

# 基于改进萤火虫寻优支持向量机的 PM2.5 预测<sup>①</sup>



范文婷, 王 晓

(太原科技大学 计算机科学与技术学院, 太原 030024)

通讯作者: 范文婷, E-mail: 674574677@qq.com

**摘 要:** 针对现有 PM2.5 浓度预测误差较大的问题, 提出一种基于改进萤火虫寻优支持向量机的预测模型 (IFA-SVM). 该模型引入邻域搜索和可变步长策略改进萤火虫算法, 利用改进 FA 对 SVM 的参数  $C$ 、 $\varepsilon$  和  $\gamma$  寻优, 用最优参数 SVM 模型预测太原市 PM2.5 值. 其中邻域搜索策略能为参数优化提供更多更精确的候选解; 可变步长可动态调整算法搜索步长, 加速收敛, 平衡 FA 的全局和局部搜索能力. 将 IFA-SVM 预测值与萤火虫算法—支持向量机 (FA-SVM)、遗传算法—支持向量机 (GA-SVM)、粒子群算法—支持向量机 (PSO-SVM) 相比较. 结果表明较其他方法, IFA-SVM 模型对太原市未来一天和三天的 PM2.5 值都取得了更精确的预测性能.

**关键词:** 萤火虫算法; 支持向量机; 邻域搜索策略; 可变步长; 参数优化; PM2.5 预测

引用格式: 范文婷, 王晓. 基于改进萤火虫寻优支持向量机的 PM2.5 预测. 计算机系统应用, 2019, 28(1): 134-139. <http://www.c-s-a.org.cn/1003-3254/6718.html>

## PM2.5 Forecasting Based on Improved Firefly Optimization SVM

FAN Wen-Ting, WANG Xiao

(School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

**Abstract:** Aiming at the problem of large deviation in existing PM2.5 concentration prediction, a novel model based on Improved Firefly Algorithm optimization SVM (IFA-SVM) was proposed. In this model, two neighborhood search strategies and variable step size mechanism were employed to improve FA. The IFA was applied to optimize the SVM parameters ( $C$ ,  $\varepsilon$ , and  $\gamma$ ), and an outstanding model was constructed to forecast PM2.5 concentrations in Taiyuan. The neighborhood search strategies can provide better candidate solutions; search step size was dynamically tuned by using variable step size strategy to accelerate convergence and obtain a trade-off between exploration and exploitation. The performance of the proposed IFA-SVM model has been compared with FA-SVM, Genetic Algorithm (GA)-SVM, and Particle Swarm Optimization (PSO)-SVM. Experimental results show that the proposed IFA-SVM model has achieved more accurate performance for PM2.5 forecasts in 1 day ahead and 3 days ahead compared to other method.

**Key words:** Firefly Algorithm (FA); Support Vector Machine (SVM); neighborhood search strategies; variable step size; parameter optimization; PM2.5 forecasting

作为雾霾的主要元凶之一, PM2.5 对人类生活、身心健康、经济发展等都造成了严重影响<sup>[1,2]</sup>. 如何高效、准确地预测 PM2.5 值, 进而制定科学合理的雾霾防控方案具有重要意义.

目前, 常用的 PM2.5 预测方法包括回归模型<sup>[3-5]</sup>、神经网络<sup>[6-8]</sup>、支持向量机<sup>[9,10]</sup>等. 其中, 支持向量机 (Support Vector Machine, SVM) 以其强大的学习泛化能力, 结构风险小的特点, 在小样本、非线性预测领

① 基金项目: 太原科技大学博士基金 (20152044)

Foundation item: Doctorate Fund of Taiyuan University of Science and Technology (20152044)

收稿时间: 2018-05-01; 修改时间: 2018-05-24, 2018-07-25; 采用时间: 2018-08-08; csa 在线出版时间: 2018-12-26

域取得了良好的应用效果. 在实际应用中, SVM 参数选择直接影响预测性能, 找到一个合适的算法对 SVM 参数寻优至关重要. 常用 SVM 参数优化方法包括: 遗传算法 (Genetic Algorithm, GA)、粒子群算法 (Particle Swarm Optimization, PSO)、萤火虫算法 (Firefly algorithm, FA) 等. 如朱霄珣等<sup>[11]</sup>使用 GA 优化 SVM 参数, 建立风速预测模型. 戴李杰等<sup>[12]</sup>将 GA 的变异因子引入到 PSO 中对 SVM 参数寻优建立未来 24 小时 PM2.5 滚动预报模型. 杨孟英<sup>[13]</sup>使用 FA 优化 SVM 参数, 进行中文文本分类. 相比于 GA 和 PSO, FA 结构简单、调节参数少, 易于计算, 能更好地平衡全局和局部性能<sup>[14-16]</sup>, 但标准 FA 也存在缺陷, 如迭代后期收敛速度慢, 固定步长造成振荡现象等<sup>[17-19]</sup>. 为此, 本文对标准 FA 进行改进, 利用改进 FA 对 SVM 参数寻优, 构建 IFA-SVM 的 PM2.5 预测模型, 用此模型预测太原市未来一天和三天的 PM2.5 含量, 分析该模型用于 PM2.5 浓度预测的可行性和高效性.

## 1 PM2.5 预测原理

PM2.5 预测是依据气象条件、污染源等因素和历史数据建立数学模型来预测未来 PM2.5 值, 表示为如下非线性关系:

$$y = f(x_1, x_2, \dots, x_n) \quad (1)$$

其中,  $\{x_1, x_2, \dots, x_n\}$  为影响因子,  $n$  为影响因子个数,  $f()$  为预测模型,  $y$  为 PM2.5 预测值.

由于 PM2.5 是气象条件、污染源、地理空间等多种因素共同作用的结果<sup>[20]</sup>, 呈现高度非线性, 本文使用 SVM 来预测 PM2.5 值.

## 2 算法理论基础

### 2.1 SVM 基本原理

SVM 基本思想是针对 PM2.5 非线性样本, 引入径向基 (Radial Basis Function, RBF) 核函数, 将样本映射到高维空间, 在高维空间求解超平面, 使得两类样本的间隔最大<sup>[21]</sup>, 即求解如下约束优化问题:

$$\min_{w, b, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^i \varepsilon_i \quad (2)$$

其中,  $C$  是惩罚参数, 它控制对错分样本的惩罚程度. RBF 核函数表达式为:

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (3)$$

因此影响 SVM 性能的主要参数即为  $C$ 、 $\varepsilon$ 、 $\gamma$ , 本文用改进萤火虫算法来优化 SVM 参数.

### 2.2 标准 FA 基本原理

FA 受自然界萤火虫生物特性启发<sup>[22]</sup>, 其对 SVM 参数优化思想为:

将所有 SVM 可选参数模拟为萤火虫个体, 将优化过程看作对最优个体的寻找, 即通过个体的吸引和位置移动实现目标参数寻优, 在迭代过程中用好的可行解淘汰较差可行解, 直到得到最优解, 即亮度最高的萤火虫的位置即为最优参数.

算法涉及两个因素, 萤火虫的发光亮度  $I$  和相对吸引度  $\beta$ :

$$I = I_0 \times e^{-\gamma r_{ij}} \quad (4)$$

$$\beta = \beta_0 \times e^{-\gamma r_{ij}^2} \quad (5)$$

其中,  $I_0$  为萤火虫的最大发光亮度, 即  $r=0$  处的亮度;  $\beta_0$  为最大吸引度, 即  $r=0$  处的吸引度;  $\gamma$  为光强吸收系数;  $r_{ij}$  为萤火虫  $i$  和  $j$  之间的欧式距离.

寻优过程中萤火虫  $i$  被吸引向萤火虫  $j$  移动的位置更新公式为:

$$x_i = x_i + \beta \times (x_j - x_i) + \alpha \times \left( rand - \frac{1}{2} \right) \quad (6)$$

其中,  $x_i, x_j$  为萤火虫  $i$  和  $j$  所处的位置,  $\alpha$  为步长因子, 是  $[0, 1]$  上的常数,  $rand$  为  $[0, 1]$  上服从均匀分布的随机因子.

FA 主要步骤包括:

- (1) 根据目标函数计算萤火虫的亮度;
- (2) 亮度较暗的萤火虫按公式 (6) 向较亮的萤火虫移动;
- (3) 对萤火虫按亮度值从大到小排序, 找到最亮的萤火虫;
- (4) 重复迭代, 直到达到最大迭代次数.

### 2.3 基于邻域搜索策略的改进 FA

标准 FA 迭代中亮度较暗的萤火虫按公式 (6) 向较亮的萤火虫移动, 当迭代后期亮度相同时, 萤火虫将随机运动, 此时收敛速度下降且很难找到最优解. 同时若萤火虫个体搜索半径内没有更亮个体, 萤火虫也将随机运动, 此时最优亮度为搜索半径内局部最优而非全局最优, 算法搜索能力下降, 精度较差. 根据文献<sup>[19]</sup>, 设计两种邻域搜索策略为参数寻优提供更多更精确的候选解, 帮助萤火虫跳出局部最优并加速收敛, 具体邻域搜索策略为:

$$x_i^1 = \lambda_1 \times x_i + \lambda_2 \times pbest + \lambda_3 \times (x_{i1} - x_{i2}) \quad (7)$$

其中,  $x_{i1}$ 、 $x_{i2}$ 是 $x_i \pm 2$ 阶半径内随机选取的两个萤火虫,  $pbest$ 是第 $t$ 次迭代中最大的目标函数值,  $\lambda_1$ 、 $\lambda_2$ 、 $\lambda_3$ 是 $[0, 1]$ 上的常数,  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ , 此邻域搜索策略在当前萤火虫 $x_i \pm 2$ 阶半径内寻找候选解。

$$x_i^2 = \lambda_4 \times x_i + \lambda_5 \times gbest + \lambda_6 \times (x_{i3} - x_{i4}) \quad (8)$$

其中,  $x_{i3}$ 、 $x_{i4}$ 是从整个种群中随机选取的两个萤火虫,  $gbest$ 是目前为止全局最大的目标函数值,  $\lambda_4$ 、 $\lambda_5$ 、 $\lambda_6$ 是 $[0, 1]$ 上的常数,  $\lambda_4 + \lambda_5 + \lambda_6 = 1$ , 此邻域搜索策略在当前迭代次数全局范围内选择候选解。

如果迭代过程中, 第 $t$ 次迭代萤火虫亮度值与 $t-1$ 次亮度值相等, 执行上述两种策略, 找到其它两个候选解 $x_i^1$ 、 $x_i^2$ , 然后从 $x_i$ 、 $x_i^1$ 、 $x_i^2$ 中选择最亮的作为 $x_i$ 。显然邻域搜索策略可以提供更多更精确的候选解, 提高萤火虫算法的搜索能力并加速收敛。

### 2.4 基于可变步长的改进 FA

标准 FA 中, 步长因子 $\alpha$ 是固定的, 固定步长使得算法早期萤火虫个体不能以最快速度靠近亮度高的伙伴; 而后期萤火虫间亮度差值减小, 位置比较靠近, 固定步长又使得萤火虫个体始终无法向最亮的那只进一步移动, 产生振荡, 算法精度较差。根据文献[20], 采用可变步长因子 $\alpha$ , 在迭代初期, 采用较大 $\alpha$ 探索新的搜索空间, 提高全局搜索能力且加速收敛; 而在后期用较小 $\alpha$ 进行局部开采, 避免振荡。可变 $\alpha$ 可动态调整算法搜索步长来平衡 FA 的全局搜索和局部开采能力, 具体 $\alpha$ 的调整公式为:

$$\alpha = 0.4 / (1 + \exp(0.015 \times (t - \max G) / 3)) \quad (9)$$

其中,  $t$ 为当前迭代次数,  $\max G$ 为最大迭代次数。

## 3 萤火虫寻优支持向量机 (FA-SVM)PM2.5 预测模型

### 3.1 IFA-SVM PM2.5 预测原理

IFA-SVM PM2.5 预测过程如下:

(1) 收集太原市 PM2.5 浓度实验数据, 划分训练集和测试集, 并归一化预处理;

(2) IFA-SVM 参数迭代寻优

1) 初始化算法各基本参数, 随机分布萤火虫;

2) 计算萤火虫的目标函数值, 以 SVM 对训练集的 PM2.5 预测性能作为目标函数值;

3) 对萤火虫的目标函数值进行亮度排序, 找到当前最优的目标函数值及其对应萤火虫, 并根据公式 (6) 更新萤火虫;

4) 如果迭代过程中, 第 $t$ 次迭代目标函数值与 $t-1$ 迭代目标函数值相等, 执行 2.3 节中两种邻域搜索策略;

5) 若达到最大迭代次数, 或满足停止迭代的条件, 则转至步骤 6), 否则转至步骤 2) 继续迭代;

6) 输出最大目标函数值及其对应的萤火虫, 即得到最优参数。

(3) 使用最优参数预测测试集 PM2.5 值, 并将预测结果反归一化, 得到实际 PM2.5 预测值, 输出结果。

具体流程如图 1 所示。

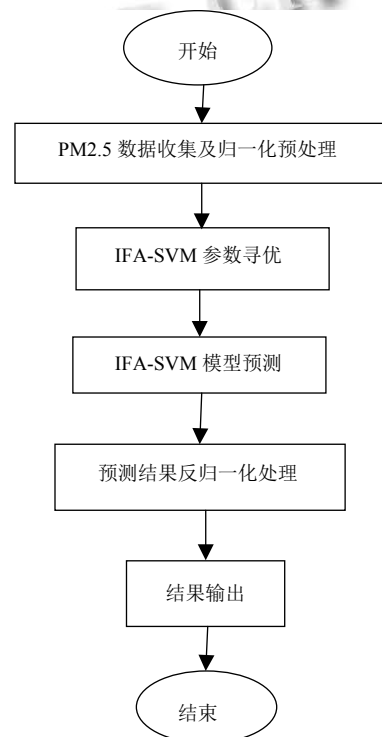


图 1 IFA-SVM PM2.5 预测流程

### 3.2 数据收集及预处理

统计发现, 太原市秋冬季节出现雾霾频率更高, 综合考虑气象条件和污染源, 收集太原市 2015.10~2016.3、2016.10~2017.3 和 2017.11 的湿度、露点、气压、风速、 $O_3$ 、CO、 $NO_2$ 、PM10 共 8 项 PM2.5 影响因子进行建模。具体如表 1 所示, 其中, 2015.10~2016.3、2016.10~2017.3 作为训练样本, 2017.11 作为测试样本。

为了消除各影响因子量纲不同对预测速度和精度的影响, 建模之前进行归一化预处理, 归一化公式为:

$$x'_i = (x_i - x_{\min}) / (x_{\max} - x_{\min}) \quad (10)$$

其中,  $x_i$ 为影响因子初始值,  $x'_i$ 为归一化后的值,  $x_{\max}$ 和  $x_{\min}$ 表示每个因子的最大和最小值。

表1 实验数据

日期	湿度 (°C)	露点 (°C)	气压 (hPa)	风速 (km/h)	O <sup>3</sup>	CO	NO <sub>2</sub>	PM10	PM2.5
2015-10-1	38	-1	1021.62	23	56	0.6	15	36	13
2015-10-2	54	2	1017.8	6	66	1	41	99	40
2015-10-3	59	4	1023	4	70	1.4	53	148	68
...	...	...	...	...	...	...	...	...	...
2016-3-1	29	-14	1022.95	6	59	2.3	49	111	60
2016-3-2	32	-11	1019.23	7	72	2.2	46	107	57
...	...	...	...	...	...	...	...	...	...
2017-11-29	41	-12	1027.6	6	23	1.4	70	180	68
2017-11-30	44	-14	1029.54	6	41	1.3	57	168	63

注: O<sup>3</sup>、CO、NO<sub>2</sub>、PM10、PM2.5 的单位均为 μg/m<sup>3</sup>.

### 3.3 评价标准

采用平均绝对误差 (MAE)、均方根误差 (RMSE) 对预测结果进行评价, 计算公式如下:

$$MAE = \frac{1}{N} \sum_{i=1}^N |O_i - P_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2} \quad (12)$$

其中,  $O_i$  为观察值, 即实际值;  $P_i$  为预测值.

### 3.4 参数优化

为了验证 IFA 的性能, 选取 Ackley 函数、Sphere 函数进行仿真测试. 种群数量设置为:  $N=50$ ,  $maxG=200$ ,  $\gamma=0.01$ , 初始步长  $\alpha=0.2$ . 每个函数独立运行 30 次, 得到 FA、IFA 的仿真曲线, 如图 2 与图 3 所示.

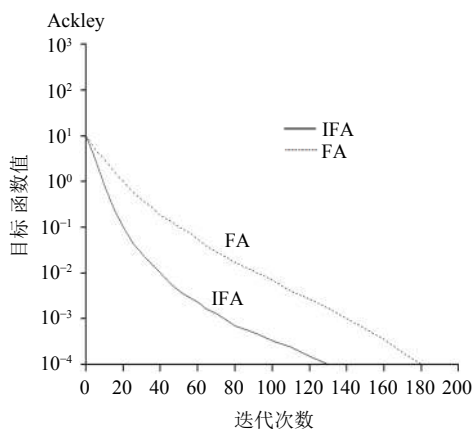


图2 算法对 Ackley 函数的收敛曲线对比

从图 2 与图 3 看到, 针对函数 Ackley 与 Sphere, IFA 收敛速度更快, 精度更高.

为了比较性能, 分别用标准 FA、GA、PSO、

IFA 选取 SVM 参数, 各优化算法得到的 SVM 参数如表 2 所示.

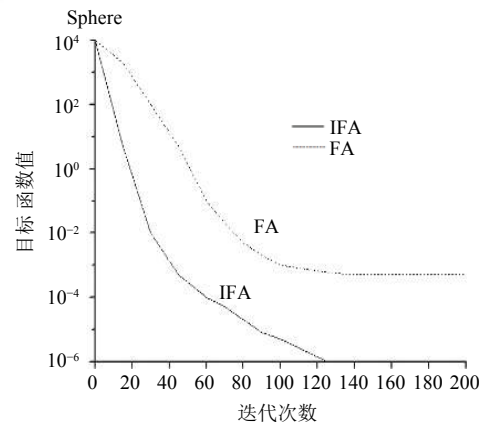


图3 算法对 Sphere 函数的收敛曲线对比

表2 优化得到的各 SVM 模型参数

模型	参数 $C, \epsilon, \gamma$
IFA-SVM	$C=2.11, \epsilon=0.54, \gamma=0.75$
FA-SVM	$C=1.95, \epsilon=0.68, \gamma=0.87$
PSO-SVM	$C=1.44, \epsilon=0.37, \gamma=0.56$
GA-SVM	$C=1.59, \epsilon=0.43, \gamma=0.44$

注: O<sup>3</sup>、CO、NO<sub>2</sub>、PM10、PM2.5 的单位均为 μg/m<sup>3</sup>.

### 3.5 实验结果与分析

#### 3.5.1 IFA-SVM 实验结果与分析

利用最优 IFA-SVM 参数预测 2017 年 11 月太原市两种 PM2.5 值:

- (1) 预测未来一天 PM2.5 浓度值;
- (2) 预测未来第三天 PM2.5 浓度值.

实验得到最终预测值, 将实际值与 IFA-SVM 的预测值进行对比, 结果如图 4 所示.

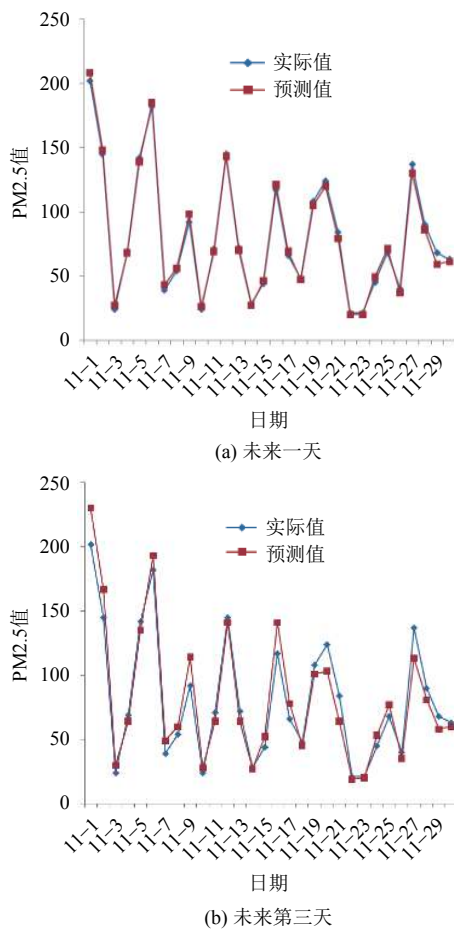


图4 IFA-SVM PM2.5 预测结果对比图

图4的结果显示,预测太原市2017年11月的PM2.5浓度,得到未来一天的预测结果与实际非常接近,而未来第三天稍有偏差,但整体变化趋势相同,预测值与实际值的曲线拟合度很高.此结果表明 IFA-SVM 模型是有效的,该模型可以帮助预测太原市 PM2.5 值.而预测未来第三天的精度比未来一天的精度低,这主要是由于长期预测会对短期的预测错误进行积累.

### 3.5.2 实验结果比较与分析

分别用模型 GA-SVM、PSO-SVM、FA-SVM 和 IFA-SVM 对 PM2.5 浓度进行预测,预测误差结果如表3.

由表3可知,使用四个模型对太原市2017年11月的PM2.5未来一天和第三天浓度预报后,得到的预测误差从低到高依次是 IFA-SVM、FA-SVM、PSO-SVM.即与模型 GA-SVM、PSO-SVM 相比,FA-SVM 的预测结果更准确,偏离实际值较小,而与 FA-SVM 相比,IFA-SVM 误差更小,即 IFA-SVM 模型的 MAE 和

RMSE 都是4个预测模型中最低的,其中性能最好的 IFA-SVM 预测未来一天的 PM2.5 浓度值,MAE 和 RMSE 分别为 3.85 和 4.07.

表3 IFA-SVM、FA-SVM、GA-SVM 和 PSO-SVM 预测结果性能比较

模型		时间	
		一天	三天
IFA-SVM	MAE	3.85	9.84
	RMSE	4.07	12.09
FA-SVM	MAE	4.0	10.2
	RMSE	4.68	12.77
GA-SVM	MAE	7.02	11.58
	RMSE	7.79	15.35
PSO-SVM	MAE	6.1	11.12
	RMSE	6.72	14.49

综合考虑 3.4 节迭代曲线和 3.5 预测误差结果,可以看到与标准 FA 相比,IFA 收敛速度更快,精度更高,同时 IFA-SVM PM2.5 预测值也最接近实际值.

## 4 结论与展望

针对 PM2.5 与影响因子间的高度非线性关系,提出一种基于改进萤火虫寻优支持向量机的 PM2.5 预测模型 IFA-SVM,并以太原市未来一天和第三天的 PM2.5 值作为实验样本进行模型测试,并对比 GA-SVM、PSO-SVM、FA-SVM 的预测效果,经实验得到如下结论:

(1) IFA-SVM 模型对未来一天和三天的 PM2.5 值都可以有效预测,由于预测误差会不断积累,一天的预测精度更高.

(2) FA 能够跳出局部最优且计算简单,FA-SVM 模型比 GA-SVM 和 PSO-SVM 方法预测更准确.

(3) 引入邻域搜索和可变步长策略改进 FA,可加速算法收敛,平衡局部和全局性能,使得 IFA-SVM 模型预测结果更接近实际的 PM2.5 变化趋势,为雾霾预测提供了一种新思路.

由于 PM2.5 浓度值不仅与本文所选因子有关,还受到城市资源能耗、道路布局、空间位置等多种因素影响.在今后的研究工作中,尝试将这些信息和技术应用进来,进一步提高 PM2.5 预测性能.

### 参考文献

- 孙志豪,崔燕平. PM<sub>2.5</sub> 对人体健康影响研究概述. 环境科技, 2013, 26(4): 76-78.
- 周艳军. PM<sub>2.5</sub> 的污染危害及预防对策. 产业与科技论坛,

- 2013, 12(11): 114–115. [doi: 10.3969/j.issn.1673-5641.2013.11.064]
- 3 Jian L, Zhao Y, Zhu YP, *et al.* An application of ARIMA model to predict submicron particle concentrations from meteorological factors at a busy roadside in Hangzhou, China. *Science of the Total Environment*, 2012, 426: 336–345. [doi: 10.1016/j.scitotenv.2012.03.025]
- 4 郎艺超, 肖璐, Christakos G. 基于 SARIMA 模型和普通 Kriging 法对杭州市主城区 PM<sub>2.5</sub> 短期预测和制图. *环境科学学报*, 2018, 38(1): 62–70.
- 5 程兴宏, 刁志刚, 胡江凯, 等. 基于 CMAQ 模式和自适应偏最小二乘回归法的中国地区 PM<sub>2.5</sub> 浓度动力-统计预报方法研究. *环境科学学报*, 2016, 36(8): 2771–2782.
- 6 Pérez P, Trier A, Reyes J. Prediction of PM<sub>2.5</sub> concentrations several hours in advance using neural networks in Santiago, Chile. *Atmosphere Environment*, 2000, 34(8): 1189–1196. [doi: 10.1016/S1352-2310(99)00316-7]
- 7 Feng X, Li Q, Zhu YJ, *et al.* Artificial neural networks forecasting of PM<sub>2.5</sub> pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, 2015, 107: 118–128. [doi: 10.1016/j.atmosenv.2015.02.030]
- 8 艾洪福, 石莹. 基于 BP 人工神经网络的雾霾天气预测研究. *计算机仿真*, 2015, 32(1): 402–405, 415. [doi: 10.3969/j.issn.1006-9348.2015.01.085]
- 9 陈俏, 曹根牛, 陈柳. 支持向量机应用于大气污染物浓度预测. *计算机技术与发展*, 2010, 20(1): 250–252. [doi: 10.3969/j.issn.1673-629X.2010.01.065]
- 10 Wu ZH. SVR smoggy forecast model based on data mining. *International Journal of Science*, 2015, 2(9): 120–125.
- 11 朱霄珣, 徐搏超, 焦宏超, 等. 遗传算法对 SVR 风速预测模型的多参数优化. *电机与控制学报*, 2017, 21(2): 70–75.
- 12 戴李杰, 张长江, 马雷鸣. 基于机器学习的 PM<sub>2.5</sub> 短期浓度动态预报模型. *计算机应用*, 2017, 37(11): 3057–3063.
- 13 杨孟英. 基于支持向量机的中文文本分类研究[硕士学位论文]. 保定: 华北电力大学, 2017.
- 14 杨海, 丁毅, 沈海斌. 基于改进萤火虫算法的 SVM 核参数选取. *计算机应用与软件*, 2015, 32(6): 256–259. [doi: 10.3969/j.issn.1000-386x.2015.06.063]
- 15 Kanimozhi T, Latha K. An integrated approach to region based image retrieval using firefly algorithm and support vector machine. *Neurocomputing*, 2015, 151: 1099–1111. [doi: 10.1016/j.neucom.2014.07.078]
- 16 Ch S, Sohani SK, Kumar D, *et al.* A support vector machine-firefly algorithm based forecasting model to determine malaria transmission. *Neurocomputing*, 2014, 129: 279–288. [doi: 10.1016/j.neucom.2013.09.030]
- 17 王晓静, 彭虎, 邓长寿, 等. 基于均匀局部搜索和可变步长的萤火虫算法. *计算机应用*, 2018, 38(3): 715–721.
- 18 Wang H, Cui ZH, Sun H, *et al.* Randomly attracted firefly algorithm with neighborhood search and dynamic parameter adjustment mechanism. *Soft Computing*, 2017, 21(18): 5325–5339. [doi: 10.1007/s00500-016-2116-z]
- 19 Yu SH, Zhu SL, Ma Y, *et al.* A variable step size firefly algorithm for numerical optimization. *Applied Mathematics and Computation*, 2015, 263: 214–220. [doi: 10.1016/j.amc.2015.04.065]
- 20 谢金鑫. 基于数据挖掘的雾霾天气影响因子研究[硕士学位论文]. 成都: 电子科技大学, 2014.
- 21 Zhang CH, Tian YJ, Deng NY. The new interpretation of support vector machines on statistical learning theory. *Science in China Series A: Mathematics*, 2010, 53(1): 151–164.
- 22 Fister I, Fister Jr I, Yang XS, *et al.* A comprehensive review of firefly algorithms. *Swarm and Evolutionary Computation*, 2013, 13: 34–46. [doi: 10.1016/j.swevo.2013.06.001]