

# DSCE-GEP 算法在 PM<sub>2.5</sub> 浓度预测中的应用

王超学, 贾晓莉, 孙嘉诚

(西安建筑科技大学 信息与控制工程学院, 西安 710055)

**摘要:** 雾霾防治是目前空气质量保护问题研究的热点, PM<sub>2.5</sub> 浓度预测是雾霾防治的关键之一; 文章采用一种双系统协同进化的基因表达式编程算法 (DSCE-GEP) 进行 PM<sub>2.5</sub> 浓度预测, 该算法在 GEP 算法中引入人工干预操作来提高算法进化速度以及解的质量; DSCE-GEP 算法是对人类进化的模拟, 不仅具有强大的模型学习能力, 而且能得到模型的显式函数表达式; 文中以西安地区逐日 PM<sub>2.5</sub> 浓度预测为例, 将 DSCE-GEP 算法与传统基因表达式编程算法 (GEP)、文献中分类回归树和极限学习机组合模型 (CART-EELM) 以及卷积神经网络和长短期记忆神经网络组合模型 (CNN-LSTM) 进行了对比实验; 实验结果表明, DSCE-GEP 算法拟合度更高, 是一种具有竞争力的智能预测算法。

**关键词:** PM<sub>2.5</sub> 浓度预测; 协同进化; 基因表达式编程; 显式表达式; 西安地区

## Application of Double System Co-evolutionary Gene Expression Programming in PM<sub>2.5</sub> Concentration Prediction

WANG Chaoxue, JIA Xiaoli, SUN Jiacheng

(College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

**Abstract:** Haze prevention and control is a hot topic of air quality protection research at present, and PM<sub>2.5</sub> concentration prediction is one of the keys to haze prevention and control. In this paper, a dual system co-evolution gene expression programming algorithm (DSCE-GEP) was used to predict PM<sub>2.5</sub> concentration. In this algorithm, manual intervention was introduced into the GEP algorithm to improve the algorithm evolution speed and the quality of the solution. DSCE-GEP algorithm is a simulation of human evolution. It not only has strong ability of model learning, but also can get explicit function expression of the model. In this paper, the daily PM<sub>2.5</sub> concentration prediction in Xi'an area is taken as an example, and the DSCE-GEP algorithm is compared with the traditional gene expression programming algorithm (GEP), the classification regression tree and extreme learning machine combination model (CART-EELM) in literature, and the convolutional neural network and long and short-term memory neural network combination model (CNN-LSTM). The experimental results show that the DSCE-GEP algorithm has higher fitting degree and is a competitive intelligent prediction algorithm.

**Keywords:** PM<sub>2.5</sub> concentration prediction; cooperative co-evolution; gene expression programming; explicit expression; Xi'an area

## 0 引言

近年来, 雾霾已成为一种频繁出现的天气, 对人们生活产生了很大影响。医学期刊公布的研究表明, 2015 年的大气污染事件已经在全世界范围内造成了约 640 万名患者死亡, 其中约 420 万名患者由于户外大气污染罹难, 2017 年刚开始, 雾霾就已经席卷了中国的大部分地区<sup>[1]</sup>。人们对于空气悬浮物 PM<sub>10</sub> 和 PM<sub>2.5</sub> 的污染表现出空前未有的忧虑和关注, 已有的研究结论表示, PM<sub>10</sub> 和 PM<sub>2.5</sub> 是当前对公众健康造成危害的主要污染物。尤其颗粒物 (PM<sub>2.5</sub>) 由于体积小, 可黏附于深呼吸道, 通过穿透肺细胞影响血液循环, 从而影响人类健康<sup>[2-3]</sup>。研究表明,

PM<sub>2.5</sub> 浓度的上升将使人们增加患各种呼吸道疾病的风险<sup>[4]</sup>。因此, 对 PM<sub>2.5</sub> 浓度的准确预测, 不仅对于降低患病风险和保护日常健康都具有重要意义, 而且可以为 PM<sub>2.5</sub> 浓度的防治提供依据。

基因表达式编程算法<sup>[5]</sup> (GEP, gene expression programming) 是 Ferreira 受基因学中开放阅读框的启发提出的一种新型的智能进化算法, 是对遗传算法和遗传编程的融合与升华。它吸取了遗传算法 (GA) 的编码特点和遗传编程 (GP) 的树形特征, 同时又将个体的基因型和表现型相分离。它的优点在于在不了解事物内部机理、只有实验数据的情况下, 依靠它的搜索和进化能力, 找出最符合建

收稿日期: 2021-03-22; 修回日期: 2021-04-22。

基金项目: 国家自然科学基金项目 (62072363); 陕西省自然科学基金基础研究计划 (S2019-JC-YB-1191)。

作者简介: 王超学 (1967-), 男, 陕西西安人, 博士, 硕士生导师, 教授, 主要从事人工智能和数据挖掘方向的研究。

通讯作者: 贾晓莉 (1994-), 女, 陕西渭南人, 硕士生, 主要从事智能建模及预测方向的研究。

引用格式: 王超学, 贾晓莉, 孙嘉诚. DSCE-GEP 算法在 PM<sub>2.5</sub> 浓度预测中的应用[J]. 计算机测量与控制, 2021, 29(10): 71-76.

模数据的表达式。目前已经成功应用于软件可靠性<sup>[6]</sup>、气象露点<sup>[7]</sup>、函数发现<sup>[8]</sup>、能源损耗<sup>[9]</sup>、建筑<sup>[10]</sup>等领域的预测建模中。该算法不仅具有强大的模型学习能力,而且还可以得到显式的函数关系式。本文基于 GEP 算法对西安地区的 PM<sub>2.5</sub> 浓度进行预测,实验表明不仅预测精度高,而且可以得到 PM<sub>2.5</sub> 浓度与各影响因素间的函数关系。

## 1 相关工作

随着人工智能技术的快速发展,许多学者采用基于神经网络的相关算法进行 PM<sub>2.5</sub> 浓度相关的预测。谢劭峰等<sup>[11]</sup>利用鲸鱼优化算法(WOA)以及狼群算法(WPA)混合优化 BP 神经网络的权值和阈值,构建 WPA-WOA-BP 神经网络模型,实验结果表明预测稳定性高且优于 BP 网络。王勇等<sup>[12]</sup>利用 LSTM 构建融合大气污染物、GNSS 水汽和风速的 PM<sub>2.5</sub> 浓度预测模型,结果表明该模型预测精度较好。段大高等<sup>[13]</sup>提出 LSTM 预测模型,能较准确地预测未来 5 小时的 PM<sub>2.5</sub> 浓度。Y. B. Sun 等<sup>[14]</sup>提出了一种深度神经网络模型(DNN),可以提高 PM<sub>2.5</sub> 浓度的预测精度。赵文芳等<sup>[15]</sup>提出卷积神经网络和长短时记忆相结合的预测方法,能有效提高 PM<sub>2.5</sub> 浓度未来 24 h 预测精度,并具有较高的泛化能力。陈成等<sup>[15]</sup>采用多示例遗传神经网络预测室内 PM<sub>2.5</sub> 浓度,其结果优于线性回归、支持向量回归、随机森林等方法。郑国威等<sup>[16]</sup>针对 PM<sub>2.5</sub> 浓度变化的非线性、时变性等特点,建立的基于支持向量机-小波神经网络的组合预测模型,其预测结果优于支持向量机单一模型。Z. G. Shang 等<sup>[17]</sup>针对每小时 PM<sub>2.5</sub> 浓度的多种变化模式采用分类回归树和极限学习机的集成模型(CART-EELM)预测未来一小时的 PM<sub>2.5</sub> 浓度,提高了处理多种 PM<sub>2.5</sub> 变化模式的能力。T. Y. Li 等<sup>[18]</sup>提出卷积神经网络和长短期记忆神经网络组合模型(CNN-LSTM),通过单变量 CNN-LSTM 模型和多变量 CNN-LSTM 模型对比,证明了多变量 CNN-LSTM 模型预测效果更好。

综上所述,神经网络算法已经是 PM<sub>2.5</sub> 浓度预测的热点方法,但是它并不能得到显式的函数表达式。基因表达式编程是受基因学中 DNA \ 蛋白质系统提出的一种新型的智能优化算法。GEP 不但有和神经网络算法一样强大的泛函学习能力,而且可以得到显式的数学模型。

刘小生等<sup>[19]</sup>采用 GEP 算法对北京地区的 PM<sub>2.5</sub> 浓度进行逐日预测,结果发现其预测精度高于灰色理论、BP 神经网络。S. Samadianfard 等<sup>[20]</sup>采用 GEP 算法对 o<sub>3</sub> 浓度进行预测,结果发现 GEP 在预测 o<sub>3</sub> 浓度上是一个有前景的算法。但是传统 GEP 算法存在着收敛速度慢,容易早熟等问题,本文采用模拟人类进化的基因表达式编程算法(DSCE-GEP, double system co-evolutionary gene expression programming),通过人工干预与自然进化协同来提高收敛速度和全局寻优能力。目前 DSCE-GEP 已经成功应用于函数发现问题、小麦蚜虫预测问题中<sup>[21]</sup>。

本文探索双系统协同进化基因表达式编程算法在 PM<sub>2.5</sub> 浓度预测中的研究。以西安市 PM<sub>2.5</sub> 浓度预测为案例进行逐日预测模拟,以拟合度( $R^2$ )、平均绝对误差(MAE)和均方根误差(RMSE)作为模型预测评价指标。实验中通过与传统 GEP 算法、文献中分类回归树与极限学习机组合模型(CART-EELM)和卷积神经网络与长短期记忆神经网络组合模型(CNN-LSTM)对比证明 DSCE-GEP 预测模型的有效性和先进性。

## 2 基于 DSCE-GEP 的 PM<sub>2.5</sub> 浓度预测算法

文章采用双系统协同进化 DSCE-GEP 算法,该算法模拟人类进化思想引入人工干预操作与自然进化协同进化进行。人工干预由两部分组成,个体干预和种群干预。

### 2.1 自然进化系统

自然进化系统和传统基因表达式编程算法进化操作相同,包括变异、转座、重组操作。

变异:由于 GEP 算法的编码特性,根据设定的变异概率在种群中随机选择位置改变其原来的符号,按照 GEP 的变异法则进行变异运算。

转座:转座又称插串,是 GEP 算法中特有的遗传算子,其中转座元素为染色体的基因片段。在 GEP 算法中,根据插入位置的不同分为 IS 转座、RIS 转座和基因转座。

IS 转座:起始位置上是函数或终结点的短片段(IS 元素)转座到基因的头部中除根部以外的任何位置,IS 元素为种群中随机选取的基因片段。

RIS 转座:和 IS 转座的操作类似,转座片段的选取以及转座目标位置不同。转座的片段的起始位应是基因头部中的一个元素,转座的目标位置应在基因的首位。

基因转座:把个体中的某个基因转座至个体的首位作为第一个基因。

重组:又名交叉。根据基因片段不同的截取点和截取方式分为单点、两点和基因重组。

单点重组:在当代种群随机选择两个个体,然后选择同一位置将个体截断,两个个体交换截断点之后的基因片段形成两个新的个体。

两点重组:在当代种群中随机选取两个个体,接着任意选取两个基因位置,将个体切断。两个个体相互交换切断点之间的基因片段,形成两个新的个体。

基因重组:父代种群中随机选取两个染色体中,然后随机选择某一个基因进行整个基因相互交换,形成两个新的子代染色体,形成的两个子代染色体含有来自两个父体的基因。

### 2.2 人工干预系统

#### 2.2.1 个体干预

个体干预通过去掉种群中的劣质基因,增加种群中的优质基因来提高算法的进化速度以及解的质量。个体干预由“去劣”和“增优”两部分组成。“去劣”删掉 PM<sub>2.5</sub> 浓度与影响因素错误的函数表达;“增优”增加与 PM<sub>2.5</sub> 浓

度相关度高的函数和影响因素在染色体中的表达。具体操作如下:

**去劣:** 对种群中非法表达包含的劣质基因位, 如使得除数为 0、对数函数真数为 0、二次根号下小于 0 等使得表达式无意义的基因位, 通过随机选择改变相应位置的基因使其成为可行解。

**增优:** 保存每一代种群中较优的  $m$  个个体作为优质基因库。在“去劣”操作完成后, 根据增优概率, 对当前种群的第  $j$  个个体的  $s$  位置到  $t$  位置采用优质基因库中的个体通过随机选择选中的第  $i$  个个体中从  $s$  位置到  $t$  位置的基因片段  $[s: t]$  移植替换, 形成新个体  $k$ 。并评价  $k$  的适应度, 如果适应度大于原个体  $j$  的适应度, 则用新个体  $k$  替换原个体  $j$ , 否则保持原个体  $j$  不变。

### 2.2.2 种群干预

种群干预的目的是增加种群遗传信息多样性, 在算法进化过程中, 通过自然选择、优胜劣汰使得种群多样性减少, 即基因种类减少, 使得算法表达过早收敛, 陷入局部最优。因此算法通过种群干预来增加种群遗传信息的多样性, 以提高最优解的质量。种群干预使得 DSCE-GEP 算法能在较大范围的函数集中更准确地表达 PM<sub>2.5</sub> 浓度和各影响因素之间的关系。

本文选取信息熵作为种群遗传信息是否丰富的评判准则, 依据设定信息熵值评判当前种群遗传信息是否丰富。如果当前种群信息熵值大于或等于设定值则当前种群不变; 小于设定值则对进行干预操作。具体操作如下:

把种群根据适应度从大到小排序, 对于倒数  $a$  个个体, 即适应度较低的  $a$  个个体, 采用镜像个体替换; 对于倒数第  $a$  个到倒数第  $a+b$  个个体, 即适应度更低的  $b$  个个体, 采用随机个体替换形成新种群。

1) 镜像个体: 函数符集  $F = \{+, -, *, /, \text{sqrt}, x^2, \exp, \cos, \sin, \ln, \lg, \sim (\text{以 } 10 \text{ 为底的指数}), \text{abs}, C (\text{常数})\}$ , 镜像函数符集  $\text{mr}_F = \{-, +, /, *, x^2, \text{sqrt}, \ln, \sin, \cos, \exp, \sim, \log, C, \text{abs}\}$ , 对需要镜像替换的个体遍历, 如果第  $i$  位基因为  $F$  中第  $j$  个元素, 则替换后个体第  $i$  位基因为  $\text{mr}_F$  中第  $j$  个元素。依据上述规则操作遍历完个体所有基因位后形成新个体。

2) 随机个体: 和初始生成个体规则相同。

### 2.3 DSCE-GEP 算法具体步骤

算法流程如图 1 所示。

**步骤 1: 始化参数:** 确定种群规模、头部长度、尾部长度、基因个数、函数集、终点集、遗传概率等参数。

**步骤 2: 码/评价个体:** 针对每个染色体, 根据适应度求解公式计算其适应度。

**步骤 3: 建立优质基因库:** 选出种群中适应度较大的  $m$  个个体作为优质基因库成员。

**步骤 4: 判断是否满足终止条件:** 判断是否存在个体适应度达到最大值或进化代数达到最大值, 如果是, 则结束;

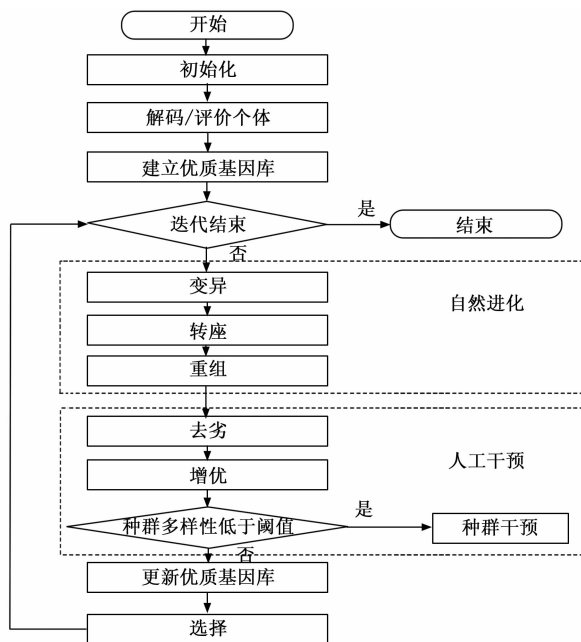


图 1 DSCE-GEP 流程图

否则, 进入下一步。

**步骤 5: 遗传操作:**

1) 变异: 根据变异概率随机选择变异个体进行变异操作。

2) 转座 (IS RIS 基因): 根据不同的转座位置分为 IS、RIS、基因转座, 依据转座概率随机选择个体进行相应的转座操作。

3) 重组 (单点两点基因): 根据不同基因片段的组合分为单点、两点、基因重组, 依据重组概率随机选择个体进行相应的重组操作。

**步骤 6: 人工干预操作**

1) 个体干预: (1) 去劣, 对种群个体遍历, 对于造成不可行解的基因通过随机选择替换使其成为可行解; (2) 增优, 对于适应度较低的个体, 通过优质基因库中的优质基因片段替换适应度较低个体相应位置的基因片段。

2) 种群干预: 根据当前种群的信息熵值判断是否需要种群干预。如果当前种群信息熵值小于理想种群信息熵值就进行种群干预, 否则不干预。

**步骤 7: 更新优质基因库:** 合并当前优质基因库和当前种群的所有个体, 从中选出适应度较高的前  $m$  个个体作为新的优质基因库成员。

**步骤 8: 选择:** 采用轮盘赌选择策略进行选择操作形成新种群并进入下一次迭代。

## 3 实验

本文设计了 3 个对比实验来验证 DSCE-GEP 算法在 PM<sub>2.5</sub> 浓度预测中的有效性和先进性。实验程序用 python 实现, 实验环境为 Intel i5 处理器、16 GB 内存、Windows10 操作系统。

### 3.1 数据集

本文采用的空气质量数据 (PM2.5, PM10, SO<sub>2</sub>, NO<sub>2</sub>, CO, O<sub>3</sub>) 与气象数据 (温度, 湿度, 风向, 风速, 气压, 露点), 通过中国环境监测网站获取的 2017.1.1~2018.12.31 的西安地区每天监测数据, 所采集的数据为每日平均值。70% 数据作为训练集, 30% 数据作为测试集。

### 3.2 适应度函数

文中采用拟合度  $R^2 = 1 - SSE/SST$  作为适应度函数, 即统计学中的复相关系数。其中,  $SSE$  计算如公式 (1) 所示,  $SST$  计算如公式 (2) 所示:

$$SSE = \sum_{j=1}^n (y_j - \hat{y}_j)^2 \quad (1)$$

$$SST = \sum_{j=1}^n (y_j - \bar{y})^2 \quad (2)$$

其中:  $y_j$  表示真实数据,  $\hat{y}_j$  表示预测数据,  $\bar{y}$  表示真实数据的均值。当  $R^2$  越接近 1 时, 表明预测准确度越高。

### 3.3 预测精度评价指标

本文采用拟合度 ( $R^2$ )、均方根误差 (RMSE)、平均绝对误差 (MAE) 3 个指标来对预测结果进行评价。均方根误差 (RMSE) 计算如公式 (3) 所示, 平均绝对误差 (MAE) 计算如公式 (4) 所示:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

其中:  $n$  为预测集的长度,  $y_i$  为预测集中第  $i$  个真实值,  $\hat{y}_i$  为模型得出的第  $i$  个预测值。

### 3.4 初始化参数设定

DSCE-GEP 算法参数如表 1 所示。

表 1 初始化参数设置

参数	设置值
最大进化代数	200
种群大小	80
优质基因库规模	15
函数符集合 F	+ - * / lg exp ln ~ x2 sqrt sin cos abs
终止符集合 T	自变量个数
连接符	+
头部长度	10
基因个数	6
倒串率	0.3
点变异率	0.3
重组率	0.2
插串概率	0.1
IS 的长度	{1,2,3,4,5}
RIS 的长度	{1,2,3,4,5}
增优概率	0.2
镜像替换个体 a	25
随机替换个体 b	15

### 3.5 对比实验与结论

#### 3.5.1 GEP、DSCE-GEP 预测精度对比

1) 预测精度对比: GEP 算法、DSCE-GEP 算法算法预测曲线分别如图 2、图 3 所示。拟合度、均方根误差、平均绝对误差比较如表 2 所示。

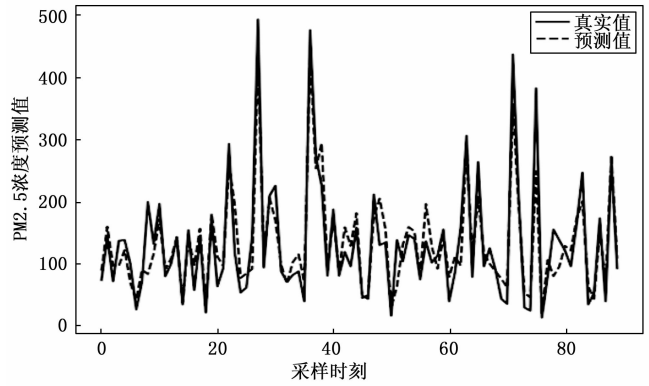


图 2 GEP 预测曲线

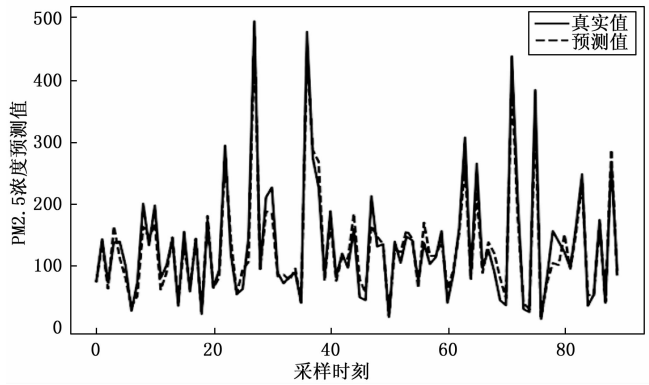


图 3 DSCE-GEP 预测曲线

从图 2、图 3 可看出 GEP 算法、DSCE-GEP 算法预测效果都比较好, 由表 2 得出 DSCE-GEP 较 GEP 拟合度提高 0.03, 均方根误差降低 4.29, 平均绝对误差降低 5.47, 较传统 GEP 提高了预测精度, 证明了 DSCE-GEP 的有效性。

表 2 DSCE-GEP 和 GEP 性能比较

模型	$R^2$	RMSE	MAE
DSCE-GEP	0.85	17.43	22.36
GEP	0.82	21.72	27.83

2) GEP、DSCE-GEP 进化速度对比: GEP 算法、DSCE-GEP 算法进化曲线分别如图 4、图 5 所示。

通过图 4、图 5 进化曲线对比发现, GEP 在 175 代找到最优解, DSCE-GEP 在 100 代左右找到最优解 DSCE-GEP 较 GEP 提高了进化速度, DSCE-GEP 通过个体干预及种群干预, 缩小了解的搜索空间, 提高了进化速度。通过实验对比 DSCE-GEP 算法能有效提高收敛速度和预测

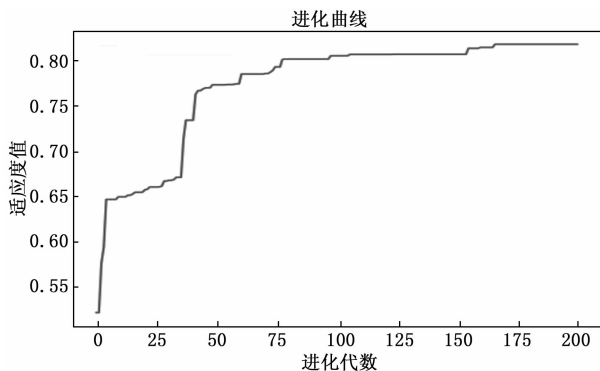


图 4 GEP 最优进化曲线

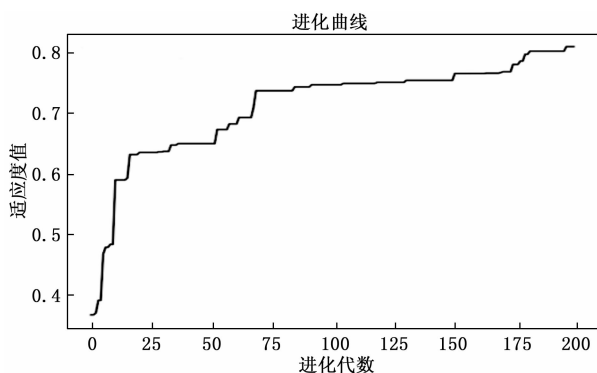


图 5 DSCE-GEP 最优进化曲线

精度。

### 3.5.2 DSCE-GEP、CART-EEL 和 CNN-LSTM 对比实验

1) DSCE-GEP 算法的显式表达式:

$$y = \cos 10^{\cos(\sin x_{11} * x_{11}) - x_1} + \cos 10^{x_7} + 2x_6 - \cos(x_{11} * x_{10}) + \sqrt{x_5^2 - x_{10}} + 10^{|x_1|} + x_{11} + \cos x_8 \quad (5)$$

通过实验得到 DSCE-GEP 算法的预测模型为公式 (5)。其中  $y$  为当前一天 PM2.5 浓度,  $x_1, x_3, x_6, x_7, x_8, x_{10}, x_{11}$  分别为前一天 SO<sub>2</sub>, O<sub>3</sub>, 温度, 湿度, 风向, 露点, PM2.5。通过实验模型可以看出, 西安地区当天 PM2.5 浓度主要与前一天 SO<sub>2</sub>, O<sub>3</sub>, 温度, 湿度, 风向, 露点, PM2.5 相关, 具体关系体现在公式 (5) 中。

2) DSCE-GEP 算法、CART-EELM 模型、CNN-LSTM 模型预测精度对比

DSCE-GEP 算法与比较分类回归树和极限学习机的组合模型和 CNN-LSTM 模型, 证明其先进性, 结果表 3 所示。

表 3 DSCE-GEP 和 CART-EELM、CNN-LSTM 预测结果对比

模型	R <sup>2</sup>	RMSE	MAE
DSCE-GEP	0.85	17.43	22.36
CART-EELM	0.84	16.24	21.86
CNN-LSTM	0.83	18.96	23.90

通过表 3 发现, DSCE-GEP 算法较 CART-EELM 拟合度高了 0.01, 均方根误差和平均绝对误差稍高于 CART-EELM; 较 CNN-LSTM 拟合度高了 0.01, 均方根误差低了 1.53, 平均绝对误差低了 1.54。拟合度均优于其他两种模型, 证明了 DSCE-GEP 算法在 PM2.5 浓度预测中的竞争力与先进性。CART-EELM 和 CNN-LSTM 分别都是基于神经网络的模型, 最终得出的模型为参数矩阵, 无法看到 PM2.5 浓度与各影响因素之间的关系, 而 DSCE-GEP 算法最终可以得到 PM2.5 浓度与其影响因素之间的显式函数关系, 可以明确表达 PM2.5 浓度与各影响因素之间的关系。

### 3.5.3 实验结论

经过上述实验分析发现, DSCE-GEP 算法不仅较传统 GEP 算法、分类回归树和极限学习机组合算法 (CART-EELM) 以及卷积神经网络和长短期记忆神经网络算法 (CNN-LSTM) 在西安地区逐日的 PM2.5 浓度中性能较优, 拟合度均高于其他 3 种模型, 而且得到了 PM2.5 浓度预测的显式函数关系。证明 DSCE-GEP 算法在 PM2.5 浓度预测中的有效性与竞争力, 对于准确的控制 PM2.5 浓度具有重要意义。

## 4 结束语

通过对西安地区空气质量数据进行分析, 根据过去一天的空气质量数据及气象数据预测当天的 PM2.5 浓度, 利用基于双系统协同进化算法得到预测模型, 不仅可以有较高的预测精度, 而且可以明确地看出各影响因素与 PM2.5 浓度之间的关系。未来还可以把 DSCE-GEP 算法应用于其他智能预测领域中。

### 参考文献:

[1] 陈晓雪. 无人幸免: 权威研究揭示空气污染的多重健康危害 [EB/OL]. (2017-01-08) [2021-01-05]. <http://www.casrs.cn/NewsShow.asp?bid=198>.

[2] PERRONE M G, GUALTIERI M, CONSONNI V, et al. Particle size, chemical composition, seasons of the year and urban, rural or remote site origins as determinants of biological effects of particulate matter on pulmonary cells [J]. Environmental Pollution, 2013, 176: 215-227.

[3] BONO R, TASSINARI R, BELLISARIO V, et al. Urban air and tobacco smoke as conditions that increase the risk of oxidative stress and respiratory response in youth [J]. Environmental Research, 2015, 137: 141-146.

[4] HE J J, GONG S L, YU Y, et al. Air pollution characteristics and their relation to meteorological conditions during 2014-2015 in major Chinese cities [J]. Environmental Pollution, 2017, 223: 484-496.

[5] FERREIRA C. Gene expression programming: a new adaptive algorithm for solving problems [J]. Complex System, 2001,

13 (2): 87-129.

- [6] 李海峰, 陆民燕, 曾敏, 等. 基因表达式编程在软件可靠性建模中的应用 [J]. 计算机科学与探索, 2011 (6): 534-546.
- [7] MEHDIZADEH S, BEHMANESH J, KHALILI K. Application of gene expression programming to predict daily dew point temperature [J]. Applied Thermal Engineering, 2017, 112: 1097-1107.
- [8] 王超学, 吴书玲, 张婧菁. 采用混合策略的改进基因表达式编程 [J]. 计算机科学与探索, 2017, 11 (1): 163-170.
- [9] KABOLI S H A, FALLAHPOUR A, SELVARAJ J, et al. Long-term electrical energy consumption formulating and forecasting via optimized gene expression programming [J]. Energy, 2017, 126: 144-164.
- [10] IQBAL M F, LIU Q F, AZIM I, et al. Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming [J]. Journal of Hazardous Materials, 2020, 384: 121322.
- [11] 谢劭峰, 赵云, 李国弘, 等. 基于 WPA-WOA-BP 神经网络的 PM<sub>2.5</sub> 浓度预测 [J]. 大地测量与地球动力学, 2021, 41 (1): 12-16.
- [12] 王勇, 王泓易, 刘严萍, 等. 融合 GNSS 水汽、风速与大气污染物的河北省冬季 PM<sub>2.5</sub> 浓度预测研究 [J]. 大地测量与地球动力学, 2020, 40 (11): 1145-1152.
- [13] 段大高, 赵振东, 梁少虎, 等. 基于 LSTM 的 PM<sub>2.5</sub> 浓度预测模型 [J]. 计算机测量与控制, 2019, 27 (3): 215-219.

(上接第 70 页)

- [7] 李雅晶, 辛妍丽. 基于 SVR 的燃煤机组 NO<sub>x</sub> 含量的软测量模型 [J]. 计算机测量与控制, 2020, 28 (8): 62-66.
- [8] 方贤, 铁治欣, 崔仕文, 等. 基于粒子群优化双支持向量机的 SCR 烟气脱硝效率预测模型 [J]. 热力发电, 2018, 47 (1): 53-58.
- [9] 杨碧源, 赵金笑, 魏宏鸽, 等. 基于 BP 神经网络的 SCR 蜂窝状催化剂脱硝性能预测 [J]. 中国电力, 2016, 49 (10): 127-131.
- [10] YANG Y Q, WANG G J, YANG Y. Parameters optimization of polygonal fuzzy neural networks based on GA-BP hybrid algorithm [J]. International Journal of Machine Learning and Cybernetics, 2014, 5 (5): 815-822.
- [11] 刘春艳, 凌建春, 寇林元, 等. GA-BP 神经网络与 BP 神经网络性能比较 [J]. 中国卫生统计, 2013, 30 (2): 173-176.
- [12] LI D Y, HUANG F M, YAN L X, et al. Landslide susceptibility prediction using particle-swarm-optimized multilayer perceptron; comparisons with multilayer-perceptron-only, BP neural network, and information value models [J]. Applied Sciences, 2019, 9 (18): 3664.
- [13] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer [J]. Advances in Engineering Software 2014, 69: 46

- [14] SUM Y B, ZENG Q L, GENG B, et al. Deep learning architecture for estimating hourly ground-level PM<sub>2.5</sub> using satellite remote sensing [J]. IEEE Geoscience and Remotesensing Letters, 2019, 16 (9): 1343-1347.
- [15] 赵文芳, 林润生, 唐伟, 等. 基于深度学习的 PM<sub>2.5</sub> 短期预测模型 [J]. 南京师大学报 (自然科学版), 2019, (3): 32-41.
- [16] 郑国威, 王腾军. 基于支持向量机小波神经网络的 PM<sub>2.5</sub> 预测模型 [J]. 四川环境, 2018 (6): 141-144.
- [17] SHANG Z G, DENG T, HE J Q, et al. A novel model for hourly PM<sub>2.5</sub> concentration prediction based on CART and EELM [J]. Science of the Total Environment, 2019, 651 (2): 3043-3052.
- [18] LI T Y, HUA M, WU X. A Hybrid CNN-LSTM model for forecasting particulate matter (PM<sub>2.5</sub>) [J]. IEEE ACCESS, 2020 (8): 26933-26940.
- [19] 刘小生, 李胜, 赵相博. 基于基因表达式编程的 PM<sub>2.5</sub> 浓度预测模型研究 [J]. 江西理工大学学报, 2013 (5): 1-5.
- [20] SAMADIANFARD S, DELIRHASANNIA R, KISI O, et al. Comparative analysis of ozone level prediction models using gene expression programming and multiple linear regression [J]. Geofizika, 2013, 30 (1): 43-73.
- [21] 王超学, 吴书玲, 张婧菁. 基于人类进化基因表达式编程的麦蚜数量预测方法及系统 [P]. 中国, 201510014836.1, 2015-05-06.
- [14] AHMED H M, YOUSSEF B A B, ELKORANY A S, et al. Hybrid gray wolf optimizer-artificial neural network classification approach for magnetic resonance brain images [J]. Applied Optics, 2018, 57 (7): B25-B31.
- [15] LU C, GAO L, LI X Y, et al. A hybrid multi-objective grey wolf optimizer for dynamic scheduling in a real-world welding industry [J]. Engineering Applications of Artificial Intelligence, 2017, 57 (C): 61-79.
- [16] ZHU A J, XU C P, LI Z, et al. Hybridizing grey wolf optimization with differential evolution for global optimization and test scheduling for 3D stacked SoC [J]. Journal of Systems Engineering & Electronics, 2015, 26 (2): 317-328.
- [17] 方一鸣, 赵晓东, 张攀, 等. 基于改进灰狼算法和多核极限学习机的铁水硅含量预测建模 [J]. 控制理论与应用, 2020, 37 (7): 1644-1654.
- [18] 万立明, 李德波, 周杰联, 等. SCR 脱硝系统流场优化改造关键技术研究 [J]. 浙江电力, 2019, 38 (7): 103-108.
- [19] 谭传玉. 火电厂氮氧化物排放测量研究 [D]. 北京: 华北电力大学 (北京), 2019.
- [20] 姜焯, 高翔, 吴卫红, 等. 选择性催化还原脱硝催化剂失活研究综述 [J]. 中国电机工程学报, 2013, 33 (14): 18-31.