

基于动态集成 LSSVR 的超短期风电功率预测*

刘荣胜¹, 彭敏放^{1†}, 张海燕¹, 万勋², 沈美娥³

(1. 湖南大学 电气与信息工程学院, 湖南 长沙 410082; 2. 国网湖南省电力公司电力科学研究院, 湖南 长沙 410007; 3. 北京信息科技大学 计算机学院, 北京 100101)

摘要:针对最小二乘支持向量回归(Least Square Support Vector Regression, LSSVR)建模风电功率时变特性的局限性,提出了一种基于动态集成 LSSVR 的超短期风电功率预测模型.首先利用风电场监测控制与数据采集(Supervisory Control And Data Acquisition, SCADA)与数值天气预报(Numerical Weather Prediction, NWP)系统的历史数据建立离线单体 LSSVR 模型库,然后根据预测时段与训练时段 NWP 序列的相似度从单体 LSSVR 模型库中动态选择候选集成成员,再后综合考虑正确性与多样性确定集成成员.最后由预测时段与训练时段 NWP 序列间的相似度分配集成 LSSVR 成员的权重.通过对湖南省某风电场输出功率进行预测,验证了动态集成 LSSVR 预测模型的有效性,与持续法、自回归求和移动平均法、单体 LSSVR 模型、常权重 LSSVR 组合模型及 BPNN 动态集成模型相比,动态集成 LSSVR 模型具有更高的精度,在天气非平稳变化阶段更加明显.

关键词:超短期风电功率预测;最小二乘支持向量回归;动态集成;动态时间弯曲距离;数值天气预报

中图分类号:TM614

文献标志码:A

Ultra-short-term Wind Power Prediction Based on Dynamical Ensemble Least Square Support Vector Regression

LIU Rongsheng¹, PENG Minfang^{1†}, ZHANG Haiyan¹, WAN Xun², SHEN Meie³

(1. College of Electrical and Information Engineering, Hunan University, Changsha 410082, China;
2. State Grid Hunan Electric Power Company Electric Power Research Institute, Changsha 410007, China;
3. Computer School Beijing Information Science & Technology University, Beijing 100101, China)

Abstract: For the limitation of least square support vector regression (LSSVR) in modeling the time varying feature of wind power, an ultra-short-term wind power prediction (USTWPP) model based on dynamical ensemble LSSVR was proposed. Firstly, the off-line LSSVR model library was created by making use of the historical data which were obtained from Numerical Weather Prediction (NWP) and supervisory control and data acquisition (SCADA) system of wind farm. Then, the candidate members of ensemble LSSVR were selected from off-line LSSVR model library dynamically according to the similarity between

* 收稿日期:2016-08-12

基金项目:国家自然科学基金资助项目(61472128;61173108), National Natural Science Foundation of China(61472128;61173108)

作者简介:刘荣胜(1981-),男,湖南娄底人,湖南大学博士研究生

† 通讯联系人, E-mail:921173232@qq.com

the NWP of forecasting period and the NWP of training period. The ensemble members were decided by considering the accuracy and diversity. Finally, the weights of ensemble LSSVR members were assigned according to the similarity between the NWP of training and NWP of prediction period. The validity of the dynamical ensemble LSSVR based predictor was verified by predicting the wind power of a wind farm in Hunan Province. Compared with persistence method (PM), auto regressive integrated moving average (AGIMA), LSSVR, constant weight ensemble LSSVR, and ensemble artificial neural networks (ANN), the dynamical ensemble LSSVR is more accurate, especially when the weather changes severely.

Key words: ultra-short-term wind power prediction; least square support vector regression; dynamical Ensemble; dynamical time warp; numerical weather prediction

风力发电是目前技术最成熟、最具商业开发前景的可再生能源发电技术之一^[1],受风向、风速、气压、温度、海拔、地形地貌及风机特性等因素影响,其输出功率具有随机性和间歇性^[1-2].风电大规模并入电力系统后,其随机性和间歇性会给电力系统的调度及稳定运行带来新的挑战^[3-4].及时、准确的风电功率预测是大力发展风力发电的关键技术之一.国家能源局颁布的《风电场功率预测预报管理暂行办法》要求风电场必须及时上报短期与超短期风电功率预测(Ultra-short-term Wind Power Prediction, USTWPP)报告^[5].

根据预测时效不同,风电功率预测可分为中长期、短期与超短期预测^[6].USTWPP的预测时效为未来0~4 h,因时效短,精度要求高,对预测方法有更高的要求.风电功率预测的方法包括统计法^[7-8]、物理法^[9]、人工智能法^[10]及组合预测法等.统计法依据风电场输出功率的变化趋势外推下一时刻的风电功率,常用的统计方法有持续法、移动平均法及自回归求和移动平均法等^[11].统计法适合于气象信息有限的USTWPP,其建模简单、计算量少,可根据风电场的特点及位置调整预测模型,准确性较高,应用广泛^[8];但统计法信息源单一,无法按不同的边界条件修正预测模型,健壮性差,在非平稳变化时段预测效果差.物理法基于数值天气预报(Numerical Weather Prediction, NWP)预测未来的气象数据,然后结合风电场周围的地形、地貌等,计算风电机组轮毂高度处的风速与风向,最后参照风机的功率转换曲线预测风电场的输出功率.物理法不需要大量历史数据训练模型,但我国用于风电的商业NWP起步较晚,需提高NWP的更新频率与时间分辨率,才能将物理法应用于USTWPP^[7,10].风电场输出功率受诸多因素影响,机理复杂,难以建立准确的数学模型,为了提高USTWPP预测的准确性,人工神经网络(Artificial Neural

network, ANN)、支持向量机(Support Vector Machine, SVM)及模糊逻辑等人工智能方法相继被用于USTWPP. ANN方法不需按机理建立数学模型,具有较好的非线性逼近能力和泛化能力,但需要大量的历史数据训练模型^[12], ANN的层数及各层的节点数等只能依赖经验确定. SVM方法在处理小样本学习问题时具有良好的性能,但其适应性有待进一步研究^[13],且不能建模时变系统.模糊逻辑可以处理风电功率的不确定性,但隶属函数选择依赖于历史数据与专家经验^[14-15].

单一预测模型有其内在局限性,组合预测方法利用不同方法优势互补,提高预测模型的准确性和健壮性,已成为研究USTWPP的一种趋势.文献[16]提出了一种利用小波变换与SVM相结合的时频综合预测方法,减少了采样误差和噪声对预测结果的影响,但其没有利用NWP信息提高风电功率预测的准确性. ANN与SVM等模型的参数会影响USTWPP的准确性,借助遗传算法或粒子群优化算法可提高参数选取的效率和预测的准确性^[17].基于权重组合的USTWPP方法将不同单体预测方法的预测值加权平均^[18],其利用不同单体模型的优点,克服某一或多个因素对预测误差的集中影响,已成为国内外的研究热点.文献[18]在6种单体风电功率预测方法的基础上采用等权平均组合预测法、均方差倒数法和多层感知器智能算法作为组合方法的融合策略确定各单体模型所占权重.文献[19]通过最大信息熵原则确定每个单体模型的权重建立组合预测模型.文献[20]将风电功率组合预测看作一个信息融合问题,应用交叉熵理论设置各个预测模型的权重,建立风电功率组合预测模型.上述组合模型有效的提高了风电功率预测的准确性,但是没有考虑风电场所处环境的时变特性,上一时刻的最优组合,下一时刻未必最优.根据风电场的时变特性,建立动态组合预测模型,可提高USTWPP预测结果的准确性.本文提出一种基于动态集成最小二乘支

持向量回归(Least Squares Support Vector Regression, LSSVR)的 USTWPP 模型. 首先基于 NWP 与风电场监测控制与数据采集(Supervisory Control And Data Acquisition, SCADA)系统的历史数据离线建立单体 LSSVR 模型库;然后根据预测时段 NWP 特征从单体 LSSVR 库中动态选择候选集成成员;再后综合考虑多样性与正确性确定集成 LSSVR 成员;最后各集成成员的权重由预测时段与训练时段 NWP 的动态弯曲时间(Dynamic Time Warping, DTW)距离给定. 将动态集成 LSSVR 模型应用于湖南某风电场的输出功率预测, 研究及实验证明了基于动态集成 LSSVR 的超短期风电功率预测模型的正确性和有效性, 是计及风电场输出功率时变特性进行 USTWPP 的有益尝试.

1 基本理论

1.1 时间序列相似性

DTW 对同步问题不敏感、计算精度高、可度量长度不同的时间序列间的相似性, 下文基于 DTW 定义时间序列相似度.

定义 1 DTW 距离. 时间序列 $\mathbf{X} = \{x_1, \dots, x_m\}$ 与 $\mathbf{Y} = \{y_1, \dots, y_n\}$ 的 DTW 距离 $D_{\text{dtw}}(\mathbf{X}, \mathbf{Y})$ 递归定义为:

如果 \mathbf{X} 与 \mathbf{Y} 均为空序列 $\{\}$, 则 $D_{\text{dtw}}(\mathbf{X}, \mathbf{Y}) = 0$;

如果 \mathbf{X} 与 \mathbf{Y} 中有且仅有一个空序列 $\{\}$, 则 $D_{\text{dtw}}(\mathbf{X}, \mathbf{Y}) = \infty$;

如果 \mathbf{X} 与 \mathbf{Y} 均不为 $\{\}$, 则 $D_{\text{dtw}}(\mathbf{X}, \mathbf{Y})$ 如式(1):

$$D_{\text{dtw}}(\mathbf{X}, \mathbf{Y}) = d(x_1, y_1) + \min \begin{cases} D_{\text{dtw}}(\text{rest}(\mathbf{X}), \mathbf{Y}) \\ D_{\text{dtw}}(\text{rest}(\mathbf{X}), \text{rest}(\mathbf{Y})) \\ D_{\text{dtw}}(\mathbf{X}, \text{rest}(\mathbf{Y})) \end{cases} \quad (1)$$

其中 $d(x_1, y_1) = \|x_1 - y_1\|_2$; $\text{rest}(\mathbf{X}) = \{x_2, \dots, x_m\}$; $\text{rest}(\mathbf{Y}) = \{y_2, \dots, y_n\}$; m, n 分别为时间序列 \mathbf{X} 与 \mathbf{Y} 的长度. 一般采用动态规划算法计算时间序列间的 DTW 距离, 时间复杂度为 $O(N^2)$ [21].

定义 2 时间序列相似度. 时间序列 \mathbf{X} 与 \mathbf{Y} 的相似度 $S(\mathbf{X}, \mathbf{Y})$ 如下:

$$S(\mathbf{X}, \mathbf{Y}) = \begin{cases} \theta - D_{\text{dwt}}(\mathbf{X}, \mathbf{Y}), & \text{if } \theta > D_{\text{dwt}}(\mathbf{X}, \mathbf{Y}) \\ 0, & \text{else} \end{cases} \quad (2)$$

其中 θ 为给定的相似度阈值, $D_{\text{dwt}}(\mathbf{X}, \mathbf{Y})$ 为时间序列 \mathbf{X} 与 \mathbf{Y} 的动态时间弯曲距离.

1.2 最小二乘支持向量回归

设最小二乘支持向量回归问题的训练集 $S_{\text{tr}} =$

$\{(x_1, y_1), \dots, (x_k, y_k), \dots, (x_K, y_K)\}$, 其中 $\mathbf{x}_k \in R^n$ 为输入量, $y_k \in R$ 为与 x_k 对应的输出, $k = 1, \dots, K$, $K = |S_{\text{tr}}|$ 为训练集的样本数. 对训练集 S_{tr} , LSSVR 优化问题可表示如式(3)~(4) [22]:

$$\min_{\mathbf{w}, b, e} J(\mathbf{w}, b, e) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\gamma}{2} \sum_{k=1}^K e_k^2 \quad (3)$$

$$\text{s. t. } y_k = \mathbf{w}^T \varphi(\mathbf{x}_k) + b + e_k, k = 1, \dots, K \quad (4)$$

其中 \mathbf{w} 为权向量, γ 为正则化参数, e_k 为误差变量, b 为偏置量, $\varphi(\cdot)$ 为从输入空间到高维特征空间的非线性映射.

对应于最优化问题(3)~(4)的 Lagrange 函数如式(5):

$$L(\mathbf{w}, b, e, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\gamma}{2} \sum_{k=1}^K e_k^2 - \sum_{k=1}^K \alpha_k (\mathbf{w}^T \varphi(\mathbf{x}_k) + b + e_k - y_k) \quad (5)$$

其中 α_k 为 Lagrange 乘子, 对应于 $\alpha_k \neq 0$ 的样本点称为支持向量.

式(5)的 KKT 条件如下:

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \sum_{k=1}^K \alpha_k \varphi(\mathbf{x}_k) \quad (6)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{k=1}^K \alpha_k = 0 \quad (7)$$

$$\frac{\partial L}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k, k = 1, \dots, K \quad (8)$$

$$\frac{\partial L}{\partial \alpha_k} = 0 \rightarrow \mathbf{w}^T \varphi(\mathbf{x}_k) + b + e_k - y_k = 0 \quad (9)$$

其中 $k = 1, \dots, K$.

式(6)~(9)转换为式(10)所示的方程组

$$\begin{pmatrix} 0 & \mathbf{I}^T \\ \mathbf{I} & \boldsymbol{\Omega} + \gamma^{-1} \mathbf{I} \end{pmatrix} \begin{pmatrix} b \\ \boldsymbol{\alpha} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{Y} \end{pmatrix} \quad (10)$$

其中 $\mathbf{Y} = (y_1, \dots, y_K)^T$, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_K)^T$, $\mathbf{I} = (1, \dots, 1)^T$, $\boldsymbol{\Omega}_{i,j} = \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j) = H(\mathbf{x}_i, \mathbf{x}_j)$, H 为核函数, 本文取高斯核函数, 其如式(11).

$$H(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right) \quad (11)$$

其中 σ 为自变量 \mathbf{x} 的分布方差.

解方程组(10)求得 $\boldsymbol{\alpha}$ 与 b , 对于新的输入向量 \mathbf{x} , 其对应的输出值 $y(\mathbf{x})$ 如式(12):

$$y(\mathbf{x}) = \sum_{k=1}^l \alpha_k H(\mathbf{x}, \mathbf{x}_k) + b \quad (12)$$

2 基于动态集成 LSSVR 的 USTWPP 建模

计及风电场的时变特性, 提高超短期风电功率预测的准确性, 本文提出一种动态集成 LSSVR 模型, 并将其应用于超短期风电功率预测, 该模型包括

3个步骤:

- 第1步, 离线构建单体 LSSVR 模型库;
- 第2步, 动态选择集成 LSSVR 成员;
- 第3步, 基于 DTW 的权重分配.

2.1 构建单体 LSSVR 模型库

基于动态集成 LSSVR 的 USTWPP 根据风电场的时变特性, 动态重构集成 LSSVR 模型. 单体 LSSVR 是集成 LSSVR 的基础, 为提高集成 LSSVR 动态重构的速度, 利用历史数据构建离线单体 LSSVR 模型库. 单体 LSSVR 模型库构建流程如图 1 所示. 包括数据预处理、子序列划分、NWP 降维、样本集构造及单体 LSSVR 创建等 5 个步骤.

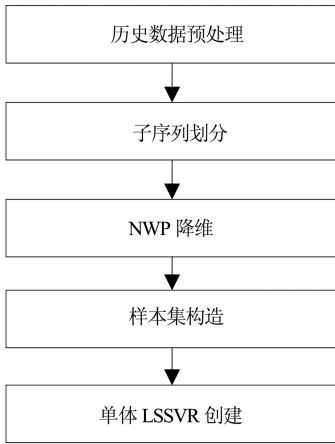


图 1 LSSVR 模型创建流程图

Fig. 1 The flowchart of creating LSSVR

1) 数据预处理: 风电场 SCADA 系统以 15 min 1 次的频率采集风电场测风塔的风速及输出功率. 图 2 为湖南某风电场 SCADA 系统测得的 2013 年 9 月风速与输出功率关系图. 风速低于切入风速时, 输出功率为 0; 风速高于额定风速后, 风电场的输出功率为额定装机容量; 风速在切入风速与额定风速之间时, 功率与风速之间呈 3 次函数关系. 但是受设备故障或计划检修等因素影响, 输出功率在某些时刻远离风速-功率拟合曲线, 该类数据会影响风电场功率预测的准确性, 用其临近 2 个时刻的平均风速与输出功率矫正. 网络和数采系统故障可能导致数据丢失, 如果连续丢失的数据点在 5 个以上, 则将其从历史数据中删除, 否则用均值法对缺失的数据进行补全.

2) 子序列划分: NWP 及 SCADA 系统积累了大量的历史数据, 以全部历史数据训练 LSSVR 模型, 会因为训练样本规模过大, 使得式(10)求解困难, 导致过拟合, 泛化能力差. 因此将历史数据划分成长度为 L 的等长子序列, 用各子序列单独训练模型.

3) NWP 降维: NWP 信息中包含风速的 x 轴分量 u , y 轴分量 v , 气压 p 和温度 t 等, USTWPP 预测时效短, 温度变化较小, 建模时只选取 u, v 与 p . 每个子序列的 NWP 信息可表示成一个如式(13)所示的 $L \times 3$ 矩阵 M , 其中 L 为子序列的长度; 经式(14), 式(15)将矩阵 M 变换成向量 Z .

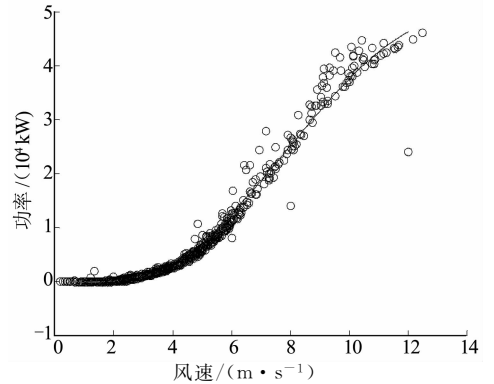


图 2 风电功率与风速关系图

Fig. 2 The relation between wind power and wind speed

$$M = \begin{pmatrix} u_1, v_1, p_1 \\ u_2, v_2, p_2 \\ \dots \\ u_L, v_L, p_L \end{pmatrix} \quad (13)$$

$$C = M^T \times M \quad (14)$$

$$Z = M \times e = (z_1, z_2, \dots, z_L)^T \quad (15)$$

其中 e 为与矩阵 C 最大特征值对应的特征向量.

4) 构建样本集: 单体 LSSVR 模型的样本集表示如下:

$$S_{\text{smp}} = \{(x_i, y_i) \mid i = 1, \dots, N\} \quad (16)$$

其中 $N = |S_{\text{smp}}|$ 为样本集 S_{smp} 的样本数, x_i 为第 i 个样本的输入, y_i 为与 x_i 对应的输出. 创建样本集 S_{smp} 的关键是确定 x_i 与 y_i .

影响风电场输出功率的因素包括地形地貌、风电机组排列、风机型号、海拔、风向、风速、温度、气压及空气密度等, 其中地形地貌、风机排列、风机型号及海拔等为时变因素, 风向、风速、温度、气压及空气密度为时变因素. 将风电场抽象成如图 3 所示的能量转换系统, 时变因素作为系统输入, 在时变因素作用下, 向外输出电能, 输出功率由输入(时变因素)与系统共同(非时变因素)决定. 不同输入作用于系统, 系统上可获得不同的监测信息, 监测信息可以反映系统的功率输出. 而风电场的 NWP 与 SCADA 系统信息就是能量转换系统的监测信息, 其反映了风电场输出功率的变化规律. 计及风速与功率成强相关性, 只将风电场当前的 NWP 与上一时刻的输出功率作为输入, 下一时刻的输出功率作为输出. 样本的输入 $x_i = \langle \text{NWP}_i, P_{i-1} \rangle$, 输出 $y_i = P_i$, 其

中 NWP_i 为经降维处理后预测时刻的 NWP 信息, P_i 为预测时刻的输出功率。

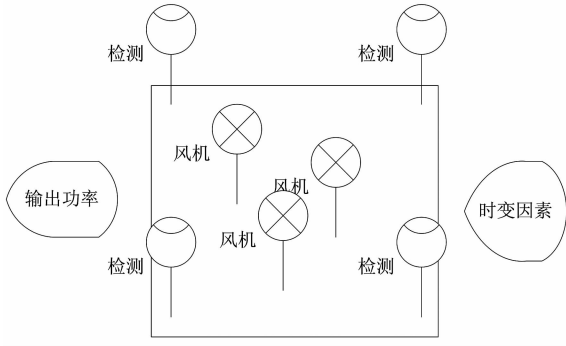


图3 风电场能量转换原理图

Fig. 3 The energy transforming schematic diagrams of wind farm

5) 单体 LSSVR 创建: 用各子序列生成的样本集 S_{smp} 分别训练 LSSVR 模型, 将全体单体 LSSVR 模型及其训练子序列存入数据库建立单体 LSSVR 预测模型库。

2.2 集成成员选择

集成成员的多样性与正确性是构建高质量集成 LSSVR 模型的基础, 为建模风电场的时变特性, 根据预测时段的 NWP 信息, 动态的从单体 LSSVR 模型库中选择预测时段与训练时段 NWP 相似度最高的 Q 个单体 LSSVR 作为集成成员候选集 \mathbf{D} , 集成成员集 \mathbf{D}' 必须满足: $\mathbf{D}' \subseteq \mathbf{D}$, 且 $\psi(\mathbf{D}')$ 有最大值^[23]。

$$\psi(\mathbf{D}') = \lambda \sum_{f_i \in \mathbf{D}'} A(f_i) + (1 - \lambda) \sum_{f_i, f_j \in \mathbf{D}'} \text{kp}(f_i, f_j) \quad (17)$$

其中 $\lambda \in [0, 1]$ 为多样性和正确性的调节因子, 若 $\lambda = 1$, 则只考虑正确性, 若 $\lambda = 0$, 则只考虑多样性。 f_i 与 f_j 均为单体 LSSVR 模型, $A(f_i)$ 为单体 LSSVR 模型 f_i 的预测准确率, 其计算如下式:

$$A(f_m) = 1 - \frac{1}{|\mathbf{S}_{\text{te}}|} \sum_{i=1}^{|\mathbf{S}_{\text{te}}|} \frac{|y_i - f(\mathbf{x}_i)|}{y_i} \quad (18)$$

其中 y_i 为风电功率实测值, $f(\mathbf{x}_i)$ 为与输入 \mathbf{x}_i 对应的预测值, \mathbf{S}_{te} 为测试集。

$\text{kp}(f_i, f_j)$ 为单体 LSSVR 模型 f_i 与 f_j 的差异性度量指标——成对 Kappa (Pairwise-kappa, kp)^[24], 其计算如式(19)。

$$\text{kp}(f_i, f_j) = \frac{2(a_{ij}d_{ij} - b_{ij}c_{ij})}{(a_{ij} + b_{ij})(b_{ij} + d_{ij}) + (a_{ij} + c_{ij})(c_{ij} + d_{ij})} \quad (19)$$

经典 kp 指标用于度量分类器的差异性, 为使其可度量回归模型差异性, 对 f_i 与 f_j 的联合进行了改进, 改进后的联合分布如表 1 所示。

表 1 LSSVR 模型 f_i 与 f_j 的联合分布
Tab. 1 The joint distribution of f_j and f_i

	$\frac{ y - f_i(x) }{y} \leq e$	$\frac{ y - f_i(x) }{y} > e$
$\frac{ y - f_j(x) }{y} \leq e$	a_{ij}	b_{ij}
$\frac{ y - f_j(x) }{y} > e$	c_{ij}	d_{ij}

a_{ij} 为 f_i 与 f_j 的相对误差均不大于允许误差 e 的样本数; b_{ij} 为 f_i 的相对误差大于允许误差 e , 而 f_j 的相对误差不大于允许误差 e 的样本数; c_{ij} 为 f_j 的相对误差大于允许误差 e , 而 f_i 的相对误差不大于允许误差 e 的样本数; d_{ij} 为 f_i 与 f_j 的相对误差均大于允许误差 e 的样本数。

2.3 权重分配

集成员权重分配是将多个单体 LSSVR 组合成高效集成预测模型的关键。风电场具有时变特性, 上一时刻最适合的权重分配, 在下一时刻不一定适应, 为了适应风电场的变化, 根据 NWP 的时变特性动态分配 LSSVR 集成预测模型的权重。

预测时段的 NWP 序列经降维后记为 NWP_0 , 单体 LSSVR _{i} 训练时段的 NWP 序列经降维后记为 NWP_i , 由式(2)计算 NWP_0 与 NWP_i 的相似度 $S(NWP_0, NWP_i)$, 集成预测成员 LSSVR _{i} 的权重如下:

$$w_i = S(NWP_0, NWP_i) / \sum_{r=1}^{|\mathbf{D}'|} S(NWP_0, NWP_r) \quad (20)$$

对于新的输入 \mathbf{x}_i , 其对应的风电功率预测结果为:

$$P_i = \sum_{r=1}^{|\mathbf{D}'|} w_r \times \text{LSSVR}_r(\mathbf{x}_i) \quad (21)$$

3 实验及其结果分析

3.1 实验环境与数据

试验平台采用 Matlab 7. 10. 0(R2010a), 在其上安装 libsvm 软件包用于单体 LSSVR 建模。数据包括两部分: 数值天气预报与风电场历史监测数据。NWP 来自丹麦某气象公司, 时间分辨率为 15 min, 某时刻的历史 NWP 数据为其最后一次更新的值。历史 NWP 信息如表 2 所示, 包含了风速的 x 轴方向分量、 y 轴方向分量及气压。

表 2 某风电场 NWP 历史数据
Tab. 2 The historical NWP data of a wind farm

时间	x 轴方向风速	y 轴方向风速	气压
0:00	7.43	6.45	0.980
0:15	8.60	5.36	0.980
0:30	6.07	9.05	0.986
...

风电场历史监测数据来自于其 SCADA 系统,时间间隔为 15 min. 风电场的装机容量为 48 MW,2013 年 9 月某日经归一化处理后的日功率曲线如图 4 所示. 误差评价指标采用平均百分比误差(MAPE)^[25].

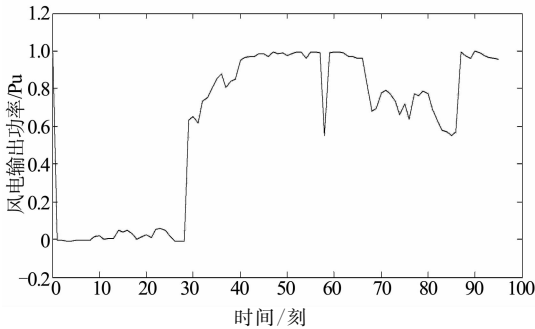


图 4 某风电场日功率曲线

Fig. 4 The daily power curve of a wind farm

3.2 实验结果与分析

以湖南某风电场 2012 年 9 月 1 日至 2013 年 8 月 31 日的 NWP 及其实际输出功率为历史数据训练单体 LSSVR 模型. 共 35 040 组数据,经数据预处理删除连续丢失数据的情形,得 33 574 组有效数据,子序列长度分别为 4 h,8 h,12 h,24 h,3 d 及 7 d 时,训练时段与预测时段 NWP 序列间 DTW 距离与单体 LSSVR 模型预测误差的关系如图 5 所示. 虽然划分子序列的长度不同,但预测误差均随训练时段与预测时段 NWP 序列间 DWT 距离的增大而增大,两者成正相关,且只有 DWT 较小的单体 LSSVR 模型的预测结果能够满足规范的要求. 试验结果说明预测时段与训练时段间 NWP 相似度越高,越有利于提高单体 LSSVR 预测模型的精度.

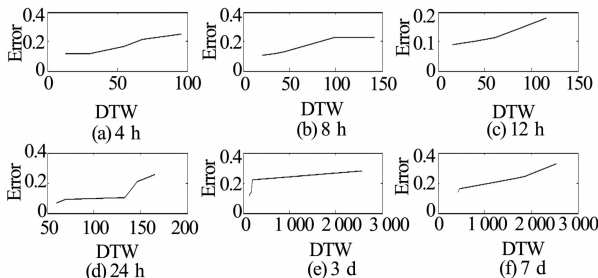


图 5 LSSVR 预测误差与数值天气预报 DTW 距离关系图

Fig. 5 The relation between the DTW of NWP and the error of LSSVR

当子序列时段长度 L 分别为 4 h,8 h,12 h,24 h,3 d 及 7 d 时,与预测时段 NWP 相似度最高的单体 LSSVR 模型预测的结果如图 6 所示,当训练时段与预测时段的 NWP 相似度较高时,各 LSSVR 模型预测值围绕实测值上下波动,LSSVR 模型能够较好地学习预测时段风电输出功率的变化规律,说明了基于 NWP 相似度构建单体 LSSVR 模型的有效性.

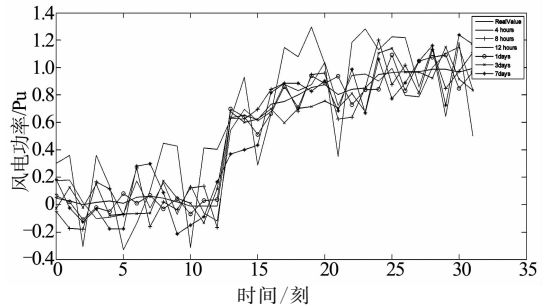


图 6 不同 L 单体 LSSVR 预测试验结果
Fig. 6 The experimental result of single LSSVR with different L

子序列时段长度 L 分别取 4 h,8 h,12 h,24 h,3 d 及 7 d 时,最优单体 LSSVR 模型预测误差的统计结果如表 3 所示.

表 3 不同 L 单体 LSSVR 比较
Tab. 3 The comparison of single LSSVR with different L

L	Min-MAPE/%	Max-MAPE/%	Avg-MAPE/%
4 h	2.33	49.44	28.31
8 h	1.67	25.29	14.68
12 h	0.99	18.27	13.53
24 h	0.68	15.84	12.31
3 d	1.36	17.49	13.08
7 d	2.73	27.03	15.79

表 3 显示: L 过小,单体 LSSVR 模型的训练样本过少,预测精度不高,增加 L ,单体 LSSVR 的预测精度逐步提高,但是当 L 太大时,NWP 的准确率会下降. 当长度 L 为 24 h 时,单体 LSSVR 模型具有最优的预测效果. 构建离线 LSSVR 库时,设定子序列的长度 L 为 24 h.

集成成员多样性与正确性调节因子 λ 是影响集成预测结果的另一个重要因素,当 L 为 24 h 时, λ 与预测误差的统计结果如表 4 所示,统计结果显示 λ 取 0 或 1 时,只顾及了集成成员的正确性或者多样性,预测误差比较大,当其取值为 0.6 时,正确性与多样性取得最好的平衡,预测效果最佳.

取 $\lambda = 0.6, L = 24$ h 时,集成预测试验结果如图 7 所示,动态集成预测模型与单体模型都能有效地预测风电场输出功率. 进一步地对 2013 年 9 月输

出功率预测结果进行统计,统计结果如表 5 所示,表 5 中 Best 表示集成模型中最好的单体 LSSVR 预测结果,Worst 为集成成员中最差的单体 LSSVR 预测结果,Average 为集成成员预测结果的平均误差,DELSSVR 为动态集成 LSSVR 预测结果的误差统计结果.统计结果显示:DELSSVR 预测模型的 MAPE 小于集成成员的平均 MAPE,也小于最好的单体 LSSVR 的 MAPE.验证了基于动态集成 LSSVR 的 USTWPP 模型的有效性,其比最好的单体 LSSVR 模型预测误差更小.

表 4 λ 与动态集成预测误差关系

Tab. 4 The relation between the error of dynamical ensemble model and λ

λ	Min-MAPE/%	Max-MAPE/%	Avg-MAPE/%
0.0	3.33	26.44	12.31
0.1	2.64	22.29	11.68
0.2	1.99	22.27	11.53
0.3	1.68	24.84	10.94
0.4	1.96	23.49	10.17
0.5	3.73	24.07	9.79
0.6	1.10	22.25	9.03
0.7	2.31	27.49	9.76
0.8	3.36	26.71	11.08
0.9	2.92	24.29	12.63
1.0	2.67	23.29	14.65

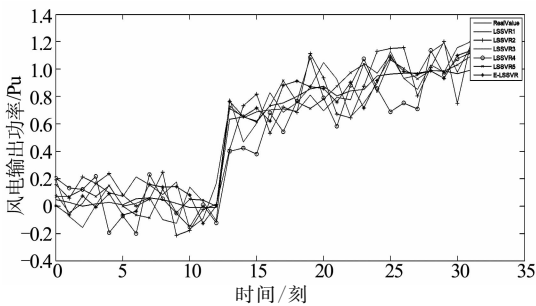


图 7 集成 LSSVR 预测试验结果

Fig. 7 The experimental result of ensemble LSSVR

表 5 集成预测误差分析

Tab. 5 The error analysis of ensemble LSSVR model %

WPP Model	Min-MAPE	Max-MAPE	Avg-MAPE
Best	1.33	21.19	12.31
Worst	0.39	27.70	15.73
Average	0.99	23.27	13.53
DELSSVR	1.12	22.25	9.03

基于 DELSSVR 的 USTWPP 模型与其它常用的 USTWPP 预测模型的对比试验结果如表 6 所示.其中 PM 表示持续法,ARIMA 为自回归移动平

均法,CWLSSVR 为常权重 LSSVR 集成预测模型,DEBPNN 为文献[25]中神经网络动态集成预测模型,DELSSVR 为本文的动态集成 LSSVR 模型. DELSSVR 模型利用了风电场 NWP 序列的时变特性,动态实现集成成员选择与权重分配,与其它方法相比,提高了 USTWPP 的准确性,在非平稳阶段,尤为明显.

表 6 不同预测模型对比试验

Tab. 6 The comparison of different forecasting models

WPP Model	Min-MAPE/%	Max-MAPE/%	Avg-MAPE/%
PM	0.32	33.41	15.36
ARIMA	0.42	28.56	13.65
CWLSSVR	0.98	27.70	10.43
ANN	1.23	32.56	10.57
DEBPNN	2.19	25.27	9.68
DELSSVR	1.93	23.25	9.03

4 结论

计及风电场的时变特征,本文提出了一种基于动态集成 LSSVR 的 USTWPP 模型,并将该模型应用于超短期风电功率预测.研究和实验得出如下结论:

1)风电场输出功率规律受外部天气环境影响,具有时变特性,建立时变模型有利于提高风电功率预测的准确性.

2)由历史数据建立离线单体 LSSVR 模型库的方式,可减少集成 LSSVR 动态重构的计算量,提高构建动态集成 LSSVR 的速度.

3)动态集成 LSSVR 模型基于预测时段 NWP 的时变特征动态选择集成成员与分配权重,其能够更好地建模风电功率的时变特性,提高 USTWPP 的预测精度.

参考文献

[1] 薛禹胜,雷兴,薛峰,等.关于风电不确定性对电力系统影响的评述[J].中国电机工程学报,2014,34(29):5029-5036.
XUE Yusheng, LEI Xing, XUE Feng, *et al.* A review on impacts of wind power uncertainties on power systems [J]. Proceedings of the CSEE, 2014, 34(29): 5029-5036. (In Chinese)

[2] KUSIAK Andrew, ZHANG Zijun. Short-horizon prediction of wind power: A data-driven approach [J]. IEEE Transactions on Energy Conversion, 2010, 25(4):1112-1122.

[3] 王成山,高菲,李鹏,等.低压微网控制策略研究[J].中国电机工程学报,2012,32(25):2-8.
WANG Chengshan, GAO Fei, LI Peng, *et al.* Control strategy research on low voltage microgrid [J]. Proceedings of the CSEE, 2012, 32(25): 2-8. (In Chinese)

[4] 李培强,王继飞,李欣然,等.双馈与直驱风电机组的小干扰稳定性对比分析[J].湖南大学学报:自然科学版,2014,41(1):92-97.
LI Peiqiang, WANG Jifei, LI Xinran, *et al.* Analysis on the impact

- of DFIG and DDSG wind plant on the small signal stability of the power system[J]. Journal of Hunan University: Natural Sciences, 2014, 41(1): 92-97. (In Chinese)
- [5] 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 Electrical Power Systems, 2015, 39(6): 141-151. (In Chinese)
- [6] SHI Jie, DING Zhaohao, LEE Weijen, *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.
- [7] 王贺, 胡志坚, 陈珍, 等. 基于集合经验模态分解和小波神经网络的短期风功率组合预测[J]. 电工技术学报, 2013, 28(9): 137-144.
WANG He, HU Zhijian, CHEN Zhen, *et al.* A hybrid model for wind power forecasting based on ensemble empirical mode decomposition and wavelet neural networks[J]. Transactions of China Electrotechnical Society, 2013, 28(9): 137-144. (In Chinese)
- [8] 周松林, 茆美琴, 苏建徽. 风电功率短期预测及非参数区间估计[J]. 中国电机工程学报, 2011, 31(25): 10-16.
ZHOU Songlin, MAO Meiqin, SU Jianhui. Short-term forecasting of wind power and non-parametric confidence interval estimation [J]. Proceedings of the CSEE, 2011, 31(25): 10-16. (In Chinese)
- [9] NIMA Amjadi, FARSHID Keynia, HAMIDREZA Zareipour. Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization[J]. IEEE Transactions on Sustainable Energy, 2011, 2(3): 265-276.
- [10] CHEN Niya, QIAN Zheng, NABNEY Ian T, *et al.* Wind power forecasts using gaussian processes and numerical weather prediction [J]. IEEE Transactions on Power Systems, 2014, 29(2): 656-665.
- [11] 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.
TAO Jia, ZHANG Hong, ZHU Guorong, *et al.* Wind power prediction based on technology of advanced phase space reconstruction [J]. Proceedings of the CSEE, 2011, 31(28): 9-14. (In Chinese)
- [12] 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.
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. (In Chinese)
- [13] 陈妮亚, 钱政, 孟晓风, 等. 基于空间相关法的风电场风速多步预测模型[J]. 电工技术学报, 2013, 28(5): 15-21.
CHEN Niya, QIAN Zheng, MENG Xiaofeng, *et al.* Multi-step a-head wind speed forecasting model based on spatial correlation and support vector machine[J]. Transactions of China Electrotechnical Society, 2013, 28(5): 15-21. (In Chinese)
- [14] Subhagata Chattopadhyay, Dilip Kumar Pratihari, Sanjib Chandra De Sarkar, *et al.* Fuzzy-logic-based screening and prediction of adult psychoses: A novel approach[J]. IEEE Transactions on System, Man, and Cybernetics—Part A: System and Humans, 2009, 39(2): 381-387.
- [15] 王贺, 胡志坚, 仇梦林, 等. 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, *et al.* 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. (In Chinese)
- [16] 田中大, 李树江, 王艳红, 等. 基于小波变换的风电场短期风速组合预测[J]. 电工技术学报, 2015, 30(9): 112-120.
TIAN Zhongda, LI Shuijiang, WANG Yanhong, *et al.* Short-term wind speed combined prediction for wind farms based on wavelet transform [J]. Transactions of China Electrotechnical Society, 2015, 30(9): 112-120. (In Chinese)
- [17] 王振树, 卞绍润, 刘晓宇, 等. 基于混沌与量子粒子群算法相结合的负荷模型参数辨识研究[J]. 电工技术学报, 2014, 29(12): 211-217.
WANG Zhenshu, BIAN Shaorun, LIU Xiaoyu, *et al.* Research on load model parameter identification based on the CQDPSO algorithm [J]. Transactions of China Electrotechnical Society, 2014, 29(12): 211-217. (In Chinese)
- [18] BOUZOU H, BENOUDJIT N. Multiple architecture system for wind speed prediction [J]. Applied Energy, 2011, 88(7): 2463-2471.
- [19] HAN S, LIU Y, LI J. Wind power combination prediction based on the maximum information entropy principle [C] // World Automation Congress. June 24-28, 2012, Puerto Vallarta, Mexico: 4p.
- [20] 陈宁, 沙倩, 汤奕, 等. 基于交叉熵理论的风电功率组合预测方法[J]. 中国电机工程学报, 2012, 32(4): 29-34.
CHEN Ning, SHA Qian, TANG Yi, *et al.* A combination method for wind power prediction based on cross entropy theory [J]. Proceedings of the CSEE, 2012, 32(4): 29-34. (In Chinese)
- [21] 李海林, 杨丽彬. 基于增量动态时间弯曲的时间序列相似性度量方法[J]. 计算机科学, 2013, 40(4): 227-230.
LI Hailin, YANG Libin. Similarity measure for time series based on incremental dynamic time warping [J]. Computer Science, 2013, 40(4): 227-230. (In Chinese)
- [22] WOJIN Kim, JAEMANN Par, JAEHYUN Yoo, *et al.* Target localization using ensemble support vector regression in wireless sensor networks [J]. IEEE Transactions on Cybernetics, 2013, 43(4): 1189-1198.
- [23] 杨春, 殷绪成, 郝红卫, 等. 基于差异性的分类器集成: 有效性分析及优化集成[J]. 自动化学报, 2013, 40(4): 661-674.
YANG Chun, YIN Xucheng, HAO Hongwei, *et al.* Classifier ensemble with diversity: effectiveness analysis and ensemble optimization [J]. Acta Automatica Sinica, 2013, 40(4): 661-674. (In Chinese)
- [24] 李洪超, 王伟刚, 董雪梅, 等. 基于 M-LS-SVR 的变压器油中溶解气体浓度预测[J]. 电气技术, 2016, 1(2): 76-80.
LI Hongchao, WANG Weigang, DONG Xuemei, *et al.* Concentration prediction of dissolved gases in transformer oil based on M-LS-SVR [J]. Electrical Technology, 2016, 1(2): 76-80. (In Chinese)
- [25] 何东, 刘瑞叶. 基于主成分分析的神经网络动态集成风功率超短期预测[J]. 电力系统保护与控制, 2013, 41(4): 50-54.
HE Dong, LIU Ruiye. Ultra-short-term wind power prediction using ANN ensemble based on the principal components analysis [J]. Power System Protection and Control, 2013, 41(4): 50-54. (In Chinese)