

文章编号:0559-9350(2017)07-0816-09

地质统计学法在地下水污染溯源中的应用及参数敏感性分析

龙玉桥^{1,2,3}, 崔婷婷^{1,2}, 李伟^{1,2}, 吴春勇^{1,3}, 李砚阁¹(1. 南京水利科学研究院, 江苏 南京 210029; 2. 水文水资源与水利工程科学国家重点实验室, 江苏 南京 210098;
3. 南京瑞迪建设科技有限公司, 江苏 南京 210029)

摘要: 为了确定地下水污染物来源, 将地质统计学方法应用于追溯一维均质含水层中污染源的排放过程, 利用数值实验分析了观测点布设位置、污染物浓度观测误差、水动力弥散参数和地下水流速对地质统计学法溯源效果的影响。结果表明, 观测点提供的浓度信息能准确反映观测时刻污染物浓度的分布时, 根据这些浓度信息可得到好的溯源效果。反演出的污染物排放过程与真实排放过程相似度较高时, 污染物浓度测量误差的数量级介于 $10^{-6} \sim 10^{-3.5}$ 之间, 水动力弥散系数的变化幅度在 $-10\% \sim 5\%$ 内, 实际平均流速的变化幅度在 $\pm 2\%$ 间。在本次数值实验条件下, 地质统计学法给出了较好的溯源效果, 与观测浓度误差相比, 水动力弥散参数和地下水流速对溯源效果影响更显著。

关键词: 地质统计学; 地下水; 污染; 溯源; 敏感性**中图分类号:** X143**文献标识码:** A**doi:** 10.13243/j.cnki.slxb.20161173

1 研究背景

由于地下水污染具有极强的隐蔽性, 当发现污染时, 往往不易确定污染物的来源, 这为锁定污染肇事者, 判定其责任大小和污染治理带来了巨大困难。据全国118座大城市浅层地下水调查结果, 97.5%的城市受到不同程度的污染, 其中40%的城市受到严重污染^[1]。2011年, 全国城市55%的地下水呈现较差至极差的水质^[2]。可是在面对严峻的地下水污染危机的同时, 在我国人口较为密集、人类活动干扰大、工农业生产发达的平原地区, 地下水正遭受着越来越严重的污染^[3]。因此我国地下水污染治理工作任重道远。治理地下水污染的先决条件和必要基础就是确定地下水污染源的位置及其排放历史。地下水污染源(地下水污染源解析)是通过有限的观测数据, 估计污染源的位置及污染物迁移转化的历史^[4]。它主要包括追溯污染物排放历史、确定污染源位置和估计污染物排放量^[5]。地下水污染源的研究有助于判别不同污染源或污染肇事者的责任大小, 有助于确定污染治理成本在不同污染肇事者间的分配比例, 也有助于制定和选择经济的污染治理策略与方法^[6]。在当前我国地下水资源供需矛盾突出, 地下水环境不断恶化的情况下, 加强地下水污染溯源的有关研究对于保护地下水资源、防治地下水污染有着重要的实际意义。

溯源方法可分为地球化学足迹法和数学模拟法^[7]。本研究关注的是数学模拟法, 即通过概化实际问题并建立数学模型, 进而反演出污染源的位置与排放量(排放历史过程)等信息, 是一种典型的逆问题(不适定问题)。该类问题是地下水科学及相关学科的研究难点与热点。采用数学方法追溯地下水污染源的问题在国外有近30年的研究历史, 而我国也有学者发表了相关研究成果^[6,8-10]。

目前, 地下水污染溯源的主要方法有解析方法、优化方法、直接法、随机理论和地质统计学方

收稿日期: 2016-11-10; 网络出版日期: 2017-06-20

网络出版地址: <http://www.cnki.net/kcms/detail/11.1882.TV.20170620.1345.001.html>

基金项目: 国家自然科学基金项目(51409161, 51509157); 江苏省自然科学基金项目(BK20140080); 水利部技术示范项目(SF-201706)

作者简介: 龙玉桥(1984-), 男, 广东英德人, 博士, 主要从事地下水数值模拟、水资源管理方面的研究。E-mail: yqlong@nhri.cn

法^[11]。Snodgrass 等^[12]采用地质统计学方法确定污染物排放的历史过程，这一方法的解更具有普适性，且不需要对未知污染源的特性和结构做出盲目的假设，但是此方法的限制是污染源的位置必须是先验的^[11]。Butera 等^[13]将这一方法推广到二维溯源问题，Michalak 等^[14]将地质统计分析与伴随法相结合，建立了追溯三维非均质污染物迁移的方法。Butera 等^[15]于 2012 年在其研究基础上将该方法拓展到同时追溯污染物排放过程和污染源位置。尽管诸多学者对地质统计学方法进行了大量的研究，但这些研究多集中于方法的基本原理，缺乏观测数据、水动力学参数对溯源效果影响的探讨。

本文将地质统计学方法应用于追溯一维均质含水层中污染源的排放过程，并通过数值实验探讨观测点布置、观测数据误差、水动力弥散系数和实际平均流速对溯源结果的影响。

2 基本原理

Snodgrass 等^[12]采用地质统计学方法确定了污染物排放的历史过程，并将污染物的排放过程作为关于时间的未知函数。由于函数存在不确定性，因此它代表了一个随机过程。符合污染物观测数据和溯源问题附加条件的函数构成了一个集合。集合中每个函数对应一个概率，每个概率代表“相应函数是溯源问题的解”的可能性。集合的数学期望是未知函数的最佳估计，其协方差度量了该估计的不确定性。

2.1 地质统计学模型 Snodgrass 等^[12]将污染物排放过程与污染物浓度观测值间的关系概化为式(1)：

$$\mathbf{z} = \mathbf{H}\mathbf{s} + \mathbf{v} \quad (1)$$

式中： $\mathbf{z}$ 为一个由观测数据组成的 $m \times 1$ 维向量； $\mathbf{s}$ 为待求未知函数经离散后形成的一个 $n \times 1$ 维向量； $\mathbf{v}$ 为测量误差组成的 $m \times 1$ 维向量； $\mathbf{H}$ 为传递函数构成的灵敏度矩阵。

假定 $\mathbf{v}$ 的均值为 0，其协方差矩阵为 $\mathbf{R}$ ， $\mathbf{s}$ 为一个随机向量，其数学期望和协方差分别由式(2)和式(3)表示。

$$E[\mathbf{s}] = \mathbf{X}\boldsymbol{\beta} \quad (2)$$

$$\mathbf{Q}(\theta) = E[(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T] \quad (3)$$

式中： $\mathbf{X}$ 为一个已知的 $n \times p$ 的矩阵； $\boldsymbol{\beta}$ 为 p 个漂移系数组成的向量； $\mathbf{Q}(\theta)$ 为关于未知参数 θ 的已知函数，通常取高斯分布函数。

2.2 地质统计学模型的求解方法 待求未知函数离散形式 $\mathbf{s}$ 的求解过程可分成两个步骤，首先是确定未知参数 θ ，而后估计未知函数 $\mathbf{s}$ 。根据文献[12, 16]，当式(4)表示的概率取最大值时， θ 的取值即为待求的参数值。

$$p(\mathbf{z}|\theta) \propto |\boldsymbol{\Sigma}|^{-1/2} |\mathbf{X}^T \mathbf{H}^T \boldsymbol{\Sigma}^{-1} \mathbf{H} \mathbf{X}|^{-1/2} \exp\left[-\frac{1}{2} \mathbf{z}^T \boldsymbol{\Xi}^{-1} \mathbf{z}\right] \quad (4)$$

$$\boldsymbol{\Sigma} = \mathbf{H} \mathbf{Q} \mathbf{H}^T + \mathbf{R} \quad (5)$$

$$\boldsymbol{\Xi} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \mathbf{H} \mathbf{X} (\mathbf{X}^T \mathbf{H}^T \boldsymbol{\Sigma}^{-1} \mathbf{H} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{H}^T \boldsymbol{\Sigma}^{-1} \quad (6)$$

式(4)描述的极值问题，等同于求式(7)的最小值。

$$L(\theta) = \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{1}{2} \ln |\mathbf{X}^T \mathbf{H}^T \boldsymbol{\Sigma}^{-1} \mathbf{H} \mathbf{X}| + \frac{1}{2} \mathbf{z}^T \boldsymbol{\Xi}^{-1} \mathbf{z} \quad (7)$$

$L(\theta)$ 对 θ 的导数取 0 值时， $L(\theta)$ 取最小值。采用 Gauss-Newton 迭代法求解此极值问题。当迭代计算收敛后，再利用 θ 求 $\mathbf{Q}$ ，并代入式(8)，解得到 $n \times m$ 的系数矩阵 $\mathbf{A}$ 和 $p \times n$ 的拉格朗日乘子阵 $\mathbf{M}$ 。

$$\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{H} \mathbf{X} \\ (\mathbf{H} \mathbf{X})^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{H} \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix} \quad (8)$$

对未知函数 $\mathbf{s}$ 的最佳估计 $\hat{\mathbf{s}}$ 及其协方差分别如式(9)和式(10)所示。

$$\hat{\mathbf{s}} = \mathbf{A} \mathbf{z} \quad (9)$$

$$\mathbf{V} = -\mathbf{X} \mathbf{M} + \mathbf{Q} - \mathbf{Q} \mathbf{H}^T \mathbf{A}^T \quad (10)$$

2.3 污染物浓度非负性约束 上述求解方法未考虑污染物浓度的非负性，为体现污染物浓度的非负性，对未知函数 s 进行如下变换^[12]：

$$\tilde{s} = \alpha \left(s^{1/\alpha} - 1 \right) \quad (11)$$

将式(11)代入式(1)得到

$$z = h \left[\left(\frac{\tilde{s} + \alpha}{\alpha} \right)^\alpha \right] + v = \tilde{h}(\tilde{s}) + v \quad (12)$$

实施上述变换后， $\tilde{h}(\tilde{s})$ 是非线性的，需要通过迭代方法求得 θ 和 $\tilde{s}$ ^[12,16]。最终利用式(13)求出未知函数 s 的最佳估计 $\hat{s}$ 。

$$\hat{s} = \left(\frac{\tilde{s}_l + \alpha}{\alpha} \right)^\alpha \quad (13)$$

3 数值算例与实验方案

3.1 数值算例 本文以一维均质饱和多孔介质中惰性污染物的迁移问题为例^[17]讨论观测点布置、观测数据误差、水动力弥散系数和实际平均流速对溯源结果的影响。算例中的地下水流场为等速均匀流场，实际平均流速为常数；初始时刻，研究域中无污染物；从初始时刻开始，在研究域的左端连续注入污染物，污染物是保守的，不发生吸附和衰变，其对流弥散是一维的。该问题可由式(14)——(17)表示。

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - v \frac{\partial C}{\partial x} \quad (14)$$

$$C(x, T) = g(x) \quad (15)$$

$$C(0, t) = 0 \quad 0 \leq t \leq T \quad (16)$$

$$C(\infty, t) = 0 \quad 0 \leq t \leq T \quad (17)$$

式中： C 为污染物浓度(ML^{-3})； D 为水动力弥散系数(L^2T^{-1} ，文中取1)； v 为实际平均流速(LT^{-1} ，文中取1)； x 为距离(L)，算例中 $x \in [0, 300]$ ； t 为时间(T)。

上述问题在 T 时刻 x 处的解析解为：

$$C(x, T) = \int_0^T s(\tau) f(x, T - \tau) d\tau \quad (18)$$

$$f(x, T - \tau) = \frac{x}{2\sqrt{\pi D(T - \tau)^3}} \exp \left[-\frac{(x - v(T - \tau))^2}{4D(T - \tau)} \right] \quad (19)$$

污染物的真实排放过程由式(20)描述(图1)。沿 x 方向均匀布置有25个观测点(图2)，各观测点在 $t = 300$ 时刻的观测值 z 所形成的污染物浓度曲线如图2所示，其测量误差的协方差阵 $\mathbf{R} = \sigma_R^2 \mathbf{I}$ ，其中 $\sigma_R^2 = 1 \times 10^{-12}$ 。待估计排放过程 s 的协方差阵 $\mathbf{Q}$ 由式(21)表示^[12]。

$$s(t) = \exp \left[-\frac{(t - 130)^2}{50} \right] + 0.3 \exp \left[-\frac{(t - 150)^2}{200} \right] + 0.5 \exp \left[-\frac{(t - 190)^2}{98} \right] \quad (20)$$

$$Q(t_i, t_j | \theta) = \sigma^2 \exp \left[-\frac{(t_i - t_j)^2}{l^2} \right] \quad (21)$$

3.2 数值实验方案

(1)观测点布置对溯源效果的影响。观测数据的详实程度是影响求解溯源问题难易程度的重要因素^[18]。在实际应用中，观测点(井、站)的数量受到建设或维护成本的制约，其位置受到建设场地的

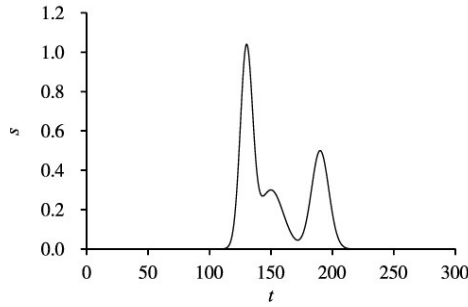


图1 污染物浓度排放曲线

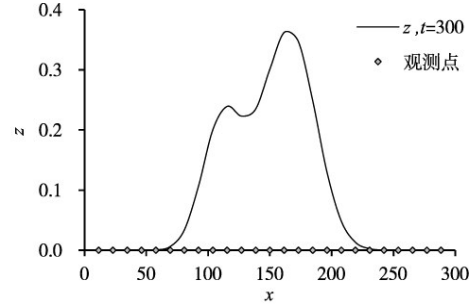


图2 观测点分布及 $t=300$ 时刻的浓度分布

限制。分析观测点数量和位置对溯源效果的影响,有助于设计适当观测点数量和位置为反演污染排放过程提供足够的、有效的浓度观测信息。本研究先在 $0 \sim 300$ 的范围内均匀布置不同数量的观测点,分析观测点数量对溯源结果的影响,而后将固定数量的观测点均匀布置在不同宽度的观测域内,并使各观测域以 $x=0$ 为起点逐渐移动至 $x=300$ 处,分析观测域宽度及位置对溯源效果的影响。

(2)溯源效果对观测数据误差的敏感性。溯源问题的解对观测数据的误差有极强的敏感性,即便是很小的污染物浓度测量误差也能使溯源结果发生巨大的变化^[17],因此分析溯源方法对观测数据误差的敏感性有助于确定适当的测量误差的范围以保证能够取得较好溯源结果。文献[12]中将 σ_R^2 设定为 10^{-12} ,以控制测值测量误差,并取得了良好的溯源效果。本文将 σ_R^2 从 10^{-12} 增大到 10^{-1} ,通过调节观测值测量误差的协方差阵 $\mathbf{R}$,分析地质统计学方法的溯源效果对观测数据误差的敏感性。

(3)溯源效果对水动力弥散系数的敏感性。水动力弥散系数由介质参数弥散度、水流实际速度和污染物在介质中的分子扩散系数确定。虽然可通过实验方法得到水动力弥散系数^[19],但多数溶质迁移研究中,常在数学模型参数校准过程中确定弥散度^[20]。实验和参数校准方法产生的误差将通过水动力弥散系数引入溯源问题,并将对污染物溯源效果产生影响。因此分析溯源方法对观测数据误差的敏感性有助于确定适当的水动力弥散系数的范围以保证能够取得较好溯源结果。上述数值算例中,水动力弥散系数 D 的实际值为 1,本文调节 D 在 $\pm 15\%$ 间变化,分析溯源方法对弥散参数的敏感性。

(4)溯源效果对实际平均流速敏感性。实际平均流速影响到污染物的对流与弥散迁移。通过实验方法或数学模型计算可得到地下水流速,相应的测量误差与模型不确定性^[21]将通过地下水流速引入溯源问题中,从而影响到污染物溯源的效果。因此分析溯源方法对观测数据误差的敏感性有助于确定适当的实际平均流速范围以保证能够取得较好溯源结果。上述数值算例中,实际平均流速 v 的实际值为 1,本文调节 v 在 $\pm 15\%$ 间变化,分析溯源方法对实际平均流速误差的敏感性。

3.3 评价指标 本文采用线性相关系数和置信区间宽度作为评价地质统计学方法溯源效果对观测数据误差、水动力弥散系数和实际平均流速敏感性的指标。线性相关系数 r 反映反演出的排放过程 cal 与实际排放过程 Rea 的相似程度,其计算公式如式(22)所示。 r 越近于 1, cal 与 Rea 相似程度越高。置信区间宽度反映了根据观测数据和模型参数反演出的污染物排放过程的不确定性的。本文采用反演出的排放过程的 95% 置信区间上限 σu 和下限 σl 间的欧氏距离 de 评价其置信区间的宽度,其计算公式如式(23)所示。 de 越小,说明反演出的排放过程的不确定性越小。指标 I 对第 i 个参数 a_i 的敏感度 β_i 的定义如式(24)所示,其中 Δa_i 为参数 a_i 的变化量。

$$r = \frac{\sum_{i=1}^n (cal_i - \overline{Rea})(Rea_i - \overline{Rea})}{\sqrt{\sum_{i=1}^n (cal_i - \overline{Rea})^2} \sqrt{\sum_{i=1}^n (Rea_i - \overline{Rea})^2}} \quad (22)$$

$$de = \sqrt{\sum_{i=1}^n (\sigma u_i - \sigma l_i)^2} \quad (23)$$

$$\beta_i = \frac{(I(a_i + \Delta a_i) - I(a_i)) / I(a_i)}{\Delta a_i / a_i} \quad (24)$$

4 分析与讨论

4.1 观测点布置对溯源效果的影响

4.1.1 数量的影响 由上述算例可知, 在 $x \in [0, 300]$ 的范围内, 随着观测点数量的增加, r 趋向于 1, de 趋向于 5.44 (图 3(a))。计算的空间范围固定时, 观测点数量影响到观测点布置密度。观测点数量小于 10 时, 相应的布置密度小于 0.033, r 值小于 0.92, de 值大于 10.2。此时, 较少的观测点提供的浓度信息难以反映出 $t=300$ 时污染物浓度的分布形态 (图 4), 根据这些浓度信息反演出的污染物排放过程不准确。随着观测点数量的增加, 观测点提供的浓度信息逐渐接近 $t=300$ 时污染物浓度的分布形态。当观测点数量大于 25 时, 观测点提供的浓度信息已与 $t=300$ 时污染物浓度的分布形态差异很小。再增加观测点数量, 浓度信息的改善并不显著, 因此反演的效果也没有较大的提升, 此时 r 和 de 均趋于稳定。这说明在 $x \in [0, 300]$ 的范围内, 均匀布置的 25 个观测点能为该数值算例提供足够有效的浓度信息观测数据。

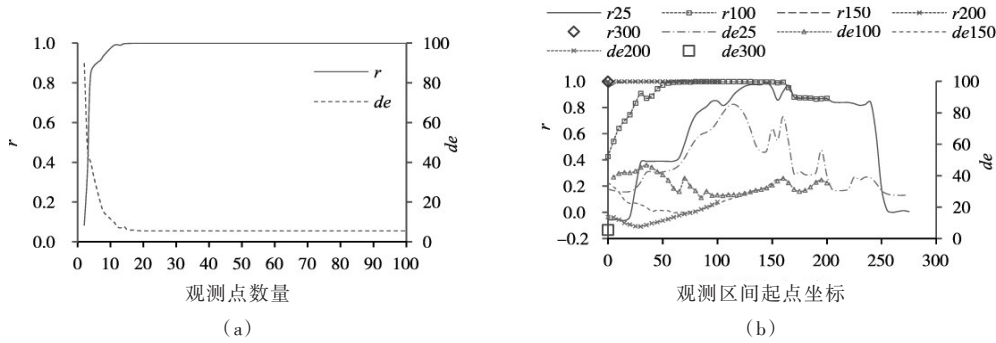


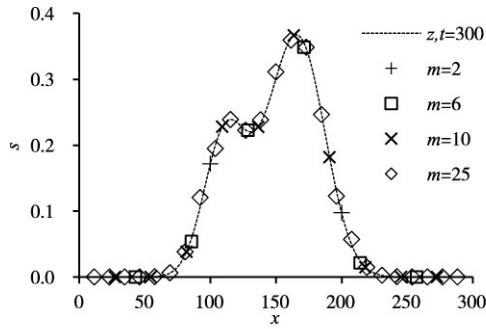
图3 观测点布置对 r 及 de 的影响

4.1.2 位置的影响 在不同宽度的观测域内均匀布置 25 个观测点, 并使各观测域以 $x=0$ 为起点逐渐移动至 $x=300$ 处, 不同宽度观测域对应的溯源效果如图 3(b) 所示。为使图形清晰, 本文仅在图形中显示了宽度为 25、100、150、200、300 的观测域对应的 r 和 de 的曲线。由图 3(b) 可知, 观测域较小时, 观测域布置在溯源计算范围的中部得到的溯源效果相对较好, 如观测域宽度为 25 且起点位于 $x \in [130, 230]$ 内时, r 取值大于 0.8, 相应的 $de \in [30, 70]$, 而观测域起点在其他范围时, r 取值小于 0.8, 相应的 $de \in [30, 80]$ 。随着观测域宽度由 25 增大至 300, r 越容易接近于 1.0, de 也更易接近于 5.44, 反演的效果逐渐变好。

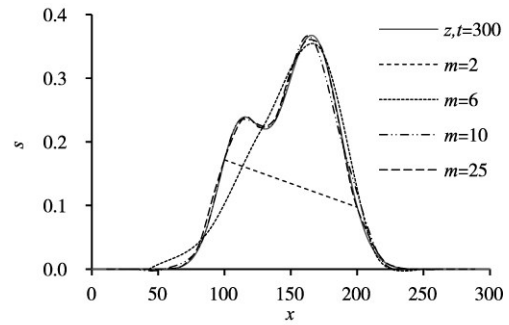
观测域宽度较窄时, 观测点难以捕捉到全面的浓度信息, 根据残缺的信息难以反演出污染物排放过程, 因此越窄的观测域对应的溯源效果越差。浓度分布的峰值与低谷是描述浓度分布的关键信息, 根据这些关键信息能够得到一定准确程度的浓度分布, 观测域较窄但涵盖了提供关键信息的观测点时, 这些关键信息对应的溯源效果相对较好。在本例中, 浓度变化的峰值与低谷集中在 $t=300$ 时刻浓度分布曲线的中部, 因此观测域布置在溯源计算范围的中部得到的溯源效果相对较好。

4.2 溯源效果对观测数据误差的敏感性 文献 [12] 将 σ_R^2 取为 10^{-12} , 得到了很好的溯源效果。本文将 10^{-12} 作为评价 σ_R^2 值对溯源效果影响的基准值, 此时 r 接近于 1.0, de 接近于 5.44。图 5 显示了 r 和 de 随着 σ_R^2 变化的情况及敏感性。由数值实验结论可知, 当 σ_R^2 值小于 10^{-12} 时, 再减小 σ_R^2 值, 并不能使溯源效果有显著的提高。随着 σ_R^2 值不断变大, 溯源效果逐渐变差, 在 σ_R^2 值大于 10^{-8} 后, 随着 σ_R^2 值的增大溯源效果急剧变差, 排放浓度曲线的峰值大小逐渐减小, 峰值出现的位置逐渐向时间轴的中部靠近。

根据文献 [22], 污染物的检出限的数量级一般大于或等于 10^{-7} , 因此误差的数量级大于或等于



(a) 不同数量的观测点位置及 $t=300$ 时刻的浓度分布



(b) 根据不同数量观测点插值出的浓度曲线

图4 不同数量观测点分布及其反映的 $t=300$ 时刻浓度分布情况

(注: 各浓度曲线根据相应浓度点插值得到。)

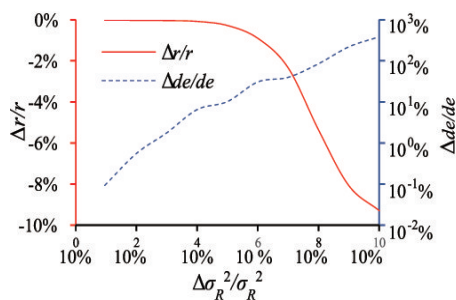


图5 观测数据误差对 r 及 de 的影响

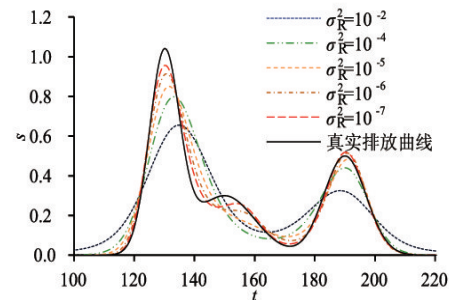


图6 不同 σ_R^2 值对应的计算浓度排放曲线与真实排放曲线对比

(注: 为使图形清晰, 仅显示 $t \in [100, 220]$ 内的浓度排放曲线。)

10^{-7} 时, 溯源结果对误差敏感性更具有实际意义。算例中观测值 z 的测量误差的协方差阵为 $\mathbf{R} = \sigma_R^2 \mathbf{I}$, σ_R^2 取 10^{-12} 时, 观测值 z 的测量误差的数量级为 10^{-6} , 说明测量误差在该数量级时能取得很好的溯源效果。当 σ_R^2 值增大至 10^{-7} 时, 测量误差的数量级为 $10^{-3.5}$, r 值为 0.996, de 值为 6.011, 较 σ_R^2 值为 10^{-12} 时, 分别变化了 -0.28 % 和 10.24 %, 此时计算的排放过程与实际排放过程较为相似, 较好地拟合了实际排放过程的浓度峰、谷值(图 6)。当 σ_R^2 值为 10^{-6} 时, 测量误差的数量级为 10^{-3} , r 值为 0.990 和 de 值为 7.130, 较 σ_R^2 值为 10^{-12} 时, 分别变化了 -0.90 % 和 30.77 %, 此时计算的排放过程较好的拟合了实际排放过程两端的浓度峰值, 但几乎反映不出中间浓度峰值。再增大 σ_R^2 值, 反演出的排放过程仅能拟合实际排放过程两端的浓度峰值, 已无法反映出中间浓度峰值。说明对于此算例, 当测量误差的数量级大于 $10^{-3.5}$ 时, 难以取得较好的溯源效果。

4.3 溯源效果对水动力弥散系数的敏感性 图 7 表明当 D 的变化幅度在 -15 % 与 15 % 间时, 相应的 r 变化幅度在 -35 % 与 5 % 间, de 变化幅度在 -15 % 与 65 % 间。当 D 的变化幅度在 -10 % 与 5 % 间时, 相应的 r 变化幅度在 -5 % 与 5 % 间, de 变化幅度在 ± 10 % 间, 且 de 出现了随 D 值减小而减小至 5.00 附近。尽管此时 r 值均在 0.9 以上, de 值均在 7 以下, 但根据图 8 可以看出 $D=0.9$ 对应的计算排放曲线在 $t=150$ 时刻的峰值已经消失, 而在 $t=240$ 附近出现了峰值, 反演出的排放曲线与真实的排放曲线已出现较大的不同, 而 $D=1.05$ 对应的反演出的排放曲线的峰、谷值较真实排放曲线更明显, 曲线动态的相似程度较高。 D 值增大时, 计算的污染物排放浓度峰值增大; D 值减小时, 计算的污染物排放浓度峰值减小。小的 de 值表明反演结果的不确定性变小了, 说明基于已知的观测数据和模型参数, 对污染物排放过程 s 的识别作用增强了, 但是因为模型参数信息有可能是错误的, 因此较小的 de 值并不意味着反演结果更接近于真实的排放过程。

4.4 溯源效果对的实际平均流速敏感性 图 9 表明当实际平均流速 v 变化幅度在 ± 5 % 间时, r 变化幅度在 -40 % 与 0 % 间, de 对 v 的敏感性较低, de 变化幅度在 ± 10 % 间。 v 值小于真实的实际平均流速, 反演出的污染物排放过程中的浓度峰、谷值将提早出现, v 值大于真实的实际平均流速, 反演出的污染

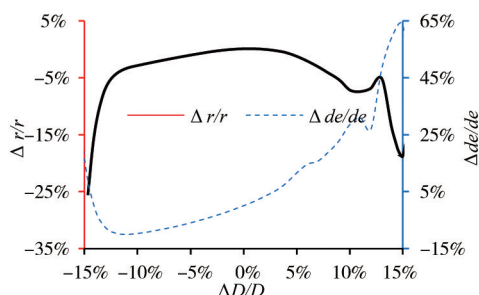


图7 水力弥散系数对 r 及 de 的影响

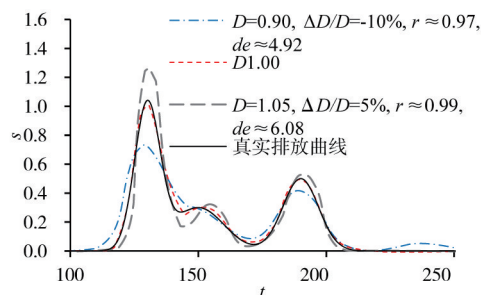


图8 不同 D 值对应的计算浓度排放曲线与真实排放曲线对比
(注:为使图形清晰,仅显示 $t \in [100, 250]$ 内的浓度排放曲线。)

物排放过程中的浓度峰、谷值将提早延迟出现(图10)。实际平均流速 v 变化幅度在 $\pm 2\%$ 间时,反演出的污染物排放过程与真实排放过程在曲线的变化上相似度较高,当 v 变化幅度超出 $\pm 2\%$ 时,由于浓度峰、谷值的偏差造成反演结果与真实排放过程相似性较差,反映出 r 对 v 的敏感性较高,故在上述一维算例中实际平均流速对是否能反演出正确的污染物排放过程是十分重要的。 v 变化幅度在 $\pm 5\%$ 间时,随着 v 值增大 de 值减小至5.00附近,说明基于此时的观测数据和模型参数,对污染物排放过程 s 的识别作用较强,但是因 v 值可能不准确,所以较小的 de 值对应的反演结果可能不同于真实的排放过程。

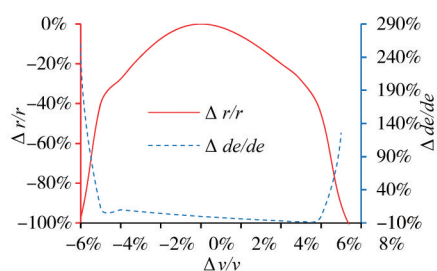


图9 实际平均流速对 r 及 de 的影响

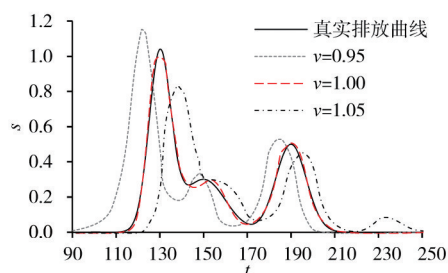


图10 不同 v 值对应的计算浓度排放曲线与真实排放曲线对比
(注:为使图形清晰,仅显示 $t \in [90, 250]$ 内的浓度排放曲线。)

4.5 不同参数敏感度比较 σ_R^2 控制着测量误差的协方差矩阵,该值与污染物浓度检测有关,是影响地质统计学方法溯源效果的外因。水力弥散系数 D 和实际平均流速 v 是描述污染物迁移的对流弥散方程的重要组成部分,是影响地质统计学方法溯源效果的内因。图11显示 r 和 de 对 σ_R^2 、 D 和 v 的敏感度。因敏感度存在负值且数值差异较大,为将其清楚地表达在同一张图上,图11将敏感度的绝对值绘制在对数坐标系中。由图11可知,当参数在 $\pm 6\%$ 间变化时, r 对参数 v 的敏感度最大,对 σ_R^2 的敏感度最小;参数在 $\pm 5\%$ 间变化时, de 对参数 v 和 D 的敏感度较接近,对 σ_R^2 的敏感度最小,参数变化范围超出在 $\pm 5\%$ 后, de 对参数 v 的敏感度急剧增大。在正向模拟计算时,包含 v 的对流计算越能准确反映实际的溶质迁移情况,弥散迁移项就会越小^[20],这说明准确的 v 在正向溶质迁移计算中是十分重要的。

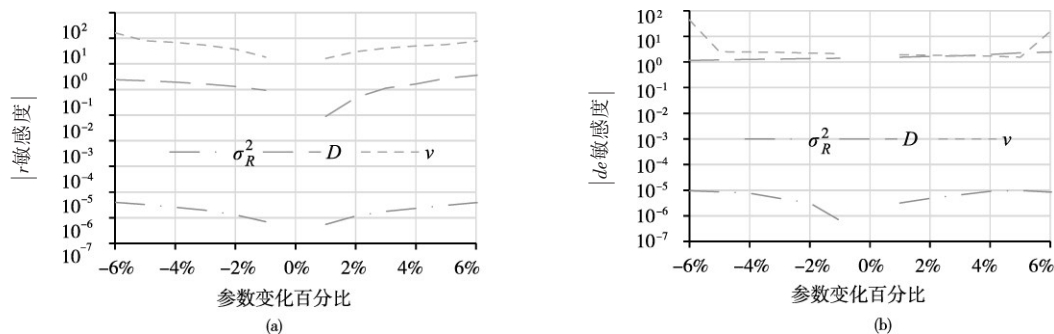


图11 r 和 de 对 σ_R^2 、 D 、 v 的敏感度

溯源问题是典型的逆问题, 重要因素的微小变化会引起计算结果产生较大的差异, 因此 r 和 de 对 v 最敏感, v 是影响溯源效果的重要因素。

5 结论

(1)观测点提供的浓度信息越能准确地反映观测时刻污染物浓度的分布, 根据这些浓度信息得到的溯源效果越好。设置观测点数量与位置的最终目的都是为了尽可能地准确反映观测时刻污染物浓度的分布。(2) σ_r^2 介于 10^{-12} 与 $10^{-3.5}$ 之间时, 观测值 z 的测量误差的数量级介于 10^{-6} 与 $10^{-3.5}$ 之间, 该方法在这一误差范围内能够取得较好的溯源结果。(3)水动力弥散系数 D 的变化量在 -10% 与 5% 之内时, 反演出的排放曲线与真实排放曲线较为一致, 溯源结果较好。(4)实际平均流速 v 的变化量在 $\pm 2\%$ 间时, 反演出的污染物排放过程与真实排放过程在曲线的变化上相似度较高, 溯源结果较好。(5) D 和 v 是影响地质统计学方法溯源效果的内因, 浓度测量误差影响地质统计学方法溯源效果的外因, 溯源结果对 D 和 v 最敏感, D 和 v 是影响溯源效果的关键因素。(6)研究成果对提高地质统计学溯源效果, 准确判定地下水污染源, 有效控制和防治地下水污染有重要应用价值。

参 考 文 献:

- [1] 姜建军, 文冬光. 合理开发利用地下水缓解水资源紧缺状况[J]. 中国水利, 2005(13): 36-39.
- [2] 新浪网. 数据显示全国城市 55 %地下水为较差至极差水质[EB/OL]. (2013-02-22)[2013-02-23]. http://news.sina.com.cn/c/2013-02-22/001726323489.shtml?bsh_bid=195660861.
- [3] 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(3): 474-481.
- [4] SUN A-y, PAINTER S L, WITTMAYER G W. A constrained robust least squares approach for contaminant source release history identification[J]. Water Resources Research, 2006, 42(4): W04414.
- [5] ZOI D, GEORGE F P. Optimal search strategy for the definition of a DNAPL source[J]. Journal of Hydrology, 2009, 376(3/4): 542-556.
- [6] LI G-s, TAN Y-j, CHENG J, et al. Determining magnitude of groundwater pollution sources by data compatibility analysis[J]. Inverse Problems in Science and Engineering, 2006, 14(3): 287-300.
- [7] MILNES E, PERROCHET P. Simultaneous identification of a single pollution point-source location and contamination time under known flow field conditions[J]. Advances in Water Resources, 2007, 30(12): 2439-2446.
- [8] 李功胜, 姚德, 马昱, 等. 一维溶质运移源(汇)项系数反演的迭代正则化算法[J]. 地球物理学报, 2008, 51(2): 582-588.
- [9] LONG Y-q, LI W, HUANG J. Advance of optimization methods for identifying groundwater pollution source properties[J/CD]. Applied Mechanics and Materials, 2012, 178: 603-608.
- [10] LONG Yu-qiao, WU Chun-yong, WANG Jian-ping. The influence of estimated pollution range on the groundwater pollution source identification method based on the simple genetic algorithm[J]. Applied Mechanics and Materials, 2014, 587: 836-841.
- [11] JULIANA A, AMVROSSIOS C B. State of the art report on mathematical methods for groundwater pollution source identification[J]. Environmental Forensics, 2001, 2(3): 205-214.
- [12] SNODGRASS M F, KITANIDIS P K. A geostatistical approach to contaminant source identification[J]. Water Resources Research, 1997, 33(4): 537-546.
- [13] BUTERA I, TANDA M G. A geostatistical approach to recover the release history of groundwater pollutants[J]. Water Resources Research, 2003, 39(12): 1372.
- [14] MICHALAK A M, KITANIDIS P K. Application of geostatistical inverse modeling to contaminant source identification at Dover AFB, Delaware[J]. Journal of Hydraulic Research, 2004, 42: 9-18.
- [15] BUTERA I, TANDA M G, ZANINI A. Simultaneous identification of the pollutant release history and the source

- location in groundwater by means of a geostatistical approach[J]. *Stochastic Environmental Research and Risk Assessment*, 2013, 27(5): 1269–1280.
- [16] KITANIDIS P K. Quasi-linear geostatistical theory for inversing[J]. *Water Resources Research*, 1995, 31(10): 2411–2419.
- [17] SKAGGS T H, KABALA Z J. Recovering the release history of a groundwater contaminant[J]. *Water Resources Research*, 1994, 30(1): 71–79.
- [18] MIRGHANI B Y, MAHINTHAKUMAR K G, MICHAEL E T, et al. A parallel evolutionary strategy based simulation-optimization approach for solving groundwater source identification problems[J]. *Advances in Water Resources*, 2009, 32(9): 1373–1385.
- [19] FETTER C W. 污染水文地质学[M]. 周念清, 黄勇, 译. 北京: 高等教育出版社, 2011.
- [20] ZHENG Chun-miao, BENNETT G D. 地下水污染物迁移模拟[M]. 第2版. 孙晋玉, 卢国平, 译. 北京: 高等教育出版社, 2009.
- [21] 薛禹群, 谢春红. 地下水数值模拟[M]. 北京: 科学出版社, 2007.
- [22] 国家环境保护总局《水和废水监测分析方法》编委会. 水和废水监测分析方法[M]. 第4版. 北京: 中国环境科学出版社, 2002.

Application and sensitivity analysis of geostatistical approach to groundwater pollution source identification

LONG Yuqiao^{1, 2, 3}, CUI Tingting^{1, 2}, LI Wei^{1, 2}, Wu Chunyong^{1, 3}, LI Yange¹

(1. *Department of Hydrology and Water Resources, Nanjing Hydraulic Research Institute, Nanjing 210029, China;*

2. *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing 210098, China;*

3. *Nanjing R&D Tech Group Co., Ltd., Nanjing 210029, China*)

Abstract: To identifying the groundwater pollution source, a geostatistical approach is applied on the groundwater pollution identification (PSI) in a 1D homogeneous aquifer. Numerical test is employed to analyze the impact of observation point configuration on the PSI and the sensitivity analysis of pollution concentration observation error and model parameters. It is found out that if concentration observation data could precisely describe the real concentration distribution at the observed time point, a nice identification of the pollution discharge process could be obtained. If the calculated pollution discharge process is similar with the real discharge process, the dimension of the observation error is between 10^{-6} and $10^{-3.5}$, the dispersion coefficient varies between -10% and 5% , and the actual mean velocity varies between $\pm 2\%$. In the numerical case, the geostatistical approach shows a good performance for identifying the contaminant release process. Comparing to the observation error of concentration, the influence of dispersion coefficient and actual mean velocity are more outstanding in this case.

Keywords: geostatistical; groundwater; pollution; identification; sensitivity

(责任编辑: 韩 昆)