

基于EMD和SVM的风电功率组合预测方法

高准确度的短期风电功率预测对大规模风电的安全并网运行有着重要意义。为了改进功率快速波动时预测系统的准确度，针对风电场输出功率序列随时间、空间呈现非平稳性变化的特征，提出基于经验模态分解（Empirical Mode Decomposition, EMD）和支持向量机（Support Vector Machine, SVM）的组合预测模型。该方法先利用EMD将建模样本中的功率序列按不同波动尺度分解为相对平稳的独立正交分量，以减少不同特征分量的相互影响；然后对分解出来的每个正交分量分别建立预测模型，通过网格寻优法优化SVM参数，提高对不同尺度功率波动的预测准确度；最后采用改进的IOWA加权方式将各分量预测模型的预测值加权得到总的预测功率。短期预测算例结果表明，EMD-SVM模型下采用改进的IOWA加权方式与单一多项式和支

持向量机相比，具有更高的功率预测准确度。

模糊逻辑（Fuzzy Logic）^[5-6]等。然而统计模型对异常样本对（尤其是矛盾样本对）的处理十分困难，处理不当会造成准确度下降。

于是，结合不同模型的优点是提高预测准确度的有效途径^[7]。针对功率序列的非线性和非平稳性的特点，提出一种基于改进的IOWA加权的EMD-SVM短期风电功率预测方法，该方法利用EMD将功率序列分解为相对平稳独立的正交分量；然后对每个分量分别建立相应的预测模型，优化模型参数，提高单项分量预测准确度；最后通过改进的IOWA加权策略对各个分量预测结果组合成最终预测结果，以期取得比单一模型更好的预测结果。

黄 焘¹ 陈建国² 范 强² 雷金勇¹ 杨 苹²/1.南方电网科学研究院有限责任公司 2.贵州电力试验研究院

风电的随机波动特性对所接入的地区电网造成冲击^[1]，准确的风电功率预测能够使调度部门根据风电场的出力调整计划，减少备用容量，降低电网的运行成本^[2]。

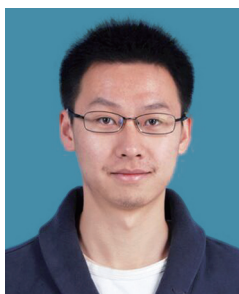
短期功率预测方法分为物理方法及统计方法两类。国外应用物理预测方法的有丹麦Risø实验室^[3]的Landberg等人研发的Prediktor风电功率预测系统。其建模难点在于获得高准确度完善的数值天气预报。而统计方法的预测原理是基于统计学理论，建立起数值天气预报与风机功率或风速输出之间的统计关联。目前常用人工智能等因果关系分析方法建立预测模型^[4]，主要包括人工神经网络（ANN）^[14-17]、支持向量机（SVM）和

基于经验模态分解（EMD）的风电功率样本数据处理

风电场输出功率序列具有随机性、非线性和非平稳的特征，为了改进统计模型对快速波动功率的预测准确度，通过EMD方法将功率序列分解为不同尺度序列，针对每个序列进行建模。

经验模式分解（Empirical Mode Decomposition, EMD）能够对非线性、非平稳的时间序列进行平稳化处理，该方法是一种自适应信号分解方法，分解的过程保留数据本身特性，避免了小波分解需要提前设定基函数与分解尺度缺点^[8]。经验模态分解假设任一个确定的时间序列都可以分解为简单的固有模态函数（IMF），且IMF之间互相正交^[9]。

利用EMD分解功率时间序列 $x(t)$ 的主



黄 焘/工程师

关键词/Keywords

风电功率 ·
支持向量机 ·
参数优化 ·
改进IOWA ·

国家科技支撑计划（2013BAA02B02）。

要思路如下。

1) 提取时间序列信号 $x(t)$ 的所有极大值、极小值, 分别对其做三次样条插值形成上包络线、下包络线。

2) 计算上下包络线均值 $m_1(t)$, 将原始时间序列 $x(t)$ 与 $m_1(t)$ 的差定义为

$$h_1(t) = x(t) - m_1(t) \quad (1)$$

3) 将 $h_1(t)$ 视为新序列信号重复前两个步骤, 经过 k 次之后满足IMF条件时停止。记 $c_1(t) = h_1(t)$ 为第一个IMF分量, 包含原始时间序列中最高频分量。

4) 从原始序列去掉 $c_1(t)$ 的差值为

$$r_1(t) = x(t) - c_1(t) \quad (2)$$

5) 将差值 $r_1(t)$ 作为初始时间序列重复步骤1) ~ 步骤4), 得到 n 个IMF分量以及最终剩余量 $r_n(t)$ 。

EMD过程后, 原始时间序列被分解为 n 个IMF分量以及剩余量 $r_n(t)$ 。

终止条件为 $r_n(t) \leq \delta(t)$, $\delta(t)$ 为限定值。

经过EMD分解将功率时间序列分解为不同尺度数据序列和对应的历史输入数值天气预报, 形成模型样本库, 提高模型对相应尺度分量的预测准确度。

基于改进IOWA加权的SVM预测模型的建立

通过EMD分解建立预测数据样本库后, 需要建立不同尺度数据的预测模型。采用SVM方法进行建模, 并且对各个分量模型中的SVM建模参数进行优化, 得到预测功率的不同尺度分量。

1. SVM预测模型建立及参数优化

支持向量机理论基础为统计学习VC维理论和结构风险最小化原则^[10-12], 具有较强的泛化能力。支持向量回归机用于功率预测原理如下。

假设给定有限样本 $\{(x_1, y_1), (x_2, y_2), \dots, (x_t, y_t)\}$, 其中 $x_i \in R_n$, $y_i \in R$, $i = 1, 2, \dots, t$ 。设对应样本的回归函数为

$$F = \{f | f(x) = w^T x + b, w \in R_n\} \quad (3)$$

问题可等价如下最优问题, 即

$$\min R(w, \xi, \xi^*) = \frac{1}{2}(w^T w) + C \sum_{i=1}^t (\xi_i + \xi_i^*) \quad (4)$$

s. t.

$$\begin{cases} f(x_i) - y_i \leq \xi_i + \varepsilon \\ y_i - f(x_i) \leq \xi_i + \varepsilon \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (5)$$

式中, C 为惩罚因子; ξ_i 、 ξ_i^* 为模型引入的松弛变量; ε 为估计准确度。

模型为非线性时, 需要引入核函数 $K(x_i, x)$ 。采用使用最广泛的径向基核函数, 函数如下, 即

$$K(x_i, x_j) = \exp(-g \|x_i - x_j\|^2) \quad g > 0 \quad (6)$$

针对EMD分解不同尺度的数据序列建模, 需要对不同模型参数进行寻优。

评价模型性能主要用交叉验证 (Cross Validation, CV), 一般采用 k -fold CV, 将模型数据集均分成 k 份, 每个子集分别做一次验证集, $k-1$ 组子集作为训练集, k 个模型最终平均准确度作为此 k -CV的性能指标。

2. IOWA加权

对不同分量建模后, 需要将不同分量加权得到最终预测结果, 先介绍IOWA加权方法。IOWA (诱导有序加权平均算子) 是由Yager提出的一种信息集成方法^[13]。

设有 $(v_1, a_1), (v_2, a_2), \dots, (v_m, a_m)$ m 个二维数组, 令

$$\text{IOWA}[(v_1, a_1), (v_2, a_2), \dots, (v_m, a_m)] = \sum_{i=1}^m w_i a_{v_{\text{index}(i)}} \quad (7)$$

式中, $\text{IOWA}[(v_1, a_1), (v_2, a_2), \dots, (v_m, a_m)]$ 简称为IOWA算子; v_i 为 a_i 的诱导值; $w = (w_1, w_2, \dots, w_m)^T$ 为与IOWA有关的权值向量; $v_{\text{index}(i)}$ 为 $v_1, v_2, \dots, v_m$ 按从大到小顺序排列后第 i 个数的下标。且满足

$$\begin{cases} \sum_{i=1}^m w_i = 1 \\ w_i \geq 0 \quad i = 1, 2, \dots, m \end{cases} \quad (8)$$

设有 m 个分量预测模型对时间 t 的预测值为一 $m \times N$ 矩阵, 即

$$\begin{pmatrix} x_{11} & \cdots & x_{1t} & \cdots & x_{1N} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{it} & \cdots & x_{iN} \\ \vdots & & \vdots & & \vdots \\ x_{m1} & \cdots & x_{mt} & \cdots & x_{mN} \end{pmatrix}$$

该矩阵 $i=1,2,\dots,m$, $t=1,2,\dots,N$ 。 x_{it} 为第 i 个分量预测模型对时间 t 的预测值。令 a_{it} 为第 i 个分量预测模型对时间 t 的预测准确度, 即

$$a_{it} = \begin{cases} 1 - |(x_t - x_{it}) / x_t| & |(x_t - x_{it}) / x_t| < 1 \\ 0 & |(x_t - x_{it}) / x_t| \geq 1 \end{cases} \quad (9)$$

式中, x_t 为时间 t 的实际功率, $a_{it} \in [0,1]$ 。

若将 a_{it} 看作预测值 x_{it} 的诱导值, 根据IOWA的定义

$$\text{IOWA}[(a_{1t}, x_{1t}), (a_{2t}, x_{2t}), \dots, (a_{it}, x_{it})] = \sum_{i=1}^m w_i x_{a_{\text{index}(i)}} \quad (10)$$

令 $e_{a_{\text{index}(i)}} = x_t - x_{a_{\text{index}(i)}}$, 于是所有时间 N 上总的预测误差平方和 F 为

$$F = \sum_{t=1}^N (x_t - \sum_{i=1}^m w_i x_{a_{\text{index}(i)}})^2 \quad (11)$$

则

$$\min F(w_1, w_2, \dots, w_m) = \sum_{i=1}^m \sum_{j=1}^m w_i w_j \left(\sum_{t=1}^N e_{a_{\text{index}(i)}} e_{a_{\text{index}(j)}} \right) \quad (12)$$

记

$$E_{ij} = \sum_{t=1}^N e_{a_{\text{index}(i)}} e_{a_{\text{index}(j)}} \quad i, j = 1, 2, \dots, m \quad (13)$$

E_{ij} 为 $m \times m$ 矩阵, 即阶诱导有序加权平均组合预测误差信息矩阵。

因此, 式(12)可表示为

$$\begin{cases} \min F = w^T E w \\ R^T w = 1 \\ w \geq 0 \\ R = (1, 1, \dots, 1)^T \end{cases} \quad (14)$$

式(14)实际上是一个二次规划问题, 可以通过标准二次型优化技术来求解, 本文直接使用Matlab求解。IOWA加权确认不同尺度的功率分量的加权系数, 所有分量组合加权结果即为预测结果。

3. 改进IOWA加权的EMD-SVM短期风电功率预测

不同分量模型预测值加权叠加时, 传统的方式是为各分量预测模型的预测值设置固定权值系数, 固定的权值系数无法随各分量模型在各目标预测时刻点上的预测准确度做动态调整, 因而加

权性能较差。为此, 提出一种改进的IOWA加权策略以提高风电功率预测准确度。

整个预测的流程具体步骤如下。

1) 首先选择输入参数。

2) 对数值天气预报 (Numerical Weather Prediction, NWP) 数据进行EMD分解, 得到 n 个数据分量。

3) 将这些分量作为SVM输入得到 n 个输出功率。

4) 最后通过IOWA加权得到最终的预测功率值, 并输出。

采用的NWP为WRF格式, 而WRF模式中有第一层、第二层和第三层水平风速、垂直风速、上下风速、气压、位势高度、温度以及比例湿度共21个输入矢量。一般风电机组的高度为80 m, 位于WRF模式中的第一层和第二层之间, 非常接近WRF模式的第二层的高度, 所以选用第二层的参数作为输入参数。

采用EMD分解SVM建模如图1所示。具体步骤如下。

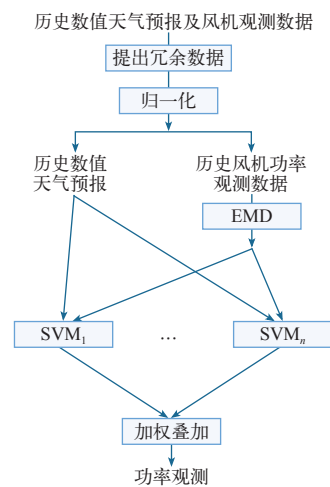


图1 建模流程

1) 对输入历史数据剔除冗余部分, 提高数据的准确性。

2) 将归一化后的历史观测功率数据按EMD分解成 n 个单独的功率时间序列分量并作为SVM分量预测模型的输出, 归一化的历史数值天气预报作为分量预测模型的输入并建模。

3) 通过改进的IOWA加权方式将各分量的预

测值加权叠加得到最终的预测功率。

改进的IOWA加权策略是依目标功率预测时刻 t 逐步将该时刻的分量模型预测值用于改进权值向量 w ，从而实现 w 的更新。该策略如图2所示。

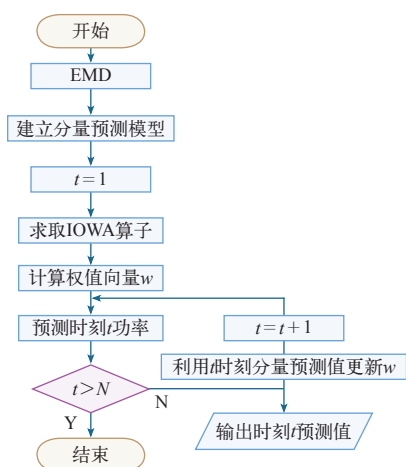


图2 改进IOWA预测流程图

算例及分析

应用算例对上述方法的有效性进行验证，采用建模数据为辽宁某风电场60天的风场SCADA观测数据，以及从气象台获取的相应历史时期的WRF预报数据。该风电场风机额定功率为1.5 MW，额定风速为11.5 m/s，切入风速为3 m/s，叶轮直径为 40.3×2 m，轮毂高度为70 m。用于建模的历史数据中前45天数据用做训练集，共有4 532个时间间隔10 min的数据点；测试集由后15天数据构成，共1 372个时间间隔10 min的数据点。

评价风电功率预测的误差较多采用均方根误差（RMSE），本文采用均方根误差（RMSE）占风机额定功率的百分比，即

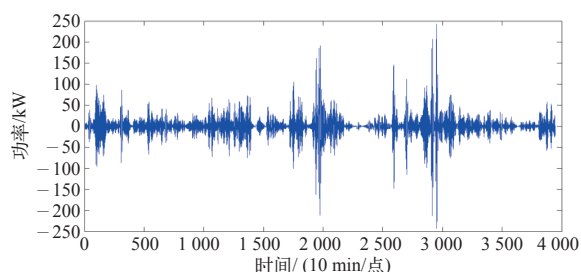
$$\alpha = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - p_i)^2}}{P_{\text{额定}}} \quad (15)$$

式中， p_i 为实际值； x_i 为预测值； $P_{\text{额定}}$ 为风机额定功率。

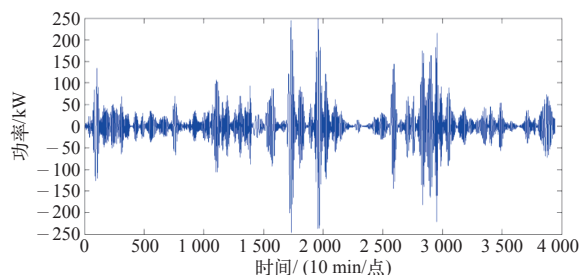
由图3可知，该功率时间序列被自动分解为10个IMF分量及一个剩余分量 C_{11} 。

第一个分量 IMF_1 参数寻优结果如图4所示，其中， C 和 g 的寻优范围均为 $[2-10, 23]$ ，寻优步

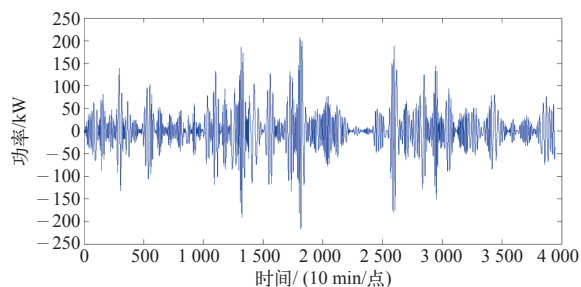
长为0.1。模型的最优参数分别为 $C=0.541$ ， $g=0.00089$ 。



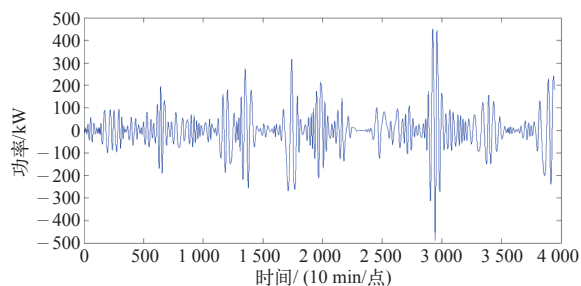
(a) IMF_1



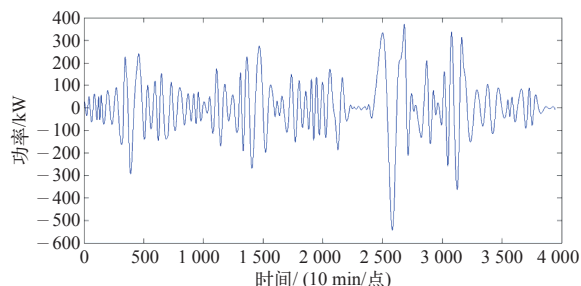
(b) IMF_2



(c) IMF_3



(d) IMF_4



(e) IMF_5

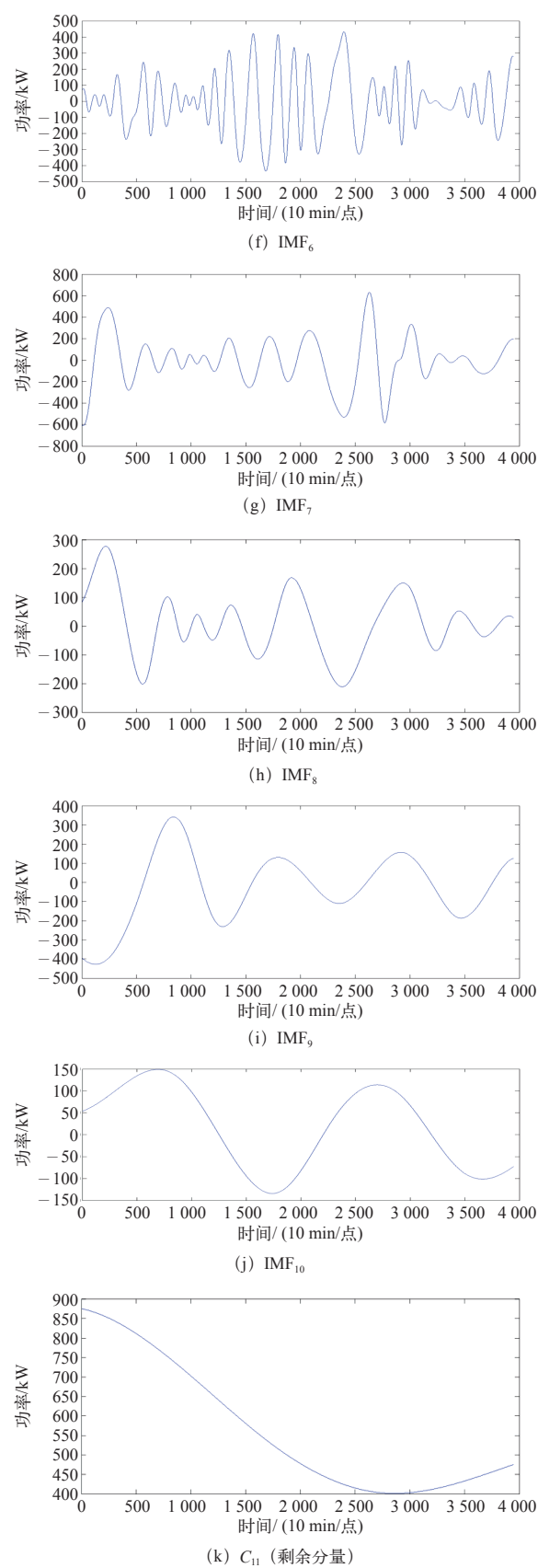


图3 EMD分解的功率分量

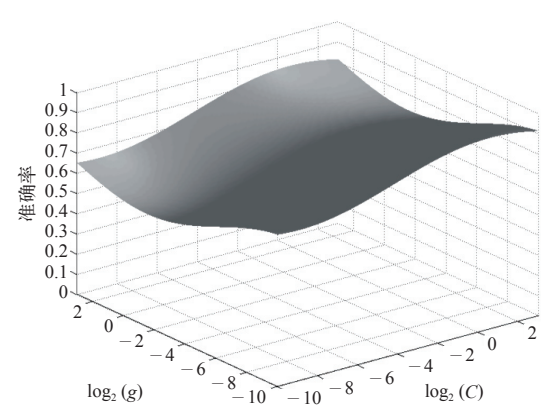


图4 IMF₁的SVM模型参数寻优图

从图3中可以看出， C ， g 在一定的范围内，模型预测准确度都较高。每个分量的模型都需要进行寻优，直接全局寻优计算量过大，为了节约计算时间，可以先扩大步长，找到对应的区间，再在小区间中用小步长寻优。从图3中可以看到局部最大点有两个，所以变步长的寻优有可能会陷入局部极值点，可能是全局次优，但是能够减少计算量，节约计算时间，用较少准确度减少换取计算时间的明显减少。

在建立SVM模型并对建模参数进行寻优的基础上，利用IOWA加权方式在测试集上进行预测，并将其与多项式和支持向量机^[18-19]两种建模方法进行了对比，误差对比如表1所示。

表1 两种算法的预测误差比较

SVM误差/kW	EMD-SVM的IOWA加权RMSE误差/kW
246.5	230.1

IOWA加权预测RMSE误差占风机容量百分比为15.3%，与多项式核支持向量机相比减小16.4 kW的RMSE误差，即减小的RMSE误差占风机容量的1.1%。

进一步利用改进IOWA加权策略进行了功率预测。改进后的权值如表2所示。权值向量 w 中， w_3 、 w_5 两个维度权值为0，除此之外，其他维度权值均大于0。

改进IOWA加权策略RMSE误差占风机容量的14.8%。其与IOWA加权比较如表3所示。

改进IOWA策略下显然具有更高的预测准确度，这验证了改进IOWA加权策略的有效性。

表2 改进IOWA加权策略及标准IOWA加权方式下的权值向量

权值向量		$W_1, W_2, \dots, W_6$				
IOWA	0.041 1	0.005 2	0	0.009 7	0	0.042 6
改进IOWA	0.036 5	0.004 4	0	0.007	0	0.038 6
权值向量		$W_7, W_8, \dots, W_{11}$				
IOWA	0.073 1	0.143 8	0.145 4	0.161 5	0.377 6	
改进IOWA	0.075 7	0.141 3	0.152 5	0.145 6	0.398 6	

表3 改进IOWA加权策略及标准IOWA加权方式误差比较

EMD-SVM的 IOWA加权RMSE误差/kW	EMD-SVM的 改进IOWA加权RMSE误差/kW
230.1	222.4

结束语

通过EMD分解将非平稳的功率序列分解为不同尺度的功率分量，然后针对不同分量分别建立对应的SVM模型进行预测，并对模型参数进行全局寻优，最后通过改进的IOWA加权预测将各个分量的预测结果组合起来，与直接建模预测比较，结果表明该方法能有效地提高预测准确度。主要结论如下。

- 1) 基于EMD-SVM建模的IOWA加权预测与单一模型多项式核支持向量机相比，能够有效提高准确度。
- 2) 对各个分量模型的参数寻优预测准确度好于单一模型默认参数的预测准确度。
- 3) EMD分解能够有效地改善实际功率快速波动时预测失准的问题，提高预测准确度。

参考文献

[1] 张新房, 徐大平, 吕跃刚, 等. 风力发电技术的发展及若干问题 [J]. 现代电力, 2003, 20 (5) : 30-34.

[2] 袁铁江, 晁勤, 李建林, 等. 风电并网技术 [M]. 北京: 机械工业出版社, 2012.

[3] Troen I, Petersen E L. European wind atlas [M]. Roskilde: Risø National Laboratory, 1989.

[4] 黄金花. 基于神经网络的风电场短期风电功率预测研究 [D]. 南京: 东南大学, 2010.

[5] Damousis I G, Dokopoulos P. A fuzzy expert system for the forecasting of wind speed and power generation in wind farms [C]. IEEE Power Industry Computer Applications Conference, 2001.

[6] 张吉礼. 模糊-神经网络控制原理与工程应用 [M]. 哈尔滨: 哈尔滨工业大学出版社, 2004.

[7] Bernhard L, Kurt R, Bernhard E, et al. Wind power prediction in Germany—recent advances and future challenges [C]. European Wind Energy Conference, Athens, 2006.

[8] 徐争光. 经验模式分解的数学理论研究 [D]. 武汉: 华中科技大学, 2009.

[9] Huang N E, Shen Z, Long S R. The empirical mode decomposition and the Hilbert spectrum or nonlinear and non-stationary time series analysis [J]. Proceedings of the Royal Society SocLond, 1998, 454 (1971) : 903-995.

[10] 刘华煜. 基于支持向量机的机器学习研究 [D]. 大庆: 大庆石油学院, 2005.

[11] Vapnik V. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1995.

[12] Cortes C, Vapnik V. Support vector networks [J]. Machine Learning, 1995 (20) : 273-297.

[13] 黎静华, 韦化. 基于诱导有序加权平均算子的最优组合短期负荷预测 [J]. 电网技术, 2011, 35 (10) : 177-181.

[14] 付伟, 等. 遗传神经网络在风电场风速估计中的应用 [J]. 现代科学仪器, 2009 (6) : 48-52.

[15] 韩力群. 人工神经网络教程 [M]. 北京: 北京邮电大学出版社, 2006.

[16] 田景文, 高美娟. 人工神经网络算法研究及应用 [M]. 北京: 北京理工大学出版社, 2006.

[17] 卢斌等. 基于GRNN的网络出口流量预测 [J]. 黑龙江科技信息, 2008 (13) : 63.

[18] 张臻. 风电功率短期预测方法研究 [D]. 广州: 华南理工大学, 2010.

[19] 丁志勇. 风电场短期功率预测方法研究 [D]. 广州: 华南理工大学, 2012.

(收稿日期: 2015-03-10) EA