

DOI: 10.14042/j.cnki.32.1309.2017.06.015

# 地下水污染源识别的数学方法研究进展

王景瑞<sup>1,2</sup>, 胡立堂<sup>1,2</sup>

(1. 北京师范大学水科学研究院, 北京 100875; 2. 地下水污染控制与修复教育部工程研究中心, 北京 100875)

**摘要:** 地下水污染源识别模型可利用有限的观测资料估计污染源位置、污染物泄露强度及其泄露过程, 是制定地下水污染修复方案的依据。在阐明地下水污染源识别基本问题基础上, 综述了污染源识别研究的两大类数学方法, 一类为直接方法, 包括反向追踪法和基于正则化的方法; 另一类为间接方法, 包括基于优化的方法和基于概率统计的方法。同时指出了当前污染源识别数学方法应用中存在的主要问题: 地下水污染源识别问题的复杂性、地下水有机污染问题和模型求解效率的低下性。对土壤-地下水的联合管理、基于物联网的地下水污染监测、对非水相流体(Non-aqueous Phase Liquid, NAPL)类污染源识别以及基于图形处理器(GPU)的异构并行计算将是未来研究的重点方向。

**关键词:** 地下水污染源识别; 非适定性; 优化算法; 贝叶斯推理; 非水相流体

**中图分类号:** P641.2; X523; G353.11      **文献标志码:** A      **文章编号:** 1001-6791(2017)06-0943-10

地下水污染问题已成为制约国家经济发展和危及人类健康以及社会稳定的重要因素<sup>[1-2]</sup>。2011年, 中国环境保护部颁发的《全国地下水污染防治规划(2011—2020年)》, 旨在通过地下水污染状况调查, 对重点污染地区的地下水进行治理与修复。地下水污染源调查是开展地下水污染修复工作的前提, 但是通过观测获取污染物信息所需要的费用较高, 且观测的数据往往不足以确定污染源的位置以及污染羽的分布范围, 因此, 造成了地下水污染修复的困难<sup>[3]</sup>。地下水污染源识别可以根据已有的调查数据, 对污染源的位置以及污染物的泄露过程进行识别, 并以此来还原污染物在地下水中的迁移转化过程, 很大程度上弥补了上述不足。由于地下水污染具有多源性和复杂性, 地下水污染源识别的理论和方法是热点和难点, 本文将综述地下水污染源识别数学方法的研究进展, 并分析这些研究方法在实际应用中存在的主要问题, 为中国正在开展的地下水污染调查与修复工作提供一定参考。

## 1 地下水污染源识别的基本问题

污染物在地下水中的迁移理论建立在地下水动力学基础上, 早期主要用于海水入侵问题研究。20世纪70年代, 随着计算机技术的发展和大规模应用, 数值模型开始被用来解决大规模水流、溶质迁移问题, 而且研究重点逐渐转向人为污染问题。一般地, 控制污染物在地下水中运移的主要作用包括对流、机械弥散、分子扩散以及化学反应等。以三维的地下水溶质运移为例, 其数学表达式为

$$\frac{\partial C}{\partial t} = \nabla [D \cdot \nabla C] - \nabla [vC] - R(C) + E \quad (1)$$

式中:  $\nabla$  为梯度算子,  $[L^{-1}]$ ;  $C$  为溶质浓度,  $[M/L^3]$ ;  $t$  为时间,  $[T]$ ;  $D$  为弥散系数,  $[L^2/T]$ ;  $v$  为渗流

收稿日期: 2017-04-10; 网络出版时间: 2017-11-15

网络出版地址: <http://kns.cnki.net/kcms/detail/32.1309.P.20171115.1310.016.html>

基金项目: 国家自然科学基金资助项目(41572220); 北京市自然科学基金资助项目(J150002)

作者简介: 王景瑞(1990—), 男, 山东滨州人, 博士, 主要从事多相流数值模拟研究。E-mail: wangjr@mail.bnu.edu.cn

通信作者: 胡立堂, E-mail: litanghu@bnu.edu.cn

速度,  $[LT^{-1}]$ ;  $R$  为反应项,  $[ML^{-3}T^{-1}]$ ;  $E$  为源汇项,  $[ML^{-3}T^{-1}]$ 。当  $D$  或  $v$  为浓度  $C$  的函数时, 模型为非线性模型, 否则为线性模型。在给定初始条件以及边界条件基础上, 通过解析法或数值法等对式(1)进行求解, 得到污染物在地下水中的时空分布。

假定已通过水文地质调查等手段获取到相应的水文地质参数信息, 在污染源位置、污染物泄露强度以及污染物泄露的时间过程等信息已知的条件下, 求解污染物在地下水中的迁移转化过程称之为正问题; 而在已知污染羽在地下水中时空分布等条件下, 反求污染源的位置、污染物泄露强度以及污染物泄露的时间过程问题即为反问题, 地下水污染源识别即为反问题。Neuman<sup>[4]</sup>最早将反问题求解理论引入水文水资源领域, 并将反演方法分为直接法和间接法。对于反问题, 如果同时满足解是存在的、唯一的、稳定的3个条件, 称该问题是适定的, 但就大多数反问题而言, 往往是不适定的。毛德强<sup>[5]</sup>概括了地下水流反演模型唯一解的约束条件: ① 模型必须含有流量边界和源汇项; ② 模型中每个节点必须有3个时间点的观测数据, 其中一个为初始水头, 不同时间点的水头变化不可为0; ③ 对于变饱和含水层, 还需要具备包括非饱和的压力观测数据; ④ 渗透系数  $K(h)$  和孔隙度  $\theta(h)$  的表达式已知; ⑤ 在估计饱和含水量  $\theta_s$  时, 指定残余含水量  $\theta_r$ <sup>[5]</sup>。由此可见, 对于一个理想的模型, 其反问题的求解尚且如此复杂, 在实际应用中, 很难获得污染物在每个节点任意时间点的分布, 而且观测数据一般含有误差, 因此, 不适定性一直是反问题求解的难点。

地下水模型反问题一般分为5类, 即识别模型参数、边界条件、初始条件、源汇项以及同时识别上述4个条件中两个以上条件。从代数方程的角度分析, 求解方程中的源汇项  $E$  比求解方程中的系数项(例如弥散系数  $D$ )简单, 但是对于实际问题, 污染源存在的时间一般比较长, 且由于土地利用类型等变化, 污染物泄露过程通常非恒定。地下水污染物迁移模型多为非线性模型, 通过解析求解反问题的偏微分方程很难实现, 因而无法实现对污染源的很好识别。自反演理论被引入该领域以来, 对反问题的求解方法的研究发展迅速, 并被认为是未来环境模型发展过程中极具挑战性的内容<sup>[6]</sup>。

## 2 地下水污染源识别主要研究方法

地下水污染源识别方法最早源于热传导反问题的研究。热传导问题的控制方程与不考虑对流项的对流-弥散方程非常相似。在污染源识别研究之初, 热传导反问题的研究已开展了几十年, 因此, 许多热传导反问题的研究方法被直接借鉴用来解决地下水污染源识别问题<sup>[7]</sup>。地下水污染源识别的研究方法主要分为两类: 直接方法和间接方法。

### 2.1 直接方法

#### 2.1.1 反向追踪法

反向追踪法是指在已知当前地下水中污染物分布条件下, 沿地下水流向的反方向, 估求之前某个时间点的污染物分布, 从而确定污染源的位置和泄露强度。Bagtzoglou 等<sup>[8]</sup>提出了一种基于反向粒子追踪的方法, 计算过程中通过将速度场变为相反的方向实现逆向的目的, 并通过算子分裂法将对流项进行逆向求解, 但是对于弥散项保持不变。弥散项包含了污染物迁移过程中的重要信息, 但弥散是熵增过程, 其物理过程具有不可逆性, 因此, 该方法无法还原污染物泄露的真实过程。Skaggs 和 Kabala<sup>[9]</sup>采用了拟反向追踪的方法(Quasi-reversibility, QR), 该方法避开求解非适定性的对流弥散方程, 而是改求与对流弥散方程非常相似的适定性方程, 研究发现该方法具有计算效率高的特点, 但是其精度相对较差。根据逆梁方程法(Backward Beam Equation), Atmadja 和 Bagtzoglou<sup>[10]</sup>提出了 Marching-Jury 逆梁方程法, 用于求解反向追踪的一维对流弥散方程, 发现该方法的结果优于 QR 法<sup>[11]</sup>, 但该方法无法扩展到二维或三维模型。Neupauer 和 Wilson<sup>[12]</sup>提出了一种基于泛函分析理论的伴随状态方法, 并认为该方法具有求解效率高的特点, 曹彤彤等<sup>[13]</sup>通过二维均质、非均质模型验证了该方法在地下水污染源识别中的适用性。

#### 2.1.2 基于正则化的直接求解法

不适定性是反问题固有的本质, 理论上来说无法对反问题进行直接求解, 在认识到反问题不适定性产生

的原因之后,很多研究尝试通过正则化的方法进行求解,即给定相应的数学假设,将非适定问题转变为适定问题,从而可以对反问题直接求解,但是,正则化方法的应用是以牺牲解的精度为前提的。目前比较常用的方法包括奇异值分解法(Singular Value Decomposition, SVD)、截断奇异值分解法(Truncated Singular Value Decomposition, TSVD)以及吉洪诺夫正则化(Tikhonov Regularization, TR)。SVD用于解决线性反问题解的不唯一性问题,而TSVD用于解决线性反问题解的不稳定性问题。TR法不仅适用于解决线性反问题解的不稳定性问题,还可以解决非线性反问题解的不稳定问题。Skaggs和Kabala<sup>[14]</sup>在假定污染源位置已知的情况下,通过一维的均质稳定流模型,利用TR法反求了污染源释放的过程。Liu和Ball<sup>[15]</sup>将Skaggs的方法成功应用于Dover空军基地某低渗场地。

## 2.2 间接方法

### 2.2.1 基于优化的方法

最早被用于污染源识别的研究方法是基于优化算法的方法,其基本原理是借助运筹学中的优化算法,通过反复调用地下水污染物迁移模型,对模型参数进行优化,使得模型的输出与观测数据尽可能拟合。由于正问题一般是适定的,因此,该方法克服了直接求解反问题的不适定性。Gorelick等<sup>[16]</sup>最早将优化方法用于地下水污染源识别问题,在借助响应矩阵描述污染源对观测井处污染物浓度的影响基础上,利用线性规划以及最小二乘回归对模型参数进行优化,识别了一维稳定流模型潜在污染源位置和污染物泄露强度以及二维非稳定流模型潜在污染源位置、污染物泄露强度和泄露的时间过程。线性规划适用于线性问题的优化问题,但是在面对复杂的非线性问题时,很难求得最优解。

基于梯度的优化算法当前已被广泛用于非线性模型的优化问题。如果初值选择比较好,该类方法可以快速地完成优化过程得到最优解。Sonnenborg等<sup>[17]</sup>使用Levenberg-Marquardt算法识别了Denmark地区某污染场地的污染源强度。Sciortino等<sup>[18]</sup>认为基于Levenberg-Marquardt法建立的反演模型对弥散系数非常敏感。Mahar和Datta<sup>[19]</sup>通过将地下水模型嵌入包含多种优化算法的优化模型,提出了一种嵌入式的反演模型(Embedded Simulation-Optimization),可以同时完成监测网设计、污染源识别以及含水层等参数估计,该方法区别于外部调用方法,可以自动完成反问题的多目标优化过程。

当非线性模型同时为非凸模型时,如果待估参数的初值选择不当,基于梯度的优化算法非常容易陷入局部极小值,而无法求得最小值,因此,被认为是局部优化算法。启发式算法是相对于局部优化算法提出的,不依赖于求解目标函数的梯度,可以解决非凸模型的优化问题,在污染源识别领域应用也非常广泛。当前应用于地下水污染源识别的启发式算法主要包括遗传算法、模拟退火法、人工神经网络以及差分进化算法等。Aral和Guan<sup>[20]</sup>提出了改进的遗传算法(Improved Genetic Algorithm)以及逐步遗传算法(Progressive Genetic Algorithm)<sup>[21]</sup>,分别用于提高计算结果的精度和计算过程的效率。Jha和Datta<sup>[22]</sup>提出了一种自适应模拟退火法(Adaptive Simulated Annealing),并与遗传算法的计算结果进行对比,发现当待估参数初值取值较好时,自适应模拟退火法收敛更快,但当观测数据存在较大误差时,该方法也出现了计算效率低的问题。人工神经网络(Artificial Neural Network, ANN)用于污染源识别,实质是建立污染物迁移模型的输入-输出函数关系模型,在此基础上寻求最优解,对函数方程进行计算求解的效率显然要高于迭代求解偏微分方程(组)的数值法<sup>[23-24]</sup>。Gurarslan和Karahan<sup>[25]</sup>通过设置不同程度的观测数据误差,首次将微分进化算法用于地下水污染源识别。

相对于局部优化算法,启发式算法对污染物迁移模型进行调用次数更多,因此,计算效率一般比较低。混合法将两种或两种以上的优化算法联合使用,充分利用不同优化算法各自的优点。根据各优化算法在优化过程中的使用顺序,混合法可分为序列式混合算法以及嵌入式混合算法<sup>[26]</sup>。Yeh等<sup>[27]</sup>提出了一种基于模拟退火法和禁忌搜索法的序列式混合算法,并识别了理想模型中污染源的位置以及污染物释放历史过程,其中禁忌搜索法用于确定污染源的位置,模拟退火法用于识别污染物的释放历史过程,并证明该方法具有较强的鲁棒性。Mahinthakumar和Sayeed<sup>[28]</sup>则提出了一种基于遗传算法与局部优化算法的序列式混合算法,不同于Yeh等<sup>[27]</sup>的方法,其首先使用遗传算法快速获得一个接近全局最优解的结果,然后使用局部优化算法对结

果进行微调,既克服了遗传算法求解速度慢的缺点,也克服了局部优化算法难以求得全局最优解的缺点。Newman 等<sup>[29-30]</sup>将遗传算法与最小相对熵算法结合,该方法的计算过程与 Mahinthakumar 的方法类似,不同之处在于并非将未知变量作为确定的值,而是服从某种概率分布的不确定值,最终不仅可以反演得到未知变量的结果,还兼具了不确定性分析功能。Mirghani 通过将 Hooke-Jeeves' Search 局部优化算法嵌入遗传算法,并通过理想案例证明该方法在污染源位置已知的条件下,可以高效识别污染物泄露过程<sup>[30]</sup>。特别地,嵌入式方法优于序列式方法在于嵌入式算法可以进行并行化处理。

基于优化算法的反问题求解效率比较低的主要原因是污染物迁移模型计算速度较慢,并行计算<sup>[31]</sup>和替代模型<sup>[32]</sup>即通过对污染物迁移模型进行改变来提高其计算效率。Mirghani 等<sup>[33]</sup>结合遗传算法与并行化的污染物迁移模型,用于均质、非均质模型的污染源识别,结果认为该方法可以将计算效率提高 100 倍以上。替代模型实质是通过支持向量机、神经网络等方法建立污染物迁移模型输入-输出之间的函数关系,通过利用已有的数据或实验设计生成的数据对替代模型进行训练,使其在一定程度上具有污染物迁移模型等效的功能,因此,替代模型是没有物理意义的数据驱动模型,用其取代地下水污染物迁移模型进行污染源识别,能够大幅度减少基于空间网格剖分与时间离散的地下水数值模型所带来的计算负荷,从而提高污染源识别的效率<sup>[34-35]</sup>。

### 2.2.2 基于概率统计的方法

最先被应用于污染源识别问题的基于概率统计的方法包括多元非线性回归以及与其相关的极大似然法<sup>[36]</sup>等。根据大数定律,一般假定观测数据误差服从均值为零的正态分布,由极大似然法建立观测数据与模型输出之间的目标函数,从而将问题转化为最小二乘问题,但是最小二乘法对奇异值非常敏感,Sun 等<sup>[37]</sup>提出了一种稳健的最小二乘法(Robust Least Squares, RLS),并用于污染源识别,结果证明该方法可以降低污染源识别结果对观测数据误差的敏感性,当模型是非适定的或者观测数据误差较大时,该方法展示出了较强的稳健性。回归法本身可以用于解决简单的污染源识别问题,但是对于复杂的非线性模型,回归法往往与优化算法结合使用。

贝叶斯理论是数理统计的一个分支,被广泛用于反问题求解,与经典的统计理论中将分布参数  $\theta$  作为确定性的未知常量不同,贝叶斯理论认为由于观测数据包含不确定性的信息(例如观测误差),将  $\theta$  看作随机变量,并以概率密度函数表示,因此,该方法求解反问题,解是存在的且唯一的。在获得样本之前,对  $\theta$  的认识称为先验信息,采用先验分布描述(先验概率密度函数);获得样本后,样本中又包含了  $\theta$  新的信息,采用似然函数描述;通过贝叶斯公式将具有主观性的先验分布与具有客观性的似然函数结合,得到  $\theta$  的后验分布(后验概率密度函数),显然,后验分布中包含了  $\theta$  的先验和样本信息,由后验分布对  $\theta$  进行统计或推断更合理;而且,获得新样本后,原来的后验分布又变成先验分布,与新样本综合又得到新的后验分布,从而不断更新对  $\theta$  的认识,这也是贝叶斯理论的精髓所在<sup>[38]</sup>。

当待估参数的先验概率密度函数以及极大似然函数的形式比较简单时,最终得到的后验概率密度函数形式也相对简单,比如高斯分布或类高斯分布,因此,可以通过点估计的方法直接求解得到最优估计值,特别是当污染物迁移模型为线性模型时,通过后验概率密度函数估计最优值可以转化为线性回归问题。此外,基于贝叶斯推理得到参数的概率密度函数,可以直接对概率密度函数进行区间估计,从而分析参数的不确定性。Snodgrass 和 Kitanidis<sup>[39]</sup>最先将贝叶斯推理的方法用于地下水污染源识别问题,基于一维的稳定流模型,在只考虑一个位置已知污染源情况下,识别了污染物的泄露过程,但是该方法无法保证求解过程中污染物浓度值是非负的。Butera 和 Tanda<sup>[40]</sup>将该方法应用于有多个污染源存在的二维模型,并考虑了观测数据含有不同程度误差情况下对结果的影响。Michalak 和 Kitanidis<sup>[41]</sup>提出了一种基于反射布朗运动(Reflected Brownian Motion)的严格统计方法,保证了贝叶斯推理过程中污染物浓度为非负值,还提出了一种基于地质统计的伴随状态方法,认为当待估参数维度较高时,该方法可以很大程度上提高计算效率<sup>[42]</sup>。Sun<sup>[43]</sup>提出一种鲁棒地质统计法用于解决数学模型为线性模型的污染源识别问题,认为该方法相对于传统的地质统计方法,在面对模型与参数的不确定性问题时,具有更强的鲁棒性。最小相对熵法(Minimum Relative Entropy)是一种基于

贝叶斯推理方法,可以充分利用先验信息进行反问题的求解。Woodbury 和 Ulrych<sup>[44]</sup>通过只考虑对流和弥散的一维污染物迁移模型,在给定污染源释放强度的取值范围等先验信息条件下,得到了污染物释放强度的后验概率密度函数,然后利用拉格朗日乘数法得到其最优估计值。

基于贝叶斯推理的卡尔曼滤波技术是一种反问题的求解方法,主要包括传统的卡尔曼滤波、扩展的卡尔曼滤波以及用于非线性问题的集合卡尔曼滤波。Dokou 和 Pinder<sup>[45]</sup>结合蒙特卡罗抽样法与卡尔曼滤波技术,识别了给定污染源位置的污染物泄露强度。江思珉等<sup>[46]</sup>提出一种基于卡尔曼滤波技术和模糊集合理论的地下水污染羽识别方法。崔尚进等<sup>[47]</sup>提出一种基于 U-D 分解的卡尔曼滤波法,用以解决卡尔曼滤波方程中协方差矩阵可能出现的不对称或非正定的情况,并通过二维稳定流模型验证了其可靠性。Huang 等<sup>[48]</sup>采用了基于集合卡尔曼滤波技术的数据同化方法,识别二维稳定流模型污染物泄露强度。卡尔曼滤波技术假定先验分布为多维高斯分布,因此,在解决非高斯问题时精度不够好。

实际问题中,先验分布的形式很多,且由于数值模型的参数与模型输出多为非线性关系,使得似然函数解析式非常复杂,因此,即使服从高斯分布的变量在经过非线性系统转换后也会呈现非高斯分布特征,最终得到的后验概率密度函数也非常复杂(例如非凸函数),导致反问题仍然是非适定的且无法通过点估计求解最优值。Xu 和 Gómez-Hernández<sup>[49]</sup>采用了一种改进的集合卡尔曼滤波法用于污染源识别研究,可用于待估参数服从非高斯分布的问题,同时也意识到该方法当前无法应用于实际问题。基于蒙特卡罗随机采样的马尔科夫链蒙特卡罗法(MCMC),根据抽样便可统计出参数的后验分布,避开了求解似然函数的解析式,已成为贝叶斯计算最流行的方法<sup>[50]</sup>,Metropolis-Hastings 算法和 Gibbs 算法是最著名的两种 MCMC 方法,被广泛应用于基于理想模型的污染源识别研究<sup>[51-52]</sup>。

基于 MCMC 的抽样方法往往需要充分的样本(通常为几千个)才能获得相对准确的统计结果。虽然当前有大量研究通过优化 MCMC 取样方法来提高其效率<sup>[53-54]</sup>,但与基于优化算法的地下水污染源识别问题类似,其主要原因在于物理模型计算速度较慢,因此,在应用过程中也存在着求解效率较低的问题。替代模型作为解决该方法的方法,同样适用于基于贝叶斯推理的地下水污染源识别,提高计算效率的同时,也得到了精度较高的结果<sup>[55-56]</sup>。

### 3 地下水污染源识别方法应用中的主要问题

#### 3.1 地下水污染源识别问题的复杂性

地下水污染源识别问题的复杂性包括两个方面:污染物泄露过程的复杂性和由此造成的观测数据的难以获取性。由于地下水特殊的赋存条件,其污染过程具有很强的隐蔽性,长时间来难以被人类察觉,在此期间,土地利用方式变更等使得污染物的种类和泄露强度也随之发生变化,利用有限且时间上不连续的观测数据,对泄露过程比较复杂的单一组分甚至多组分的污染源进行识别将变得非常棘手。宋昕等<sup>[57]</sup>认为由于中国工业企业早期管理不具系统性,大多数污染场地的运行记录没有备案,使得场地污染源描述、污染羽的迁移转化规律分析非常困难。对北京平原区地下水主要污染源进行解析的结果显示,工业企业生产过程中的废水排放对可能进入地下水的污染物质影响很大,但涉及资料难以获取,此外,单一污染源释放的污染物往往是多种的<sup>[58]</sup>。杨悦锁等<sup>[59]</sup>综述了土壤和地下水中重金属和多环芳烃复合污染的研究进展,认为中国当前复合污染已经非常普遍,且由于其复杂的物理化学过程,为污染治理修复带来很大困难。

理想模型的研究一般假定地下水模型完全准确,但实际上地下水模型存在大量的不确定性,其来源包括了水文地质参数等驱动数据的不确定性、观测资料的不确定性以及模型结构等的不确定性<sup>[60]</sup>。特别地,弥散度具有明显的尺度效应,这些不确定性的存在将对污染源识别形成严重干扰。此外,地下水污染源识别依赖于先验信息,以贝叶斯推理为例,在缺乏先验信息的情况下,一般假定待估参数在某一区间内服从线性分布,这一主观假设与地下水模型本身的不确定性有可能使得污染源识别得出错误的结论,从而影响决策者制定修复方案。

### 3.2 地下水有机污染问题

由于人类早期的地下水污染问题主要针对溶解性污染物(海水入侵、农业灌溉等),因此,污染源识别研究大多建立在对流弥散方程基础之上。近年来,非水相流体(Non-aqueous Phase Liquid, NAPL)类污染物被广泛关注,它与溶解性污染物相比,具有不同的性质和特征,作为一种独立的相态存在,多相流理论的发展为研究其在多孔介质中的运移提供了重要工具<sup>[61-62]</sup>。当前对京津冀地区的地下水中有有机污染物的初步调查表明,该地区地下水中有有机污染物种类高达133种,长期滞留于地下水中的有机污染以及经缓慢降解后的中间产物,一般毒性很强。但单纯地通过取样检测其在水中的浓度等方法很难用于识别该类污染源的信息,特别对于重质非水相流体(DNAPL),一旦进入潜水含水层后便持续向下运移,累积于潜水含水层底板,难以对其量化。由于含水层介质的非均质性以及不同种类NAPL类污染物在地下水中迁移转化和归趋等的差异,也使得对地下水有机污染问题的研究成为世界性难题。Kashuk等<sup>[63]</sup>认为在获取到部分观测资料基础上,仅通过构建数学模型或数值模型等对有机污染物的迁移转化过程进行研究,无法对模型进行有效识别,从而无法保证其高仿真,并提出了一种基于色彩空间表示有机污染物浓度的非侵入式地下水有机污染物监测方法。其他用于可视化有机污染物在土壤、地下水中等运移的方法可参考文献[63-65]。

### 3.3 求解方法计算效率低

随着计算机技术的发展,对地下水污染源识别方法的研究发展迅速,但不难发现,上述研究多以理想模型为主,对于场地尺度的实际问题,如何在保证模型精度的前提下,高效地完成计算求解过程仍然是无法回避的问题。基于优化算法的方法需要反复调用正演模型,该过程将非常耗时,而且,随着待估参数维度增加,其效率下降非常快;基于概率统计的方法,特别是基于贝叶斯推理的方法,虽然MCMC方法在估计后验分布时,避免了直接求解积分的困难,但将其应用于实际问题时是否有足够的计算效率,能否满足实际要求,仍待进一步研究。替代模型作为没有物理意义的数据驱动模型,虽然求解效率非常高,但其训练过程仍需要通过实验设计(Design of Experimental, DOE),反复调用物理模型进行取样,因此,替代模型的训练过程非常耗时。直接求解法虽然不存在计算效率的问题,然而,由于其自身的局限性,多用于求解理想模型的污染源识别。因此,当前对于污染源识别问题的研究大多停留在理论层面,可供实际应用的成果相对较少。

随着以贝叶斯推理为核心的人工智能技术的发展,基于贝叶斯推理的方法用于地下水污染源识别已展现出良好的发展前景。从数学角度分析,基于优化的方法求解反问题更加严格,但是考虑到当前观测数据的匮乏以及模型、数据等不确定性的存在,通过贝叶斯推理所建立的学习系统,从统计的角度对该问题进行求解,以数据同化的方式反复地将新的观测数据融入模型,提高模型的仿真性,因此,可以更加充分地利用已有的观测数据,同时兼顾考虑数据等的不确定性。特别对于地下水有机污染问题,当前所采用的研究方法大多以该方法为主<sup>[66]</sup>。以大数据和云计算为背景发展起来的人工智能,同时也为环境、水文相关模型反问题的研究提供了很好的发展方向,随着未来监测技术的提高,大量的观测数据将很大程度上提高物理模型的仿真性<sup>[67-68]</sup>,因此,如何提高物理模型的计算效率将成为反问题求解的主要障碍。

## 4 结论与展望

(1) 从污染物的传导过程来看,地下水污染与土壤污染关系非常密切,在一定程度上可以归为同一问题,地下水主要存贮于地下含水层中,与非饱和带直接连通,二者之间发生着频繁的物质交换,地下水污染经常是由土壤污染转移造成的。因此,与当前地下水污染源识别只将饱和带作为研究对象相比,未来将饱和带和非饱和带进行统筹考虑,其物理意义更具合理性。

(2) 大数据背景下,数据的价值越来越被重视,而地下水污染源识别作为污染物迁移问题的反问题,所需的驱动数据更多,而且数据的获取难度更大。近年来,物联网技术的发展将使得地下水污染监测数据的采集、管理和实时分析成为可能。因此,全国范围内布设地下水监测网将成为中国地下水污染防治规划的重要数据来源,是开展研究工作的基础,利用这些数据进行污染源识别的同时,可以优化监测网的设计,而新

增加的观测井获取到的观测数据又可以进一步用于污染源识别。

(3) 将地球化学方法(比如同位素实验和污染物迁移可视化实验)与数学模型相结合的方法将是识别地下水有机污染物(特别是非水相流体)比较可行的方法,借助室内、室外实验等获取足够数据的基础上,利用数学模型等方法进行污染源识别,可以保证模型的高仿真性和求解的精度。

(4) 基于 GPU 的异构并行计算已在深度学习领域展现了较好的性能,因此,基于 GPU 的并行方法对物理模型进行并行化以及对构建替代模型的过程进行并行化将很大程度上提高地下水污染源识别这类反问题的求解效率,也是将当前的理论研究应用于实际的关键。

#### 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. Twofold significance of groundwater pollution prevention in China's water pollution control [J]. Acta Scientiae Circumstantiae, 2009, 29(3): 474-481. (in Chinese))
- [2] 刘庄,晁建颖,张丽,等.中国非点源污染负荷计算研究现状与存在问题[J].水科学进展,2015,26(3):432-442. (LIU Z, ZHAO J Y, ZHANG L, et al. Current status and problems of non-point source pollution load calculation in China [J]. Advances in Water Science, 2015, 26(3): 432-442. (in Chinese))
- [3] 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.
- [4] NEUMAN S P. Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty [J]. Water Resources Research, 1973, 9(4): 1006-1021.
- [5] 毛德强.地下水反演模型解的唯一性和高精度反演方法研究[D].北京:中国地质大学,2013. (MAO D Q. A study of uniqueness issue and high resolution methods in groundwater inverse modeling [D]. Beijing: China University of Geosciences, 2013. (in Chinese))
- [6] POETERE P, HILL M C. Inverse models: a necessary next step in ground-water modeling [J]. Ground Water, 1997, 35(2): 250-260.
- [7] ATMADJA J, BAGTZOGLOW A. State of the art report on mathematical methods for groundwater pollution source identification [J]. Environmental Forensics, 2001, 2(3): 205-214.
- [8] BAGTZOGLOU A C, DOUGHERTY D E, TOMPSON A F B. Application of particle methods to reliable identification of groundwater pollution sources [J]. Water Resources Management, 1992, 6(1): 15-32.
- [9] SKAGGS T H, KABALA Z J. Recovering the history of a groundwater contaminant plume-method of quasi-reversibility [J]. Water Resources Research, 1995, 31(11): 2669-2673.
- [10] ATMADJA J, BAGTZOGLOU A C. Pollution source identification in heterogeneous porous media [J]. Water Resources Research, 2001, 37(8): 2113-2125.
- [11] BAGTZOGLOU A C, ATMADJA J. Marching-jury backward beam equation and quasi-reversibility methods for hydrologic inversion: application to contaminant plume spatial distribution recovery [J]. Water Resources Research, 2003, 39(2): 1038.
- [12] NEUPAUER R M, WILSON J L. Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant [J]. Water Resources Research, 1999, 35(11): 3389-3398.
- [13] 曹彤彤,曾献奎,吴吉春,等.基于伴随状态方法的地下水污染源识别研究[J].高校地质学报,2016,22(3):563-571. (CAO T T, ZENG X K, WU J C, et al. Identification of groundwater contaminant source based on adjoint-state method [J]. Geological Journal of China Universities, 2016, 22(3): 563-571. (in Chinese))
- [14] SKAGGS T H, KABALA Z J. Recovering the release history of a groundwater contaminant [J]. Water Resources Research, 1994, 30(1): 71-79.
- [15] LIU C, BALL W P. Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware [J]. Water Resources Research, 1999, 35(7): 1975-1985.
- [16] GORELICK S M, EVANS B, REMSON I. Identifying sources of groundwater pollution-an optimization approach [J]. Water Resources Research, 1983, 19(3): 779-790.

- [17] SONNENBORG T O, ENGESGAARD P, ROSBJERG D. Contaminant transport at a waste residue deposit: 1: inverse flow and nonreactive transport modeling [J]. *Water Resources Research*, 1996, 32(4): 925-938.
- [18] SCIORTINO A, HARMON T C, YEH W. Inverse modeling for locating dense non-aqueous pools in groundwater under steady flow conditions [J]. *Water Resources Research*, 2000, 36(7): 1723-1735.
- [19] MAHAR P S, DATTA B. Optimal identification of ground-water pollution sources and parameter estimation [J]. *Journal of Water Resources Planning and Management-ASCE*, 2001, 127(1): 20-29.
- [20] ARAL M M, GUAN J. Genetic algorithms in search of groundwater pollution sources [M]//ARAL M M. *Advances in Groundwater Pollution Control and Remediation*. Dordrecht: Kluwer Academic Publishers, 1995: 347-369.
- [21] ARAL M M, GUAN J B, MASLIA M L. Identification of contaminant source location and release history in aquifers [J]. *Journal of Hydrologic Engineering*, 2001, 6(3): 225-234.
- [22] JHA M, DATTA B. Three-dimensional groundwater contamination source identification using adaptive simulated annealing [J]. *Journal of Hydrologic Engineering*, 2013, 18(3): 307-317.
- [23] SINGH R M, DATTA B, JAIN A. Identification of unknown groundwater pollution sources using artificial neural networks [J]. *Journal of Water Resources Planning and Management-ASCE*, 2004, 130(6): 506-514.
- [24] BORAH T, BHATTACHARJYA R K. Development of an improved pollution source identification model using numerical and ANN based simulation-optimization model [J]. *Water Resources Management*, 2016, 30: 1-14.
- [25] GURARSLAN G, KARAHAN H. Solving inverse problems of groundwater-pollution-source identification using a differential evolution algorithm [J]. *Hydrogeology Journal*, 2015, 23(6): 1109-1119.
- [26] MIRGHANI B Y A. *Evolutionary Algorithms-based parallel simulation-optimization framework for solving inverse problems* [D]. Raleigh: North Carolina State University Civil Engineering Department, 2007.
- [27] YEH H, CHANG T, LIN Y. Groundwater contaminant source identification by a hybrid heuristic approach [J]. *Water Resources Research*, 2007, 43(9): W09420.
- [28] MAHINTHAKUMAR G K, SAYEED M. Hybrid genetic algorithm-local search methods for solving groundwater source identification inverse problems [J]. *Journal of Water Resources Planning and Management-ASCE*, 2005, 131(1): 45-57.
- [29] NEWMAN M A, HATFIELD K, HAYWORTH J, et al. Inverse characterization of NAPL source zones [J]. *Environmental Science & Technology*, 2006, 40(19): 6044-6050.
- [30] NEWMAN M, HATFIELD K, HAYWORTH J, et al. A hybrid method for inverse characterization of subsurface contaminant flux [J]. *Journal of Contaminant Hydrology*, 2005, 81(1/2/3/4): 34-62.
- [31] COMMER M, KOWALSKY M B, DOETSCH J, et al. MPiTOUGH2: a parallel parameter estimation framework for hydrological and hydrogeophysical applications [J]. *Computers & Geosciences*, 2014, 65: 127-135.
- [32] RAZAVI S, TOLSON B A, BURN D H. Review of surrogate modeling in water resources [J]. *Water Resources Research*, 2012, 48(7): W07401.
- [33] MIRGHANI B Y, MAHINTHAKUMAR K G, TRYBY M E, 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.
- [34] BASHI-AZGHADI S N, KERACHIAN R, BAZARGAN-LARI M R, et al. Characterizing an unknown pollution source in groundwater resources systems using PSVM and PNN [J]. *Expert Systems with Applications*, 2010, 37(10): 7154-7161.
- [35] 王宇, 卢文喜, 卞建民, 等. 基于小波神经网络的地下水流数值模拟模型的替代模型研究 [J]. *中国环境科学*, 2015, 35(1): 139-146. (WANG Y, LU W X, BIAN J M, et al. Surrogate model of numerical simulation model of groundwater based on Wavelet Neural Network [J]. *China Environmental Science*, 2015, 35(1): 139-146. (in Chinese))
- [36] WAGNER B J. Simulations parameter-estimation and contaminant source characterization for coupled groundwater-flow and contaminant transport modeling [J]. *Journal of Hydrology*, 1992, 135(1/2/3/4): 275-303.
- [37] SUN A Y, PAINTER S L, WITTMAYER G W. A constrained robust least squares approach for contaminant release history identification [J]. *Water Resources Research*, 2006, 42(4): W04414.
- [38] 梁忠民, 戴荣, 李彬权. 基于贝叶斯理论的水文不确定性分析研究进展 [J]. *水科学进展*, 2010, 21(2): 274-281. (LIANG Z M, DAI R, LI B Q. A review of hydrological uncertainty analysis based on Bayesian theory [J]. *Advances in Water Science*, 2010, 21(2): 274-281. (in Chinese))



- [39] SNODGRASS M F, KITANIDIS P K. A geostatistical approach to contaminant source identification [J]. *Water Resources Research*, 1997, 33(4): 537-546.
- [40] BUTERA I, TANDA M G. A geostatistical approach to recover the release history of groundwater pollutants [J]. *Water Resources Research*, 2003, 39(12): 1372.
- [41] MICHALAK A M, KITANIDIS P K. A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification [J]. *Water Resources Research*, 2003, 39(2): 1033.
- [42] MICHALAK A M, KITANIDIS P K. Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling [J]. *Water Resources Research*, 2004, 40(8): W08302.
- [43] SUN A Y. A robust geostatistical approach to contaminant source identification [J]. *Water Resources Research*, 2007, 43(2): W02418.
- [44] WOODBURY A D, ULRYCH T J. Minimum relative entropy and probabilistic inversion in groundwater hydrology [J]. *Stochastic Hydrology and Hydraulics*, 1998, 12(5): 317-358.
- [45] DOKOU Z, PINDER G F. Optimal search strategy for the definition of a DNAPL source [J]. *Journal of Hydrology*, 2009, 376(3/4): 542-556.
- [46] 江思珉, 张亚力, 周念清, 等. 基于卡尔曼滤波和模糊集的地下水污染羽识别[J]. *同济大学学报(自然科学版)*, 2014(3): 435-440. (JIANG S M, ZHANG Y L, ZHOU N Q, et al. Groundwater contaminant plume identification by using Kalman filter technique and fuzzy set theory [J]. *Journal of Tongji University (Natural Science)*, 2014(3): 435-440. (in Chinese))
- [47] 崔尚进, 卢文喜, 顾文龙, 等. 基于 U-D 分解的卡尔曼滤波法在地下水污染源识别中的应用[J]. *中国环境科学*, 2016, 36(9): 2633-2637. (CUI S J, LU W X, GU W L, et al. Application of U-D factorization-based Kalman filter to identify the groundwater pollution source [J]. *China Environmental Science*, 2016, 36(9): 2633-2637. (in Chinese))
- [48] HUANG C, HU B X, LI X, et al. Using data assimilation method to calibrate a heterogeneous conductivity field and improve solute transport prediction with an unknown contamination source [J]. *Stochastic Environmental Research and Risk Assessment*, 2009, 23(8): 1155-1167.
- [49] XU T, GÓMEZ-HERNÁNDEZ J J. Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering [J]. *Water Resources Research*, 2016, 52(8): 6587-6595.
- [50] YUSTRES Á, ASENSIO L, ALONSO J, et al. A review of Markov Chain Monte Carlo and information theory tools for inverse problems in subsurface flow [J]. *Computational Geosciences*, 2012, 16(1): 1-20.
- [51] HAZART A, GIOVANNELLI J, DUBOST S, et al. Inverse transport problem of estimating point-like source using a Bayesian parametric method with MCMC [J]. *Signal Processing*, 2014, 96(5): 346-361.
- [52] 顾文龙, 卢文喜, 张宇, 等. 基于贝叶斯推理与改进的 MCMC 方法反演地下水污染源释放历史 [J]. *水利学报*, 2016, 47(6): 772-779. (GU W L, LU W X, ZHANG Y, et al. Reconstructing the release history of groundwater contamination sources based on the Bayesian inference and improved MCMC method [J]. *Journal of Hydraulic Engineering*, 2016, 47(6): 772-779. (in Chinese))
- [53] SMITH T J, MARSHALL L A. Bayesian methods in hydrologic modeling: a study of recent advancements in Markov Chain Monte Carlo techniques [J]. *Water Resources Research*, 2008, 44(12): W00B05.
- [54] ZHANG J J, LI W X, WU L S, et al. An adaptive Gaussian process-based method for efficient Bayesian experimental design in groundwater contaminant source identification problems [J]. *Water Resources Research*, 2016, 52(8): 5971-5984.
- [55] ZENG L, SHI L, ZHANG D, et al. A sparse grid based Bayesian method for contaminant source identification [J]. *Advances in Water Resources*, 2012, 37(1): 1-9.
- [56] ZHANG J, ZENG L, CHEN C, et al. Efficient Bayesian experimental design for contaminant source identification [J]. *Water Resources Research*, 2015, 51(1): 576-598.
- [57] 宋昕, 林娜, 殷鹏华. 中国污染场地修复现状及产业前景分析 [J]. *土壤*, 2015, 47(1): 1-7. (SONG X, LIN N, YIN P H. Contaminated site remediation industry in China: current state and future trends [J]. *Soils*, 2015, 47(1): 1-7. (in Chinese))
- [58] 陆燕, 何江涛, 王俊杰, 等. 北京平原区地下水污染源识别与危害性分级 [J]. *环境科学*, 2012, 33(5): 1526-1531. (LU Y, HE J T, WANG J J, et al. Groundwater pollution sources identification and grading in Beijing Plain [J]. *Environmental Science*, 2012, 33(5): 1526-1531. (in Chinese))

- [59] 杨悦锁, 陈煜, 李盼盼, 等. 土壤和地下水中重金属和多环芳烃复合污染和修复研究进展[J]. 化工学报, 2017, 68(6): 2219-2232. (YANG Y S, CHEN Y, LI P P, et al. Research progress on co-contamination and remediation of Heavy Metals (HMs) and Polycyclic Aromatic Hydrocarbons (PAHs) in soil and groundwater [J]. CIESC Journal, 2017, 68(6): 2219-2232. (in Chinese))
- [60] SINGH A, MISHRA S, RUSKAUFF G. Model averaging techniques for quantifying conceptual model uncertainty[J]. Ground Water, 2010, 48(5): 701-715.
- [61] CHRYSIKOPOULOS C V, LEE K Y. Contaminant transport resulting from multi-component non-aqueous phase liquid pool dissolution in three-dimensional subsurface formations [J]. Journal of Contaminant Hydrology, 1998, 31(1/2): 1-21.
- [62] 施小清, 姜蓓蕾, 吴吉春, 等. 非均质介质中重非水相污染物运移受泄漏速率影响数值分析[J]. 水科学进展, 2012, 23(3): 376-382. (SHI X Q, JIANG B L, WU J C, et al. Numerical analysis of the effect of leakage rate on dense non-aqueous phase liquid transport in heterogenous porous media [J]. Advances in Water Science, 2012, 23(3): 376-382. (in Chinese))
- [63] KASHUK S, MERCURIO S R, ISKANDER M. Visualization of dyed NAPL concentration in transparent porous media using color space components [J]. Journal of Contaminant Hydrology, 2014, 162/163(5): 1-16.
- [64] PPWER C, GERHARD J I, KARAOULIS M, et al. Evaluating four-dimensional time-lapse electrical resistivity tomography for monitoring DNAPL source zone remediation[J]. Journal of Contaminant Hydrology, 2014, 162/163: 27-46.
- [65] POWER C, GERHARD J I, TSOURLLOS P, et al. Improved time-lapse electrical resistivity tomography monitoring of dense non-aqueous phase liquids with surface-to-horizontal borehole arrays[J]. Journal of Applied Geophysics, 2015, 112: 1-13.
- [66] KOCH J, NOWAK W. Identification of contaminant source architectures-A statistical inversion that emulates multiphase physics in a computationally practicable manner[J]. Water Resources Research, 2016, 52(2): 1009-1025.
- [67] LETTENMAIER D P. Observational breakthroughs lead the way to improved hydrological predictions[J]. Water Resources Research, 2017, 53(4): 2591-2597.
- [68] OJHA R, RAMADAS M, GOVINDARAJU R S. Current and future challenges in groundwater: I: modeling and management of resources [J]. Journal of Hydrologic Engineering, 2015, 20(1): A40140071.

## Advances in mathematical methods of groundwater pollution source identification \*

WANG Jingrui<sup>1,2</sup>, HU Litang<sup>1,2</sup>

(1. College of Water Sciences, Beijing Normal University, Beijing 100875, China;

2. Engineering Research Center of Groundwater Pollution Control and Remediation of Ministry of Education, Beijing 100875, China)

**Abstract:** Using limited observation data of groundwater quality, models of groundwater pollution source identification can be used to estimate the locations, leakage rate, and the dominant processes of the pollution sources, which thus can provide a reference for formulating remedial schemes for groundwater pollution. Based on the principles and theories of pollutant movement and source identification, this paper presents an overview of existing mathematical methods, including direct methods (inverse particle tracing methods and regularization-based methods) and indirect methods (optimization-based and probability-based methods). The main problems in the application of these methods are ① the complexity of groundwater pollution source identification, ② groundwater organic contamination, and ③ the low efficiency of model calculation. Integrated research of soil-groundwater systems, instrumentation-based groundwater pollution monitoring, identification of NAPL pollutants, and GPU-based heterogeneous parallel computing will be the keys to groundwater pollution source identification in the future.

**Key words:** groundwater pollution source identification; Ill-posedness; optimization; Bayesian inference; non-aqueous phase liquid

\* The study is financially supported by the National Natural Science Foundation of China (No. 41572220).