

基于粗糙集的支持向量机地下水质量评价模型

黄 鹤¹,梁秀娟^{1*},肖 霄¹,邱淑伟¹,肖长来¹,王 重²(1.吉林大学环境与资源学院,地下水资源与环境教育部重点实验室,吉林 长春 130021; 2.吉林森工开发建设集团有限公司,吉林 长春 130000)

摘要: 以辽宁绥中县第四系松散岩类孔隙水的 10 组水质监测数据为基础,选取 pH 值、Cl⁻、SO₄²⁻、NH₄⁺、NO₂⁻、NO₃⁻、F⁻、总硬度、总溶解固体等 14 项水质评价指标,采用粗糙集对指标进行约简,将基于属性依赖度和信息熵的启发式算法结合,获得属性约简集,应用支持向量机分别评价约简前后的地下水质量.结果表明,属性约简将 14 项水质指标精简为 8 项,水质评价结果与约简前保持一致,区域地下水普遍在 III 类以上,部分地区铁、“三氮”等超标,不适宜饮用.粗糙集和支持向量机的联合应用,在保证分类能力的前提下有效地减少冗余指标,降低运算维度,保证水质评价的合理性.

关键词: 粗糙集; 支持向量机; 地下水; 水质评价

中图分类号: X824 **文献标识码:** A **文章编号:** 1000-6923(2016)02-0619-07

Model of groundwater quality assessment with support vector machine based on rough set. HUANG He¹, LIANG Xiu-juan^{1*}, XIAO Xiao¹, QIU Shu-wei¹, XIAO Chang-lai¹, WANG Zhong² (1.Key Laboratory of Groundwater Resources and Environment, Ministry of Education, College of Environment and Resources, Jilin University, Changchun 130021, China; 2.Jilin Forest Industry Development and Construction Group Corporation Limited, Changchun 130000, China). *China Environmental Science*, 2016,36(2): 619-625

Abstract: A total of 10 quaternary loose rock pore water samples were collected from Suizhong County, Liaoning. The pH, Cl⁻, SO₄²⁻, NH₄⁺, NO₂⁻, NO₃⁻, F⁻, total hardness, total dissolved solids, iron, manganese, zinc, cyanide and volatile phenols were considered as the water quality parameters. Rough set theory was employed for data reduction. Meanwhile, to find attribute reduction set, the attribute dependence degree and information entropy heuristic algorithms were combined. Support vector machine was employed to evaluate groundwater quality for all parameters before and after reduction, respectively. The results showed that rough set theory reduced the number of chemical parameters from 14 to 8, and assessment results with attribute reduction were the same as those without attribute reduction. The groundwater quality in the study area was mainly class II and III, which meets the permissible limits. However, iron and three nitrogen were exceeded drinking water quality standard. Although the combination of rough set and support vector machine reduced redundant indices, the accuracy of water quality classification remained effective, while the complexity of calculation was reduced and the rationality of assessment results was guaranteed.

Key words: rough set; support vector machine; groundwater; water quality assessment

地下水污染具有过程缓慢、不易发现和难以治理的特点,地下水一旦受到污染,水质复原一般需要十几年甚至几十年.在预防和治理地下水污染过程中,需进行合理的水质评价^[1].目前常用的水质评价方法除加附注评分法外还有主成分分析法^[2]、因子分析法^[3]、层次分析法^[4]、模糊理论法^[5]、神经网络法^[6]等.由于水质指标众多,污染源多样,无论单独选用哪种评价方法,都会面临处理大量水质监测数据的问题,就特定研究区而言,

考虑所有的指标和污染源不仅识别效率低且效果不佳^[7].因此,如何方便、快速、有效地精简数据,确保水质评价结果准确已成为众多环境学者研究的热点^[8-9].科学的识别方法是进行属性约简,降低运算维度.Pai 等^[10]选择 9 项影响人类健

收稿日期: 2015-08-25

基金项目: 吉林省科技攻关项目(20100452);吉林省自然科学基金(20140101164JC);吉林省水资源专项(0773-1441GNJL00390)

* 责任作者, 教授, lax64@126.com

康和环境的水质因素,结合回归分析法,对指标进行属性约简并提出决策规则,将规则约简了62.5%.Li 等^[11]针对干旱地区的水质进行了约简,并耦合 TOPSIS 法,快速有效地评价区域水质.

粗糙集不依靠附加信息和先验知识,仅从数据集本身出发,揭示数据间存在的规律,有效地刻画出数据中的关键因素.近年来,因其具有模糊处理多分类问题的能力和独特的数据分析观点而备受关注^[10],在数据的特征选择、特征提取、决策支持与分析、机器学习、数据挖掘等领域均有成功的应用^[12].支持向量机借助核函数,以结构风险最小化为原则,成功的将非线性分类转变为高维空间中的线性问题^[13],其在文本分类、遥感图像分析、人脸图像识别、信号处理、语音识别等领域均表现出优越的分类性能^[14].目前,粗糙集与支持向量机的联合应用已在很多领域展开,在水质分析方面,学者们的研究主要集中在河湖、水产养殖等地表水的水质预警和预测^[15-16],将二者共同应用于地下水质量评价的实例较少.本研究将粗糙集和支持向量机联合应用于地下水质量分析,在粗糙集属性约简的基础上,充分利用支持向量在非线性数据集分类能力,合理地评价地下水质量,为地下水质量评价的研究提供一种新的方法.

1 研究方法

1.1 粗糙集

粗糙集(Rough Set,RS)由 Pawlak 于 20 世纪 80 年代提出,是一种刻画不完备性和不确定性的数学工具^[17].该理论以代数学的等价关系和集合运算为基础,定义信息系统和可辨识矩阵,通过寻求信息系统的约简集来获得数据的决策规则.

粗糙集以信息系统为研究对象,表示为:

$$IS = (U, A, V, f) \tag{1}$$

式中: U 表示论域; A 表示属性集, $A = C \cup D$, C 为条件属性, D 为决策属性; $V = \bigcup_{a \in A} V_a$, 表示属性值的集合; f 为信息函数, $f : U \times A \rightarrow V, f(x, a) \in V_a$, 指定了论域 U 中每一个对象 x 的属性值^[18].信息系统中的对象由于缺乏知识而无法被准确地分类,则称二者之间具有不可辨识关系,表示为:

$I(B) = \{(x_i, x_j) | (x_i, x_j) \in U^2\}$, 根据不可辨识关系导出论域的等价划分,即 $UI(B)$.

根据对象间的不可辨识关系,可得到信息系统的可辨识矩阵,设属性集的维数为 n ,则可辨识矩阵为 $n \times n$ 阶的对称矩阵,矩阵中任一元素表示为 $\alpha(x, y) = \{a \in A | f(x, a) \neq f(y, a)\}$.对非空元素进行布尔合取,得到可辨识函数 $f(A)$ ^[18]:

$$f(A) = \prod_{x, y \in (U \setminus \text{Ind}(A))} \sum \alpha(x, y) \tag{2}$$

式中: $\sum \alpha(x, y) = a_1 \vee a_2 \vee \dots \vee a_k \sum \alpha(x, y)$ 表示元素的布尔合取,

若去除属性 a_i 后信息系统的不可辨识关系与未去除时保持一致,则表明 a_i 为冗余属性.设 $B \subseteq A$, 如果 $I(B) = I(A)$, 则 B 是 A 的一个约简,所有约简的集合记为 $\text{red}(A)$,所有约简的交集构成属性的核,记为 $\text{core}(A)$.寻找信息系统的核并不容易,若系统呈现一致性,即决策属性 D 依赖于条件属性 C 时,则该信息系统存在核,若不一致,则各属性均为相对重要属性,该信息系统没有核,需计算最小约简.

1.2 支持向量机

支持向量机(SVM)由 Vapnik 于 20 世纪 90 年代提出,是一种基于统计学习理论的机器学习方法^[19].该方法以结构风险最小化为准则,构建最优分类超平面,将样本数据正确无误地分开,且分类间隔最大.在多维数据空间中,通过核函数将向量映射到高维特征空间,在特征空间中寻找最优分类超平面^[20],提高非线性数据分类的准确性.

设样本集为 $(x_i, y_i), i = 1, 2, \dots, n$, 输入 n 维向量,通过非线性变换函数 $\Psi(x)$ 将输入空间映射到高维特征空间中,定义最优分类超平面方程及相应的优化问题为^[21]:

$$\begin{cases} \omega \cdot x_i + b = 0 \\ y_i [(\omega \cdot x_i) + b] \geq 1 & i = 1, 2, \dots, n \\ \min \frac{1}{2} (\omega \cdot \omega) + c \sum_{i=1}^n \xi_i \end{cases} \tag{3}$$

式中: ω 为权向量; b 为偏置; ξ_i 为松弛变量; c 为惩罚因子, $c > 0$.

根据 Lagrange 乘子法,引入乘子 α, β , 得到优化问题的对偶问题:

$$\max_{\alpha, \beta} Q(\alpha) = L(\omega, \xi, b, \alpha, \beta)$$

$$= \sum_{i=1}^n \alpha_i \cdot \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (4)$$

求解该对偶问题,得到判断函数:

$$f(x) = \text{sgn}(\sum_{i=1}^n \alpha_i^* y_i K(x_i, x_j) + b) \quad (5)$$

式中: $K(x_i, x_j) = \psi(x_i) \cdot \psi(x_j)$, $K(x_i, x_j)$ 表示核函数.核函数的选取对分类结果有一定的影响,常用核函数包括:线性核函数、多项式核函数、径向基函数核函数、二层神经网络核函数.

1.3 属性重要度

采用启发式算法确定属性重要度,常用的启发式算法有基于属性依赖度的属性重要度计算和基于信息熵的属性重要度计算.

基于属性依赖度的属性重要度:决策属性 D 对条件属性 C 的依赖度为 $r(C,D)$,对于 $\forall a_i \in A$,则属性 a_i 的重要度 $\text{SGF}(a_i)$ 定义为去除 a_i 时引起的决策属性对条件属性依赖度的变化^[22]:

$$\text{SGF}(a_i) = r(C,D) - r(C - \{a_i\}, D) \quad (6)$$

基于信息熵的属性重要度:信息量的大小由所消除的不确定性的来度量,设 $R \subset C$,向 R 中增加一个属性 a_i 后信息的增量即为 a_i 的重要度,表示为^[22]:

$$\text{SGF}(a_i, R, D) = H(D|R) - H(D|R \cup \{a_i\}) \quad (7)$$

其中, $H(D|R)$ 表示已知 R 时, D 的条件熵.

2 研究区概况及数据来源

2.1 研究区自然地理

研究区位于辽宁省绥中县狗河下游平原区,面积 71.4km^2 .狗河是绥中县境内一条独流入海河流,发源于绥中县加碑岩乡窝岭村大锥子山板石顶,源头海拔高程 946.3m ,河长为 86.7km .该区地处中纬度,属于温带季风气候,年均降水量为 671mm ,年均蒸发量为 1755mm ,多年平均气温在 9.5°C ,地势平坦,自西北微向东南倾斜,主要为冲洪积倾斜平原以及河谷冲积河床漫滩.

2.2 水文地质条件

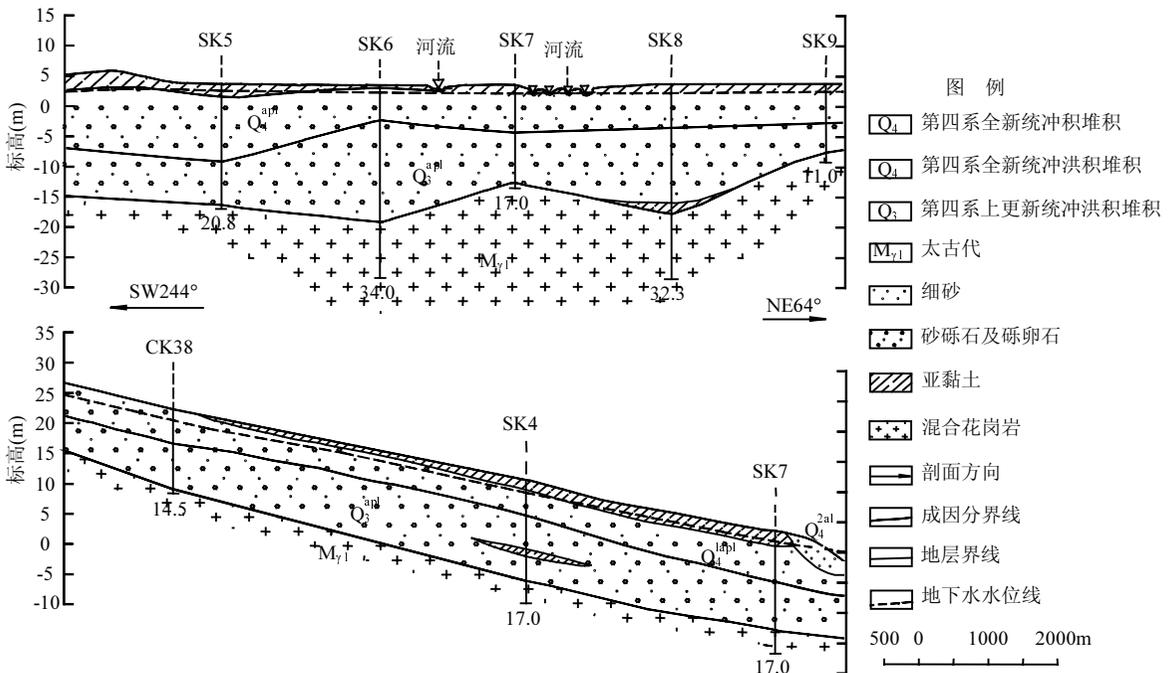


图1 水文地质剖面

Fig.1 Hydrogeologic profile

地下水主要赋存于第四系松散岩类孔隙含水层中,以全新统上部冲积砂砾石卵石层和全新

统洪积—冲积亚黏土及砂砾石层为主.狗河左岸单井涌水量 $1000\sim 3000\text{m}^3/\text{d}$,右岸单井涌水量

500~1000m²/d.从研究区水文地质剖面图(图 1)可以看出,含水层岩性以砂砾石及砾卵石为主,顶部覆盖一层 2~3m 厚的壤土,底部太古界混合花岗岩构成含水层底板.沿河流纵向剖面地势逐渐降低,地下水位逐渐下降;横向剖面地势平稳,含水层厚度变幅较大;地下水流向自西北至东南方向.区域地下水的主要补给来源为大气降水入渗补给,沿河地带丰水期可接受河水的季节性补给,排泄方式主要为潜水蒸发和人工开采,近海 1.5km

宽度向海排泄.

2.3 数据来源

选择位于研究区的 10 个第四系松散岩类孔隙水监测孔,作为水质分析基础数据,监测孔位置如图 2 所示.实验室测定各水质指标浓度,以地下水环境质量标准(GB/T14848-93)^[23]为参照,选取包含 pH、NH₄⁺、NO₂⁻、NO₃⁻、Cl⁻、SO₄²⁻、F⁻、总硬度 (TH)、总溶解固体(TDS)、铁、锰、锌、氰化物和挥发性酚类在内的 14 项水质指标进行分析评价.

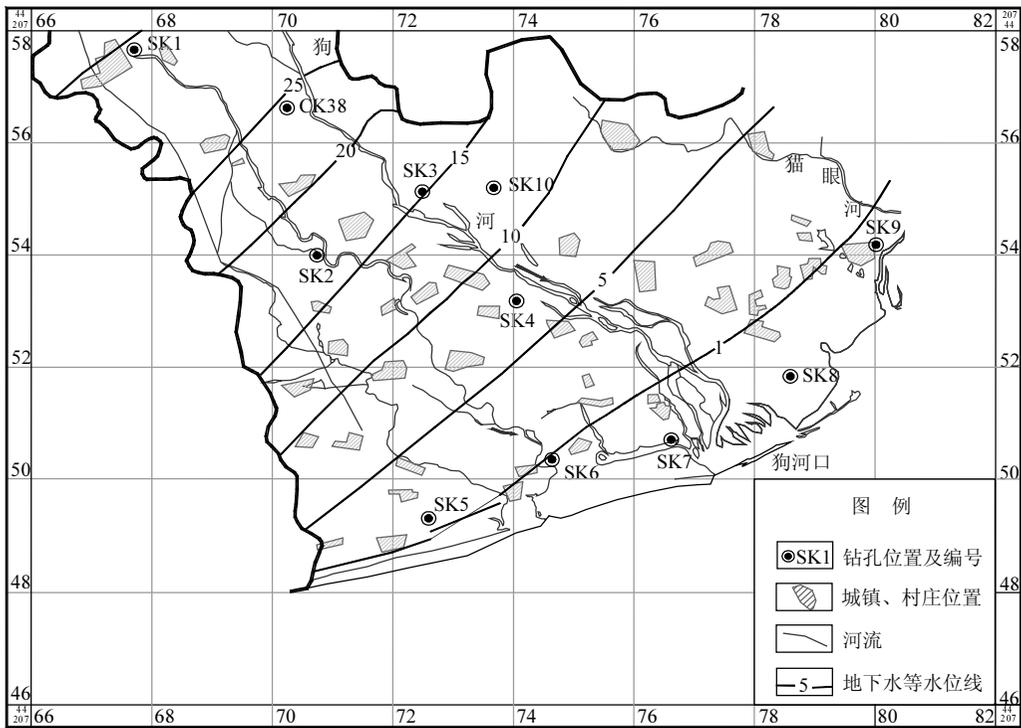


图 2 监测孔位置

Fig.2 Locations of monitoring drilling

3 基于 RS 的 SVM 水质评价

水质评价涉及化学指标众多,各指标对综合水质影响大小存在差异,粗糙集根据数据间的关联性去除冗余指标,是进行知识约简的有效途径.

支持向量机对非线性数据的分类具有优势.将粗糙集与支持向量机联合运用,以达到有效约简冗余指标,提高学习效率,保证水质评价结果合理性的目的.基于 RS 的 SVM 水质评价步骤为:

Step 1:建立水质等级评价的信息系统,进行

数据离散化;

Step 2:判断信息系统的一致性,依据 RS 理论进行属性约简;

Step 3:按式(8)对约简后的样本进行归一化处理,控制其在 0~1 范围内;

$$X = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (8)$$

Step 4:确定核函数并寻找最优参数,建立基于 RS 的 SVM 水质评价模型;

Step 5:将测试样本代入模型,实现分类评价.

4 结果与分析

4.1 属性约简

在水质评价的信息系统中,评价因子作为属

性集,各因子的属性值作为集合 V ,构成论域 U .以地下水环境质量标准(GB/T14848-93)^[23]为原则划分单因子水质级别,将连续的水质浓度以 I~V 类划分,使其离散化,离散结果见表 1.

表 1 地下水水质离散化

Table 1 Discretized for groundwater quality

样本点	pH 值 (a ₁)	TH (a ₂)	TDS (a ₃)	NO ₃ ⁻ (a ₄)	NH ₄ ⁺ (a ₅)	NO ₂ ⁻ (a ₆)	F ⁻ (a ₇)	Cl ⁻ (a ₈)	SO ₄ ²⁻ (a ₉)	酚 (a ₁₀)	铁 (a ₁₁)	锰 (a ₁₂)	锌 (a ₁₃)	氰化物 (a ₁₄)
SK1	I	I	I	II	I	I	I	I	I	I	III	I	III	I
SK2	I	III	III	III	I	I	I	I	IV	IV	II	I	III	II
SK3	I	I	I	II	I	I	I	I	I	I	IV	I	I	I
SK4	I	I	I	III	I	I	I	I	I	III	IV	I	II	I
SK5	I	I	I	III	I	I	I	I	I	I	II	I	II	I
SK6	I	I	I	II	I	I	I	I	I	I	V	I	I	II
SK7	I	I	I	III	I	I	I	I	I	IV	I	I	I	I
SK8	IV	I	IV	III	V	IV	I	IV	III	I	V	IV	I	I
SK9	I	I	I	IV	I	I	I	I	II	I	V	I	II	I
SK10	I	I	I	II	I	IV	IV	I	I	I	I	I	II	I

经分析,上述信息系统呈现不一致性,由于求解最小约简属于 NP-hard 问题^[24],本次通过计算各属性的重要度来确定最小约简.根据文献[22],基于属性依赖度和基于信息熵的属性重要度算法在刻画重要度时都是不完备的.因此,对两种算法加权平均,计算出各属性的重要度,以属性依赖度作为主要决策标准,赋权值 0.9,信息熵作为辅助决策标准,赋权值 0.1.

编写关于该信息系统的属性约简程序,借助 Matlab R2011b 软件,精简属性集,获得信息系统的最优约简集合为 {a₃,a₄,a₆,a₉,a₁₀,a₁₁,a₁₃,a₁₄},即 TDS、NO₃⁻、NO₂⁻、SO₄²⁻、酚、铁、锌和氰化物是评价研究区水质的关键性指标.

4.2 SVM 水质分析

将水质指标作为输入集,等级标准作为输出集,构建水质等级评价体系.以 GB/T14848-93^[23]为水质参照标准,在 14 项指标的 5 个等级标准中分别线性内插 4 个值,构成 25 组学习样本,属性约简前后的信息系统分别作为测试样本.对归一化的样本进行多次模拟试验,根据试验结果选择径向基函数 $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ (γ 为内核参数, $\gamma > 0$) 作为本次评价的核函数.采用交叉验证

法对惩罚因子(c)和内核参数(γ)进行寻优,利用 Grid.py 函数及制图软件 Gnuplot,绘制样本在约简前后的交叉验证曲线,如图 3 所示:(a)为约简前指标的交叉验证曲线,(b)为约简后指标的交叉验证曲线.由图可知,约简前最优惩罚因子 $c=8.0$,内核参数 $\gamma=64.0$;约简后最优惩罚因子 $c=4.0$,内核参数 $\gamma=32.0$.

分别将约简前的全部水质指标和约简后的 TDS、NO₃⁻、NO₂⁻、SO₄²⁻、酚、铁、锌、氰化物指标,及其对应的惩罚因子(c)和内核参数(γ)代入 Matlab 水质评价程序,输出结果如图 4.水质评价结果见表 2.

通过对约简前后信息系统的分别评价,两次评价结果一致,约简前共有 14 项评价指标,程序运行时间 1.53s,约简后共有 8 项评价指标,程序运行时间仅用 1.23s.粗糙集将 pH 值、TH、NH₄⁺、F⁻、Cl⁻和锰指标去除,在本信息系统中,以上 6 项指标的重要度小于其他指标,据表 1 可知,这些指标的水质等级相对简单,对分类结果的影响较弱.可见,经过粗糙集的属性约简,去除了冗余的评价指标,同时较好地保持了原信息系统的分类能力.

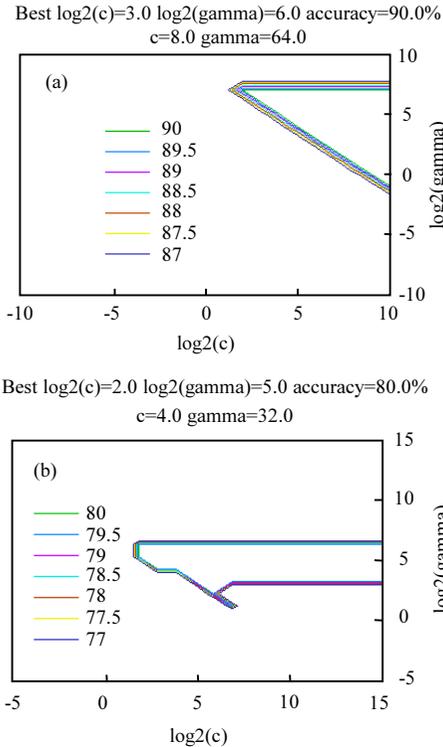


图3 参数交叉验证曲线

Fig.3 Parametric cross validation curve

(a)约简前;(b)约简后;c 惩罚因子;gamma 内核参数

表2 属性约简前后评价结果

Table 2 Assessment results of before and after reduction

样本点	约简前	约简后	样本点	约简前	约简后
SK1	II	II	SK6	IV	IV
SK2	III	III	SK7	III	III
SK3	III	III	SK8	V	V
SK4	III	III	SK9	IV	IV
SK5	II	II	SK10	III	III

根据各监测孔水质综合评价结果(图 5)可以看到,研究区地下水水质以 II、III 类居多,其中,II 类水占 20%,III 类水占 50%,IV 类水占 20%,V 类水占 10%.根据水质标准 GB/T14848-93^[23],该地下水有 70%(II、III 类)满足饮用条件,主要分布在研究区域河流上游西部及北部地区,20%(IV 类)的地下水经处理后方可饮用,主要分布在河流下游,研究区域东部及东南部,10%(V 类)的地下水不适宜人类饮用,分布在河流下游入海口处.IV 类水在监测孔 SK6 和 SK9 处出现,SK6 铁

超标 8.3 倍,SK9 铁超标 36.7 倍, NO_3^- 超标 1.4 倍,两处地下水水质均较差,未经处理不可作为饮用水;在监测孔 SK8 中,水质呈现出 V 类极差水,其中铁超标 18.7 倍, NH_4^+ 超标 2.8 倍, NO_2^- 超标 1.5 倍,Cl⁻ 超标 1.3 倍,此外 pH 和 TDS 亦少量超标,可见 SK8 处地下水不宜作为饮用水。

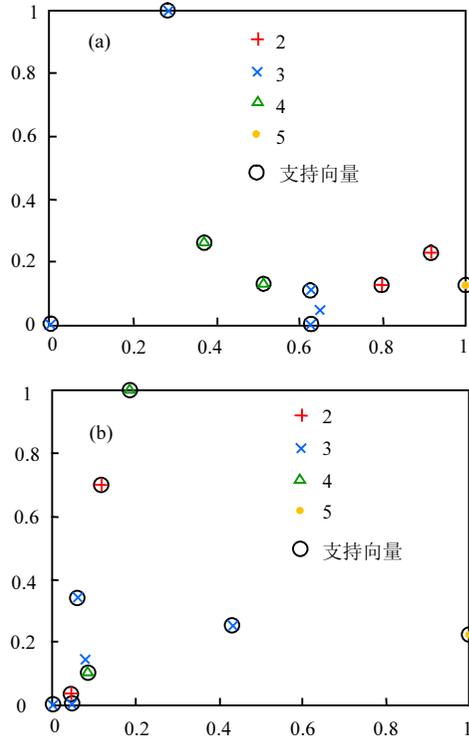


图4 Matlab 水质分类输出结果

Fig.4 Matlab output for water quality classification

(a)约简前;(b)约简后;2,3,4,5 分别为 II,III,IV,V 类

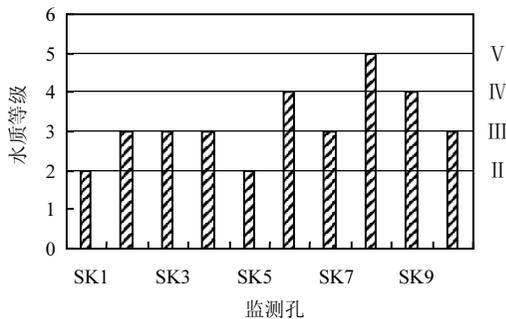


图5 监测孔综合评价结果

Fig.5 Results of monitoring drilling comprehensive assessment

研究区主要污染因子为铁和“三氮”,分析其原因是由于狗河上游存在黄铁矿,致使下游地下水铁超标,此外化肥和农药等农村非点源污染,导致地下水“三氮”超标.从监测井的平面位置来看,IV、V类水主要分布在沿海地区,考虑到研究区位于狗河下游入海口处,地势由高变低明显,水力梯度较大,上游溶解于地下水中的铁、“三氮”等污染物随水流运移,在下游积聚,致使沿海地区水质恶化.

5 结论

5.1 研究区地下水部分适宜饮用,以 II、III 类为主.部分地区出现 IV 类较差水、甚至 V 类极差水,主要污染因子为铁、“三氮”等,其中铁超标严重,多在 8.3 倍以上.

5.2 影响地下水环境质量的化学指标众多,其中不乏冗余指标,粗糙集的属性约简方法将 14 项指标精简为 8 项,约简前后的指标经支持向量机法评价,水质等级一致.粗糙集与支持向量机的联合使用,不仅在保证分类能力的前提下有效地减少冗余指标,降低运算的复杂程度,同时机器算法缩减了计算时间,且其在多维计算中的优势使评价结果更加准确.

5.3 粗糙集与支持向量机的运算过程均依据于数据间的相关性,在揭示数据规律的同时实现指标约简和水质评价,在地下水质量评价中具有普遍的适用性.

参考文献:

- [1] Viala E. Water for food, water for life a comprehensive assessment of water management in agriculture [J]. *Irrigation and Drainage Systems*, 2008,22(1):127-129.
- [2] Shrestha S, Kazama F. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan [J]. *Environmental Modelling and Software*, 2007, 22(4):464-475.
- [3] 刘 博,肖长来,梁秀娟,等.吉林市城区浅层地下水污染源识别及空间分布 [J]. *中国环境科学*, 2015,35(2):457-464.
- [4] 梁小俊,张庆庆,许月萍,等.层次分析法-灰关联分析法在京杭运河杭州段水质综合评价中的应用 [J]. *武汉大学学报(工学版)*, 2011,44(3):312-316,325.
- [5] Dahiya S, Singh B, Gaur S, et al. Analysis of groundwater quality using fuzzy synthetic evaluation [J]. *Journal of Hazardous Materials*, 2007,147(3):938-946.
- [6] Maier H R, Morgan N, Chow C W K. Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters [J]. *Environmental Modelling and Software*, 2004,19(5):485-494.
- [7] Huang H, Liang XJ, Xiao CL, et al. Analysis and assessment of confined and phreatic water quality using a rough set theory method in Jilin City, China [J]. *Water Science and Technology: Water Supply*, 2015,15(4):773-783.
- [8] 董海彪,卢文喜,安永凯,等.基于对应分析法的鄂尔多斯盆地东北部地下水污染分析 [J]. *中国环境科学*, 2015,35(11):3371-3378.
- [9] 肖传宁,卢文喜,安永凯,等.基于两种耦合方法的模拟-优化模型在地下水污染源识别中的对比 [J]. *中国环境科学*, 2015,35(8):2393-2399.
- [10] Pai P F, Lee F C. A rough set based model in water quality analysis [J]. *Water Resources Management*, 2010,24(11):2405-2418.
- [11] Li Peiyue, Wu Jianhua, Qian Hui. Groundwater quality assessment based on rough sets attribute reduction and TOPSIS method in a semi-arid area, China [J]. *Environmental Monitoring and Assessment*, 2012,184(8):4841-4854.
- [12] 胡可云,陆玉昌,石纯一.粗糙集理论及其应用进展 [J]. *清华大学学报:自然科学版*, 2001,41(1):64-68.
- [13] 李祚泳,张正健.基于回归支持向量机的指标规范值的水质评价模型 [J]. *中国环境科学*, 2013,33(8):1502-1508.
- [14] 刘双印,徐龙琴,李道亮,等.基于时间相似数据的支持向量机水质溶解氧在线预测 [J]. *农业工程学报*, 2014,30(3):155-162.
- [15] Zhang D Z, Xia B K. Soft measurement of water content in oil-water two-phase flow based on RS-SVM Classifier and GA-NN Predictor [J]. *Measurement Science Review*, 2014,14(4):219-226.
- [16] Liu Shuangyin, Xu Longqin, Li Daoliang. Water quality early-warning model based on support vector machine optimized by rough set algorithm [J]. *Systems Engineering - Theory and Practice*, 2015,35(6):1617-1624.
- [17] Pawlak Z, Skowron A. Rudiments of rough sets [J]. *Information Sciences*, 2007,177(1):3-27.
- [18] Swiniarski R W, Skowron A. Rough set methods in feature selection and recognition [J]. *Pattern recognition letters*, 2003, 24(6):833-849.
- [19] Vapnik V N. The nature of statistical learning theory [M]. Berlin: Springer-Verlag, 1995.
- [20] 牛瑞卿,彭 令,叶润青,等.基于粗糙集的支持向量机滑坡易发性评价 [J]. *吉林大学学报(地球科学版)*, 2012,42(2):430-439.
- [21] 梁雪春,龚艳冰,肖 迪.一种多核加权支持向量机的水质预测方法 [J]. *东南大学学报(自然科学版)*, 2011,41(s):14-17.
- [22] 石 峰,娄臻亮,张永清.一种改进的粗糙集属性约简启发式算法 [J]. *上海交通大学学报*, 2002,36(4):478-481.
- [23] GB/T14848-93 地下水质量标准 [S].
- [24] 乌兰图雅,李东魁.求解一类 NP-HARD 问题的一个快速算法 [J]. *内蒙古大学学报(自然科学版)*, 2012,43(1):85-88.

作者简介: 黄鹤(1988-),女(满族),吉林长春人,吉林大学博士研究生,主要从事水资源与水环境研究.发表论文 7 篇.