

GeoChem AI：利用 PHREEQC 和 COMSOL Multiphysics 协同研究复杂地质和化学现象的分析

包庆君
aibaoqj@163.com

摘要

地质化学模拟在地下环境研究中具有重要作用，尤其在放射性废物地质处置、地下水污染治理和碳捕获与地质封存等领域。PHREEQC 和 COMSOL Multiphysics 是地质化学和多物理场建模领域的重要工具。本研究探讨了如何利用两者的强大功能协同研究复杂的地质与化学现象，提出了一种基于动态反馈的多阶段研究流程。通过整合化学反应建模与多物理场模拟，结合深度学习技术，实现对长时间尺度和高复杂性地质系统的预测与优化。本研究为地球化学研究提供了一种高效、精准且动态的研究方法，并且可以落地实施为 GeoChem AI 软件平台，大幅提升科研生产效率。

关键词：GeoChem, AI, PHREEQC, COMSOL Multiphysics, 地球化学模拟, 动态反馈, 多物理场建模

1. 引言

地球化学过程研究是地下环境模拟的重要部分，对核废物地质处置、地下水污染扩散预测及矿物成因研究等领域具有重要意义。然而，传统模拟方法面临计算效率低、动态性不足、多工具难以协同等挑战。PHREEQC 专注于化学反应建模，COMSOL Multiphysics 则具有强大的多物理场建模能力。如何整合两者的优势，协同研究复杂地质与化学现象，已成为学术界关注的焦点。

本研究提出了一种利用 PHREEQC 和 COMSOL 联合建模的方法，通过动态反馈机制和深度学习技术，提高模拟精度与效率。本文首先回顾两者的功能特点，然后分析其协同研究的具体方法，最后以典型案例验证其有效性。

2. 问题的提出与整合研究的课题示例

作为地球化学学科，地下环境模拟中常见化学对于地质的长期影响，可以利用 PHREEQC 和 COMSOL Multiphysics 的强大功能，通过将 PHREEQC 的精确化学平衡和反应计算与 COMSOL 的物理场耦合与空间动态模拟结合，可以深入研究复杂的地质化学现象，揭示自然系统中多物理场交互作用的核心机制。这种跨软件的协作提供了更全面的研究能力。

但是如何跨软件平台的协同研究一直是业界亟待解决的问题，不仅需要多专业协作的知识技能，更依赖软件的操作易用性。当前的解决方案主要以人工传输接口数据，根据专家经验设定参数，经反复调整获取最优迭代结果。为发挥各软件的优势协同工作，考虑创建多场物理数据接口，结合 LSTM（Long Short-Term Memor）长短期记忆循环神经网络，CNN（Convolutional Neural Networks）卷积神经网络，进行自动调参和长期影响预测，更佳快捷有效的获取计算结果。

以下列举若干跨软件整合研究的课题示例，以供选择：

2.1. 岩溶地貌形成过程模拟

研究目标：模拟地下水与碳酸盐岩（如石灰岩）之间的化学反应，分析溶蚀作用及其对岩溶地貌发展的影响。

- PHREEQC 的作用：
 - 模拟地下水中碳酸盐的溶解-沉淀平衡，计算饱和指数和溶质浓度。
 - 分析 pH、CO₂ 分压、矿物饱和状态对溶蚀速率的影响。
- COMSOL 的作用：
 - 建立三维流体流动模型，模拟地下水流速和流向。
 - 耦合化学反应模型，模拟溶蚀作用对孔隙和裂隙扩展的动力学影响。

2.2. 地下水污染物迁移与化学反应耦合

研究目标：研究重金属或污染物（如铬、砷）在地下水系统中的迁移、扩散及其与土壤矿物的吸附、沉淀反应。

- PHREEQC 的作用：
 - 模拟污染物在地下水中的溶解、吸附、沉淀、氧化还原反应。
 - 计算不同条件下的污染物移动性和化学形态。
- COMSOL 的作用：
 - 模拟地下含水层的流体流动、污染物扩散和对流。
 - 耦合物质传输和化学反应动力学，预测污染物的空间分布和长期行为。

2.3. 火山岩浆与地下水相互作用

研究目标：模拟高温岩浆接触地下水时的热传递、矿物溶解、沉淀以及气体逸出的复杂过程。

- PHREEQC 的作用：
 - 模拟高温下水-岩反应的化学平衡，预测新生成矿物和溶液成分。
 - 模拟硫化物、碳酸盐等的溶解及气体（如 CO₂、H₂S）的释放。
- COMSOL 的作用：
 - 模拟热传递、地下水循环以及岩浆冷却过程。
 - 耦合热流、化学反应和流体流动，分析高温区域的地质-化学变化。

2.4. 二氧化碳地质封存（CCS）效率评估

研究目标：研究 CO₂ 在深部储层中的溶解、矿化及其与岩石/流体的长期相互作用。

- PHREEQC 的作用：
 - 模拟 CO₂ 注入后溶解在地下水中的化学平衡，预测碳酸盐矿物沉淀及其封存效率。
 - 分析储层矿物饱和状态和水化学变化。
- COMSOL 的作用：
 - 模拟 CO₂ 在储层中的注入、扩散、对流和溶解动力学。
 - 耦合流体流动、化学反应和多相流模型，评估注入稳定性和长期封存能力。

2.5. 深部热液矿床的成矿过程模拟

研究目标：研究热液流体与围岩的相互作用过程，预测矿物沉淀位置及成矿条件。

- PHREEQC 的作用：
 - 模拟热液中金属离子和配体的溶解-沉淀反应。
 - 计算不同温度、压力条件下的矿物饱和状态。
- COMSOL 的作用：
 - 模拟深部热液流体的流动路径和热传递。
 - 耦合流体流动和化学反应，预测成矿位置和规模。

2.6. 盐湖沉积物的成因与演化

研究目标：模拟蒸发作用下盐湖水体的化学变化和矿物沉淀过程。

- PHREEQC 的作用：
 - 模拟蒸发过程中的水-盐平衡，预测盐类矿物的沉淀顺序。
 - 分析水化学成分对盐类形成的影响。
- COMSOL 的作用：
 - 模拟盐湖表层的蒸发速率和水流循环。
 - 耦合水动力学和化学反应，预测不同区域的盐矿沉积特征。

2.7. 放射性废物地质处置的长期安全性

研究目标：评估放射性废物的地下处置对环境的长期影响，包括放射性核素的迁移与化学行为。

- PHREEQC 的作用：
 - 模拟放射性核素（如铀、钚）的溶解、沉淀和吸附反应。
 - 分析不同条件下的核素稳定性和迁移风险。
- COMSOL 的作用：
 - 模拟废物容器腐蚀、地下流体对废物的侵蚀和扩散。
 - 耦合地下水流动、热传递和放射性物质迁移模型。

3. 跨软件协同研究基本流程与适用场景

3.1. PHREEQC 和 COMSOL 的功能特点

1. PHREEQC：
 - 专用于化学平衡和动力学模拟，支持溶解、沉淀、氧化还原、离子交换等反应。
 - 提供多种数据库（如 phreeqc.dat, wateq4f.dat），广泛用于地质化学研究。
2. COMSOL Multiphysics：
 - 以有限元法为核心，支持流体流动、热传递、化学反应、多相流等多物理场耦合建模。
 - 提供灵活的几何建模工具和强大的求解器。

3.2. 协同研究流程

1. 化学建模（PHREEQC）：
 - 确定地下水的初始化学条件，包括 pH、氧化还原电位（Eh）、离子浓度等。
 - 模拟核素迁移、矿物饱和状态和化学反应速率，生成初步化学结果。
2. 物理场建模（COMSOL）：
 - 根据 PHREEQC 的化学输出，定义流体流动、热传递或溶质迁移的初始条件。
 - 构建三维几何模型，耦合对流-扩散方程与化学反应速率。
3. 动态反馈机制：
 - 利用监测数据更新模型参数。
 - 应用深度学习（如 LSTM）实现核素迁移的时间序列预测。

4. 敏感性分析与优化：

4. 敏感性分析与优化：

- 采用 Sobol 方法量化关键参数对模型结果的影响。
- 提供优化建议，改进工程设计。

3.3. 协同研究适用场景

场景 1：深地岩石中的放射性废物处置库

研究目标：研究花岗岩或岩盐地层中核素的长期迁移行为和屏障稳定性。

分析内容：

- 核素的扩散和吸附行为。
- 热效应对地层裂隙渗透率的影响。
- 地下水与屏障材料（如膨润土）之间的化学相互作用。

场景 2：膨润土屏障的性能退化

研究目标：研究膨润土在长期水化、化学反应和放射性核素影响下的劣化过程。

分析内容：

- 膨润土吸水膨胀后渗透率变化。
- 高温下膨润土矿物的溶解与转化。
- 吸附核素能力的动态变化。

场景 3：容器腐蚀与核素释放

研究目标：评估金属容器腐蚀引起的核素释放速率及其迁移路径。

分析内容：

- 容器材料的腐蚀速率与裂纹形成。
- 腐蚀产物对核素迁移的化学屏障作用。
- 核素释放后在裂隙和孔隙介质中的扩散。

场景 4：长时间尺度的地质演化

研究目标：预测数千年到数百万年时间尺度上地质和化学过程对核素迁移的影响。

分析内容：

- 气候变化导致的地下水动力学变化。
- 地壳应力变化引起的裂隙网络重组。
- 长期矿物反应对屏障性能的改变。

4. 研究课题示例

4.1. 课题示例 1：放射性废物的地质处置长期安全性

4.1.1. 研究题目设定

放射性废物地质处置是一种将放射性废物深埋于稳定地质构造中的技术，利用地质屏障和工程屏障的协同作用，阻止放射性物质迁移到生物圈。在理论研究中，结合 PHREEQC 和 COMSOL Multiphysics 的能力，可以对核素迁移、化学反应、流体流动和长期环境影响进行全面模拟，长期安全性可以被更全面、精确地评估。这种跨软件的研究框架有助于指导工程设计和政策制定，为放射性废物的安全处置提供科学依据。

4.1.2. 研究的核心问题

1. 放射性核素迁移行为

- 核素通过扩散、对流和化学反应迁移。
- 核素与周围矿物或水体发生吸附、溶解、沉淀、配位和氧化还原反应。

2. 工程屏障和地质屏障的性能

- 工程屏障（如金属容器、膨润土屏障）的耐腐蚀性、膨胀性、渗透性和隔离性能。
- 地质屏障（如岩盐、花岗岩、黏土岩）的长期稳定性和对核素迁移的限制能力。

3. 热-水-化学-力学耦合效应

- 废物热衰变产生热量影响地下水流动、矿物溶解和裂隙变化。
- 多物理场（热、流体、化学、应力）的相互作用对处置库长期安全性的影响。

4. 长期稳定性与环境风险

- 模拟数千至数百万年的地质演化过程。
- 预测核素释放后可能的扩散范围和对周边环境的影响。

4.1.3. 详细理论分析

1. 放射性核素的化学行为

- PHREEQC 可用于精确计算放射性核素（如铀、钚、镅、锔）在地下水中的溶解度、配位状态和吸附行为。
- 吸附与溶解平衡：核素与膨润土或地质矿物（如伊利石、高岭石、氧化铁等）之间的表面吸附。

- 氧化还原反应：核素如 U(IV)/U(VI) 的氧化还原状态变化对其迁移性影响。
- 沉淀反应：如 UO_2CO_3 的沉淀形成降低迁移速率。

2. 核素迁移与屏障作用

- PHREEQC：
 - 评估核素在复杂化学体系中的迁移路径。
 - 考虑多种反应：溶解-沉淀、氧化还原、离子交换和吸附。
- COMSOL：
 - 结合流体力学（Darcy 流动模型）和溶质传输方程（对流-扩散方程），模拟核素在裂隙网络和孔隙介质中的迁移。
 - 耦合流体流动和化学反应，分析屏障的阻隔效应。
 - 模拟屏障的劣化过程（如膨润土的干裂和容器的腐蚀）。

3. 热效应对系统的影响

- 影响预设：放射性废物的热衰变会引起周围岩石温度升高，影响化学反应和地下水流动。
- PHREEQC：
 - 模拟高温条件下的矿物溶解度和反应动力学。
- COMSOL：
 - 使用热传递模块模拟放射性废物的热分布。
 - 耦合热流与流体流动，分析热-化学-流体效应。

4. 多物理场耦合

- COMSOL 可整合热-水-化学-力学（THMC）耦合模拟：
 - 热效应：废物衰变产生热量影响地下水流动和矿物反应。
 - 水力效应：地下水通过岩石孔隙和裂隙迁移的流体动力学特性。
 - 化学效应：核素的化学反应和矿物溶解-沉淀动态。
 - 力学效应：热膨胀引起岩石应力变化，导致裂隙扩展或闭合。

4.2. 课题示例 2：放射性废物处置中的核素迁移模拟

4.2.1. 研究题目设定

放射性废物处置中的核素迁移模拟是评估放射性废物长期安全性的关键问题。通过结合 PHREEQC 和 COMSOL Multiphysics 的能力，可以对核素的扩散、对流、化学反应及其与屏障材料的交互作用进行全面建模，并通过深度学习实现动态预测。本课题旨在模拟不同条件下核素迁移的动态行为，验证屏障设计的有效性，并提出优化建议。

具体案例如：预测铀（U）在含水层中的迁移行为，评估膨润土屏障的长期性能。

4.2.2. 研究的核心问题

1. 核素迁移路径与浓度分布：
 - 核素通过对流、扩散和化学反应迁移的动力学特性。
 - 环境因子（pH、Eh、流速）对迁移行为的影响。
2. 屏障材料的长期性能：
 - 膨润土对核素吸附的稳定性。
 - 屏障的劣化过程（如裂隙扩展、材料渗透率变化）。
3. 多物理场耦合效应：
 - 热衰变导致的温度场与流体场相互作用。
 - 化学反应与物理迁移之间的动态耦合。

4.2.3. 研究方法与流程

1. PHREEQC 化学模拟：
 - 输入初始地下水条件（pH: 7.5, Eh: 300 mV、核素浓度等）。
 - 模拟 UO_2^{2+} 的吸附、溶解-沉淀过程，输出吸附等温线和矿物饱和状态的沉淀行为。
 - 输出矿物饱和状态和吸附等温线。
2. COMSOL 物理场建模：
 - 构建三维含水层几何模型，模拟地下水的对流与扩散。
 - 构建多物理场耦合模型：对流-扩散、热传递与化学反应。
 - 耦合 PHREEQC 的化学反应速率，预测核素浓度的空间分布。
 - 定义屏障材料的初始渗透率与吸附特性。
 - 模拟核素在屏障与含水层中的时空分布。
3. 动态反馈与深度学习预测：
 - 铀的迁移范围受吸附系数和屏障渗透率显著影响。
 - 基于 LSTM 模型，训练核素迁移的时间序列预测模型。
 - 利用监测数据更新模拟输入校正模型，动态反馈分析显示屏障的劣化将显著加速核素迁移。
4. 敏感性分析与优化设计：
 - 采用 Sobol 方法评估关键参数的敏感性。
 - 提出基于模型预测的屏障优化建议。
5. 结果与讨论
 - 核素迁移行为显著受 pH、吸附系数和屏障完整性影响。
 - 屏障劣化将显著提高核素释放风险，需设计多层屏障以增强可靠性。

- 深度学习方法有效预测了长期迁移趋势，显著提升了模型精度与效率。

5. 机器学习与深度学习的应用

通过引入机器学习和深度神经网络，可以显著提高地质处置长期安全性分析的效率和准确性。GeoChem AI 与 PHREEQC 和 COMSOL 的结合，不仅可以弥补传统模拟方法的不足，还能在不确定性分析、复杂过程建模和实时预测中发挥重要作用，为放射性废物处置提供可靠的科学支持。

5.1. 应用方案举例

1. 模拟和预测复杂非线性过程

- 问题：地质处置系统中的物理、化学和力学过程高度非线性，传统模拟方法可能耗时且不够灵活。
- 解决方案：
 - 通过机器学习模型（如支持向量机、随机森林、神经网络），训练快速预测工具。
 - 使用深度神经网络（DNN）建立数据驱动模型，从历史模拟数据中学习复杂关系，预测长期行为。

2. 不确定性分析

- 问题：地下环境的参数（如渗透率、矿物分布、核素吸附系数）具有高不确定性。
- 解决方案：
 - 利用贝叶斯方法或生成对抗网络（GAN），生成参数分布的样本并进行不确定性量化。
 - 构建代理模型，通过神经网络快速分析大量参数组合的结果。

3. 降低计算成本

- 问题：传统的数值模拟（如耦合热-水-化学-力学模型）需要高计算成本。
- 解决方案：
 - 使用机器学习模型代替高维数值模拟的部分功能，例如通过元模型（metamodeling）简化复杂模拟。
 - 训练 DNN 替代部分 COMSOL/PHREEQC 模块，在大规模计算中快速逼近结果。

4. 数据挖掘与模式识别

- 问题：长期监测数据和模拟结果中可能隐藏关键信息或模式。
- 解决方案：
 - 使用卷积神经网络（CNN）从空间分布数据中提取核素迁移的关键特征。
 - 使用聚类算法（如 K-means 或 DBSCAN）发现核素迁移或屏障劣化中的潜在模式。

5. 长期行为的外推预测

- 问题：实验室和现场实验的时间尺度远小于地质处置所需的预测时间（数千到数百万年）。
- 解决方案：
 - 使用循环神经网络（RNN）或长短期记忆网络（LSTM），模拟时间序列数据并外推未来行为。

5.2. 机器学习与深度学习结合 PHREEQC 和 COMSOL 的方法

1. 数据生成

- 使用 PHREEQC 和 COMSOL 生成高质量的模拟数据，作为机器学习模型的训练数据：
- 模拟核素在不同地质条件下的迁移路径、化学反应。
- 模拟屏障材料的长期性能变化。

2. 数据预处理

- 整理和处理模拟数据：
- 对模拟结果进行归一化和降维。
- 提取关键物理和化学参数作为模型输入特征。

3. 建立机器学习模型

- 使用以下模型进行训练：
- 回归模型（线性回归、随机森林、梯度提升树）：预测单一物理量（如核素迁移速率）。
- 深度学习模型（如多层感知机、卷积神经网络）：用于处理复杂非线性关系或空间分布。
- 时间序列模型（如 RNN 或 LSTM）：预测长期演化过程。

4. 结果验证与校准

- 通过以下步骤验证机器学习模型：
- 使用实验数据或现场监测数据验证预测结果的准确性。
- 与 PHREEQC 和 COMSOL 的高精度模拟结果对比，校准模型。

5. 集成到 workflow

- 将机器学习模型嵌入到 PHREEQC 和 COMSOL 的工作流中，实现快速预估和实时反馈。
- 例如：在模拟过程中，机器学习模型可以预测某些参数变化对结果的影响，而无需重复运行全模型。

5.3. 具体应用场景

1. 参数敏感性分析

- 场景：分析核素迁移对不同地质参数（如地下水流速、吸附系数、矿物饱和状态）的敏感性。
- 解决方法：
 - 使用神经网络构建代理模型，快速分析大量参数组合的结果。
 - 确定对长期安全性影响最大的关键参数。

2. 实时监测与预测

- 场景：结合地下监测系统，实时预测核素迁移或屏障性能。
- 解决方法：
 - 使用 LSTM 模型对监测数据进行实时分析，预测未来的核素分布。
 - 根据模型预测结果调整监测策略。

3. 数据驱动的设计优化

- 场景：优化屏障材料的性能（如膨润土的吸附能力、容器的抗腐蚀性）。
- 解决方法：
 - 利用生成对抗网络（GAN）生成新材料设计的候选方案。
 - 使用 DNN 模型快速评估不同材料组合的长期性能。

4. 长时间尺度预测

- 场景：预测数千到数百万年时间尺度上的核素行为。
- 解决方法：
 - 使用 RNN 或 LSTM 进行时间序列外推预测。
 - 将地质演化（如裂隙扩展、地下水变化）纳入预测模型。

5.4. 挑战与解决策略

1. 模型训练数据的不足

- 挑战：高质量的训练数据获取困难。
- 解决策略：
 - 使用 PHREEQC 和 COMSOL 生成合成数据进行模型训练。
 - 结合实验室数据和现场监测数据进行模型校准。

2. 长时间尺度的准确性

- 挑战：时间尺度过长可能导致模型外推误差积累。
- 解决策略：
 - 使用多模型融合技术，将机器学习预测结果与传统模拟结果结合。
 - 定期更新模型输入，降低外推误差。

3. 多物理场的复杂性

- 挑战：多物理场耦合的复杂性难以完全捕捉。
- 解决策略：
 - 结合物理约束的神经网络（Physics-Informed Neural Networks, PINNs），在深度学习模型中加入物理规律约束。

6. 跨软件人工智能综合应用示例

以下是结合 PHREEQC 和 COMSOL 的应用举例，利用 GeoChem AI 的深度学习模型建立代理模型以加速分析。

6.1. 应用示例 1：参数敏感性分析的详细过程

6.1.1. 研究目标

在放射性废物地质处置中，参数敏感性分析的目标是评估不同地质和化学参数（如地下水流速、核素吸附系数、矿物溶解速率等）对核素迁移行为的影响，识别对处置安全性最关键的参数。

通过结合 PHREEQC 和 COMSOL 的数值模拟，生成高质量数据用于深度学习模型训练，再利用神经网络进行敏感性分析和快速预测，可以显著提高参数分析的效率和精度。这种方法不仅能降低计算成本，还能为放射性废物处置提供科学依据和设计优化建议。

6.1.2. 分析步骤

1. 确定关键参数与初始条件

- 参数选择：
 - 地质参数：
 - 地下水流速
 - 渗透率
 - 孔隙率
 - 化学参数：
 - 核素的吸附分配系数（Kd）
 - 溶解-沉淀速率
 - pH、氧化还原电位（Eh）
 - 屏障性能参数：
 - 膨润土的渗透率
 - 容器腐蚀速率
- 场景设置：
 - 放射性废物处置库位于花岗岩或岩盐地层，含膨润土屏障。
 - 模拟核素（如铀、钚）从容器释放后的迁移行为。

2. 使用 PHREEQC 模拟化学反应

- 任务描述：模拟地下水中核素的化学反应，包括溶解-沉淀、吸附、氧化还原等。
- 任务目的：为每个参数组合提供精确的化学反应结果，生成用于后续敏感性分析的化学特征数据。
- 模拟过程：
 - PHREEQC 输入文件设计：
 - 定义地质和化学初始条件：

```
SOLUTION 1
  pH 7.5
  pe 4.0
  U(6) 1e-6 # 浓度 (mol/L)
  Ca 2.5
  Cl 1.0
```

- 定义吸附表面与反应模型：

```
SURFACE 1
  Hfo_wOH
  sites 0.005
  -equilibration with solution 1
```

- 模拟矿物溶解：

```
EQUILIBRIUM_PHASES 1
  Calcite 0 0.1
```

- 输出结果：
 - 核素的化学形态、溶解度。
 - 矿物饱和指数（如方解石、铀酰碳酸盐沉淀）。

3. 使用 COMSOL 模拟核素迁移

- 任务描述：**将 PHREEQC 输出的化学结果（如吸附系数、溶解度）输入 COMSOL，用于模拟核素在含水层中的迁移。
- 任务目的：**生成包含多参数、多时间尺度的核素迁移行为数据。
- 模拟过程：**

- COMSOL 模型设置：
 - 几何与网格：
 - 定义三维岩层几何，包含核废物容器、膨润土屏障和地质层。
 - 物理场设置：
 - 流体流动（Darcy 模型）： $\nabla \cdot (\kappa \nabla P) = Q$
 - 溶质传输（对流-扩散方程）： $\frac{\partial C}{\partial t} + \nabla \cdot (-D \nabla C + v C) = R$
 - C ：核素浓度。
 - D ：扩散系数。
 - v ：地下水流速。
 - R ：化学反应速率（来自 PHREEQC 数据）。
 - 耦合化学反应：
 - 使用 PHREEQC 接口或定义局部反应速率。
- 参数化模拟：
 - 针对每个参数组合（如不同流速、吸附系数），运行迁移模拟。
 - 输出核素在不同时间节点的空间分布。

4. 深度学习模型的建立 为降低模拟成本，建立深度学习模型作为替代（代理模型），用于快速预测参数对迁移行为的影响。建立步骤如下：

- 数据准备：
 - 从 PHREEQC 和 COMSOL 模拟中提取关键数据：
 - 参数值（输入）：流速、Kd、矿物溶解速率、孔隙率等。
 - 结果（输出）：核素浓度随时间和空间的分布。
- 数据预处理：
 - 将数据归一化或标准化，以适应深度学习模型训练。
- 模型设计：
 - 使用神经网络建模，建议架构：
 - 输入层：参数（如流速、吸附系数）。
 - 隐藏层：多层全连接层（ReLU 激活）。
 - 输出层：核素浓度预测（可随时间/空间变化）。
- 示例结构：

```
Input: [流速, Kd, pH, 矿物溶解速率]
Hidden Layers: [128, 64, 32] (ReLU Activation)
Output: 核素浓度分布
```

- 模型训练：
 - 使用模拟数据训练模型，优化损失函数（如均方误差）。
 - 训练完成后验证模型预测精度。
- 模型部署：
 - 将模型嵌入分析流程，快速预测不同参数条件下的核素迁移。

5. 敏感性分析

- 分析目标：识别对核素迁移最敏感的参数。
- 分析方法：
 - 全局敏感性分析：
 - 使用 Sobol 指数或 Morris 方法，量化每个参数对输出的影响。
 - 局部敏感性分析：

- 固定其他参数，逐步调整单一参数，观察输出变化。

- 深度学习辅助：
 - 使用训练好的代理模型进行大规模参数扫描。
 - 高效计算每个参数变化对核素分布的影响。

6.1.3. 结果与应用

1. 识别关键参数：
 - 例如，渗透率、吸附系数可能对核素迁移范围和浓度峰值影响最大。
2. 优化设计：
 - 根据敏感性分析结果优化屏障设计（如调整膨润土渗透率、提高矿物吸附能力）。
3. 不确定性量化：
 - 使用代理模型快速评估参数不确定性对长期安全性的影响。

6.2. 应用示例 2：参数敏感性分析的详细过程

6.2.1. 研究目标

GeoChem AI 通过调用 PHREEQC 和 COMSOL Multiphysics 模拟核素和乏燃料的迁移与分布过程，通过 LSTM（长短期记忆网络）模型实现核素迁移行为的时间序列预测，结合监测数据动态调整策略以优化放射性废物地质处置的长期安全性。

6.2.2. 分析步骤

1. 初始数据准备

1. 地质条件定义
 - 岩石类型：如花岗岩、岩盐、或黏土岩层。
 - 地质参数：孔隙率、渗透率、裂隙分布。
 - 水化学环境：地下水流速、pH、氧化还原电位（Eh）、溶解矿物类型。
2. 核素定义
 - 核素种类：如铀（U）、铯（Cs）、钚（Pu）、锶（Sr）。
 - 核素释放源：乏燃料容器的腐蚀速率、核素释放速率。
3. 屏障条件
 - 工程屏障：膨润土层的渗透性、吸附性能。
 - 地质屏障：岩层对核素的阻隔能力。

2. PHREEQC 模拟：化学反应

- 目的：模拟核素与地下水和岩石之间的化学反应过程。

1. 输入文件设计

- 定义水化学初始条件：

```
SOLUTION 1
  pH 7.0
  pe 4.0
  U(6) 1e-6
  Na 10
  Cl 10
```

- 定义矿物平衡：

```
EQUILIBRIUM_PHASES 1
  Calcite 0 0.1
  Goethite 0 0.05
```

- 定义核素的吸附过程：

```
SURFACE 1
  Hfo_wOH
    sites 0.01
    -equilibration with solution 1
```

2. 输出结果

- 核素的化学形态、溶解度。
- 矿物与核素的吸附容量、饱和状态。
- 数据用于 COMSOL 模拟核素迁移。

3. COMSOL 模拟：核素迁移与分布

- 目的：模拟核素在含水层中的迁移行为，并耦合 PHREEQC 的化学反应结果。

1. 几何建模与物理场设置

- 几何模型：构建三维含水层结构，包括核废物容器、屏障层和周围地质。
- 流体流动：使用 Darcy 流动模拟地下水的对流。

- 溶质迁移：对流-扩散方程描述核素迁移：
 - 溶质传输（对流-扩散方程）：
$$\frac{\partial C}{\partial t} + \nabla \cdot (D \nabla C + v C) = R$$
- 化学反应耦合
 - 利用 COMSOL 的 Chemical Reaction Engineering Module，导入 PHREEQC 数据模拟吸附、溶解-沉淀反应的空间动态。
 - 定义化学反应速率为函数：
$$R = k_f \cdot (1 - S/S_{\text{sat}})$$
 - k_f ：反应速率常数。
 - S ：当前矿物饱和状态。
 - S_{sat} ：矿物饱和指数。
- 输出结果
 - 核素在不同时空的浓度分布。
 - 屏障层的核素累积情况。

4. LSTM 模型：时间序列预测

- 目的：利用 LSTM 对核素迁移进行时间序列预测，结合监测数据实时调整预测和策略。

- 数据收集与预处理
 - 数据来源：
 - PHREEQC 和 COMSOL 的模拟输出。
 - 实地监测数据（地下水核素浓度、流速、pH 等）。
 - 数据处理：
 - 时间序列构建：将模拟和监测数据组织为时间序列形式。
 - 归一化：对浓度、流速等数据进行归一化处理以提高模型训练效果。
- LSTM 模型结构
 - 输入层：包含时序数据特征（如时间、核素浓度、流速、pH）。
 - 隐藏层：
 - 多层 LSTM 单元，捕捉时间序列中的长期依赖关系。
 - Dropout 层防止过拟合。
 - 输出层：
 - 输出未来时刻的核素浓度分布。
- 模型架构示例：

```
model = Sequential()
model.add(LSTM(128, input_shape=(time_steps, features), return_sequences=True))
model.add(LSTM(64, return_sequences=False))
model.add(Dense(32, activation='relu'))
model.add(Dense(1, activation='linear'))
```

- 模型训练
 - 损失函数：均方误差 (MSE)。
 - 优化器：Adam 优化器。
 - 训练数据：使用 PHREEQC 和 COMSOL 模拟数据进行训练，监测数据验证。
- 模型预测
 - 输入最新的监测数据，预测未来核素浓度的时空分布。

5. 动态调整策略

- 实时监测与反馈
 - 收集实时监测数据（地下水核素浓度、流速等），更新 LSTM 模型输入。
 - 比较预测与监测结果，校正 PHREEQC 和 COMSOL 模型的初始条件。
- 调整模拟参数
 - 根据预测结果调整关键参数：
 - 流速变化：更新 COMSOL 的流体流动模型。
 - 化学参数：更新 PHREEQC 的矿物饱和状态和反应速率。
- 优化处置策略
 - 增强屏障性能：根据核素迁移预测结果调整膨润土厚度或材料成分。
 - 改进监测网络：增加监测井布设密度，提高数据精度。

6.2.3. 结果分析与决策支持

1. 核素迁移预测

- 确定核素可能的高风险区域及迁移时间。
- 提供屏障性能改进建议。

2. 实时优化策略

- 实现监测与预测的闭环反馈，动态调整处置策略。
- 提高地质处置系统的长期安全性。

3. 效率与成本优化

- 使用 LSTM 模型显著减少高成本的物理实验和全模型模拟运行次数。
- 实现精准的长期核素迁移预测和风险评估。

6.2.4. 效果总结

GeoChem AI 通过对 PHREEQC 和 COMSOL 的耦合模拟，提供高精度的核素迁移与分布数据；结合 LSTM 模型的时间序列预测能力，实现对迁移行为的实时分析与优化。整个过程充分结合监测数据与模拟结果，为放射性废物地质处置的长期安全提供科学依据和动态调整策略。

7. 大语言模型 (LLM) 在研究中的潜力

利用 LLM (Large Language Model)，如 Llama 3，通过设计高质量 Prompt，可以在地质处置研究中显著提升研究成果的效率和创新性。以下是结合 LLM 的应用场景、优点，以及如何通过 Prompt 设计推动研究的详细分析。

7.1. 大语言模型的潜力分析

1. 快速数据生成与分析：
- 自动生成核素化学反应输入（如 PHREEQC 文件）。
 - 提供优化后的多物理场模型设置建议（如 COMSOL 模拟参数）。
2. 复杂现象建模的灵感来源：
- 基于现有数据和背景知识，生成跨学科的建模假设和场景。
 - 提供针对核素迁移、屏障劣化等新视角和方案。
3. 自动化工作流与集成：
- 优化模型交互流程，例如将 PHREEQC 与 COMSOL 数据快速对接。
 - 生成敏感性分析脚本或序列预测模型的代码。
4. 解释与优化策略：
- 提供核素迁移行为的语义解释，帮助非专业领域合作伙伴理解研究结果。
 - 动态生成更优的实验与监测策略。

7.2. 设计 Prompt 的关键策略

1. 明确问题背景与上下文
- 提供清晰的地质背景、核素种类和目标。
 - 包括关键物理化学参数和研究限制条件。
 - 示例 Prompt：

目标：研究铀（U）在地下水系统中的迁移行为。
地质背景：岩盐储层，孔隙率为0.15，渗透率为 $1e-15\text{ m}^2$ ，地下水流速为 0.01 m/d 。
核素化学：主要考虑 UO_2^{2+} 的溶解与吸附。
请生成 PHREEQC 输入文件，考虑以下约束：

1. pH 范围为6.5-8.5。

2. 孔隙介质含有高岭石和赤铁矿。

3. 输出溶液中铀的浓度与矿物饱和指数。

2. 分步式问题拆解
- 指导模型生成逐步分析方案。
 - 如：PHREEQC 化学模拟 → COMSOL 流体场设置 → LSTM 模型预测。
 - 示例 Prompt：

我正在研究 Cs-137 在膨润土屏障中的迁移问题。
请分步骤指导我：

1. 如何用 PHREEQC 模拟 Cs-137 的吸附和溶解过程？

2. 如何在 COMSOL 中设置渗透率为 $1e-12\text{ m}^2$ 的膨润土屏障模型？

3. 如何结合 LSTM 模型预测迁移行为的长期趋势？

提供每一步的详细方法与输入示例。

3. 复杂现象的语义解读
- 让模型基于数据或模拟结果提供直观的解释。
 - 生成便于理解的结果报告或决策建议。
 - 示例 Prompt：

这是 PHREEQC 的输出结果：

- 溶液中 U(6) 的浓度为 $1.2e-7\text{ mol/L}$ 。

- 赤铁矿的饱和指数为-0.5。

- 方解石的饱和指数为0.3。

请根据这些结果解释：

1. 核素迁移的主要驱动因素是什么？

2. 方解石沉淀对铀迁移的影响。

3. 如果增加溶液的 pH，迁移行为可能如何变化？

4. 结合最新的 AI 技术提供代码生成
- 自动生成用于 PHREEQC 和 COMSOL 的集成代码，或 LSTM 模型的训练代码。
 - 示例 Prompt：

我需要 Python 脚本，整合 PHREEQC 和 COMSOL 的数据。

1. 调用 PHREEQC 的输出文件（如 “output.txt”），提取核素浓度。

2. 将浓度数据格式化并输入 COMSOL 模型。

3. 同时，基于时间序列训练一个 LSTM 模型预测浓度变化。

请生成完整的代码。

5. 多场景优化与敏感性分析
- 要求模型生成优化的实验设计或模拟参数组合。
 - 指导用户高效开展敏感性分析。

- 示例 Prompt :

为了研究地下水流速对 Pu-239 迁移的影响, 请帮助我:

1. 提供一个敏感性分析框架, 流速范围从0.001 m/d到0.01 m/d。
2. 针对每种流速条件生成 PHREEQC 输入文件示例。
3. 提供流速和浓度变化的可视化建议。

7.3. 结合 LLM 的研究优化点

1. 自动化建模与数据生成

- PHREEQC: 基于关键化学参数, 自动生成优化输入文件, 避免人工出错。
 - COMSOL: 生成优化的几何与物理场设置建议, 并快速对接化学结果。
 - LSTM 模型: 自动化生成高效、优化的时间序列预测代码。
2. 敏感性分析与策略优化 使用 LLM 提供的多场景模拟建议 (如 Sobol 或 Morris 方法) 提高分析效率。通过 LLM 动态生成输入文件, 快速迭代分析。
 3. 结果解读与沟通支持 通过自然语言生成数据报告, 直观展示核素迁移行为的核心影响因素。动态调整模型, 响应监测数据变化。

7.4. 应用实例: Prompt 驱动的监测与预测 workflow

1. Prompt 驱动输入

我需要构建一个研究 Pu-239 在地下水系统中的迁移预测 workflow:

1. 地质环境: 孔隙率 0.2, 渗透率 $1e-13$ m²。
2. 化学条件: pH 7.5, Eh 300 mV, 含赤铁矿。
3. 时间范围: 0 到 500 年。

请生成:

- PHREEQC 输入文件。
- COMSOL 的初始条件与流体场设置建议。
- 一个 LSTM 模型的训练代码框架。

2. Prompt 生成的具体输出

- PHREEQC 输入文件: 核素吸附、溶解-沉淀反应模型。
- COMSOL 建模建议: 包括几何建模、流体流动设置、反应动力学。
- LSTM 模型框架: 训练与验证代码, 包含数据预处理和可视化。

3. 动态调整策略

根据 100 年监测数据 (流速、浓度):

1. 生成一个新的 PHREEQC 输入文件, 更新初始化学条件。
2. 调整 LSTM 模型的输入, 重新预测 200 年后的浓度分布。
3. 提供一个优化的监测布点建议。

7.5. 效果评估

通过结合 Llama 3 等 LLM 的强大自然语言处理能力, 可以显著加速地质处置研究的分析与优化过程。得益于灵活的 Prompt 设计, 研究者能够快速生成高质量的建模方案、代码和优化策略。更重要的是, 这种方法不仅提高了研究效率, 还能通过持续迭代的动态调整, 提高模型预测的准确性, 为放射性废物处置的长期安全提供强有力的支持。

8. 结论与展望

GeoChem AI 采用了一种基于 PHREEQC 和 COMSOL 的协同研究方法, 并结合动态反馈机制和深度学习技术, 实现了复杂地质与化学现象的高效模拟与预测。以放射性废物处置为例, 验证了该方法在核素迁移模拟中的应用潜力。

GeoChem AI 的理论创新点和学术贡献主要体现在将多物理场建模、地球化学反应模拟和人工智能技术 (AI) 深度融合, 提供了一种全新的一体化研究框架。这种方法从根本上解决了传统地球化学模拟中计算效率低下、参数不确定性分析复杂和缺乏动态反馈机制的难题, 同时在解决当前热门的地球化学研究课题中实现了关键突破。

8.1. 方法创新分析

1. 将物理模型与数据驱动相结合

- 现状问题:
 - 地球化学和多物理场模拟依赖传统数值方法 (如有限元法), 计算成本高, 尤其在长时间尺度或大空间域模拟时效率低下。
 - 实验条件有限, 无法涵盖地质系统中所有复杂的参数组合。
- 创新点:
 - 本项目通过深度学习 (LSTM、DNN) 建立代理模型, 将物理模拟与数据驱动方法相结合:
 - 深度学习加速长时间尺度的模拟预测。
 - 提供动态的时间序列预测能力, 超越静态的数值模拟。
 - 突破传统物理模拟的局限, 实现实时预测与调整。

2. 动态反馈的闭环系统

- 现状问题:
 - 当前的模拟工具 (如 PHREEQC、COMSOL) 以静态建模为主, 缺乏与现场监测数据的实时交互。
- 创新点:

- 本平台通过 LLM 和深度学习技术，将现场监测数据动态整合到模拟流程中：
 - 结合实时数据，校正模型输入与参数。
 - 实现 "模拟-预测-反馈-优化" 的闭环工作流。
- 提供了地球化学过程动态分析的新视角。

3. 敏感性分析与不确定性量化的创新方法

- 现状问题：
 - 核素迁移和污染物扩散的敏感性分析耗时且缺乏系统化工具。
 - 不确定性分析多依赖于参数扫描，效率低。
- 创新点：
 - 利用深度学习（DNN）和代理建模快速实现敏感性分析，显著降低了计算成本。
 - 应用 Sobol 指数、贝叶斯方法等高级统计工具，量化参数对预测结果的影响。

4. 跨学科融合

- 现状问题：
 - 地球化学研究通常与工程、人工智能的结合较少，多依赖于领域专家手动调节参数和建模。
- 创新点：
 - GeoChem AI Suite 打破学科壁垒：
 - 通过 LLM（如 Llama 3）提供智能建模建议，降低专业门槛。
 - 集成了多学科方法论（地球化学、多物理场、AI），提出了全新的研究范式。

8.2. 地球化学理论问题分析

1. 核素迁移的时间与空间动态机制

- 研究长时间尺度下核素迁移的动力学规律。
- 定量评估化学反应（如吸附、溶解、沉淀）与物理场（如流体流动、热传递）的交互作用。

2. 地下水与屏障材料的相互作用

- 模拟地下水中化学组分与屏障材料（如膨润土、岩盐）的反应。
- 动态分析屏障劣化对核素迁移的影响。

3. 多参数不确定性在复杂地质系统中的传播效应

- 系统分析参数（如孔隙率、吸附系数、流速）变化对核素迁移的整体影响。
- 提供参数不确定性对预测结果的全局敏感性定量评估方法。

4. 动态反馈与实时优化的新范式

- 建立了一种实时动态耦合监测数据与模拟模型的方法。
- 为复杂地质系统的长期预测提供更高精度和动态调整能力。

8.3. 当前热门课题

1. 放射性废物地质处置

- 研究放射性核素（如铀、钚、铯）在地质屏障中的迁移与长期安全性。
- 应用案例：核能废物处置场设计、国际深地实验室研究。

2. 地下水污染物扩散

- 分析重金属（如铬、砷）或有机污染物在地下含水层中的迁移机制。
- 应用案例：污染场地修复、水资源保护。

3. 碳捕获与地质封存（CCS）

- 研究二氧化碳在深地储层中的溶解、扩散及矿化过程。
- 应用案例：温室气体减排与碳封存技术。

4. 矿床成因与资源预测

- 模拟热液系统与矿物沉淀的动力学过程。
- 应用案例：矿产资源的勘探与开发。

8.4. 相关领域软件差异与协同

在地球化学领域比较常用的科研分析软件有 TOUCH系列：Touch2 iTouch2 Touch3 与 Modelflow。TOUCH系列软件专注于模拟多相流体和热传导过程，广泛应用于地质储层研究、二氧化碳封存和核废物处置等领域。而Modelflow软件专注于单相地下水流动的数值模拟，主要用于水资源管理和污染物迁移分析。以下针对这两款软件与 GeoChem AI 的相关性展开分析。

1. GeoChem AI 平台与其他软件的相关性分析

• TOUGH2 和 TOUGH3

- 功能：
 - TOUGH2 和 TOUGH3 是用于模拟多相流体流动和热传导的工具，广泛用于地质储层研究（如二氧化碳封存、核废料处置）。
 - TOUGH3 相较于 TOUGH2，增加了更高的并行计算能力，支持复杂地质模型的模拟。
- 与 GeoChem AI 的相关性：
 - 相似点：
 - 都可以处理地下流体流动和热传递问题。
 - 涉及核废料处置中的地下水流动和热影响。
 - 差异点：
 - GeoChem AI 注重化学反应的模拟与动态反馈（如通过 PHREEQC 和深度学习的集成）。
 - TOUGH 系列更偏向于物理场流体和热力学模拟，缺乏对复杂化学反应的专注。
 - 协作潜力：
 - 可将 TOUGH 模型的流体流动结果导入 GeoChem AI，结合化学反应建模，模拟热-化学耦合效应。

• MODFLOW

- 功能：
 - 专注于地下水流动的数值模拟，应用于水资源管理和污染物迁移研究。
 - 模拟单相流动，适合低渗透性含水层环境。
- 与 GeoChem AI 的相关性：
 - 相似点：
 - 模拟含水层中的地下水流动和污染物扩散。
 - 可用于核素迁移与地下水污染的研究。
 - 差异点：
 - MODFLOW 更适用于纯物理流动场，而 GeoChem AI 强调化学过程与物理场的耦合。
 - GeoChem AI 结合 LSTM 等 AI 技术，可动态预测核素迁移行为，而 MODFLOW 偏静态分析。
 - 协作潜力：
 - MODFLOW 的流场计算结果可作为 GeoChem AI 的输入，用于模拟地下水与核素的化学交互。

2. 协作关系

- 差异化：
 - GeoChem AI 更注重化学反应的建模（PHREEQC 集成）和动态预测（深度学习结合）。
 - 它通过多工具协同和动态反馈机制，解决了传统工具的静态性问题。
- 互补性：
 - TOUGH 系列和 MODFLOW 的流体和热力场模拟结果可作为 GeoChem AI 的输入，用于深入分析化学反应和核素迁移。
 - GeoChem AI 的化学建模结果可以反过来优化 TOUGH 和 MODFLOW 的参数设置。
- 集成工具链：
 - GeoChem AI 可作为一个整合平台，将 MODFLOW 和 TOUGH3 的结果动态集成，实现多物理场与化学过程的全面模拟。

3. 总结

GeoChem AI 平台在功能定位上与 TOUGH 系列和 MODFLOW 既有竞争，也具备协作潜力。其优势在于动态反馈和多工具整合能力，能够有效补充传统软件在化学反应建模和动态预测中的不足。通过接口开发和联合模拟，GeoChem AI 平台可以成为多物理场与化学过程的协作中心，为地球化学和环境工程提供更加全面的解决方案。

8.5. 未来研究拓展领域

未来研究可进一步扩展至：

1. 不同地质条件下的动态行为预测。
2. 将强化学习用于优化监测与模拟策略。
3. 扩展至矿床成因分析和污染物迁移治理等领域。

参考文献

1. Parkhurst, D. L., & Appelo, C. A. J. (2013). Description of Input and Examples for PHREEQC Version 3. US Geological Survey Techniques and Methods.
2. COMSOL Inc. (2022). COMSOL Multiphysics® User Guide. COMSOL AB.
3. Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation, 55(1-3), 271-280.
4. Zhang, Y., & Zheng, C. (2013). Reactive transport modeling of concentrated solution using COMSOL and PHREEQC. Environmental Modelling & Software, 43, 38-50.
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
6. Section A, Groundwater. Description of Input and Examples for PHREEQC Version 3—A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations
7. Yoojin Jung, George Shu Heng Pau, Stefan Finsterle, Christine Doughty. TOUGH3 User's Guide Version 1.0