

基于样本熵和极端学习机的 超短期风电功率组合预测模型

张学清, 梁军, 张熙, 张峰, 张利, 徐兵

(电网智能化调度与控制教育部重点实验室(山东大学), 山东省 济南市 250061)

Combined Model for Ultra Short-term Wind Power Prediction Based on Sample Entropy and Extreme Learning Machine

ZHANG Xueqing, LIANG Jun, ZHANG Xi, ZHANG Feng, ZHANG Li, XU Bing

(Key Laboratory of Power System Intelligent Dispatch and Control of Ministry of Education (Shandong University),
Jinan 250061, Shandong Province, China)

ABSTRACT: An ultra short-term wind power combined prediction approach based on empirical mode decomposition (EMD)-sample entropy (SE) and extreme learning machine (ELM) was proposed. Firstly, the wind power time series was decomposed into a series of wind power subsequences with obvious differences in complex degree by using EMD-SE. Secondly, the prediction models of each subsequence were constructed with least squares support vector machine (LSSVM), extreme learning machine (ELM) and ELM improved by primal ridge regression (PRR-ELM), of which the parameters and the input vector dimensions were determined by cross validation and chaotic phase space theory to improve the forecasting accuracy of each prediction model. Finally, taking the actual collecting data of certain a wind farm for an example, the simulation results illustrate that ELM and PRR-ELM prediction model based on EMD-SE are much better than the combined LSSVM model based on EMD-SE on forecasting accuracy and training speed, and the prediction results of ELM are closer to the actual value, by which it is possible to achieve the online ultra short-term wind power combined prediction with higher precision.

KEY WORDS: wind power prediction; sample entropy; extreme learning machine; ridge regression; support vector machine

摘要: 该文提出一种经验模态分解(empirical mode

decomposition, EMD)-样本熵(sample entropy, SE)和极端学习机(extreme learning machine, ELM)相结合的风电功率超短期预测方法。该方法首先利用 EMD-SE 将风电功率时间序列分解为一系列复杂度差异明显的风电子序列;其次利用最小二乘支持向量机(least squares support vector machine, LSSVM)、极端学习机和经原始岭回归(primal ridge regression, PRR)改进的极端学习机(PRR-ELM)对各子序列建立组合预测模型,并采用交叉验证法和重构相空间法确定各模型的参数和输入向量维数,以提高各组合模型的预测精度;最后以某一风电场实际采集的数据为算例,结果表明基于 EMD-SE 理论的 ELM 和 PRR-ELM 组合预测模型在预测精度和训练速度上都明显优于 EMD-SE 理论和 LSSVM 的组合模型,且其预测结果更接近于真实值,为实现风电功率在线的较高精度超短期预测提供了可能。

关键词: 风电预测; 样本熵; 极端学习机; 岭回归; 支持向量机

0 引言

由于风电固有的间歇性和随机性,风电大规模并网必将给电网运行与控制带来难度^[1-2]。为了解决风电并网的不确定性问题,可对风电场所发出的功率进行准确的短期预测,这样不但可以节约常规能源和降低旋转备用的容量,提高电网运行的经济性,还可有效提高电网接纳风电的能力^[3-4]。

最初风电功率预测大都基于单一的预测模型^[5-10],从而导致在某些测量点容易产生较大的误差。文献[11]综合多种预测方法,在多嵌入维的基础上提出了风电组合预测模型,取得了比单体预测更好的效果。从此组合预测模型在风电功率短期预

基金项目:国家自然科学基金项目(51177091);山东省自然科学基金项目(ZR2010EM055)。

Projects Supported by National Natural Science Foundation of China (51177091); Projects Supported by Shandong Province Natural Science Foundation (ZR2010EM055).

测得到了普及。目前在风电功率组合预测中应用最广泛的是神经网络和支持向量机(support vector machine, SVM)理论^[12-13]。但是传统的神经网络如BP神经网络等主要基于梯度下降的学习算法,收敛速度慢,容易陷入局部最优,从而制约了其广泛的应用。SVM是一种基于结构风险最小的机器学习算法,具有较好的泛化能力,虽然最小二乘支持向量机(least squares support vector machine, LSSVM)把SVM的学习问题转化为求解一组线性方程组的问题,其在保持原求解算法精度几乎不变的条件下有效的减少了计算时间,但是LSSVM仍要求其核函数必须满足Mercer定理,从而限制了其在时间序列预测中的应用。极端学习机(extreme learning machine, ELM)是最近几年发展起来的一种新型的前馈神经网络方法,具有训练过程简单,抗干扰能力强等优点^[14],且ELM克服了传统神经网络的缺陷,并已被证明了在许多时间序列的预测上优于SVM^[15-16],但ELM在某些条件下训练的过程中自相关矩阵易产生奇异,所以利用原始岭回归(Primal ridge regression, PRR)方法改进ELM(PRR-ELM),其在保持原ELM测试精度几乎不变的条件下具有更快的学习训练速度和较好的算法稳定性,本文尝试将ELM应用于风电功率的超短期组合预测。

本文提出一种基于样本熵和极端学习机的风电功率组合预测模型。针对风电序列的非平稳特性,首先利用经验模态分解(empirical mode decomposition, EMD)将风电功率分解为一系列具有不同特征尺度的子序列;其次利用样本熵(sample entropy, SE)对不同尺度的子序列进行复杂度分析,根据子序列的不同熵值进行归类叠加产生新的子序列。最后利用交叉验证法和重构相空间法确定了学习机的各种参数和输入维数,再利用PRR-ELM、ELM和LSSVM分别对各子序列进行建模预测分析,结果表明基于EMD-SE和ELM组合预测模型无论在预测精度和训练速度上都明显优于EMD-SE理论和LSSVM的组合预测模型,且与真实值的预测误差更小,其为实现风电功率在线的较高精度预测提供了可能。

1 经验模态分解

经验模态分解非常适用于非线性非平稳信号的分析^[13]。为了分析风电功率的局部特征,本文将

经验模态分解应用于风电功率时间序列的分析。文献[17]假设任一信号都是由许多固有模态函数(intrinsic mode function, IMF)组成的,详细内容参见文献[17]。EMD分解具体的处理步骤:

1)先根据序列 $x(t)$ 的极大点和极小点求出其上包络 $v_1(t)$ 及下包络 $v_2(t)$ 之平均值:

$$m = \frac{1}{2}(v_1(t) + v_2(t)) \quad (1)$$

然后考察 $x(t)$ 与 m 之差 h ,即

$$x(t) - m = h \quad (2)$$

2)将 h 视为新的 $x(t)$ 重复以上操作,直到 h 满足IMF条件时,记 $c_1 = h$, c_1 为第一个IMF分量。

$$x(t) - c_1 = r \quad (3)$$

视 r 为新的 $x(t)$,重复以上过程,依次得到 c_2, c_3, \dots ,直到 $r(t)$ 基本成单调趋势或 $|r(t)|$ 很小,视为测量误差时即可停止。于是

$$x(t) = \sum_{i=1}^n c_i + r \quad (4)$$

即把原序列分解成 n 个IMF即 c_1, c_2, \dots, c_n ,和1个剩余分量 r 。式(4)可表征EMD分解的完备性。

通过对2006年6月份美国科罗拉多州某一大型风电场实际采集的风电功率时间序列进行分析,共采集19d的数据,其中每天24h,每10min采集一次数据,风电场的额定装机容量是267MW,为了方便计算,将风电功率转化为标么值并作归一化处理,风电功率时间序列如图1所示。

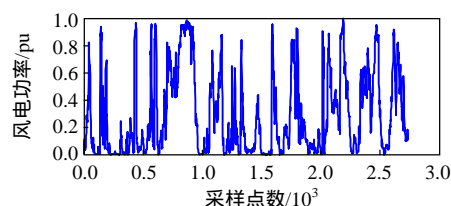


图1 风电功率时间序列

Fig. 1 Wind power time series

2 样本熵

近似熵^[18]可以度量序列的复杂性,且只需要较少的数据就可以得出稳定的数值,但是由于近似熵存在固有的对自身数据段的比较而导致其计算会产生偏差;同时近似熵的值与数据长度有关,一致性较差,所以针对近似熵的缺陷,由Richman^[19]于2000年提出来的与近似熵类似,但精度更好的样本

熵理论能很好的降低近似熵的误差。若序列的自相似性越高，样本熵值就越小，序列越复杂，其样本熵值就越大。样本熵是条件概率的严格自然对数，可用 $\text{SampEn}(N, m, r)$ 表示。其中 N 为长度， r 为相似容限， m 为维数。其具体算法如下^[18]：

假设时间序列 $\{x_i\}$ 为 $x(1), x(2), \dots, x(N)$ (N 为数据总数)：

1) 将序列 $\{x_i\}$ 按顺序组成 m 维矢量，即 $X(i) = [x(i), x(i+1), \dots, x(i+m-1)]$ ，其中 $i = 1, 2, \dots, N-m+1$ 。

2) 定义两者 $X(i)$ 与 $X(j)$ 之间的距离 $d_m(X(i), X(j))$ 为两者对应元素差值最大的一个，即

$$d_m(X(i), X(j)) = \max_{0 \leq k \leq m-1} |x(i+k) - x(j+k)| \quad (5)$$

对于每一个 i 值计算 $X(i)$ 与其余矢量 $X(j)$ ($j = 1, 2, \dots, N-m+1$ ，且 $j \neq i$) 之间的 $d_m(X(i), X(j))$ 。

3) 给定相似容限 r ($r > 0$)，对每一个 i 值统计 $d_m(X(i), X(j)) < r$ 的数目，然后计算其与距离总数 $N-m$ 的比值，记作 $B_i^m(r)$ ，即

$$B_i^m(r) = \frac{1}{N-m} \text{num}\{d_m(X(i), X(j)) < r\} \quad (6)$$

式中 $i = 1, 2, \dots, N-m+1$ ， $j \neq i$ ， num 为 $d_m(X(i), X(j)) < r$ 数目。该过程称为 $X(i)$ 模板匹配过程， $B_i^m(r)$ 表示任一个 $X(j)$ 与模板的匹配概率。

4) $B_i^m(r)$ 的平均值为

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (7)$$

5) 增加维数为 $m+1$ ，重复步骤(1)——(3)，则 $B_i^{m+1}(r)$ 的平均值为

$$B^{m+1}(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^{m+1}(r) \quad (8)$$

样本熵定义为

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \{-\ln(\frac{B^{m+1}(r)}{B^m(r)})\} \quad (9)$$

当 N 取有限值时，上述得出的是样本熵估计值，为

$$\text{SampEn}(N, m, r) = -\ln[B^{m+1}(r)/B^m(r)] \quad (10)$$

SampEn 的取值与 m 和 r 的值有关，但样本熵具有良好的一致性，其熵值增大与减小的趋势并不受 m 和 r 的影响，一般情况下 m 取为 2， r 为 0.1~0.25SD，SD 为时间序列的标准差，本文取 $m=2$ ， $r=0.25\text{SD}$ 。

3 极端学习机

3.1 极端学习机模型

极端学习机是一种新型的前馈神经网络，其表示如下^[20]：设有 N 个训练样本 $\{(x_k, t_k)\}_{k=1}^N$

$$o_k = \omega^T f(W_{\text{in}} x_k + b), \quad k = 1, 2, \dots, N \quad (11)$$

式中： x_k 为输入向量； W_{in} 为连接输入节点和隐层节点的输入权值； b 为隐层偏置； o_k 为网络输出； ω 为连接隐含层与输出层的输出权值； f 为隐含层激活函数，一般取为 Sigmoid 函数； N 为样本数^[14]。

在训练开始时， W_{in} 和 b 随机生成并保持不变，仅需训练确定输出权值 ω 。假设单隐层前馈网络以零误差逼近训练样本，即 $\sum_{k=1}^N \|o_k - t_k\| = 0$ 。则存在

W_{in} 、 b 和 ω 使式(12)成立：

$$\omega^T f(W_{\text{in}} x_k + b) = t_k, \quad k = 1, 2, \dots, N \quad (12)$$

将式(12)写成矩阵形式为 $H\omega = T$ ，其中，

$$H = \begin{bmatrix} f(W_{\text{in}} x_1 + b_1) & \cdots & f(W_{\text{in}} x_1 + b_m) \\ \vdots & \ddots & \vdots \\ f(W_{\text{in}} x_N + b_1) & \cdots & f(W_{\text{in}} x_N + b_m) \end{bmatrix}_{N \times m} \quad (13)$$

式中： H 为极端学习机的隐含层输出矩阵； m 为隐含层节点数； $T = [t_1, t_2, \dots, t_N]^T$ 为期望输出向量。文献[21]已经证明，给定任意小的正数 $\varepsilon > 0$ ，只要单隐层前馈神经网络的隐层节点激活函数无限可微，且隐层节点的个数为 $m \leq N$ ，则直接对网络的输入权值 W_{in} 和偏置 b 随机赋值，就可使其以任意小的正数 ε 的训练误差逼近 N 个训练样本，而通常采用伪逆算法求取 ω 的值。

单隐层前馈神经网络的训练过程等价于求取线性系统 $H\omega = T$ 的最小二乘解：

$$\|H\hat{\omega} - T\| = \min_{\omega} \|H\omega - T\| \quad (14)$$

式中 $\hat{\omega} = H^+ T$ 为 $H\omega = T$ 方程组的极小范数最小二乘解； H^+ 为隐含层输出矩阵 H 的 Moore-Penrose 广义逆。对于前馈神经网络，权值越小泛化能力越强。而在方程 $H\omega = T$ 的所有最小二乘解中， $\hat{\omega}$ 有最小范数， $\|\hat{\omega}\| = \|H^+ T\| \leq \|\omega\|$ ，且

$$\forall \omega \in \{\omega \mid \|H\omega - T\| \leq \|H\hat{\omega} - T\|, \forall z \in \mathbf{R}^{N \times m}\} \quad (15)$$

从式(15)可以看出，ELM 不仅能达到最小训练误差，而且比传统的梯度下降算法泛化能力更强，而且由于矩阵 H 的 Moore-Penrose 广义逆 H^+ 是唯一的，所以解 $\hat{\omega}$ 也是唯一的。

3.2 ELM 与最小二乘支持向量机

LSSVM 是由 SVM 演变而来的,但由于标准 SVM 的训练为二次规划过程,训练复杂且占用时间较长,由文献[17]提出的 LSSVM 把 SVM 的学习问题转化为求解线性方程组问题,极大地提高了运算速度。但是 LSSVM 跟 SVM 一样仍是把非线性输入变量从低维空间映射到高维特征空间进行某种运算,由于在映射的过程中会引起空间维数增加,为了降低计算规模需要采用核函数理论来简化计算,故选择合适的核函数成为 LSSVM 回归问题的关键。

ELM 作为一种新型的前馈神经网络,已被证明具有优良的预测性能^[15-16]。由于在某些情况下,如训练样本数小于隐层节点数、输入权值较大等情况可能会导致 ELM 的自相关矩阵出现奇异,所以常规的伪逆算法不再适用,本文采用岭回归法来代替伪逆算法求解输出权值向量^[22],其主要思路如下。

由式(12)知可以转化为如下的优化问题:

$$\begin{cases} \min_{\omega, \xi} L(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + \frac{C}{2} \sum_{i=1}^N \xi_i^2 \\ \text{s.t. } f(x_i)\omega = t_i - \xi_i, \quad i=1, 2, \dots, N \end{cases} \quad (16)$$

式中 C 为正则项系数,基于 KKT 条件可以表示为如下的优化问题:

$$L = \frac{1}{2} \|\omega\|^2 + \frac{C}{2} \sum_{i=1}^N \xi_i^2 - \sum_{i=1}^N \alpha_i (f(x_i)\omega + \xi_i - t_i) \quad (17)$$

式中 α_i 为 Lagrange 乘子。由 KKT 条件易得

$$\omega = CH^T \xi \Rightarrow \xi = \frac{1}{C} (H^T) \omega \quad (18)$$

$$H\omega - T + \xi = 0 \quad (19)$$

可以推得

$$\omega = \left(\frac{I}{C} + H^T H \right)^{-1} H^T T \quad (20)$$

可见,ELM 与 LSSVM 都是将原始训练数据映射到某一高维特征空间,但是 ELM 的特征空间是可知的而 LSSVM 是未知的,从而导致了 LSSVM 严重依赖核函数,与此相反 ELM 则完全独立于核函数,是一种无核机器学习方法,且经过原始岭回归改进的算法的训练速度比原 ELM 以及 LSSVM 训练更方便快捷,不仅不会出现自相关矩阵奇异的现象,而且具有较好的算法稳定性。

4 基于 SE 和 ELM 的风电功率预测模型

1) 由图 1 可知,风电功率时间序列具有明显的非线性和非平稳性,同时由于 EMD 分解具有较好的处理非线性数据能力,本文采用 EMD 分解对风电功率时间序列进行分解,以产生一系列的不同尺度的固有模态信号(intrinsic mode function, IMF)和剩余分量 r 。

2) 由于 EMD 分解后的 IMF 分量很多,如果直接利用 ELM 对每一个分量分别进行建模预测,会增大计算的规模。采用样本熵理论,对每一 IMF 信号分别进行复杂性评估,得到的结果如图 2 所示。

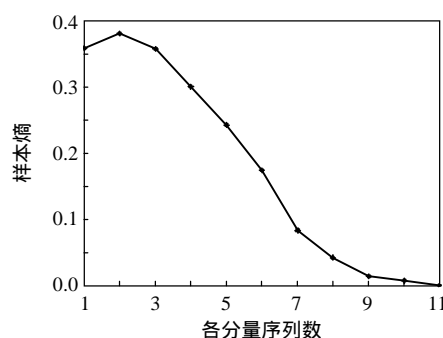


图 2 各 IMF 分量的样本熵

Fig. 2 Sample entropy of each IMF component

从图 2 可知,前 3 个 IMF 分量的样本熵相差不多,说明各 IMF 分量产生新模式的概率大小基本一致,可以把这 3 项叠加作为一个时间序列输入 ELM 进行训练和预测。同理其余 IMF 分量叠加结果如表 1 所示,叠加后的各子时间序列如图 3 所示。

表 1 各 IMF 分量合并为新子序列的结果

Tab. 1 Results of the new subsequences with merged IMF components

新序列号	原 IMF 分量序列号	新序列号	原 IMF 分量序列号
1	1,2,3	4	7,8
2	4,5	5	9,10
3	6	6	11

3) 对上述得到的各子时间序列分别建立 ELM 预测模型进行预测,最后将各子序列的预测结果进行叠加即可得到风电功率预测值。具体预测流程见图 4 所示。

利用 ELM 模型进行预测,需要确定隐层节点的个数,因为隐层节点个数直接决定 ELM 的学习能力^[14];同时利用原始岭回归算法训练 ELM 输出权值需要确定正则项等参数,本文以第 19 日的风电功率各子序列前 24 点作为预测对象。其处理过

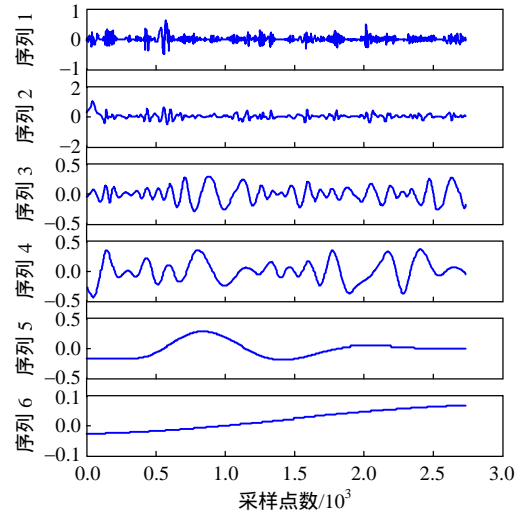


图 3 经 EMD-SE 处理后的风电子序列

Fig. 3 Wind power subsequences processed by EMD-SE

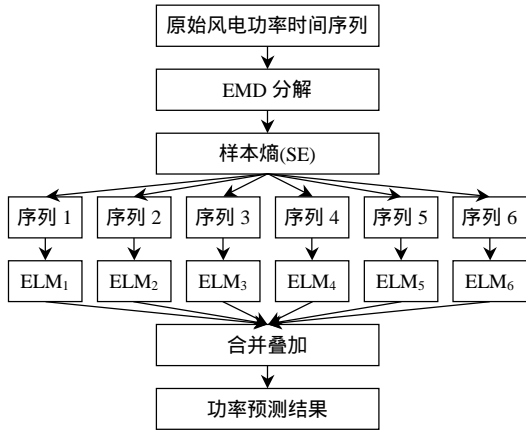


图 4 基于 EMD-SE 和 ELM 组合预测流程

Fig. 4 Combined prediction flowchart based on EMD-SE and ELM

程是每次进行一步预报，再将预测值作为新息加入原子时间序列并重复一步 ELM 即可得 24 点的多步预报，并以一次 24 步预报的均方根误差和训练时间作为评价指标来获取合适的隐层节点个数和正则项系数。均方根误差公式为

$$E_{\text{rmse}} = \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} (x_i - \hat{x}_i)^2} \quad (21)$$

式中： x_i 为实际的风电功率子序列； \hat{x}_i 为预测值； N_m 为预测的时间点数，取为 24。选择 ELM 和 PRR-ELM 参数的主要依据如下：

ELM 的隐层节点数的确定主要依据隐层节点数不同时各子序列的均方根误差，选取一次 24 步最小预测误差的隐层节点数，而且使隐层节点数尽量选择在误差波动比较平缓的区域；PRR-ELM 正则项的确定则固定已确定的隐层节点数，观察不同

正则项时各子序列的均方根误差，并选取一次 24 步最小预测误差时的正则项，各子序列的 ELM 与 PRR-ELM 模型参数如表 2 所示。

表 2 各子序列预测模型的参数

Tab. 2 Parameter of prediction model for each subsequence

序列数	LSSVM 预测		ELM 预测		PRR-ELM 预测		输入维数
	正则项系数	核参数	隐层节点数	正则项系数	隐层节点数	隐层节点数	
1	420	4.6	30	7.2	30		8
2	90	6.4	230	30.6	230		9
3	90	0.8	80	24.4	80		7
4	50	1.6	40	1 336.9	40		5
5	1 200	2.4	340	3.1	340		6
6	120	0.1	420	27.4	420		4

5 算例分析

为验证本文所提出的风电功率组合预测模型的可信性，以美国科罗拉多州某一大型风电场在 2006 年 6 月 1 日到 19 日实际采集的风电功率时间序列进行分析，首先对数据进行了归一化处理，转化为[-1,1]之间的数据，其中前 18 日数据进行模型训练，即前 2 592 个点为训练样本，第 19 日的数据进行测试验证；利用 EMD-SE 处理过的样本数据对 ELM 和 PRR-ELM 进行训练后建立预测模型，以 EMD-SE 和 LSSVM 的风电功率组合预测模型和风电功率实际值进行对比预测分析，说明了本文预测模型的有效性和计算的快速性。本文所有的训练和仿真均在 Matlab 环境下进行，采用 AMD Athlon(tm) 7450 2.4 GHz 双核处理器，2.0 G 内存的计算机平台。

EMD-SE 和 LSSVM 的风电功率组合预测模型流程与前述图 4 的 EMD-SE 和 ELM 的组合预测模型流程类似，只不过把各 ELM 换成 LSSVM 再进行组合预测即可。其也需要对每个子序列进行学习机参数的选择以及输入向量维数的确定。LSSVM 的参数选择也可仿照前面确定 ELM 参数的方法进行确定，LSSVM 的参数见表 2 所示；而输入向量的维数利用重构相空间理论进行确定^[24]，各子序列的输入维数见表 2 所示。

为便于与文献[13]的组合预测进行对比分析，本文也采用提前 12 步的一次多步预测，由于 19 日的测试数据是 144 点，所以要进行 12 次多步预测，即 12 次提前 2 h 的风电功率预测。具体方法就是进行完一次 12 步的多步预测之后，把刚测试完的实

际风电子序列再送入学习机进行训练,再继续下一轮的 12 步预测,如此反复 12 次即可得到 19 日一天的各风电子序列的预测结果,各子序列的各预测模型与真实值的均方根误差以及建立初始模型的训练时间如表 3 所示。最后再将各子序列预测结果进行叠加就可以得到风电功率预测值。利用 ELM、PRR-ELM 组合模型以及 EMD-SE 和 LSSVM 的组合预测模型分别进行风电功率预测,12 次 12 步的多步预测结果以及实际风电功率值如图 5 所示。

表 3 各子序列的预测模型训练时间和测试误差

Tab. 3 Training time and testing errors of prediction model for each subsequence

序列数	LSSVM 预测		ELM 预测		PRR-ELM 预测	
	训练时间/s	E_{rmse}	训练时间/s	E_{rmse}	训练时间/s	E_{rmse}
1	307.141	0.098 520	0.172 0	0.058 39	0.056 21	0.056 21
2	121.859	0.050 720	1.981 0	0.021 75	0.639 60	0.029 33
3	126.265	0.014 570	0.296 4	0.001 06	0.140 40	5.223×10^{-3}
4	39.500	0.000 688	0.218 4	5.060×10^{-4}	0.202 80	5.990×10^{-5}
5	39.624	1.872×10^{-5}	3.541 2	8.489×10^{-5}	0.686 40	7.648×10^{-5}
6	56.8125	4.489×10^{-4}	3.369 6	1.507×10^{-6}	1.216 80	1.386×10^{-5}

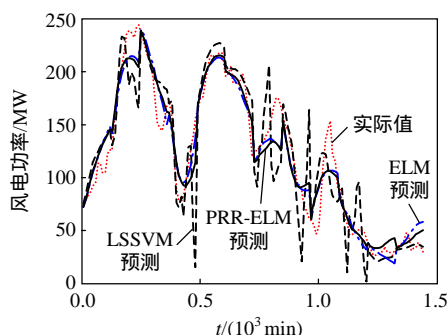


图 5 风电功率实际值以及各模型的提前 2 h 预测值

Fig. 5 Actual value of wind power and prediction results ahead 2 hours of each prediction model

为定量的评价各方法的预测结果与真实值接近程度,以国际上普遍采用的归一化绝对平均误差、归一化均方根误差和最大绝对误差为依据来评价各模型的预测性能。

1) 归一化绝对平均误差 E_{nmae}

$$E_{nmae} = \frac{1}{P_{cap}} \frac{1}{N_t} \sum_{i=1}^{N_t} |x_i - \hat{x}_i| \quad (22)$$

2) 归一化均方根误差 E_{nrmse}

$$E_{nrmse} = \frac{1}{P_{cap}} \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} (x_i - \hat{x}_i)^2} \quad (23)$$

3) 最大绝对误差 E_{mae}

$$E_{mae} = \frac{1}{P_{cap}} \max(|\hat{x}_i - x_i|), \quad i = 1, 2, \dots, N_t \quad (24)$$

式中: N_t 为测试的时间点数为 144; P_{cap} 为风电场的额定装机容量。表 4 分别列出了 3 种预测模型的各种误差性能指标。

表 4 各组合预测模型的预测误差比较

Tab. 4 Compared errors of each prediction model

方法名称	归一绝对平均 误差/%	归一均方根 误差/%	最大绝对 误差/%
LSSVM 组合预测	7.920 404 6	11.622 032 5	52.665 240 4
ELM 组合预测	4.878 986 4	6.392 145 9	17.362 475 8
PRR-ELM 组合预测	4.983 528 8	6.606 015 8	17.712 809 0

从图 5、表 3 和表 4 可以得知,基于 EMD-SE 的 PRR-ELM 组合模型和 ELM 组合模型作提前 2h 的风电功率预测值与实际值吻合较好,2 种模型的 3 种误差相差不大,且均比目前在风电预测中采用的 EMD 和 LSSVM 的组合预测模型的预测误差要小,归一化均方根误差仅为 6.39% 和 6.6%,而 LSSVM 组合预测则为 11.6%。2 种模型在风电波动的拐点处能紧跟风电功率的变化趋势,而且在风电急剧变化时也能很快的做出反应,而 LSSVM 组合模型虽然也能跟踪风电的变化趋势,但是效果明显不及 ELM 和 PRR-ELM 组合模型,而且 PRR-ELM 和 ELM 组合模型在序列的训练时间上都要明显小于 LSSVM 组合模型;由此可见本文提出的方法不仅在整体上提高了预测精度,而且在样本的训练时间上也远小于目前普遍采用的 SVM 组合预测模型,其为实现风电在线的较高精度超短期预测提供了可能。

虽然本文的方法实现了较高精度的风电功率预测,但是在风电波动特别剧烈的时刻仍然会产生较大的误差,例如图 5 最后一个波峰的时段,这是由于风电功率本身的非平稳性造成的,如果只分析历史数据可以获得风电波动的趋势,但是在风电剧烈波动时要获得较为准确的波动大小较为困难,若能在预测时加入天气预报等风电场周围的物理信息就能有效的减小误差,提高风电预测的精度。

6 结论

针对风电功率的超短期预测问题,本文提出了一种样本熵和 ELM 的组合预测模型,其比目前在风电超短期预测中采用的 EMD 和 LSSVM 的组合

模型具有更高的预测精度和更少的样本学习时间，主要有以下 3 点：

1) 利用小波变换对风电功率进行局部分析需解决优先确定小波基的困难，本文采用 EMD 分解将风电序列作多尺度分解；同时 EMD 分解的 IMF 分量数目较多且不固定，利用样本熵对各 IMF 分量进行复杂度分析，取熵值相近的 IMF 分量进行叠加，产生一组复杂度差异明显的子序列，减小计算规模。

2) 对复杂度差异明显的子序列分别建立预测模型。分别采用交叉验证法和重构相空间法确定各预测模型的隐层节点数、正则项系数以及输入维数等参数，以利于提高风电预测精度。

3) 通过与目前采用的 EMD 和 LSSVM 的风电组合预测模型对比分析发现，该方法不仅整体上提高了预测精度，预测结果更接近于真实值，而且减少了样本的学习时间，为实现风电功率的在线较高精度超短期预测提供了可能。

参考文献

- [1] 高爽, 冬雷, 高阳, 等. 基于粗糙集理论的中长期风速预测[J]. 中国电机工程学报, 2012, 32(1): 32-37.
Gao Shuang, Dong Lei, Gaoyang, et al. Mid-long term wind speed prediction based on rough set theory[J]. Proceedings of the CSEE, 2012, 32(1): 32-37(in Chinese).
- [2] Ernst B, Oakleaf B, Ahlstrom M L, et al. Predicting the wind[J]. IEEE Transactions on Power & Energy Magazine, 2007, 11(2): 79-89.
- [3] 张丽英, 叶廷路, 辛耀中, 等. 大规模风电接入电网的相关问题及措施[J]. 中国电机工程学报, 2010, 30(25): 1-9.
Zhang Liying, Ye Tinglu, Xin Yaozhong, et al. Problems and measures of power grid accommodating large scale wind power[J]. Proceedings of the CSEE, 2010, 30(25): 1-9(in Chinese).
- [4] 刘吉臻, 柳玉, 曾德良, 等. 单一风电场的短期负荷调度优化策略[J]. 中国科学: 技术科学, 2012, 42(4): 437-442.
Liu Jizhen, Liu Yu, Zeng Deliang, et al. Optimal short-term load dispatch strategy in wind farm[J]. Sci. China Tech. Sci., 2012, 55: 1140-1145.
- [5] 范高峰, 王伟胜, 刘纯, 等. 基于人工神经网络的风电功率预测[J]. 中国电机工程学报, 2008, 28(34): 118-123.
Fan Gaofeng, Wang Weisheng, Liu Chun, et al. Wind power prediction based on artificial neural network[J]. Proceedings of the CSEE, 2008, 28(34): 118-123(in Chinese).
- [6] 冯双磊, 王伟胜, 刘纯等. 风电场功率预测物理方法研究[J]. 中国电机工程学报, 2010, 30(2): 1-6.
Feng Shuanglei, Wang Weisheng, Liu Chun, et al. Study on the physical approach to wind power prediction[J]. Proceedings of the CSEE, 2010, 30(2): 1-6(in Chinese).
- [7] 杨秀媛, 肖洋. 风电场风速和发电功率预测研究[J]. 中国电机工程学报, 2005, 25(11): 1-5.
Yang Xiuyan, Xiao Yang. Wind speed and generated power forecasting in wind farm[J]. Proceedings of the CSEE, 2005, 25(11): 1-5(in Chinese).
- [8] 潘迪夫, 刘辉. 风电场风速短期多步预测改进算法[J]. 中国电机工程学报, 2008, 28(26): 87-91.
Pan Difu, Liu Hui. Optimization algorithm of short-term multi-step wind speed forecast. Proceedings of the CSEE, 2008, 28(26): 87-91(in Chinese).
- [9] 蔡凯, 谭伦农, 李春林, 等. 时间序列与神经网络法相结合的短期风速预测[J]. 电网技术, 2008, 32(8): 82-85.
Cai Kai, Tan Lunnong, Li Chunlin, et al. Short-term wind speed forecasting combining time series and neural network method[J]. Power System Technology, 2008, 32(8): 82-85(in Chinese).
- [10] 孙春顺, 王耀南, 李欣然. 小时风速的向量自回归模型及应用[J]. 中国电机工程学报, 2008, 28(14): 112-117.
Sun Chunshun, Wang Yaonan, Li Xinran. A vector autoregression model of hourly wind speed and its application in hourly wind speed forecasting[J]. Proceeding of the CSEE, 2008, 28(14): 112-117(in Chinese).
- [11] 王丽婕, 冬雷, 胡国飞, 等. 基于多嵌入维数的风力发电功率组合预测模型[J]. 控制与决策, 2010, 25(4): 576-580.
Wang Lijie, Dong Lei, Hu Guofei, et al. Combined prediction of wind power generation in multi-dimension embedding phase space[J]. Control and Decision, 2010, 25(4): 576-580(in Chinese).
- [12] 王丽婕, 冬雷, 廖晓钟, 等. 基于小波分析的风电场短期发电功率预测[J]. 中国电机工程学报, 2009, 29(28): 30-33.
Wang Lijie, Dong Lei, Liao Xiaozhong, et al. Short-term power prediction of a wind farm based on wavelet analysis[J]. Proceeding of the CSEE, 2009, 29(28): 30-33(in Chinese).
- [13] 叶林, 刘鹏. 基于经验模态分解和支持向量机的短期风电功率组合预测模型[J]. 中国电机工程学报, 2011, 31(31): 102-108.
Ye Lin, Liu Peng. Combined model based on EMD-SVM for short-term wind power prediction[J]. Proceeding of the CSEE, 2011, 31(31): 102-108(in Chinese).
- [14] 王新迎, 韩敏. 基于极端学习机的多变量混沌时间序列预测[J]. 物理学报, 2012, 6(8): 080507-1-080507-9.

- Wang Xinying, Han Min. Multivariate chaotic time series prediction based on extreme learning machine[J]. Acta Phys. Sin., 2012, 6(8): 080507-1-080507-9(in Chinese).
- [15] Huang G B, Ding X J, Zhou H M. Optimization method based extreme learning machine for classification[J]. Neurocomputing, 2010, 74(1-3): 155-163.
- [16] Huang G B, Zhou H M, Ding X J, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Trans. on System, Man and Cybernetics-Part B: Cybernetics, 2012, 42(2): 513-529.
- [17] 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 Lond., 1998, 454(1971): 903-995.
- [18] 彭涛, 陈晓宏, 庄承彬. 基于样本熵的东江月径流序列复杂性分析[J]. 生态环境学报, 2009, 18(4): 1379-1382. Peng Tao, Chen Xiaohong, Zhuang Chengbin. Analysis on complexity of monthly runoff series based on sample entropy in the Dongjiang river[J]. Ecology and Environmental Sciences, 2009, 18(4): 1379-1382(in Chinese).
- [19] Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy[J]. American Journal of Physiological Heart and Circulatory Physiology, 2000, 278(6): 2039-2049.
- [20] 袁琦, 周卫东, 李淑芳, 等. 基于 ELM 和近似熵的脑电信号检测方法[J]. 仪器仪表学报, 2012, 33(3): 514-519. Yuan Qi, Zhou Weidong, Li Shufang, et al. Approach of EEG detection based on ELM and approximate entropy[J]. Chinese Journal of Scientific Instrument, 2012, 33(3): 514-519(in Chinese).
- [21] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and application[J]. Neurocomputing, 2006, 70(1-3): 489-501.
- [22] Hoerl A E, Kennard R W. Ridge regression: biased estimation for nonorthogonal problems[J]. Technometrics, 1970, 12(1), 55-67.
- [23] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.
- [24] 韩敏. 混沌时间序列预测理论与方法[M]. 北京: 中国水利水电出版社, 2007: 42-93. Han Min. Prediction theory and method of chaotic times series[M]. Beijing: China Waterpower Press, 2007: 42-93(in Chinese).



张学清

收稿日期: 2012-09-10。

作者简介:

张学清(1982), 男, 博士研究生, 主要从事电力系统运行、调度与控制技术的研究, zxq_wy@163.com;

梁军(1956), 男, 教授, 博士生导师, 主要从事电力系统运行与控制技术的研究;

张熙(1982), 男, 博士研究生, 主要从事电力系统运行与控制方面的研究;

张峰(1983), 男, 讲师, 博士, 主要从事电力系统运行和新能源调度控制方面的研究;

张利(1967), 女, 副教授, 硕士生导师, 主要从事电力系统运行与控制以及电力市场方面的研究;

徐兵(1988), 男, 硕士研究生, 主要从事电力系统运行、控制及建模方面的研究。

(责任编辑 王剑乔)

Combined Model for Ultra Short-term Wind Power Prediction Based on Sample Entropy and Extreme Learning Machine

ZHANG Xueqing, LIANG Jun, ZHANG Xi, ZHANG Feng, ZHANG Li, XU Bing

(Shandong University)

KEY WORDS: wind power prediction; sample entropy; extreme learning machine; ridge regression; support vector machine

Short-term wind power prediction is critical to economic operation of power systems. Accurate prediction can not only save conventional energy sources and reduce the spinning reserve capacity, but also enlarge wind power accommodation in power systems. Therefore, more efforts need to be made to improve the performance of short-term wind power prediction.

Conventional wind power forecasting methods such as neural network and Least Squares Support Vector Machine are not suitable for the current wind power prediction. In order to improve the accuracy of wind power prediction, a combined approach for ultra short-term wind power prediction is proposed based on empirical mode decomposition (EMD)-sample entropy (SE) and extreme learning machine (ELM). Using EMD, the wind power time series set is decomposed into several subsequences, which should be made complex degree analysis by Sample Entropy (SE), then on the basis of entropy for each subsequence, a new series are formed by classifying and superimposing the primal wind power subsequences.

Extreme Learning Machine (ELM) is a new kind of feedforward neural network, which can be expressed in the matrix form as: $H\omega = T$,

where:

$$H = \begin{bmatrix} f(W_{in}x_1 + b_1) & \cdots & f(W_{in}x_1 + b_m) \\ \vdots & \ddots & \vdots \\ f(W_{in}x_N + b_1) & \cdots & f(W_{in}x_N + b_m) \end{bmatrix}_{N \times m} \quad (1)$$

is the hidden layer output matrix of ELM, m is the number of hidden layer nodes, $T = [t_1, t_2, \dots, t_N]^T$ is the desired output vector. Generally, the autocorrelation matrix of ELM is made singular in some cases, in which the number of training samples is less than the hidden nodes or input weight is much large, thereby, the output weight vector is solved by primal ridge regression (PRR) instead of the conventional pseudo-inverse algorithm.

$$\omega = (I / C + H^T H)^{-1} H^T T \quad (2)$$

where C is regularization coefficient.

The wind power time series set of a certain large wind farm in Colorado, USA, which consists of data sampled from June 1st 0:00 to 19th 24:00, 2006, is used as

raw data to verify the proposed method. Data of the first 18 days is used for model training and data of the 19th day is used for tests. The sample data sets processed by EMD-SE are used to train LSSVM, ELM and ELM improved by primal ridge regression (PRR-ELM) to constitute the prediction model, and their parameters and input vector dimensions are determined by the cross validation and chaotic phase space theory. Combined prediction models of LSSVM, ELM and PRR-ELM are adopted to forecast wind power respectively. The results of the 2-hour ahead prediction for 12 times and the actual wind power value are shown in Fig.1 and Tab.1.

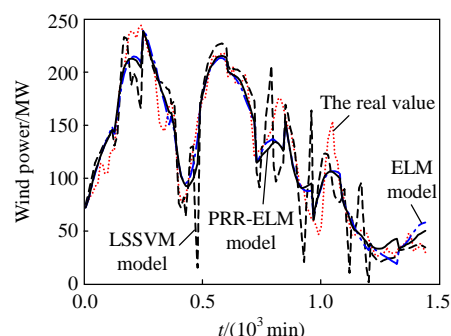


Fig. 1 The actual value of wind power and the 2-hour ahead prediction results of each prediction model

Tab. 1 Compared errors of each prediction model

Model name	$E_{mae}/\%$	$E_{nmse}/\%$	$E_{mse}/\%$
LSSVM model	7.920 404 6	11.622 032 5	52.665 240 4
ELM model	4.878 986 4	6.392 145 9	17.362 475 8
PRR-ELM model	4.983 528 8	6.606 015 8	17.712 809 0

From Fig 1 and Tab.1, it can be seen that the 2-hour ahead prediction results produced by the combined models of PRR-ELM and ELM based on EMD-SE are in good agreement with the real wind power. The three kinds of errors of the above two models have little difference, which are all less than those of the combined prediction models of EMD and LSSVM that are currently used in wind power prediction. Meanwhile, the training time of the combined models of PRR-ELM and ELM is also significantly less than that of LSSVM. Thereby, it is possible to use the proposed method online to achieve more accurate ultra short-term wind power prediction.