

# 基于二次自适应支持向量机的 光伏输出功率预测

殷 豪, 陈云龙, 孟安波, 林艺城

(广东工业大学自动化学院, 广州 510006)

**摘 要:** 针对传统相似度函数在聚类过程中所存在的问题, 提出改进灰色-欧氏距离相似度函数, 并将历史样本日模糊聚类分为若干类。另外, 考虑到光伏输出数据的复杂性, 该文结合小波分解(WD)和集成经验模态分解(EEMD)的各自优势, 对光伏数据作双分解处理, 得到趋势分量与细节分量, 然后采用二次自适应支持向量机模型分别进行光伏功率预测。最后, 采用美国俄勒冈州某处光伏发电场的实测数据实验, 验证模型的实用性和可行性。

**关键词:** 光伏功率预测; 小波分解; 集成经验模态分解; 二次自适应支持向量机; 相似度

**中图分类号:** TM615

**文献标识码:** A

## 0 引 言

光伏预测技术对电网的稳定运行具有重大意义, 近几年来国内外学者对光伏功率预测的研究越来越深入。文献[1]综合考虑气象影响因素, 建立基于支持向量机理论的预测模型, 可该方法在历史样本日的相似度分析上存在不容小觑的缺陷。文献[2]考虑到影响光伏出力的主要因素, 提出一种基于灰色神经网络的组合模型, 具有潜在工程价值, 然而它忽视了对原始光伏数据的预处理。文献[3]提出一种基于集合经验模态分解和支持向量机(SVM)的EEMD-SVM组合模型预测方法, 在一定程度上降低了预测误差。

本文以具有相似气候特征的历史日为样本集, 对预测模型进行训练与测试, 故相似日聚类效果将很大程度上影响到预测精度。文献[4]依据日特征向量的欧氏距离选取相似日, 而欧氏距离仅能反映样本序列间的“值相似”, 在复杂天气情况下效果不佳。文献[5]则采用灰色关联度方法, 而灰色关联度仅表征研究对象在空间形状上的相似性, 即“型相似”。实际上, 这两种方法都不足以描述天气状况的相似性。

针对以上问题, 本文建立基于双分解(EEMD-

WD)的二次自适应支持向量机的短期光伏出力预测模型。另外, 提出一种改进灰色-欧氏距离的新型相似度函数, 并根据日气候特征向量对历史样本进行相似日模糊聚类, 从而构建训练样本和测试样本。

## 1 改进相似度的模糊聚类分析

### 1.1 相似度函数

聚类就是依据各指标的相似度, 可将数据信息分成多个类, 使属于同一类的数据尽可能相似, 而不同类的数据尽可能相异。设归一化后的基准数据序列  $i$  与第  $j$  个数据序列分别为  $x_0 = [x_0(1), x_0(2), \dots, x_0(n)]$ ,  $x_j = [x_j(1), x_j(2), \dots, x_j(n)]$  常用以下 2 种方法计算它们之间的相似度:

1) 欧氏距离公式法<sup>[6]</sup>

$$d_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_0(i) - x_j(i))^2} \quad (1)$$

式中,  $d_j$  ——  $x_0$  和  $x_j$  之间的欧氏距离;  $n$  —— 总元素数。

2) 灰色关联度法<sup>[7]</sup>

首先, 根据式(2)计算  $x_0$  和  $x_j$  第  $i$  维元素的关联系数  $\varepsilon_j(i)$ :

$$\varepsilon_j(i) = \frac{\min_j \min_i |x_0(i) - x_j(i)| + \rho \min_j \max_i |x_0(i) - x_j(i)|}{|x_0(i) - x_j(i)| + \rho \max_j \max_i |x_0(i) - x_j(i)|} \quad (2)$$

式中,  $\rho$  ——分辨系数, 一般取值 0.5。则数据序列  $x_0$  和  $x_j$  的灰色关联度  $F_j$  定义为:

$$F_j = \frac{1}{n} \sum_{i=1}^n \varepsilon_j(i) \quad (3)$$

## 1.2 改进灰色-欧氏距离相似度

式(1)所计算的欧氏距离反映 2 个数据序列在空间上的距离大小, 而式(3)的灰色关联度则反映它们在空间上的形状相似性。显然, 2 种方法在相似度表述上都欠缺完整性。鉴于此, 论文提出一种综合相似系数和距离系数的统一计量相似度。

考虑到式(2)的关联度系数  $\varepsilon_j(i)$  恒非负, 即不能反映出两序列间的正负关联性, 本文首先对灰色关联系数的计算方法做出以下改进。

设序列  $X=[x_1, x_2 \cdots x_n]$  和  $Y=[y_1, y_2 \cdots y_n]$ , 并有  $D_{xi}=(x_i - x_{i-1})$ ,  $D_{yi}=(y_i - y_{i-1})$  分别表示序列  $X$  和  $Y$  的第  $i$  维元素相对于第  $i-1$  维元素的增量, 其中,  $i=2, 3, 4 \cdots n$ 。分以下情况求 2 个序列第  $i$  维元素的关联系数  $\varepsilon_j(i)$ :

1) 若  $D_{xi}$ ,  $D_{yi}$  同时为 0, 则:

$$\varepsilon(i) = 0 \quad (4)$$

2) 若  $D_{xi}$ ,  $D_{yi}$  不同时为 0, 则:

$$\varepsilon(i) = \frac{\lambda_i}{1 + \frac{1}{2} \left[ \left| |D_{xi}| - |D_{yi}| \right| + \left( 1 - \frac{\min(|D_{xi}|, |D_{yi}|)}{\max(|D_{xi}|, |D_{yi}|)} \right) \right]} \quad (5)$$

式(5)引入符号函数  $\lambda_i$ , 即:

$$\lambda_i = \begin{cases} 1, & D_{xi} \times D_{yi} > 0 \\ -1, & D_{xi} \times D_{yi} < 0 \end{cases} \quad (6)$$

若  $\varepsilon(i) > 0$ , 表示  $X$  与  $Y$  的第  $i$  维元素相对第  $i-1$  维元素变化的正负方向一致; 若  $\varepsilon(i) < 0$ , 则表示变化方向相反。

结合(1)式欧氏距离的定义, 得出本文所提出的改进灰色-欧氏距离相似度函数  $S_{xy}$ :

$$S_{xy} = \mu \frac{1}{n-1} \sum_{i=2}^n \omega_i \varepsilon(i) + \nu \left[ 1 - \sqrt{\frac{1}{n} \sum_{i=1}^n (\omega_i x_i - \omega_i y_i)^2} \right] \quad (7)$$

式中,  $S_{xy}$  ——数据序列  $X$  与  $Y$  之间的相似度;  $\mu$ 、 $\nu$  ——权重系数 (其中  $\mu + \nu = 1$ , 本文取  $\mu = 0.5$ ,  $\nu = 0.5$ );  $\omega_i$  ——数据序列第  $i$  维元素的权重,  $\omega_i \in [0, 1]$ 。显然,  $S_{xy}$  包含 2 部分, 第 1 部分表示序列  $X$  与  $Y$  之间的“型相似”度, 第 2 部分表示“值相似”度。

## 1.3 日特征向量的选取

影响光伏系统功率输出的天气因素有很多, 而实际上, 有些天气因素与光伏功率输出相关性很小。在相似日聚类过程中, 如果将所有这些因素都考虑进去, 反而会干扰聚类效果。基于此, 本文选取与光伏功率输出具有较大相关性的因素构成用以相似日聚类的日特征向量。

光伏阵列系统单位面积输出力可表示为<sup>[5]</sup>:

$$P_{PV}(t) = I(t) A \eta_{pc} [1 - \beta(t_c - T_{cref})] \quad (8)$$

式中,  $I(t)$  —— $t$  时刻的光的辐照度,  $W/m^2$ ;  $A$  ——为光伏组件面积,  $m^2$ ;  $\eta_{pc}$  ——直流变换环节 MPPT 效率;  $\beta$  ——温度系数;  $t_c$ 、 $T_{cref}$  ——环境温度和参考温度。

很明显, 式(8)表明光照强度和温度对光伏发电功率输出具有不容忽视的影响。另外, 考虑到太阳辐照强度与太阳时角有关, 而太阳时角则可通过日照时间长短反映。故选取日特征向量为:

$$X = (t, I, I_{max}, t_1, T, \Delta T, t_2, t_3) \quad (9)$$

式中,  $t$  ——日照时长, min;  $I$ 、 $I_{max}$  ——平均太阳辐照强度和日最大辐照强度,  $W/m^2$ ;  $t_1$  ——最大太阳辐照值出现时刻;  $T$ 、 $\Delta T$  ——平均气温和日最大温差,  $^{\circ}C$ ;  $t_2$ 、 $t_3$  ——日最高气温和最低气温出现时刻,  $^{\circ}C$ 。

## 2 模型理论基础

### 2.1 小波分解与集成经验模式分解<sup>[9]</sup>及其比较

光伏功率预测就是通过历史数据所隐含的内在规律, 推断出光伏阵列未来的输出功率。考虑到光伏功率数据呈非线性, 且具有较高的随机性和突变性, 因此, 为降低数据的非平稳性, 提高预测精度, 本文结合小波分解和集成经验模式分解的优势, 对光伏数据进行双分解。

小波变换是一种多尺度的信号分解方法, 通过

小波基的伸缩、平移,并与被分析信号进行比较,可实现信号的时频分析局部化。最终能达到在高频处时间细分,低频处频率细分的效果。

连续小波变换可定义为:

$$WT_x(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-b}{a}\right) dt = \langle x(t), \psi_{a,b}(t) \rangle \quad (10)$$

式中,  $\psi(t)$  —— 基本函数;  $\psi_{a,b}(t)$  —— 小波基;  $x(t)$  —— 一个平方可积的信号,  $x(t) \in L^2(R)$ 。

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (11)$$

式中,  $a$ 、 $b$  —— 表征其尺度和位移的参数。将  $a$ 、 $b$  进行离散化处理,可得  $x(t)$  的离散小波变换:

$$WT_x(j,k) = \int_{-\infty}^{+\infty} x(t) \psi_{j,k}(t) dt, \quad j, k \in \mathbf{Z} \quad (12)$$

而 EEMD 是针对经验模态分解(EMD)的不足而提出的一种噪声辅助数据分析法。通过将白噪声加入到待分解的信号中,消除原始间歇现象,抑制分解结果中噪声所产生的影响。EEMD 将时间序列数据自适应地分解成若干个内禀模态函数(imf)和剩余分量(RES),从而更加有效地揭露原始历史数据的内在变化规律。

为对比 WD 与 EEMD 在数据序列中的趋势检索能力,本文用这 2 种方法对含噪的斜坡信号作分解,并进行定量分析。图 1 是加噪前后的斜坡信号。

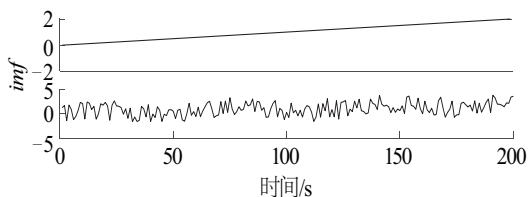


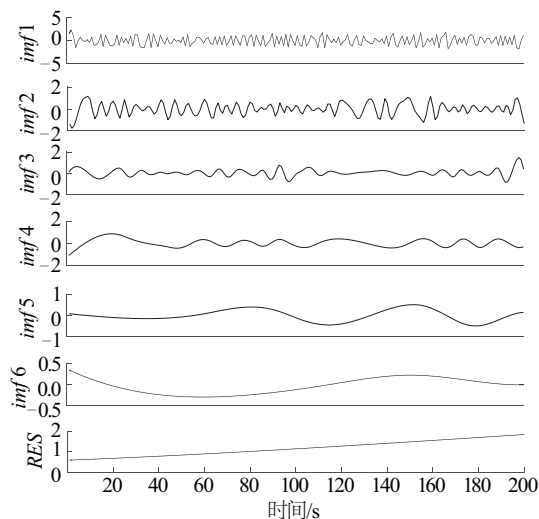
图1 原始斜坡信号与含噪斜坡信号

Fig. 1 Original ramp signal and the noisy signal

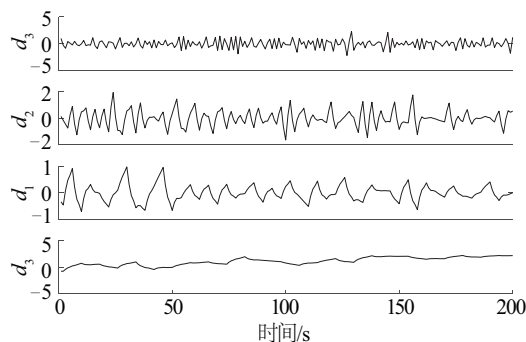
图 2 分别为该信号的 EEMD 与 WD 分解结果(图中  $d_1 \sim d_3$  为高频信号)。表 1 为趋势分量 RES 和低频信号  $a_3$  与原始斜坡信号的均方差(RESE)和相关系数( $R$ )。

另外,为比较上述 2 种分解方法对信号局部特征的表现能力,依次采用它们分解突变波形信号,分解结果如图 3 所示,从图 3 可看出,WD 的分解分

量更能有效表现出原始波形信号变化处的局部波动特性。)



a. EEMD 分解结果



b. WD 分解结果

图2 EEMD与WD对加噪音斜坡信号分解结果对比

Fig. 2 EEMD and WD results of ramp signal with noisy

表1 WD与EEMD趋势检索能力的定量对比

Table 1 Quantitative comparison for trend retrieval ability between WD and EEMD

评估指标	均方差 RESE	相关系数 $R$
RES	0.3335	0.9701
$a_3$	0.5886	0.6632

注:RES—剩余分量; $a_3$ —低频信号。

小波变换的时频窗口具有可调性,当尺度较小时,窗口的时域部分窄而分析频率高,因此可反映数据局部特性,而 EEMD 依据信号本身的时间尺度进行分解,未经过频域转换。因而,集成经验模态分解在信号局部特征的表现能力上不如小波分解效果好。

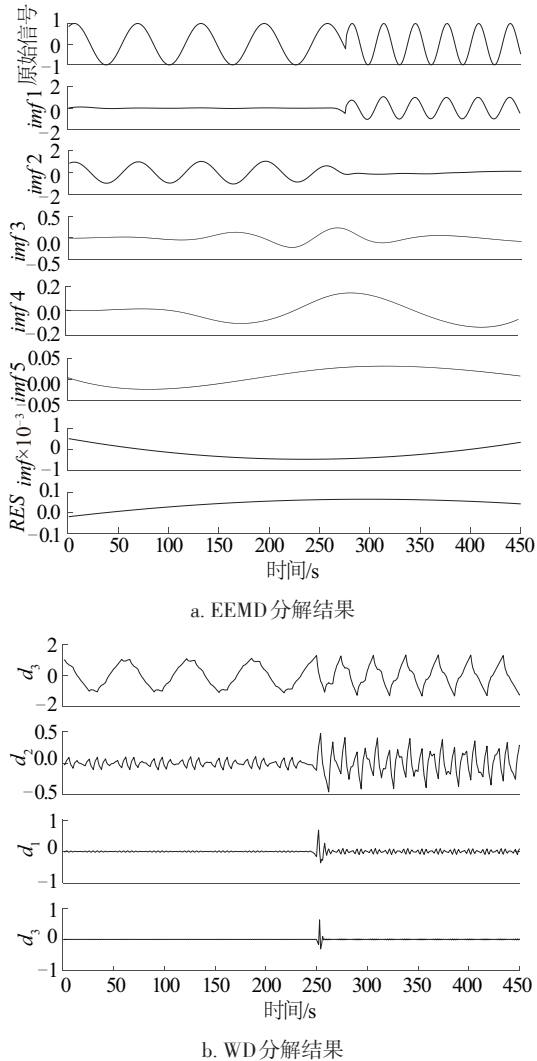


图3 EEMD与WD对突变波形信号分解结果的对比  
Fig. 3 EEMD and WD results of waveforms with sudden change

## 2.2 光伏功率的双分解理论

光照的强度是光伏阵列输出的最大决定因素。整体而言,由于地球自传和公转,地表光照强度具有相对比较稳定的规律性变化。另外,受环境因素影响,光伏功率表现出波动特性。基于此,本文结合 EEMD 和 WD 这 2 种分解方法的优势,对光伏输出数据进行双分解,具体步骤为:

1) 运用集成经验模态分解方法对原始光伏数据进行分解,得各本征模函数分量  $imf$  和剩余分量  $RES$ ,即趋势分量;

2) 将步骤 1) 的各  $imf$  分量进行等权叠加,得新序列  $S$ ;

3) 运用小波分解方法对序列  $S$  进行再分解,得各小波分量序列,称之为细节分量。

其中,步骤 1) 的趋势分量  $RES$  揭示了光伏输出趋势;而步骤 3) 的各细节分量的特征信息相对更为集中,既能保持各个时段的局部信息,又能进一步降低数据的不稳定性。

## 2.3 二次自适应支持向量机预测模型理论

支持向量机基本模型定义为特征空间上的间隔最大的线性分类器,其学习策略便是间隔最大化,最终可转换为一个二次规划问题的求解<sup>[10]</sup>。在光伏预测的非线性回归问题上,SVM 首先通过非线性函数将向量从低维空间映射到高维特征空间中,再在高维空间进行线性回归,取得在原空间进行线性回归的效果。具体回归函数为:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b \quad (13)$$

式中,  $\alpha_i$ 、 $\alpha_i^*$ ——均不为 0,对应的样本点就是支持向量;  $b$ ——偏移量;  $K(x, x_i)$ ——核函数,常用的核函数有线性核函数、多项式核函数、高斯基 BRF 核函数、Sigmoid 核函数等。预测过程中,核参数和惩罚系数的大小对算法的表现有非常关键的作用,为避免过学习或欠学习的现象,本文采用纵横交叉 (CSO)<sup>[11]</sup>算法对其进行优化。

为提高预测精度,本文在传统支持向量机预测理论上提出二次自适应支持向量机预测模型。根据前文分析,EEMD 分解后的趋势分量  $RES$  反映了历史数据的总体趋势,相对比较平稳,故运用一次支持向量机模型预测其未来走势,选用线性核函数:

$$K(x, x_i) = x \times x_i \quad (14)$$

而经过双分解后的细节分量集中反映了数据的局部特性,波动性相对较强。因此,论文对各细节分量采用二次预测方式,即将各细节分量的第一次支持向量机预测拟合值分别作为该分量第二次支持向量机模型的输入值,第二次预测输出值即为该分量最终预测结果。由于细节分量波动较大,故选用高斯核函数:

$$K(x, x_i) = \exp \left\{ -\frac{\|x - x_i\|^2}{2\delta^2} \right\} \quad (15)$$

式中,  $\delta$ ——核宽度。



最后,将各分量(包括趋势分量)的预测值进行重构,得出最终预测结果。由于二次向量机针对性地学习第一次向量机输出值与实际光伏出力之间的对应关系,因此,最终结果将具有更高的拟合度。二次自适应支持向量机模型采用了递进融合多个层次上的基分类器的集成学习核心思想,而集成学习已经被证明在机器学习诸方法中具有更好的预测精度和泛化能力<sup>[12]</sup>。

### 3 预测模型设计

#### 3.1 模型的输入、输出量

设计一次支持向量机输入变量由相似日输出功率和待预测日天气数据等构成,输入量有 11 个。预测某天  $T_i$  时刻(其中,  $T_i$  为 07:00~17:00 整点时刻)的一次自适应支持向量机的各输入变量如表 2 所示。

表 2 一次预测模型输入变量

Table 2 Input variable of first model

输入变量	变量含义
$x_1 \sim x_5$	与待预测日相似度最大 5 d 的 $T_i$ 时刻光伏功率的对应分解量
$x_6, x_7$	与待预测日相似度最大 1 d 的 $T_i$ 时刻前后两时刻光伏功率的对应分解量
$x_8 \sim x_{11}$	待预测日预报的平均气温、最低气温、最高气温和平均太阳辐照度的强度

输出值是待预测日  $T_i$  时刻光伏出力的对应分解量;而二次 SVM 模型输入量分别为该时刻分解量的一次 SVM 输出拟合值  $x_{b1}$ 、其前后 2 个整点时刻的一次 SVM 拟合值  $x_{b2}$ 、 $x_{b3}$  和一次 SVM 的最大  $x_{b4}$  与最小输出拟合值  $x_{b5}$ ,即输入量为 5 维向量  $X_i(x_{b1}, x_{b2}, x_{b3}, x_{b4}, x_{b5})$ 。输出量与一次 SVM 模型的对应输出量相同,即为该待预测时刻的分解分量。

#### 3.2 预测流程分析

预测流程如图 4 所示。

首先,采用改进灰色-欧氏距离作为相似度函数对历史样本日进行模糊聚类,并根据待预测日的气象信息,进行相似日识别,从而构建训练样本和测试样本。再将该样本的光伏数据进行双分解,得到

一系列细节分量和趋势分量。然后用本文所提的二次自适应支持向量机模型分别进行趋势预测和细节预测,最后将各分量的预测结果叠加,得出最终光伏预测值。

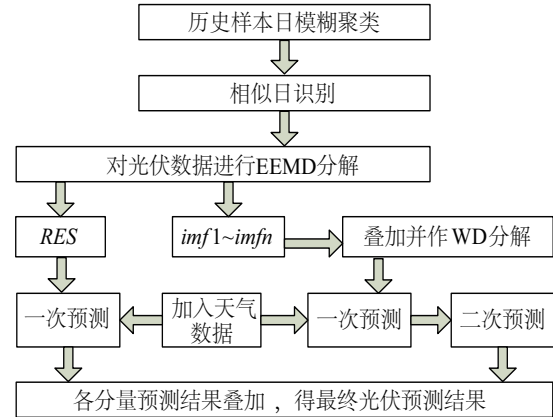


图 4 光伏输出功率短期预测模型框图

Fig. 4 Diagram of short-term PV output power forecasting model

### 4 实例验证

#### 4.1 实验数据

为验证本文所提的改进灰色-欧氏距离法的有效性,采用美国俄勒冈州本德一处光伏发电场(西经 121.31°,北纬 44.06°,海拔 1124 m)2010~2013 年的实测数据和当地气象站发布的环境数据进行实验。光伏功率数据每 15 分钟采样一次,环境数据包括辐照强度、温度、风速、湿度等,本文根据模型的设计选取部分相关数据,其中光伏输出功率取 07:00~17:00 整点时刻的小时力。

#### 4.2 聚类分析

本文以(9)式的日特征向量为指标对历史样本日进行模糊聚类,并将聚类为“突变型天气”的样本日以上、下午形式统计为图 5、图 6 所示的三维直方图。图中的  $X$  轴和  $Y$  轴分别为该时间段的平均温度与平均辐照度,  $Z$  轴为频数。

为证明本文所提的改进灰色-欧氏距离相似度函数在实际应用中具备可行性,将其聚类结果与以欧氏距离为相似度函数的聚类结果作对比。

经分析可知,除了各别奇异数据外,本文所提方法聚类划分的“突变型天气”的上午的辐照度主要集中

在  $440 \sim 590 \text{ W/m}^2$  之间, 比较密集分布于  $450 \sim 520 \text{ W/m}^2$  之间, 温度集中分布区间为  $[3 \text{ } ^\circ\text{C}, -4 \text{ } ^\circ\text{C}]$ 。

而基于欧氏距离法聚类的该类型天气上午的辐照度主要分布在  $330 \sim 580 \text{ W/m}^2$  区间, 而相对分布比较密集区间为  $360 \sim 530 \text{ W/m}^2$ , 温度集中分布区间为  $[5 \text{ } ^\circ\text{C}, -5 \text{ } ^\circ\text{C}]$ 。显然, 本文所提方法具有更好的聚类效果。下午情况亦然。

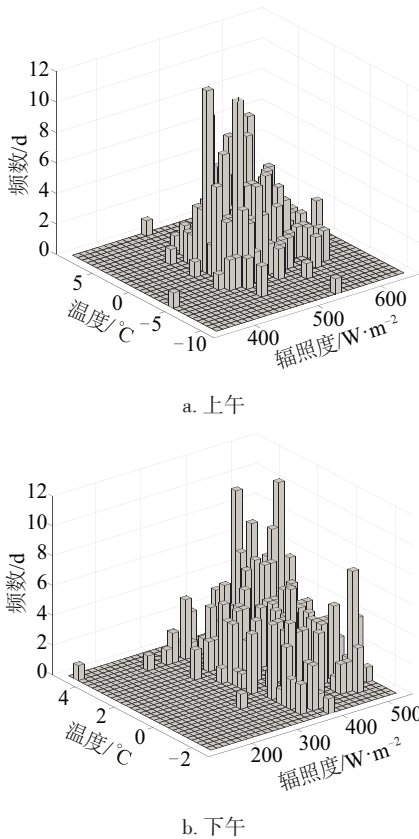


图5 改进相似度的“突变型天气”上、下午指标统计  
Fig. 5 Index statistics in morning/afternoon of mutant type by improved function improved function

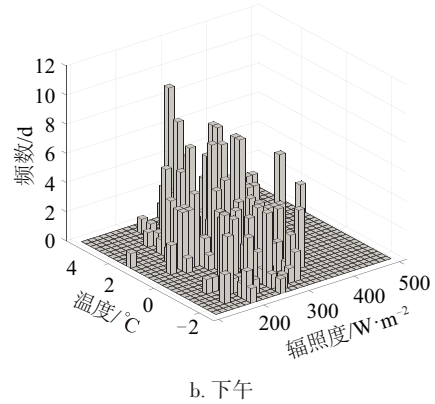
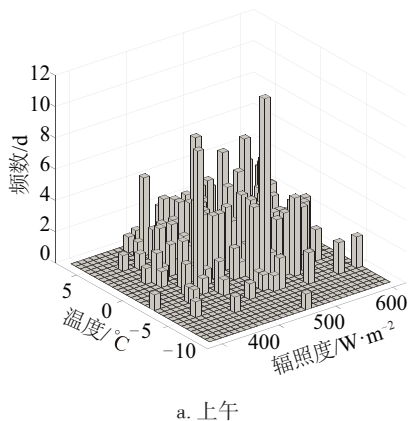


图6 欧氏距离法的“突变型天气”上、下午指标统计  
Fig. 6 Index statistics in morning/afternoon of mutant type by Euclidean distance function

### 4.3 预测结果与分析

论文选取具有代表性的晴天、雨天和多云阴天作为研究对象, 并采用平均绝对百分比误差 (MAPE) 和均方差 (RMSE) 这 2 种误差指标对预测结果作评估, 其定义公式为<sup>[13]</sup>:

$$MAPE = \frac{100}{N} \sum_{i=1}^N \frac{|P_i^f - P_i^a|}{P_i^a} \times 100\% \quad (16)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i^f - P_i^a)^2} \quad (17)$$

式中,  $P_i^f$  —— 预测功率值;  $P_i^a$  —— 实际光伏输出功率;  $N$  —— 数据总数。

为验证本文所提方法的有效性, 论文选取当地 2014 年 10 月 13 日 (晴天)、2014 年 10 月 21 日 (雨天) 和 2014 年 11 月 4 日 (多云) 作为待预测日, 并将分别与基于小波分解的支持向量机模型 (WD-SVM)、EEMD-SVM 模型和基于双分解的支持向量机模型 (WD-EEMD-SVM) (以欧氏距离作为相似度) 进行对比。仿真结果如图 7 所示。

通过表 3 的误差统计可知, 在不同种天气类型下, 本文所提的基于双分解的二次自适应 BP 神经网络模型的预测精度较其他 3 种模型的预测精度都要高。在晴天天气类型, 平均绝对百分比误差达到 8.51%。虽然雨天的光伏预测误差为 18.80%, 比晴天和云天的精度差较多, 但是相比其他 3 种模型而言, 其 MAPE 误差指标已优胜了 2 倍以上。表明本文所提的方法在一定程度上能提高光伏输出功率的预测精度, 尤其在突变的天气类型情况下。从而验证了本文方法的有效性。

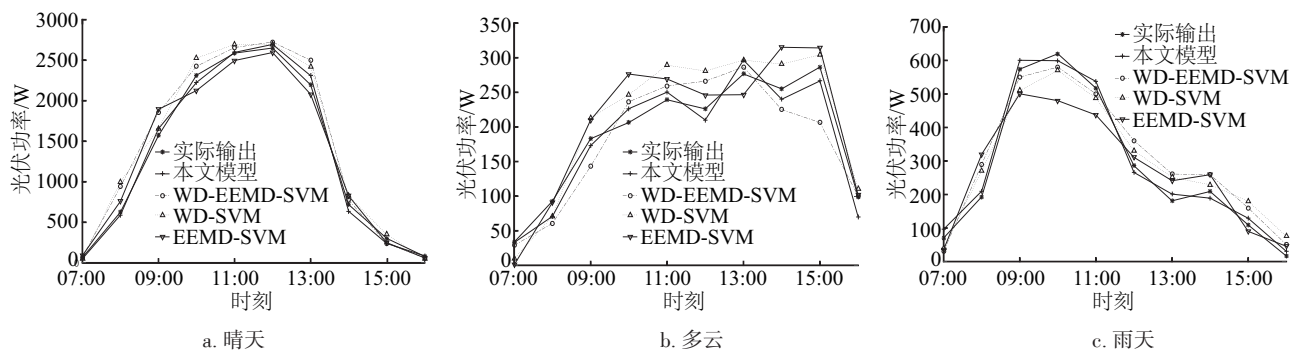


图7 3种气类型光伏输出功率短期预测结果图

Fig. 7 Forecasting results of short-term PV output power in three kinds of weather

表3 4种预测模型预测结果的误差对比

Table 3 Error comparison of results from the four forecasting models

模型	EEMD-SVM		WD-SVM		WD-EEMD-SVM		本文模型	
	MAPE/%	RMSE/%	MAPE/%	RMSE/%	MAPE/%	RMSE/%	MAPE/%	RMSE/%
晴天	19.52	140.92	18.17	164.36	15.22	173.60	8.51	64.94
多云	21.59	36.45	21.35	33.40	15.51	35.97	10.18	17.96
雨天	42.47	75.30	59.22	54.86	46.69	56.32	18.80	20.70

## 5 结 论

本文综合考虑数据序列间的灰色关联度和欧氏距离,并引入符号函数,得出改进灰色-欧氏距离相似度函数。接着将历史样本日进行模糊聚类。预测过程中,针对光伏数据的特点,论文将光伏输入数据作双分解处理,并运用集成学习机中叠加法的递进思想,提出自适应二次支持向量机预测模型。验证结果表明,本文所提的方法能够较好地提高光伏预测精度,具有一定的实用性和可行性。

### 〔参考文献〕

[1] 金鑫,袁越,傅质馨,等.天气类型聚类的支持向量机在光伏系统输出功率预测中的应用[J].现代电力,2013,30(4):13—18.

[1] Jin Xin, Yuan Yue, Fu Zhixin, et al. Application of support vector machine based on weather type clustering in power output forecasting of photovoltaic generation system[J]. Modern Electric Power, 2013, 30(4): 13—18

[2] 王守相,张娜.基于灰色神经网络组合模型的光伏短期出力预测[J].电力系统自动化,2012,36(19):1—5.

[2] Wang Shouxiang, Zhang Na. Short-term output power

forecast of photovoltaic based on a grey and neural network hybrid model[J]. Automation of Electric Power Systems, 2012, 36(19): 1—5.

[3] 茆美琴,龚文剑,张榴晨,等.基于EEMD-SVM方法的光伏电站短期出力预测[J].中国电机工程学报,2013,33(34):17—24.

[3] Mao Meiqin, Gong Wenjian, Zhang Liuchen, et al. Short-term photovoltaic generation forecasting based on EEMD-SVM combined method[J]. Proceedings of the CSEE, 2013, 33(34): 17—24.

[4] 丁明,王磊,毕锐.基于改进BP神经网络的光伏发电系统输出功率短期预测模型[J].电力系统保护与控制,2012,40(11):93—99.

[4] Ding Ming, Wang Lei, Bi Rui. A short-term prediction model to forecast output power of photovoltaic system based on improved BP neural network[J]. Power System Protection and Control, 2012, 40(11): 93—99.

[5] 王晓兰,葛鹏江.基于相似日和径向基函数神经网络的光伏阵列输出功率预测[J].电力自动化设备,2013,33(1):100—103.

[5] Wang Xiaolan, Ge Pengjiang. PV array output power forecasting based on similar day and RBFNN[J]. Electric Power Automation Equipment, 2013, 33(1): 100—103.

[6] 王勃,冯双磊,刘纯.基于天气分型的风电功率预测方法[J].电网技术,2014,38(1):93—98.

- [6] Wang Bo, Feng Shuanglei, Liu Chun. Study on weather typing based wind power prediction[J]. Power System Technology, 2014, 38(1): 93—98.
- [7] 刘文颖, 门德月, 梁纪峰, 等. 基于灰色关联度与LSSVM组合的月度负荷预测[J]. 电网技术, 2012, 36(8): 228—232.
- [7] Liu Wenying, Men Deyue, Liang Jifeng, et al. Monthly load forecasting based on grey relational degree and least squares support vector machine[J]. Power System Technology, 2012, 36(8): 228—232.
- [8] 韩晓娟, 陈跃燕, 张 浩, 等. 基于小波包分解的混合储能技术在平抑风电场功率波动中的应用[J]. 中国电机工程学报, 2013, 33(19): 8—13.
- [8] Han Xiaojuan, Chen Yueyan, Zhang Hao, et al. Application of hybrid energy storage technology based on wavelet packet decomposition in smoothing the fluctuations of wind power[J]. Proceedings of the CSEE, 2013, 33(19): 8—13.
- [9] 王军辉, 贾 嵘, 谭 泊. 基于EEMD和模糊C均值聚类的风电机组齿轮箱故障诊断[J]. 太阳能学报, 2015, 36(2): 319—324.
- [9] Wang Junhui, Jia Rong, Tan Bo. Fault diagnosis of wind turbine's gearbox based on EEMD and fuzzy C means clustering[J]. Acta Energaie Solaris Sinica, 2015, 36(2): 319—324.
- [10] 罗 文, 王莉娜. 基于小波分解与遗传算法和支持向量机的短期风速预测[J]. 太阳能学报, 2012, 33(8): 1327—1333.
- [10] Luo Wen, Wang Lina. Short-term wind speed forecasting based on wavelet decomposition and genetic algorithm as well as support vector machine[J]. Acta Energaie Solaris Sinica, 2012, 33(8): 1327—1333.
- [11] 彭显刚, 林利祥, 刘 艺, 等. 基于纵横交叉-拉丁超立方采样蒙特卡洛模拟法的分布式电源优化配置[J]. 中国电机工程学报, 2015, 35(16): 4077—4085.
- [11] Peng Xiangang, Lin Lixiang, Liu Yi, et al. Optimal distributed generator allocation method based on correlation latin hyper cube sampling monte carlo simulation embedded crisscross optimization algorithm[J]. Proceedings of the CSEE, 2015, 35(16): 4077—4085.
- [12] 张春霞, 张讲社. 选择性集成学习算法综述[J]. 计算机学报, 2011, 34(8): 1399—1410.
- [12] Zhang Chunxia, Zhang Jianshe. A survey of selective ensemble learning algorithms[J]. Chinese Journal of Computers, 2011, 34(8): 1399—1410.
- [13] Pinson P, Nielsen H, Moller J, et al. Non parametric probabilistic forecasts of wind power: Required properties and evaluation[J]. Wind Energy, 2007, 10(6): 497—516.

## FORECASTING PHOTOVOLTAIC POWER BASED ON QUADRIC SELF-ADAPTIVE SVM MODEL

Yin Hao, Chen Yunlong, Meng Anbo, Lin Yicheng

(College of Automation, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** To make up for the defect of the traditional similarity function in the fuzzy clustering, a new function called improved gray-Euclidean distance is proposed, and then the history day samples are divided into several categories. In view of the complexity of photovoltaic power data, the paper decomposes it into the trend term and the detail term based on the respective advantages of wavelet transform (WD) and ensemble empirical mode decomposition (EEMD), then the quadric self-adaptive SVM model is first employed in the paper, forecasting and modeling the data sequence respectively. Finally, the forecasting model is verified by the field data of a PV generation system in Oregon of America. The research results show that the proposed forecasting model possesses high accuracy, proven feasible and practical.

**Keywords:** photovoltaic prediction; wavelet transform; ensemble empirical mode decomposition; quadric self-adaptive SVM model; similarity function