

基于 IFOA-FSVM 的 WRF 模式台风期阵风精细化预报

何彩芬¹, 钱斌凯², 金 炜^{2*}, 张国超¹

(1. 镇海区气象局, 浙江 宁波 315202; 2. 宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 为提高天气预报模式(Weather Research and Forecasting Model, WRF)输出中台风期阵风的预测精度, 将 WRF 模式输出与某观测站实况数据相结合, 提出一种台风期阵风精细化预报方法. 针对影响台风风速的因素众多, 而传统依据人工经验预判的风速存在较大误差的现状, 该方法构建了台风期阵风预测的模糊支持向量回归模型, 同时为解决模糊支持向量回归模型中惩罚因子 C 和核参数 g 难于确定的问题, 将果蝇优化算法(Fruit Fly Optimization Algorithm, FOA)引入到模糊支持向量机(Fuzzy Support Vector Machine, FSVM)的参数寻优中, 并根据风速回归的特点, 把果蝇优化算法引入到三维空间, 结合增强因子 γ 以提高传统果蝇优化算法的全局寻优能力. 实验结果表明, 本文构建的模型预测风速与实际风速基本一致, 相关性达到 99%, 不仅提高了 WRF 模式风速的适用性, 而且风速预测精度明显优于传统 FOA-FSVM 和 FOA-SVM 方法, 具有更强的泛化能力.

关键词: 果蝇优化算法; 模糊支持向量机; 风速预测; 台风

中图分类号: TP391

文献标志码: A

文章编号: 1001-5132 (2018) 06-0020-07

台风灾害是我国东南部沿海地区最主要的自然灾害, 通常伴有狂风暴雨, 进而引发洪涝、滑坡、泥石流等次生灾害, 给人民生命财产安全、环境保护和经济发展等带来巨大危害. 台风具有灾害性强、破坏性大和突发性强的特点, 及时掌握台风信息对抗灾救灾至关重要. 由于影响台风风速的因素众多, 传统的依据人工经验预判的风速存在较大的误差, 因此, 如何有效准确地预测台风风速成为众多学者的研究内容.

目前, 国内外对风速预测方法的研究众多, 主要有: 随机时间序列法、卡尔曼滤波法、人工神经网络法和支持向量机法(Support Vector Machine, SVM)等. 随机时间序列法^[1]利用有序的随机数据进行分析和处理, 随着模型阶数不同, 预测结果会有较大差异. 卡尔曼滤波法^[2-3]具有动态修改权值的优点, 但是建立状态方程和预测方程较为困难. 人工神经网络法^[4-5]有很好的鲁棒性和泛化能力,

但是存在学习记忆不稳定、收敛速度慢和容易陷入局部极小化的问题. 支持向量机法^[6-7]可以对小样本数据进行学习, 并获得较好的泛化能力, 引入核函数后, 能有效地解决维数灾难, 然而传统支持向量机无法处理噪声和模糊信息. 因此, 本文引入模糊隶属度函数, 将模糊理论与支持向量机相结合, 构建台风期阵风预测模型, 以期提高支持向量机法处理噪声和模糊信息的能力.

模糊支持向量机(Fuzzy Support Vector Machine, FSVM)的学习性能和泛化能力很大程度上取决于惩罚因子 C 和核参数 g , 因此首先对果蝇优化算法(Fruit Fly Optimization Algorithm, FOA)进行改进, 将其引入到三维空间计算, 并通过一个增强因子 γ 来脱离局部最优解; 其次将改进后的果蝇优化算法应用到模糊支持向量机的参数选择中; 最后建立基于改进型果蝇优化算法的模糊支持向量机(Improved FOA-SVM, IFOA-FSVM)的台风期阵风

收稿日期: 2018-05-22.

宁波大学学报(理工版)网址: <http://journalg.nbu.edu.cn/>

基金项目: 国家自然科学基金(61471212); 浙江省自然科学基金(LY16F010001); 宁波市自然科学基金(2016A610091, 2017A610297).

第一作者: 何彩芬(1974—), 女, 浙江宁海人, 高级工程师, 主要研究方向: 短临、短期天气预报. E-mail: 468176571@qq.com

*通信作者: 金炜(1969—), 男, 浙江兰溪人, 教授, 主要研究方向: 压缩感知、模式识别和数字图像处理. E-mail: xyjw1969@126.com

预测模型, 利用 WRF 数值预报和自动观测站的实况数据进行风速预测, 以此改进数值预报, 并与未改进的 FOA-FSVM 和 FOA-SVM 进行对比. 实验结果表明, 改进后的果蝇优化算法更能提取增强模型预测精度的参数, 且 IFOA-FSVM 对风速预测具有更强的有效性、优势性以及适应性.

1 模糊支持向量机原理

1.1 ε 支持向量回归

在统计学习理论的基础上, 文献[8-9]对 SVM 算法进行了发展. SVM 作为一种以统计学习理论为基础的监督式机器学习方法, 可有效避免人工神经网络中的维数灾难问题, 且具有良好的泛化能力, 在小样本学习问题上有着优异的表现, 其主要算法思想是利用核函数将输入样本映射到更高维度的特征空间中, 在其中建立可以使数据分隔成两个相互平行的距离间隔最大化的超平面, 最终获得输入输出样本之间的非线性关系. Vapnik 等在原有 SVM 的基础上定义了 ε 不敏感函数, 提出了 ε 支持向量回归^[8-9].

在 ε 支持向量回归中, 偏离回归间隔的样本被赋予了相同的惩罚因子, 由于样本偏离回归间隔程度不同, 其损失函数值也不同, 离回归间隔越远的样本对回归函数的影响越大, 造成回归函数对孤立点异常敏感, 过度拟合这些孤立点. 为此, 通过模糊理论, 引入模糊隶属度来提高 SVM 的性能和泛化能力.

1.2 模糊聚类分析

模糊聚类^[10]是指根据客观事物之间的相似性、特征、亲疏程度建立模糊相似关系, 从而对客观事物进行聚类的分析方法. 模糊聚类方法的关键是确定模糊隶属度. 本文选用模糊聚类分析中应用最广泛、研究较成熟的模糊 C 均值聚类算法.

假设样本集为 $\{x_i, i=1, 2, \dots, n\}$, 模糊聚类损失函数的目标函数为:

$$z = \sum_{j=1}^c \sum_{i=1}^n \mu_{ij}^b \|x_i - m_j\|^2, \text{ s.t. } \mu_{ij} \in [0, 1],$$

$$\sum_{i=1}^n \mu_{ij} = 1, i=1, 2, \dots, n; j=1, 2, \dots, c, \quad (1)$$

式中: c 为样本的分类数目; $m_j, j=1, 2, \dots, c$ 为聚类

中心; 采用欧式距离计算样本 x_i 到聚类中心的距离, 即 $\|x_i - m_j\|^2$; μ_{ij} 表示第 i 个样本对第 j 类的隶属度函数.

要使式(1)达到极值, 需要目标函数达到两个必要条件:

$$\mu_{ij} = 1 / \sum_{k=1}^c \left(\frac{\|x_i - m_j\|}{\|x_i - m_k\|} \right)^{\frac{2}{b-1}},$$

$$i=1, 2, \dots, n; j=1, 2, \dots, c, \quad (2)$$

$$m_j = \sum_{i=1}^n \mu_{ij}^b x_i / \sum_{i=1}^n \mu_{ij}^b, j=1, 2, \dots, c. \quad (3)$$

模糊 C 均值聚类算法的步骤为: (1)初始化设置聚类数目 c , 阈值条件 φ , 迭代次数 T ; (2)初始化计算聚类中心; (3)重复以下步骤: (i)根据式(2)和(3)计算 c_i 与 μ_{ij} ; (ii)若 $|z_i - z_{i-1}| \leq \varphi$, 则停止计算.

1.3 模糊支持向量回归

在传统支持向量回归的基础上, 对每一个样本引入 μ_i ($0 < \mu_i \leq 1$) 模糊隶属度, 以表示第 i 个样本的正常程度. 通过引入模糊隶属度得到的样本集为:

$$\{(x_1, y_1, \mu_1), (x_2, y_2, \mu_2), \dots, (x_N, y_N, \mu_N)\} \subseteq$$

$$R^D \times R, 0 < \mu_i < 1,$$

式(3)的目标函数变为:

$$\min_{w, \xi, \xi^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \mu_i (\xi_i + \xi_i^*),$$

$$\text{s.t. } f(x_i) - y_i \leq \xi_i^* + \varepsilon, i=1, 2, \dots, N,$$

$$y_i - f(x_i) \leq \xi_i + \varepsilon, i=1, 2, \dots, N,$$

$$\xi_i, \xi_i^* \geq 0, i=1, 2, \dots, N. \quad (4)$$

引入 Lagrange 函数和核函数后, 将目标函数转化为其对偶形式:

$$\max_{a, a^*} \left(-\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (a_i - a_i^*)(a_j - a_j^*) K(x_i \cdot x_j) + \right.$$

$$\left. \sum_{i=1}^N (a_i - a_i^*) y_i - \sum_{i=1}^N (a_i + a_i^*) \varepsilon \right),$$

$$\text{s.t. } \sum_{i=1}^N (a_i - a_i^*) = 0, 0 \leq a_j, a_j^* \leq \mu_i C,$$

$$i=1, 2, \dots, N. \quad (5)$$

从式(5)约束条件可以看出 Lagrange 系数 a_i 和

a_i^* 的上限为模糊隶属度的函数. 模糊支持向量机最终的回归决策函数形式为:

$$f(x) = \sum_{i=1}^N (a_i - a_i^*) K(x, x_i^*) + b. \quad (6)$$

有研究表明^[11-12], 模糊支持向量机的学习性能和泛化能力很大程度上取决于惩罚因子 C 和核参数 g . 惩罚因子 C 影响模型的泛化能力, 越高说明越不容忍出现误差, 容易过度拟合; 越小则容易欠拟合. 核参数 g 为核函数中的 gamma 函数设置(针对多项式/RBF/sigmoid 核函数), 决定数据映射到新的特征空间后的分布, g 越大支持向量越少; 反之, 支持向量越多. 但目前对于两者的选择并没有有效的理论依据. 果蝇优化算法具有算法简单、参数少、适应性强、不容易陷入局部最优解等特点, 因此将果蝇优化算法引入到了 FSVM 参数的自动寻优中.

1.4 改进型果蝇优化算法

果蝇优化算法是一种基于果蝇觅食行为特征推演出的一种全局寻优的新方法^[13-15]. 果蝇拥有优于其他物种的感官知觉, 尤其在嗅觉和视觉上. 当果蝇在觅食过程中, 首先嗅觉器官会很好地收集空气中食物的气味, 分析味道浓度, 并根据味道浓度的极值方向觅寻食物源. 然后飞近食物源后利用敏锐的视觉定位食物和同伴聚集的位置.

传统果蝇优化算法是在二维空间寻找最优解, 而果蝇觅食过程是在三维空间上, 因此有可能无法搜寻到三维空间中的最优解. 同时在计算味道浓度判定值 S 时, 由于果蝇与原点之间的距离 d 变化范围大且随机取值, 因此味道浓度判定值 S 可能限制在很小的范围内, 这样果蝇优化算法容易生成局部最优解.

为了解决上述两个问题, 把果蝇优化算法引入到三维空间进行计算, 并引入一个增强因子 γ 来脱离局部最优解. 具体算法步骤为:

(1) 随机初始化果蝇群体的位置:

$$(X_{\text{axis}}, Y_{\text{axis}}, Z_{\text{axis}}).$$

(2) 赋值定义果蝇个体搜寻食物的随机方向和距离:

$$X_i = X_{\text{axis}} + R_{\text{Random Value}x},$$

$$Y_i = Y_{\text{axis}} + R_{\text{Random Value}y},$$

$$Z_i = Z_{\text{axis}} + R_{\text{Random Value}z}. \quad (7)$$

(3) 计算味道浓度判定值 S . 由于无法确定食物的具体位置, 因此需要先估计果蝇个体与原点之间的距离 D , 再根据 D 计算味道浓度判定值 S :

$$D_i = \sqrt{X_i^2 + Y_i^2 + Z_i^2}, \quad S_i = 1 / D_i. \quad (8)$$

为了优化味道浓度判定值 S 可能限制在很小范围的问题, 引入一个增强因子 γ , 修改味道浓度判定值 S 为:

$$S_i^* = S_i + \gamma, \quad \gamma = D_i \theta, \quad 0 \leq \theta \leq 1, \quad (9)$$

式中 θ 应为接近 0 的较小的数, 防止当果蝇与原点之间的距离 D 很大时, 过度扩大浓度判定值 S 的分布范围, 从而造成模型过度拟合, 影响算法计算效率.

优化味道浓度判定值 S 后, 对 SVM 的参数惩罚因子 C 和核参数 g 进行赋值:

$$C = 20S_i^*(i, 1), \quad g = S_i^*(i, 2), \quad (10)$$

式中 $S_i^*(i, t)$ 表示 S_i^* 的第 i 行第 t 列的元素.

(4) 计算味道浓度 S_{Smell} . 将味道浓度判定值 S 代入味道浓度判定函数, 求出该果蝇个体位置的味道浓度 S_{Smell} :

$$S_{\text{Smell}_i} = F(S_i). \quad (11)$$

(5) 找出果蝇群体中味道浓度最高的果蝇:

$$[S_{\text{bestSmell}}, I_{\text{bestIndex}}] = \max(S_{\text{Smell}_i}). \quad (12)$$

(6) 保留味道浓度最佳的果蝇个体的坐标值, 同时果蝇群体利用视觉飞往该位置:

$$S_{\text{Smellbest}} = S_{\text{bestSmell}}, \quad (13)$$

$$X_{\text{axis}} = X(I_{\text{bestIndex}}),$$

$$Y_{\text{axis}} = Y(I_{\text{bestIndex}}),$$

$$Z_{\text{axis}} = Z(I_{\text{bestIndex}}). \quad (14)$$

(7) 进入迭代寻优阶段, 重复执行步骤(2)到步骤(5), 并判断味道浓度是否优于前一迭代的味道浓度, 若是, 则用新一迭代最佳味道浓度的果蝇位置代替前一迭代的位置, 直到循环结束.

2 IFOA-FSVM 阵风预测模型

2.1 IFOA-FSVM 阵风预测模型的构建

本文构建了基于 IFOA-FSVM 的台风期阵风预测模型, 如图 1 所示. 首先选取样本集, 确定训练

集和预测集,并进行归一化预处理;其次利用改进后的果蝇优化算法优化惩罚因子 C 和核参数 g ;然后利用优化好的参数构建 FSVM 预测模型;最后用模型对预测集进行测试,并评价分析预测结果。

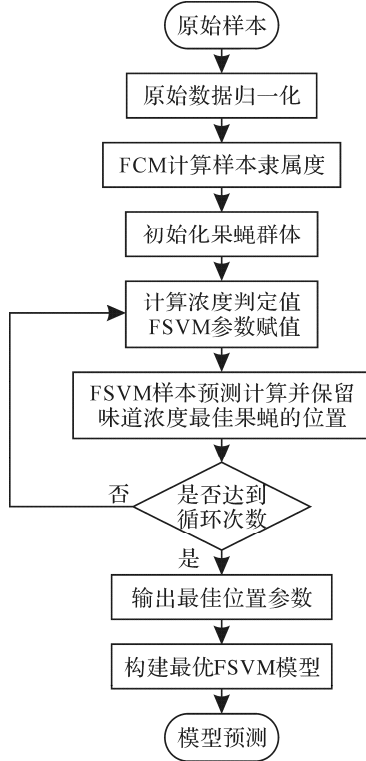


图1 IFOA-FSVM 模型

具体步骤为:

(1)样本数据选取.通过相关性分析选取预报因子并提取数据生成样本集,将样本集分为训练集和预测集并进行归一化预处理.其中,训练集由特征分量和类别标签两部分组成,WRF 模式预报特征因子作为模型的特征分量,类别标签是与特征分量一一对应的风速真值.

(2)计算样本集隶属度.利用模糊 C 均值聚类算法对样本数据进行模糊聚类分析,计算出样本数据的隶属度矩阵和聚类中心,获得每个样本的隶属度,建立新的模糊样本集.

(3)初始化模型参数.设定 FSVM 的核函数和参数、果蝇优化算法的迭代次数 $N_{\max\text{gen}}$ 和种群规模 N_{sizepop} ,算法终止条件 $S_{\text{bestSmell}}$ 和味道浓度判定函数等参数.本文选取的味道浓度判定函数为:

$$\min_{c,g} F = \sqrt{\sum_{i=1}^N (y_i - \tilde{y}_i)^2 / N}, \quad (15)$$

式中: y_i 为第 i 个样本的真实值; $\tilde{y}_i$ 为第 i 个样本的预测值; N 为样本个数.

(4)利用果蝇优化算法迭代优化 FSVM 的惩罚因子 C 和核参数 g .根据式(12)计算出果蝇味道浓度判定值 S_i ,并将其代入式(14)计算出果蝇味道浓度 S_{Smell_i} ,迭代循环,直到寻找出满足终止条件的最佳果蝇味道浓度 $S_{\text{bestSmell}}$,并取得最佳味道浓度值对应的 C 和 g .

(5)建立最优预测模型.利用最优的参数 C 和 g 构建 FSVM 预测模型,对台风期阵风进行预测.

2.2 阵风预测模型评价标准

预测结果的评价指标为平均百分比误差 (M_{MAPE})和均方根误差(R_{RMSE}),其计算公式为:

$$R_{\text{RE}} = \frac{v - \tilde{v}}{v} \times 100\%, \quad (16)$$

$$M_{\text{MAPE}} = \frac{1}{N} \sum_{i=1}^N \frac{|v - \tilde{v}|}{|v|} \times 100\%, \quad (17)$$

$$R_{\text{RMSE}} = \sqrt{\sum_{i=1}^N (R_{\text{RE}} - M_{\text{MAPE}})^2 / N} \times 100\%, \quad (18)$$

其中: v 为实际真实风速; $\tilde{v}$ 为模型预测风速.

3 实验结果与分析

3.1 数据预处理和预测模型参数确定

样本数据来源于 2015 年 7 月到 8 月(台风活跃时期)某避风港锚地的 WRF 模式预报数据和自动站观测数据(采样间隔为 1 h).选取 2015 年 7 月 1 日至 8 月 7 日 912 个样本作为训练集,8 月 8 日至 8 月 11 日 96 个样本作为台风期间预测集.通过相关性分析从 WRF 模式预报数据中选取的预报因子为极大风风向、降雨量、气温、相对湿度、露点温度和气压,自动观测站数据中选取的预报因子为阵风风速.阵风风速又称极大风速,是指某个时间段内出现的最大瞬时风速值.瞬时风速是指 3 s 的平均风速.本文选取径向基核函数(RBF)作为阵风预测模型的核函数.

在阵风预测模型中,惩罚因子 C 和核参数 g 对预测性能有着直接的影响.利用 FOA 对预测模型中的参数 C 和 g 进行迭代优化,设置果蝇优化算法的最大迭代次数 $N_{\max\text{gen}}$ 为 200,种群规模 N_{sizepop} 为 20,惩罚系数 C 的取值范围为[1,100],核

参数 g 的取值范围为 $[1, 100]$, θ 为 0.001. 经过 IFOA 优化后的惩罚因子 C 为 5.009 5, 核参数 g 为 0.062 2, 代入阵风预测模型对样本集进行训练与预测.

3.2 阵风预测模型结果分析与评价

台风期间的阵风预测结果如图 2 所示. 从图 2 可以看出, 与 WRF 模式预报中的风速相比, 本文模型具有很好的预测效果, 预测结果与真实风速误差较小, 能够改进风速预报产品. 计算预测风速与真实风速之间的皮尔森相关系数可知, 两者的相关性可达 99%, 所选的预测因子可以很好地刻画风速.

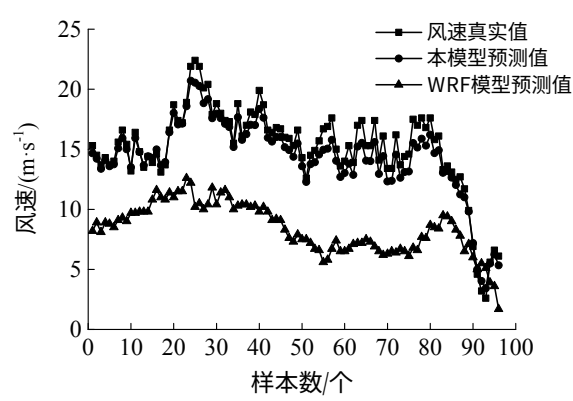


图2 IFOA-FSVM 模型预测结果比较

将模型预测的台风期阵风预测结果与真实风速的相对误差对比, 结果见表 1. 分析 96 个预测点的相对误差可知, 相对误差超过 15% 的预测点仅为 2.08%, 平均百分比误差为 5.99%, 说明该模型具有较高的精度和良好的推广价值.

表 1 台风期阵风预测结果

t/h	实际风 速/($m \cdot s^{-1}$)	预测风 速/($m \cdot s^{-1}$)	相对 误差/%	t/h	实际风 速/($m \cdot s^{-1}$)	预测风 速/($m \cdot s^{-1}$)	相对 误差/%
1	15.3	14.66	-4.17	11	16.4	15.93	-2.85
2	14.5	14.22	-1.92	12	14.8	14.77	-0.19
3	13.7	13.37	-2.42	13	13.5	13.67	1.23
4	14.3	14.08	-1.53	14	14.4	14.42	0.11
5	13.7	13.59	-0.81	15	13.9	14.28	2.74
6	14.0	13.75	-1.80	16	14.6	14.98	2.62
7	15.6	15.07	-3.38	17	13.1	13.67	4.34
8	16.6	15.93	-4.07	18	13.7	13.93	1.65
9	15.4	15.00	-2.59	19	16.6	16.41	-1.14
10	13.2	13.51	2.33	20	18.7	18.02	-3.62

续表

t/h	实际风 速/($m \cdot s^{-1}$)	预测风 速/($m \cdot s^{-1}$)	相对 误差/%	t/h	实际风 速/($m \cdot s^{-1}$)	预测风 速/($m \cdot s^{-1}$)	相对 误差/%
21	17.4	17.06	-1.95	59	13.6	12.69	-6.68
22	17.2	17.13	-0.43	60	14.0	13.02	-7.00
23	18.9	18.59	-1.65	61	15.3	13.80	-9.81
24	21.9	20.71	-5.45	62	13.9	12.87	-7.39
25	22.4	20.55	-8.28	63	17.0	15.23	-10.41
26	21.9	20.27	-7.46	64	17.4	15.52	-10.79
27	20.1	18.84	-6.26	65	15.3	14.05	-8.16
28	20.4	19.17	-6.02	66	15.3	14.02	-8.39
29	17.9	17.59	-1.75	67	17.4	15.60	-10.34
30	18.8	17.95	-4.50	68	14.1	12.95	-8.14
31	18.0	17.58	-2.36	69	16.1	14.43	-10.38
32	17.4	17.12	-1.61	70	13.4	12.33	-8.01
33	17.3	16.85	-2.62	71	13.4	12.37	-7.71
34	15.5	15.19	-2.02	72	16.2	14.55	-10.21
35	18.8	17.66	-6.05	73	13.7	12.62	-7.85
36	16.1	15.78	-2.01	74	14.4	13.11	-8.97
37	17.0	16.26	-4.36	75	14.6	13.14	-9.97
38	18.1	17.03	-5.90	76	17.5	15.52	-11.34
39	17.9	17.01	-4.95	77	17.0	15.15	-10.88
40	19.9	18.35	-7.78	78	17.6	15.88	-9.80
41	18.7	17.63	-5.73	79	16.8	15.30	-8.94
42	16.6	15.96	-3.87	80	17.6	16.20	-7.93
43	16.4	15.64	-4.62	81	15.8	14.68	-7.07
44	16.8	16.02	-4.67	82	16.1	14.93	-7.30
45	16.7	16.02	-4.06	83	13.3	13.04	-1.97
46	16.0	15.17	-5.17	84	13.6	13.18	-3.12
47	15.9	14.93	-6.08	85	13.1	12.65	-3.47
48	15.3	14.37	-6.06	86	12.5	12.03	-3.77
49	16.6	15.48	-6.76	87	12.7	11.24	-11.49
50	14.3	13.58	-5.01	88	11.7	11.00	-6.00
51	12.6	12.28	-2.53	89	9.9	9.81	-0.87
52	14.5	13.76	-5.08	90	6.9	7.18	4.05
53	14.9	13.92	-6.57	91	4.6	4.98	8.29
54	15.7	14.56	-7.29	92	3.2	4.03	25.84
55	16.7	14.95	-10.48	93	2.6	3.42	31.71
56	16.9	15.04	-10.98	94	5.6	5.49	-2.02
57	17.6	15.78	-10.36	95	6.6	6.23	-5.61
58	15.0	14.04	-6.43	96	6.1	5.34	-12.44

为评估本文提出的模型的合理性和优势,将 IFOA-FSVM 模型风速预测结果与未改进的 FOA-FSVM 和 FOA-SVM 的预测结果进行对比。3 种算法的种群规模、迭代次数以及参数寻优范围等参数的设置基本相同,具体预测结果对比如图 3 和表 2 所示。

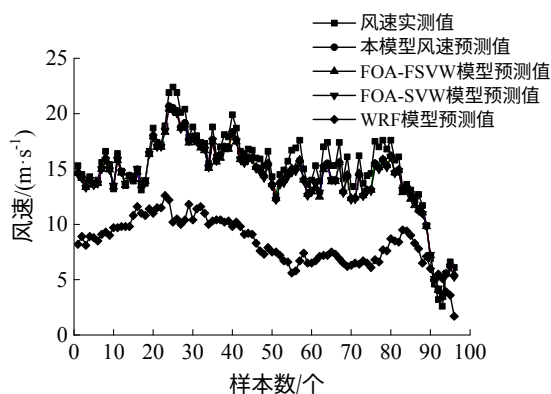


图3 3种模型预测结果比较

表2 3种模型误差对比

误差	预测模型		
	IFOA-SVM	FOA-FSVM	FOA-SVM
$M_{MAPE}/\%$	5.99	6.70	9.52
$R_{RMSE}/\%$	4.61	5.50	6.91

从图3可知,与WRF模式预报的风速值相比,3种模型对台风期阵风的预测都比较有效,其中 IFOA-FSVM 模型预测结果更接近真实风速。另外,从表2可知,IFOA-FSVM 模型平均百分比误差和均方根误差均小于 FOA-FSVM 和 FOA-SVM 模型,说明改进后的果蝇优化算法对 FSVM 参数优化选择的有效性。综上所述,IFOA-FSVM 模型具有较强的泛化能力,预测精度更高,可以更好地预测台风期阵风风速。

4 结论

通过利用本文提出的 IFOA-FSVM 模型对台风期阵风预测,并进行结果分析,可以得到以下结论:

(1)为了解决模糊支持向量回归模型中惩罚因子 C 和核参数 g 的选择问题,选择全局寻优、计算简单、精度高的果蝇优化算法对所需参数进行优化寻优,并将果蝇优化算法引入到三维空间计算,同时引入一个增强因子 γ 来脱离局部最优解以改进果蝇优化算法。实验结果表明,改进后的果蝇优化

算法能够更好地克服模糊支持向量回归模型参数选择的盲目性,可以自动、快速地提取收敛速度快和识别率高的参数。

(2)将本文提出的 IFOA-FSVM 台风期阵风预测模型与未改进的 FOA-FSVM 和 FOA-SVM 预测模型结果进行对比可知,IFOA-FSVM 模型可行,且具有更高的预测精度,能较好地解决传统人工经验预判风速存在较大误差的问题。另外,预测风速与真实风速之间的相关性达到 99%,说明所选的预测因子可以很好地刻画风速,能够提高 WRF 模式风速的适用性。

IFOA-FSVM 模型的优越性为进一步提高台风期阵风预测精度提供了新思路,也为其他领域预测提供理论参考。

参考文献:

- [1] Kerem A, Kirbas I, Saygin A. Performance analysis of time series forecasting models for short term wind speed prediction[EB/OL]. [2018-01-10]. https://www.researchgate.net/publication/312891204_Performance_Analysis_of_Time_Series_Forecasting_Models_for_Short_Term_Wind_Speed_Prediction.
- [2] Gauterin E, Kammerer P, Kühn M, et al. Effective wind speed estimation: Comparison between Kalman filter and Takagi-Sugeno observer techniques[J]. Isa Transactions, 2016, 62:60-72.
- [3] Galanis G, Papageorgiou E, Liakatas A. A hybrid Bayesian Kalman filter and applications to numerical wind speed modeling[J]. Journal of Wind Engineering & Industrial Aerodynamics, 2017, 167:1-22.
- [4] Finamore A R, Galdi V, Calderaro V, et al. Artificial neural network application in wind forecasting: A one-hour-ahead wind speed prediction[EB/OL]. [2018-01-10]. <https://ieeexplore.ieee.org/document/7835492>.
- [5] 解坤, 张俊芳. 基于 KMO-Bartlett 典型风速选取的 PCA-WNN 短期风速预测[J]. 发电设备, 2017, 31(2): 86-91.
- [6] Shao H, Wei H, Deng X, et al. Short-term wind speed forecasting using wavelet transformation and AdaBoosting neural networks in Yunnan wind farm[J]. IET Renewable Power Generation, 2017, 11(4):374-381.
- [7] 颜晓娟, 龚仁喜, 张千锋. 优化遗传算法寻优的 SVM 在短期风速预测中的应用[J]. 电力系统保护与控制, 2016, 44(9):38-42.
- [8] Cristianini N, Shawe-Taylor J. 支持向量机导论[M]. 李

- 国正, 译. 北京: 电子工业出版社, 2004:82-105.
- [9] 丁世飞, 齐丙娟, 谭红艳, 等. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1):2-10.
- [10] 刘畅, 孙德山. 模糊支持向量机隶属度的确定方法[J]. 计算机工程与应用, 2008, 44(11):41-42.
- [11] 王正帅, 邓喀中, 谭志祥. 导水裂缝带高度预测的模糊支持向量机模型[J]. 地下空间与工程学报, 2011, 7(4): 723-727.
- [12] 王欣, 李胜刚, 秦斌, 等. 基于模糊支持向量机的风电场功率预测[J]. 新型工业化, 2014(9):50-55.
- [13] Pan W T. A new fruit fly optimization algorithm: Taking the financial distress model as an example[J]. Knowledge-Based Systems, 2012, 26(2):69-74.
- [14] Zheng X, Wang L, Zheng H. A knowledge-based fruit fly optimization algorithm for multi-skill resource constrained project scheduling problem[J]. Soft Computing, 2017, 21(6):1537-1548.
- [15] Seghir F, Khababa A. A hybrid approach using genetic and fruit fly optimization algorithms for QoS-aware cloud service composition[J]. Journal of Intelligent Manufacturing, 2017(6):1-20.
- [13] Pan W T. A new fruit fly optimization algorithm: Taking

Precision forecasting of typhoon wind speed in WRF model based on IFOA-FSVM

HE Cai-fen¹, QIAN Bin-kai², JIN Wei^{2*}, ZHANG Guo-chao¹

(1.Zhenhai Observatory, Ningbo 315202, China;

2.Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China)

Abstract: For the sake of improving the accuracy of forecasting wind speed during typhoon strike in WRF model forecast, a new method for precision forecasting of typhoon wind speed is proposed by combining the data collected from the WRF model forecast and an automatic observation station. The method incorporates many factors influencing the typhoon wind speed. The wind speed which is obtained using the traditional human prediction produces large error when compared with actual wind speed. To address this issue, a fuzzy support vector regression model for wind forecasting is built. Considering the fact that the fuzzy support vector regression model is not adequately efficient in determining the punishment factor C and kernel parameter g , the fly optimization algorithm is introduced into optimizing the parameters of the fuzzy support vector machine. According to the characteristics of the wind speed regression, the fruit fly optimization algorithm is developed in three dimensional space, combining with the enhancement factor γ for improving the global optimization ability of traditional fruit fly optimization algorithm. The results show that the forecasting wind speed and the actual one is in good agreement with each other, and the correlation is as high as 99%. The presented method of wind speed prediction provides higher accuracy than that of traditional FOA-FSVM model and FOA-SVM model.

Key words: fruit fly optimization algorithm; fuzzy support vector machine; wind speed forecasting; typhoon

(责任编辑 史小丽)