. 为了定量

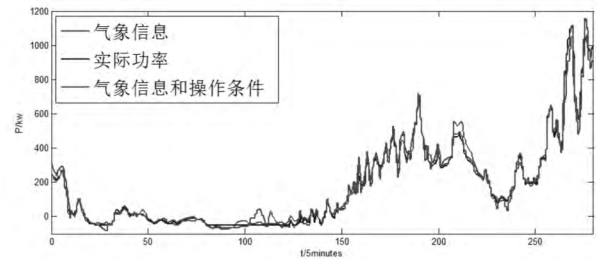


图 3 1 号风机功率预测结果

Fig. 3 Power prediction result of wind turbine No. 1

分析预测结果, 使用归一化绝对平均误差 e_{NMAE} 和归一化均方根误差 e_{NRMSE} 作为误差指数以评估预测精度. 评价指标的定义如式(6) 和(7):

$$e_{NMAE} = \frac{1}{P_{cap}} \frac{1}{N} \sum_{i=1}^N |x'(i) - x(i)|, \quad (6)$$

$$e_{NRMSE} = \frac{1}{P_{cap}} \sqrt{\frac{1}{N} \sum_{i=1}^N [x'(i) - x(i)]^2}, \quad (7)$$

其中 $x'(i)$ 是风力的预测值; $x(i)$ 是风力的真实值; N 是风力发电训练样本的数量; P_{cap} 是风力发电机的额定容量.

表 1 给出了两个模型的预测误差, 第一模型的预测误差大于第二模型的预测误差, 约为第二模型的 3 倍. 因此可以确定运行条件影响有功功率输出, 为了提高预测精度, 模型应考虑风电机组的运行条件.

表 1 两种模型下预测误差

Tab. 1 Prediction error in two models

模型	$e_{NMAE} / \%$	$e_{NRMSE} / \%$
第一模型	1.604	2.314
第二模型	0.570	0.727

将 11 台风电机组功率预测的预测误差列于表 2 中.

表 2 11 台风电机组预测误差

Tab. 2 Prediction error of 11 sets of wind turbines

风电机组 序列号	$e_{NMAE} / \%$		$e_{NRMSE} / \%$	
	气象 信息	气象信息 + 操作条件	气象 信息	气象信息 + 操作条件
风机 1	3.445	3.400	5.675	4.927
风机 2	3.192	2.811	4.71	4.167
风机 3	4.237	4.398	6.523	5.871
风机 4	4.159	3.779	6.636	5.28
风机 5	5.426	4.844	7.828	6.721
风机 6	3.478	3.567	5.333	4.755
风机 7	3.774	4.216	5.594	5.4
风机 8	3.835	4.450	6.284	6.078
风机 9	3.731	4.232	4.941	5.752
风机 10	5.356	3.608	8.588	5.821
风机 11	4.555	3.946	5.999	5.081

表 2 表明, 考虑风机运行条件的模型预测比仅

考虑气象信息的模型误差小,未考虑运行条件风机的平均 e_{NMAE} 为 4.1%, 平均 e_{NRMSE} 为 6.2%; 考虑运行条件风机的平均 e_{NMAE} 为 3.9%, 平均 e_{NRMSE} 为 5.4%, 模型精度更高, 证明了该方法的有效性。

3.2 风电场风力预测模型

本文选择每个单元的预测功率作为 SVR 模型的输入向量, 选择风场的实际输出功率作为训练 SVR 模型的相应输出。由于风电机组数量众多, 每个都考虑并不现实, 因此随机选择风电机组, 构成风电场。它们的 SCADA 数据用于验证情况, 作为对比, 还给出了风电机组的预测值的直接求和的结果。实验表明, 两种算法的预测结果可以表示有功功率的波动, SVR 模型的预测性能优于另一种。

表 3 提供了上述两种模型的 e_{NMAE} 和 e_{NRMSE} 的预测误差。结果表明, 来自 SVR 模型的预测误差小于来自每个风电机组模型的总和的预测误差, 结果符合“大数定律”。因此, 本文提出的 SVR 模型对超短期风力发电有较好的预测精度。

表 3 风电场的功率与预测误差

Tab. 3 Wind power and prediction error

方法	$e_{NMAE}/\%$	$e_{NRMSE}/\%$
SVR	1.794	2.483
风力发电总和	2.664	3.544

4 结论

在本文中, 通过计算 SCADA 监测项目与输出有功功率之间的皮尔逊相关系数, 证明了风力发电机的运行条件可以影响风电转换效率。多个风电机组的相关曲线趋势是一致的, 这表明各种监测项目和有功功率之间的相关系数适合于实际风电场。

本文提出了两种利用 SVR 方法建立发电机组模型。一种是只考虑气象信息, 另一个考虑气象信息和操作条件。当使用相同的 SCADA 数据集训练时, 后者模型的预测误差小于前一模型, 后者模型的误差指数 e_{NMAE} 和 e_{NRMSE} 分别为 3.9% 和 5.4%, 预测精度高于当前预测模型的总体水平, 证明了该方法

的有效性。

当预测每个风电机组的功率时, 建立 SVR 模型使用单个风电机组的预测功率作为输入来预测风力发电场的功率, 这可以获得比利用一个模型预测所有风电机组的功率更高的精度。

参 考 文 献

- [1] 刘波, 贺志佳, 金昊. 风力发电现状与发展趋势[J]. 东北电力大学学报, 2016 (02): 7-13.
- [2] 米增强, 白俊, 刘力卿, 等. 黑启动火电机组后储能型风电场的调控策略研究[J]. 储能科学与技术, 2017(1): 147-153.
- [3] 郭振海. 基于 PLC 的风电混合动力控制系统研究[D]. 阜新: 辽宁工程技术大学, 2011.
- [4] 王松岩, 于继来. 风速与风电功率的联合条件概率预测方法[J]. 中国电机工程学报, 2011(7): 7-15.
- [5] 朱建红, 潘文霞, 马飞. 超短期最优预报下风电场功率预测修正策略[J]. 可再生能源, 2016, 34(3): 401-407.
- [6] 陶佳, 张弘, 朱国荣, 等. 基于优化相空间重构技术的风电场发电功率预测研究[J]. 中国电机工程学报, 2011, 31(28): 9-14.
- [7] 崔嘉, 杨俊友, 杨理践, 等. 基于改进 CFD 与小波混合神经网络组合的风电场功率预测方法[J]. 电网技术, 2017, 41(1): 79-85.
- [8] Han S, Li J, Liu Y. Tabu Search algorithm optimized ANN model for wind power prediction with NWP[J]. Energy Procedia, 2011, 12(39): 733-740.
- [9] Alves M M, Pirmez L, Rossetto S, et al. Damage prediction for wind turbines using wireless sensor and actuator networks[J]. Journal of Network & Computer Applications, 2017, 80: 123-140.
- [10] 刘长良, 吴家佳. 基于核极端学习机的短期风电功率预测研究[J]. 热能动力工程, 2017(1): 95-100.
- [11] 董小念, 文卫蔚. 利用 MapReduce 模型训练支持向量机的人脸识别方法[J]. 中南民族大学学报(自然科学版) 2013, 32(1): 83-86.