

## 地下水污染强度及渗透系数场的反演识别研究

吴延浩, 江思珉, 吴自军

### Identification of groundwater pollution intensity and hydraulic conductivity field

WU Yanhao, JIANG Simin, and WU Zijun

在线阅读 View online: <https://doi.org/10.16030/j.cnki.issn.1000-3665.202208042>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 考虑预报偏差的迭代式集合卡尔曼滤波在地下水水流数据同化中的应用

Application of the bias aware Ensemble Kalman Filter with Confirming Option (Bias-CEnKF) in groundwater flow data assimilation

杨运, 吴吉春, 骆乾坤, 钱家忠 水文地质工程地质. 2022, 49(6): 13-23

#### 地下水污染风险预警等级及阈值确定方法研究综述

Review on the determination methods for early warning grade and threshold of groundwater pollution risk

王嘉瑜, 蒲生彦, 胡, 李博文 水文地质工程地质. 2020, 47(2): 43-50

#### 基于多源数据的弱透水层水文地质参数反演研究

Research on hydrogeological parameter inversion of an aquitard based on multi-source data: A case study of a silt layer in the Hohhot Basin

石鸿蕾, 郝奇琛, 邵景力, 崔亚莉, 张秋兰 水文地质工程地质. 2021, 48(2): 1-7

#### 基于多点地质统计与集合平滑数据同化方法识别非高斯渗透系数场

Characterization of non-Gaussian hydraulic conductivity fields using multiple-point geostatistics and ensemble smoother with multiple data assimilation method

宗成元, 康学远, 施小清, 吴吉春 水文地质工程地质. 2020, 47(2): 1-8

#### 高盐强酸性地下水中复合苯系污染物原位芬顿氧化实验研究

In-situ Fenton oxidation experiment of compound benzene pollutants in high salt and strong acid groundwater

郭威, 袁放, 张佳, 谢斌, 冯学洋, 陈鸿汉 水文地质工程地质. 2021, 48(2): 182-189

#### 基于iTOUGH2的生物降解模型全局敏感性时变分析

Temporal variation of global sensitivity analysis for biodegradation model using iTOUGH2

杜建雯, 施小清, 徐红霞, 吴吉春 水文地质工程地质. 2020, 47(2): 35-42



关注微信公众号, 获得更多资讯信息

DOI: 10.16030/j.cnki.issn.1000-3665.202208042

吴延浩, 江思珉, 吴自军. 地下水污染强度及渗透系数场的反演识别研究 [J]. 水文地质工程地质, 2023, 50(4): 193-203.

WU Yanhao, JIANG Simin, WU Zijun. Identification of groundwater pollution intensity and hydraulic conductivity field[J]. Hydrogeology & Engineering Geology, 2023, 50(4): 193-203.

## 地下水污染强度及渗透系数场的反演识别研究

吴延浩<sup>1</sup>, 江思珉<sup>1</sup>, 吴自军<sup>2</sup>

(1. 同济大学土木工程学院水利工程系, 上海 200092;  
2. 海洋地质国家重点实验室(同济大学), 上海 200092)

**摘要:** 在制定地下水污染修复方案时, 污染源参数和渗透系数场是最重要的地下水数值模型参数, 但前人研究多集中于单一类型参数的识别。文章中采用地下水污染物运移模型(MT3DMS)和数据同化方法(迭代局部更新集合平滑器, ILUES)构成地下水污染源识别的求解框架, 并利用 Karhunen-Loève 展开技术实现渗透系数场的参数降维, 最后通过同化水头与浓度数据实现地下水污染源强和渗透系数场的联合反演。结果表明: (1)ILUES 算法能精确识别污染源参数和渗透系数场, 并且具有很高的普适性; (2)精确表征渗透系数在空间上呈现出的非均质性, 是预测污染物迁移路径、反演污染强度的关键; (3)ILUES 算法参数影响着反演效果, 综合考虑计算效率和计算精度等, 可以得到算例的最佳样本集合大小( $N_e=4\ 000$ )和 ILUES 算法最佳参数组合(局部临近样本集合占比  $\alpha=0.4$ , 相对权重  $b=4$ )。但在实际工程案例中, 如果对精度的要求不是过高, 经验组合( $\alpha=0.1, b=1$ )更值得推荐。研究结果对于区域地下水资源调查、评价和管理等工作具有较强的实践意义, 并可作为后期地下水污染预测及地下水监测井网优化提供技术支撑。

**关键词:** 地下水污染; 参数反演; 数据同化; 集合平滑器

中图分类号: X523

文献标志码: A

文章编号: 1000-3665(2023)04-0193-11

## Identification of groundwater pollution intensity and hydraulic conductivity field

WU Yanhao<sup>1</sup>, JIANG Simin<sup>1</sup>, WU Zijun<sup>2</sup>

(1. Department of Hydraulic Engineering, School of Civil Engineering, Tongji University, Shanghai 200092, China; 2. State Key Laboratory of Marine Geology, (Tongji University), Shanghai 200092, China)

**Abstract:** The pollution source parameters and hydraulic conductivity field are the most important parameters of groundwater numerical models when making groundwater pollution remediation plans. However, previous studies focused mainly on the identification of single type parameters. The groundwater pollutant transport model (MT3DMS) and data assimilation method (iterative local updating ensemble smoother, ILUES) are used to form a solution framework for groundwater pollution source identification, and Karhunen-Loève expansion technique is used to realize parameter dimension reduction of the hydraulic conductivity field. The joint inversion of groundwater pollution source intensity and hydraulic conductivity field are also realized by assimilating hydraulic heads and concentration data. The results show that (1) the ILUES algorithm can accurately identify pollution

收稿日期: 2022-08-17; 修订日期: 2022-09-24

投稿网址: [www.swdzgcdz.com](http://www.swdzgcdz.com)

基金项目: 国家自然科学基金项目(41976057; 41676061)

第一作者: 吴延浩(1999-), 男, 硕士研究生, 主要研究方向为污染水文地质学。E-mail: [630436663@qq.com](mailto:630436663@qq.com)

通讯作者: 江思珉(1980-), 男, 博士, 副教授, 主要从事地下水数值模拟等方面研究。E-mail: [jiangsimin@tongji.edu.cn](mailto:jiangsimin@tongji.edu.cn)

source parameters and permeability coefficient field, and it is of high universality. (2) Accurate characterization of spatial heterogeneity of the coefficient of permeability is the key to predict pollutant migration path and inversion of pollution intensity. (3) The ILUES algorithm parameters affect the inversion results. By considering the computational efficiency and accuracy, the optimal sample set size ( $N_e=4\ 000$ ) and the optimal parameter combination of ILUES algorithm ( $\alpha=0.4$ ,  $b=4$ ) can be obtained. However, in practical engineering cases, the empirical combination ( $\alpha=0.1$ ,  $b=1$ ) is more recommendable if the requirement for accuracy is not too high. The results of this study have strong practical significance for regional groundwater resources investigation, evaluation and management, and can provide technical support for later groundwater pollution prediction and optimization of groundwater monitoring well networks.

**Keywords:** groundwater pollution; parameter inversion; data assimilation; ensemble smoother

地下水污染的隐蔽性对污染修复带来了巨大挑战<sup>[1-2]</sup>。为了应对这一挑战,地下水污染源识别(groundwater contaminant source identification, GCSI)受到了越来越多的关注<sup>[3-4]</sup>。GCSI能够通过有限的地下水实际观测数据(包括水头和污染物浓度)和其他先验信息来获取地下水污染源的有关信息<sup>[5]</sup>。此外,地下介质以及污染物自身化学反应的空间异质性也给GCSI模型参数带来了新的挑战<sup>[6]</sup>。GCSI模型对含水层性质(如孔隙度、渗透系数等)及生化参数十分敏感,所以在地下水污染的实际应用中需要对渗透系数等空间异质性参数进行准确估计<sup>[7-8]</sup>。因此地下水污染源强和渗透系数的同步识别在GCSI中是亟需解决的问题。

目前,常用的地下水污染溯源及参数识别方法分为以下2种。第1种是模拟-优化(S-O)方法<sup>[9]</sup>,包括遗传算法<sup>[10-11]</sup>、模拟退火算法<sup>[12-13]</sup>和粒子群优化算法<sup>[14-15]</sup>。然而,对于观测数据点较少、未知参数较多的情况,问题的不确定性大大增加,S-O方法效率可能较低<sup>[16]</sup>。第2种方法则是数据同化(DA)方法,包括集合卡尔曼滤波器(EnKF)和集合平滑器(ES)算法,它不仅可以考虑未知参数的不确定性,还考虑了模型结构和模型输入的不确定性<sup>[17]</sup>。如Xue等<sup>[18]</sup>将EnKF集合到贝叶斯模型平均框架,实现了地下水流模型边界和渗透系数场不确定条件下的污染源识别;Li等<sup>[19]</sup>在基于正态得分集合卡尔曼滤波(NS-EnKF)的基础上提出基于正态得分集合平滑器(NS-ES),高效反演出同精度的多峰非高斯渗透系数场。EnKF虽被广泛应用,但其对模型参数和状态是同时进行更新的,在诸如地下水模型这样非线性的模型中,就会引起参数和状态的不一致,并且EnKF需要频繁的同化数据。相比较而言,ES的批处理方法只更新参数,不更新状态,并且一次性使用所有观测数据进行同化,所以

ES可能比EnKF更适用于地下水模型的参数反演。为了进一步提高ES对求解非线性问题的能力,许多学者对该方法进行了改进。Zhang等<sup>[20]</sup>提出了一种迭代局部更新集合平滑器(ILUES),通过对每个样本使用ES进行局部更新并进行多次同化,提高了ES求解性能的同时,也验证了其能量化复杂水文模型的参数不确定性;贺楠楠<sup>[21]</sup>在ILUES的基础上提出多残差深度卷积神经网络耦合迭代局部更新集合平滑器反演框架(MRDCN-ILUES)反演框架,以更低的计算代价,实现了地下水中铬污染源的泄漏速率反演。

基于以上分析,本研究中采用ILUES算法作为反演框架来估计污染源特征和渗透系数场。在以往的研究中,大多集中于污染源强或渗透系数等单一类型参数的识别,而对二者的同步反演则少有研究。本文首先将渗透系数视为已知参数,说明了渗透系数场的准确刻画在污染物信息识别的重要性,在此基础上利用ILUES算法实现了地下水污染源强和渗透系数的同步反演,并确认了最佳参数组合。

## 1 研究方法

### 1.1 地下水流与溶质运移模型

#### (1) 地下水流动方程

根据Darcy定律以及质量守恒定律,在不考虑水的密度变化条件下,三维空间孔隙介质中地下水流动的偏微分方程:

$$\frac{\partial}{\partial x} \left( K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_y \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_z \frac{\partial h}{\partial z} \right) + W = S_s \frac{\partial h}{\partial t} \quad (1)$$

式中: $K_x$ 、 $K_y$ 、 $K_z$ ——分别为渗透系数在 $x$ 、 $y$ 、 $z$ 方向上的分量/ $(\text{m} \cdot \text{d}^{-1})$ ;

$W$ ——源汇项,用以表示流进汇或来自源的水量/ $\text{d}^{-1}$ ;

$h$ ——含水层水位/m;

$S_s$ ——含水介质的贮水率/ $m^{-1}$ ;

$t$ ——时间/d。

(2) 污染物迁移的数学表达式

$$\frac{\partial}{\partial x_i} \left( \theta D_{ij} \frac{\partial C^p}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (\theta v_i C^p) + q_s C_s^p + \sum R_n = \frac{\partial (\theta C^p)}{\partial t} \quad (2)$$

式中:  $\theta$ ——有效孔隙度;

$D_{ij}$ ——水动力弥散系数张量/ $(m^2 \cdot d^{-1})$ ;

$C^p$ ——污染质 p 的质量浓度/ $(kg \cdot m^{-3})$ ;

$v_i$ —— $i$  分量的渗流速度/ $(m \cdot d^{-1})$ ;

$C_s^p$ ——源汇污染质 p 的质量浓度/ $(kg \cdot m^{-3})$ ;

$q_s$ ——源汇单位面积上通量/ $d^{-1}$ ;

$R_n$ ——化学反应项/ $(kg \cdot m^{-3} \cdot d^{-1})$ ;

$x_i, x_j$ ——距离在相应坐标方向上的分量/m。

在建立地下水流三维运动方程和溶质运移模型的基础上, 分别采用 MODFLOW 程序<sup>[22]</sup> 求解地下水流模型, 用 MT3DMS 程序<sup>[23]</sup> 求解溶质运移方程。

## 1.2 ILUES 算法

相比于 ES, 由 Zhang 等<sup>[20]</sup> 提出的 ILUES, 可通过利用扰动后的观测误差的协方差矩阵对观测数据进行多次同化, 解决了当参数的后验分布不服从高斯分布时, ES 方法不适用的问题, 而且提高了反演高维参数的计算效率。ILUES 中采用的协方差矩阵扰动策略来源于 Emerick 等<sup>[24]</sup> 所提出的 ES-MDA 算法。扰动协方差矩阵  $C_D$  通常由下式得到:

$$C_D = N_{iter} C_D \quad (3)$$

式中:  $C_D$ ——观测误差的原始协方差矩阵;

$N_{iter}$ ——ILUES 算法的迭代次数。

传统的集合平滑器方法采用集合内所有参数样本来更新样本, 为了更好的解决高维、强非线性反演问题, ILUES 方法与传统集合平滑方法不同, 采用了一种局部更新策略, 仅使用局部样本来更新整体参数样本<sup>[25]</sup>。具体来说, 对于地下水模型中的参数  $m$ , 不同于 ES 直接更新参数, ILUES 是先更新参数的  $N_e$  个样本的局部集合  $m_j^f (j = 1, \dots, N_e)$ :

$$J(m) = \frac{J_1(m)}{J_1^{\max}} + b \cdot \frac{J_2(m)}{J_2^{\max}} \quad (4)$$

$$J_1(m) = [f(m) - d]^T C_D^{-1} [f(m) - d] \quad (5)$$

$$J_2(m) = [m - m_m^f]^T C_{MM}^{-1} [m - m_m^f] \quad (6)$$

式中:  $J_1(m)$ ——用来度量模型预测值  $f(m)$  与观测值  $d$  之间的距离;

$J_2(m)$ ——用来度量模型参数  $m$  与  $m_m^f$  对应样本之

间的距离;

$b$ ——表示两个距离度量  $J_1(m)$  和  $J_2(m)$  在综合度量标准中的相对权重,  $b \in (0, \infty)$ ;

$J_1^{\max}, J_2^{\max}$ ——分别表示  $J_1(m)$  和  $J_2(m)$  的最大取值;

$C_{MM}$ ——维度为  $N_m \times N_m$  的模型参数集合  $M$  的自协方差矩阵。

具体在更新局部样本集合时, ILUES 首先选择  $J(m)$  值最低的样本  $N_{loc} = \alpha N_e$ ,  $\alpha \in (0, 1]$ , 作为样本  $m$  的局部临近样本集合, 再用集合平滑器更新每个样本的局部临近样本集合  $m_j^{\text{loc},f}$ :

$$m_{j,k}^a = m_{j,k}^f + C_{MD}^{\text{loc},f} (C_{DD}^{\text{loc},f} + C_D)^{-1} [d_k - f(m_{j,k}^f)] \quad (7)$$

式中:  $m_{j,k}^a$ ——更新后的局部集合;

$m_{j,k}^f$ ——更新前的局部集合;

$k$ ——选取的  $J(m)$  最低的样本个数;

$N_{loc}$ ——局部临近集合中样本的个数;

$\alpha$ ——局部临近样本集合  $m_j^{\text{loc},f}$  占全局样本集合  $m$  的比例;

$C_{MD}^{\text{loc},f}$ ——维度为  $N_m \times N_d$  的模型参数局部临近样本集合  $m_j^{\text{loc},f}$  与模型预测值样本集合  $D_j^{\text{loc},f} = [f(m_{j,1}^f), f(m_{j,2}^f), \dots, f(m_{j,N_{loc}}^f)]$  的互协方差矩阵;

$C_{DD}^{\text{loc},f}$ ——维度为  $N_d \times N_d$  的  $D_j^{\text{loc},f}$  的自协方差矩阵;

$d_k$ ——表示在观测值  $d$  基础上引入随机观测误差扰动的样本;

$f(m_{j,k}^f)$ ——预测值。

最终得到经过更新的全部样本集合  $M^a = [m_1^{\text{loc},a}, m_2^{\text{loc},a}, \dots, m_{N_e}^{\text{loc},a}]$

## 2 算例研究

本研究中, 在稳态地下水流条件下, 对流和弥散是二维污染物迁移系统中的主导过程。假设含水层为  $200 m \times 150 m$  的非均匀各向异性承压含水层, 具有二维稳态地下水流, 初始模型网格尺寸为 30 行 40 列, 单元尺寸为 5 m。南北边界为二类隔水边界, 西部和东部边界是 20 m 和 15 m 的恒定水头分边界, 含水层厚度为 1 m (图 1)。含水层参数值见表 1。

初始时刻, 研究区无污染物。含水层渗透系数 ( $K$ ) 满足对数正态分布, 其均值 ( $\overline{\ln K}$ ) 与方差 ( $\sigma_{\ln K}^2$ ) 分别为 3.0 和 0.5, 沿  $x$  和  $y$  方向的相关长度均为 50 m, 变差函数为指数型, 真实渗透系数场见图 2(a)。模拟总时长为 800 d, 共 8 个应力期 (每个应力期 100 d)。为了验证 ILUES 算法反演的有效性, 本算例中考虑 3 个

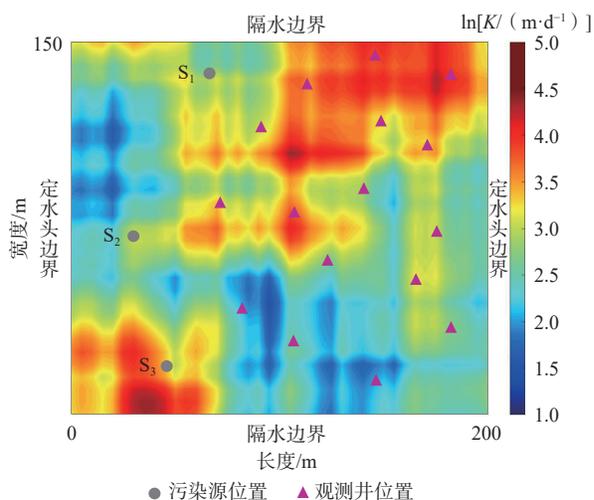


图 1 含水层结构平面图 (S<sub>1</sub>、S<sub>2</sub> 和 S<sub>3</sub> 为 3 个污染源位置)

Fig. 1 Planar structure of the aquifer (S<sub>1</sub>, S<sub>2</sub> and S<sub>3</sub> are the three contamination source locations)

表 1 含水层的水文地质特征

Table 1 Hydrogeological features of aquifers

| 参数名称  | 单位  | 参数值 |
|-------|-----|-----|
| 含水层厚度 | m   | 1.0 |
| 孔隙度   | 无量纲 | 0.3 |
| 纵向弥散度 | m   | 2.0 |
| 横向弥散度 | m   | 0.6 |

污染源, 位置分别为 S<sub>1</sub>、S<sub>2</sub> 和 S<sub>3</sub>。其中 S<sub>1</sub> 在应力期内无规律性释放; S<sub>2</sub> 经过人工修复后, 释放强度不断减小; 而 S<sub>3</sub> 为新污染源, 释放强度持续增大。污染源参数由 24 个参数进行量化(3 个污染源在 8 个应力期的污染释放强度), 表示为 S<sub>ij</sub><sup>i</sup>, 其中 i=1,2,3, 代表第 i 个污

染源, j=1,2,⋯,8, 代表第 j 个应力期, 8 个应力期所指示的时间为 [100(j-1), 100j]。假设表征污染源释放强度的 24 个参数的先验分布是均匀分布, 其先验区间和真实值见表 2。

表 2 污染源释放强度参数的真实值及先验区间

Table 2 True values and a priori interval of the pollution source release intensity parameter

| 应力期 | 先验区间     | 污染源源强/(kg·d <sup>-1</sup> ) |                |                |
|-----|----------|-----------------------------|----------------|----------------|
|     |          | S <sub>1</sub>              | S <sub>2</sub> | S <sub>3</sub> |
| SP1 | [20,100] | 98.0                        | 54.0           | 30.0           |
| SP2 | [20,100] | 87.0                        | 78.0           | 39.0           |
| SP3 | [20,100] | 74.0                        | 64.0           | 46.0           |
| SP4 | [20,100] | 61.0                        | 85.0           | 57.0           |
| SP5 | [20,100] | 47.0                        | 71.0           | 66.0           |
| SP6 | [20,100] | 39.0                        | 61.0           | 71.0           |
| SP7 | [20,100] | 31.0                        | 45.0           | 79.0           |
| SP8 | [20,100] | 24.0                        | 89.0           | 83.0           |

为了估计上述未知参数, 在研究区域布设 16 个观测井(图 1), 并由 MODFLOW 和 MT3DMS 正演得到水头和浓度的模拟值, 由正态分布的测量噪声计算得到观测值。

$$C = C_s + \varepsilon NC_s \quad (8)$$

式中: C——模拟值;

C<sub>s</sub>——观测值;

N——观测噪声(均值为 0, 标准差分别为 0.005 m 和 0.005 kg/m<sup>3</sup>)的高斯分布);

ε——满足标准高斯分布的随机偏差。

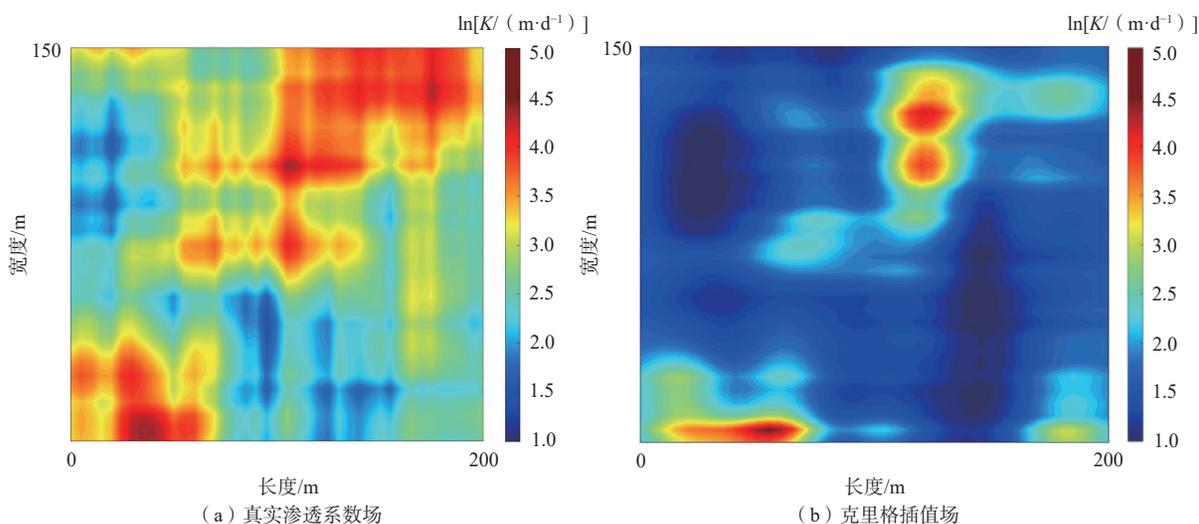


图 2 含水层渗透系数场

Fig. 2 Hydraulic conductivity field of the aquifer

### 3 结果

#### 3.1 算例场景设置

在影响地下水系统的所有不确定因素中,渗透系数是表征含水层特性的一个关键要素,其变异性主要是由自然界含水介质的非均质各向异性、取样和测试的失真以及试验量测误差造成的,而渗透系数是影响污染物在地下水中运移产生不确定性的主要原因。因此,渗透系数的准确获取对于解决污染物反演问题至关重要。在实际场地中,对于含水层的渗透系数场分布的认知往往是零散的,即仅有从井中测量得到的该点处的渗透系数值。为了研究 ILUES 算法对地下水污染强度识别的准确性,对渗透系数场分别采用 3 种方式进行处理(对应于 3 种场景),观测数据都是真实渗透系数场条件下获取的。

1) 场景 1。渗透系数场设置为恒定值是最常用的简化方法。实际上,渗透系数是通过拟合试验结果得到的,但是样本数量有限,无法覆盖整个场地。一般情况下,利用有限的样本得到的渗透系数值的平均值便成为场地渗透系数的代表性定值。本场景案例的渗透系数场取真实渗透系数场的均值( $\ln \bar{K} = 3$ )。

2) 场景 2。采用 Kriging 方法对渗透系数进行插值,得到具有空间变异性的渗透系数场。受技术和经济条件制约,现实中获取的渗透系数数据往往较少,不能充分反映整个含水层渗透系数空间分布特征。因此,大多数情况下需要根据有限的对渗透系数进行空间插值。普通克里格法属线性平稳地质统计学范畴,是地质统计学中最基本的方法,也是最常用的方法之一。设区域化变量满足二阶平稳假设,其数学期望为未知常数,协方差函数和变异函数存在且平稳,以任一未知点为中心点的某一邻域范围内的  $n$  个样本点,赋予相应权重计算未知点的属性值,形成估计领域内  $n$  个信息值的线性组合,该方法即为普通克里格方法(ordinary kriging, OK)。本场景中,已知研究区观测位置的  $K$  值,利用普通克里格插值方法得到整个研究区的渗透系数场,如图 2(b)所示。

3) 场景 3。引入数据降维的方法表征具有空间变异性的渗透系数场。渗透系数场的维度与模型的网格剖分程度有关(本算例为  $30 \times 40 = 1200$  维),通过降维,可以寻找到高维数据的低维子空间,并保留原始数据的大部分变异性。Karhunen-Loève 展开(以下简称 KL 展开)具有离散效率高、得到的随机变量数较少等优点,是一种较为理想的降维方法,因此在这里

采用 KL 展开对地下水模型的高维渗透系数场进行降维:

$$\ln K(x, y) = \overline{\ln K}(x, y) + \sum_{i=1}^{N_{KL}} \xi_i \sqrt{\tau_i} S_i(x, y) \quad (9)$$

式中:  $\xi_i$ ——独立标准高斯随机变量;

$\tau_i$ ——相关函数的特征值;

$S_i(x, y)$ ——相关函数的特征函数;

$N_{KL}$ ——KL 展开项的数目。本算例中选择  $N_{KL} = 200$ , 可以保留渗透系数场约 95.20% 的变异性。

#### 3.2 不同场景下源强的反演结果

在上述 3 种场景中,分别获得整个研究区的渗透系数场,然后写入 MODFLOW 模型中的 BCF(包括其中 LPF、HUF 等)模块并利用 ILUES 算法反演污染强度,因此本节中,渗透系数场不会被视为一个不确定的模型参数。为了对参数不确定性进行充分估计,ILUES 算法参数集合大小  $Ne=500$ ; 为了充分考虑解的收敛性,迭代次数  $N_{iter}=10$ 。同时根据张江江<sup>[26]</sup>的建议,局部集合参数  $\alpha=0.1$ , 距离权重参数  $b=1$ 。图 3 分别是 3 种情境下,随着迭代次数的增加,3 个污染源释放强度反演结果的变化趋势图(从表征污染源信息的 24 个未知参数分别选取了  $S_1$  的  $S_{s1}^1$ 、 $S_{s1}^4$ 、 $S_{s1}^7$ ,  $S_2$  的  $S_{s2}^1$ 、 $S_{s2}^4$ 、 $S_{s2}^7$  和  $S_3$  的  $S_{s3}^1$ 、 $S_{s3}^4$ 、 $S_{s3}^7$ )。

在场景 1 中,通过 ILUES 算法得到准确的污染源特征识别结果,得到的污染源特征参数反演结果偏离于参考值较多,较为明显的参数有  $S_{s1}^1$ 、 $S_{s1}^4$ 、 $S_{s1}^7$  和  $S_{s3}^7$ , 如图 3(a)所示,主要原因在于在该情境下,仅利用恒定的渗透系数值来描述整个场地渗透系数的分布情况,无法充分表征污染物运移的不确定性。在场景 2 中,9 个污染源参数的识别结果,如图 3(b)所示,明显优于场景 1 的结果;仅在个别参数(如  $S_{s1}^1$  和  $S_{s3}^4$ )与真实值之间出现明显偏差,其余 7 个参数识别结果接近于真实值,但是整体识别的准确性和集合样本的稳定性略逊于场景 3 的识别结果,如图 3(c)所示。可能的原因是在场景 2 中,利用若干已知渗透系数(30 个观测点)通过 Kriging 插值方法得到场地渗透系数场的分布情况,在一定程度上考虑了污染物运移的不确定性,但作用是有限的。

综上所述,利用 ILUES 作为反演算法且选用相同的算法参数,从准确性和稳定性角度看,场景 3 的污染源强度识别结果都优于其他 2 种场景。结果表明:ILUES 是能够解决此类反演问题的有效算法;精确表

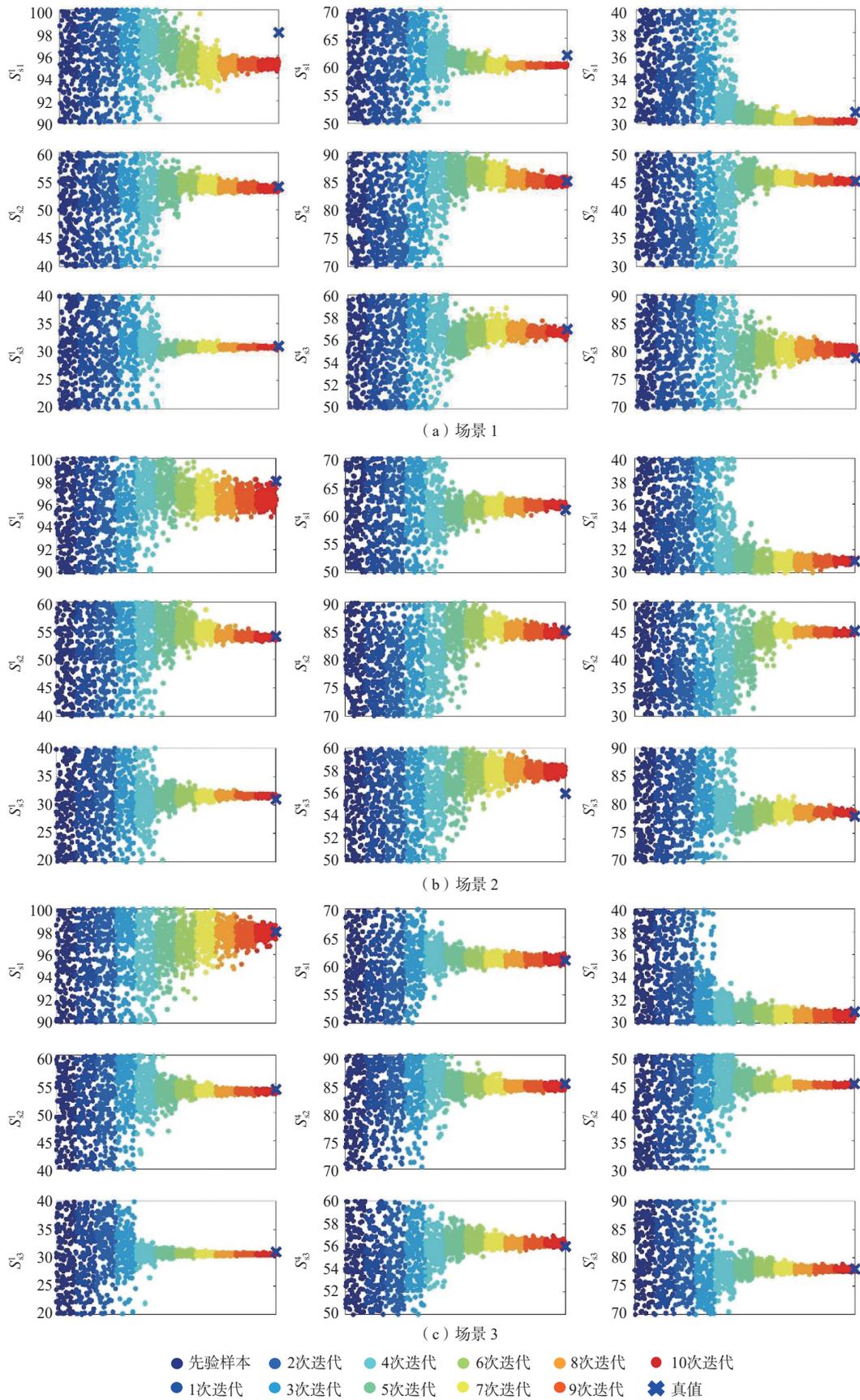


图 3 污染源释放强度反演图

Fig. 3 Inverted intensity of pollution source release

征渗透系数在空间上呈现出的非均质性,是预测污染物迁移路径、反演污染强度的关键,在此基础上才能制定有效的场地修复措施。

## 4 讨论

### 4.1 集合数量对渗透系数场同步反演的影响

在对空间上高变异性的渗透系数场进行同步反演时,对于不同于污染强度的单一反演,将渗透系数同样视为未知参数,同时更新污染源源强和渗透系数两个参数,因此面临更高维地下水逆问题的求解。场景3中的降维方法可以有效减少参数数量,在后续反演中只需反演降维后的参数,即可估计整个渗透系数

场,显著降低了计算压力。因此本节在场景3的基础上,反演24个污染源参数和200个含水层参数。由于未知参数较多,往往会选择更大集合样本数( $N_e$ )进行准确估计,但更大的 $N_e$ 值计算时会更多的调用原始模型,显著增加运算时间和结果存储空间。为了平衡计算效率与计算精度,选用不同的 $N_e$ 值( $N_e=1\ 000, 2\ 000, 3\ 000, 4\ 000, 5\ 000$ )并比较其在ILUES算法下反演的结果。由图4可知迭代次数 $N_{iter}=8$ 时基本收敛,因此为了减短计算时间,本节中取 $N_{iter}=8$ ,其他参数不变( $\alpha=0.1$ 和 $b=1$ )。由于 $N_e=500$ 时24个污染源参数的反演已经比较精确,因此下面仅讨论200个含水层参数(渗透系数场)的反演精度。

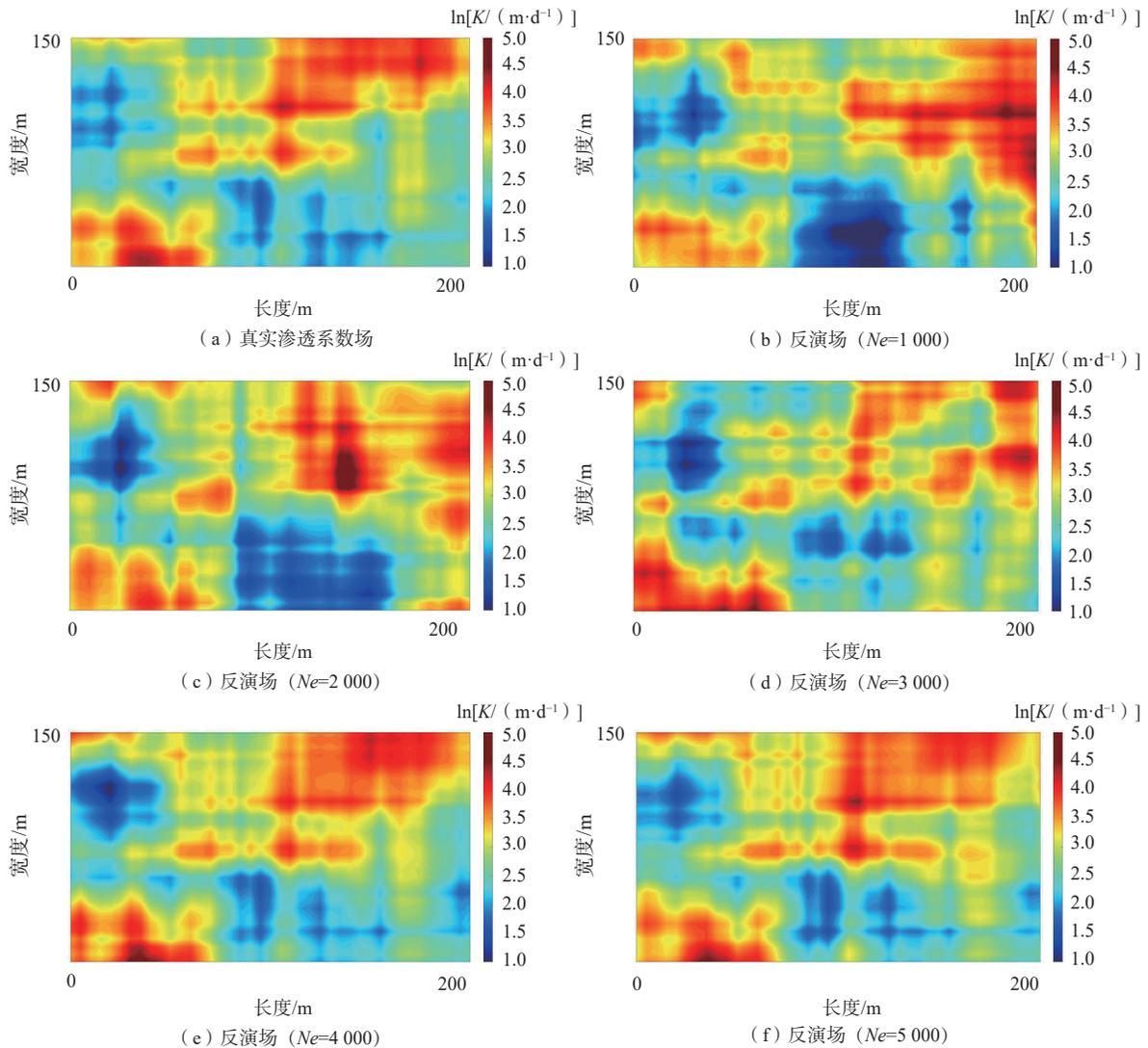


图4 不同 $N_e$ 值下渗透系数场的估计结果

Fig. 4 Estimation of the hydraulic conductivity field at different  $N_e$  values

图4为不同 $N_e$ 值下的渗透系数场的估计结果,由图可以看出, $N_e$ 值从1 000增加至4 000的过程中,

ILUES算法估计的渗透系数场与参考场的相似度逐渐上升,见图4(a)—(d)。但当 $N_e$ 值从4 000继续增

加到 5 000 时, 渗透系数场的估计水平没有明显的提升, 见图 4(e)–(f), 可能原因是  $Ne=4\ 000$  和 5 000 经 8 次迭代, 渗透系数场估计结果的不确定性都达到很低的水平。进一步定量评价 ILUES 方法的参数估计结果, 引入均方根误差 (RMSE) 和平均集合离散度 (AES) 作为定量评价指标。其中, RMSE 用来度量待估计参数的估计值和真实值之间的匹配程度, AES 用来度量参数估计值的不确定性或置信度, 2 者值越低代表参数反演效果越好。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i - M_i)^2} \quad (10)$$

$$AES = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{Ne} \sum_{j=1}^{Ne} (S_{ij} - \widehat{S}_i)^2} \quad (11)$$

式中:  $S_i$ ——第  $i$  个参数的估计值;  
 $M_i$ ——第  $i$  个参数的真实值;  
 $S_{ij}$ ——第  $i$  个参数在第  $j$  个集合中的估计值;  
 $\widehat{S}_i$ ——第  $i$  个参数集合的均值,  $N$  为参数的总个数。

图 5 为不同  $Ne$  值下渗透系数场反演结果的评价指标 RMSE 和 AES 趋势图。可以看出,  $Ne$  由 1 000 到 4 000 时, AES 值在这个过程中由 0.012 5 下降至 0.009, 降幅为 28%, RMSE 值由 0.549 6 下降至 0.211 5, 降幅为 61.5%, 说明用 ILUES 方法估算出的渗透系数场与真实场的相似性逐步增大。但当  $Ne$  由 3 000 到 4 000 的过程中, AES 值和 RMSE 值的变化幅度很小, 降幅分别为 9.1% 和 4.4%, 证实了  $Ne$  值到达 4 000 及以上时候, 反演结果的不确定性已到达一个较低的水平。为了使运算的效率和准确性达到均衡, 认为本场景中  $Ne=4\ 000$  为最佳集合样本大小。

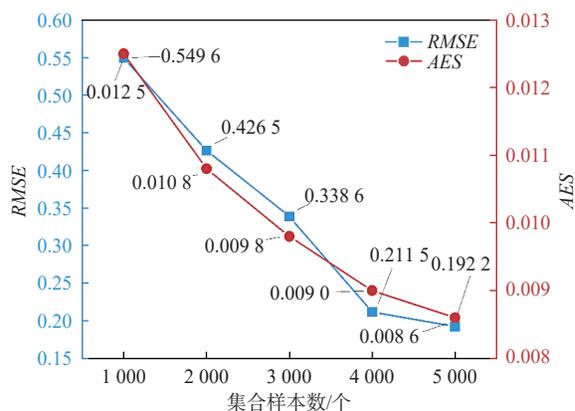


图 5 不同  $Ne$  下的 RMSE 和 AES 的变化趋势图  
 Fig. 5 Trends in RMSE and AES under different  $Ne$  values

#### 4.2 最优参数值的选取

在上述模拟中, 根据经验选择 ILUES 算法的系数

$\alpha=0.1, b=1$  (经验组合), 因为在前人的研究中它具有最好的综合效果。但在实际的情况中, 由于场地的不同, 导致最佳模拟结果的参数组合往往并不是经验组合。为了讨论本算例中最优参数组合, 本节基于上节的场景, 固定  $Ne=4\ 000$  并设置  $\alpha=(0.01, 0.1, 0.2, 0.3, 0.4, 0.5)$  和  $b=(0.01, 0.1, 1, 2, 3, 4, 5)$ , 组合后进行  $6 \times 7=42$  次同步反演, 并计算渗透系数场估计结果的 RMSE 值, 结果如图 6 所示。由图可知, 当  $\alpha < 0.1$  时, 若  $b$  取值较小 (例如 0.01), ILUES 算法可以得到相对较好的拟合结果, 但当选用  $\alpha \geq 0.1$  时, 往往需要相对较大的  $b$  值 (例如  $b \geq 1$ ), 才能得到更好的估计结果, 这与张江江<sup>[26]</sup>得出的结论一致, 因为当  $\alpha$  大于等于 0.1 时, 局部集合的样本数量也会比较大, 因而更可能保留接近观测值的样本, 这时保留和  $m_j^f$  距离更近的样本, 即选择较大的  $b$  值, 会更有利于参数反演<sup>[27]</sup>。特别地, 最佳参数组合 ( $\alpha=0.4, b=4$ ) 也体现了上述规律。

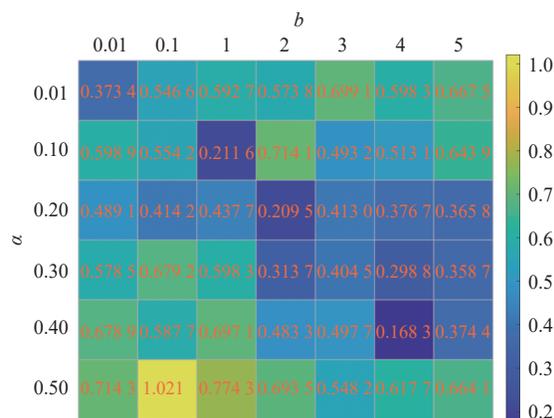
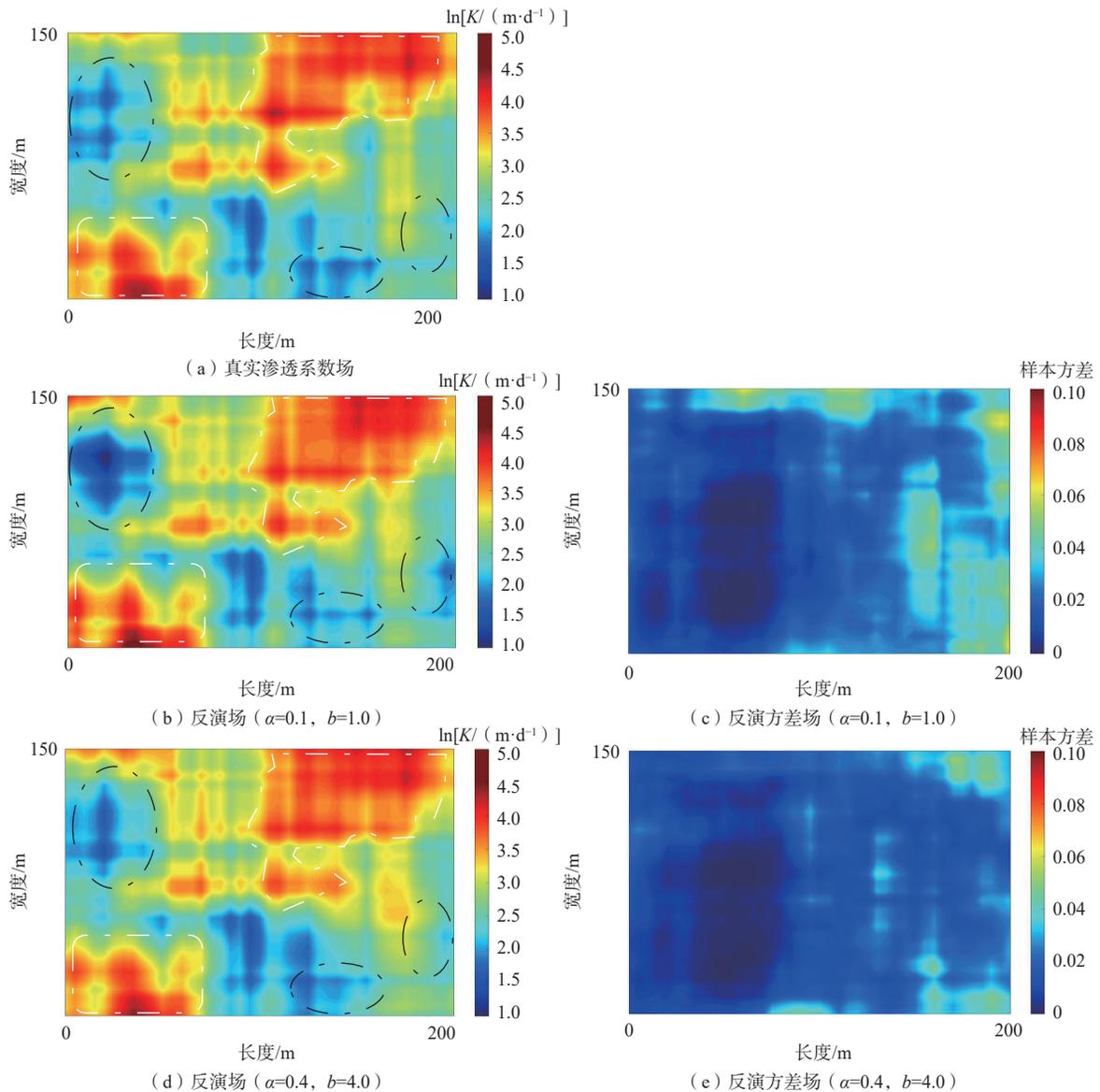


图 6 不同 ( $\alpha, b$ ) 参数组合下的估计渗透系数场的 RMSE 值  
 Fig. 6 RMSE values of an estimated hydraulic conductivity field under different ( $\alpha, b$ ) parameter combinations

进一步类比最佳组合 ( $\alpha=0.4, b=4$ ) 与经验组合 ( $\alpha=0.1, b=1$ ) 的反演结果, 分别进行场景 3 中渗透系数场估计。考虑到 2 种参数组合下反演精度较高, 因此重点比较高低渗透系数区的估计精度。图 7 为 2 种参数组合下的渗透系数场的估计结果, 从形态上看两者均得到与真实场相近的估计结果。对比估计场和真实场的高、低值区, 最佳组合对低渗透性区域刻画更好, 即图 7 中黑色虚线圈定的区域, 并且最佳组合的估计方差场也相对较小, 说明估计的渗透系数场离散型较小, 因此对于本算例而言, 最佳组合 ( $\alpha=0.4, b=4$ ) 即为最优参数组合。

虽然本算例中最优组合为 ( $\alpha=0.4, b=4$ ), 但在实际工程案例中, 由于无法获取研究区域真实的渗透系数

图 7 2 种  $(\alpha, b)$  参数组合下渗透系数场的估计结果Fig. 7 Estimation of the hydraulic conductivity field under the combination of two  $(\alpha, b)$  parameters

场分布,也就无法比较不同的参数组合。鉴于本文中经验组合( $\alpha=0.1, b=1$ )同样取得了不错的精度,并在过往研究中均具有最好的综合效果<sup>[26-27]</sup>,因此在实际应用过程中,如果对精度的要求不是过高,经验组合的普适性或许更值得推荐。

## 5 结论

(1)使用 ILUES 算法同化水头与浓度数据,可以在 8 个迭代步骤内实现多点污染源源强的精确反演,这表明 ILUES 算法在识别污染源方面具有很高的普适性。但同时精确表征渗透系数在空间上呈现出的非均质性,是预测污染物迁移路径、反演污染强度的

关键。

(2)本研究中的场地参数大小为  $30 \times 40 = 1\ 200$  维,如果直接反演渗透系数场将造成极大的运算负荷。通过场景 3 的测试,依托于 KL 展开方法,可以将渗透系数场的参数维度降低至 200 维的同时保留大部分的含水层参数变异性(95.20%),在后续反演中只需反演降维后的参数,即可估计整个渗透系数场,显著降低了计算压力。

(3)ILUES 算法能同步精确识别污染源参数和渗透系数场,其中集合样本数量  $N_e$  和 ILUES 算法参数  $\alpha$  和  $b$  影响着反演效果。综合考虑计算效率和计算精度,本研究中选择  $N_e=4\ 000$  为最佳样本集合大小,

( $\alpha=0.4$ ,  $b=4$ )为算法最佳参数组合。但在实际工程案例中,如果对精度的要求不是过高,经验组合( $\alpha=0.1$ ,  $b=1$ )更值得推荐。

(4)实际上,水文地质问题不仅包括渗透系数的不确定性,还包括着弥散度、孔隙度等参数的不确定性。另外预测污染物运移不能忽略污染物自身的变化过程,因此化学反应参数(如速率常数、分配系数等)的准确获取对于地下水溶质运移模型同样十分重要。未来识别地下水污染源的研究可以集中在这些特征的影响上,进一步验证数据同化方法的普适性。

#### 参考文献 (References) :

- [ 1 ] 刘长礼,王秀艳,赵悦文,等.一种岩溶地下水污染风险的源径标值评价方法[J].地球学报,2021,42(3):363-372. [LIU Changli, WANG Xiuyan, ZHAO Yuewen, et al. A method for evaluation of pollution risk of Karst groundwater by source runoff[J]. *Acta Geoscientica Sinica*, 2021, 42(3): 363-372. (in Chinese with English abstract) ]
- [ 2 ] 王嘉瑜,蒲生彦,胡玥,等.地下水污染风险预警等级及阈值确定方法研究综述[J].水文地质工程地质,2020,47(2):43-50. [WANG Jiayu, PU Shengyan, HU Yue, et al. Review on the determination methods for early warning grade and threshold of groundwater pollution risk[J]. *Hydrogeology & Engineering Geology*, 2020, 47(2): 43-50. (in Chinese with English abstract) ]
- [ 3 ] CHANG Zhenbo, LU Wenxi, WANG Zibo. A differential evolutionary Markov chain algorithm with ensemble smoother initial point selection for the identification of groundwater contaminant sources[J]. *Journal of Hydrology*, 2021, 603: 126918.
- [ 4 ] MO Shaoxing, ZABARAS N, SHI Xiaoqing, et al. Deep autoregressive neural networks for high-dimensional inverse problems in groundwater contaminant source identification[J]. *Water Resources Research*, 2019, 55(5): 3856-3881.
- [ 5 ] 郝辰宇.基于自适应替代模型的地下水数据同化研究[D].杭州:浙江大学,2021. [HAO Chenyu. An adaptive surrogate based method for groundwater data assimilation[D]. Hangzhou: Zhejiang University, 2021. (in Chinese with English abstract) ]
- [ 6 ] 王波,王宇,张贵,等.滇东南泸江流域岩溶地下水质量及污染影响因素研究[J].地球学报,2021,42(3):352-362. [WANG Bo, WANG Yu, ZHANG Gui, et al. A study of quality and pollution factors of Karst groundwater in Lujiang River Basin in southeast Yunnan[J]. *Acta Geoscientica Sinica*, 2021, 42(3): 352-362. (in Chinese with English abstract) ]
- [ 7 ] 宗成元,康学远,施小清,等.基于多点地质统计与集合平滑数据同化方法识别非高斯渗透系数场[J].水文地质工程地质,2020,47(2):1-8. [ZONG Chengyuan, KANG Xueyuan, SHI Xiaoqing, et al. Characterization of non-Gaussian hydraulic conductivity fields using multiple-point geostatistics and ensemble smoother with multiple data assimilation method[J]. *Hydrogeology & Engineering Geology*, 2020, 47(2): 1-8. (in Chinese with English abstract) ]
- [ 8 ] ZHANG Ruicheng, ZHOU Nianqing, XIA Xuemin, et al. Joint estimation of hydraulic and biochemical parameters for reactive transport modelling with a modified ILUES algorithm[J]. *Water*, 2020, 12(8): 2161.
- [ 9 ] THEODOSSIOU N P. Application of non-linear simulation and optimisation models in groundwater aquifer management[J]. *Water Resources Management*, 2004, 18(2): 125-141.
- [ 10 ] 胡鹏,文章,胡新丽,等.基于遗传算法-支持向量机的滑坡渗透系数反演[J].水文地质工程地质,2021,48(4):160-168. [HU Peng, WEN Zhang, HU Xinli, et al. Estimation of hydraulic conductivity of landslides based on support vector machine method optimized with genetic algorithm[J]. *Hydrogeology & Engineering Geology*, 2021, 48(4): 160-168. (in Chinese with English abstract) ]
- [ 11 ] 宋健,杨蕴,吴剑锋,等.混合多目标遗传算法求解地下水污染修复管理模型[J].环境科学学报,2016,36(9):3428-3435. [SONG Jian, YANG Yun, WU Jianfeng, et al. A new hybrid multi-objective genetic algorithm for optimal design of groundwater remediation systems[J]. *Acta Scientiae Circumstantiae*, 2016, 36(9): 3428-3435. (in Chinese with English abstract) ]
- [ 12 ] 江思珉,朱国荣,胡西嘉,等.单纯形-模拟退火混合算法反求水文地质参数及其并行求解[J].地质论评,2007,53(1):92-97. [JIANG Simin, ZHU Guorong, HU Xijia, et al. Hybrid simplex-simulated annealing method for solving hydrogeologic parameters and its parallel implementation[J]. *Geological Review*, 2007, 53(1): 92-97. (in Chinese with English abstract) ]
- [ 13 ] 罗腾,冯晖,郭智奇,等.基于模拟退火粒子群优化算法的裂缝型储层各向异性参数地震反演[J].吉林大学学报(地球科学版),2019,49(5):1466-1476. [LUO Teng, FENG Xuan, GUO Zhiqi, et al. Seismic

- inversion of anisotropy parameters of fractured reservoirs by simulated annealing and particle swarm optimization[J]. *Journal of Jilin University (Earth Science Edition)*, 2019, 49(5): 1466 – 1476. (in Chinese with English abstract)
- [14] 江思珉, 朱国荣, 施小清, 等. 水文地质参数反演的 Hooke-Jeeves 粒子群混合算法[J]. *水科学进展*, 2010, 21(5): 606 – 612. [JIANG Simin, ZHU Guorong, SHI Xiaoqing, et al. Inverse analysis of hydrogeological parameters using hybrid Hooke-Jeeves and particle swarm optimization method[J]. *Advances in Water Science*, 2010, 21(5): 606 – 612. (in Chinese with English abstract)]
- [15] 李小兰, 曾献奎, 王栋, 等. 基于优化-自适应稀疏网格替代模型的地下水模拟参数不确定性分析[J]. *吉林大学学报(地球科学版)*, 2022, 52(4): 1234 – 1243. [LI Xiaolan, ZENG Xiankui, WANG Dong, et al. Uncertainty analysis of groundwater simulation parameters based on optimized-adaptive sparse grid surrogate model[J]. *Journal of Jilin University (Earth Science Edition)*, 2022, 52(4): 1234 – 1243. (in Chinese with English abstract)]
- [16] ZHANG Jiangjiang, ZENG Lingzao, CHEN Cheng, et al. Efficient Bayesian experimental design for contaminant source identification[J]. *Water Resources Research*, 2015, 51(1): 576 – 598.
- [17] LI Liangping, PUZEL R, DAVIS A. Data assimilation in groundwater modelling: ensemble Kalman filter versus ensemble smoothers[J]. *Hydrological Processes*, 2018, 32(13): 2020 – 2029.
- [18] XUE Liang, ZHANG Dongxiao. A multimodel data assimilation framework via the ensemble Kalman filter[J]. *Water Resources Research*, 2014, 50(5): 4197 – 4219.
- [19] LI Liangping, STETLER L, CAO Zhendan, et al. An iterative normal-score ensemble smoother for dealing with non-Gaussianity in data assimilation[J]. *Journal of Hydrology*, 2018, 567: 759 – 766.
- [20] ZHANG Jiangjiang, LIN Guang, LI Weixuan, et al. An iterative local updating ensemble smoother for estimation and uncertainty assessment of hydrologic model parameters with multimodal distributions[J]. *Water Resources Research*, 2018, 54(3): 1716 – 1733.
- [21] 贺楠楠. 基于污染源泄漏速率反演识别的地下水监测网优化研究[D]. 长春: 吉林大学, 2021. [HE Nannan. Study on optimization of groundwater monitoring network based on identification of pollution source leak rate[D]. Changchun: Jilin University, 2021. (in Chinese with English abstract)]
- [22] HARBAUGH A W, BANTA E, HILL M, et al. MODFLOW-2000, the U. S. geological survey modular ground-water flow model-User guide to modularization concepts and the ground-water flow process[R]. U. S. Geological Survey Open-File Report 00-92. USA: U. S. Geological Survey, 2000.
- [23] ZHENG C, WANG P P. MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide[R]. U. S. Army Engineer Research and Development Center Contract Report SERDP-99-1. USA: U. S. Army Engineer Research and Development Center, 1999.
- [24] EMERICK A A, REYNOLDS A C. Ensemble smoother with multiple data assimilation[J]. *Computers & Geosciences*, 2013, 55: 3 – 15.
- [25] 张瑞城, 周念清, 江思珉, 等. ILUES 算法融合 ERT 数据反演污染源参数与渗透系数场[J]. *同济大学学报(自然科学版)*, 2022, 50(2): 223 – 230. [ZHANG Ruicheng, ZHOU Nianqing, JIANG Simin, et al. Estimation of contaminant source parameters and hydraulic conductivity field by assimilating electrical resistance tomography data with iterative local updating ensemble smoother algorithm[J]. *Journal of Tongji University (Natural Science)*, 2022, 50(2): 223 – 230. (in Chinese with English abstract)]
- [26] 张江江. 地下水污染源解析的贝叶斯监测设计与参数反演方法[D]. 杭州: 浙江大学, 2017. [ZHANG Jiangjiang. Bayesian monitoring design and parameter inversion for groundwater contaminant source identification[D]. Hangzhou: Zhejiang University, 2017. (in Chinese with English abstract)]
- [27] 夏学敏. 融合机器学习与数据同化反演地下水污染源与模型参数[D]. 上海: 同济大学, 2021. [XIA Xuemin. Groundwater contamination source and model parameter inversion based on machine learning and data assimilation[D]. Shanghai: Tongji University, 2021. (in Chinese with English abstract)]