

link

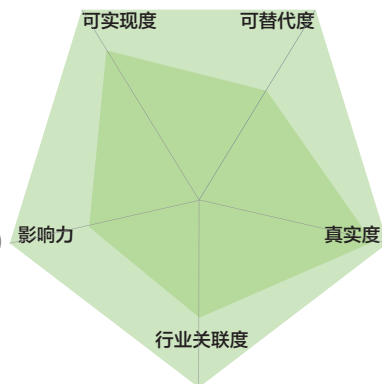
appraisalment

industry

方彦 (1981-) 女, 讲师。

1. 泉州师范学院 数学与计算机科学学院; 2. 福建省大数据管理新技术与知识工程重点实验室

基金项目: 泉州市科技局计划项目 (2014Z135); 泉州师范学院校级自选项目 (2013KJ11)



方彦 基于灰色 RBF 神经网络的 空气质量预测

针对神经网络建模预测时, 其建模精度往往受到数据随机性的影响, 以及灰色系统具有减小数据随机性的特点, 提出了灰色 RBF 神经网络组合模型。根据中国环境监测总站数据中心采集到的泉州市 2017 年 8 月 PM_{2.5}、PM₁₀、CO、NO₂、SO₂ 和 O₃ 六个表征空气污染程度的数据作为训练样本构建模型, 并选取 2018 年同期数据作为测试样本以对预测模型进行精度检验, 最后利用该模型预测泉州市空气质量状况。结果表明, 所建立的灰色 RBF 神经网络组合模型相对模拟误差小, 预测结果更为可靠, 可以用于空气质量预测。

随着城市经济社会的不断发展, 在经历一些严重的大气污染导致的突发事件之后, 人们对大气生态环境安全的问题日益重视。近年来, 国家也加大了对大气环境污染治理力度。为了寻找分析空气污染源及预测空气质量异常变化态势, 必须对各地空气质量进行有效监测与预测。

由于空气质量受到地区经济发展布局、污染源排放以及气象条件等众多因素影响, 其数据是复杂的非线性时间序列, 预测模型必然受到多变量耦合作用且具有明显的非线性特征, 运用传统的线性方法难以取得良好的预测结果, 因此, 构造非线性模型对空气质量进行预测, 是该问题的主要研究趋势。张楠等将灰狼优化算法结合支持向量机回归算法预测空气质量指数, 蒲国林等提出改进人工蜂群算法和反向传播神经网络相结合的空气质量预测方法, 林开春等基于随机森林和神经网络优化建立了预测空气质量指数的模型, 张梦瑶等引入模糊级别特征值算法和空气质量指数评价方法作预测分析, 王坤等建立季节时间序列模型预测空气质量。

本文在已有的研究基础上将灰色预测模型和 RBF 神经网络理论结合构建组合预测模型, 既减少数据中随机性, 又提高复杂非线性函数逼近能力。而后运用模型研究空气质量预测的问题, 建立多维预测数据, 挖掘数据间的内在联系, 实现对空气质量的实时、有效、准确预测。

基本原理

灰色系统理论

灰色系统理论是 1982 年由我国学者邓聚龙教授创立的, 该理论以不确定性系统为研究对象, 将部分信息已知、部分信息未知的小样本、贫信息视为灰信息, 通过对部分已知信息的生成和开发, 挖掘有价值的内在规律。其模型可描述为如下。

设系统特征数据序列:

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \quad (1)$$

为原序列, 其中 $x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)$ 表示决策属性的因变量的 n 个观测值, 其一次累加生成序列 (1-AGO):

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)), \quad (2)$$

其中 $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$, $k=1, 2, \dots, n$ 对生成序列 $X^{(1)}$ 建立微分方程:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (3)$$

方程的参数列 $c = (a, b)^T$, 由最小二乘法可以得到近似值

$$c = (B^T B)^{-1} B^T Y_n \quad (4)$$

其中, B 、 Y_n 为:

$$B = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \dots & \dots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix}, \quad Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix}^T \quad (5)$$

该方程的解为:

$$x^{(1)}(t+1) = \left(x^{(1)}(1) - \frac{b}{a} \right) e^{-at} + \frac{b}{a} \quad (6)$$

其中 $x^{(1)}(t+1)$ 是累加值的预测值, 将其做一次累减还原成原数据得:

$$x^{(0)}(t+1) = x^{(1)}(t+1) - x^{(1)}(t) = -a \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-at}, \quad (t=1, 2, 3, \dots, n) \quad (7)$$

RBF 神经网络理论

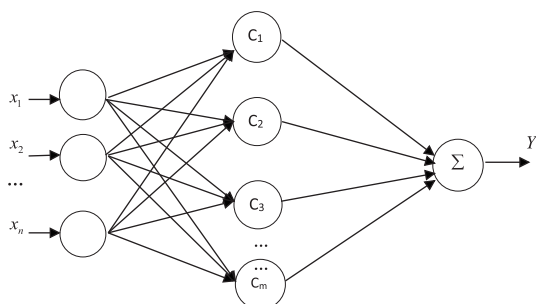


图 1 RBF 神经网络结构图

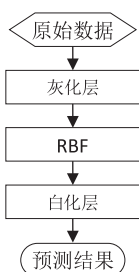


图 2 灰色 - RBF 神经网络组合模型结构

人工神经网络是研究非线性科学问题的模型，RBF 神经网络是以径向基函数作为隐含层结点激活（核）函数的三层前向型神经网络，包括输入层、隐含层和输出层，其拓扑结构如图 1 所示。

其中， $X = (x_1, x_2, \dots, x_n)$ 为输入向量， $C = (c_1, c_2, \dots, c_m)$ 为隐含层，通过高斯基函数 $c_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{2\delta_i^2}\right)$ 实现输入层到隐含层的非线性映射，其中 c_i 为第 i 个基函数的中心， $\|x - c_i\|$ 为向量 $x - c_i$ 的范数，即 x 与 c_i 之间的距离， δ_i 为第 i 个感知变量， Y 为输出向量，是隐含层 C 的线性映射， $Y = \sum_{j=1}^m w_j c_j$ ， $w_j (j=1 \dots m)$ 为第 j 个结点到输出层的权值。

灰色 - RBF 神经网络组合模型

灰色系统与 RBF 神经网络的结合形式主要有并联型、串联型、嵌入型，本文采用嵌入型灰色 - RBF 神经网络，在 RBF 神经网络的输入层之前加入灰化层对输入数据做了灰化处理，而在输出层后加入白化层对输出数据做还原处理，组合模型结构如图 2 所示。

实验结果与分析

数据整理

本文自中国环境监测总站数据中心采集泉州市 2017 年 8 月 $PM_{2.5}$ 、 PM_{10} 、 CO 、 NO_2 、 SO_2 和 O_3 六个表征空气污染程度的数据作为训练样本构建模型，并选取 2018 年同期数据作为测试样本以对预测模型进行精度检验。

模型构建与结果分析

本文采用 PYTHON 与 RBF - MASTER 编程实现，分别针对 $PM_{2.5}$ 、 PM_{10} 、 CO 、 NO_2 、 SO_2 和 O_3 的构建预测模型进行仿真模拟实验，预测结果如图 3 ~ 图 9 所示。

由上述对比图及误差图可以看出，灰色 - GBF 神经网络模型预测结果与测试样本实际检测结果基本一致，预测结

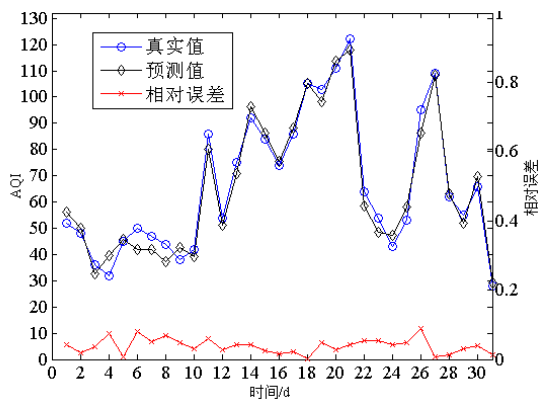
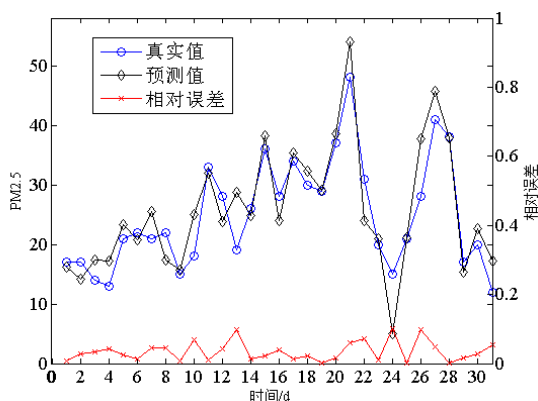
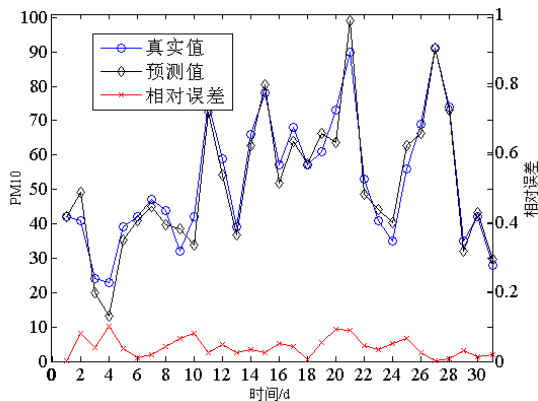
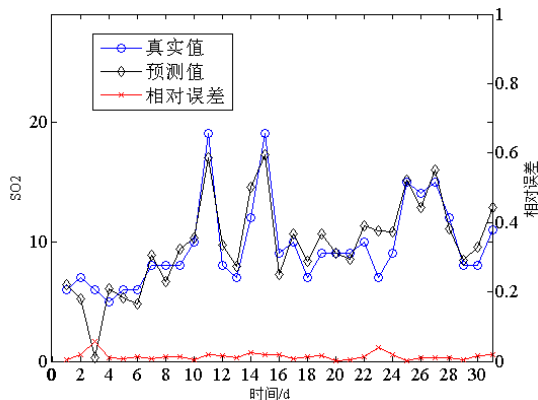


图 3 AQI 预测曲线

图 4 $PM_{2.5}$ 预测曲线图 5 PM_{10} 预测曲线图 6 SO_2 预测曲线

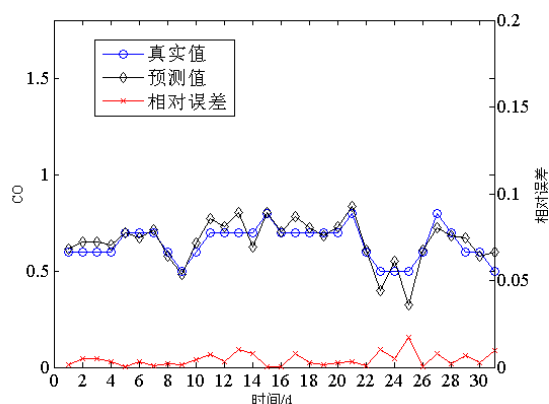


图 7 CO 预测曲线

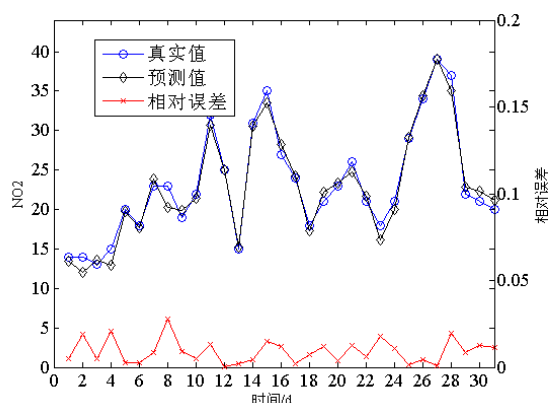


图 8 NO₂ 预测曲线

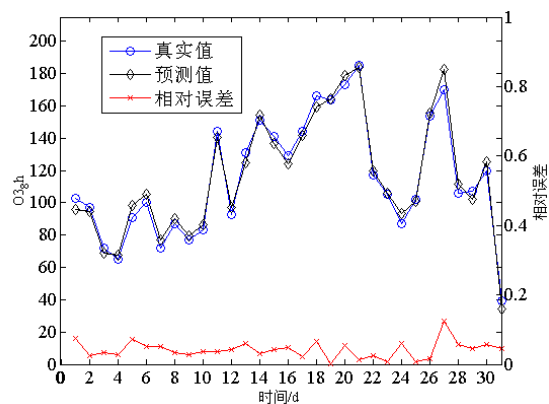


图 9 O₃ 预测曲线

果表明本文所建立的模型能够为空气质量预测工作作出一定的指导。

结束语

本模型将灰色系统与 RBF 神经网络组合, 从仿真结果来看具有以下几个特点:

- (1) 采用灰色系统对数据预处理, 消除了数据的随机性, 并一定程度上减少数据样本较少的影响;
- (2) RBF 神经网络具有高度非线性的拟合能力, 拟合精度能够达到预计目标;
- (3) 灰色系统能减少 RBF 神经网络建模精度容易受数据随机性影响的问题。

(上接第 99 页)

此冲击缺陷的初始损伤, 参数见表 5, 初始损伤见式 (19)。

$$D_0 = \left[\frac{\sigma_{eqmax}^2 R_p}{2ES(1-D)^2} \right]^s \Delta p = \left[\frac{874^2}{2 \times 107000 \times 4.426} \right]^{4.65} \times 0.2755 = 0.147 \quad (19)$$

根据式 (15) 计算相同试验载荷下的损伤力学预估寿命, 计算结果及误差见表 6。

表 6 计算结果及误差

载荷幅值 σ_a (Mpa)	损伤力学预估寿命 (周次)	与试验寿命误差
480	54454.272	13.56%
530	45225.523	5.17%

将一系列应力值代入到式 (15) 中可得到一系列预估寿命, 将得出的数据列于一个坐标系中观察分布规律, 见图 7。

图 7 表明对于此 TB6 板试样, 由理论计算得出的载荷与寿命的关系曲线与试验拟合的 S - N 曲线的变化趋势一致, 故在计算疲劳极限时可采用表 2 中的形状参数。

将估算的寿命与载荷代入式 (17) 中可得到基于损伤力学的带冲击缺陷 TB6 板试样的疲劳极限 $S'_{\infty flsw}$ 及基于损伤力学的 TB6 冲击缺陷影响系数 $K'_{\infty flsw}$, 将计算结果与试验结果列于表 7 中作对比。对比结果表明两者的误差为 2.8%。

表 7 算例与试验对比结果

$S_{\infty flsw}$ (Mpa)	$S'_{\infty flsw}$ (Mpa)	$K'_{\infty flsw}$	$K_{\infty flsw}$	误差
380	391	1.074	1.105	2.8%

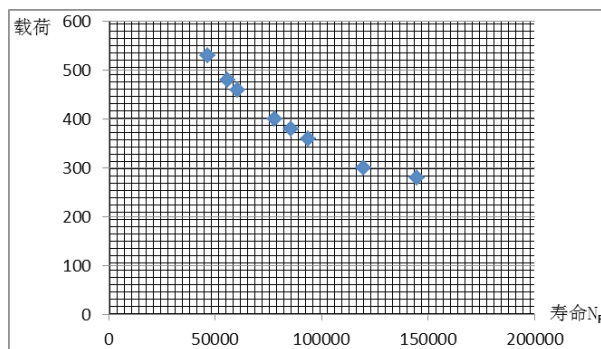


图 7 损伤力学寿命预估分布规律

结束语

基于经典的金属疲劳损伤力学演化方程建立了一种计算带冲击缺陷 TB6 疲劳寿命的预估模型, 模型中考虑了冲击残余应力及初始损伤。通过此模型得出的载荷与寿命的关系曲线与试验拟合的 S - N 曲线的变化趋势一致, 故可通过疲劳三参数公式得出基于损伤力学的带冲击缺陷 TB6 疲劳极限, 进而可得到基于损伤力学的 TB6 冲击缺陷影响系数。

利用此种方法对 TB6 板试样进行算例验证并与试验结果对比得到的 TB6 冲击缺陷影响系数误差为 2.8%, 误差水平较低, 该理论方法为工程上对带冲击缺陷 TB6 的缺陷影响系数估算提供了一种有效途径。