

# 基于 KPCA-BAS-GRNN 的埋地管道 外腐蚀速率预测

骆正山<sup>1</sup>, 姚梦月<sup>1</sup>, 骆济豪<sup>2</sup>, 王小完<sup>1</sup>

(1.西安建筑科技大学 管理学院, 西安 710055; 2.西安交通大学附属中学, 西安 710056)

**摘 要:** **目的** 提高埋地管道外腐蚀速率的预测精度。**方法** 建立基于核主成分分析法 (KPCA) 和天牛须搜索 (BAS) 算法优化的广义回归神经网络 (GRNN) 腐蚀速率预测模型, 通过 KPCA 对原始数据进行预处理, 提取影响管道外腐蚀的主要因素, 应用 GRNN 建立埋地管道外腐蚀速率预测的数学模型, 并采用 BAS 算法对模型进行优化, 减小了人为设置参数的影响。以川气东送埋地管段为例, 分析选取出 12 种关键影响因素, 建立了埋地管道外腐蚀指标体系, 借助 MATLAB-R2014a 编写程序进行仿真, 并与实际值进行对比。**结果** 模型的预测结果与实际值基本一致, KPCA 可有效降低指标体系的维度, 提取出包含原始信息 97.9% 的 3 个主因素——土壤电阻率、氧化还原电位、氯离子含量, 简化了运算过程。采用的 BAS-GRNN 模型将预测精度提高到 7.83% 以内, 平均相对误差 5.21%, 决定系数取值 0.93。与其他模型相比, 该模型性能较好, 预测精度更高。**结论** 采用 KPCA 提取的主要影响因素符合工程实际, 建立的 BAS-GRNN 模型预测精度高, 有较好的适应性, 为埋地管道外腐蚀速率预测提供了新思路, 对管道的维护更新工作提供了参考依据。

**关键词:** 埋地管道; 外腐蚀速率预测模型; 核主成分分析法 (KPCA); 天牛须搜索算法 (BAS); 广义回归神经网络 (GRNN)

中图分类号: TG172 文献标识码: A 文章编号: 1001-3660(2018)11-0173-08

DOI: 10.16490/j.cnki.issn.1001-3660.2018.11.025

## Prediction of External Corrosion Rate of Buried Pipeline Based on KPCA-BAS-GRNN

LUO Zheng-shan<sup>1</sup>, YAO Meng-yue<sup>1</sup>, LUO Ji-hao<sup>2</sup>, WANG Xiao-wan<sup>1</sup>

(1.School of Management, Xi'an University of Architecture & Technology, Xi'an 710055, China;

2.Affiliated Middle School of Xi'an Jiaotong University, Xi'an 710056, China)

**ABSTRACT:** The work aims to improve the prediction accuracy of the external corrosion rate of buried pipeline. The corrosion rate prediction model of buried pipeline was established based on kernel principal component analysis (KPCA) and the general regression neural network (GRNN) optimized by Beetle antennae search (BAS) algorithm. The main factors affecting external corrosion of buried pipeline were extracted by preprocessing the original data through KPCA. GRNN was used to build a

收稿日期: 2018-08-13; 修订日期: 2018-09-23

Received: 2018-08-13; Revised: 2018-09-23

基金项目: 国家自然科学基金资助 (41877527); 陕西省社科基金项目 (2018S34); 陕西省重点学科建设专项资金资助项目 (E08001); 陕西省教育厅自然专项基金 (16JK1465)

**Fund:** Supported by the National Natural Science Foundation of China (41877527), Social Science Fund Project in Shaanxi province (2018S34), Special Fund for Key Disciplines Construction in Shaanxi Province (E08001), Education Department Nature Special Fund in Shaanxi Province (16JK1465)

作者简介: 骆正山 (1969—), 男, 教授, 博士生导师, 主要研究方向为油气管道腐蚀、风险评估与建模。邮箱: luozhengshan@163.com

**Biography:** LUO Zheng-shan (1969—), Male, Professor, Doctoral supervisor, Research focus: oil and gas pipeline corrosion, risk assessment and modeling. E-mail: luozhengshan@163.com

mathematical model to predict the external corrosion rate of buried pipeline and BAS algorithm was adopted to optimize the model to reduce the effects of artificially set parameters. In addition, the pipelines buried in natural gas transmission project from Sichuan to East were utilized as an example to analyze 12 key influencing factors and establish the external corrosion index system of buried pipeline. MATLAB-R2014a software was used for simulation processing, and compared with the actual values. The predicted results of the model were basically consistent with the actual values. KPCA could effectively reduce the dimensions of the indicator system and extract three main factors with 97.9% original information, including soil resistivity, oxidation-reduction potential and  $\text{Cl}^-$  content. Thus, the calculation process was simplified. BAS-GRNN model was adopted to improve the prediction accuracy to 7.83%. The average relative error was 5.21%, and the determination coefficient was 0.93. Compared with other models, this model had better performance and higher prediction accuracy. Thus, the main influencing factors extracted by KPC Accord with engineering practice. BAS-GRNN model provides a new idea for the prediction of external corrosion rate of buried pipeline and a reference basis for the maintenance and updating of buried pipeline by higher precision and better adaptability.

**KEY WORDS:** buried pipeline; external corrosion rate prediction model; kernel principle component analysis (KPCA); beetle antennae search algorithm (BAS); generalized regression neural network (GRNN)

近年来,我国油气资源的开发不断深入,管道作为重要的能源传输系统,其安全可靠的运行是防止事故、减少损失的关键<sup>[1]</sup>。管道的运行受多种腐蚀因素的影响,以往学者常考虑单因素下管道的内腐蚀情况。如:李斌等<sup>[2]</sup>分析了在  $\text{H}_2\text{S}/\text{CO}_2$  酸性条件下,不同溶解氧浓度对低合金管道点蚀的影响;李兆玲等<sup>[3]</sup>通过实验研究了氯离子和直流杂散电流共同作用下,对 X65 钢的均匀腐蚀速率的影响;谭才渊等<sup>[4]</sup>以渤海某油田油管为例,分别分析了  $\text{H}_2\text{S}$  分压、 $\text{CO}_2$  分压、含水率对 L80 管道电化学性能的影响。但土壤所处环境复杂,存在多种腐蚀因素<sup>[5]</sup>,导致埋地管道的外腐蚀问题日渐突出。因此,研究管道的外腐蚀因素及规律,建立可靠的管道外腐蚀预测模型具有重要意义。

目前,国内外学者对管道的腐蚀情况开展了大量研究。如:毕傲睿等<sup>[6]</sup>利用主成分分析法提取埋地管道土壤腐蚀的主要影响因素,消除了相关联指标的重叠信息,减少了问题的复杂性;陈健飞等<sup>[7]</sup>利用灰色关联度分析方法,找出了影响管道防腐层质量的主要因素;张杰<sup>[8]</sup>利用主成分-聚类分析法降低了管道风险评价指标的冗余性和主观性信息,并根据风险指标值对管段排序和归类。但聚类分析法根据相似系数或距离得出最终结果,带有很大的主观性,并且仅适用于小样本。灰色关联度分析法中各种指标值需要人为确定,主观性过强。PCA 仅适用于线性映射问题,而埋地管道外腐蚀影响因素多,采用非线性映射方法将原始数据映射到特征空间,并在特征空间进行线性操作,由此提取的主要影响因素更符合工程实际。随着计算机行业的蓬勃兴起,大量算法应用于管道的腐蚀预测。Swati Jain 等<sup>[9]</sup>提出了一种基于 Bayesian networks 定量评价油气管道外腐蚀速率的模型,考虑了管道外腐蚀的影响因素,将部分定性信息量化,提高了腐蚀预测的准确度。李丽等<sup>[10]</sup>运用了 BP 人工神经网络,对埋地管道土壤腐蚀的速率进行预测分析,

为防腐工作提供了参考依据。MengQian 等<sup>[11]</sup>建立 PSO-SVR 模型,对煤矿安全进行了深入研究。贾义鹏等<sup>[12]</sup>利用 PSO-GRNN 模型对岩爆进行了预测。但以上使用的算法存在诸多不足之处, Bayesian networks 训练过程过于繁琐,难以适应多因素复杂问题; BP 人工神经网络结构较为复杂,收敛速度相对较慢,并易得到局部的最优解; PSO 收敛速度快,但精度较低,容易发散。而 SVR 难以适应多分类问题,不满足埋地管道腐蚀因素多的条件。对此,本文通过 GRNN 对埋地管道外腐蚀速率进行预测,应用 BAS 优化光滑因子,建立了基于 KPCA-BAS-GRNN 的埋地管道外腐蚀速率预测模型,并在 Matlab 中对新模型进行仿真,将预测值与实际测量值对比,验证了该模型的可靠性,为腐蚀防护工作提供了理论参考依据。

## 1 KPCA 及原理

核主成分分析<sup>[13-15]</sup> (KPCA, Kernel Principal Components Analysis) 是一种非线性推广的主成分分析法。采用非线性映射方法将原样本集由数据空间映射到特征空间,进而在特征空间进行相应的线性操作。与主成分分析法<sup>[16]</sup> (PCA) 相比, KPCA 有较好的非线性数据处理能力。

设数据集  $X = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in R^p$  ( $R^p$  为输入空间,  $p$  为数据的维数,  $i=1, \dots, n$ ), 利用映射函数对输入空间进行映射  $\Phi$ , 即  $R^p \rightarrow F$ ,  $x \rightarrow X$ , 且满足  $\sum_{i=1}^m \Phi(x_i) = 0$ , 则特征空间中的协方差矩阵为:

$$C = \frac{1}{m} \sum_{i=1}^m \Phi(x_i) \Phi(x_j)^T \quad (1)$$

求  $C$  的特征值  $\lambda (\lambda \geq 0)$  及特征向量  $v$ ,  $C$  的特征向量为非负。设  $C$  的特征值为  $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ , 对应的特征向量为  $v_1, v_1, \dots, v_n$ , 记:

$$\mathbf{v}_r = \sum_{i=1}^n \alpha_r^i \Phi(x_i) \quad (2)$$

式 (1) — (2) 代入  $\Phi(x_i) \cdot C\mathbf{v} = \lambda(\Phi(x_i) \cdot \mathbf{v})$ , 令  $\mathbf{K} = (k_{ij})_{n \times n} = (\Phi(x_i) \cdot \Phi(x_j)) (i, j = 1, 2, \dots, n)$ , 得  $\mathbf{K}\alpha = n\lambda\alpha$ , 其中  $\mathbf{K}$  为核矩阵, 其特征值为  $n\lambda_i$ , 特征向量为  $\alpha_1, \alpha_2, \dots, \alpha_n$ 。取前  $m (m < n)$  个特征值对应的标准化的特征向量  $\alpha_1, \alpha_2, \dots, \alpha_m$ , 其中,  $\alpha_r = \{\alpha_r^1, \alpha_r^2, \dots, \alpha_r^m\}$ ,  $(r = 1, 2, \dots, m)$ , 对  $\Phi(x_i) (i = 1, 2, \dots, n)$  在  $\mathbf{v}_r$  投影:

$$g_r(x_j) = \sum_{i=1}^m \alpha_r^i [\Phi(x_i) \cdot \Phi(x_j)] \quad (3)$$

通常,  $\sum_{i=1}^m \Phi(x_i) = 0$  难以满足条件, 则需要进行变换, 设:

$$\tilde{\Phi}(x_i) = \Phi(x_i) - \frac{1}{m} \sum_{j=1}^m \Phi(x_j), i = 1, 2, \dots, m \quad (4)$$

则:

$$\begin{aligned} \tilde{K}_{ij} &= (\tilde{\Phi}(x_i) \cdot \tilde{\Phi}(x_j)) = K_{ij} - \frac{1}{m} \sum_{p=1}^m K_{ip} - \\ &\quad \frac{1}{m} \sum_{q=1}^m K_{qj} + \frac{1}{m^2} \sum_{p,q=1}^m K_{pq} \end{aligned} \quad (5)$$

最后, 根据具体情况, 选取式 (3) 前几个分量作为主成分。KPCA 与 PCA 的本质区别在于: 前者基于样本, 而后者基于输入向量的维数。

## 2 基于天牛须搜索算法的广义回归神经网络模型 (BAS-GRNN 模型)

### 2.1 广义回归神经网络

广义回归神经网络<sup>[17-18]</sup> (Generalized Regression Neural Network, GRNN) 是一种基于非线性回归理论的前期反馈神经网络, 是 RBF 的重要分支, 具有训练速度快、全局收敛性好、调节参数少等优点, 其结构图如图 1 所示, 包含输入层、隐含层和线性输出层。

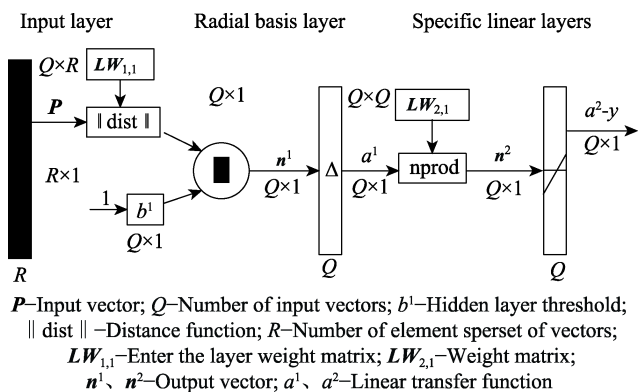


图 1 GRNN 神经网络结构示意图

Fig.1 Schematic diagram of GRNN neural network structure

在输入层, 仅将输入矢量传送给隐含层, 并不参与运算; 在隐含层, 神经元的数量等同于学习样本的

数量, 并用  $\| \text{dist} \|$  距离公式计算网络输入与第一层的权重  $LW_{1,1}$  之间的距离, 该层  $Q$  个神经元的节点函数

$$\text{为高斯函数, } p_i = \exp \left[ -\frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \right], i = 1, 2, \dots, n,$$

$n$  为样本个数,  $\sigma$  为光滑因子; 在输出层, 采用线性函数作为该层的传递函数。最终, GRNN 网络输出的预测值表达式为:  $y = a^2 = \text{purelin}(LW_{2,1} \times a^1 / \text{sum} a^1)$ 。

由以上分析可知, GRNN 网络结构简单, 其性能好坏较大幅度地依赖于光滑因子  $\sigma$ , 减少了参数选取时人为因素的影响。目前,  $\sigma$  的选取主要采用手工调整方法, 存在效率低、精度差的问题。本文引入 BAS 对 GRNN 网络进行优化, 寻求一个最优的  $\sigma$ 。

### 2.2 BAS 优化的 GRNN 模型

天牛须搜索 (beetle antennae search, BAS) 算法<sup>[19-20]</sup> 是 2017 年提出的一种受到天牛觅食原理启发而开发的智能优化算法, 其原理为: 天牛每次需要寻找食物时, 并不知道其具体位置, 而是通过食物气味的强弱来觅食。如图 2 所示, 两只长触角是天牛觅食的主要方式, 当左边触角收到的气味强度比右边大, 那下一步天牛就往左飞, 否则往右飞, 依此原理天牛可以找到食物<sup>[21]</sup>。BAS 原理由以下公式描述。

1) 创建天牛须朝向的随机向量且做归一化处理。

$$\vec{b} = \frac{\text{rands}(k, 1)}{\| \text{rands}(k, 1) \|} \quad (6)$$

式中:  $\text{rands}()$  为随机函数;  $k$  表示空间维度。

2) 创建天牛左右须空间坐标:

$$\begin{cases} x_r = x' + d_0 * \vec{b} / 2 \\ x_l = x' - d_0 * \vec{b} / 2 \end{cases} \quad (t = 0, 1, 2, \dots, n) \quad (7)$$

式中:  $x_r$  为天牛右须第  $t$  次迭代时的位置坐标;  $x_l$  为天牛左须第  $t$  次迭代时的位置坐标;  $x'$  为天牛第  $t$  次迭代时的质心坐标;  $d_0$  为两须之间的距离。

3) 根据适应度函数判断左右须气味强度, 即  $f(x_r)$  和  $f(x_l)$  的强度,  $f(x)$  函数为适应度函数。

4) 迭代更新天牛位置:

$$x^{t+1} = x' - \delta' * \vec{b} * \text{sign}(f(x_r) - f(x_l)) \quad (8)$$

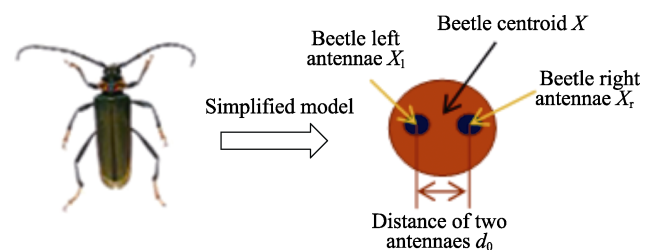


图 2 天牛简化模型

Fig.2 Beetle antennae simplified model

式中:  $\delta'$  为第  $t$  次迭代的步长因子;  $\text{sign}()$  为符号函数。

本文选取均方差作为 GRNN 模型的适应度函数, 并采用 BAS 对  $\sigma$  进行优化。结合优化过程, 相应的伪代码如下:

- 
- Initialization of algorithm parameters
    - eta=s;     % initial value
    - c=a;     % ratio between step and d0
    - step=b;   % initial step set as the largest input range
    - n=c;     % iterations
    - k=d;     % space dimension
  - Set the fitness function
    - Fitness(i)=f(spread);
  - Take the learning samples into the GRNN network to calculate the fitness value
    - net=newgrnn(A,B,spread);
    - y=sqrt(sum(A-B).^2)/N;
  - Compare the left and right antennae fitness values, update the beetle position, and calculate the fitness value under the current position to obtain the optimal value
    - If f<fbest
    - xbest=x;
    - fbest=f
    - end
  - Adjust the search step length, update the beetle's left and right antennae position, and the position of the beetle according to the initial conditions
    - xleft=x+dir\*d0/2;
    - fleft=f(xleft);
    - xright=x-dir\*d0/2;
    - fright=f(xright);
    - x=x-step\*dir\*sign(fleft-fright);
    - fbest=f(x);
  - Satisfy the iteration stop condition, then take the optimum value into GRNN network.
    - xbest=fbest;
    - net=newgrnn(P,T,xbest);
- 

综上, 基于 BAS 优化的 GRNN 预测模型流程如图 3 所示。

## 2.3 模型验证方法

为了验证 BAS-GRNN 模型的预测精度, 采用平均相对误差 (Average Relative Error, ARE) 和决定系数 (Coefficient of determination,  $R^2$ ) 2 个指标来评价模型的优劣性, 计算公式如下:

$$\text{ARE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (9)$$

$$R^2 = \frac{\left( n \sum_{i=1}^n \hat{y}_i y_i - \sum_{i=1}^n \hat{y}_i \sum_{i=1}^n y_i \right)^2}{\left[ n \sum_{i=1}^n \hat{y}_i^2 - \left( \sum_{i=1}^n \hat{y}_i \right)^2 \right] \left[ n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right]} \quad (10)$$

式中:  $n$  为样本数;  $\hat{y}_i$  为第  $i$  个样本的预测值;  $y_i$  为第  $i$  个样本的实际值。决定系数取值区间为 [0,1], 取值越靠近 1, 表明模型的性能越好, 相应地, 误差就越小。

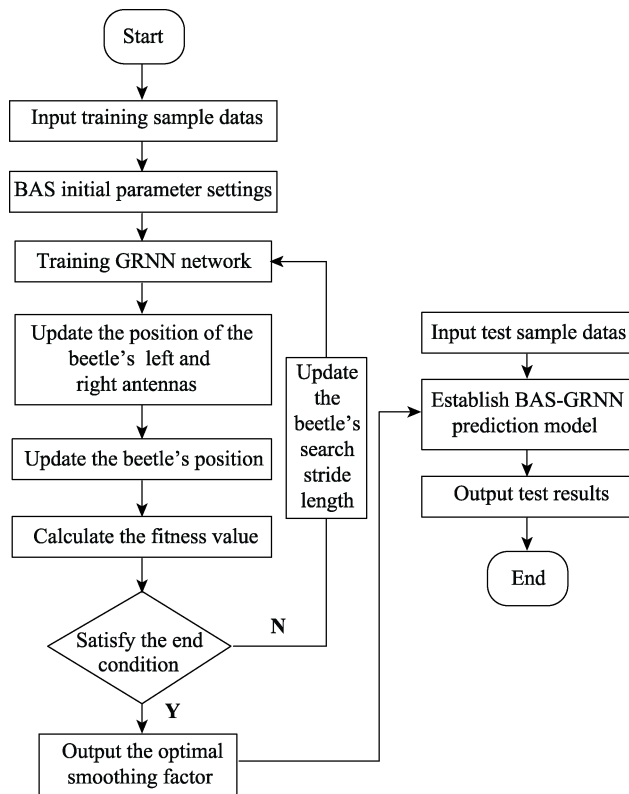


图 3 腐蚀预测流程

Fig.3 Flow chart of corrosion prediction

## 3 实例应用

### 3.1 实验与结果

川气东送是在我国西气东输工程投入使用后又一项天然气远距离管网输送工程。管道设计输量为  $1.2 \times 10^{10} \text{ m}^3/\text{a}$ , 设计输气压力为 10.0 MPa, 管径为 1016 mm, 钢管材质为 X70, 全长 2170 km, 沿线土质复杂, 腐蚀因素多, 管道外腐蚀情况严重<sup>[22]</sup>。本文采用实地埋片法, 通过试片的质量变化计算腐蚀速率, 试片表面未设置腐蚀防护措施。沿线选取监测点中的 60 组试片, 现场记录试片外观、腐蚀产物描述及土壤参数, 并采集土样。图 4a、b 分别为 5 号测试点埋地 X70 钢试片在 1 a 后的腐蚀情况和去除产物后的形貌。经过实验分析得到 60 组数据, 并确定出 12 项主要腐蚀因素作为外腐蚀指标体系, 见图 5。最后, 选取其中具有代表性的 15 组数据进行预测分析, 详见表 1。

### 3.2 核主成分分析

由于埋地管道外腐蚀影响因素的数据量纲各不相同, 首先对数据进行标准化处理, 结果见表 2。通常, KPCA 对核函数的确定无固定的形式要求, 本文选取 RBF 函数, 即  $K(x, y) = \exp(-\|x - y\|^2 / \sigma^2)$ 。对标准化后的 12 个变量进行非线性特征提取, 计算得到的累计贡献率如表 3 所示。

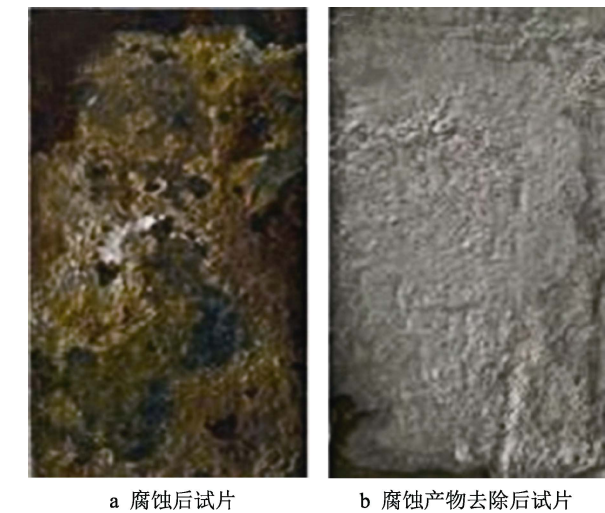


图 4 5 号测试点 X70 钢试片埋地 1 a 后的形貌  
Fig.4 Appearance diagram of X70steel specimens at No.5 test point after buried for 1 year: (a) Specimen after corrosion; (b) Specimen without corrosion product

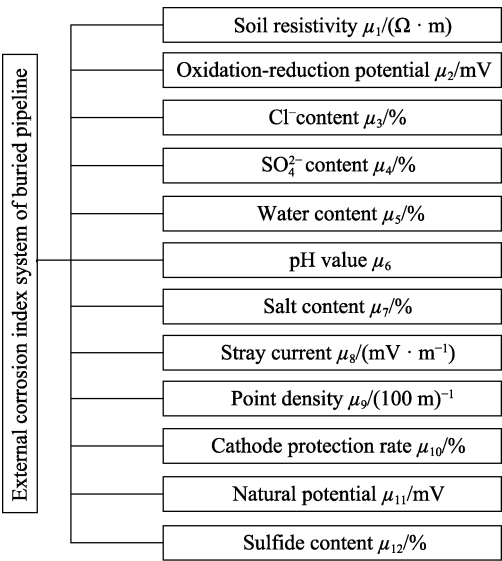


图 5 埋地管道外腐蚀指标体系  
Fig.5 External corrosion indicator system of buried pipelines

表 1 管外腐蚀点监测数据  
Tab.1 Monitoring data of corrosion points of external pipelines

Samples	$\mu_1/(\Omega \cdot m)$	$\mu_2/mV$	$\mu_3/\%$	$\mu_4/\%$	$\mu_5/\%$	$\mu_6$	$\mu_7/\%$	$\mu_8/(mV \cdot m^{-1})$	$\mu_9/(100 m)^{-1}$	$\mu_{10}/\%$	$\mu_{11}/mV$	$\mu_{12}/\%$	$V/(g \cdot dm^{-2} \cdot a^{-1})$
1	62.10	320.53	0.0141	0.0172	18.28	7.43	0.06	2.38	3.2	55.15	301	0.0091	4.2136
2	205.08	299.05	0.0127	0.0136	32.45	7.54	0.11	0.80	2.9	78.32	325	0.0211	3.1588
3	94.08	333.92	0.0128	0.0356	44.45	7.55	0.18	1.55	2.1	47.55	400	0.0101	3.2531
4	38.73	307.48	0.0242	0.0168	44.12	6.98	0.06	2.76	1.9	50.21	421	0.0078	5.2877
5	34.15	381.02	0.0121	0.0211	22.35	7.12	0.18	2.04	1.4	35.34	360	0.0035	6.0001
6	59.89	310.11	0.0191	0.0112	10.06	7.19	0.27	0.78	1.8	40.42	335	0.0184	4.7217
7	28.01	419.86	0.0202	0.0614	9.75	7.26	0.12	1.91	2.1	36.18	406	0.0097	7.2455
8	113.05	499.05	0.0173	0.0348	31.09	7.32	0.43	1.11	3.2	29.76	368	0.0086	4.8486
9	77.33	167.99	0.0154	0.0246	16.58	6.83	0.38	0.38	1.8	39.89	394	0.0006	5.8749
10	144.99	461.02	0.0132	0.0364	20.79	7.54	0.21	0.75	2.0	41.09	334	0.0045	4.1204
11	24.23	523.12	0.0142	0.0356	36.54	7.11	0.09	0.56	0.9	51.43	411	0.0264	4.3562
12	60.11	412.03	0.0149	0.0356	18.72	6.79	0.14	2.47	1.5	48.14	387	0.0012	6.3621
13	36.29	257.21	0.0112	0.0372	24.85	7.03	0.22	1.81	1.2	51.45	452	0.0171	6.1165
14	28.24	479.64	0.0194	0.0368	33.45	7.12	0.11	0.56	1.1	60.72	377	0.0039	4.6968
15	102.12	169.12	0.0274	0.0492	25.58	6.84	0.08	0.87	2.3	44.63	364	0.0058	5.0002

表 2 各因素标准化数据  
Tab.2 Standard data of various factors

Samples	$\mu_1$	$\mu_2$	$\mu_3$	$\mu_4$	$\mu_5$	$\mu_6$	$\mu_7$	$\mu_8$	$\mu_9$	$\mu_{10}$	$\mu_{11}$	$\mu_{12}$
1	-0.0177	2.1864	-0.5473	-0.5473	-0.3915	-0.4840	-0.5469	-0.5271	-0.5201	-0.0770	2.0198	-0.5473
2	1.0151	1.7733	-0.6396	-0.6395	-0.3778	-0.5788	-0.6388	-0.6332	-0.6163	-0.0077	1.9827	-0.6395
3	0.1185	1.8446	-0.5586	-0.5584	-0.2387	-0.5043	-0.5574	-0.5475	-0.5435	-0.2164	2.3202	-0.5586
4	-0.2439	1.6817	-0.5213	-0.5213	-0.2053	-0.4714	-0.5210	-0.5017	-0.5078	-0.1617	2.4951	-0.5214
5	-0.2566	2.2049	-0.4988	-0.4987	-0.3403	-0.4484	-0.4976	-0.4844	-0.489	-0.2481	2.0558	-0.4989
6	-0.0319	2.0116	-0.5209	-0.5209	-0.4389	-0.4623	-0.5188	-0.5146	-0.5063	-0.1909	2.2148	-0.5209
7	-0.3036	2.1785	-0.4809	-0.4806	-0.4193	-0.4350	-0.4803	-0.4689	-0.4677	-0.2518	2.0907	-0.4810
8	0.1515	2.4629	-0.5254	-0.5253	-0.3393	-0.4817	-0.5229	-0.5189	-0.5063	-0.3473	1.6782	-0.5254
9	0.1590	0.9356	-0.5033	-0.5032	-0.3614	-0.4449	-0.5002	-0.5002	-0.4880	-0.1617	2.8716	-0.5034
10	0.3929	2.4411	-0.5467	-0.5466	-0.4121	-0.4979	-0.5454	-0.5419	-0.5338	-0.2805	1.6179	-0.5468
11	-0.3548	2.4245	-0.4897	-0.4896	-0.2862	-0.4502	-0.4893	-0.4867	-0.4848	-0.2033	1.7999	-0.4897
12	-0.1186	2.2032	-0.5150	-0.5149	-0.3916	-0.4703	-0.5142	-0.4988	-0.5052	-0.1975	2.0381	-0.5151
13	-0.2351	1.3364	-0.4932	-0.4930	-0.3165	-0.4433	-0.4917	-0.4804	-0.4847	-0.1273	2.7220	-0.4931
14	-0.3295	2.4200	-0.5014	-0.5013	-0.2977	-0.4581	-0.5008	-0.4981	-0.4948	-0.1316	1.7948	-0.5015
15	0.3885	1.0012	-0.5451	-0.5449	-0.3114	-0.4828	-0.5446	-0.5374	-0.5243	-0.1372	2.7832	-0.5453



表3 核主成分提取后的变量

Tab.3 Extracted variables of kernel principal component

Principal components	Eigenvalue	Contribution rate/%	Cumulative contribution rate/%
1	2.7757	68.07	68.07
2	0.6670	16.36	84.42
3	0.5497	13.48	97.90
4	0.0764	1.87	100
5	0.0085	0.21	100
6	0.0005	0.01	100

由表3可知,前3项的累计贡献率已经超过了85%,因此选取前3个主成分作为埋地管道外腐蚀影响因素,设前3个主成分为 $F_1$ 、 $F_2$ 、 $F_3$ ,则表达式为:

$$\begin{cases}
 F_1 = 0.0014x_1 + 0.6385x_2 - 0.1473x_3 - \\
 \quad 0.1473x_4 - 0.1389x_5 - 0.1485x_6 - \\
 \quad 0.1473x_7 - 0.1477x_8 - 0.1479x_9 - \\
 \quad 0.1075x_{10} + 0.6395x_{11} - 0.1473x_{12} \\
 F_2 = -0.8227x_1 + 0.0562x_2 + 0.1701x_3 + \\
 \quad 0.1700x_4 - 0.0909x_5 + 0.1030x_6 + \\
 \quad 0.1687x_7 + 0.1592x_8 + 0.1500x_9 - \\
 \quad 0.3522x_{10} + 0.1185x_{11} + 0.1701x_{12} \\
 F_3 = 0.0361x_1 - 0.7107x_2 - 0.0075x_3 - \\
 \quad 0.0075x_4 + 0.0046x_5 - 0.0043x_6 - \\
 \quad 0.0075x_7 - 0.0070x_8 - 0.0065x_9 + \\
 \quad 0.0157x_{10} + 0.7021x_{11} - 0.0075x_{12}
 \end{cases} \quad (11)$$

### 3.3 BAS-GRNN 模型与其他模型对比

分析 GRNN 模型预测效果,分别选取 0.1、0.2、0.3、0.4、0.5 作为 spread 值,构建了 10 组训练值(见图6)和 5 组预测值(见图7)。结果表明,GRNN 模型的预测结果不稳定,无法实现管道腐蚀速率的精确预测,而采用 BAS 算法寻找最优 spread 值,可有效提高 GRNN 网络预测的稳定性和精确性。

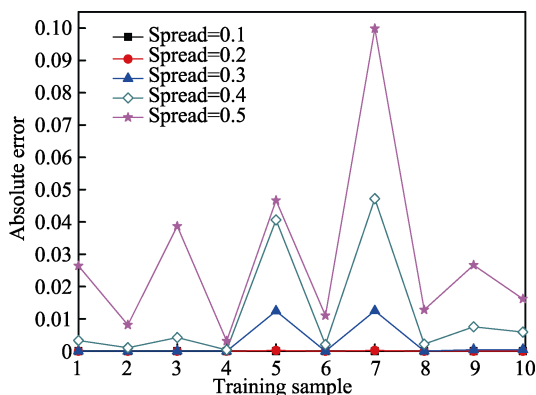


图6 GRNN 网络训练误差效果图

Fig.6 Training error effect of GRNN network

由式(11)可得到 KPCA 重构的 15 组数据,设置 BAS 参数:初始步长  $\delta^0 = 3$ ,迭代次数  $n = 100$ ,并将提取后的 15 组数据分为训练样本集(样本序号为 1—

10)和测试样本集(样本序号为 11—15),分别代到 BAS-GRNN 模型进行训练与预测。由图8结果可知,BAS 寻优后的最小 RMSE 为 0.020 42,迭代 54 次达到收敛状态,并且 spread 取值为 0.6747。

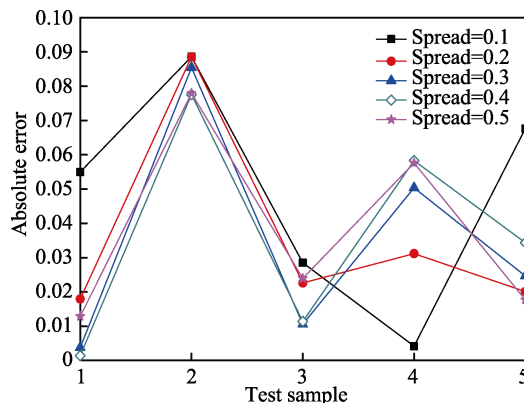


图7 GRNN 网络预测误差效果图

Fig.7 Prediction error effect of GRNN network

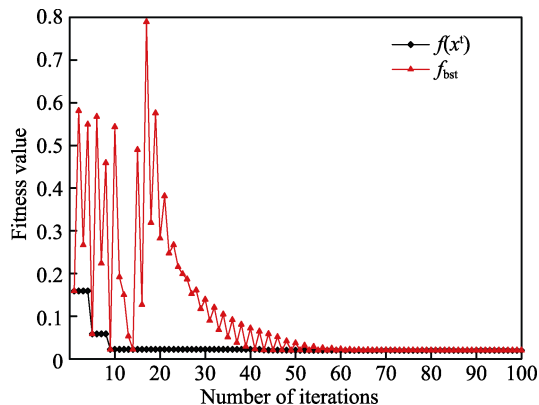


图8 BAS-GRNN 模型迭代过程

Fig.8 Iterative process of BAS-GRNN model

为验证 KPCA-BAS-GRNN 模型预测的准确性,将输出结果反归一化处理,并与文献[10]中的 BP 网络和文献[23]中的 PCA-SVM 模型进行对比,三种模型的训练结果如图9。为直观表达模型预测效果,用残差表示实际值与预测值的差,结果见图10。三种模型的预测结果对比及模型性能评价见表4。

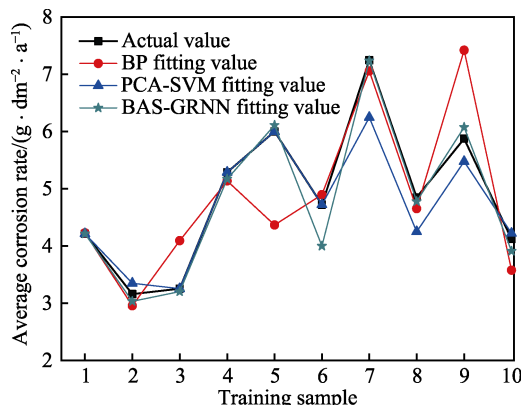


图9 模型训练值对比

Fig.9 Comparison of model training value

表 4 预测结果相对误差分析  
Tab.4 Relative error analysis of prediction results

Test samples	Average corrosion rate/ (g·dm <sup>-2</sup> ·a <sup>-1</sup> )	BP model				PCA-SVM model				KPCA-BAS-GRNN model			
		Prediction value	RE/%	ARE/%	R <sup>2</sup>	Prediction value	RE/%	ARE/%	R <sup>2</sup>	Prediction value	RE/%	ARE/%	R <sup>2</sup>
1	4.3562	3.9687	8.90			4.0299	7.49			4.1435	4.88		
2	6.3621	5.6355	11.42			6.9524	9.28			6.8602	7.83		
3	6.1165	6.7012	9.56	9.84	0.81	5.7008	6.80	8.09	0.84	6.0765	0.65	5.21	0.93
4	4.6968	4.0645	13.46			4.3529	7.32			5.0133	6.74		
5	5.0002	4.7084	5.84			5.4792	9.58			4.7026	5.95		

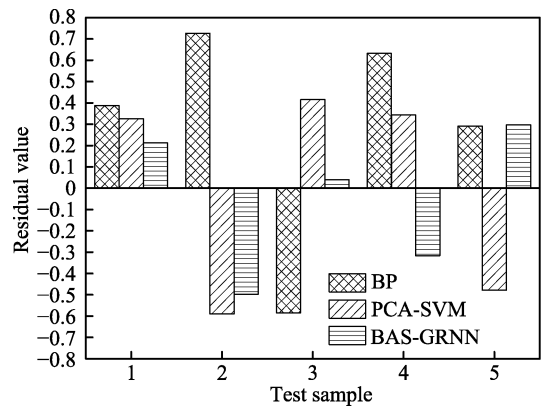


图 10 模型预测值残差对比  
Fig.10 Comparison of model prediction residual

从表 4 可以看出，BP、PCA-SVM、KPCA-BAS-GRNN 的平均相对误差分别为 9.84%、8.09%、5.21%，BAS-GRNN 拟合结果更好，KPCA-BAS-GRNN 模型的决定系数达到 0.93，模型性能得到很大的提高，预测效果较好。

4 结论

- 1) 本文分析了影响管道腐蚀的关键因素，建立了埋地管道外腐蚀预测指标体系,通过 KPCA 对指标因子进行降维，提取出包含原始信息 97.9%的三个变量，降低了输入向量的维数，简化了分析过程。
- 2) 提出经 BAS 优化的 GRNN 预测模型，BAS 优化后的 GRNN 参数克服了光滑因子  $\sigma$  难以选取的问题，大大提高了预测精度和迭代速率。基于 BAS-GRNN 预测模型用于川气东送埋地管段的外腐蚀速率预测上，平均相对误差为 5.21%，可靠性较高。
- 3) 本模型选用 KPCA 对数据进行预处理，为非线性问题的研究提供了参考，并且将天牛须搜索 (BAS)算法用于管道腐蚀预测上，是一种新的尝试，为埋地管道的腐蚀研究拓宽了思路。同时，本文还存在有待进一步提高的工作，如何改进 BAS 初始步长的选取，解决过早收敛的问题，是下一步研究的方向。

参考文献：

[1] VANAEIH R, ESLAMI A, EGBEWANDE A. A review on

pipeline corrosion, in-line inspection (ILI), and corrosion growth rate models[J]. International journal of pressure vessels and piping, 2017, 149: 43-54.

[2] 李斌, 邢希金, 张鑫, 等. 酸性环境下溶解氧对低合金管材点蚀的影响[J]. 表面技术, 2017, 46(3): 246-252.

LI Bin, XING Xi-jin, ZHANG Xin, et al. Effects of dissolved oxygen on pitting of low alloy tubes in acidic environment[J]. Surface technology, 2017, 46(3): 246-252.

[3] 李兆玲, 杨任继, 陈浩. 氯离子和直流电流密度对 X65 钢的腐蚀研究[J]. 表面技术, 2017, 46(8): 254-258.

LI Zhao-ling, YANG Ren-ji, CHEN Hao. Corrosion of X65 steel under influence of Cl<sup>-</sup> and direct current density[J]. Surface technology, 2017, 46(8): 254-258.

[4] 谭才渊, 殷启帅, 杨进, 等. 渤海某油田 L80 油管腐蚀机理研究[J]. 表面技术, 2017, 46(3): 236-245.

TAN Cai-yuan, YIN Qi-shuai, YANG Jin, et al. Corrosion mechanism of L80 tubing in a Bohai oilfield[J]. Surface technology, 2017, 46(3): 236-245.

[5] VELAZQUEZ J C, CALEYO F, VALOR A, et al. Predictive model for pitting corrosion in buried oil and gas pipelines[J]. Corrosion, 2009, 65(5): 332-342.

[6] 毕傲睿, 骆正山, 王小完, 等. 基于土壤腐蚀主成分的金属管道退化维纳过程研究[J]. 材料保护, 2018, 51(1): 37-42.

BI Ao-rui, LUO Zheng-shan, WANG Xiao-wan, et al. Wiener process of degradation of metal pipelines based on principal components of soil corrosion[J]. Materials protection, 2018, 51(1): 37-42.

[7] 陈健飞, 王孟, 江文军, 等. 基于灰色关联度的埋地管道检测数据分析[J]. 表面技术, 2017, 46(4): 264-269.

CHEN Jian-fei, WANG Meng, JIANG Wen-jun, et al. Buried pipeline detection data analysis based on grey correlation[J]. Surface technology, 2017, 46(4): 264-269.

[8] 张杰. 基于主成分-聚类分析法的管道风险评价方法[J]. 油气储运, 2014, 33(2): 139-143.

ZHANG Jie. Pipeline risk assessment method based on principle component-clustering analysis[J]. Oil & gas storage and transportation, 2014, 33(2): 139-143.

[9] JAIN S, SANCHEZ A N, GUAN S, et al. Probabilistic assessment of external corrosion rates in buried oil and gas pipelines[C]//2015 International Corrosion Conference Series. Naperville: NACE, 2015.

[10] 李丽, 李晓刚, 邢士波, 等. BP 神经网络对国内典型地区碳钢土壤腐蚀的预测研究[J]. 腐蚀科学与防护技术, 2013, 25(5): 372-376.

LI Li, LI Xiao-gang, XING Shi-bo, et al. Research on soil corrosion rate prediction of carbon steel in typical Chinese

- cities based on BP artificial neural network[J]. Corrosion science and protection technology, 2013, 25(5): 372-376.
- [11] MENG Q, MA X, ZHOU Y. Forecasting of coal seam gas content by using support vector regression based on particle swarm optimization[J]. Journal of natural gas science & engineering, 2014, 21(21): 71-78.
- [12] 贾义鹏, 吕庆, 尚岳全. 基于粒子群算法和广义回归神经网络的岩爆预测[J]. 岩石力学与工程学报, 2013, 32(2): 343-348.
- JIA Yi-peng, LV Qing, SHANG Yue-quan. Rock burst prediction using particle swarm optimization algorithm and general regression neural network[J]. Chinese journal of rock mechanics and engineering, 2013, 32(2): 343-348.
- [13] WANG Hai-xian, HU Zi-lan. An unified EM algorithm for PCA and KPCA[J]. Neurocomputing, 2007, 71(1): 459-462.
- [14] 程砚秋, 迟国泰. 基于核主成分分析的生态评价模型及其应用研究[J]. 中国管理科学, 2011, 19(3): 182-192.
- CHENG Yan-qiu, CHI Guo-tai. The ecosystem evaluation model based on kernel principle component analysis and its applications[J]. Chinese journal of management science, 2011, 19(3): 182-192.
- [15] 张钦礼, 王兢, 王新民. 基于核主成分分析法与 PSO-SVW 的填充管道失效风险性分级评价模型[J]. 黄金科学技术, 2017, 25(3): 70-76.
- ZHANG Qin-li, WANG Jing, WANG Xin-min. Failure risk assessment model of filling pipeline based on KPCA and PSO-SVW[J]. Gold science and technology, 2017, 25(3): 70-76.
- [16] 王晓丽, 魏志兵, 彭士涛, 等. 基于主成分分析法的液体管道泄漏后果综合评价模型[J]. 中国安全生产科学技术, 2014, 10(5): 84-89.
- WANG Xiao-li, WEI Zhi-bing, PENG Shi-tao, et al. Comprehensive assessment model of liquid pipeline leakage consequences based on principle component analysis[J]. Journal of safety science and technology, 2014, 10(5): 84-89.
- [17] WANG Shao-fu, ZHANG Jin-lei, ZHAO Shi-jun, et al. The generalized regression network and its application to predictive control[J]. Microelectronics & computer, 2009, 26(6): 32-35.
- [18] 林喆, 兰生, 张宇航. 基于广义回归神经网络的油纸绝缘变压器的寿命预测[J]. 高压电器, 2015, 51(2): 125-130.
- LIN Zhe, LAN Sheng, ZHANG Yu-hang. Life prediction of oil paper insulation of transformer based on GRNN neural network[J]. High voltage apparatus, 2015, 51(2): 125-130.
- [19] JIANG X Y, LI S. BAS: beet lean tennae search algorithm for optimization problems[J]. Neural and evolutionary, 2017, 1710: 10724.
- [20] JIANG X Y, LI S. Beetle antennae search without parameter tuning(BAS-WPT) for multi-objective optimization[J]. Neural and evolutionary, 2017, 1711: 02395.
- [21] 王甜甜, 刘强. 基于 BAS-BP 模型的风暴潮灾害损失预测[J]. 海洋环境科学, 2018, 37(3): 457-463.
- WANG Tian-tian, LIU Qiang. The assessment of storm surge disaster loss based on BAS-BP model[J]. Marine environmental science, 2018, 37(3): 457-463.
- [22] 任帅, 张琪, 王东. 基于主分量法的管道腐蚀评价——以川气东送管道为例[J]. 油气储运, 2015, 34(5): 519-523.
- REN Shuai, ZHANG Qi, WANG Dong. Pipeline corrosivity assessment based on principal component analysis-take the Sichuan-east gas transmission pipeline as an example[J]. Oil & gas storage and transportation, 2015, 34(5): 519-523.
- [23] 骆正山, 毕傲睿, 王小完. 基于 PCA-SVM 的高含硫油气混输管道腐蚀预测[J]. 中国安全科学学报, 2016, 26(2): 85-90.
- LUO Zheng-shan, BI Ao-rui, WANG Xiao-wan. Corrosion prediction of high sulfur gas-oil mixed transmission pipelines based on PCA-SVM[J]. China safety science journal, 2016, 26(2): 85-90.