

基于集合经验模态分解和改进极限学习机的 短期风速组合预测研究

张翌晖¹, 王贺², 胡志坚², 王凯¹, 黄东山¹, 宁文辉¹, 张承学²

(1. 广西电力科学研究院, 广西 南宁 530023; 2. 武汉大学电气工程学院, 湖北 武汉 430072)

摘要: 提出一种基于集合经验模态分解 (Ensemble empirical mode decomposition) 和改进极限学习机 (Improved Extreme Learning Machine, IELM) 的新型短期风速组合预测模型。采用集合经验模态分解将风速序列分解成不同频段的分量, 以降低序列的非平稳性。使用改进极限学习机对各分量分别建模预测, 为避免极限学习机输入维数选取的随意性和分量信息丢失等问题, 先对各分量重构相空间, 最后将各分量预测结果叠加得到最终预测结果。实例研究表明, 所提的组合预测模型具有较高的预测精度。

关键词: 风速; 预测; 改进极限学习机; 集合经验模态分解; 相空间重构

A hybrid short-term wind speed forecasting model based on ensemble empirical mode decomposition and improved extreme learning machine

ZHANG Yi-hui¹, WANG He², HU Zhi-jian², WANG Kai¹, HUANG Dong-shan¹, NING Wen-hui¹, ZHANG Cheng-xue²

(1. Guangxi Electric Power Research Institute, Nanning 530023, China;

2. School of Electrical Engineering, Wuhan University, Wuhan 430072, China)

Abstract: This paper proposes a new short-term combination prediction model of wind speed by means of ensemble empirical mode decomposition (EEMD) and improved extreme learning machine (IELM). Firstly, wind speed series is decomposed into several components with different frequency bands by EEMD to reduce the series non-stationary. Secondly, the phase space of each component is reconstructed in order to solve the randomness and component information lost of input dimensionality selection of extreme learning machine, and then an IELM model of each component is established. Finally, the forecast result of each component is superimposed to get the final result. The simulation result verifies that the hybrid model has higher prediction accuracy of wind speed.

This work is supported by Specialized Research Fund for Doctoral Program of Higher Education (No. 20110141110032) and the Fundamental Research Funds for the Central University (No. 20112072020008).

Key words: wind speed; forecasting; improved extreme learning machine; ensemble empirical mode decomposition; phase space reconstruction

中图分类号: TM715

文献标识码: A

文章编号: 1674-3415(2014)10-0029-06

0 引言

风能作为一种用之不竭、环境友好的可再生能源, 近年来装机容量一直保持较高的年增长速度。然而风能的混沌特性和随机特性使其输出功率的波动速度较快, 波动范围较大。风电并网后对电力系

统调度、电压及无功控制造成困难^[1]。如果能及时有效地对风速进行预测, 不仅可以减少电力系统备用容量、降低系统运行成本, 还可以减轻风电对电网造成的不利影响, 从而提高风电竞争力^[2]。

目前, 国内对风速预测进行了大量研究, 建立的模型^[3-9]主要有: 时间序列模型, 支持向量机模型和神经网络模型等, 取得了较好的预测效果, 然而这些模型同样存在着一些问题。时间序列模型高阶模型参数估计难度大、低阶模型预测精度低。高斯

基金项目: 博士点基金项目 (20110141110032); 教育部中央高校基本科研业务费专项资金资助 (20112072020008)

过程回归模型和支持向量机模型的预测结果在一定程度上受到核函数的选择和参数优化等问题的影响。需要采用较好的智能算法进行模型优化才能取得较好的预测效果^[5]。人工神经网络模型应用较为成熟,作为一种数据驱动算法,具有逼近任意非线性函数的能力,可以映射出序列间复杂的非线性关系,从而在风速和风功率预测中得到广泛应用^[6-8]。然而传统的人工神经网络方法存在一些问题,如算法运行时间长,容易陷入局部极小等。

极限学习机^[10]是huang 等于2006年提出的一类性能优良的新型单隐层前向神经网络(Single-hidden layer feedforward neural networks, SLFNs),该算法在随机选择输入层权值和隐层神经元阈值的前提下,仅通过一步计算即可求得网络输出权值,同传统神经网络相比,极限学习机极大地提高了网络的泛化能力和学习速度,具有较强的非线性拟合能力。经过学者努力,极限学习机已在模式分类^[11]、暂态稳定评估^[12]、非线性拟合预测^[13]等方面得到广泛应用。然而风速序列是一种具有非线性和非平稳性的特殊序列信号,极限学习机虽可以很好地拟合风速的非线性部分,但是风速的非平稳部分会对预测效果造成较大影响,所以降低风速的非平稳性显得尤为重要。集合经验模态分解可以将一个复杂的非平稳性信号分解成不同频率段的信号,从而有效降低序列的非平稳性。针对风速的非平稳性和非线性特性,本文提出一种基于集合经验模态分解和极限学习机的新型风速组合预测模型。实例研究表明,本文所提模型取得了较好的预测效果。

1 改进极限学习机预测模型

1.1 极限学习机原理

极限学习机是一种新型的单隐层前向神经网络^[10-11],对于 N 个不同样本 (x_i, t_i) , $x_i \in R^p$, $t_i \in R^p$, $i=1, \dots, N$, 含有 L 个节点且激发函数为 $f(x)$ 的极限学习机数学模型可表示为

$$\sum_{i=1}^L \beta_i f(w_i \cdot x_j + b_i) = y_j \quad j=1, \dots, N \quad (1)$$

式中: β_i 为连接第 i 个隐层节点和输出神经元之间的连接权向量; $w_i = [w_{i1}, w_{i2}, \dots, w_{im}]^T$ 是连接第 i 个隐层节点和输入节点的权重; b_i 为第 i 个隐层节点的偏差; y_j 为第 j 个节点的输出值。 $w_i \cdot x_j$ 构成 w_i 和 x_j 的内积。网络的训练相当于零误差逼近 N

个训练样本,即存在 \hat{w}_i , \hat{b}_i , $\hat{\beta}$ 。使得

$$\sum_{i=1}^L \hat{\beta}_i f(\hat{w}_i \cdot x_j + \hat{b}_i) = t_j, j=1, \dots, N \quad (2)$$

上式可表示为

$$H\beta = T \quad (3)$$

式(3)中, $H = \begin{bmatrix} f(w_1 x_1 + b_1) & \dots & f(w_L x_1 + b_L) \\ \vdots & & \vdots \\ f(w_1 x_N + b_1) & \dots & f(w_L x_N + b_L) \end{bmatrix}_{N \times L}$;

$\beta = [\beta_1^T, \dots, \beta_L^T]^T_{L \times m}$; $T = [t_1^T, \dots, t_N^T]^T_{N \times m}$ 。 H 为网络隐层输出矩阵, H 的第 i 列表示第 i 个隐层节点对应于输入 x_1, x_2, \dots, x_N 的第 i 个隐层神经元的输出向量。

本文选择基函数神经网络(Radical Basis Function, RBF)作为激励函数,其表达式为 $f(x, w, b) = \exp(-b \|x - w\|^2)$,文献[10]已经证明:当激励函数无限可微时,网络参数并不需要全部进行调整,输入连接权值 w 和隐层节点偏置 b 在训练开始时可随机选择,而输出连接权值可通过求解线性方程组(4)的最小二乘解来求得。

$$\min_{\beta} \|H\beta - T\| \quad (4)$$

其解为

$$\hat{\beta} = H^2 T \quad (5)$$

其中, H^2 为隐层输出矩阵 H 的Moore-Penrose广义逆。

极限学习机的学习算法过程可分为三步:

- (1) 随机设置输入权值 w_i 以及偏置 b_i , $i=1, \dots, N$;
- (2) 计算隐层输出矩阵 H ;
- (3) 计算输出权值 $\hat{\beta}$ 。

相比于传统的神经网络,极限学习机在训练过程中不需要调整 w_i 和 b_i 的值,只需相应的拟合出 β 值,便可获得一个全局最优解,训练速度显著提升,且不会陷入局部最优。

1.2 极限学习机网络结构的确定

极限学习机的拟合性能在一定程度上受到网络结构的影响^[14],对于本文用于风速预测的极限学习机模型来说,输入层神经元个数为相空间重构的饱和输入维数,输出层有一个神经元,唯一不能确定的是隐层神经元个数。如果隐层神经元设计过少,则不能对序列进行很好的拟合,设计过多会造成运算成本的提升。本文运用误差最小极限学习机

算法^[14] (Error Minimized Extreme Learning Machine, EM-ELM) 来确定极限学习机的结构。

给定一个训练集 $D = (x_i, t_i)$, $x_i \in R^n$, $t_i \in R$, $i = 1, \dots, N$ 。激活函数 $g(x)$, 最大的隐节点个数 \tilde{N}_{\max} , 期望的学习误差 ε : 运用 EM-ELM 算法对极限学习机网络结构进行优化的步骤如下^[14]:

Step 1: 初始化过程

(1) 初始化一个有 L_0 个隐节点的单隐层前向神经网络, L_0 为一人为设定小的正整数;

(2) 计算隐藏层输出矩阵 H_1 ;

(3) 计算相应残差 $E(H_1)$ 。

Step 2: 学习过程

本文设置 $k = 0$, $\tilde{N}_{\max} \leq 100$, $\varepsilon = 0.01$,

当 $L_k < \tilde{N}_{\max}$ 且 $E(H_k) > \varepsilon$ 时进行以下迭代:

(1) $k = k + 1$;

(2) 随机增加 δL_{k-1} 个隐节点, 总隐节点数目为: $L_k = L_{k-1} + \delta L_{k-1}$, 相应隐藏层输出矩阵 $H_{k+1} = [H_k, \delta H_k]$;

(3) 计算残差 $E(H_{k+1})$ 。

2 样本设计

2.1 集合经验模态分解^[16]

集合经验模态分解是一种新型的自适应序列分析技术, EEMD改进了经验模态分解^[15] (EMD)的“模态混叠 (mode mixing)”问题。首先介绍经验模态分解的原理和信号分解步骤, 经验模态分解可将存在于原序列中不同特征的趋势逐级筛选出来, 得到具有相同特征的固有模态分量 (Intrinsic mode function, IMF), 固有模态分量需满足如下两个条件: 1) IMF序列中的零点与极点数相差0个或1个; 2) IMF序列中局部极小值点定义的包络线和局部极大值点定义的包络线均值为0。

对于某风功率序列 $\{x(t)\}$, 求得序列 $\{x(t)\}$ 中所有局部极大值和极小值。采用三次样条函数进行插值求出上包络线 $u_{\max}(t)$ 和下包络线 $u_{\min}(t)$ 的局部平均值 $m(t)$, 其中 $m_1(t) = [b_{\max}(t) + b_{\min}(t)] / 2$, 提取 $h_1(t) = x(t) - m_1(t)$, 判断 $h_1(t)$ 是否满足固有模态分量条件, 是则 $h_1(t)$ 就是第一个固有模态分量, 不是则将 $h_1(t)$ 作为原始序列重复上述分解步骤。直到经过 n 次筛选后的差值 $h_n(t)$ 满足固有模态分量条件, 称为一个IMF, 记为 $C_1(t) = h_n(t)$ 。得到 $C_1(t)$ 后, 根据式 (6) 从信号 $x(t)$ 中得到剩余分量 $r_1(t)$:

$$r_1(t) = x(t) - C_1(t) \quad (6)$$

将 $r_1(t)$ 作为原始序列, 重复上述步骤可得到其余的IMF分量, 当余量 $r_n(t)$ 为单调函数或 $|r_n(t)|$ 很小时终止。则原有风速信号可分解为

$$x(t) = \sum_{i=1}^n C_i(t) - r_n(t) \quad (7)$$

经验模态分解具有一定的先进性, 但上述算法在一些情况下会出现模态混叠现象^[16]。集合经验模态分解^[16]的基本思想是利用噪声的统计特性来有效避免混叠现象。集合经验模态分解步骤如下: 首先向原始风速序列 $\{x(t)\}$ 中加入白噪声序列, 加入的白噪声序列应服从 $(0, (\alpha\varepsilon)^2)$ 的正态分布, 其中 α 为噪音的强度参数, ε 为信号的标准差。文献[16]认为: 当 r 为100, α 从[0.1, 0.3]之间选择时能够取得较好的分解结果。因此在本文中 r 等于100, α 等于0.25; 然后使用EMD将加入白噪声后的风功率序列分解为若干个本征模态分量 $c_n(t)$ 和一个剩余分量 $r_N(t)$; 重复上述两个步骤共 r 次, 每次加入的白噪声序列的幅值不同; 最后将 r 次分解得到的IMF值求整体平均, 将其作为原信号的最终IMF分量。

为避免极限学习机模型输入维数选取的任意性和序列信息丢失问题, 对集合经验模态分解得到的各分量进行相空间重构, 并以重构后相空间的饱和嵌入维数 m 作为极限学习机预测模型的输入维数。

2.2 相空间重构

Takens 已经证明了适当选取延迟时间和嵌入维数会使重构后的相空间反映出系统状态随时间演化的规律。对于风速序列 $\{x(i)\}, i = 1, 2, \dots, N$, 求取延迟时间 τ 和嵌入维数 m 是相空间重构的关键所在, 重构相空间已有很多较为成熟的方法。本文运用互信息法^[17]求时间延迟, 虚假紧邻法^[18]求嵌入维数。

3 实例分析

3.1 评判指标

合理的选择预测误差指标对评定模型性能有着重要作用。本文选取平均绝对误差 e_{mac} 、均方根误差 e_{rms} 作为误差评判指标。 e_{mac} 和 e_{rms} 如式 (8)、式 (9) 所示。

$$e_{\text{mac}} = \frac{1}{N} \sum_{i=1}^N |y'(i) - y(i)| \quad (8)$$

$$e_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y'(i) - y(i))^2} \quad (9)$$

3.2 样本设计

以广西金紫山风电场某号机组数据采集与监视控制系统 (Supervisory Control And Data Acquisition, SCADA) 连续 720 h 风速序列作为研究对象。该系统采样周期为 10 min, 为满足研究需要, 对数据进行小时平均化处理。处理后的风速序列如图 1 所示。

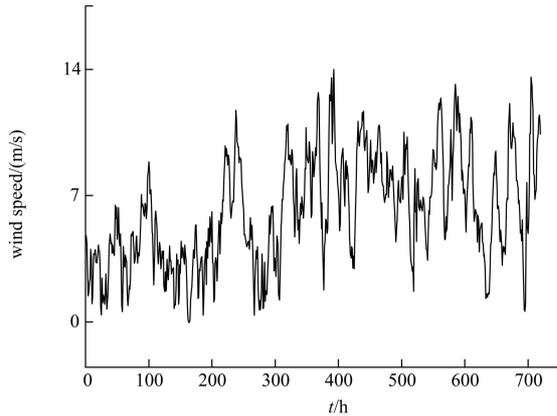


图1 风速时间序列

Fig. 1 Wind speed time series

首先对风速序列进行集合经验模态分解, 分解结果如图 2 所示。

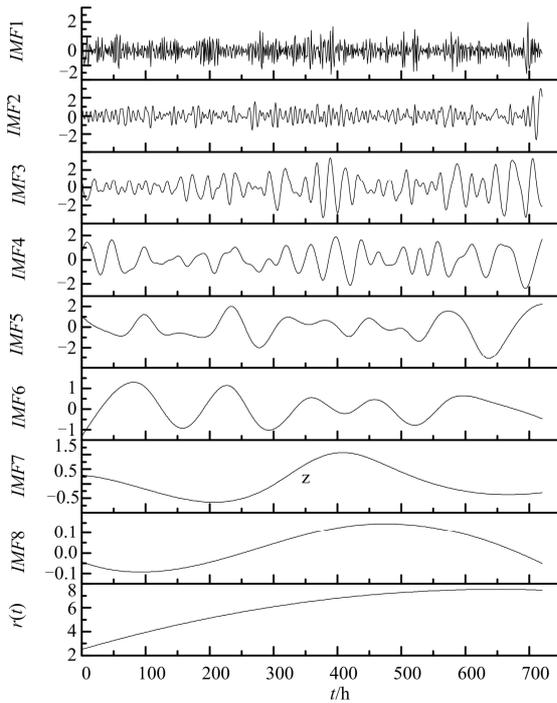


图2 集合经验模态分解结果

Fig. 2 Results of ensemble empirical mode decomposition

然后对各个分量进行相空间重构, 对各个分量分别运用互信息法^[17]求时间延迟, 虚假紧邻法^[18]

求嵌入维数, 求得的结果如表 1 所示。重构相空间后, 每个分量可以构成 T 个相点, $T = N - (m - 1)\tau$, N 为原始数据样本数量。 T 作为重构相空间向量数, 本文选取前 $(T - 100)$ 组相空间重构后的数据作为训练样本, 后 100 组相空间重构后数据作为测试样本来检验预测模型的性能, 进行提前一个小时的风速预测研究。运用 EM-ELM 算法优化各子序列极限学习机的网络结构, 得到的隐含层神经元个数如表 2 所示。

表 1 各分量相空间重构参数表

Table 1 Parameters of phase space reconstruction for each subsequence

分量	延迟时间	嵌入维数
IMF1	7	5
IMF2	8	6
IMF3	11	7
IMF4	13	6
IMF5	11	4
IMF6	14	6
IMF7	17	5
IMF8	19	8
$r(t)$	13	7

表 2 各分量隐含层神经元数

Table 2 Number of hidden layer neurons of each subsequence

分量	隐含层神经元个数
IMF1	24
IMF2	31
IMF3	23
IMF4	19
IMF5	17
IMF6	24
IMF7	16
IMF8	33
$r(t)$	21

3.3 预测结果分析

对各个分量分布采用建立的 EM-ELM 预测模型进行预测, 将各个分类的预测结果叠加得到风速预测结果, 如图 3 所示。

为了对比研究, 本文还同时采用以下两种模型进行风速预测。

第一种模型: 对集合经验模态分解和相空间重构后的各分量分别采用 ELMAN 神经网络进行预测, 然后对各个分量预测值叠加得到风速预测结果。

第二种模型: 采用最小二乘支持向量 (Least

squares support vector machine, LSSVM) 机对各分量分别建模预测, 然后对集合经验模态分解和相空间重构后各分量的预测值叠加得到风速预测结果。其中采用径向基 (RBF) 函数作为 LSSVM 核函数, 并运用网格搜索算法优化 LSSVM 模型超参数。最后将各分量预测结果进行叠加得到风速预测值, 本文称为 LSSVM 模型。

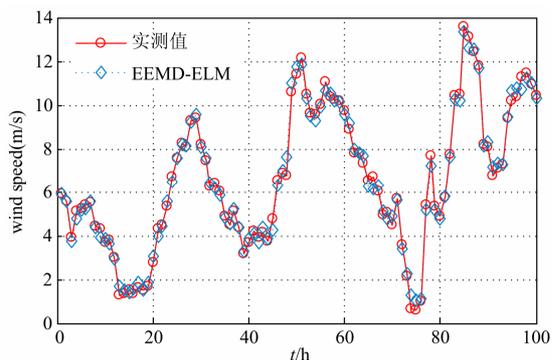


图3 ELM模型风速预测结果

Fig. 3 Wind speed forecasting of the ELM model

对三种模型的预测结果进行误差性能分析, 三种模型的性能指标如表3所示。

表3 模型性能

Table 3 Model performance indicators

模型	e_{mae}	e_{rms}
ELMAN	0.315 8	0.439 5
LSSVM	0.205 0	0.237 4
ELM	0.197 8	0.203 5

从图3可以看出, 本文所提的集合经验分解和改进极限学习机的组合预测模型可以对风速序列进行很好的跟踪, 取得了较好的预测效果。从表3可以看出, 本文模型与其他两种模型的对比研究表明了本文模型的先进性。

4 结论

风速的非平稳性和非线性加大了风速预测模型的拟合难度, 本文结合集合经验模态分解可有效分析非平稳序列的特点和极限学习机良好的非线性拟合能力, 提出一种基于集合经验模态分解和极限学习机的短期风速预测方法, 通过实例研究得到以下结论:

1) 本文提出的集合经验模态分解和极限学习机的短期风速组合预测模型可以有效降低风速序列的非平稳性特征, 使模型可以更好地拟合风速信号, 从而进一步提高算法精度。

2) 极限学习机具有较强的非线性学习能力,

用于风速预测取得了较好的预测结果。

3) 本文所提的预测模型与另外两种模型的对比研究表明, 本文提出的预测模型在风速预测方面的优势, 具有较高的工程应用价值。

参考文献

- [1] 周玮, 孙辉, 顾宏, 等. 含风电场的电力系统经济调度研究综述[J]. 电力系统保护与控制, 2011, 39(24): 148-154.
ZHOU Wei, SUN Hui, GU Hong, et al. A review on economic dispatch of power system including wind farms[J]. Power System Protection and Control, 2011, 39(24): 148-154.
- [2] 胡国伟, 别朝红, 王锡凡. 考虑运行可靠性的含风电电力系统优化调度[J]. 电工技术学报, 2013, 28(5): 58-65.
HU Guo-wei, BIE Zhao-hong, WANG Xi-fan. Optimal dispatch in wind integrated system considering operation reliability[J]. Transactions of China Electrotechnical Society, 2013, 28(5): 58-65.
- [3] CHEN Pei-yuan, PEDERSEN T, BAK-JENSEN B, et al. ARIMA-based time series model of stochastic wind power generation[J]. IEEE Trans on Power Systems, 2010, 25(2): 667-676.
- [4] 罗文, 王莉娜. 风场短期风速预测研究[J]. 电工技术学报, 2011, 26(7): 68-74.
LUO Wen, WANG Li-na. Short-term wind speed forecasting for wind farm[J]. Transactions of China Electrotechnical Society, 2011, 26(7): 68-74.
- [5] 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 Ni-ya, QIAN Zheng, MENG Xiao-feng, et al. Multi-step ahead wind speed forecasting model based on spatial correlation and support vector machine[J]. Transactions of China Electrotechnical Society, 2013, 28(5): 15-21.
- [6] 孙国强, 卫志农, 翟玮星. 基于RVM与ARMA误差校正的短期风速预测[J]. 电工技术学报, 2012, 27(8): 187-193.
SUN Guo-qiang, WEI Zhi-nong, ZHAI Wei-xing. Short term wind speed forecasting based on RVM and ARMA error correcting[J]. Transactions of China Electrotechnical Society, 2012, 27(8): 187-193.

- [7] 吴栋梁, 王扬, 郭创新, 等. 基于改进 GMDH 网络的风电场短期风速预测[J]. 电力系统保护与控制, 2011, 39(2): 88-93.
WU Dong-liang, WANG Yang, GUO Chuang-xin, et al. Short-term wind speed forecasting in wind farm based on improved GMDH network[J]. Power System Protection and Control, 2011, 39(2): 88-93.
- [8] 孙斌, 姚海涛. 基于 PSO 优化 LSSVM 的短期风速预测[J]. 电力系统保护与控制, 2012, 40(5): 85-89.
SUN Bin, YAO Hai-tao. The short-term wind speed forecast analysis based on the PSO-LSSVM predict model[J]. Power System Protection and Control, 2012, 40(5): 85-89.
- [9] 杨洪, 古世甫, 崔明东, 等. 基于遗传优化的最小二乘支持向量机风电场风速短期预测[J]. 电力系统保护与控制, 2011, 39(11): 44-48.
YANG Hong, GU Shi-fu, CUI Ming-dong, et al. Forecast of short-term wind speed in wind farms based on GA optimized LS-SVM[J]. Power System Protection and Control, 2011, 39(11): 44-48.
- [10] HUANG Guang-bin, ZHU Qin-yu, SIEW C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1-3): 489-501.
- [11] HUANG Guang-bin, ZHOU Hong-ming, DING Xiao-jian, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Trans on Systems, Man and Cybernetics-Part B: Cybernetics, 2012, 42(2): 513-528.
- [12] XU Y, DONG Z Y, MENG K, et al. Real-time transient stability assessment model using extreme learning machine[J]. IET Generation, Transmission & Distribution, 2011, 5(3): 314-322.
- [13] 高光勇, 蒋国平. 采用优化极限学习机的多变量混沌时间序列预测[J]. 物理学报, 2012, 61(4): 1-9.
GAO Guang-yong, JIANG Guo-ping. Prediction of multivariable chaotic time series using optimized extreme learning machine[J]. Acta Physica Sinica, 2012, 61(4): 1-9.
- [14] FENG Guo-ru, HUANG Guang-bin, LIN Qing-ping, et al. Error minimized extreme learning machine with growth of hidden nodes and incremental learning[J]. IEEE Trans on Neural Networks, 2009, 20(8): 1352-1357.
- [15] HUANG N E, SHEN Z, LONG S R. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society Soc Land, 1998, 454(1971): 903-995.
- [16] WU Z, HUANG N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009, 1(1): 1-41.
- [17] 吕涛, 唐巍, 所丽. 基于混沌相空间重构理论的风电场短期风速预测[J]. 电力系统保护与控制, 2010, 38(21): 113-117.
Lü Tao, TANG Wei, SUO Li. Prediction of short-term wind speed in wind farm based on chaotic phase space reconstruction theory[J]. Power System Protection and Control, 2010, 38(21): 113-117.
- [18] 廖志强, 李太福, 余德均, 等. 基于相空间重构的神经网络短期风速预测[J]. 江南大学学报: 自然科学版, 2012, 11(1): 15-18.
LIAO Zhi-qiang, LI Tai-fu, YU De-jun, et al. Short-term wind speed forecasting of neural network based on phase space reconstruction[J]. Journal of Jiangnan University: Natural Science Edition, 2012, 11(1): 15-18.

收稿日期: 2013-08-09; 修回日期: 2013-09-27

作者简介:

张翌晖 (1975-), 男, 博士, 高级工程师, 主要研究方向为绿色能源并网技术、全控直流输电、电力系统稳定与控制; E-mail: yhzgh@hotmail.com

王贺 (1984-), 男, 通信作者, 博士研究生, 主要研究方向为风电功率预测及风电并网后电力系统经济调度; E-mail: wanghe0396@126.com

胡志坚 (1969-), 男, 博士, 教授, 博士生导师, 主要研究方向为新能源与分布式发电、互联电力系统鲁棒协调控制等。