

# 基于模拟-优化方法的地下水污染源溯源辨识

潘紫东<sup>1,2</sup>, 卢文喜<sup>1,2\*</sup>, 范越<sup>1,2</sup>, 李久辉<sup>1,2</sup>, 王涵<sup>1,2</sup> (1. 吉林大学地下水与资源环境教育部重点实验室, 吉林 长春 130012; 2. 吉林大学新能源与环境学院, 吉林 长春 130012)

**摘要:** 以抚顺市某煤矸石堆放场为研究区, 根据研究区的实际条件建立地下水污染质运移模拟模型, 预测地下水污染质未来时空变化特征。基于正演预报结果构建了假想例子, 应用模拟-优化方法对地下水污染源源强及场地的渗透系数进行反演识别。为减小优化模型反复调用模拟模型所产生的计算负荷, 分别采用 Kriging 方法和 BP 神经网络方法建立了模拟模型的替代模型。最后运用模拟退火法求解优化模型, 得到反演识别结果。研究表明: 应用 Kriging 方法建立的替代模型输出结果的平均相对误差为 0.3%; 应用 BP 神经网络方法建立替代模型的输出结果平均相对误差为 1.5%, 应用两种替代模型对污染源源强识别的相对误差均小于 0.5%, 对场地两个参数分区渗透系数识别的相对误差均不大于 5%。综上, 应用 Kriging 方法建立的替代模型精度高于 BP 神经网络方法, 利用基于两种替代模型的模拟-优化方法对污染源源强和渗透系数进行同步识别精度可以满足实际需求, 是有效可行的。

**关键词:** 污染源反演识别; 模拟-优化方法; 替代模型; Kriging 方法; BP 神经网络方法

**中图分类号:** X523      **文献标识码:** A      **文章编号:** 1000-6923(2020)04-1698-08

**Inverse Identification of groundwater pollution source based on simulation-optimization approach.** PAN Zi-dong<sup>1,2</sup>, LU Wen-xi<sup>1,2\*</sup>, FAN Yue<sup>1,2</sup>, LI Jiu-hui<sup>1,2</sup>, WANG Han<sup>1,2</sup> (1. Key Laboratory of Groundwater Resources and Environment Ministry of Education, Jilin University, Changchun 130012, China; 2. College of New Energy and Environment, Jilin University, Changchun 130012, China). *China Environmental Science*, 2020, 40(4): 1698~1705

**Abstract:** A coal gangue pile in Fushun City was selected as the study area, and the groundwater numerical simulation model for this area was established based on physical condition. After that, we applied the model in predicting possible spatial and temporal variation of groundwater pollutant in research area. Based on forward modelling, a hypothetical example was designed to explore the application of simulation-optimization method and simultaneously carried out inverse identification of groundwater pollution sources and model parameters. The Kriging method and BP Neural Network method were proposed to establish a surrogate model of the groundwater numerical model so as to reduce the computational load caused by the repeated invocation of the simulation model by the optimization model, and the surrogate model was then integrated to simulation-optimization model where simulated annealing method was used. The study reached the following conclusions: As the surrogate model of the simulation model established by Kriging method, the mean relative error of the output concentration is 0.3%, by contrast with 1.5% with BP Neural Network method. The identification error of the release intensity of pollution source with two methods are both below 0.5%, and the identification error of hydraulic conductivity both partitions is no larger than 5%. Above all, the accuracy of surrogate model established by Kriging method is higher than BP neural network method. It is proved that the inverse identification by using the simulation-optimization method based on two surrogate models is effective and accurate.

**Key words:** inverse identification of pollution source; simulation-optimization model; surrogate model; Kriging method; BP neural network method

随着人口的日益增长和科技水平的提高,我国在社会进步的同时也出现了一系列环境污染问题,其中地下水污染为较为严重的问题之一<sup>[1]</sup>。由于地下水埋藏于地下环境,其受污染情况不易被人们察觉,这就使得地下水污染问题具有隐蔽性和发现滞后性<sup>[2]</sup>,所以十分有必要对地下水污染进行提早防治。而在地下水污染防治过程中,其先行工作则需要探测污染源的各类特征,如位置、释放历史(不同时段污染源的释放强度)等,以实现污染源及污染羽的有效控制。污染源特征的识别属于一类反演问题。

地下水污染源反演识别是指运用有限且离散的地下水动态观测数据,对地下水污染数值模拟模型进行反演求解,识别确定污染源的个数、位置及释放历史<sup>[3]</sup>。本文主要进行污染源源强及渗透系数的联合识别。

地下水污染反演识别在数学上属于数理方程反问题。在地下水污染反演问题的解决方法中,模拟

收稿日期: 2019-09-10

基金项目: 国家自然科学基金资助项目(41672232)

\* 责任作者, 教授, luwx999@163.com

优化方法的应用较为广泛。Triptimoni Borah 等<sup>[4]</sup>利用基于人工神经网络的模拟-优化方法成功识别出了污染源;肖传宁等<sup>[5]</sup>针对地下污染源识别分别采用了两种方法(响应矩阵法和状态转移方法)耦合模拟-优化模型,并采用遗传算法准确识别了污染源的释放强度。模拟-优化方法把模拟模型与优化模型耦合,通过多次调用模拟模型使得每一次的模拟结果与观测结果之间的差异函数极小,以达到对模型参数的不断优化并最终识别。Ayvaz<sup>[6]</sup>等利用模拟-优化方法,结合采用二进制遗传算法和广义降阶方法较为准确地求解出了场地的水文地质参数。目前对于地下水污染的反演识别多为对污染源或模型参数的单一识别,即此二者其一为已知的前提下,对另一因素进行识别,本文旨在解决污染源及模型参数均未知的情况下,如何对二者进行联合识别。

从运筹学的角度上来看,对污染源和模型参数的反演问题可以转化为一个优化问题,以污染源和模型参数为待求变量,以污染质浓度监测值与模拟计算值尽可能接近为优化目标,以地下水模拟模型为约束条件,形成一个有约束条件的极小化问题。

对于优化模型的求解多采用启发式算法进行求解,常见的启发式算法有遗传算法和模拟退火法,其中模拟退火法通过随机模型扰动,选择最优解区间,不断更新状态模型,并以一定概率选择领域中能量值大的状态避免困于局部最优解,不断迭代直至出现最小能量状态即为最优解<sup>[7]</sup>。模拟退火法具有收敛速度较快和不易陷入局部最优解等优点,因此广泛应用于优化模型的求解中<sup>[8-11]</sup>。

利用模拟-优化方法解决地下水污染反演识别问题,需要对地下水数值模拟模型进行多次调用,较大的运算负荷也会伴随着产生。基于此则需要引入模拟模型的替代模型,以提高运算效率以及减小运算负荷。已应用于地下水领域的替代模型构建方法有:人工神经网络法(ANN)<sup>[12]</sup>;径向基函数法<sup>[13]</sup>;Kriging 法<sup>[14]</sup>等。根据前人研究,Kriging 方法和神经网络方法建立的模拟模型的替代模型均有着较高的拟合精度,可以考虑推行于地下水反演识别问题中,然而前人研究都只构建一种替代模型,并没有对多种替代模型构建方法进行对比。故本文将针对同一假设例子采用两种方法建立替代模型,并比较二者的性能及精度。

综上,本文以模拟-优化方法为框架,提出了对地下水污染源和模型参数进行联合辨识的方法。本次研究以抚顺市某煤矸石堆放场为研究区,构建了地下水数值模拟模型,再以正演预报结果进行假设例子的构建,以验证该方法的有效性。为减小反演过程中的计算负荷,对比使用了 Kriging 和 BP 神经网络两种方法建立模拟模型的替代模型。最后采用模拟退火法求解优化模型,得到辨识结果。本次研究为地下水污染信息的综合识别提供了一种稳定可靠的方法。

## 1 模拟-优化方法

模拟-优化法分为两个部分,模拟部分即为建立地下水污染质运移模拟模型,用以建立输入-输出响应系统,模拟地下水污染质时空分布特征。优化部分即为将污染质浓度的模拟计算值与实际观测值之差的平方和作为目标函数,建立优化模型。本文的模拟-优化模型求解旨在将水流、水质模型耦合于优化模型,形成模拟-优化模型,通过不断调用模拟-优化模型,将每一次模拟的结果与实际观测数据值进行比对,直至二者的差异函数极小,从而达到对模型参数的持续优化,最终得到使模拟结果最接近观测数据的解。

### 1.1 地下水污染质运移模拟模型

1.1.1 地下水水流模型的建立 根据研究区水文地质概念模型,建立研究区地下水流的数值模拟模型,如下所示:

$$\begin{cases} \frac{\partial}{\partial x}(K(H-B)\frac{\partial H}{\partial x}) + \frac{\partial}{\partial y}(K(H-B)\frac{\partial H}{\partial y}) + w \\ = \mu \frac{\partial H}{\partial t} \quad (x, y) \in S, \quad t \geq 0 \\ H(x, y, t)|_{t=0} = H_0(x, y) \\ H(x, y, t)|_{\Gamma_1} = \varphi(x, y, t) \quad (x, y) \in \Gamma_1 \\ K \frac{\partial H}{\partial n} |_{\Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5} = q(x, y, t) \quad (x, y) \in \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5 \end{cases} \quad (1)$$

式中:  $K$  为渗透系数;  $H$  为水位高程;  $B$  为底板高程;  $w$  为源汇项;  $\mu$  为给水度;  $\Gamma_1$  为已知水头边界(第一类边界);  $\Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5$  为已知流量边界(第二类边界);  $\vec{n}$  为边界上某点  $(x, y)$  处外法线方向上的单位向量,  $\varphi(x, y, t)$  和  $q(x, y, t)$  为已知函数。

1.1.2 地下水污染质运移模型的建立 在地下水

水流数值模型基础之上,建立地下水污染质运移数学模拟模型:

$$\begin{cases} \frac{\partial}{\partial x}(D_x \frac{\partial c}{\partial x}) + \frac{\partial}{\partial y}(D_y \frac{\partial c}{\partial y}) - \frac{\partial}{\partial x}(u_x c) - \frac{\partial}{\partial y}(u_y c) + I \\ = R_d \frac{\partial c}{\partial t}, & (x, y) \in S, t \geq 0 \\ c(x, y, t)|_{t=0} = c_0(x, y) & (x, y) \in S \\ c(x, y, t)|_{\Gamma_3} = c_1(x, y, t) & (x, y) \in \Gamma_3, t > 0 \\ -nbD \text{grad } c \cdot \vec{n}|_{\Gamma_2, \Gamma_4} = c_2(x, y, t) & (x, y) \in \Gamma_2, \Gamma_4, t > 0 \\ nb(\bar{c} - D \text{grad } c) \cdot \vec{n}|_{\Gamma_1, \Gamma_5} = c_3(x, y, t) & (x, y) \in \Gamma_1, \Gamma_5, t > 0 \end{cases} \quad (2)$$

式中:  $t$  为时间变量;  $c$  为地下水中溶质浓度;  $D_x$ 、 $D_y$  是  $x$ 、 $y$  方向上的水动力弥散系数;  $u_x$ 、 $u_y$  分别为实际平均流速向量  $\bar{u}$  在  $x$ 、 $y$  轴方向上的分量;  $n$  为孔隙度;  $b$  为含水层厚度;  $I$  为源汇项;  $R_d$  为滞留因子;  $S$  为研究区;  $\Gamma_3$  为已知浓度边界(第一类边界),  $\Gamma_2, \Gamma_4$  为已知水动力弥散通量边界(第二类边界);  $c_0(x, y, t)$ ,  $c_1(x, y, t)$ ,  $c_2(x, y, t)$ ,  $c_3(x, y, t)$  为已知函数。

### 1.2 替代模型

通过训练,替代模型可以得到与模拟模型相近的输入-输出响应关系<sup>[15-17]</sup>。作为模拟模型的近似替代,替代模型更易于解算,能够大幅度地减少程序计算负荷和时间<sup>[18-20]</sup>。替代模型通过对一定数量的已知样品特征的输入-输出拟合,来构建模拟模型的拟合函数,用于预测未知样品的特征输出响应<sup>[21]</sup>。本文采用 Kriging 方法及 BP 神经网络方法建立模拟模型的替代模型。

#### 1.2.1 Kriging 替代模型 其数学表达式如下:

$$y(x) = f(x)^T \beta + z(x) \quad (3)$$

式中:  $y(x)$  为替代模型中所输出的污染物浓度,  $\hat{y}(x)$  为  $y(x)$  的估计值,由线性回归部分  $f(x)^T \beta$  和随机部分  $z(x)$  组成。其中  $f(x) = [f_1(x), f_2(x), \dots, f_k(x)]$  为已知回归模型的基函数; 待定参数  $\beta = [\beta_1, \beta_2, \dots, \beta_k]$  为基函数相对应的系数。随机部分  $z(x)$  满足下列条件:

$$\begin{cases} E(z(x)) = 0 \\ D(z(x)) = \sigma^2 \\ \text{cov}[z(x_i), z(x_j)] = \sigma^2 R(x_i, x_j) \end{cases} \quad (4)$$

式中:  $R(x_i, x_j)$  为任意 2 个采样点  $x_i$  和点  $x_j$  之间的空间相关关系方程,即关联函数,其数学表达形式为:

$$R(x_i, x_j) = \exp\left(-\sum_{k=1}^n \theta_k |x_k^i - x_k^j|\right) \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, n) \quad (5)$$

式中:  $\theta_k$  为待定参数;  $x_k^i$  为第  $i$  个样本的  $k$  维坐标。

根据 Kriging 模型,预测点  $x$  处的响应值  $y(x)$  的预测估计值为:

$$\hat{y}(x) = f^T \beta + r^T(x) R^{-1} (y - f \beta) \quad (6)$$

式中:  $r(x)$  为点  $x$  与  $n$  个训练样本采样点  $(x_1, x_2, \dots, x_n)$  之间的相关向量,  $r(x) = [R(x, x_1), R(x, x_2), \dots, R(x, x_n)]$ ;  $y$  是  $n$  个采样点对应的污染物浓度响应值,为  $n \times 1$  的向量;  $\beta$  为线性回归部分的待定参数,可以通过最优线性无偏估计求得:

$$\beta = (f^T R^{-1} f)^T f^T R^{-1} y \quad (7)$$

$R$  为  $n$  个采样点相关系数组成的  $n \times n$  阶相关矩阵:

$$R = \begin{bmatrix} R(x_1, x_1) & \dots & R_3(x_1, x_n) \\ \vdots & \ddots & \vdots \\ R_3(x_n, x_1) & \dots & R_3(x_n, x_n) \end{bmatrix} \quad (8)$$

方差  $\sigma^2$  估计值利用以下式子确定:

$$\sigma^2 = \frac{1}{n} (y - f \beta)^T R^{-1} (y - f \beta) \quad (9)$$

所以 Kriging 替代模型的建立实际上就是求解上面的非线性无约束优化问题。待求参数  $\theta_k$  求出后,通过建立的 Kriging 模型可获得污染物浓度响应值,其中  $\theta_k$  可以通过无约束优化式子求得:

$$\max_{\theta_k} \{-n \ln \sigma^2 + \ln |R|\} \quad (10)$$

1.2.2 BP 神经网络替代模型 BP 神经网络为多层前向网络,是一种多隐层堆叠或是多模型组合构造的具有强大计算能力的神经网络。每一层由数个神经元组成,通过神经元中的计算与权重更新来实现正向和反向传播,该模型中的神经元的计算方式模拟生物神经元突触信号传递方式,按照一定的函数计算方式来获得输出,将前一层的输出通过激活函数作为下一层的输入,最终获得预测值(正向传播),再通过反向传播使误差极小化,即选择神经网络权

值使损失函数极小化(反向传播).其结构与求解过程如下所示:

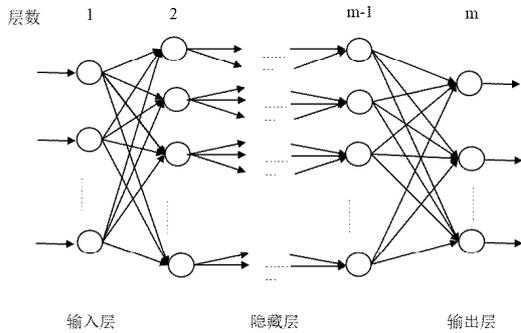


图1 BP神经网络结构

Fig.1 Structure of BP Neural Network

(1)正向传播

假设建立的BP神经网络有  $m$  层,隐层  $k$  与输出层的各个神经元的非线性输入输出关系记为  $f_k (k=2,3,\dots,m)$ ,由第  $k-1$  层的第  $j$  个神经元到第  $k$  层的第  $i$  个神经元的连接权值为  $W_{ij}^k$ ,另设第  $k$  层第  $i$  个神经元输入的总和为  $u_i^k$ 、输出为  $y_i^k$ ,则各变量之间的关系为:

$$\begin{aligned} y_i^k &= f_k(u_i^k) \\ u_i^k &= \sum_j w_{ij}^{k-1} y_j^{k-1} \\ k &= 2, 3, \dots, m \end{aligned} \tag{11}$$

当该神经网络输入数据  $X = [x_1 \ x_2 \ \dots \ x_{p_1}]^T$  (设输入层有  $p_1$  个神经元),则从输入层依次经过各隐层节点可得到输出数据  $Y = [y_1^m \ y_2^m \ \dots \ y_{p_m}^m]^T$  (设输出层有  $p_m$  个神经元).

(2)反向传播

反向传播过程通过梯度下降法不断沿目标函数负梯度方向更新权重  $W_{ij}^k$  使期望输出  $Y_{si}$  与神经网络实际输出  $Y_j^m$  的差异函数最小.具体公式如下:

$$\begin{aligned} \Delta w_{ij}^{k-1} &= -\varepsilon d_i^k y_j^{k-1} \\ d_i^m &= y_i^m (1 - y_i^m) (y_i^m - y_{si}) \\ d_i^k &= y_i^k (1 - y_i^k) \sum_l d_l^{k+1} w_{li}^k \quad (k = m - 1, \dots, 2) \end{aligned} \tag{12}$$

式中:  $\varepsilon$  为学习步长,一般小于 0.5.

本文采用的BP神经网络结构为3-30-5,如图2所示,其中输入层包含3个神经元(对应于源强  $S$  及两个分区的渗透系数  $K_1$ 、 $K_2$ ),隐含层包含30个神经元,输出层包含5个神经元(对应于5口监测井的

污染质浓度).本文利用 MATLAB 代码,将上述数学公式应用于替代模型的建立.

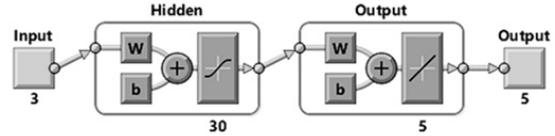


图2 单隐层神经网络结构

Fig.2 Structure of Single-Hidden-Layer Neural Network

1.3 优化模型

优化模型通常由3个部分组成:决策变量、目标函数和约束条件.本文所对应的决策变量即为待求的地下水污染源源强及场地的渗透系数,目标函数则为各井污染质浓度实际监测值与模拟计算值之间差异函数的极小化,约束条件则需要各监测井污染质满足地下水溶质运移规律的同时,其浓度也处于合理的范围内.具体表达形式如下:

$$\begin{aligned} z &= \min \sum_{m=1}^n (C_m - \hat{C}_m) ^2 \\ s.t. &\begin{cases} C = f(s, k) \\ C_L \leq C \leq C_U \\ s_l \leq s \leq s_u \end{cases} \end{aligned} \tag{13}$$

式中:  $z$  为目标函数;  $C_m$  为第  $m$  口监测井污染质浓度实际监测值;  $\hat{C}_m$  为第  $m$  口监测井污染质浓度模拟计算值;  $C$  为污染物浓度;  $s$  为污染源源强;  $k$  为场地渗透系数;  $C_L$ 、 $C_U$  为污染物浓度的上下限值;  $s_l$ 、 $s_u$  为污染源源强上下限值.

2 案例应用

2.1 实际例子的正演预报

2.1.1 研究区概况 根据本次地下水数值模拟的研究目的,将研究区圈定在抚顺市某煤矸石堆至浑河的地区,东西长约 9700m,南北宽约 6100m,总面积约为 52.35km<sup>2</sup>.模拟研究区示意图详见图 3.本次研究的主要对象为一层潜水含水层,厚约 8m.研究区根据含水层渗透性能可分为 2 个区域,概化为非均质各向同性含水层,水流为二维非稳定流.研究区北部边界  $\Gamma_1$  为浑河,概化为已知水头边界;东北部边界  $\Gamma_2$  和南部边界  $\Gamma_4$  概化为零通量边界;将研究区东部边界  $\Gamma_3$  概化为侧向径流补给边界;西南部边界  $\Gamma_5$  概化为侧向径流排泄边界.侧向径流边界概化为已知流

量边界.在溶质运移模型中,研究区北部边界  $\Gamma_1$  和西南部边界  $\Gamma_5$  概化为已知对流-弥散通量边界;东北部边界  $\Gamma_2$  和南部边界  $\Gamma_4$  概化为零通量的水动力弥散通量边界;东部边界  $\Gamma_3$  概化为已知浓度边界.

为监测研究区水位及污染质浓度动态变化,依据前期踏勘结果,本次研究沿地下水水流大致流向(东南流向西北)设立 3 个水位控制点(J1~J3)及 5 个水质控制点(J4~J8).

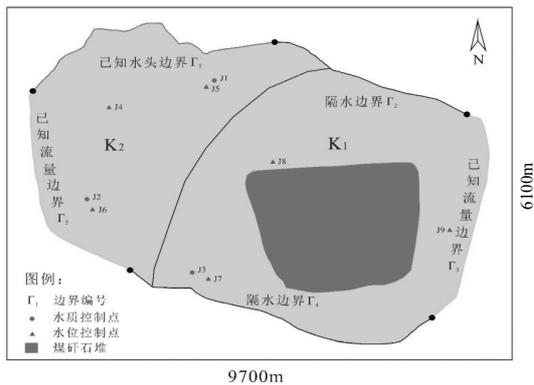


图3 研究区概况

Fig.3 Overview of the study area

表 1 模型校正后水文地质参数取值

Table 1 Hydrogeology parameters through calibration of simulation model

区域	渗透系数 (m/d)	弥散度(m)	孔隙度	吸附分配系数 (cm <sup>3</sup> /g)	滞留因子
东部地区	82.7	4.55	0.22	$6.5 \times 10^{-2}$	1.34
西部地区	97.4	9.8	0.25		

表 2 各井水位实测值与模拟值对比

Table 2 Contrast of observed and simulated water levels

井号	实测值(m)	模拟值(m)	绝对误差(m)
J1	67.35	67.17	0.18
J2	65.07	65.48	0.41
J3	71.23	70.79	0.44

2.1.2 模型的建立与求解 本次数学模拟模型使用 GMS 软件中的 MODFLOW 和 MT3D 模块来求解.本次模拟的求解方法是在计算区域内采用矩形剖分,应用有限差分法进行求解.经过细致的调参拟合,观测井计算水位与实测水位误差的平均绝对值为 0.34m,各井拟合误差均小于 0.5m;硫酸根计算浓度与实测浓度的平均相对误差为 7.38%,所有

观测井的相对误差均小于 20%,可见模型校正取得了较好的结果.最终确定煤矸石堆向地下水释放硫酸根离子的强度为 7000mg/d,研究区的水文地质参数如下:

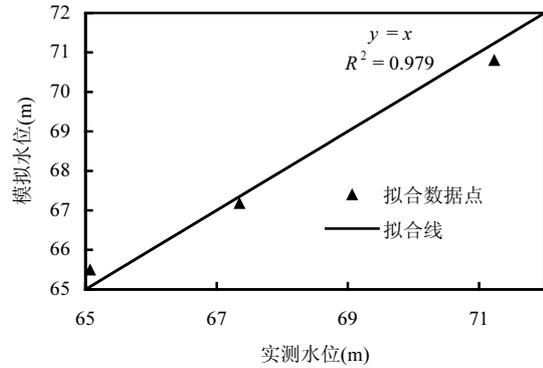


图 4 各井水位拟合情况

Fig.4 Contrast of observed and simulated water levels

表 3 各井硫酸根浓度实测值与模拟值对比

Table 3 Contrast of observed and simulated concentration of sulfate

井号	实测值(mg/L)	模拟值(mg/L)	绝对误差(mg/L)	相对误差(%)
J4	312.25	319.96	7.71	2.47
J5	270.46	252.14	18.32	6.77
J6	175.83	150.16	25.67	14.59
J7	156.44	146.29	10.15	6.48
J8	3060.00	3034.36	25.64	0.83
J9	217.25	188.73	28.52	13.13

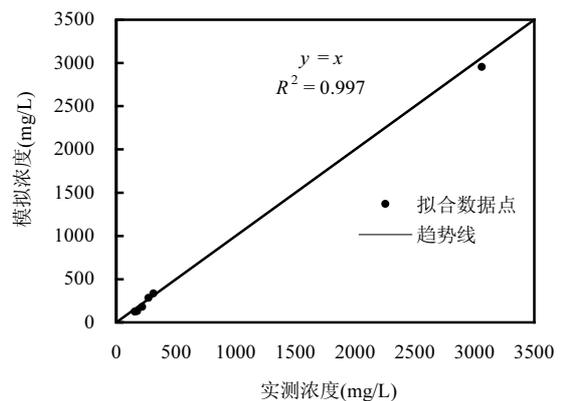


图 5 各井硫酸根浓度拟合情况

Fig.5 Contrast of observed and simulated concentration of sulfate

2.1.3 模型的预报 对模型进行校正后,下一步是应用模型对研究区污染质未来的时空分布特征进行预测.本次模型预报以当前污染质浓度现状为初

始条件,预测未来 5a、10a 和 20a 后污染质的时空分布特征.利用 MODFLOW 预测研究区流场分布情况,如图 6 所示,研究区地下水流向总体沿东南流向西北,利用 MT3D 模块预测 5a、10a、20a 后研究区硫酸根浓度分布情况如图 7.

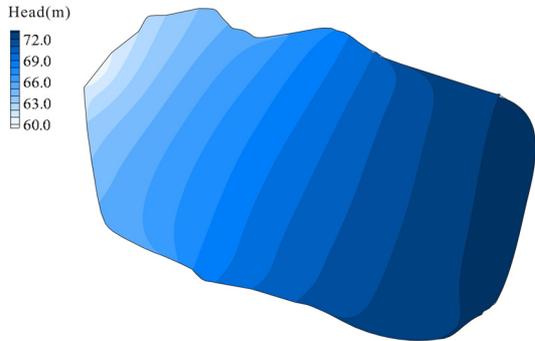


图 6 研究区地下水流场分布  
Fig.6 Distribution of groundwater flow

如模型预测污染质分布图所示,由于污染质在迁移中主要受对流因素的水动力影响,因此硫酸根离子迁移方向与地下水流向基本一致(东南流向西北).整个模拟期内污染羽面积随时间逐渐扩大:5a 后硫酸根污染羽最大运移距离 4275m,污染面积 21.7km<sup>2</sup>;10a 后污染羽前沿到达浑河地表水体,届时浑河下游硫酸根浓度必定会有所提高;20a 后污染羽将以更大规模侵入浑河,届时污染影响面积将达 26.5km<sup>2</sup>.

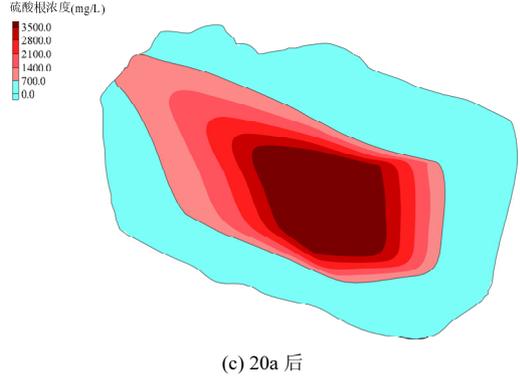
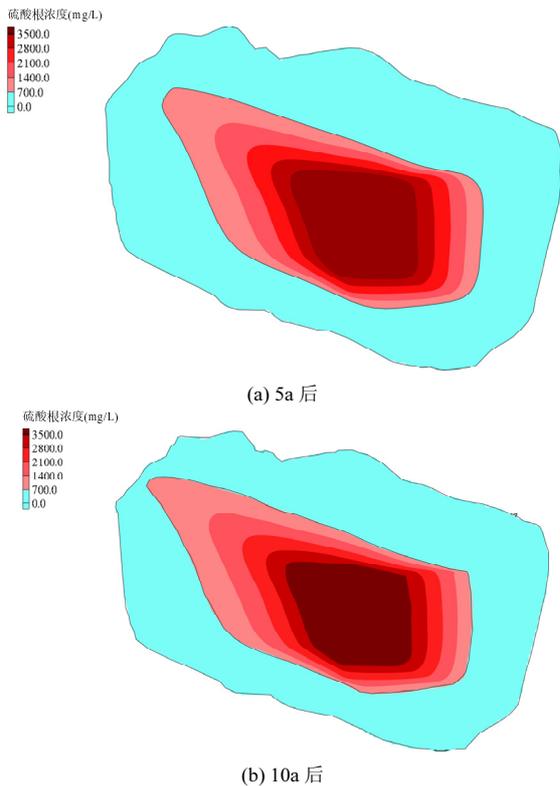


图 7 预测未来 5~20a 研究区硫酸根浓度分布  
Fig.7 Predicted distribution of sulfate ions after 5~20a

### 2.2 假想例子的反演求解

根据前文所述的研究区情况,本次研究构建了一个假想例子.假设研究区内存在如图 8 所示的 5 口监测井,其监测数据为已知条件,待求煤矸石堆的污染源源强和 2 个分区的渗透系数.

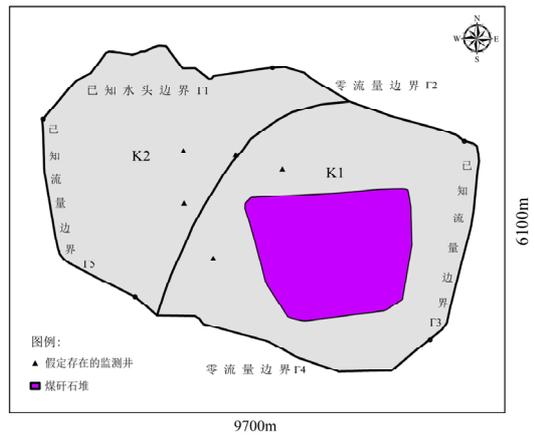


图 8 研究假设区监测井分布  
Fig.8 Distribution of monitoring wells in the study area

2.2.1 替代模型的建立 根据 Jin(2005)的研究<sup>[22]</sup>,为建立精度较高的 Kriging 替代模型,抽样组数 P 应满足以下公式:

$$P \geq (n+1)(n+2)/2 \tag{14}$$

式中:n 为待求变量的个数.

对两个分区的渗透系数 K 和污染源强 S 利用拉丁超立方抽样方法抽样 40 组,组合后输入模拟模型,以 5 个监测点的污染质浓度作为输出,由此建立了 40 组模拟模型的输入-输出数据集作为训练样本.另抽样 10 组检验组样本以验证替代模型精度.本文采用 R<sup>2</sup> 和平均相对误差 MRE 两个指标衡量替代模型

精度.其中 MRE 的数学表达式如下:

$$MRE = \frac{1}{n} \sum_{i=1}^n |y_i - y_{si}| / y_i \quad (15)$$

式中: $n$  为样本组数; $y_i$  为模拟模型输出值; $y_{si}$  为替代模型输出值.

表 4 替代模型与模拟模型拟合精度

Table 4 Fitting of surrogate model and simulation model

衡量指标	替代模型	
	Kriging 方法	BP 神经网络方法
$R^2$	0.9987	0.9956
MRE(%)	0.327	1.534

由表 4 可知,Kriging 替代模型输出的平均相对误差为 0.3%,BP 神经网络替代模型输出的平均相对误差为 1.5%.替代模型模拟结果拟合情况如图 9 所示.由此可见,两种方法建立的替代模型与模拟模型拟合精度较高,且 Kriging 方法优于 BP 神经网络方法.

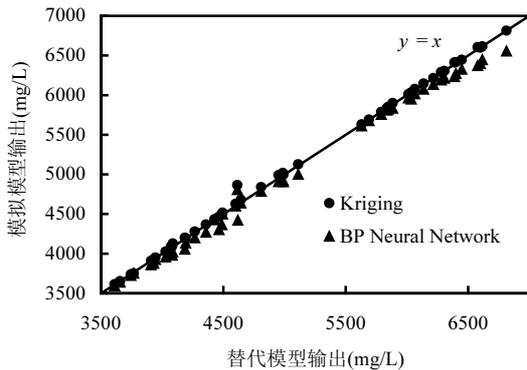


图 9 替代模型和模拟模型输出对比

Fig.9 Contrast of output between surrogate model and simulation model

2.2.2 优化模型的建立与求解 本研究以各监测井模拟计算值和实际观测值之间的误差极小化为目标函数,以污染源源强和渗透系数为决策变量,构建了针对污染源源强及渗透系数反演识别的优化模型.将模拟模型的替代模型以约束条件的形式嵌入到优化模型中.构建的优化模型形式如下:

$$\begin{aligned} \min Z &= \sum_{i=1}^5 (C_i^{\text{cal}} - C_i^{\text{obs}})^2 \\ \text{s.t.} &\begin{cases} C_i^{\text{cal}} = \text{Kriging}(X_0) \\ C_i^{\text{cal}} \geq 0 \\ C_i^{\text{obs}} \geq 0 \end{cases} \end{aligned} \quad (16)$$

式中: $C_i^{\text{cal}}$  为替代模型计算的各监测井浓度,表示实际浓度值满足地下水溶质运移规律; $C_i^{\text{obs}}$  为各监测点实际浓度,表示实际场地浓度的非负性; $Z$  为优化模型中的目标函数,即为各监测井的模拟计算值与实际观测值的差异函数.

利用模拟退火算法求解优化模型,两种方法计算识别值如表 5 所示.

表 5 基于 Kriging 方法和 BP 神经网络方法的计算识别值与实际场地值与对比

Table 5 Contrast of Kriging and BP neural network output between simulation model and optimization model

模型	参数	实际场地值	计算识别值	绝对误差	相对误差 (%)
Kriging 方法	源强(mg/d)	7000	6987	13.0	0.18
	东部地区渗透系数(m/d)	82.7	83.7	1.0	1.15
	西部地区渗透系数(m/d)	97.4	96.4	1.0	1.07
BP 神经网络方法	源强(mg/d)	7000	7031	31.0	0.44
	东部地区渗透系数(m/d)	82.7	78.6	4.1	4.96
	西部地区渗透系数(m/d)	97.4	96.7	0.7	0.72

由上表可知,优化模型取得了较好的反演结果.基于 Kriging 方法对污染源源强和渗透系数的反演识别相对误差均不大于 1.2%,运行时间为 63min; BP 神经网络方法对污染源源强和渗透系数的反演识别相对误差均不大于 5%,运行时间为 15min,两种方法识别精度较高.综合比较,Kriging 替代模型与模拟模型的拟合精度较高,而 BP 神经网络方法在优化模型中的调用速度更快.综上,基于两种耦合方法的模拟-优化方法均可有效完成反演的联合识别工作,为修复方案设计、污染风险评估和污染责任认定提供可靠依据.

### 3 结论

3.1 根据数值模拟的结果,预测整个模拟期内污染扩散范围随时间逐渐扩大:10a 后污染羽前沿到达浑河地表水体,浑河下游硫酸根浓度将会有所提高;20a 后污染羽将以更大规模侵入浑河,届时污染影响面积将达 26.5km<sup>2</sup>.

3.2 Kriging 替代模型输出的平均相对误差为 0.3%,BP 神经网络替代模型输出的平均相对误差为 1.5%.

两种方法建立的替代模型与模拟模型拟合精度较高,且 Kriging 方法优于 BP 神经网络方法。

**3.3** 本次研究利用基于两种耦合方法的模拟-优化方法对污染源源强和渗透系数进行反演识别,利用 Kriging 方法对污染源源强和渗透系数的反演识别相对误差均不大于 1.2%,利用 BP 神经网络方法对污染源源强和渗透系数的反演识别相对误差均不大于 5%,二者识别精度均较高。综上,利用基于两种耦合方法的模拟-优化方法对污染源源强和渗透系数进行联合识别是切实可行的。

#### 参考文献:

- [1] 薛禹群,张幼宽.地下水污染防治在我国水体污染控制与治理中的双重意义 [J]. 环境科学学报, 2009,29(3):474-481.  
Xue Y Q, Zhang Y K. Two fold significance of ground water pollution prevention in China's water pollution control [J]. Acta Scientiae Circumstantiae, 2009,29(3):474-481.
- [2] 徐凤兰,叶丹,曹德福等.浅谈地下水污染及其防治 [J]. 地下水, 2005,(1):50-52.  
Xu F L, Ye D, Cao D F, et al. Discussion on groundwater pollution and its prevention [J]. Groundwater, 2005,(1):50-52.
- [3] 张宇.基于随机统计方法的地下水污染源反演识别研究 [D]. 长春:吉林大学, 2016.  
Zhang Y. Groundwater pollution source identification based on random statistical method [D]. Changchun: Jilin University, 2016.
- [4] Borah T, Bhattacharjya R K. Development of unknown pollution source identification models using GMS ANN-based simulation optimization methodology [J]. Journal of Hazardous Toxic and Radioactive Waste, 2015,19(3):DOI:10.1061/(ASCE)HZ.2153-5515.0000242.
- [5] 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.  
Xiao C N, Lu W X, An Y K, et al. Simulation-optimization model of identification of groundwater pollution sources based on two coupling method [J]. China Environmental Science, 2015,35(8):2393-2399.
- [6] Ayvaz M T. A hybrid simulation-optimization approach for solving the areal groundwater pollution source identification problems [J]. Journal of Hydrology, 2016,538:161-176.
- [7] 陈华根,吴健生,王家林,等.模拟退火算法机理研究 [J]. 同济大学学报(自然科学版), 2004,(6):802-805.  
Chen H G, Wu J S, Wang J L, et al. Mechanism Study of Simulated Annealing Algorithm [J]. Journal of Tongji University, 2004,(6):802-805.
- [8] Jha M, Datta B. Three-dimensional groundwater contamination source identification using adaptive simulated annealing [J]. Journal of Hydrologic Engineering, 2013,18(3):307-317.
- [9] Jha M K, Datta B. Simulated annealing based simulation-optimization approach for identification of unknown contaminant sources in groundwater aquifers [J]. Desalination and Water Treatment, 2011,32(1-3):79-85.
- [10] Bahrami S F, Ardejani F D, Baafi E. Application of artificial neural network coupled with genetic algorithm and simulated annealing to solve groundwater inflow problem to an advancing open pit mine [J]. Journal of Hydrology, 2016,536:471-484.
- [11] Lin Y P, Tan Y C, Rouhani S. Identifying spatial characteristics of transmissivity using simulated annealing and kriging methods [J]. Environmental Geology, 2001,41(1/2):200-208.
- [12] Bahrami S, Ardejani F D. Prediction of pyrite oxidation in a coal washing waste pile using a hybrid method, coupling artificial neural networks and simulated annealing (ANN/SA) [J]. Journal of Cleaner Production, 2016,137:1129-1137.
- [13] 肖传宁,卢文喜,赵莹,等.基于径向基函数模型的优化方法在地下水污染源识别中的应用 [J]. 中国环境科学, 2016,36(7):2067-2072.  
Xiao C N, Lu W X, Zhao Y, et al. Optimization method of identification of groundwater pollution sources based on radial basis function model [J]. China Environmental Science, 2016,36(7):2067-2072.
- [14] 范越,卢文喜,欧阳琦,等.基于 Kriging 替代模型的地下水污染监测井网优化设计 [J]. 中国环境科学, 2017,37(10):3800-3806.  
Fan Y, Lu W X, Ouyang Q, et al. Optimum design of groundwater pollution monitoring well network based on Kriging surrogate model. China Environmental Science, 2017,37(10):3800-3806.
- [15] Park D-I, Ryou K B. Comparative studies of Kriging methods for estimation of groundwater level distribution [J]. Journal of Korean Society of Water Science and Technology, 2016,24(6):137-148.
- [16] 高云凯,孙芳,余海燕.基于 Kriging 模型的车身耐撞性优化设计 [J]. 汽车工程, 2010,32(1):17-21.  
Gao Y K, Sun F, Yu H Y. Crashworthiness optimization of carbody based on Kriging Surrogate Model [J]. Automotive Engineering, 2010,32(1):17-21.
- [17] Zhao Y, Lu W X, Xiao C N. A Kriging surrogate model coupled in simulation-optimization approach for identifying release history of groundwater sources [J]. J. Contamination Hydrology, 2016,185:51-60.
- [18] Juang K W, Lee D Y. Comparison of three nonparametric Kriging methods for delineating heavy-metal contaminated soils [J]. Journal of Environmental Quality, 2000,29(1):197-205.
- [19] Datta B, Chakrabarty D, Dhar A. Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters [J]. Journal of Hydrology, 2009,376(1/2):48-57.
- [20] Brooks R H, Corey A T. Hydraulic properties of porous media [M]. Hydrology Paper no. 3, Civil Engineering Dep., Colorado State Univ., Fort Collins, Colo. 1964:892-898.
- [21] 侯泽宇,卢文喜,王宇.基于替代模型的地下水 DNAPLs 污染源反演识别 [J]. 中国环境科学, 2019,39(1):188-195.  
Hou Z Y, Lu W X, Wang Y. Surrogate-based source identification of DNAPLs-contaminated groundwater [J]. China Environmental Science, 2019,39(1):188-195.
- [22] Jin Y. A comprehensive survey of fitness approximation in evolutionary computation [J]. Soft Computing, 2005,9(1):3-1.

**作者简介:** 潘紫东(1996-),男,辽宁沈阳人,吉林大学硕士研究生,主要从事地下水污染源反演识别研究。