

# 基于 EEMD-近似熵和储备池的风电功率混沌时间序列预测模型\*

张学清<sup>1)†</sup> 梁军<sup>1)2)</sup>

1) (山东大学电气工程学院, 济南 250061)

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

(2012年8月20日收到; 2012年10月30日收到修改稿)

针对风电功率时间序列的混沌特性, 提出了一种基于集成经验模态分解 (ensemble empirical mode decomposition, EEMD)-近似熵和回声状态网络 (echo state network, ESN) 的风电功率混沌时间序列组合预测模型. 首先为降低对风电功率局部分析的计算规模以及提高预测的准确性, 利用 EEMD-近似熵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; 然后对各子序列分别建立 ESN、经过高频分量正则化改进的 EEMD-ESN 模型和最小二乘支持向量机预测模型; 最后以某一风电场实际采集的数据为算例, 仿真结果表明 EEMD-ESN 模型在训练速度和预测精度上优于最小二乘支持向量机模型, 为实现风电功率短期预测的在线工程应用提供了新的有益参考.

**关键词:** 混沌时间序列, 风电预测, 集成经验模态分解, 近似熵

**PACS:** 05.45.Tp, 88.50.Mp

**DOI:** 10.7498/aps.62.050505

## 1 引言

风电作为一种清洁的可再生资源, 其开发和应用日益得到广泛的关注. 但风电具有间歇性和波动性, 其大规模的接入电网必将对电网的经济运行产生影响. 而对风电输出功率进行准确的预测, 不但可以提高电网接纳风电的能力, 而且还可以减轻风电接入对电网的影响, 提高电网运行的经济性<sup>[1-4]</sup>.

目前对风电功率进行短期预测主要包括两种方法, 一种是物理的方法<sup>[5]</sup>, 另一种是统计的方法<sup>[6,7]</sup>. 文献[8]通过分析风电功率时间序列的最大 Lyapunov 指数, 揭示了风电功率时间序列具有明显的混沌特性, 并利用经加权马尔科夫链校正的 Volterra 滤波器法进行多步预测, 取得了较好的结果. 由于风电功率时间序列具有明显的非线性和非平稳性, 非线性时间序列的研究方法<sup>[9-12]</sup>逐渐在

风电预测领域得到推广. 文献[13]提出了经验模态分解 (empirical mode decomposition, EMD) 与支持向量机的短期风电功率组合预测模型, 通过与单一预测模型进行对比分析, 验证了组合预测模型具有更高的预测精度. 近年来, 神经网络和 SVM 理论<sup>[14,15]</sup>在风电混沌时间序列预测中得到了广泛的应用, 但是传统的神经网络如 BP 神经网络以及较传统神经网络具有更高泛化能力的 SVM 都因为存在学习训练速度慢等问题制约了其广泛的应用. 文献[16]在 SVM 理论的基础上提出了最小二乘支持向量机 (least squares support vector machine, LSSVM) 理论, 其把标准 SVM 二次规划学习训练问题转化为求解一组线性方程组问题, 极大得提高了运算速度. 但是 LSSVM<sup>[17-19]</sup>跟 SVM 一样仍需要核函数理论且须满足 Mercer 定理, 同时其一般不具有记忆能力, 限制了其在混沌序列预测中的应用<sup>[20]</sup>.

\* 国家自然科学基金 (批准号: 51177091) 和山东省自然科学基金 (批准号: ZR2010EM055) 资助的课题.

† 通讯作者. E-mail: zxq\_wy@163.com

2004年, Jaeger发表了关于储备池 (Reservoir) 的报道后, 其立刻引起了国内外学者的广泛关注<sup>[21]</sup>. 储备池作为一种新型的递归神经网络结构, 是回声状态网络 (echo state network, ESN) 的主要组成部分, 在稳定性和网络训练方面都优于传统神经网络, 已经在水文等领域的混沌时间序列预测取得了很好的应用效果. 但是由于 ESN 在某些条件下训练过程中状态自相关矩阵易出现奇异, 近几年提出了通过加入随机噪声以及正则项等方法克服 ESN 状态矩阵的病态问题, 但是这些方法本质上都是通过引入额外干扰项来降低状态自相关矩阵特征值的离散度, 引入的干扰项不仅增加了不相关的信息, 而且相对于 ESN 的传统伪逆算法也增加了学习的复杂度<sup>[22]</sup>. 本文利用 EEMD 分解得到的风电功率高频分量作为 ESN 的干扰项, 简记为 EEMD-ESN, 起到正则化作用, 以克服 ESN 的不适定问题, 在无需引入额外的干扰条件下提高风电功率预测的精度.

基于上述, 本文提出了一种 EEMD-近似熵和 ESN 的风电功率混沌时间序列预测模型. 首先利用集成经验模态分解 (ensemble empirical mode decomposition, EEMD) 将非平稳的风电功率时间序列分解为一系列具有不同特征尺度的固有模态信号 (intrinsic mode function, IMF); 其次利用近似熵对各 IMF 分量进行复杂度分析, 依据不同的熵值进行组合叠加产生了新的子序列; 最后利用 ESN, 经过高频分量正则化改进 EEMD-ESN 以及 LSSVM 模型分别对各子序列进行建模分析, 以某一实际风电场风电功率数据为算例, 仿真结果表明 EEMD-ESN 模型在预测精度和训练速度上都优于 LSSVM 模型, 其为实现风电功率短期预测提供了新的有益参考.

## 2 集成经验模态分解 (EEMD)<sup>[23]</sup>

针对风电功率混沌时间序列的非平稳的特点, 本文采用集成经验模态分解 (EEMD) 对其进行信号局部特征的分析. EEMD 是在经验模态分解 (empirical mode decomposition, EMD) 的基础上发展起来的, EMD 分解<sup>[24]</sup>吸取了小波多分辨分析的优点, 且克服了小波变换需要实现选择小波基的困难,

是一种自适应的信号分解方法, 目前在故障诊断和生物工程等许多领域得到了广泛的应用. 但是标准 EMD 分解中, 其对所有极大值和极小值利用三次样条插值求取包络线, 由于信号的端点不一定是极值点, 所以包络线在端点会出现发散现象, 从而导致序列分解结果的不准确<sup>[25]</sup>. 对此文献<sup>[26]</sup>提出了利用集成经验模态分解来改进 EMD 分解的边界效应问题, 其通过添加辅助噪声来消除原序列的间歇现象, 并利用随机白噪声通过多次试验相抵消的特性抑制分解结果中噪声所产生的影响.

EEMD 分解步骤为:

1) 在原序列中加入相对均方根较小的高斯白噪声序列, 表示为

$$x_m(t) = x(t) + n_m(t), \quad (1)$$

式中,  $x(t)$  为原风电功率序列,  $n_m(t)$  为白噪声序列,  $x_m(t)$  为混有噪声信号后待分解的序列.

2) 将混有白噪声的序列  $x_m(t)$  利用 EMD 分解为  $k$  个 IMF 分量  $c_{im}, i = 1, 2, \dots, k$ .

3) 每次加入均方根相等的不同白噪声序列, 重复步骤 (1) 和 (2), 得到  $n$  组不同的 IMF 分量,  $m = 1, 2, \dots, n$ .

4) 计算  $n$  组 IMF 分量的均值, 把分解后得到的各 IMF 的均值作为 EEMD 最终的分解结果.

$$c_i = \sum_{m=1}^n c_{i,m}/n. \quad (2)$$

EEMD 分解利用辅助分析的白噪声信号通过统计集成平均而相互抵消, 使 EMD 方法成为真正的二进制滤波器组<sup>[23]</sup>, 是对 EMD 分解的重要改进.

## 3 近似熵

Pincus<sup>[27]</sup>提出了度量序列复杂度的方法——近似熵, 其是用来定量的描述事件序列的复杂性的非线性动力学参数, 只需要较少的数据就可以得出稳定的数值, 比较适合工程应用, 而且近似熵的大小与序列的复杂程度有关, 产生新模式的概率越大, 序列越复杂所对应的熵值也越大.

给定时间序列  $\{x(i), i = 1, 2, \dots, N\}$ , 按以下步骤计算近似熵<sup>[28]</sup>:

1) 将序列  $\{x(i)\}$  按顺序组成  $m$  维矢量, 即

$$X(i) = [x(i), x(i+1), \dots, x(i+m-1)], \quad (3)$$

其中  $i = 1, 2, \dots, N - m + 1$ .

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

$$d_m[\mathbf{X}(i), \mathbf{X}(j)] = \max_{0 \sim m-1} |x(i+k) - x(j+k)|. \quad (4)$$

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

3) 给定相似容限  $r (r > 0)$ , 对每一个  $i$  值统计  $d_m[\mathbf{X}(i), \mathbf{X}(j)] < r$  的数目, 然后计算其与距离总数  $N - m + 1$  的比值, 记作  $C_i^m(r)$ , 即

$$C_i^m(r) = \frac{1}{N - m + 1} \text{num} \{d_m[\mathbf{X}(i), \mathbf{X}(j)] < r\}, \quad (5)$$

( $i, j = 1, 2, \dots, N - m + 1, i \neq j$ ), num 为数目.

4) 将比值  $C_i^m(r)$  取对数, 再求其对所有  $i$  的平均值, 记作  $\Phi^m(r)$  为

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r). \quad (6)$$

5) 增加维数为  $m + 1$ , 重复步骤 1)–4), 求得  $C_i^{m+1}(r)$  和  $\Phi^{m+1}(r)$ .

6) 近似熵  $ApEn(m, r)$  定义为

$$ApEn(m, r) = \lim_{N \rightarrow \infty} [\Phi^m(r) - \Phi^{m+1}(r)]. \quad (7)$$

当  $N$  取有限值时, 可得到近似熵的估计值

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r). \quad (8)$$

$ApEn$  的取值与  $m, r$  和  $N$  的值有关, 但对  $N$  的依赖程度最小, 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  $m = 2$ ,  $r = 0.2SD$ . 从上可以看出, 近似熵反映了时间序列在模式上的自相似程度.

## 4 回声状态网络 (ESN)

### 4.1 回声状态网络的模型结构

ESN 作为一种新型的递归神经网络, 其核心是储备池, 其在模型构建和学习算法方面与传统递归神经网络有很大的差异, 既具有传统神经网络的强大非线性映射能力, 又克服了传统神经网络的学习训练过程易陷入局部最优的缺陷<sup>[29]</sup>. ESN 由输入层、储备池和输出层等三层构成. 各层之间分别通

过输入连接权值和输出连接权值进行连接. 储备池位于输入层和输出层中间, 是由大量神经元稀疏连接构成的. 其网络方程为

$$x(k+1) = f[\mathbf{W}\mathbf{x}(k) + \mathbf{W}_{in}\mathbf{u}(k) + \mathbf{W}_{fb}\mathbf{y}(k) + \mathbf{v}(k)], \quad (9)$$

$$\mathbf{y}(k+1) = g\{\mathbf{W}_{out}[\mathbf{x}(k+1), \mathbf{u}(k+1)]\}, \quad (10)$$

式中  $\mathbf{x}(k)$ ,  $\mathbf{u}(k)$  和  $\mathbf{y}(k)$  分别为状态变量, 输入激励和期望输出信号;  $\mathbf{W}$ ,  $\mathbf{W}_{in}$ ,  $\mathbf{W}_{fb}$  和  $\mathbf{W}_{out}$  分别为内部连接权值矩阵、输入权值矩阵、输出反馈矩阵和输出权值矩阵;  $f$  和  $g$  分别为储备池处理单元和输出单元的激活函数, 本文取为双曲正切函数. 其中  $\mathbf{W}$ ,  $\mathbf{W}_{in}$ ,  $\mathbf{W}_{fb}$  都是随机生成, 一经产生就保持不变, 且在回声状态网络的训练过程中只需要调整输出权值. 储备池主要由以下参数表示, 包括储备池维数、内部连接矩阵的稀疏度和谱半径等. 其中网络的稳定性可以通过在网络训练之前预先设定储备池的谱半径来保证, 只要谱半径小于 1, ESN 就是渐进稳定的<sup>[20]</sup>. 而网络输出是唯一待定的参数, 使用线性回归即可计算出输出权值, 一般采用伪逆方法求解.

设储备池的状态矩阵为  $\mathbf{A} = [\mathbf{x}(n+1), \mathbf{x}(n+2), \dots, \mathbf{x}(n+m)]$ , 期望输出  $\mathbf{Y} = [y(n+1), y(n+2), \dots, y(n+m)]$ , 其中  $\mathbf{A}$  的维数为  $m \times N$ ,  $n$  为初始选择点, 则

$$\mathbf{A}\mathbf{W}_{out} \approx \mathbf{Y} \Rightarrow \mathbf{W}_{out} = (\mathbf{A}^T\mathbf{A})^\dagger \mathbf{A}^T\mathbf{Y}, \quad (11)$$

其中“ $\dagger$ ”表示矩阵的伪逆. 但是 ESN 在学习过程中经常会出现  $\mathbf{A}^T\mathbf{A}$  接近奇异, 结果使  $\mathbf{A}$  的特征值分散度较大, 则输出权值  $\mathbf{W}_{out}$  往往会过大, 影响网络的泛化能力<sup>[22]</sup>.

### 4.2 EEMD-ESN 风电功率预测模型

实际操作中  $\mathbf{A}^T\mathbf{A}$  常会接近奇异, 误差估计非常保守, 即

$$\|\mathbf{W}_{out} - \hat{\mathbf{W}}_{out}\| \leq \|(\mathbf{A}^T\mathbf{A})^\dagger \mathbf{A}^T\| \|\mathbf{A}^T\hat{\mathbf{W}}_{out} - \mathbf{Y}\|. \quad (12)$$

如果在储备池状态变量上加入小的噪声, 可以改善解的性能<sup>[31]</sup>. 为了消除奇异矩阵等不适定问题, Jeager<sup>[30]</sup> 采用噪声抖动技术来减小状态矩阵  $\mathbf{A}$  的特征值分散度, 即 (9) 式的向量  $\mathbf{v}$  为噪声项, 其

在一定程度上缓解了不适定问题,是目前认为解决 ESN 不适定问题的主要方法. 后来,文献 [20, 31] 提出储备池的正则化学习方法. 即通过截断方法直接处理病态的系数矩阵来舍弃较小的奇异值,再采用岭回归的方法改进因子化矩阵,从而使其对称正定以解决不适定问题. 但由于干扰项的加入,增加了额外的信息,同时噪声抖动技术虽然可在一定程度上改善 ESN 的性能,但是也会引入对原序列本身无关的信息. 所以本文采用 EEMD 分解得到的风电功率训练样本的高频分量 IMF1,即首先分解出来的信号序列作为 ESN 的干扰项,起到正则化作用,以克服原 ESN 的不适定问题,模型简记为 EEMD-ESN.

即状态方程变为

$$\begin{aligned} \mathbf{x}(k+1) = & f[\mathbf{W}\mathbf{x}(k) + \mathbf{W}_{in}\mathbf{u}(k) \\ & + \mathbf{W}_{fb}\mathbf{y}(k) + k_d \cdot e(k)], \end{aligned} \quad (13)$$

式中  $k_d$  为加入的高频分量幅值系数,  $e$  即为高频分量 IMF1, 其维数等于储备池状态维数. 由于训练样本的高频分量本身就包含有用的信息,代替原随机噪声能更好的提高序列的预测精度,同时有效的克服了原 ESN 的病态矩阵问题. EEMD-ESN 同原 ESN 一样也由储备池状态维数  $N$ 、输出反馈系数  $f_b$ 、内部连接矩阵的稀疏度  $c$  和谱半径  $r$  等参数构成,并且增加了高频分量幅值系数  $k_d$ . 除了  $k_d$  其它参数的选取与原 ESN 类似,本文均通过交叉验证的方法求得.

### 5 基于 EEMD-近似熵的风电预测模型

对 2006 年 7 月份美国堪萨斯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,风电场的额定装机容量是 254 MW. 样本共采集 19 天的数据,其中每天 24 h,每 10 min 采集一次数据,为了方便计算,将风电功率转化为标么值,风电功率时间序列如图 1 所示.

从图 1 可以看出,风电功率时间序列具有明显的非线性和非平稳性,为了更好的对风电功率混沌时间序列进行局部分析以获得更高的预测精度,利用 EEMD-近似熵对风电功率时间序列进行分析. 具体步骤是首先利用 EEMD 对其进行分解,得到

了一系列不同尺度的 IMF 信号,分解后的 IMF 分量和剩余分量 res 如附录图 A1 所示. 从图 A1 可以看出,由于风电波动的非平稳性非常明显,造成 EEMD 分解后的 IMF 分量很多,如果直接利用 ESN 对每一个分量分别进行建模预测,增加了计算的规模. 为了更有效的对风电进行预测,采用近似熵理论对每一个 IMF 信号进行复杂度评估,各 IMF 的近似熵如图 2 所示.

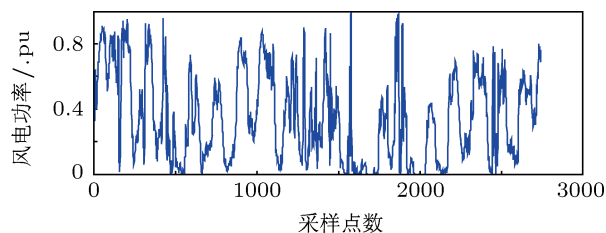


图 1 某一风电场风电功率时间序列

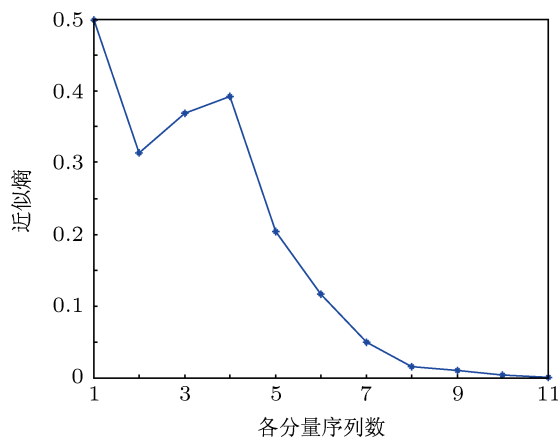


图 2 各 IMF 分量的近似熵

从图 2 可以看出,各 IMF 分量的近似熵值随着 IMF 分量频率的降低整体上呈递减趋势,说明从高频到低频分量复杂度是减小的,验证了近似熵的有效性;而且部分相邻的 IMF 的熵值相差不大,为了减少建模的计算规模,对相邻的熵值相差不大的 IMF 进行合并叠加,具体合并结果如表 1 所示,处理后的新的子序列如附录图 A2 所示.

表 1 各 IMF 分量合并为新子序列的结果

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

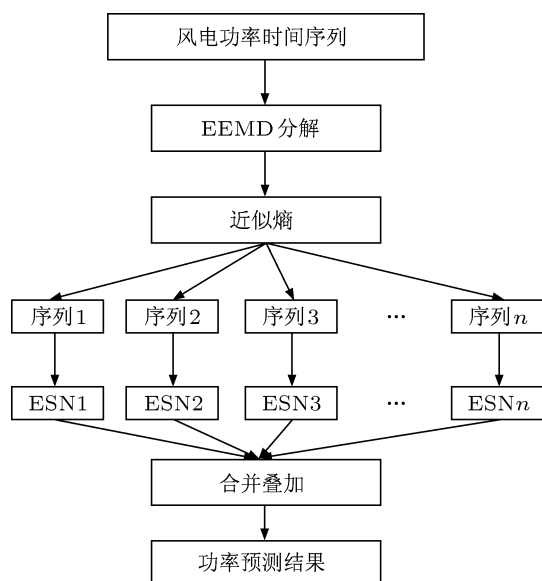


图3 近似熵-ESN组合预测流程

将得到的新的各子序列分别建立 ESN 和 EEMD-ESN 的预测模型进行预测, 再将各子序列的预测结果进行组合叠加并作反归一化处理即可得到风电功率预测值, 模型的具体预测流程如图 3 所示. 建立模型的过程中需要事先对 ESN 的各个模型进行交叉验证以确定各模型的谱半径、稀疏度、储备池维数以及储备池的反馈系数等参数, 由于 ESN 用于时间序列预测时, 几乎不需要反馈, 所

以可以设定很小的数值甚至可以设定为零, 所以主要需要确定三个参数, 其中最重要的是谱半径和储备池维数. 本文以 7 月 19 日的风电功率各子序列的前 24 点作为预测对象, 即前 4 个小时的风电功率子序列. 处理过程是每次进行一步预报, 再将得到一步预测值送入原子序列并重复一步 ESN 预测, 如此往复即可得到 24 点的多步预报. 以一次 24 步预报的均方根误差和训练时间作为评价指标, 利用交叉验证来选取合适的模型参数, 具体结果见附录表 B1.

## 6 算例分析

为验证本文所提出的风电混沌时间序列短期组合预测模型的有效性, 以美国堪萨斯州某一大型风电场在 2006 年 7 月 1 日到 19 日的实际采集的风电功率数据为算例, 首先将风电原始数据转化为标么值, 并进行  $[-1, 1]$  的归一化处理, 前 18 日的数据进行训练, 第 19 日的数据进行测试验证. 在利用 EEMD-近似熵理论得到风电功率的子序列的基础上, 同时采用目前风电功率预测领域应用较多的支持向量机组合预测模型与本文模型进行对比分析. 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.

表 2 各子序列的预测模型训练时间和测试误差

子序列号	LSSVM 预测		ESN 预测		EEMD-ESN 预测	
	训练时间/s	均方误差	训练时间/s	均方误差	训练时间/s	均方误差
1	73.3594	0.01311	0.1716	0.0173084	0.4368	0.01661
2	314.516	0.05012	0.2808	0.0626769	0.24960	0.06177
3	73.03125	0.007381	0.48360	0.012756	0.234002	0.01236
4	117.859	0.06754	0.1404	0.002052	0.140401	0.001449
5	12.3906	0.001445	0.39001	0.0012115	0.608404	0.001158
6	6.1875	$6.304 \times 10^{-4}$	0.42120	$1.9898 \times 10^{-5}$	0.234002	$1.911 \times 10^{-5}$

支持向量机模型的预测流程与前述图 3 的 ESN 预测流程相似, 也是在 EEMD-近似熵处理原风电功率时间序列的基础上对各个新的子序列分别建立预测模型. SVM 中主要考虑标准 LSSVM 模型, 其也需要进行模型参数的选择, 同样也采用交

叉验证法进行确定. 由于文献 [8] 已经论证了风电功率时间序列具有较明显的混沌特性, 所以借用混沌相空间理论确定风电各子序列的输入维数. 首先利用互信息法确定各子序列的延迟时间, 在此基础上利用 Cao 氏法确定各子序列的最大嵌入维

数, 而得到的最大嵌入维数即为各子序列的输入维数, 只要其足够大即可以刻画风电功率相空间的奇异吸引子, 有利于表达有用信息. 互信息法和Cao氏法分别参见文献 [32], 最终选取的各子序列的输入维数和模型参数具体见附录表 B1 所示. 而 EEMD-ESN 模型的幅值系数  $k_d$  取为 0.001.

本文采用提前 2 h 的风电功率预测, 即提前 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, 具体方法是进行完一次 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预测模型进行训练, 再继续下一轮的 12 步预测, 如此反复即可得到 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均方根误差以及建立初始模型的训练时间如表 2 所示, 表中 ESN 及 EEMD-ESN 预测值为 10 次计算相近结果的平均值. 从表 2 可以看出, LSSVM 模型在各序列的一次建立模型的训练时间上要远远大于 ESN 模型, 同时 EEMD-ESN 在同样的参数条件下与 ESN 模型训练时间大致相等但误差更小. 利用 ESN, EEMD-ESN 以及 LSSVM 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 4 所示.

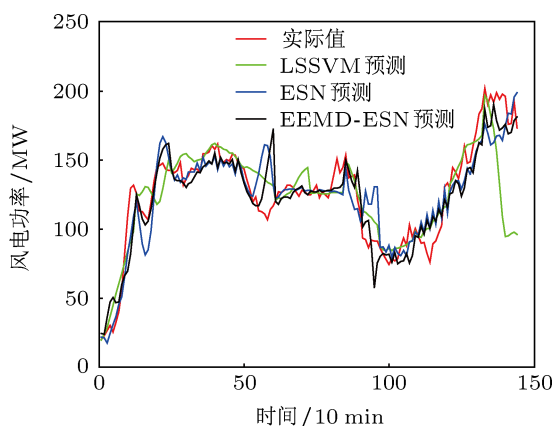


图 4 风电功率实际值以及各模型的提前 2 h 预测值

为了定量评价各模型的预测结果, 利用目前应用在风电预测比较普遍的归一化绝对平均误差  $E_{nmae}$ , 归一化均方根误差  $E_{nrmse}$  和最大绝对误差  $E_{mae}$  为依据来评价各模型的预测性能 [6], 表 3 列出了三种预测模型的各种误差性能指标.

从图 4 和表 3 可以看出, 在利用 EEMD-近似熵对风电功率时间序列处理的基础上, EEMD-ESN

模型作提前 2 h 的风电功率平均预测与实际值最接近, 其归一化绝对平均误差、归一化均方根误差以及最大绝对误差分别为 4.1722%, 5.7048%和 21.8794%, 均比 LSSVM 预测模型预测误差要小, 说明 EEMD-ESN 模型的预测值能紧跟风电功率的变化趋势, 特别在 19 日的最后阶段即 120 点以后能与风电功率进行很好的吻合, 而 LSSVM 虽然也能跟踪风电功率的变化趋势, 但是效果不及 ESN 和 EEMD-ESN 模型, 特别在 IMF 分量的低频部分预测误差较大, 以至在预测的 140 点以后出现较大的误差. 虽然在预测的第 55 点附近 ESN 和 EEMD-ESN 与实际值偏离较大, 但在其余时刻均能很好的与风电实际值进行较好的拟合, 而且对比表 2 和表 3 可知, 本文所提出的 EEMD-ESN 模型在各子序列的训练时间较小, 比 LSSVM 模型一次建模训练时间要小三个数量级, 说明本模型不但有更好的训练速度, 而且有更高的预测精度, 验证了利用风电功率本身的 EEMD 分解的高频分量作为 ESN 的正则项具有更好的预测精度, 为风电功率的预测提供一条新的有效途径.

表 3 各预测模型的预测误差比较

模型名称	$E_{nmae}$	$E_{nrmse}$	$E_{mae}$
LSSVM 预测	0.055057094	0.08999179	0.398437823
ESN 预测	0.048051245	0.066039567	0.211516036
EEMD-ESN 预测	0.041721804	0.057047592	0.218794304

## 7 结论

本文提出了一种 EEMD-近似熵和回声状态网络的风电功率混沌时间序列短期组合预测方法, 对比目前在风电短期预测中采用的 LSSVM 组合预测模型, 其具有更高的预测精度以及更短的训练时间. 首先针对标准 EMD 分解的端点效应问题会引起风电功率序列分解的不准确, 采用 EEMD 对原始风电混沌时间序列进行多尺度分解; 由于风电 EEMD 分解 IMF 分量较多, 利用近似熵对各 IMF 分量进行复杂度分析, 以熵值为指标对各 IMF 分量进行重新组合叠加产生了一组新的复杂度差异明显的子序列, 降低了预测计算规模; 其次利用交叉验证和混沌相空间理论对各预测模型的参数进行选择, 有利于提高预测精度; 最后以 ESN、经过高频分量正

则化改进 EEMD-ESN 预测模型与目前普遍采用的 LSSVM 预测模型进行对比研究发现, 本文所提出的预测模型不仅具有更高的预测精度而且具有更少的学习训练时间, 其为实现风电功率短期预测在

线工程应用提供了一条有效的途径.

### 附录 A

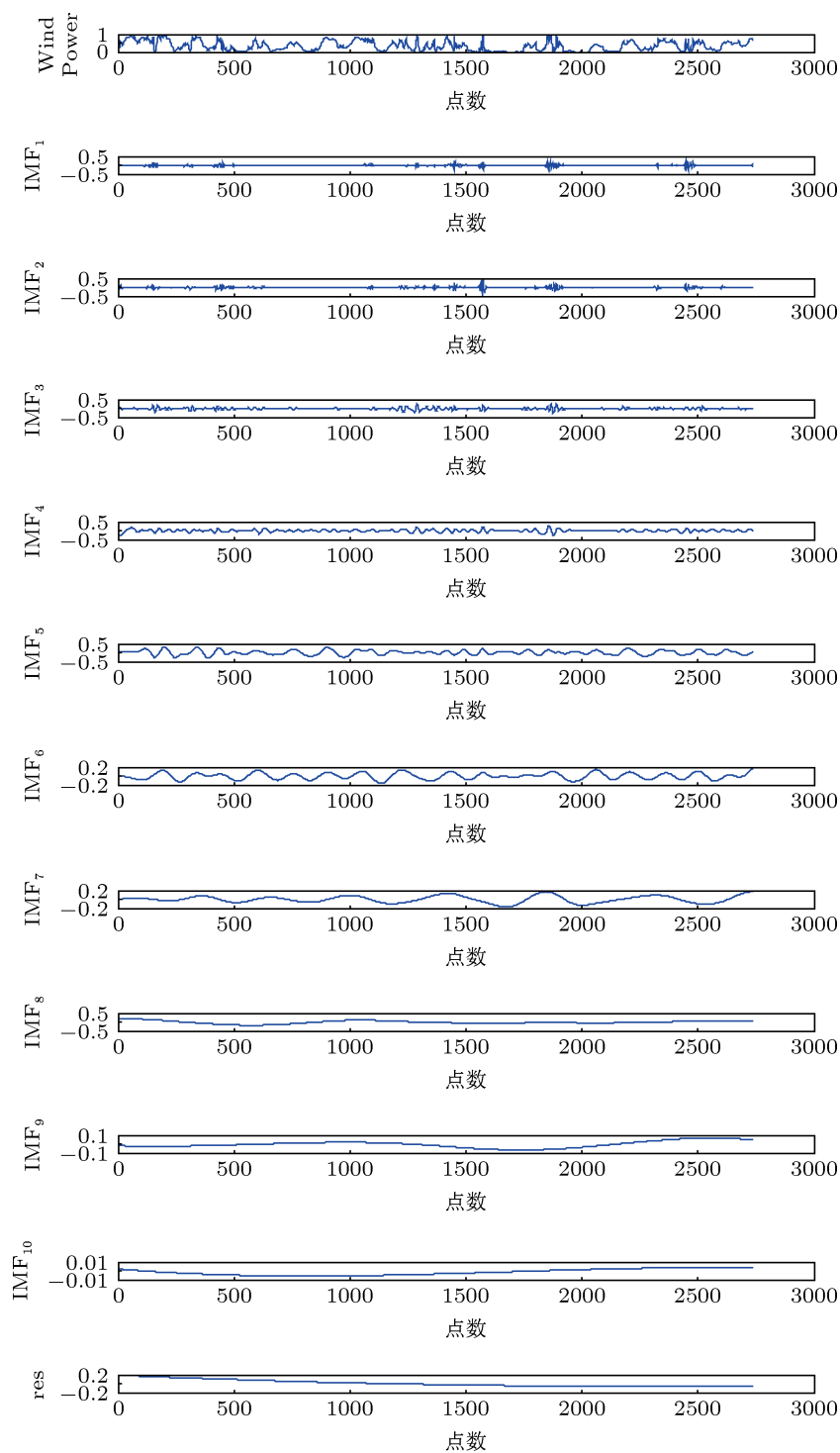


图 A1 风电功率 EEMD 分解



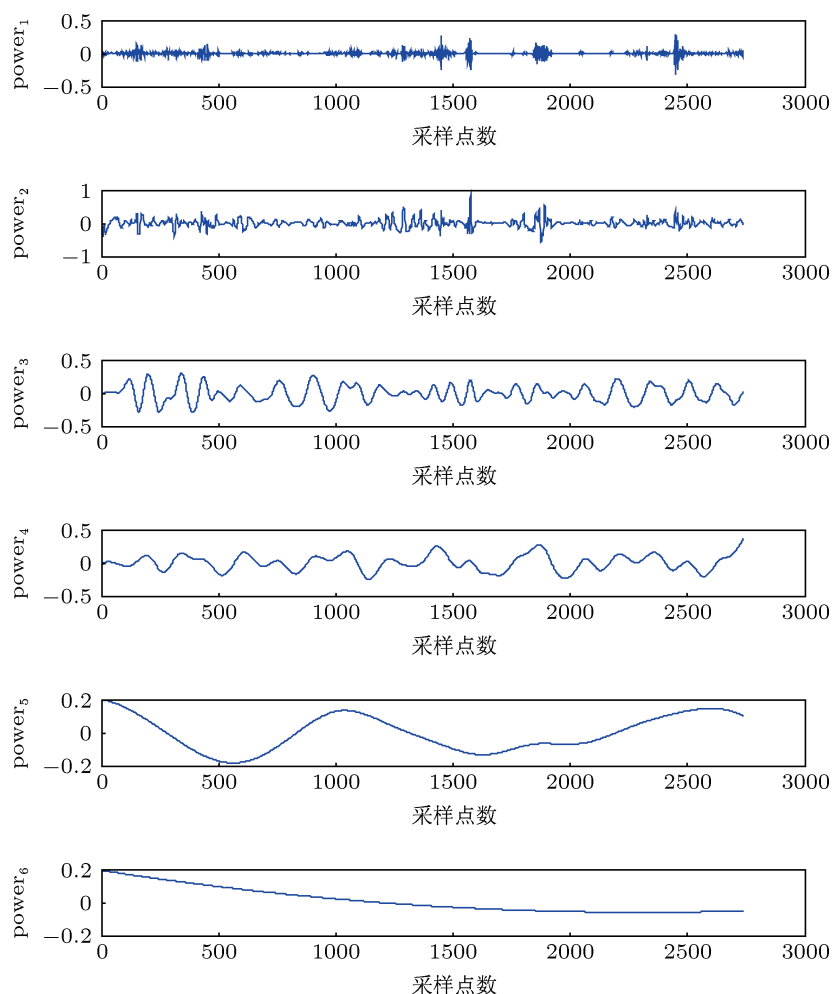


图 A2 经 EEMD-近似熵处理后的风电子序列

表 B1 各子序列预测模型的参数

序列数	LSSVM 预测		ESN			EEMD-ESN 预测				输入向量维数
	正则项系数	核参数	谱半径	反馈系数	储备池维数	幅值系数	谱半径	反馈系数	储备池维数	
1	30	10.6	0.93	0.12	23	0.001	0.93	0.12	23	12
2	300	2.8	0.98	0.02	62	0.001	0.98	0.02	62	9
3	120	5.5	0.98	0.01	20	0.001	0.98	0.01	20	9
4	90	0.6	0.98	0.02	26	0.001	0.98	0.02	26	7
5	70	8.8	0.96	0.06	50	0.001	0.96	0.06	50	5
6	30	12.2	0.9	0.14	44	0.001	0.9	0.14	44	4



- [1] Liu J Z, Liu Y, Zeng D L, Liu J W, Lü Y, Hu Y 2012 *Sci. China E* **55** 1140 [刘吉臻, 柳玉, 曾德良, 刘继伟, 吕游, 胡阳 2012 中国科学 (E 辑): 技术科学 **42** 437]
- [2] Ernst B, Oakleaf B, Ahlstrom M L, Lange M, Moehrlen C, Lange B, Focken U, Rohrig K 2007 *IEEE Transactions on Power & Energy Magazine* **5** 78
- [3] Chen S Y, Dai H Z, Bai X M, Zhou X X 2000 *Proceedings of the CSEE* **20** 26 (in Chinese) [陈树勇, 戴慧珠, 白晓民, 周孝信 2000 中国电机工程学报 **20** 26]
- [4] Gao S, Dong L, Gao Y, Liao X Z 2012 *Proceedings of the CSEE* **32** 32 (in Chinese) [高爽, 冬雷, 高阳, 廖晓钟 2012 中国电机工程学报 **32** 32]
- [5] Feng S L, Wang W S 2010 *Proceedings of the CSEE* **30** 1 (in Chinese) [冯双磊, 王伟胜 2010 中国电机工程学报 **30** 1]
- [6] Wang L J, Dong L, Liao X Z, Gao Y 2009 *Proceedings of the CSEE* **29** 30 (in Chinese) [王丽婕, 冬雷, 廖晓钟, 高阳 2009 中国电机工程学报 **29** 30]
- [7] Guo C X, Wang Y, Shen Y, Wang M, Cao Y J 2012 *Proceeding of the CSEE* **32** 24 (in Chinese) [郭创新, 王扬, 沈勇, 王媚, 曹一家 2012 中国电机工程学报 **32** 24]
- [8] Zhang X Q, Liang J 2012 *Acta Phys. Sin.* **61** 190703 (in Chinese) [张学清, 梁军 2012 物理学报 **61** 190703]
- [9] Gao Z K, Jin N D 2011 *Physica A* **390** 3541
- [10] Zhang J, Small M 2006 *Phys. Rev. Lett.* **96** 238701
- [11] Gao Z K, Jin N D 2012 *Nonlinear Analysis: Real World Applications* **13** 947
- [12] Zhang J, Zhang K, Feng J F, Small M 2010 *PLoS Comput. Biol.* **6** e1001033
- [13] Ye L, Liu P 2011 *Proceeding of the CSEE* **31** 102 (in Chinese) [叶林, 刘鹏 2011 中国电机工程学报 **31** 102]
- [14] Cui W Z, Zhu C C, Bao W X, Liu J H 2004 *Acta Phys. Sin.* **53** 3303 (in Chinese) [崔万照, 朱长纯, 保文星, 刘君华 2004 物理学报 **53** 3303]
- [15] Deng L F, Liu H, Liu D 2006 *Chin. Phys.* **15** 1196
- [16] Suykens J A K, Vandewalle J 1999 *Neural Processing Letters* **9** 293
- [17] Wang X D, Ye M Y 2004 *Chin. Phys.* **13** 454
- [18] Sun J C, Zhang T Y, Liu F 2004 *Chin. Phys.* **13** 2045
- [19] Wang H D, Zhang H Y, Ye M Y 2005 *Acta Phys. Sin.* **54** 2568 (in Chinese) [汪晓东, 张浩然, 叶美盈 2005 物理学报 **54** 2568]
- [20] Shi Z W 2008 *Ph. D. Dissertation* (Dalian: Dalian University of Technology) (in Chinese) [史志伟 2008 博士学位论文 (大连: 大连理工大学)]
- [21] Jaeger H, Haas H 2004 *Science* **304** 5667
- [22] Song T, Li H 2012 *Acta Phys. Sin.* **61** 080506 (in Chinese) [宋彤, 李菡 2012 物理学报 **61** 080506]
- [23] Hu A J, Ma W L, Tang G J 2012 *Proceeding of the CSEE* **32** 106 (in Chinese) [胡爱军, 马万里, 唐贵基 2012 中国电机工程学报 **32** 106]
- [24] Gao X Q, Wan S Q, He W P, Feng G L, Li J P, Dong W J 2005 *Chin. Phys.* **14** 628
- [25] Cai Y P, Li A H, Shi L S, He Y P, Zhao J R 2011 *Journal of Vibration and Shock* **30** 46 (in Chinese) [蔡艳平, 李艾华, 石林锁, 何艳萍, 赵静茹 2011 振动与冲击 **30** 46]
- [26] Wu Z H, Huang N E 2009 *Advances in Adaptive Data Analysis* **1** 1
- [27] Pincuss M 1991 *Proceeding of the National Academy Sciences* **88** 2297
- [28] Li R, Lu F Y, Xu H G, Zhang L Y 2010 *Proceeding of the CSEE* **30** 51 (in Chinese) [栗然, 陆凤怡, 徐宏锐, 张烈勇 2010 中国电机工程学报 **30** 51]
- [29] Han M, Mu D Y 2011 *Control and Decision* **26** 1469 (in Chinese) [韩敏, 穆大芸 2011 控制与决策 **26** 1469]
- [30] Jaeger H 2007 *Scholarpedia* **2** 2330
- [31] Shi Z W, Han M 2007 *Control and Decision* **22** 258 (in Chinese) [史志伟, 韩敏 2007 控制与决策 **22** 258]
- [32] Han M 2007 *Predict Theory and Method of Chaotic Times Series* Vol.1 (Beijing: China WaterPower Press) p168 (in Chinese) [韩敏 2007 混沌时间序列预测理论与方法 (第1卷) 第168页]

# Chaotic time series prediction model of wind power based on ensemble empirical mode decomposition-approximate entropy and reservoir\*

Zhang Xue-Qing<sup>1)†</sup> Liang Jun<sup>1)2)</sup>

1) (School of Electrical Engineering, Shandong University, Jinan 250061, China)

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

(Received 20 August 2012; revised manuscript received 30 October 2012)

## Abstract

According to the chaotic feature of wind power time series, a combined short-term wind power forecasting approach based on ensemble empirical mode decomposition (EEMD)-approximate entropy and echo state network (ESN) is proposed. Firstly, in order to reduce the calculation scale of partial analysis for wind power and improve the wind power prediction accuracy, the wind power time series is decomposed into a series of wind power subsequences with obvious differences in complex degree by using EEMD-approximate entropy. Then, the forecasting model of each subsequence is created with least squares support vector machine (LSSVM), ESN and EEMD-ESN improved with the regularized high frequency parts. Finally, the simulation is performed by using the real data collected from a certain wind farm, the results show that the EEMD-ESN model is better in the training speed and forecasting accuracy, than those obtained from the least square support vector machine (LSSVM) model, which provides a new useful reference for the short-term forecasting of wind power in online engineering application.

**Keywords:** chaotic time series, wind power prediction, ensemble empirical mode decomposition, approximate entropy

**PACS:** 05.45.Tp, 88.50.Mp

**DOI:** 10.7498/aps.62.050505

---

\* Project supported by the National Natural Science Foundation of China (Grant No. 51177091), and the Natural Science Foundation of Shandong Province (Grant No. ZR2010EM055).

† Corresponding author. E-mail: zxq\_wy@163.com