基于模态分解和支持向量机的风电功率组合预测

针对风电功率序列的不确定性和随机性特征，提出基于聚类经验模态分解（EEMD）和支持向量回归机（SVR）的风电功率预测模型。同时，为克服支持向量回归机依赖人为经验选择学习参数的弊端，采用纵横交叉算法（CSO）优化支持向量回归机学习参数。首先，利用聚类经验模态分解将原始风电功率序列分解为一系列复杂度差异明显的子序列。然后，分别对每子序列单独建立CSO－SVR预测模型。最后，叠加各子序列的预测值得到实际预测结果。实例研究表明，所提模型能获得优良的风电功率预测结果。

陈庆鸿　肖建华　陈冬洋／广东电网公司揭阳市供电局
董　朕　林艺城　孟安波／广东工业大学自动化学院

陈庆鸿／高级工程师

关键词／Keywords
聚类经验模态分解·纵横交叉算法·支持向量回归机·风电功率预测·

0　引言

近年来，风电由于其应用技术日益成熟以及能源的清洁性，受到世界各国的青睐，发展规模迅速扩大，在电网规划中的比重也逐年上升。但是风电出力固有的波动性和间歇性给电网带来的冲击严重影响电能质量和供电可靠性，成为制约大规模风电并网的最大瓶颈。对风电场输出功率进行预测，可为电网安排发电计划、停机检修提供可靠参考，有利于保证系统的安全可靠和经济运行[1-3]。故提高风电预测的准确度对大规模风电的逐步接入具有重要意义。

目前，国内外建立的风电功率预测方法包括持续法、卡尔曼滤波法、模糊逻辑法、空间相关性法、人工神经网络法、支持向量机法和物理预测法等[4-6]。其中持续法认为连续两点的功率值相差不大，某一点的风电预测值即为上一点的实际功率值，故该方法难以避免出现预测误差大的情况。卡尔曼滤波法以风功率作为状态变量建立模型，基于卡尔曼滤波算法实现功率预测。但该方法在统计特性未知时，存在很大局限性[7]。模糊逻辑法以模糊规则库为基础，使用线性模型搜寻非线性的风电功率的变化规律来进行预测，但模糊规则很难完整诠释风电的随机性规律，风电功率预测中单纯的模糊方法效果欠佳[8]。人工神经网络法模拟人脑结构和功能，通过大量数据训练确定网络参数，可用于短期风功率预测。但其网络结构很难科学确定，且学习速度慢、容易陷入局部最优，故预测准确度一般难以满足要求[9]。支持向量机法依据结构风险最小化的思想，模型建立方式大致与神经网络相同，都是以大量的样本数据训练来建立相应的预测模型，该模型虽然泛化能力强、全局寻优效果好，但人为选择其学习参数依然是该方法存在的致命缺点[10-11]。物理方法主要以气象数据和功率转化曲线为预测基础，通过分析风机周围的物理信息，计算得出发电机组轮毂高度的风速、风向等信息，带入功率转化曲线中得到实际输出功率，但目前很难获取准确的风速、风向等信息，模型建立相对困难[12]。

上述预测方法均为传统的单一预测，受风电功率的随机性以及单一预测方法本身的局限性的影响，单一的预测方法预测误差一般为25%～40%，很难满足预测需求[13]。文献[14-15]分别建立基

广东省科技计划项目(2016A010104016)；广东电网公司科技项目(GDKJQQ20152066)。

于小波包分解和经验模态分解的组合预测模型，即先将原始风电序列分解后再建立相应的预测模型，结果表明预测准确度得到有效提高。风电功率的组合预测已是大势所趋，通过各模型的优势互补来建立组合预测模型，充分挖掘风电时间序列的有效信息，提高预测准确度。

针对单一模型预测准确度低的问题，提出了一种基于聚类经验模态分解（EEMD）和支持向量回归机（SVR）的风电功率组合预测模型。同时为降低通过人为经验选择SVR学习参数给预测造成的误差，采用纵横交叉算法（CSO）优化SVR的学习参数。实例研究表明，所提出的组合预测模型的预测结果准确度较高。

1 模型原理

1.1 聚类经验模态分解

风电时间序列首先使用基于经验模态分解（Empirical Mode Decomposition，EMD）的聚类模态分解（Ensemble Empirical Mode Decomposition，EEMD）进行分解。EMD已被广泛应用于时间序列分解。通过筛选过程将原始时间序列数据$X(t)$分解成多个固有模态分量（Intrinsic Mode Function，IMF）。

$$X(t)=\sum_{i=1}^{n}c_i(t)+r_n(t) \tag{1}$$

式中，$c_i(t)$为各IMF分量；$r_n(t)$为剩余分量；n为IMF分量总数。

利用EMD方法，能将时间序列自适应分解成由高频到低频的一系列分量。然而，EMD的一个主要缺点就是模态混叠问题[16]。聚类经验模态分解则通过添加白噪声有效解决了这个问题。作为EMD的改进版，EEMD向原始数据中加入有限的白噪声序列，然后通过集成平均来获得准确的IMF分量，其可以自动消除模式混叠问题。分解步骤如下。

1）向原始序列中加入小幅度的白噪声来平衡序列，形成一个整体序列。

2）使用EMD将该整体序列分解，得到相应的IMF分量和一个剩余分量。

3）不断执行步骤1）共n次，但每次加入的小幅度白噪声序列要各不相同。

4）将n次分解得到的全部分量集成平均后作为最终的结果。

正如文献[16]所述，添加的白噪声可以由以下规则控制，即

$$\varepsilon_e=\varepsilon/\sqrt{e} \tag{2}$$

式中，e为集合总数；ε为加入白噪声的幅度；ε_e为误差的标准偏差。

1.2 支持向量回归机

采用EEMD对原始风功率序列进行多尺度平稳化分解，同时，结合支持向量回归机（Support Vector Regression Machine，SVR）模型对各IMF分量分别进行预测。

支持向量机（Support Vector Machine，SVM）理论基础源于统计学的VC维理论，最初用来解决二类线性样本的分类问题，将分类问题推广到解决回归拟合问题就形成了SVR[17]。SVR使用非线性核函数，将低维空间的非线性数据集映射到高维空间，转化为线性问题进行回归估计。

SVR依据结构风险最小化的思想构建损失函数，并使以下目标函数最小化来确定回归函数，即

$$\begin{cases}\min\left[\dfrac{1}{2}w^Tw+\dfrac{C}{n}\sum\limits_{i=1}^{n}(\xi_i+\xi_i^*)\right]\\ f(x_i)-y_i\leqslant\varepsilon+\xi_i\\ \text{s.t.}\quad y_i-f(x_i)\leqslant\varepsilon+\xi_i\\ \xi_i,\xi_i^*\geqslant 0,\quad i=1,2,\cdots,n\end{cases} \tag{3}$$

式中，C为惩罚系数；ξ_i、ξ_i^*为松弛因子；ε为不敏感损失系数。

引入拉格朗日乘子，可将式（3）转化为以下对偶问题的求解，即

$$\begin{cases}\min\left[\dfrac{1}{2}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(a_i^*-a_i)(a_j^*-a_j)K(x_i,x_j)\right]+\\ \qquad\varepsilon\sum\limits_{i=1}^{n}(a_i^*+a_i)-\sum\limits_{i=1}^{n}y_i(a_i^*-a_i)\\ \text{s.t.}\quad\sum\limits_{i=1}^{n}(a_i-a_i^*)=0\text{且}0\leqslant a_i,a_i^*\leqslant\dfrac{C}{n}\end{cases} \tag{4}$$

式中，$K(x_i, x_j)$为核函数，满足Mercer 条件；a_i、a_i^*为拉格朗日乘子。

求解式（4）可得到最优回归函数为

$$f(x)=\sum_{i=1}^{n}(a_i-a_i^*)K(x_i,x)+b \tag{5}$$

常用核函数有如下形式。

1）多项式核函数，$K(x, x_i)=x_i^{\mathrm{T}}x$。

2）径向基（radial basis function，RBF）核函数，$K(x, x_i)=\exp(-\|x-x_i\|^2/2\sigma^2)$。

3）多层感知器核函数，$K=(x, x_i)=\tanh(vx_i^{\mathrm{T}}+y)$。由于RBF核函数局部插值能力强，因此其更适合用于短期风电功率预测研究。

因此，使用SVR预测时需要选择的学习参数包括形式3）中的C和RBF核函数的核参数，本文使用纵横交叉算法来实现这两个参数的优化选择。

1.3 纵横交叉算法

纵横交叉算法（CSO）包括横向交叉和纵向交叉两个算子，每迭代一次将交替执行两种交叉算子，纵横两种交叉机制保证了CSO具有较强的全局收敛能力和较快收敛速度。

（1）横向交叉操作

执行交叉算子前先对种群中所有粒子进行两两不重复随机组合，然后对组合粒子所有维执行横向交叉来更新粒子。若父代粒子$X(i)$、$X(j)$的第d维进行横向交叉，则粒子更新方式如下，即

$$MS_{\mathrm{hc}}(i,d)=r_1X(i,d)+(1-r_1)X(j,d)+c_1[X(i,d)-X(j,d)] \tag{6}$$

$$MS_{\mathrm{hc}}(j,d)=r_2X(j,d)+(1-r_2)X(i,d)+c_2[X(j,d)-X(i,d)]$$
$$i,j\in N(1,M),\quad d\in N(1,D) \tag{7}$$

式中，r_1、r_2为[0，1]之间的随机数；c_1、c_2为[－1，1]之间的随机数；M为粒子规模；D为变量维数；$X(i，d)$、$X(j，d)$为父代粒子$X(i)$、$X(j)$的第d维；$MS_{\mathrm{hc}}(i，d)$、$MS_{\mathrm{hc}}(j，d)$为$X(i，d)$、$X(j，d)$在横向交叉后各自的第d维子代。

（2）纵向交叉操作

粒子的每一维代表一个变量，不同变量的取值范围会有差异，故交叉前须对每一维进行归一化处理，然后依交叉概率P_v对全部粒子不同维之间执行纵向交叉。为了使陷入局部最优的某一维从中跳出，但同时又不破坏其他正常维，故每次纵向交叉只须产生一个子代即可。若父代粒子$X(i)$的第d_1维和第d_2维执行纵向交叉，则粒子更新方式如式（8）所示。

$$MS_{\mathrm{vc}}(i,d_1)=rX(i,d_1)+(1-r)X(i,d_2)$$
$$d_1,d_2\in N(1,D),\quad i\in N(1,M),\quad r\in[0,1] \tag{8}$$

式中，$MS_{\mathrm{vc}}(i，d_1)$为粒子$X(i)$的第d_1维和第d_2维在纵向交叉后的子代。

由式（8）可知，纵向交叉通过粒子维度层面的交叉来产生中庸解，其能使陷入局部最优的某一维从中摆脱出来，以此来保持和增加种群多样性。由于并不是所有维都陷入停滞或是局部最优，故为纵向交叉算子设置一个交叉概率P_v，借此来控制种群中参与交叉的维的数量，有效保证算法的收敛速度。

2 基于EEMD－CSO－SVR的短期风电功率组合预测模型

风电功率时间序列的随机性和非线性特征使得单一的预测方法难以取得理想的预测效果，故为保证预测准确度须建立组合预测模型。鉴于EEMD技术在非平稳数据处理中的突出优势，以及CSO算法在寻优求解中全局搜索能力强的特点，本文提出了一种基于EEMD－CSO－SVR的短期风电功率组合预测模型。

2.1 适应度函数的选取

用均方误差公式（Mean Squared Error，MSE）（适应度函数）衡量每个粒子的适应值，计算公式为

$$fit=\frac{1}{n}\sum_{j=1}^{n}(y_j-\hat{y}_j)^2 \tag{9}$$

式中，y_j和$\hat{y}_j$分别为实际值和训练输出值；fit为适应度函数，fit越小，则选择的学习参数预测效果越好。

2.2 测误差评价函数

预测误差评价函数采用平均绝对百分比误差（Mean Absolute Percentage Error，MAPE）和

均方根误差（Root of the Mean Squared Error，RMSE）。

2.3 CSO优化SVR参数步骤

SVR的学习参数对模型预测结果影响很大，良好的学习参数将有效提高预测的准确度。SVR学习参数优化步骤如下。

1）参数和种群初始化。随机生成初始种群X；对CSO算法以下参数初始化，包括纵向交叉概率P_v、最大迭代次数maxgen和种群大小M。

2）将种群中全部粒子转换成SVR的学习参数C和σ，然后对样本数据进行预测，用式（9）即适应度函数对每个粒子进行适应度评价。

3）进入横向交叉算子，交叉后的折中解放入矩阵MS_{hc}中，并计算全部解的适应度值，与其父代种群DS_{vc}（第一代为X）进行适应度值比较，选取fit值小的粒子存在DS_{hc}中。

4）进入纵向交叉算子，交叉后的解放入矩阵MS_{hc}中，并计算MS_{hc}中全部解的适应度值，与其父代种群DS_{hc}的适应度值比较，fit值小的粒子替换掉fit值大的粒子存在DS_{vc}中。

5）判断终止条件是否满足。如果迭代次数大于设定的最大值时，则迭代终止并将DS_{vc}中适应值最好的一组解转换为支持向量回归机对应的C和σ。否则，转到步骤3）进行新一轮迭代。

建模流程图如图1所示，具体步骤如下。

1）将平均化处理后的原始风电功率数据利用聚类模态分解进行分解，得到各分量（IMF_1～IMF_n）和余量$r_N(t)$。

2）对各IMF分量和$r_N(t)$分别建立支持向量回归机预测模型。

3）采用纵横交叉算法优化选择支持向量回归机的最佳学习参数，然后对所有分量进行预测，得到各子序列的预测值。

4）汇总全部分量的预测值后求和，得到实际预测值。

5）采用预测评价函数进行误差分析。

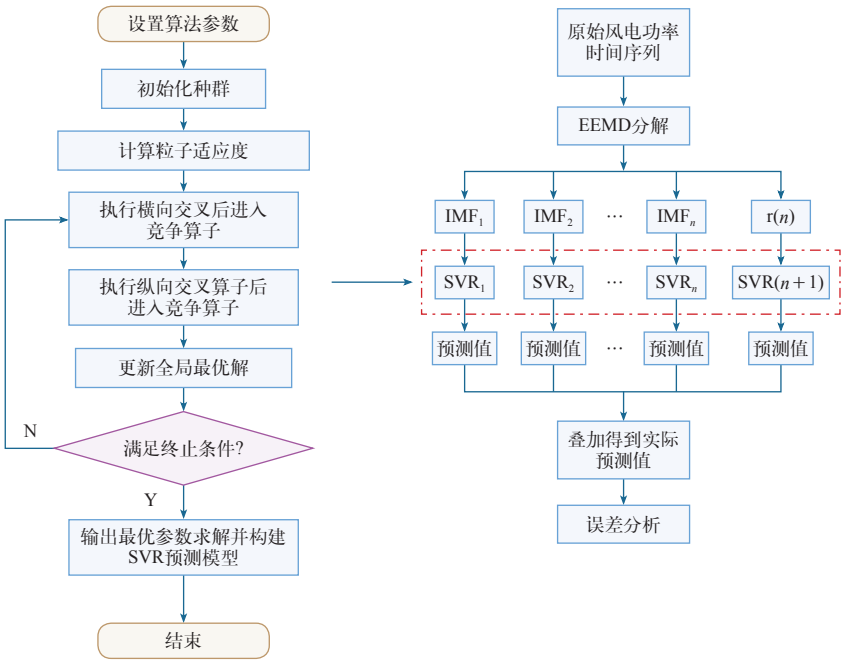

图1 EEMD－CSO－SVR组合预测流程图

3 算例及结果分析

以我国某风电场2012年的实测风电功率数据为实验样本。平均化处理后的数据为每小时1个点，共700 h的数据。训练样本为处理后前600 h的数据，测试样本则为后100 h的数据，输入维数为6维。平均化处理后的700 h风电数据时间序列如图2所示；将这700 h的原始数据采用EEMD分解后的各分量如图3所示。从图3中可以看出，与原始风电功率时间序列相比，分解后得到的各IMF分量和余量其非线性程度明显降低。对相对平稳的各IMF分量和余量分别建立SVR预测模型，并用CSO优化的SVR的学习参数进行预测。

CSO算法的纵向交叉概率设为0.8，种群大小$M=60$，最大迭代次数maxgen＝300。PSO的学习因子取2，惯性权重$w_{max}=0.9$，$w_{min}=0.5$。EEMD中的参数e和ε分别设为20和0.15。分别采用SVR、EEMD－SVR、EEMD－PSO－SVR、

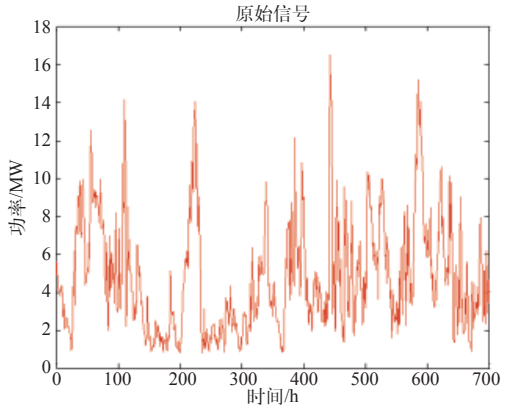

图2 原始风电功率时间序列

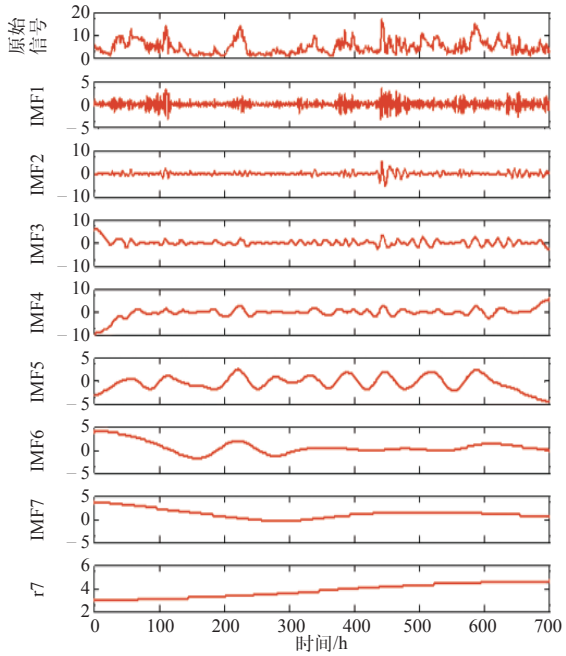

图3 风电功率时间序列EEMD分解结果

EEMD－GA－SVR以及EEMD－CSO－SVR 五种模型进行单步预测对比实验。不同模型的预测误差如表1所示。不同模型的功率预测曲线如图4、图5所示。

表1 不同风电功率预测方法的误差对比

方法	MAPE(%)	RMSE/MW
SVR	12.085	0.578 8
EEMD－SVR	6.647 6	0.329 1
EEMD－GA－SVR	4.614 9	0.215 5
EEMD－PSO－SVR	4.495 4	0.208 2
EEMD－CSO－SVR	3.294 3	0.170 8

1）对比SVR与EEMD－SVR，后者的预测准

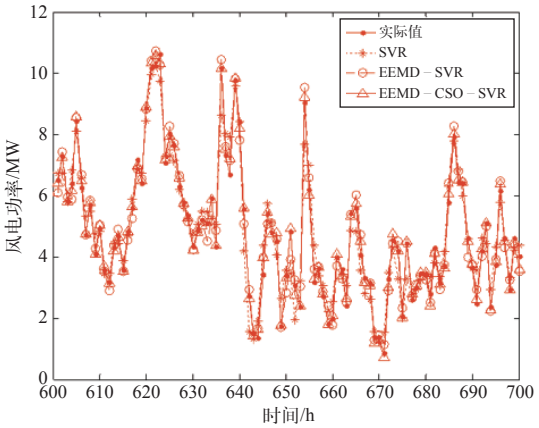

图4 三种模型的风电功率预测曲线

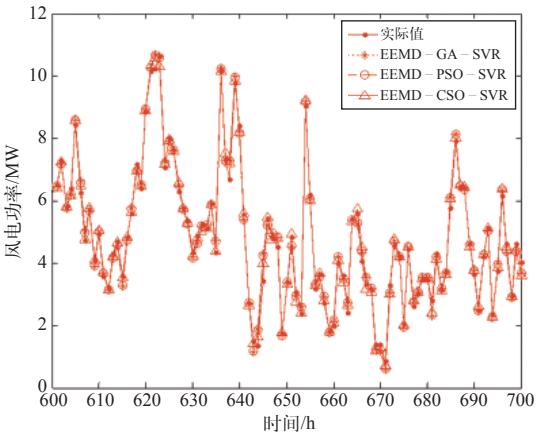

图5 不同算法模型的风电功率预测曲线

确度明显大幅度提高，说明EEMD分解有效地降低了风电功率的非线性和非平稳性，有效地提高了整体预测准确度。

2）分别对比EEMD－GA－SVR、EEMD－PSO－SVR、EEMD－CSO－SVR与EEMD－SVR，采用智能算法优化SVR的模型预测平均误差分别减小2.032 7%、3.353 3%，说明通过智能算法优化支持向量机的参数，有效提高了SVR的泛化能力，具有更高的预测准确度。

3）对比EEMD－GA-SVR、EEMD－PSO－SVR和EEMD－CSO－SVR，后者的预测准确度更高，平均误差分别减小了1.321 6%、1.201 1%，预测的RMSE仅为0.170 8 MW，在三种模型中也是最小，说明所提的纵横交叉算法算法具有更强大的全局搜索能力。

由图4可以看出，无论是EEMD－SVR模型还是EEMD－CSO－SVR模型都能很好地捕捉到风

电功率的变化规律，预测结果与SVR模型相比更接近实际值，充分表明了所提模型的有效性和高准确度。

从图5可以看出，三种方法的预测结果均能很好地拟合实际值，这主要是因为EEMD分解很好地降低了原始风电功率曲线的复杂度，有效挖掘了数据内部的变化规律，提高了模型的泛化能力和预测准确度，这种基于模态分解的组合预测方式更能满足预测准确度的需要。EEMD－CSO－SVR模型与其他两种模型相比，预测曲线更接近于真实值，尤其是在功率波动大的拐点，其准确性更加突出。这主要是因为GA和PSO在算法迭代后期容易陷入局部最优，无法优化得到最佳的支持向量回归机学习参数，CSO算法则具备全局优化能力强、鲁棒性好的独特优势，优化选择的学习参数更优良，提高了预测模型的预测准确度和泛化能力。

综合表1、图4和图5可以看出，单纯的SVR模型虽然能预测出功率的大体走势，但在风电功率突变的拐点处预测效果不佳，平均百分比误差为12.085%，已经很难满足预测准确度的需求；EEMD－SVR模型通过聚类模态分解挖掘到有效信息，提高了在风电功率突变处的预测准确度，但依靠人为经验选择SVR参数阻碍了模型预测准确度的进一步提高；EEMD－PSO－SVR、EEMD－GA－SVR和EEMD－CSO－SVR则采用智能算法优化选择SVR的学习参数，进一步提高了模型的预测准确度。其中，CSO以其全局搜索能力强、鲁棒性好的独特优势解决了常规智能算法容易陷入局部最优的问题，EEMD－CSO－SVR模型的平均百分比误差仅为3.294 3%，相对于其他模型表现出最高的预测准确度。

4 结束语

本文提出一种基于聚类经验模态分解（EEMD）和纵横交叉算法（CSO）优化支持向量回归机（SVR）学习参数的短期风电功率预测模型。经实例研究表明：

1）EEMD－SVR模型的MAPE和RMSE相对于单纯的SVR模型分别降低了5.437 4%和0.248 7 MW，说明EEMD有效降低了原始风电功率的复杂度，提高了预测准确度。

2）EEMD－GA－SVR、EEMD－PSO－SVR和EEMD－CSO－SVR模型与EEMD－SVR模型相比，MAPE分别降低了2.032 7%、2.152 2%和3.353 3%，说明采用智能算法优化SVR的学习参数，有效解决了SVR依赖人为经验选择学习参数的弊端，提高了预测准确度。

3）所提EEMD－CSO－SVR组合预测模型表现出最高的预测准确度，说明CSO算法相对于其他智能算法具有更强的全局优化能力，预测结果显示出其高准确度和有效性。

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)：5029-5040.

[2] 薛志英，周明，李庚银．大规模风电接入电力系统备用决策评述 [J]．电力系统保护与控制，2013，41 (4)：148-155.

[3] 余民，杨宸，蒋传文，等．风电并网后电力系统可靠性评估和备用优化研究 [J]．电力系统保护与控制，2012，40 (12)：100-104.

[4] 杨秀媛，肖洋，陈树勇．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.

[5] 谷兴凯，范高锋，戴慧珠．风电功率预测技术综述 [J]．电网技术，2007，31 (2)：335-338.

[6] 王丽婕，廖晓钟．风电场发电功率的建模和预测研究综述 [J]．电力系统保护与控制，2009，37 (13)：118-121.

[7] 范高峰，王伟胜，戴慧珠，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.

[8] Damousis I G，Alexiadis M C，John B．A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation [J]．IEEE Transactions on Energy Conversion，

2004，19 (2)：352-361.

[9] Cameron W P，Negnevitsky M. Very short-term wind forecasting for Tasmanian power generation [J]. IEEE Transactions on Power Systems，2006，21 (2)：965-972.

[10] Bludszuweit H，Dominguez-NavarroJ A，Llombart A. Statistical analysis of wind power forecast error [J]. IEEE Transactions on Power Systems，2008，23 (3)：983-991.

[11] James W T，Patrick E，Sharry M. Wind power density forecasting using ensemble predictions and times series models [J]. IEEE Transactions on Energy Conversion，2009，24 (3)：775-782.

[12] 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.

[13] Barthelmie R，Giebel G，Badger J. Short-term forecasting of wind speeds in the offshore environment [C]. Copenhagen Offshore Wind Conference，Copenhagen：University of Copenhagen，2005：24-35.

[14] 孟安波，卢海明，胡函武，等. 混合小波包与纵横交叉算法的风电预测神经网络模型 [J]. 太阳能学报，2015，36 (7)：1645-1651.

[15] 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.

[16] Zhang Jian，Yan Ruqiang，Gao Robert X，et al. Performance enhancement of ensemble empirical modedecomposition [J]. Mechanical Systems and Signal Processing，2010，24 (7)：2104-2123.

[17] Vapnik V. The nature of statistical learning theory [M]. New York：SpringerVerlag，1995.

（收稿日期：2016－10－20）EA

广告 | AD

Weidmüller

魏德米勒Energy Meter电力参数测量仪
—— 高效透明的生产电能监测

魏德米勒Energy Meter电力参数测量仪不仅能监测电能的消耗，它还能监测电能质量的基本参数和分析来自不同导体的电流的细微差别，使电力行业用户能够快速地了解生产设备中电能参数的变化情况，保证其稳定和高效率地使用。具有综合功能的Energy Meter电力参数测量仪使您可以将生产点的能源网络分块测量，尽可能精确地进行监测，配备有一体的液晶显示屏，重要的监测数据如电压、电流、功率和能耗都能一目了然。

魏德米勒电联接(上海)有限公司
电话: 86 21-22195008 网址: www.weidmueller.com.cn

扫一扫，关注魏德米勒官方微信

Let's connect.