

# 基于 Isomap+SVR 的瓦斯浓度预测方法

吴海波<sup>1,2\*</sup>, 施式亮<sup>3</sup>, 念其锋<sup>2</sup>

(1.中南大学 资源与安全工程学院,湖南长沙 410082;

2.湖南科技大学 煤矿安全开采技术湖南省重点实验室,湖南湘潭 411201;

3.湖南科技大学 资源环境与安全工程学院,湖南湘潭 411201)

**摘要:**对煤矿关键测点的瓦斯浓度进行科学且准确的预测,是防范瓦斯灾害的关键.为了从瓦斯监测监控系统实时采集的煤矿安全环境数据中分析和挖掘瓦斯浓度信息,从而进行关键测点的瓦斯浓度预测,本文采用等度量映射算法(Isomap)结合支持向量回归算法(SVR)来预测瓦斯浓度.该方法首先通过 Isomap 算法将非线性的高维煤矿井下安全环境数据进行维数约减,然后利用 SVR 算法进行回归预测.通过实验分析与对比,该方法行之有效,与多元线性回归(MLR)、支持向量回归(SVR)方法相比,在预测精度上有一定的优势,且在瓦斯波动异常情况下,鲁棒性更强.

**关键词:**瓦斯灾害;瓦斯浓度;预测;Isomap;SVR

中图分类号:TD712

文献标志码:A

文章编号:1672-9102(2019)02-0051-04

## Prediction of Gas Concentration Based on Isomap and SVR

Wu Haibo<sup>1,2</sup>, Shi Shiliang<sup>3</sup>, Nian Qifeng<sup>2</sup>

(1. School of Resources and Safety Engineering, Central South University, Changsha 410082, China;

2. Hunan Provincial Key Laboratory of Safe Mining Techniques of Coal Mines, Xiangtan 411201, China;

3. School of Resources, Environment and Safety Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

**Abstract:** The key to prevent gas disasters is to scientifically and accurately predict the gas concentration of key measuring points in coal mines. In order to analyze and mining the gas concentration information from the real-time data of coal mine safety environment collected by the gas monitoring system, and then predict the gas concentration of key measuring points, this paper proposes a method based on Isomap and SVR to predict the gas concentration. This method first reduces the non-linear high-dimensional mine safety environment data by Isomap, and then uses SVR algorithm for regression prediction. Through experimental analysis and comparison, this method has some advantages over MLR and SVR in prediction accuracy, and has stronger robustness in case of abnormal gas fluctuation.

**Keywords:** gas disasters; gas concentration; prediction; Isomap; SVR

矿井瓦斯是煤矿安全生产过程中的最不安全因素<sup>[1]</sup>,因此瓦斯事故一直是煤矿重大事故源.为了提高煤矿安全,预防重特大事故,根据煤矿安全规程规定,瓦斯矿井必须安装瓦斯监测监控系统以减低瓦斯风险,达到瓦斯风险可预可控.此外,规定还要求瓦斯监测监控系统必须 24 h 运行,巡检周期不大于 30 s,各种安全环境监测数据必须保存至少 2 年等.因此,如何从海量的安全环境监测数据中分析与挖掘与瓦斯相关的安全信息,成为了预防瓦斯灾害的关键.

收稿日期:2019-04-25

基金项目:国家自然科学基金资助项目(51774135);煤矿安全开采技术湖南省重点实验室开放基金资助项目(201304);湖南科技大学煤炭资源清洁利用与矿山环境保护湖南省重点实验室开放基金资助项目(E21701);湖南省高等学校科学研究优秀青年资助项目(14B058);湖南省普通高等学校教学改革研究资助项目(2017-237)

\*通信作者,E-mail: hbwu73@sina.com

为了预防瓦斯风险,对煤矿关键点的瓦斯浓度进行预测,从而进行相应的安防措施,一直是预防瓦斯灾害的主要技术手段,这也使得如何进行瓦斯浓度的科学预测成为了煤矿安全研究的主要内容.现阶段瓦斯浓度预测研究方法有:郭思雯、董丁稳等的基于时间序列的回归模型的瓦斯浓度动态预测<sup>[2-4]</sup>;张昭昭、曹博等基于神经网络预测瓦斯浓度<sup>[5,6]</sup>;付华等应用优化后的支持向量机(SVM)来进行瓦斯浓度的预测<sup>[7,8]</sup>等.这些研究方法通过分析监测数据之间的关联关系来进行瓦斯浓度的预测,为瓦斯风险预警研究提供了基础.事实上,瓦斯浓度的预测不仅与自身历史数据有关,而且还会受到巷道内风速、风量、温度、CO含量以及巷道内相邻点瓦斯浓度等历史数据的影响<sup>[9]</sup>.为了更好地挖掘煤矿井下安全环境数据与瓦斯监测数据之间的非线性关系,提高瓦斯浓度的预测精度,本文提出了一种基于非线性降维的等度量映射算法(Isomap)结合支持向量回归算法(SVR)来进行瓦斯浓度预测,该方法首先通过 Isomap 将与瓦斯浓度相关的具有非线性关系的安全环境数据降维,然后通过 SVR 算法来进行回归预测.通过实验对比分析,该方法行之有效,且预测精度有一定的提高,有助于瓦斯风险的预测预警.

## 1 关键算法

### 1.1 等度量映射算法 Isomap

Isomap 算法是一种非线性无监督降维算法<sup>[10]</sup>,是将非线性结构高维数据低维表示的方法.Isomap 算法首先采用流形空间的最短路径来求得近似测地距离,然后输入到多维尺度分析(MDS)中处理,进而发现嵌入在高维空间的低维坐标<sup>[11]</sup>.Isomap 的算法描述如下:

1) 输入:高维( $m$ 维)空间样本集  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ ; 近邻参数  $y$ ; 低维空间维数  $k(k < m)$ ;

2) 输出:样本集  $\mathbf{X}$  在  $k$  维空间的投影  $\mathbf{Z} = \{z_1, z_2, \dots, z_n\}$ ;

3) 过程:

① for  $i = 1, 2, \dots, n$  do;

② 确定  $x_i$  的  $y$  个近邻点;

③  $x_i$  与  $y$  个近邻点的距离为欧氏距离  $\Omega_{ij} = \sqrt{\sum_{t=1}^m (x_{it} - x_{jt})^2}$ , 与其他点距离设置为  $+\infty$ ;

④ end for;

⑤ 调用最短路径算法计算任意两样本点的距离  $d_{ij}$ ;

⑥ 生成距离矩阵  $\mathbf{D} = [d_{ij}]_{n \times n}$ ;

⑦ 构造矩阵  $\mathbf{A} = [a_{ij}]_{n \times n} = \left[ -\frac{1}{2} d_{ij}^2 \right]_{n \times n}$ ;

⑧ 计算内积矩阵  $\mathbf{B} = [b_{ij}]_{n \times n} = [a_{ij} - \bar{a}_{i.} - \bar{a}_{.j} + \bar{a}_{..}]_{n \times n}$ ;

⑨ 求  $\mathbf{B}$  的特征值和特征向量;

⑩ 构造矩阵  $\mathbf{Z} = [E_k \mathbf{A}_K^{1/2}]_{n \times k}$ , 其中  $E_k$  为矩阵  $\mathbf{B}$  的前  $k$  个特征向量组成的矩阵,  $\mathbf{A}_K$  是  $k$  个特征值组成的对角矩阵;

⑪: return  $\mathbf{Z}_{n \times k}$ .

### 1.2 支持向量回归算法 SVR

支持向量机(SVM)是将样本点“升维”,即将样本空间非线性数据映射到线性可分的高维空间中,解决样本空间中的高维度非线性问题<sup>[12]</sup>.SVR 则是 SVM 常用应用形式之一,用于解决回归问题.SVR 算法是将高维特征空间  $R^k$  通过转换函数  $\varphi$  映射到低维空间  $F$  中,再在  $F$  维中进行线性回归.

给定  $k$  维变量和 1 维目标变量的训练样本数据  $\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$ , 其中  $x \in R^k, y \in R$ . SVM 中的回归函数可定义为

$$f(x) = \omega \varphi(x) + b. \quad (1)$$

式中:  $\omega$  为回归系数;  $b$  为偏置变量;  $\varphi(x)$  是将非线性高维特征空间  $x$  映射到与  $\omega$  同维的函数.令  $d = f(x)$  为  $y$  的预测值,则损失函数定义为

$$L(y, d) = |y - d|. \quad (2)$$

SVR 是在高维特征空间中利用  $\varepsilon$ -不敏感损失进行线性回归,  $\varepsilon$ -不敏感损失函数为

$$L_\varepsilon(y, d) = \begin{cases} |y - d| - \varepsilon, & \text{当 } |y - d| \geq \varepsilon; \\ 0, & \text{否则.} \end{cases} \quad (3)$$

则在给定的训练样本中, 风险函数定义为

$$\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^k L_{i,\varepsilon}(y_i, d_i). \quad (4)$$

式中:  $C$  为权重.

为了最小化风险函数, 定义松弛向量  $\xi_i$  和  $\xi_i^*$  来描述  $\varepsilon$ -不敏感损失函数  $L_{i,\varepsilon}(y_i, d_i)$ , 则最小化风险函数可以写为一个凸二次优化问题:

$$\text{最小化: } \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*). \quad (5)$$

$$\text{约束条件: } \begin{cases} y_i - d_i \leq \varepsilon + \xi_i; \\ d_i - y_i \leq \varepsilon + \xi_i^*; \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, k. \end{cases} \quad (6)$$

为解决上述问题, 通过构造拉格朗日函数, 将其化为极小极大问题, 然后改写成为对偶问题, 通过对函数求偏导, 求解拉格朗日乘子  $\alpha_i$  和  $\alpha_i^*$ , 可得式(1)的解:

$$\omega = \sum_{i=1}^k (\alpha_i - \alpha_i^*) \varphi(x_i); \quad (7)$$

$$f(x) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) K(x_i, x) + b. \quad (8)$$

式中:  $0 \leq \alpha_i \leq C, 0 \leq \alpha_i^* \leq C, K(x_i, x)$  为核函数.

## 2 实验与分析

为验证结合 Isomap 算法和 SVR 算法来预测瓦斯浓度方法的有效性和精确度, 本次实验从河南某煤矿瓦斯监测监控系统中提取 4 d 的工作面实时安全环境数据作为实验数据集. 瓦斯监测监控系统采集的安全环境数据包括瓦斯浓度、一氧化碳浓度、风速、负压和温度等共计 35 项, 巡检周期为 5 min, 实验利用这些安全环境数据来预测工作面上隅角瓦斯浓度. 工作面上隅角瓦斯浓度 4 d 共计 1 152 条瓦斯浓度实测值如图 1 所示.

实验采用 1 h 共 12 条安全环境数据作为训练数据集, 用来预测下次巡检周期的上隅角瓦斯浓度值. 首先, 将安全环境训练数据集的数据通过 Isomap 进行线性降维, 然后将降维后的数据作为 SVR 模型的训练数据集, 求出相对应的 SVR 模型, 最后, 利用 SVR 模型与当前安全环境数据来预测下次巡检周期的瓦斯浓度. 图 2 为利用 Isomap+SVR 方法来预测上隅角瓦斯的结果, 其平均绝对误差 (MAE) 为 0.023 784.

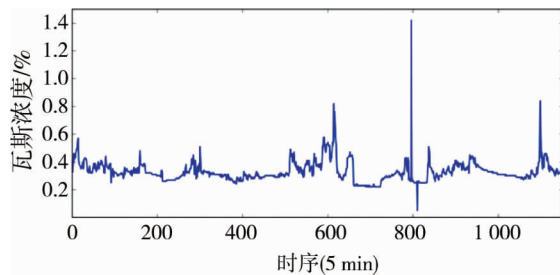


图1 上隅角瓦斯浓度实测值

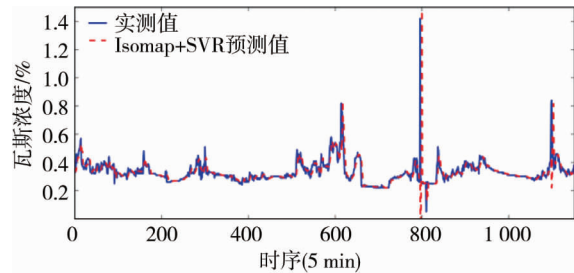


图2 Isomap+SVR方法预测值与实测值对比

为了进行对比验证, 实验分别利用多元线性回归 (MLR)、支持向量回归 (SVR) 方法来进行瓦斯浓度预测, 实验结果如图 3、图 4 所示. 经统计, MLR, SVR 的 MAE 分别为 0.029 103 与 0.024 499. 从图 2~图 4 的

预测值与实测值对比可以看出,无论采用何种预测方法,瓦斯浓度预测值与实测值的 MAE 均小于 0.03,但 Isomap+SVR 的 MAE 最小,预测效果最佳.

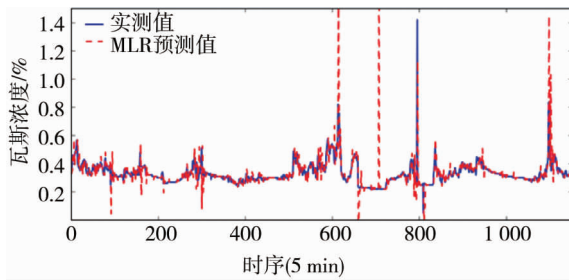


图3 MLR方法预测瓦斯浓度值与实测值对比

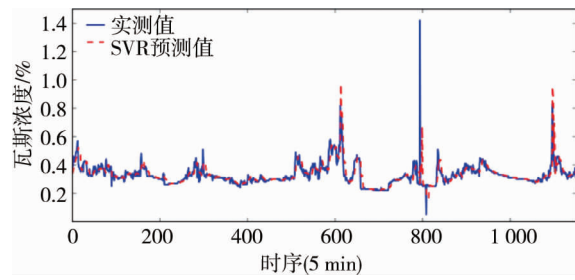


图4 SVR方法预测瓦斯浓度值与实测值对比

表1为各预测模型预测的平均绝对百分比误差(MAPE)和均方误差(MSE)值.从表1中可以看出,通过 Isomap+SVR 方法预测瓦斯浓度的 MAPE 最小,MSE 也较 MLR 更优,但略大于 SVR,具有较好的预测精度.

表1 各模型预测效果对比表

预测模型	平均绝对百分比误差(MAPE)	均方误差(MSE)
多元线性回归(MLR)	8.645 468	0.008 815
支持向量回归(SVR)	6.967 863	0.003 189
Isomap+SVR	6.737 586	0.004 904

此外,基于 Isomap+SVR 方法在提升预测精度的同时,对瓦斯异常波动预测鲁棒性也更好.从图2~图4对比分析可以看出,MLR方法在瓦斯浓度发生异常后的预测精度波动大,导致某个巡检周期的预测值与实测值有较大偏差;而SVR方法对瓦斯异常的预测不敏感,并没有有效的反映瓦斯异常的大小变化;相对而言,Isomap+SVR对异常预测的鲁棒性较其他方法要优.

### 3 结论

1) 基于 Isomap+SVR 的瓦斯浓度预测方法利用非线性降维算法 Isomap 将高维非线性的煤矿井下安全环境数据映射到低维数据中,然后再利用 SVR 算法进行回归预测,既挖掘了井下安全环境数据的非线性特征,同时又保留了环境数据与瓦斯浓度之间的回归关系,提高了瓦斯预测的精确度.

2) 基于 Isomap+SVR 的瓦斯浓度预测方法在井下安全环境数据出现异常时,能保持预测的稳定性,能准确实时的预测瓦斯异常波动大小,预测的鲁棒性较强,非常有利于瓦斯风险的有效识别.

#### 参考文献:

- [1] 谢东海,何爱宝.回采工作面瓦斯涌出量的聚类分析及归类预测模型的建立[J].矿业工程研究,2017(1):29-33.
- [2] 郭思雯,陶玉帆,李超.基于时间序列的瓦斯浓度动态预测[J].工矿自动化,2018,44(9):20-25.
- [3] 董丁稳,刘洁,王红刚.矿井瓦斯浓度自适应预测及其预警应用[J].中国安全科学学报,2013(5):88-93.
- [4] 吴海波,施式亮,念其锋.基于 SparkStreaming 流回归的煤矿瓦斯浓度实时预测[J].中国安全生产科学技术,2017(5):84-89.
- [5] 张昭昭,乔俊飞,余文.基于动态神经网络的瓦斯浓度实时预测方法[J].控制工程,2016,23(4):478-483.
- [6] 曹博,白刚,李辉.基于 PCA-GA-BP 神经网络的瓦斯含量预测分析[J].中国安全生产科学技术,2015(5):84-90.
- [7] 付华,丰盛成,刘晶,等.基于 DE-EDA-SVM 的瓦斯浓度预测建模仿真研究[J].传感技术学报,2016(2):285-289.
- [8] 付华,瞿海,孟祥云,等.一种 EKF-WLS-SVR 与混沌时间序列分析的瓦斯动态预测新方法[J].传感技术学报,2015(1):126-131.
- [9] 杨丽,刘晖,毛善君,等.基于多元分布滞后模型的瓦斯浓度动态预测[J].中国矿业大学学报,2016(3):455-461.
- [10] 李香元,蔡骋,何进荣.基于密度缩放因子的 ISOMAP 算法[J].计算机科学,2018,45(7):207-213.
- [11] 赵连伟,罗四维,赵艳敞,等.高维数据流形的低维嵌入及嵌入维数研究[J].软件学报,2005(8):1423-1430.
- [12] 邓军,雷昌奎,曹凯,等.煤自燃预测的支持向量回归方法[J].西安科技大学学报,2018,38(2):175-180.