

文章编号:1671-6833(2020)05-0044-06

基于相似日和特征提取的短期风电功率预测

张颖超<sup>1,2</sup>, 成金杰<sup>1</sup>, 邓 华<sup>1,2</sup>, 宗 阳<sup>1</sup>, 章 璇<sup>1</sup>

(1.南京信息工程大学 自动化学院,江苏 南京 210044; 2.南京信息工程大学 气象灾害预报预警与评估  
协同创新中心,江苏 南京 210044)

**摘 要:**为提高短期风电功率预测精度,增强预测模型对特定天气状况的代表性和适应性,提出一种基于离散 Fréchet 距离与核熵成分分析(KECA)相结合的数据处理方法。通过引入离散 Fréchet 距离,建立匹配相似日的数学模型,提取与预测日相似的样本,使用 KECA 从多种气象要素中提取合适的非线性主元作为支持向量机(SVM)模型的输入。实验结果表明:所提出的方法明显提高了预测精度并具有一定的适用性。

**关键词:** 离散 Fréchet 距离;相似日;核熵成分分析;支持向量机;功率预测

**中图分类号:** TM614      **文献标志码:** A      **doi:**10.13705/j.issn.1671-6833.2020.02.023

0 引言

风电机组的有功出力受自然风影响,具有较高的间歇性和波动性。当大规模风电并网运行时,传统电力系统的平衡能力受到破坏,给电力部门的调控策略带来极大的挑战<sup>[1]</sup>。要想在提高风力发电在电力系统中比重的同时,确保电力系统安全稳定运行,使用较高精度的短期风电功率预测技术是有效措施之一。

目前,短期风电功率预测技术主要分为物理方法和统计方法<sup>[2]</sup>。物理方法通过对风电场的地理位置进行物理建模,将高精度的数值天气预报(numerical weather prediction, NWP)结果和风电场的地理信息转换为风机轮毂高度处的风速、风向等信息,再结合功率曲线推算出风机的有功出力<sup>[3]</sup>;统计方法需要准确的风电场历史观测数据去训练模型,由于组合预测方法和先进的参数寻优算法能明显提高其预测精度,使得统计方法成为目前较为常用的短期风电功率预测方法<sup>[4]</sup>。

然而,统计方法的训练样本多数为风电场全年的历史观测数据,一方面,过多的训练样本存在冗余信息,会导致过拟合现象;另一方面,这种整体性地选择训练样本难以适应复杂多样的天气状

况,预测精度也往往达不到理想效果。近年来,一些学者借鉴电力负荷预测中“相似日”的原理选取训练样本,并取得了不错的预测效果。文献[5-7]选取风速、风向、温度、气压和相对湿度等气象信息的均值或最值构成天气模式向量,通过关联系数和夹角余弦等作为相似性判据,对历史日进行搜索,构成训练样本集。文献[8-11]采用聚类方式,根据相似的天气模式向量将训练样本归类,建立自适应天气变化的预测模型。以上结果表明,与没有对样本进行筛选和分类的其他模型相比,其预测精度得到明显提高。

由于天气状况比较复杂,实践中很难将其定性为特定的类别,即使在相似的天气状况下,风速变化曲线也会相差较大。而风速作为影响风机出力的主要因素,由于风机桨叶的惯性作用,风速的波动对风电机组的有功出力会有较大影响<sup>[12-13]</sup>。因此,在选择相似日时应充分考虑风速变化趋势的相似性。基于此,通过离散 Fréchet 距离作为风速、风向、温度、气压、相对湿度等时间序列的相似性判据来匹配相似日。此外,这些气象要素之间存在复杂的非线性关系,使用核熵成分分析(kernel entropy component analysis, KECA)在高维特征空间提取非线性主元作为支持向量机

收稿日期:2019-11-11;修订日期:2020-01-15  
基金项目:国家自然科学基金资助项目(41675156);江苏省六大人才高峰项目(WLW-021)  
作者简介:张颖超(1960—),男,江苏沛县人,南京信息工程大学教授,博士生导师,主要从事复杂系统建模与评估、灾害评估与损失预测等研究,E-mail:420585350@qq.com。

(SVM)的输入。实验结果表明,该方法能够较好地历史数据中选取相似日和非线性主元,且预测精度较高。

## 1 数据和方法

### 1.1 数据

以云南某风电场为例,实验数据为 13 台型号相同的风机运行数据和测风塔采集的气象数据,采样时段为 2016 年 7 月至 8 月,采样间隔为 15 min,风机信息如表 1 所示。

表 1 风机基本信息

Table 1 Basic information of wind turbines

切入 风速/ ( $\text{m}\cdot\text{s}^{-1}$ )	额定 风速/ ( $\text{m}\cdot\text{s}^{-1}$ )	切出 风速/ ( $\text{m}\cdot\text{s}^{-1}$ )	额定 功率/ MW	风轮 直径/ m	扫风 面积 / $\text{m}^2$
3	11	25	1.5	93	6 805

### 1.2 离散 Fréchet 距离

Fréchet 距离是法国数学家 Fréchet 基于空间路径相似度的方法提出的,能够度量两曲线间整体走势的相似情况,其定义<sup>[14]</sup>如下:

**定义** 设  $f, g: [0, 1] \rightarrow R^2$  是给定两条连续曲线,  $\alpha, \beta: [0, 1] \rightarrow [0, 1]$  为连续非减的实函数,且满足  $\alpha(0) = \beta(0) = 0, \alpha(1) = \beta(1) = 1$ , 则两曲线间的 Fréchet 距离为:

$$\delta_F(f, g) = \inf_{\alpha, \beta} \max_{s \in [0, 1]} \text{dist}(f(\alpha(t)), g(\beta(t))). \quad (1)$$

可见,Fréchet 距离的实质是寻找一对实函数  $\alpha(t), \beta(t)$  最小化曲线  $f$  与  $g$  之间连通的极大距离。为解决实际问题,Eiter 和 Mannila 在此基础上,通过离散化和求极限的思想提出了离散 Fréchet 距离的概念。

设给定两条离散有序序列  $P = \langle u_1, \dots, u_p \rangle$  和  $Q = \langle v_1, \dots, v_q \rangle, L$  为  $P, Q$  间各序列点组成的序列点对:

$$L = \{ (u_{a_1}, v_{b_1}), (u_{a_2}, v_{b_2}), \dots, (u_{a_k}, v_{b_k}) \},$$

其中,  $a_1 = b_1 = 1, a_k = p, b_k = q$  且对于  $i = 1, \dots, k$ , 满足  $a_{i+1} = a_i$  或  $a_{i+1} = a_i + 1, b_{i+1} = b_i$  或  $b_{i+1} = b_i + 1$ , 即序列点对  $L$  同时服从  $P, Q$  中各序列点的对应次序关系。

**定义**  $\|L\|$  为序列点对  $L$  中最长连接的长度,表示如下:

$$\|L\| = \max_{i=1,2,\dots,k} d(u_{a_i}, v_{b_i}). \quad (2)$$

则两曲线间的离散 Fréchet 距离的定义<sup>[15]</sup>为:

$$d(P, Q) = \min \{ \|L\| \}. \quad (3)$$

$d(P, Q)$  可根据以下公式递归求解:

$$d(P, Q) =$$

$$\max \left\{ \begin{array}{l} d_E(u_p, v_q); \\ \min \left\{ \begin{array}{l} d(\langle u_1, \dots, u_{p-1} \rangle, \langle v_1, \dots, v_q \rangle), \\ \forall n \neq 1; \\ d(\langle u_1, \dots, u_p \rangle, \langle v_1, \dots, v_{q-1} \rangle), \\ \forall m \neq 1; \\ d(\langle u_1, \dots, u_{p-1} \rangle, \langle v_1, \dots, v_{q-1} \rangle), \\ \forall n \neq 1 \text{ 且 } m \neq 1, \end{array} \right. \end{array} \right. \quad (4)$$

式中:  $d_E(u_p, v_q)$  是两点  $u_p$  和  $v_q$  间的欧式距离。

### 1.3 核熵成分分析 (KECA) 原理

核熵成分分析 (kernel entropy component analysis, KECA) 是数据转换与维度变换的一种新方法。该方法以二次 Renyi 熵值作为信息量的度量,在高维特征空间提取非线性主元,能够很好地表征原始特征数据关于熵的内在结构<sup>[16]</sup>。

设某一概率系统中,有概率密度函数为  $p(x)$  的数据集  $X = [x_1, x_2, \dots, x_N], x_i \in R^D, X$  对应的二次 Renyi 熵表示如下:

$$H(p) = -\log \int p^2(x) dx. \quad (5)$$

令  $V(p) = \int p^2(x) dx = E\{p(x)\}$ , 为求  $H(p)$  的值,需对  $V(p)$  进行核密度估计,核密度估计函数为:

$$\hat{p}(x) = \frac{1}{N} \sum_{i=1}^N k_\sigma(x, x_i), \quad (6)$$

式中:  $k_\sigma(x, x_i)$  为核函数;  $\sigma$  为宽度参数。

根据高斯函数卷积定理,化简可得

$$\hat{V}(p) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N k_{\sqrt{2}\sigma}(x_i, x_j) = \frac{1}{N^2} \mathbf{I}^T \mathbf{K} \mathbf{I}, \quad (7)$$

式中:  $\mathbf{K}$  为  $N \times N$  核矩阵;  $\mathbf{I}$  为各元素均为  $N \times 1$  的矩阵。将  $\mathbf{K}$  特征值分解为:  $\mathbf{K} = \mathbf{E} \mathbf{D} \mathbf{E}^T$ , 其中  $\mathbf{D}$  为特征值  $\lambda_1, \lambda_2, \dots, \lambda_N$  构成的对角矩阵,  $\mathbf{E}$  为对应的特征向量  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$  构成的特征矩阵, 则式 (7) 可表示为:

$$\hat{V}(p) = \frac{1}{N^2} \sum_{i=1}^N (\sqrt{\lambda_i} \mathbf{e}_i^T \mathbf{I})^2. \quad (8)$$

KECA 根据每个  $\sqrt{\lambda_i} \mathbf{e}_i^T$  对二次 Renyi 熵的估计,将核矩阵的特征值和相应的特征向量按估计值大小对其降序排列,由此产生的 KECA 映射  $\Phi_{eca}$  为:

$$\Phi_{eca} = U_d \Phi = \mathbf{D}_d^{\frac{1}{2}} \mathbf{E}_d^T, \quad (9)$$

式中:  $\mathbf{D}_d$  为对特征值  $\lambda_1, \lambda_2, \dots, \lambda_N$  降序排列后选取的前  $d$  个特征值所构成的对角矩阵;  $\mathbf{E}_d$  则是由相应的特征向量构成。

对于一个新的测试样本  $x_{\text{new}}$  在特征空间上的投影可表示为:

$$Y_{\text{new}} = D_d^{-\frac{1}{2}} E_d^T K(x, x_{\text{new}}). \quad (10)$$

由上述可知,核相关矩阵的维数等于训练样本的个数。因此,KECA 方法最多能够从高维特征空间提取  $N$  个非线性主元作为训练模型的输入。

#### 1.4 SVM 理论

支持向量机(support vector machine, SVM)是建立在 VC 维理论(Vapnik-Chervonenkis dimension)和结构风险最小化准则的基础上的,相较神经网络等学习方法而言具有严格的理论基础<sup>[17]</sup>。此外,SVM 能够很好地解决小样本、非线性和高维度等问题。在案例中,使用了著名的开源软件包 LibSVM<sup>[18]</sup>,选取的 SVM 核函数为径向基函数。

## 2 基于相似日和特征提取的 SVM 建模

### 2.1 相似日的匹配

影响风机出力的主要因素有风速、风向、温度、气压、相对湿度等 5 个气象要素,以这些气象要素的采样点序列匹配相似日,具体步骤如下。

(1)假设有  $N$  个历史日,首先分别将  $N$  个历史日和预测日各气象要素的采样点序列表示成  $\{(i, c_i)\}_{i=1, \dots, k}$  形式,其中  $i$  是按采样时间顺序排列的序列号,  $c_i$  是采集数据的瞬时值,每隔 15 min 采样一次,则  $k=96$ 。

(2)分别将风速、风向、温度、气压、相对湿度的采样点序列按式(11)归一化处理:

$$c'_i = \frac{c_i - \min(c_i)}{\max(c_i) - \min(c_i)}. \quad (11)$$

根据式(4)分别计算各历史日与预测日对应气象要素采样点序列的离散 Fréchet 距离  $d_{i,j}$ ,获得的距离矩阵  $D$  如下:

$$D = [d_{i,j}], \quad (12)$$

式中:  $i=1, \dots, 5$  为第  $i$  个气象要素;  $j=1, \dots, N$  为第  $j$  个历史日。

(3)根据距离矩阵  $D$  计算各历史日与预测日的整体相似度,为表征影响风机出力的关键因素在相似日判定中的主导作用,采用连乘的方式定义整体相似度公式,整体相似度公式如下:

$$F_j = \prod_{i=1}^5 \frac{1}{d_{i,j}}. \quad (13)$$

(4)选取  $N$  个历史日中相似度最高的  $m$  日作

为预测日的相似日样本。

### 2.2 特征提取

风机发电功率与各影响因素之间存在着复杂的非线性关系。KECA 方法通过核函数把输入的数据映射到高维特征空间,进行非线性主元的提取,最大化保留了原始数据中的信息熵。

数据集  $X = [x_1, x_2, \dots, x_{m \times 96}]$  为选取出来的  $m$  个相似日数据,  $x_i$  表示为同时刻风速、风向、温度、气压、相对湿度构成的向量,则基于 KECA 方法的特征提取的具体步骤分为以下 3 步。

(1)选取径向基函数作为式(6)的核函数,并构造相应的核相关矩阵  $K$ , 径向基函数表示为:

$$k_\sigma(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right). \quad (14)$$

(2)对核矩阵  $K$  进行特征值分解,并根据式(8)对二次 Renyi 熵的估计,将核矩阵的特征值和相应的特征向量按估计值大小降序排列。

(3)确定非线性主元个数  $d$ ,即将数据映射到由降序排列后的前  $d$  个特征值和特征向量所张成特征子空间中。根据式(9)求取产生的非线性映射。

### 2.3 预测模型建立

结合相似日的外推效果,KECA 方法能够表征数据关于熵的内在结构及 SVM 处理小样本、高维度的优点,提出了基于相似日和特征提取的 SVM 短期风电功率预测模型,如图 1 所示。

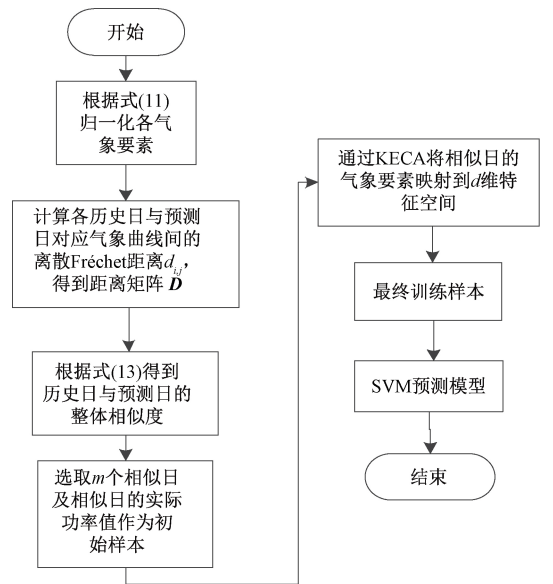


图 1 预测流程图

Figure 1 Forecasting flow chart

将  $m$  个相似日的风速、风向、温度、气压、相对湿度及相应时刻的功率作为初始样本集,通过

KECA 方法对这 5 个气象要素的初始样本进行分析,提取出的非线性主元作为 SVM 模型的输入,输出为对应时刻的实际功率值,完成 SVM 预测模型的训练。预测时,需要根据式(10)将预测日的气象数据即风速、风向、温度、压力、相对湿度映射到 KECA 产生的特征子空间中,并将其作为预测时的输入。

### 3 算例分析

以 8 月 27 日为预测日,其余为历史日,利用所提出的匹配相似日的方法,可找出历史日中与预测日相似度最高的前  $m$  天。为确定  $m$  值,将  $m$  取值为 1~15,分别使用相似日-KECA-SVM 方法对预测日的功率进行预测。实验中,通过网格法求取径向基核函数的参数和非线性主元的个数。图 2 为预测结果的相对均方根误差( $rRMSE$ )对比图。从图 2 可以看出,较少的相似日数量,无法代表样本空间的特征;当相似日选取较多时,存在冗余信息,不仅降低模型的预测性能,而且需要过多的训练时间。当  $m=7$  时,模型预测的效果相对较好,对应的径向基参数  $\sigma$  为 4.5,非线性主元  $d$  为 25。因此相似日为 7 d,表 2 为这 7 个相似日与预测日的整体相似度计算结果。

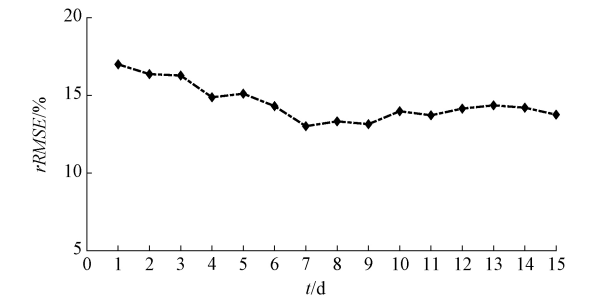


图 2 不同数量的相似日预测误差对比  
Figure 2 Comparison of prediction errors for different quantities of similar days

表 2 整体相似度计算结果						
Table 2 Calculation results of similarity degree						
相似日	07-02	07-03	07-04	07-19	08-05	08-11 08-18
相似度	0.166	0.184	0.156	0.205	0.186	0.162 0.180

图 3 为用离散 Fréchet 距离作为相似性判据提取出的相似日的风速曲线。图中每个采样点为每 15 min 采样一次。为了能清楚直观地比较,图中给出了较为相似的一条风速曲线。可以看出,提取出的相似日的风速曲线不仅在幅值上与预测日的风速曲线比较接近,而且整体的变化趋势上也大致相同。说明整体相似度的定义能够很好地

表征风速的主导作用,而离散 Fréchet 距离作为曲线相似性的判据,能够充分考虑到风速整体走势的相似情况。选取相似日后,相似日和预测日的风速有了相似的变化趋势,天气状况也较为接近。

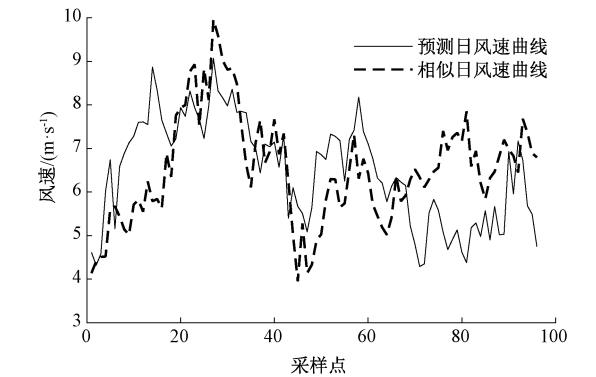


图 3 相似日的风速曲线  
Figure 3 Wind speed curve on similar days

为了验证相似日匹配和核熵成分分析方法的合理性和有效性,将以下 3 种预测方法进行对比。

方法 1:SVM 预测方法,即直接使用原始样本训练 SVM 模型,由于原始样本数据量大,模型的训练需要花费大量的时间。

方法 2:相似日-SVM 预测方法,即采用上述的相似日匹配的方法,从原始样本中提取相似日,将相似日的风速、风向、温度、气压、相对湿度和对应时刻的实际功率作为 SVM 的训练样本。

方法 3:相似日-KECA-SVM 预测方法,即在方法 2 的基础上将 7 个相似日的风速、风向、温度、气压、相对湿度及对应时刻的实际功率作为初始样本,并采用 KECA 方法提取相似日各气象要素的非线性主元,最后将  $d$  个主元作为 SVM 的输入节点和实际功率作为输出,完成 SVM 模型的训练。

上述 3 种方法的预测结果如图 4 所示。图中每个采样点为每 15 min 采样一次。结合图 3 可以看出,使用原始样本做预测时,单纯的 SVM 预测结果较差,大多数的预测值与实际值相比明显偏低。相似日-SVM 预测方法过滤了原始样本中的冗余信息,使得相似日和预测日的气象信息较为接近,而且风速变化的整体趋势也较为吻合。预测的结果表明,在风速变化较大的时间段预测值更加接近实际观测值。相似日-KECA-SVM 预测方法不仅在风速变化较大的时间段的预测精度得到提高,而且在风速变化幅度较小的时间段(即采样点为 70~90)预测精度也有一定的提高。

图 5 为 3 种方法预测的绝对误差,图中每个采样点为每 15 min 采样一次。从中可以看出单

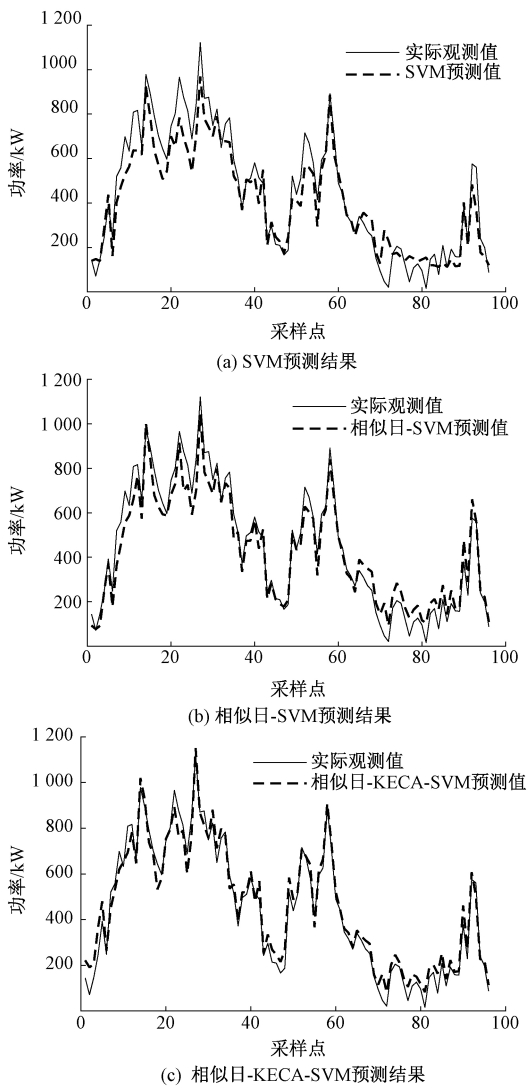


图4 各方法风电功率预测值对比

Figure 4 Comparison of predicted values of different methods

纯地使用支持向量机作预测,其绝对误差的波动范围较大,出现较大误差的次数较多。经过相似日方法处理后,误差波动范围变小,尤其在风速波动较大的时间段,出现较大误差的次数也较少。而经过相似日匹配和核熵成分分析后,预测效果最好。

分别采用相对平均绝对误差( $rMAE$ )、相对均方根误差( $rRMSE$ )和均方根误差( $RMSE$ )作为评价各方法预测效果的指标,各误差计算结果如表3所示。

表3 各方法预测误差比较

预测方法	$rMAE/\%$	$rRMSE/\%$	$RMSE/kW$
SVM 方法	15.86	20.33	90.81
相似日-SVM 方法	12.07	15.99	71.66
相似日-KECA-SVM 方法	10.27	13.10	63.92

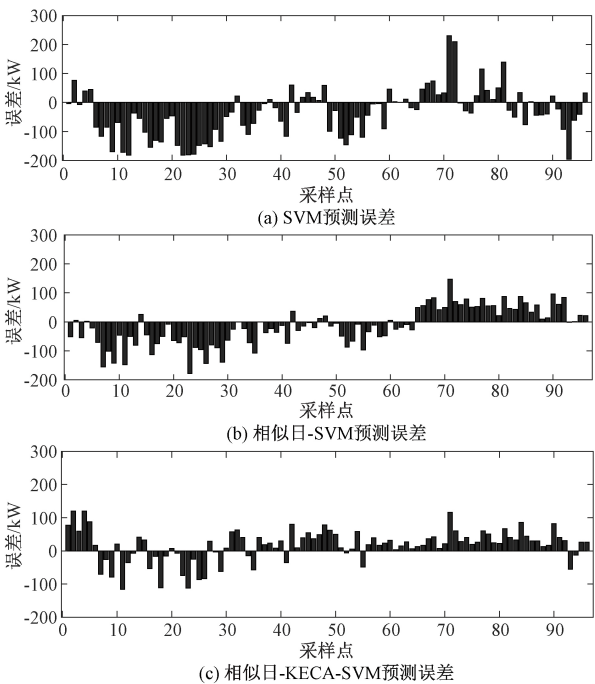


图5 各种方法预测的绝对误差

Figure 5 Prediction absolute error of various methods

从表3可以看出,经过相似日的匹配和核熵成分分析后,各个误差指标都有不同程度的降低。相比于使用全部样本,相似日-KECA-SVM 预测效果更好,说明所提出的相似日匹配方法和核熵成分分析方法能够有效提高建模数据的相似性,从而降低误差。

4 结论

结合相似日原理和核熵成分分析方法对中国云南某风电场的采集数据进行预处理,主要优点如下:

(1)风速是影响风机出力的关键因素,其变化趋势在相似日判定中至关重要,通过离散Fréchet距离判别曲线相似性的方法来匹配相似日,有效兼顾了风速的变化趋势。相似日的选取增强了数据的相关性,提高了模型的预测精度。

(2)风机出力和各影响因素之间存在着复杂的非线性关系,经过核熵成分分析后在高维特征空间能更好地呈现数据集的特点,有利于预测精度的提高。

(3)针对风力发电的周期性不太明显、相似日出现的频率较低的情况,结合SVM具有支持小样本建模和较强的非线性学习等特点,建立的短期风电功率预测模型具有较好的泛化能力。

参考文献:

[1] 王健,严干贵,宋薇,等. 风电功率预测技术综述[J]. 东北电力大学学报, 2011, 31(3):20-24.

[2] 李军,卓泽赢. 基于经验小波变换和多核学习的风电功率短期预测[J]. 信息与控制, 2018, 47(4):437-447.

[3] 孙晓燕,时良振,徐瑞东,等. 基于区间样本和回声状态网络的风电功率不确定性预测[J]. 郑州大学学报(工学版), 2017, 38(1):56-62.

[4] 阎洁,许成志,刘永前,等. 基于风速云模型相似日的短期风电功率预测方法[J]. 电力系统自动化, 2018, 42(6):53-59.

[5] 孟洋洋,卢继平,孙华利,等. 基于相似日和人工神经网络的风电功率短期预测[J]. 电网技术, 2010, 34(12):163-167.

[6] 倪鹏,孙富荣. 基于相似日和风速连续性的风电场短期功率预测[J]. 山东电力技术, 2016, 43(11):39-43.

[7] 张宜阳,严欢. 基于分段分层相似日搜索和自适应脊波神经网络的风电功率多步预测[J]. 电网与清洁能源, 2015, 31(4):124-131.

[8] YESILBUDAK M. Clustering analysis of multidimensional wind speed data using K-means approach[C]// 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA). Birmingham, United Kingdom: IEEE, 2016: 961-965.

[9] 丁志勇,杨苹,杨曦,等. 基于连续时间段聚类的支持向量机风电功率预测方法[J]. 电力系统自动化, 2012, 36(14):131-135.

[10] 郑婷婷,王海霞,吕泉,等. 基于相似日选取样本的短期风电功率混沌预测[J]. 电网与清洁能源, 2013, 29(3):74-79.

[11] 郁琛,薛禹胜,文福拴,等. 按时序特征优化模型后在线选配的超短期风电预测[J]. 电力系统自动化, 2015, 39(8):5-11.

[12] GANGER D, ZHANG J S, VITTAL V. Statistical characterization of wind power ramps via extreme value analysis[J]. IEEE transactions on power systems, 2014, 29(6):3118-3119.

[13] 叶小岭,陈浩,郭晓杰,等. 基于风速升降特征的短期风电功率预测[J]. 电力系统保护与控制, 2016, 44(19):56-62.

[14] ALT H, GODAU M. Computing the Fréchet distance between two polygonal curves[J]. International journal of computational geometry & applications, 1995, 5(1/2):75-91.

[15] EITER T, MANNILA H. Computing discrete fréchet distance[R]. Viena: Information System Department, Technical University of Vienna, 1994.

[16] 李思同,庄强,金琳,等. 基于时频特征核熵成分分析的局部放电模式识别方法[J]. 高压电器, 2018, 54(6):125-131.

[17] 彭怀午,杨晓峰,刘方锐. 基于 SVM 方法的风电场短期风速预测[J]. 电网与清洁能源, 2009, 25(7):48-52.

[18] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[J]. ACM transactions on intelligent systems and technology, 2011, 2(3): 27.

Short-Term Wind Power Prediction Based on Similar Day and Feature Extraction

ZHANG Yingchao<sup>1,2</sup>, CHENG Jinjie<sup>1</sup>, DENG Hua<sup>1,2</sup>, ZONG Yang<sup>1</sup>, ZHANG Xuan<sup>1</sup>

(1.School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2.Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology, Nanjing 210044, China)

**Abstract:** In order to improve the short-term wind power prediction accuracy and enhance the representativeness and adaptability of the prediction model to specific weather conditions, a data processing method was proposed based on discrete Fréchet distance and kernel entropy component analysis (KECA).Using Fréchet distance, this paper established a mathematical model matching similar days, extracted samples similar to the prediction date, and then used KECA to extract suitable nonlinear principal elements from various meteorological elements as input of support vector machine (SVM) model. The simulation verification showed that the proposed method could significantly improve the prediction’s accuracy and has certain applicability.

**Key words:** discrete Fréchet distance; similar day; kernel entropy component analysis; support vector machine; power prediction