陈骏骏,戴振学.基于改进串联式神经网络的反应性溶质运移参数反演研究 [J]. 中国环境科学, 2026,46(2):833-842.
Chen J J, Dai Z X. Reactive transport inverse modeling based on an improved tandem neural network architecture algorithm [J]. China Environmental Science, 2026,46(2):833-842.

基于改进串联式神经网络的反应性溶质运移参数反演研究

陈骏骏[1],戴振学[2*] (1.中国矿业大学矿山互联网应用技术国家地方联合工程实验室,江苏 徐州 221008；2.吉林大学建设工程学院,吉林 长春 130026)

摘要：串联式神经网络(TNNA)是一种通过逆向建模实现参数反演的算法框架,其可靠性高度依赖于替代模型在参数真实后验分布区域的预测精度.对于反应性溶质运移问题,构建高精度的全局替代模型通常需要大量样本数据,在一定程度上限制了反演计算效率.为此,本研究提出了一种基于多保真度(MF)替代模型自适应更新策略(AUS)的TNNA反演算法(TNNA-AUS-MF).该方法在传统TNNA算法的基础上,通过反演结果引导的局部采样建立多保真度替代模型,并在迭代过程中逐步更新模型参数反演结果.算法验证基于一个假想的阳离子交换实验开展.通过控制初始替代模型样本数目,设计了 6 种不同初始替代模型精度的反演计算场景.结果表明,TNNA-AUS-MF 算法相比于 TNNA 算法,能够显著减少正演模拟调用次数并改善反演结果.此外,少样本场景的研究结果验证了利用神经网络构建低保真和高保真模型间非线性关系的可行性,并表明结合自适应更新采样策略能够有效提升反演算法性能.

关键词：串联式神经网络；多保真度替代模型；参数反演；反应性溶质运移模拟；反演结果自适应更新
中图分类号：X523,P641 **文献标识码**：A **文章编号**：1000-6923(2026)02-0833-10

Reactive transport inverse modeling based on an improved tandem neural network architecture algorithm. CHEN Jun-jun[1], DAI Zhen-xue[2*] (1.National and Local Joint Engineering Laboratory of Internet Application Technology on Mine, China University of Mining and Technology, Xuzhou 221008, China；2.College of Construction Engineering, Jilin University, Changchun 130026, China). *China Environmental Science*, 2026,46(2)：833~842

Abstract：The Tandem Neural Network Architecture (TNNA) algorithm was an inversion framework through reverse mapping. Its reliability strongly depended on accurate surrogate modeling within the true posterior region of parameters. For reactive transport problems, establishing a globally accurate surrogate model required a large number of numerical model samples, which considerably limited the efficiency of inversion processes. To address this challenge, the TNNA-AUS-MF inversion method was proposed, in which the original TNNA was integrated with a multi-fidelity (MF) surrogate model and an adaptive updating sampling (AUS) strategy. In this method, multi-fidelity surrogate models were iteratively established based on the conventional TNNA and local samples to update the inversion results. The proposed algorithm was validated using a hypothetical cation-exchange experiment. By controlling the number of initial surrogate model samples, six scenarios with different initial surrogate model accuracies were designed. The results showed that the TNNA-AUS-MF significantly reduced the number of forward simulations while yielding more reliable inversion results compared with the TNNA method. Furthermore, it was demonstrated that the neural network could establish nonlinear relationships between low- and high-fidelity models, and that incorporating the adaptive updating sampling strategy effectively enhanced the performance of the inversion algorithm.

Key words：tandem neural network architecture；multi-fidelity surrogate model；parameter inversion；reactive transport modeling；inversion results adaptive updating

地下水反应性溶质运移模型是预测含水层环境中物质组分迁移演化规律的重要工具,对地下水资源管理和生态环境保护具有重要意义[1-2].但在实际研究中,反应性溶质运移模型的参数条件难以直接确定.通常需要利用容易获取的系统观测数据,通过反演模拟来识别未知模型参数信息[3-4].反演模拟已成为提高模型预测结果可靠性的重要工具,也是地下水数值模拟领域的重要研究方向[5].

目前的反演算法主要基于贝叶斯定理,它从条件概率的角度解释了模型参数的不确定性随系统观测信息的增加而降低的过程[6].其中一类重要的反演算法是模拟-优化算法,它们以求解模型参数后

收稿日期：2025-05-27
基金项目：国家自然科学基金项目(42402241);中央高校基本业务经费(2024QN11066);江苏省安全应急装备技术创新中心(BM2022013)
* 责任作者, 教授, dzx@jlu.edu.cn

验分布概率峰值点为目标,将模型参数的反演识别转变为优化问题.求解方法主要包括局部优化算法和全局优化算法.局部优化算法是根据目标函数的梯度或 Jacobian 矩阵确定模型参数的更新方向,然后将模型参数从初始估计值迭代更新为满足收敛条件的模型参数值.常用的算法有 Gauss-Newton 法[7]、Levenberg-Marquardt 法[8]和共轭梯度法[9]等.这类算法计算效率高,但一旦模型参数更新到梯度接近零的位置时,便停止对模型参数的继续更新.因此,局部优化算法在很大程度上依赖于模型参数起始更新点的选择.它们适用于解决满足凸优化条件的反演问题,否则容易得到局部最优解.全局优化算法通常指元启发式搜索算法,通过执行模型参数解空间的全局探索(广泛搜索参数解空间的不同区域以避免局部最优)和局部开发(针对潜在最优解区域进行细致搜索,以获得更为精确的最优解)来获取具有全局意义的模型参数估计.常用的有粒子群算法[10]、遗传算法[11]和模拟退火算法[12]等.这些算法的本质区别在于全局探索和局部开发过程实现方式的差异.这两个过程的执行程度将共同决定算法的计算效率和最终反演结果的全局可靠性.全局优化算法能够有效避免局部最优解,但通常需要更高的计算成本.并且随着反演模型参数数目的增加或者模型非线性程度的增强,全局优化算法的计算代价将呈指数增长[13].针对这一问题,可通过改进传统算法的参数更新机制或提出新的算法来应对不同场景下的反演问题[14-15].另外,还可以建立替代模型来提高正演模拟的计算效率,从而确保反演算法在规定的研究周期内执行更加充分的模型评估和参数更新[16].人工智能的发展为各领域的研究提供了新技术.其中,串联式神经网络反演算法(TNNA)是基于深度学习建立的一种新型反演算法.它在建立正演模拟替代模型的基础上,以广义最小二乘估计作为约束条件,通过训练逆向神经网络模型实现反演求解.基于逆向神经网络的反演方案能够充分发挥神经网络拟合非线性关系的优势,更高效地找到满足反演约束条件的全局最优解.并且,TNNA 算法还能够通过调整神经网络模型结构来适应不同的反演计算场景,在人工智能快速发展的时代具有巨大的应用和推广潜力[3].

但 TNNA 算法的最终反演结果在很大程度上取决于替代模型在模型参数后验分布区域(模型参数真实值)的预测精度.而后验分布在整个参数先验分布中通常仅占据很小的范围,根据参数先验分布随机采样得到的训练样本数据很难确保后验分布区域形成充足的样本覆盖.针对这一问题,一种可行的方案是引入自适应更新采样策略,逐步提高替代模型在真实参数附近的预测精度[17].我们前期的研究开发了基于自适应更新的串联式神经网络反演算法(TNNA-AUS).它通过逐步优化训练样本数据在真实值附近的分布,来提高替代模型在该区域的预测精度,最终同步提高了模型参数反演精度和算法效率[3].为了更全面地探索 TNNA 算法的改进策略,本研究建立了一种基于多保真度(MF)替代模型自适应更新策略(AUS)的串联式神经网络反演算法(TNNA-AUS-MF).与 TNNA-AUS 反演算法不同,TNNA-AUS-MF 将在每次自适应采样后建立低保真度替代模型与高保真度数值模型之间的非线性映射关系.以此来降低替代模型在真实参数值附近的近似误差,逐步改善模型参数的反演结果.本研究的算法理论将通过一个假想的阳离子交换实验算例进行验证.最终的研究成果有助于拓展智能算法在地下水科学领域的应用,为地下水模型参数反演研究提供新思路.

1 TNNA-AUS-MF 算法基本原理

本研究提出的 TNNA-AUS-MF 算法由 3 个部分组成:分别为初始替代模型,基于逆向映射建模的参数反演识别和基于多保真度替代模型的自适应更新策略.下面将详细介绍三个部分的算法原理及反演算法的整体实现步骤.

1.1 初始替代模型

初始替代模型是根据模型参数先验分布的初始采样来建立的模型参数 m_i 和模型响应 y_i 之间的非线性映射关系:

$$F_{\text{Forward}}\left(m,\theta_{\text{Forward}}\right)\approx F_{\text{HF}}\left(m\right)=\tilde{y} \tag{1}$$

式中: $F_{\text{HF}}(\cdot)$ 和 $F_{\text{Forward}}(\cdot)$ 分别表示数值模型和替代模型的算子;θ_{Forward} 表示正向神经网络模型的可训练参数; $\tilde{y}$ 表示 RTM 模型对于模型参数 m 的数值模拟结果.本研究采用的神经网络结构的主体部分为一个深度残差卷积神经网络.输入层为向量形式,通过连接一个全连接层,将任意维度的参数向量映射为

一个 6400 维的固定向量.再将该固定向量转变为80×80的矩形形状数据,输入到主体部分的残差神经网络模型中.主体部分的残差神经网络结构与文献[6]中的一致.模型输出层为与观测数据维度一致的向量,所采用的激活函数为 Sigmoid.替代模型的损失函数将基于 L1 范数建立得到:

$$\theta_{\text{Forward}}^{*}=\arg\min\frac{1}{N_{\text{Train}}}\sum_{i=1}^{N}\left|F_{\text{Forward}}\left(m_i,\theta_{\text{Forward}}\right)-\tilde{y}_i\right| \tag{2}$$

式中:$\theta_{\text{Forward}}^{*}$表示优化后的权重参数, m_i和$\tilde{y}_i$分别表示训练样本集中第 i 组样本的模型输入和模型输出, N_{Train}表示数据集中的训练样本数目.

1.2　基于逆向映射建模的参数反演识别

逆向映射建模的基本思想是希望通过深度学习建立一个由观测数据向量$\tilde{y}_{\text{obs}}$作为输入变量,m作为输出变量的深度神经网络模型映射, 本研究将其命名为逆向神经网络(Reverse network), 即:

$$m=F_{\text{Reverse}}\left(\tilde{y}_{\text{obs}},\theta_{\text{Reverse}}\right) \tag{3}$$

式中: F_{Reverse}和 θ_{Reverse}分别表示逆向神经网络的数学算子和可训练参数.

本研究中的逆向神经网络模型结构为一个包含两个隐含层的全连接神经网络.每一个隐含层的神经元数目均为 512.隐含层之间连接的激活函数为Swish,输出层的激活函数为 Sigmoid.逆向神经网络的损失函数将根据广义最小二乘来得到,具体表达式如下:

$$\theta_{\text{Reverse}}^{*}=\arg\min\sum_{i=1}^{N_{\text{obs}}}\frac{1}{\sigma_i}[\tilde{y}_{\text{obs}}[i]-F_{\text{Forward}}^{i}\left(F_{\text{Reverse}}\left(\tilde{y}_{\text{obs}},\theta_{\text{Reverse}}\right),\theta_{\text{Forward}}\right)] \tag{4}$$

式中: $\theta_{\text{Reverse}}^{*}$表示优化后的权重参数; σ_i表示第 i 个观测数据的标准差, F_{Forward}^{i}表示正向神经网络 $F_{Forward}$的输出向量中的第 i 个元素值, $\tilde{y}_{\text{obs}}[i]$表示观测数据向量$\tilde{y}_{\text{obs}}$中的第 i 个元素.在完成逆向神经网络的训练之后,根据下式得到模型参数的反演结果:

$$\tilde{m}=F_{\text{Reverse}}\left(\tilde{y}_{\text{obs}},\theta_{\text{Reverse}}^{*}\right) \tag{5}$$

传统TNNA算法的反演精度高度依赖$F_{\text{Forward}}(\cdot)$在模型参数真实值附近的预测精度.下面将介绍本研究提出的基于多保真度替代模型的自适应更新策略,来逐步提高模型参数反演精度.

1.3　基于多保真度替代模型自适应更新策略

多保真度替代模型旨在利用一个神经网络模型,建立当前精度较低的替代模型(低保真度模型 F_{LF} 与高保真度数值模型 $F_{\text{HF}}(\cdot)$之间的非线性映射关系.$F_{\text{MF}}(\cdot)$的输入向量将由模型参数 m 和它对应的低保真度模型输出 $F_{\text{LF}}(m)$一同构成.多保真度替代模型的一般形式表示为:

$$F_{\text{HF}}(m)\approx F_{\text{MF}}(m,F_{\text{LF}}(m),\theta_{\text{MF}}) \tag{6}$$

$F_{\text{MF}}(\cdot)$的训练过程中采用的损失函数与初始替代模型一致.神经网络的结构将采用全连接神经网络模型, 模型结构与逆向神经网络采用的一致.

多保真度替代模型的思想是基于低保真度模型与高保真度模型之间存在一定的相关性,且这种相关性的非线性程度一般低于高保真度模型数据集本身所具有的非线性程度.因此,理论上可以通过增加少量的额外样本结合一个新的神经网络模型来表征这些具有规律性的系统误差[18].额外增加点的位置分布在一定程度上决定多保真度替代模型在哪些区域的预测精度能够得到改善.在模型参数真实解区域未知的条件下,本研究将通过自适应更新采样策略来逐步实现模型参数真实解附近区域样本分布的改善,从而最终提高替代模型在该区域的预测精度.

TNNA-AUS-MF 算法的实现可根据步骤 1~步骤 6 来执行.图 1 展示了算法程序设计的主要模块和相互关系.

步骤 1:根据模型参数的先验信息,基于拉丁超立方抽样和高保真度正演模型 F_{HF},生成一批替代模型的样本数据集.根据该数据集,利用 1.1 节的方法建立初始替代模型,将当前的训练样本数据集和替代模型分别记为 D_{LF} 和 F_{LF}.

步骤 2:根据 1.2 节中的逆向映射建模方法,利用观测数据 $\tilde{y}_{\text{obs}}$ 和当前的 F_{LF} 模型建立反演约束条件,并通过串联式神经网络方法建立逆向神经网络,得到相应的模型参数反演结果 m_{iter}.

步骤 3:以当前的模型参数反演结果 m_{iter} 为中心点设置一个较小的采样区间,获取少量数目 N_{loc} 的局部参数样本 $M_{\text{loc}}=\{m_i\,|\,i=1,...,N_{\text{loc}}\}$,在得到对应的数值模型输出 $Y_{\text{loc}}=\{y_i\,|\,i=1,...,N_{\text{loc}}\}$ 后,将 $\{M_{\text{loc}};Y_{\text{loc}}\}$补充到训练样本数据集中,得到更新后的训练样本数据集 D_{LF}.

步骤 4:根据步骤 3 中得到的更新后的训练样本数据集 D_{LF} 和当前的低保真度替代模型 F_{LF},利用本

节介绍的方法建立一个多保真度替代模型 F_{MF}.

步骤 5:根据步骤 4 中得到的多保真度替代模型 F_{MF},基于 1.2 节中的逆向映射建模方法,以 F_{HF} 作为正演模拟求解器,联合观测数据 $\tilde{y}_{obs}$,建立反演模拟约束条件,对逆向神经网络模型的权重参数进行再次更新,重新得到新的模型参数反演结果 m_{iter}.

步骤 6:计算当前反演结果 m_{iter} 所对应的高保真模型模拟结果 y_{iter}.判断 y_{iter} 与实际观测数据 $\tilde{y}_{obs}$ 之间的拟合精度是否满足收敛条件.如果满足终止条件,则终止迭代循环过程,将当前的 m_{iter} 作为最终的模型参数反演结果.如果不满足终止条件,则重新返回步骤 3 继续执行迭代计算.本研究的终止条件将从以下两个方面来设定:

(1)判断参数估计值 m_{iter} 位置处正向神经网络的近似值 $F_{Forward(i)}(m_{iter})$与高保真正演模型的模型输出值 $F_{HF}(m_{iter})$之间的误差 $\varepsilon_{Forward\text{-}i}$ 是否小于预设阈值.

(2)AUS 迭代循环的次数是否达到最大迭代次数 I_{max}.

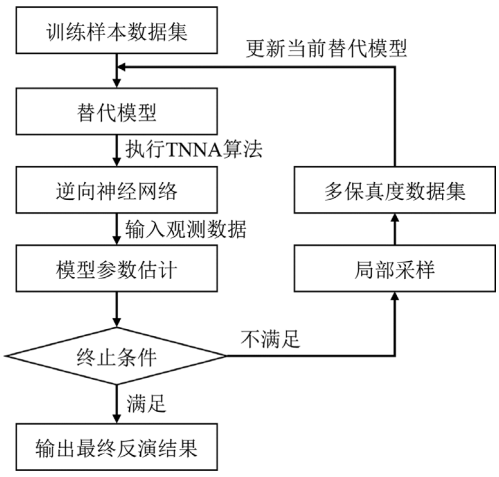


图 1 TNNA-AUS-MF 算法基本流程

Fig.1 Flowchart of the TNNA-AUS-MF algorithm

2 算例设计

2.1 模型概况

本研究提出的 TNNA-AUS-MF 反演算法将通过一个假想的阳离子交换实验模型算例来验证.这一算例最早被 Appelo 等[19]用于测试反应性溶质运移模型的正演模拟程序,也是水文地球化学模拟程序 PHREEQC 中的基准算例之一[20],在 Dai 等[21]的研究中也被用于反演计算程序的测试.

如图 2,该案例为一个 0.08m 长的多孔介质圆柱.所充填介质的平均孔隙度和密度分别为 0.3 和 2650kg/m^3.多孔介质中的所有孔隙均为饱水状态.溶液中的初始溶质组分为 1mmol/L 的硝酸钠($NaNO_3$)和 0.2mmol/L 的硝酸钾(KNO_3).在实验装置的一端以 0.1m/h 的速率注入浓度为 0.6mmol/L 的 $CaCl_2$ 溶液.在 $CaCl_2$ 溶液注入的过程中,实验柱溶液中主要发生如下两种阳离子交换反应:

$$Na^+ + K\text{-}X = K^+ + Na\text{-}X \quad (7)$$

$$Na^+ + 0.5Ca\text{-}X_2 = 0.5Ca^{2+} + Na\text{-}X \quad (8)$$

式中:X 代表阳离子交换位点.模型中的已知阳离子交换容量(CEC)为 0.01779meq/100g. K^+和 Ca^{2+}相对于 Na^+的阳离子交换选择系数分别为 $K_{Na/K}$=0.1995 和 $K_{Na/Ca}$=0.3981.

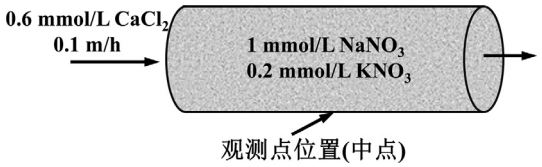


图 2 验证算例的模型示意

Fig.2 Schematic of the synthetic validation model

基于上述信息利用 TOUGHREACT 建立数值模型,模拟溶液中 Ca^{2+}, Na^+, K^+,和 Cl^- 4 种离子组分的浓度随时间的变化情况.建模过程中将多孔介质圆柱等分为 20 个网格单元,每个单元的长度为 Δx=0.004m.根据 TOUGHREACT 中的上游加权法则,本案例中的等效弥散度为 $\alpha=\Delta x/2$=0.002m[22].模型中设置的主要组分及其初始浓度和边界浓度如表 1 所示.

表 1 每种组分的初始浓度和边界浓度

Table 1 Initial and boundary concentrations of chemical species

化学组分	初始浓度(mol/L)	边界浓度(mol/L)
H^+	1×10^{-7}	1×10^{-7}
Ca^{2+}	1×10^{-10}	6×10^{-4}
Na^+	1×10^{-3}	1×10^{-10}
K^+	2×10^{-4}	1×10^{-10}
Cl^-	1×10^{-10}	1.2×10^{-3}
NO_3^-	1.2×10^{-3}	1×10^{-10}

利用上述模型模拟得到 4 种离子组分在溶质中

点位置处的浓度随时间的变化情况.在模拟结果的基础上添加标准差为 0.05 的高斯白噪声作为观测数据(图 3).在该模型中共设置了 7 个待反演参数,分别为:阳离子交换选择系数 $K_{Na/K}$ 和 $K_{Na/Ca}$;Na^+和 K^+的初始浓度 C_{0Na^+}和 C_{0K^+};Ca^{2+}和 Cl^-的边界浓度 $C_{BCa^{2+}}$和 C_{BCl^-};和阳离子交换容量 CEC.这些参数的先验区间如表 2 所示[21].需要注意的是本模型为均质介质中的稳定流模型,因此流场已由边界流量确定,无需关注渗透率参数的反演.

本研究在后续的算法评价过程中,将根据图 3 中的 4 种模拟组分的观测数据,对表 2 中的 7 个未知参数进行反演识别.这一算例模型过程较为简单,便于开展多种不同训练样本数量条件下的反演结果对比.同时,模型包含 4 种不同数量级的溶质组分,可用于评判反演算法能否完成对多种模拟组分的同步校正,实现多源数据同化.

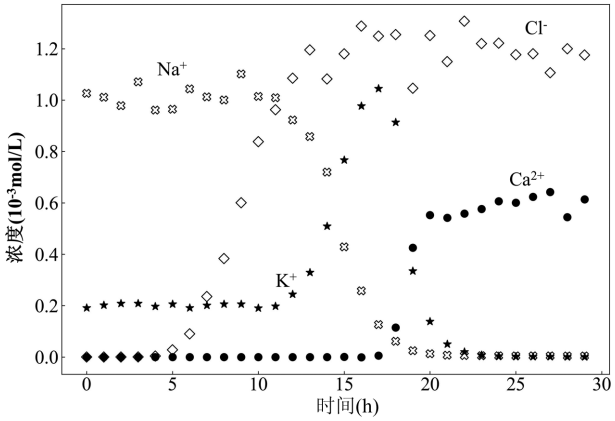


图 3　四种组分在ε~$N(1, 0.05^2)$噪声场景下的观测数据分布

Fig.3　observation data for the four species under noise condition ε~$N(1, 0.05^2)$

表 2　模型待反演参数的先验区间

Table 2　Prior ranges of uncertain model parameters

模型参数	参数先验区间	参数真实值
$K_{Na/K}$	[0.1, 0.3]	0.1995
$K_{Na/Ca}$	[0.25, 0.42]	0.3981
C_{0Na^+} (mol/L)	$[8\times10^{-4}, 8\times10^{-3}]$	1×10^{-3}
C_{0K^+} (mol/L)	$[1\times10^{-4}, 1\times10^{-3}]$	2×10^{-4}
$C_{BCa^{2+}}$ (mol/L)	$[3\times10^{-4}, 8\times10^{-4}]$	6×10^{-4}
C_{BCl^-} (mol/L)	$[5\times10^{-4}, 5\times10^{-3}]$	1.2×10^{-3}
CEC	$[5\times10^{-3}, 5\times10^{-2}]$	0.01779

2.2　验证方案与评价指标

本研究准备了 6 组样本数目 N_{Train} 分别为 100, 200, 500, 1000, 1500 和 2000 的训练样本数据集,从而训练得到 6 个不同预测精度的初始替代模型.再根据这 6 个初始替代模型,分别用 TNNA 算法和 TNNA−AUS−MF 算法进行反演计算.通过对比初始替代模型在不同精度条件下的最终反演结果,来验证多保真度替代模型自适应更新策略对于提高 TNNA 算法反演结果精度的可行性.本研究在执行这一方案过程中,将分别引入初始替代模型的预测精度量化指标和反演算法可靠性量化指标.

初始替代模型的预测精度根据其在测试数据集上预测结果的均方根误差(RMSE)和确定系数(R^2)来量化评价.其中,RMSE 用于衡量替代模型预测结果与数值模拟结果之间的差异;而 R^2 用于评估替代模型与数值模拟结果之间的相关性.RMSE 越小且 R^2 值越接近 1 时的替代模型预测精度越高.这两个指标的数学表达式分别如下:

均方根误差:

$$\mathrm{RMSE}=\sqrt{\frac{1}{N_{\mathrm{test}}}\sum_{i=1}^{N_{\mathrm{test}}}\left\|\tilde{y}_i^{\mathrm{Norm}}-\tilde{y}_i^{\mathrm{DNN}}\right\|_2^2} \tag{9}$$

确定性系数:

$$R^2=1-\frac{\sum_{i=1}^{N_{\mathrm{test}}}\left\|\tilde{y}_i^{\mathrm{Norm}}-\tilde{y}_i^{\mathrm{DNN}}\right\|_2^2}{\sum_{i=1}^{N_{\mathrm{test}}}\left\|\tilde{y}_i^{\mathrm{Norm}}-\bar{y}\right\|_2^2} \tag{10}$$

式中:N_{test} 表示测试数据集的样本数目,$\tilde{y}_i^{\mathrm{DNN}}$ 表示正向神经网络对第 i 个测试样本预测得到的响应向量,$\tilde{y}_{i(\mathrm{test})}^{\mathrm{Norm}}$ 表示第 i 个测试样本的标准化数值模拟结果向量,$\bar{y}=1/N_{\mathrm{test}}\sum_{i=1}^{N_{\mathrm{test}}}\tilde{y}_{i(\mathrm{test})}^{\mathrm{Norm}}$ 是根据所有 $\tilde{y}_{i(\mathrm{test})}^{\mathrm{Norm}}$ (i=1,…,N_{test})得到的均值向量.

在基于优化理论的反演算法研究中,反演结果的可靠性主要根据模型校正后预测结果与实际观测数据之间的拟合程度来评价.本研究的反演计算主要依靠公式(4)中的损失函数得到.因此,将通过替换公式(4)中的替代模型为数值模型来计算得到评价拟合精度的校正误差:

$$\mathrm{Loss}_{\mathrm{Fit}}=\sum_{i=1}^{N_{\mathrm{obs}}}\frac{1}{\sigma_i}\left[y_{\mathrm{obs}}[i]-F_{\mathrm{HF}}(M^*)[i]\right]^2 \tag{11}$$

式中: M^*表示模型参数反演结果.

另外,在本研究采用的假想算例中,模型参数的真实值均为已知.因此也可以通过计算每一项模型参数反演结果与真实值之间的相对误差,来量化它

们之间的差异.但采用相对误差来验证反演算法时还需要综合考虑模型参数的敏感性.模型参数不同程度的敏感性反应了它们在反演计算中的可识别性差异.当某项模型参数的敏感性较高时,代表该参数的可识别程度较高,反之亦然.图 4 为各项模型待反演参数针对损失函数的局部敏感性程度,它反映了各项模型参数的可识别性.这些参数的局部敏感性曲线的变幅数量级在 10^{-6}~10^{-3}.

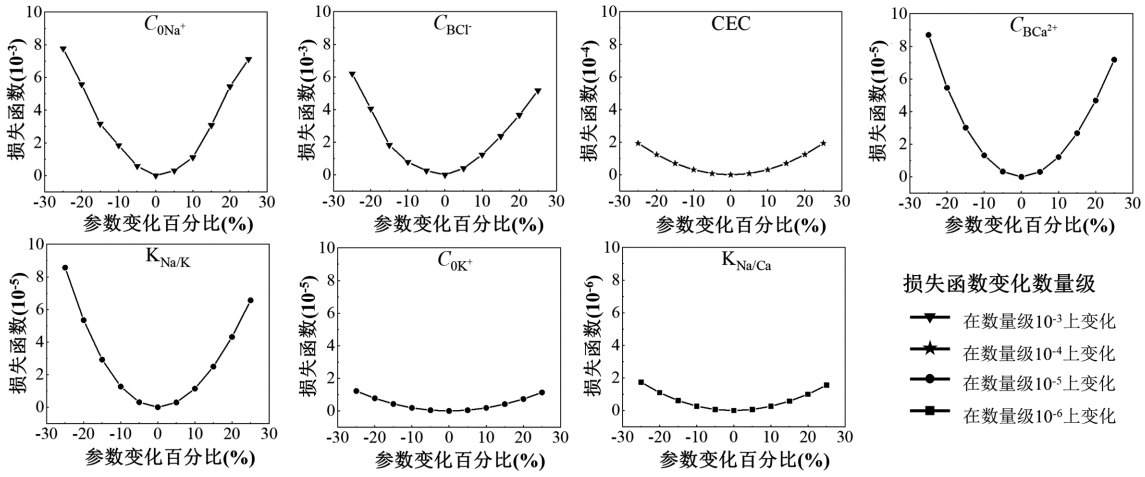


图 4 模型待反演参数敏感性分析结果

Fig.4 Sensitivity analysis of the uncertain model parameters

3 结果与讨论

图 5(a)~(f)展示了 6 组不同样本数目对应的初始替代模型在测试数据集上的预测精度.图中每个散点表示某种组分在特定时空坐标点的替代模型与数值模型预测结果的对比.散点的横纵坐标分别表示替代模型和数值模型的模拟结果.替代模型的预测精度越高,数据点将越集中分布在对角参考线附近.从图 5 中的结果可以看出,本研究采用的替代模型方法需要在训练样本数达到 1000 以上时,才能确保测试数据集上的预测结果与数值模拟结果基本一致.具体来说,当训练样本数 N_{Train} 从 100 增加到 1000 时,替代模型在测试数据集上的预测精度显著提升, RMSE 从 0.091 降至 0.021, R^2 从 0.896 升高至 0.994.当 N_{Train} 进一步增大到超过 1000 之后,替代模型的整体预测精度不再得到显著改善; N_{Train} 分别为 1000, 1500 和 2000 时所对应的 RMSE 和 R^2 分别稳定在 0.015~0.021 和 0.994~0.997.因此,可以将不同训练样本数目训练后所得到的替代模型作为具有不同初始预测精度的替代模型,用于评价多保真度替代模型自适应更新策略对改善 TNNA 算法反演精度的效果.

图 6 为基于不同初始替代模型的 TNNA 和 TNNA-AUS-MF 算法对各项模型参数的反演结果与真实值之间的对比情况.图 7 展示了各项模型参数反演结果的相对误差对比.图 8 和图 9 分别为 TNNA 和 TNNA-AUS-MF 方法反演校正后的模拟曲线与观测数据之间的拟合对比.

根据图 6~8 的结果可知,当 TNNA 算法中的替代模型训练样本数目较少时,难以给出可靠的模型参数反演结果.当 N_{Train} 分别为 100 和 200 时,TNNA 反演校正后的模拟结果显著偏离观测值,具体表现在 TNNA-100 对 Cl^-和 K^+的反演校正结果,以及 TNNA-200 对 Na^+和 K^+的反演校正结果(图 8(a)和(b)).在这两种反演算法得到的反演结果中存在多项模型参数明显偏离真实值的情况(图 6).其中各有四项模型参数的相对误差超过了 0.2(图 7(a)和(b)).特别的,对于敏感性变幅在 10^{-3} 数量级的两项模型参数 C_{0Na^+} 和 C_{BCl^-} ,TNNA-200 算法分别给出了相对误差接近 0.5 和超过 1.0 的反演识别结果.这表明初始替代模型的精度显著影响着具有较高可识别性的模型参数反演识别精度,而这些模型参数的反演误差也是导致模型校正误差的重要因素.在 N_{Train} 达到 500 后,TNNA 算法能够基本保证四种模拟组分的校正结果接近各自的观测值,但个别组分仍然存在局部拟合误差.并且进一步增加 N_{Train} 并未有效改善这些局部拟合误差.例如,在 N_{Train} 为 500, 1000 和 1500

时, 0~10h 段的 Na^+反演校正模拟结果均略低于其观测数据.根据图 8 和图 9 中的结果,每种替代模型所对应的 TNNA−AUS−MF 反演校正精度均高于 TNNA 反演算法.且 6 种不同预测精度的初始替代模型所对应的 TNNA−AUS−MF 算法均给出了能同步拟合 4 种模拟组分观测数据的反演结果,拟合精度指标 $LOSS_{Fit}$ 均在 1.5×10^{-4} 以下.除局部敏感性最低的 $K_{Na/Ca}$ 参数外,各项模型参数的反演结果相对误差均在 0.05 以下.另外,当初始替代模型的训练样本数目为 100 时,TNNA−AUS−MF 算法给出的模型参数反演结果拟合精度指标为 $LOSS_{Fit}=1.443\times10^{-4}$,同样优于 TNNA−2000 算法的反演结果 ($LOSS_{Fit}=2.342\times10^{-4}$).上述 TNNA 和 TNNA−AUS−MF 算法的反演精度对比结果表明,基于多保真度替代模型的自适应更新策略能通过增加少量的额外训练样本提升模型参数真实值附近的替代模拟预测精度.由于 TNNA−2000 算法和 TNNA−AUS−MF−100 算法在执行反演计算时分别需要 2000 次和 185 次 (100+17×5)正演数值模拟,因此 TNNA−AUS−MF 算法能够以 TNNA 算法所需计算成本的十分之一来给出精度更高的反演结果.因此,TNNA−AUS−MF 算法在计算效率方面也具有显著的优势.

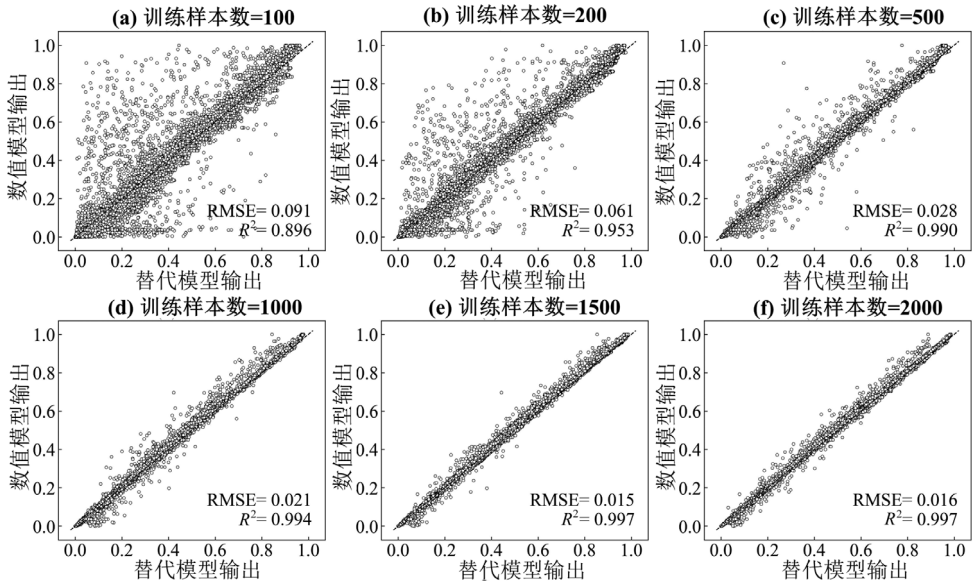


图 5　不同样本数的训练数据集得到的替代模型的预测精度

Fig.5　Prediction accuracy of surrogate models under different training sample sizes

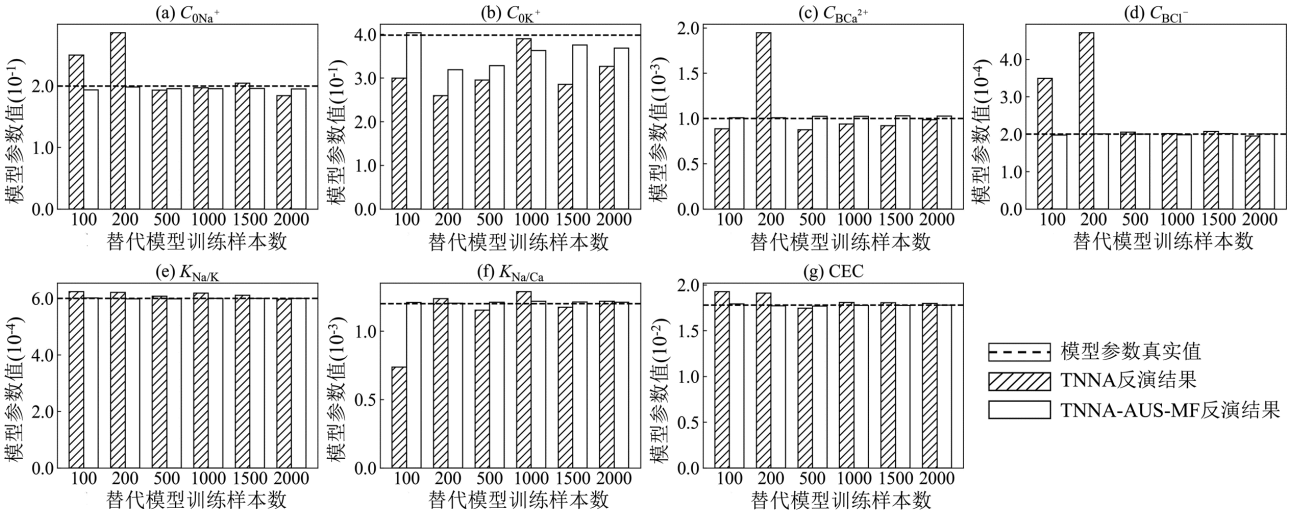


图 6　TNNA 算法和 TNNA−AUS−MF 算法得到的模型参数反演结果对比

Fig.6　Comparison of the parameter inversion results obtained by the TNNA and TNNA−AUS−MF algorithms

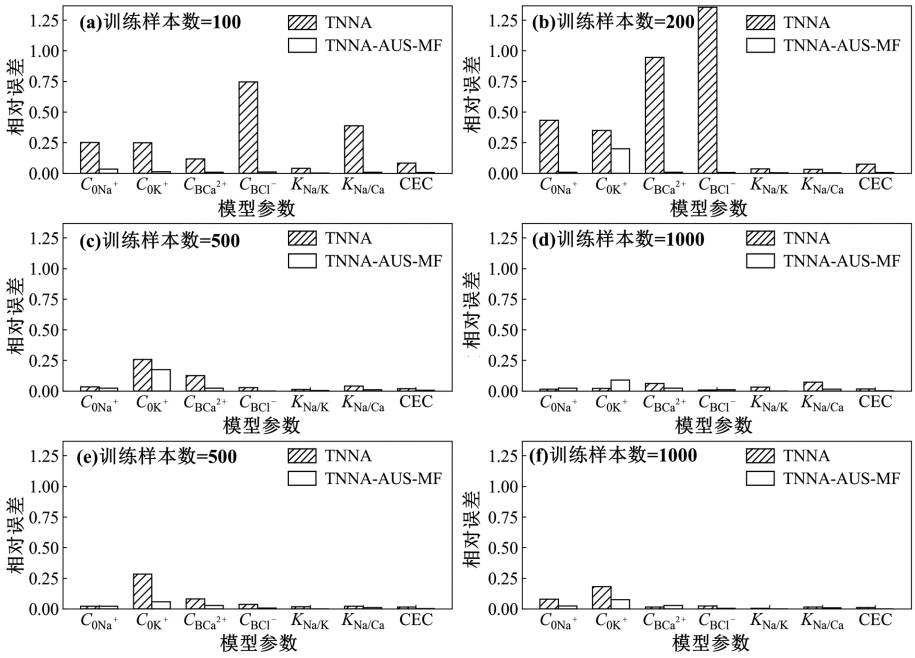


图 7 TNNA 和 TNNA−AUS−MF 算法反演结果相对误差对比

Fig.7 Comparison of relative errors between the inversion results of the TNNA and TNNA−AUS−MF algorithms

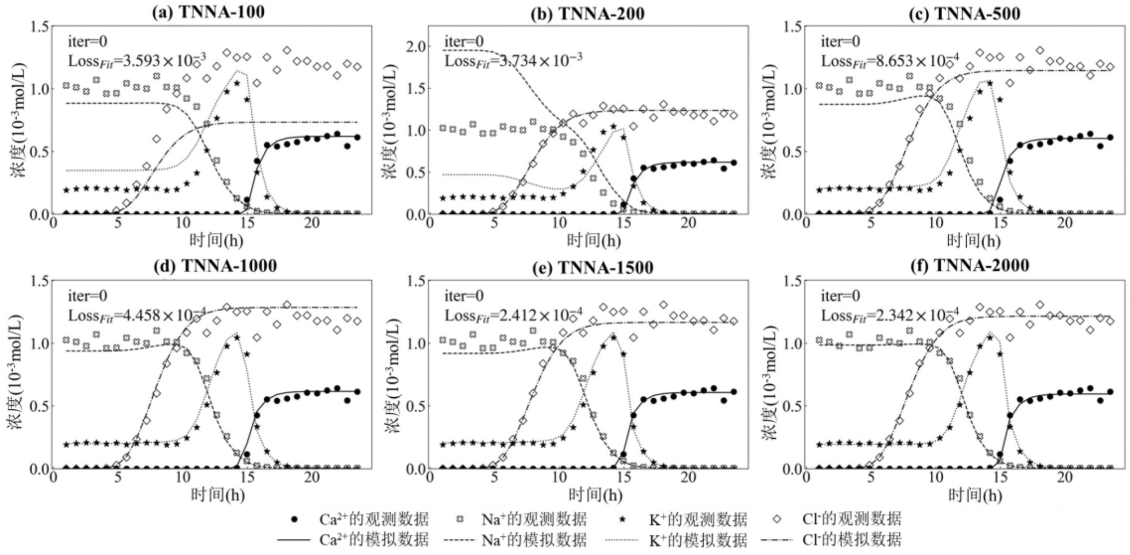


图 8 TNNA 算法反演结果得到的模型校正拟合情况

Fig.8 Model calibration based on the TNNA inversion results

上述研究结果表明,在 TNNA 算法基础上融合多保真度替代模型自适应更新策略所建立的 TNNA−AUS−MF 算法不仅能够提高 TNNA 算法反演结果的可靠性,还可以有效避免替代模型样本分布不合理而导致的模型参数反演结果偏差.相比于 TNNA 算法, TNNA−AUS−MF 算法的鲁棒性和反演结果的稳定性得到了显著提升.

在地下水环境科学研究中,一个不可回避问题是算法在实际场景中的适用性.在实际研究中,不同的模拟场景可能涉及不同类型未知参数.对于参数间数量级差异而带来的影响,一般可以通过标准化处理来得到解决.多数情况下反演问题的复杂性是由模型参数的空间特征和模型过程的非线性强度共同决定的.因此,需要明确不同参数空间特征问题和非线性程度下的应对措施.本研究提出的 TNNA−AUS−MF 算法在 TNNA 算法的基础上,引入了改进替代模型局部预测精度的方法策略,而每次自适应迭代过程中的反演计算仍依赖原始 TNNA 算法的参数更新机制.根据我们

前期的研究[3,23],原始的 TNNA 算法主要适用于低维参数的模型场景.对于高维空间相关性参数的反演问题,需要额外引入生成式机器学习等参数化降维方法,并通过对降维后的参数进行反演,间接得到高维参数的反演结果.因此, TNNA-AUS-MF 算法在处理此类问题时也可以采取类似的措施.另外,本研究提出的 TNNA 算法改进策略,本质上是通过逐步利用多个神经网络来建立初始替代模型与高保真数值模型之间的非线性关系.初始替代模型的神经网络结构和训练样本的选择共同决定了其与数值模型之间的差异;而在自适应更新过程中,多保真度神经网络结构的选择则决定了每次更新能够降低的差异程度.因此,当模型涉及更多物理化学过程并增强非线性特征时,需要根据具体问题选择合适的神经网络结构,以确保 TNNA-AUS-MF 算法能够在预设的自适应迭代次数内获得可靠的反演结果.

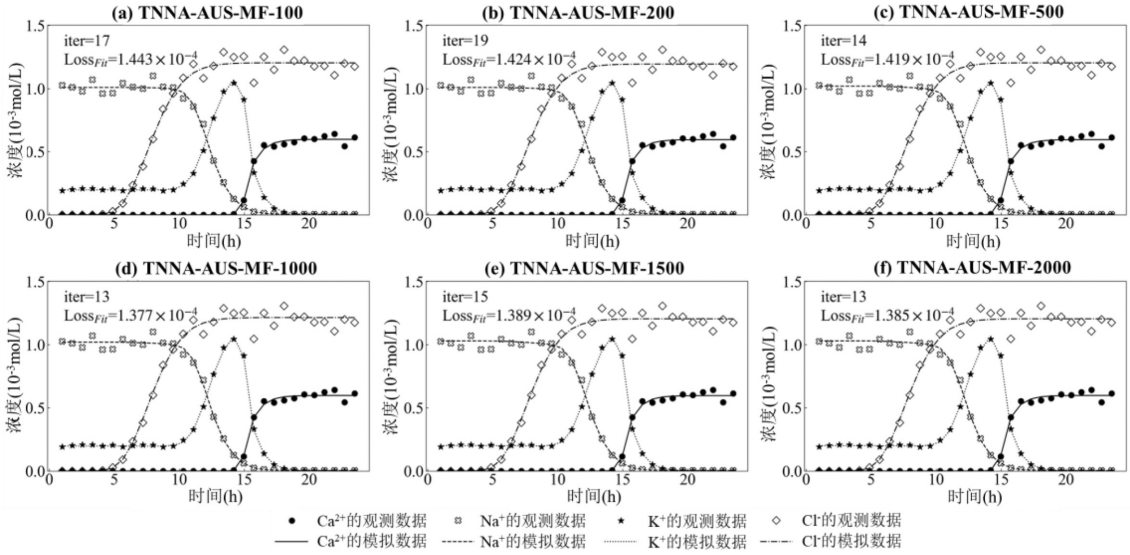


图 9　TNNA-AUS-MF 算法反演结果得到的模型校正拟合情况

Fig.9　Model calibration based on the TNNA-AUS-MF inversion results

4　结论

4.1　在利用 TNNA 算法进行反演模拟时,需要给定足够多的替代模型训练样本数据,通过充分覆盖模型参数的整个先验空间来确保参数真实值区域的替代模型精度,从而保证最终反演结果的可靠性.

4.2　TNNA-AUS-MF 算法中采用的多保真度替代模型自适应更新策略能够逐步改善替代模型训练样本数据在参数真实值区域的密集程度,提高该区域的替代模拟精度,确保最终反演结果的可靠性.同时也验证了利用神经网络建立低保真度和高保真度模型之间非线性关系的可行性.

4.3　TNNA-AUS-MF 算法能够主动搜索有利于模型校正的参数样本,显著减少反演模拟所需的高保真数值模型调用次数.相比于 TNNA 算法,TNNA-AUS-MF 降低了反演结果对初始替代模型预测精度的依赖,具有更高的鲁棒性和稳定性.

5　展望

在未来的研究中,可以根据实际问题结合针对性的机器学习方法,将本研究的算法思路拓展到更多的场景中.例如,可通过引入生成式机器学习方法将其拓展到高维参数反演研究领域.对于计算效率极低的高保真度模型(如包含矿物动力学反应的模型),还可以考虑基于物理模型驱动的替代模型方法来进一步提升整体算法的计算效率.本研究所采用的基于多保真度替代模型的自适应更新策略同样可以与其它优化算法相结合,形成更多高效准确的反演算法.

本研究的 TNNA-AUS-MF 算法本质上属于确定性反演算法,需要给定充分的约束条件确保反演结果的唯一性.为了应对实际研究中普遍存在的多模态和不适定性研究场景,模型参数的反演结果难免出现异参同效性.因此,未来还可以关注如何拓展

TNNA 算法实现随机反演模拟的研究,在给出模型参数最优估计的同时量化不确定性信息.

参考文献：

[1] 许天福,金光荣,岳高凡,等.地下多组分反应溶质运移数值模拟:地质资源和环境研究的新方法 [J]. 吉林大学学报(地球科学版), 2012,42(5):1410-1425.
Xu T F, Jin G R, Yue G F, et al. Subsurface reactive transport modeling: a new research approach for geo-resources and environments [J]. Journal of Jilin University (Earth Science Edition), 2012,45(5):1410-1425.

[2] Li L, Maher K, Navarre-Sitchler A, et al. Expanding the role of reactive transport models in critical zone processes [J]. Earth-Science Reviews, 2017,165:280-301.

[3] Chen J J, Dai Z X, Yang Z J, et al. An improved tandem neural network architecture for inverse modeling of multicomponent reactive transport in porous media [J]. Water Resources Research, 2021,57(12): 2021WR030595.

[4] Cao M X, Dai Z X, Chen J J, et al. An integrated framework of deep learning and entropy theory for enhanced high-dimensional permeability field identification in heterogeneous aquifers [J]. Water Research, 2025,268:122706.

[5] Zhan C J, Dai Z X, Yang Z J, et al. Subsurface sedimentary structure identification using deep learning: A review [J]. Earth-Science Reviews, 2023,239:104370.

[6] 陈骏骏.基于深度学习的地下水反应性溶质运移模型参数反演研究 [D]. 长春:吉林大学, 2023.
Chen J J. Deep learning based inverse modeling of groundwater reactive transport models [D]. Changchun: Jilin University, 2023.

[7] Qin Y W, Kavetski D, Kuczera G, et al. Can Gauss-Newton algorithms outperform stochastic optimization algorithms when calibrating a highly parameterized hydrological model? A case study using SWAT [J]. Water Resources Research, 2022,58(11):e2021WR031532.

[8] Lin Y Z, O'Malley D, Vesselinov V V. A computationally efficient parallel Levenberg-Marquardt algorithm for highly parameterized inverse model analyses [J]. Water Resources Research, 2016,52(9): 6948-6977.

[9] Hill M C. Solving groundwater flow problems by conjugate-gradient methods and the strongly implicit procedure [J]. Water Resources Research, 1990,26(9):1961-1969.

[10] Sadler B, Pulliam J. Modeling crustal structure in the permian basin by waveform-Matching P receiver functions and autocorrelograms with particle swarm optimization [J]. Journal of Geophysical Research: Solid Earth, 2023,128(6):2023JB026730.

[11] Leon A S, Bian L L, Tang Y. Comparison of the genetic algorithm and pattern search methods for forecasting optimal flow releases in a multi-storage system for flood control [J]. Environmental Modelling & Software, 2021,145:105198.

[12] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing [J]. Science, 1983,220(4598):671-680.

[13] Yeh W W G. Review: Optimization methods for groundwater modeling and management [J]. Hydrogeology Journal, 2015,23(6): 1051-1065.

[14] Turgut O E, Turgut M S, Kırtepe E. A systematic review of the emerging metaheuristic algorithms on solving complex optimization problems [J]. Neural Computing and Applications, 2023,35(19): 14275-14378.

[15] Rajwar K, Deep K, Das S. An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges [J]. Artificial Intelligence Review, 2023,56:13187–13257.

[16] Asher M J, Croke B F W, Jakeman A J, et al. A review of surrogate models and their application to groundwater modeling [J]. Water Resources Research, 2015,51(8):5957-5973.

[17] Gong W, Duan Q Y. An adaptive surrogate modeling-based sampling strategy for parameter optimization and distribution estimation (ASMO-PODE) [J]. Environmental Modelling & Software, 2017,95: 61-75.

[18] Yan L, Zhou T. An adaptive surrogate modeling based on deep neural networks for large-scale Bayesian inverse problems [J]. arXiv preprint arXiv:08926, 2019.

[19] Appelo C A J, Postma D. Geochemistry, groundwater and pollution [M]. Rotterdam, Brookfield: A. A. Balkema, 1993:xvi+536.

[20] Parkhurst D L, Appelo C A J. User's guide to PHREEQC (Version 2): A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations [R]. Denver, Colorado: US Geological Survey, 1999:238-242.

[21] Dai Z X, Samper J. Inverse problem of multicomponent reactive chemical transport in porous media: Formulation and applications [J]. Water Resources Research, 2004,40(7):W07407.

[22] Xu T F, Sonnenthal E, Spycher N, et al. TOUGHREACT user's guide: A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media, Version 2.0 [R]. Berkeley, California: Lawrence Berkeley National Lab (LBNL), 2012.

[23] Chen J J, Dai Z X, Yin S X, et al. Enhancing inverse modeling in groundwater systems through machine learning: a comprehensive comparative study [J]. Hydrology and Earth System Sciences, 2025, 29(17):4251-4279.

作者简介：陈骏骏(1990-),男,浙江东阳人,副教授,博士,主要从事地下水模型参数反演算法研究.发表论文 20 余篇.junjunchen@cumt.edu.cn.