

利用生成对抗网络实现地热能开发中的储层温度预测优化

邓 翟

成都理工大学能源学院(页岩气现代产业学院), 四川 成都

收稿日期: 2024年4月27日; 录用日期: 2024年5月22日; 发布日期: 2024年5月31日

摘 要

地热是一种分布广泛、稳定、清洁的零碳排放能源, 而孔隙型热储是最重要的地热资源类型之一, 从对热储中提取出来的热量, 可用于地热发电站发电、供暖等。数值模拟可以用于评估地热自然状态的生产场景。前人在地热数值模拟上已经展开大量的研究。然而数值模拟通常需要大量的计算资源和时间。与数值模拟相比, 代理模型不仅具有与数值模拟近似的精度, 而且花费的时间更少, 这是显著优于数值模拟方法的。在本研究中, 基于GAN开发了一个三维代理模型来预测储层的温度分布, 模型以渗透率分布作为输入, 以时间作为条件, 最终输出储层不同时间的温度分布。最终的结果表明, 代理模型可以非常好地预测储层温度分布。

关键词

地热能开发, 生成对抗网络, 代理模型

Utilizing Generative Adversarial Networks to Optimize Reservoir Temperature Prediction in Geothermal Energy Development

Di Deng

College of Energy (College of Modern Shale Gas Industry), Chengdu University of Technology, Chengdu Sichuan

Received: Apr. 27th, 2024; accepted: May. 22nd, 2024; published: May. 31st, 2024

Abstract

Geothermal energy is a widely distributed, stable and clean zero-carbon emission energy source,

among which pore heat storage is one of the most important types of geothermal resources, and the heat extracted from thermal storage can be used for power generation, heating and cooling in geothermal power stations. Numerical simulations can be used to evaluate production scenarios in the natural state of geothermal energy. A large number of studies have been carried out on geothermal numerical simulation. However, numerical simulations often require a lot of computational resources and time. Compared with numerical simulation, the proxy model not only has the accuracy of the same as the numerical simulation, but also takes less time, which is significantly better than the numerical simulation method. Based on GAN, a three-dimensional proxy model is developed to predict the temperature distribution of the reservoir, and the model takes the permeability distribution as the input and the time as the condition, and finally outputs the temperature distribution of the reservoir at different times. The final results show that the proxy model can predict the reservoir temperature distribution very well.

Keywords

Geothermal Energy Development, Generative Adversarial Network, Proxy Model

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着全球对清洁可再生能源需求的不断增加,地热能因其储量丰富、稳定可靠等特点在全球能源转型中扮演着重要角色[1]。我国地热资源分布广泛,尤其在大型沉积盆地中储量丰富[2] [3] [4]。其中,孔隙型热储以其渗透性好、非均质性弱等特点,成为地热开发的理想选择[5]。CO₂循环采热是一种高效的地热开发方法,与传统的注水开采相比,CO₂在超临界状态下具有更高的流动性和采热效率,同时还能在一定程度上实现CO₂的地质封存[6] [7]。然而,由于孔隙型热储CO₂循环采热过程涉及复杂的多场耦合过程,传统数值模拟方法模拟CO₂循环采热过程计算量大、耗时长。近年来,国内外学者针对CO₂循环采热开展了大量的数值模拟研究。通过建立热流耦合数值模拟模型研究了地质控制因素对地热流体储层CO₂循环采热影响[8]。建立三维的热流耦合模型模拟了CO₂在EGS中的热提取以及CO₂封存[9]。但数值模拟建模复杂,且具有复杂性和非线性的特征,计算时通常会消耗大量的时间和计算资源[10] [11]。而基于深度学习方法的代理模型因为不依赖数学公式或物理模型并可以形成具有高度预测性能模型得到了迅猛发展。

代理模型的基本思想是用一个简单的模型来近似代替一个复杂的模型,以实现更高效的计算和预测。随着计算机性能的进步,越来越多的代理模型技术在地热领域得到了发展。生成式对抗网络是最流行的深度学习框架之一,它是由Ian J. Goodfellow等人提出的[12],生成式对抗网络是一个通过对抗过程训练模型的框架。在本研究中,以渭河盆地西安凹陷高陵群这一典型的孔隙型热储为例,使用序贯高斯模拟方法生成随机渗透率场,基于生成对抗网络(GAN)开发了一种模拟注入CO₂循环采热过程热储三维温度演化的代理模型,代理模型通过输入渗透率场和时间输出热储温度分布,GAN可以学习到输入与输出之间的映射关系达到快速预测热储温度分布的目的。

2. 数据集生成

本研究基于渭河盆地西安凹陷高陵群的储层物性建立地质模型。首先采用序贯高斯方法生成多个符

合高陵群地质统计特征的渗透率分布，具体参数设定见表 1。然后根据固有的渗透率推算地层岩石的孔隙度，应用霍尔兹相关性[13]来关联孔隙度和渗透率，可用以下等式表示：

$$\varphi = \left(\frac{k}{7 \times 10^7} \right)^{\frac{1}{9.61}} \quad (1)$$

其中 k 为岩石的固有渗透率， φ 为储层岩石的孔隙度。为通过该方法随机生成的渗透率分布。

Table 1. Geological statistical parameters

表 1. 地质统计参数

参数	值
X 方向相关长度	250
Y 方向相关长度	250
平均渗透率	100
标准对数渗透率	1
区域范围 m	640 m × 640 m × 64 m
X 方向网格	64
Y 方向网格	64
Z 方向网格	16
网格大小	10 m × 10 m × 4 m

根据以上参数使用序贯高斯模拟生成 800 个渗透率分布，图 1 为其中一个生成结果。

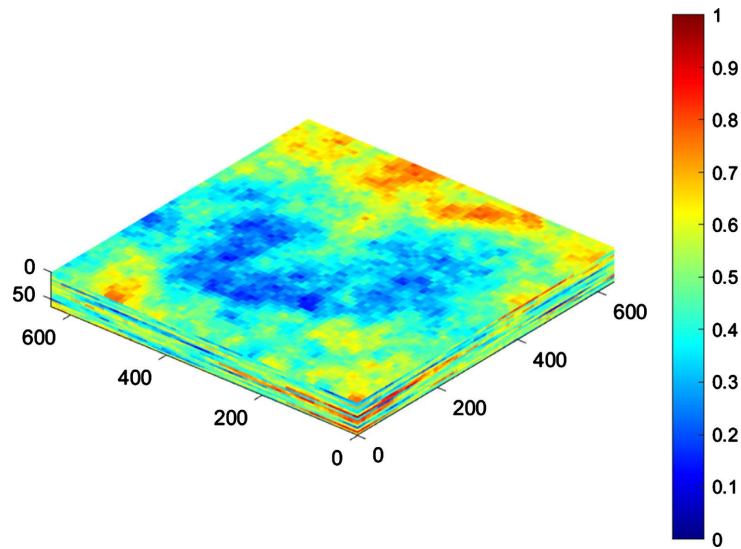


Figure 1. Permeability distribution generated by the sequential Gaussian simulation method

图 1. 序贯高斯数值模拟生成渗透率场

以渭河盆地西安凹陷高陵群储层物性为例，气水两相渗流传热数值模拟热储参数设定如表 2 所示。在三维数值模拟模型中，网格数量为 $64 \times 64 \times 16$ ，每一个网格大小为 $10 \text{ m} \times 10 \text{ m} \times 4 \text{ m}$ 。地热开采采用两口垂直井，注入井位于热储剖面网格的(7, 57)，生产井位于(57, 7)。CO₂ 在 37 MPa 的恒定井下压力约

束下注入，注入的 CO_2 温度为 40°C ， CO_2 循环采热过程持续 30 年。描述不同组分之间的质量守恒方程与 CO_2 -水两相渗流的能量守恒方程通过有限差分对方程进行离散，在每个时间步，离散后的非线性代数方程组通过牛顿迭代法进行求解。

Table 2. Numerical simulation model parameter settings

表 2. 数值模拟模型参数设置

模型参数	值
深度	3500 m
区域	640 m × 640 m × 64 m
网格大小	10 m × 10 m × 4 m
渗透率分布	Log-normal
水平方向相关长度 x	250 m
垂直方向相关长度 z	100 m
平均孔隙度	0.246
平均渗透率	100 mD
初始水饱和度	1.0
初始压力	35,000 kPa
初始储层温度	150°C
压缩系数	$1.0 \times 10^{-7} \text{ kPa}^{-1}$
水的热传导系数	$0.63 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
CO_2 热传导系数	$0.037 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
岩石热传导系数	$1.74 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
岩石比热容	$2.45 \times 10^6 \text{ J}\cdot\text{m}^{-3}\cdot\text{K}^{-1}$

为了生成用于代理模型训练和测试的储层温度分布图，需要保存 7 个时间步长的输出，首先以 1 年为间隔，保存 1 到 10 年的时间段，然后以 10 年为间隔，保存第 10 年到 50 年的储层温度分布。

数值模拟计算完成后需要对其进行预处理。预处理主要是对数据进行归一化，将所有数据缩放到一个统一的范围。归一化处理能够帮助代理模型更快地学习和收敛，减少训练时间，并防止训练过程中的梯度消失或爆炸问题。训练数据包含渗透率场、时间信息、不同时刻下渗透率场对应的温度分布数据。

归一化公式为：

$$x = \log \frac{k - k_{\min}}{k_{\max} - k_{\min}} \quad (2)$$

其中 k 为数据， k_{\min} 为每一类数据的最小值， k_{\max} 为每一类数据的最大值， x 为处理后的数据。

数据归一化后，将渗透率场与时间信息进行拼接。处理结束后，将数据的 80% 作为训练集，20% 作为测试集。

3. 代理模型构建及模型评价

3.1. 代理模型构建

生成对抗网络由生成器和判别器组成，生成器负责生成逼真的储层温度分布，而判别器则负责判断

输入的样本是真实的(来自真实数据分布)还是生成的(来自生成器) [12]。在构建模型时, 需要设计生成器和判别器的网络结构。在本研究中, 该模型由生成器(Generator)和判别器(Discriminator)两部分组成, 如图 2 所示。生成器采用 U-net 结构, 由下采样编码器和上采样解码器构成, 生成器的输入输出均为 $64 \times 64 \times 16$ 。下采样编码器由 7 个串联的卷积块组成, 每个卷积块包含卷积层、激活函数层和批标准化层, 通过逐步缩小特征图的空间分辨率和增加通道数, 提取输入数据的局部特征。上采样解码器由 7 个串联的反卷积块组成, 通过逐层上采样和通道数减少, 将特征图的空间分辨率恢复到与输入图像相同, 并引入随机丢弃正则化以减少过拟合风险。在 U-net 架构中, 上采样过程中的每一层特征图与下采样过程中对应层的特征图按通道维度拼接, 以传递不同尺度的语义信息。判别器采用 Patch GAN 结构, 由 5 个卷积块组成, 对输入的热储温度分布图像的局部区域进行真假判断, 最终通过平均池化得到全局的真假判别结果。这种设计可以显著减少参数量和计算消耗, 提供局部精细的梯度信号, 鼓励生成器合成逼真细节。

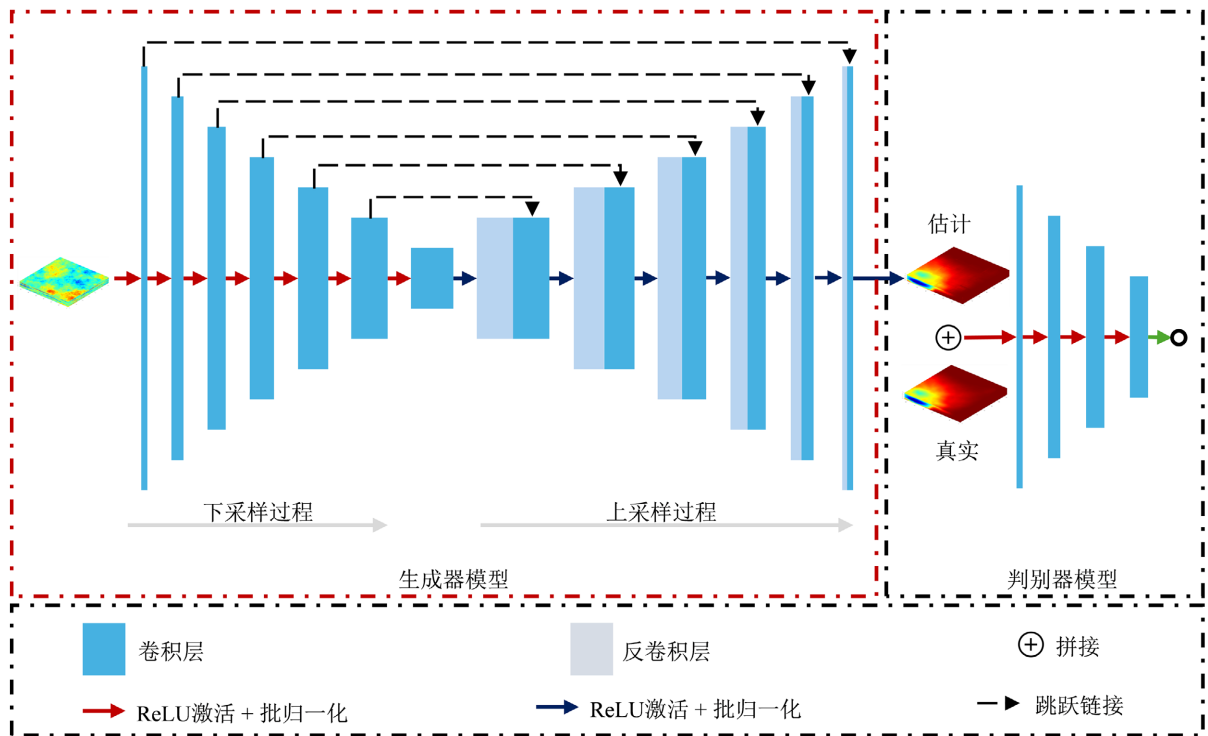


Figure 2. Proxy model framework
图 2. 代理模型框架设计

生成对抗网络模型的损失函数由两部分组成: 生成器损失函数采用二元交叉熵和 L1 损失函数相组合, 判别器损失函数采用二元交叉熵。在训练过程中将迭代地更新生成器和判别器的参数以使两个损失函数同时达到最优, 训练过程中的每一步都会根据实际数据进行参数更新, 直到模型收敛或达到停止条件为止。生成器损失函数 L_G 可以表示为:

$$L_G = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) + \lambda \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (3)$$

其中, 前半部分为二元交叉熵, y 是二元标签 0 或者 1, $p(y)$ 是属于 y 标签的概率, 二元交叉熵是用来评判一个二分类模型预测结果的好坏程度的, 通俗的讲, 即对于标签 y 为 1 的情况, 如果预测值 $p(y)$ 趋近于 1, 那么损失函数的值应该趋近于 0。反之, 如果此时预测值 $p(y)$ 趋近于 0, 那么损失函数的值

应该非常大。后半部分为 L1 损失函数，通过一个权重 λ 控制两部分的占比。

判别器损失函数 L_D 可以表示为：

$$BCELoss = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (4)$$

3.2. 模型评价标准

在本研究中，使用了两种方法评价模型性能，分别是结构相似性指数、均方根误差。SSIM 算法分别从亮度、对比度、结构三方面度量两幅图像相似性，其值越接近于 1 越相似，越接近于 0 则越不相似[14]。SSIM 定义如下：

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (5)$$

其中 μ 表示给定图像的平均值， σ 表示给定图像的标准差， x 和 y 是被比较的两个图像， C_1 与 C_2 为常数，保证分母为 0 时的稳定性。

RMSE 通过计算预测值和实际值之间距离(即误差)的平方后，开方来衡量模型优劣。即预测值和真实值越接近，两者的均方差就越小。RMSE 可以表示为：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (6)$$

其中， N 为样本数量， y_i 为真实图像， \hat{y}_i 为生成器模型生成的图像。

4. 结果分析

本研究采用气水两相渗流数值模拟，生成了 904 组 3 维图像数据集。每组数据包含从第 1 年至第 30 年(间隔 5 年)，共 7 个时间节点的储层温度分布情况。为了深入分析和验证所开发的深度学习代理模型的有效性，训练集和测试集是完全独立的，确保了评估的公正性和准确性。图 3 展示了训练集中的一组示例数据。

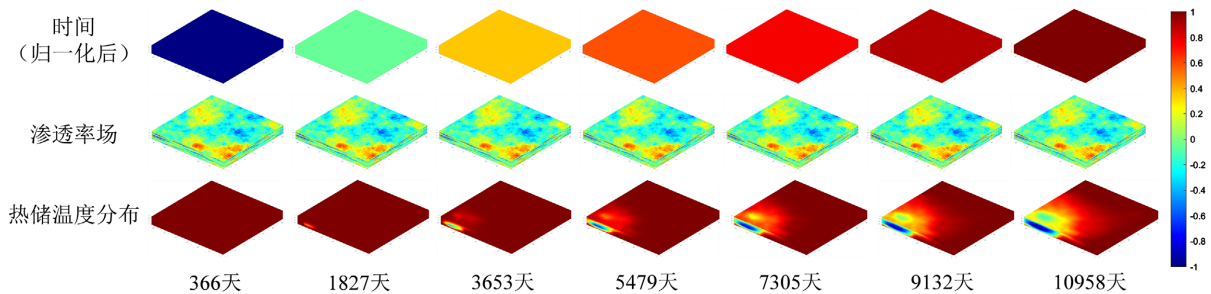


Figure 3. Training data

图 3. 训练数据

在训练完成后，使用训练时未被使用的数据测试代理模型。预测了注入 CO_2 采热后第 366 天、1827 天、3653 天、5479 天、7305 天、9132 天、10,958 天的储层温度分布情况，代理模型预测结果如图 4 所示。本节中使用 SSIM 与 RMSE 两种方法计算代理模型预测结果与数值模拟结果之间的误差。如图 5 所示，SSIM 最小值为 0.97，RSME 最大值小于 0.05，说明代理模型能够复现非均质热储 CO_2 循环采热过程中温度场时空演化规律。

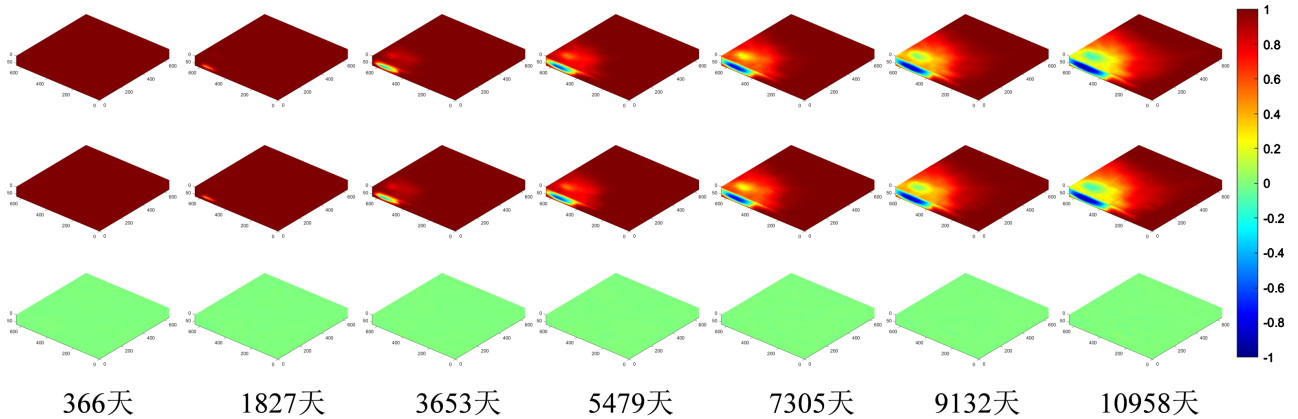


Figure 4. Dynamic temperature distribution during heat extraction process (numerical simulation results, proxy model predictions, residual plot)

图 4. 采热过程动态温度分布(数值模拟计算结果、代理模型预测结果、残差图)

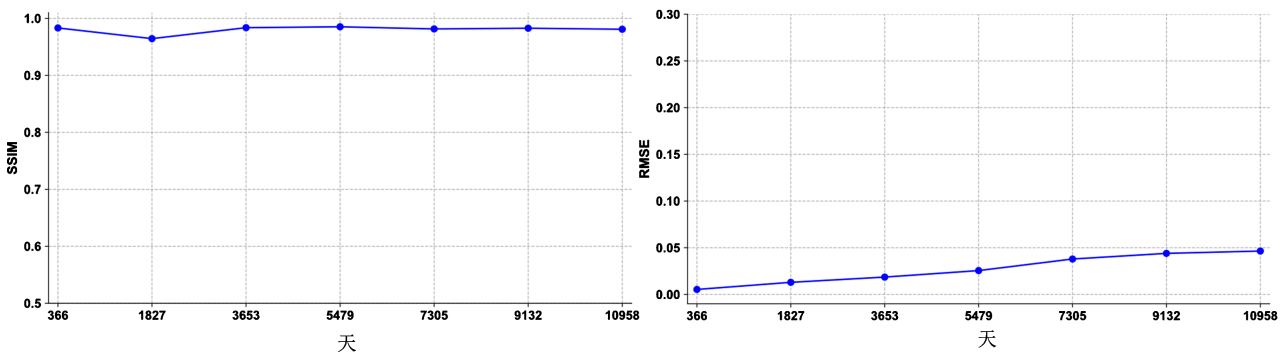


Figure 5. SSIM and RMSE comparison between proxy model results and numerical simulation results

图 5. 代理模型结果与数值模拟结果 SSIM 与 RMSE

本研究采用测试集中的所有样本，对训练后的代理模型进行了系统测试，并使用结构相似性指数(SSIM)和均方根误差(RMSE)两个定量指标对代理模型的预测结果与数值模拟结果进行了比对分析。图6展示了深度学习代理模型在整个测试集预测结果与数值模拟计算结果的SSIM分布情况。统计发现，所有测试样本的SSIM值均具有较高的数值，最大值高达99.74%，最小值也达到78%，平均SSIM指数更是高达97%，表明代理模型生成的三维温度场在结构纹理上与数值模拟结果具有极高的相似度，验证了代理模型优异的泛化预测能力。同时，本研究计算了整个测试集预测结果的均方根误差分布，如图6所示，可以看出，所有测试样本的RMSE值分布较为集中，最大值为0.2，最小值为0.002，整体平均RMSE仅为0.04。这些定量结果再次证实，代理模型能够以较小的预测误差准确预测孔隙型热储CO₂循环采热过程的三维热储温度演化，具有良好的泛化性能。

为了比较三维深度学习代理模型与气水两相渗流传热数值模拟在计算效率上的差异，本研究对两种方法的计算耗时进行了定量分析，如图7所示。在代理模型训练之前，我们首先使用数值模拟方法生成了766组结果作为训练集数据，每个数值模拟过程平均耗时约20分钟。由于训练集的生成是代理模型训练的前提，因此在模型训练开始前，代理模型与数值模拟在时间成本上并无差异，累计耗时相同。三维代理模型训练需要花费45个小时，这相当于数值模拟计算135个模型。当代理模型训练结束后，代理模型预测热储温度仅需要不到1秒，而数值模拟模型依然需要20分钟，因此，如果需要模拟超过135个，使用本代理模型将会是更优的选择。

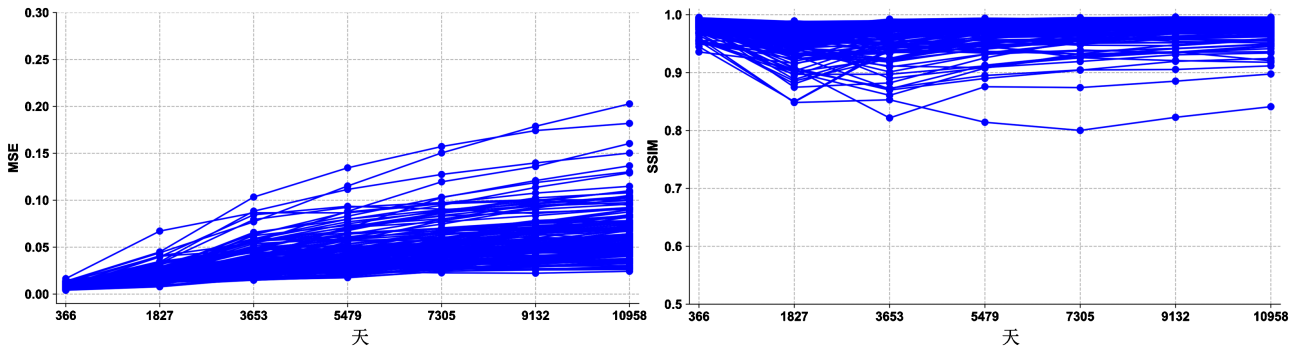


Figure 6. Overall test set SSIM and RMSE comparison between proxy model predictions and numerical simulation results
图 6. 整体测试集代理模型预测结果与数值模拟结果 SSIM 与 RMSE

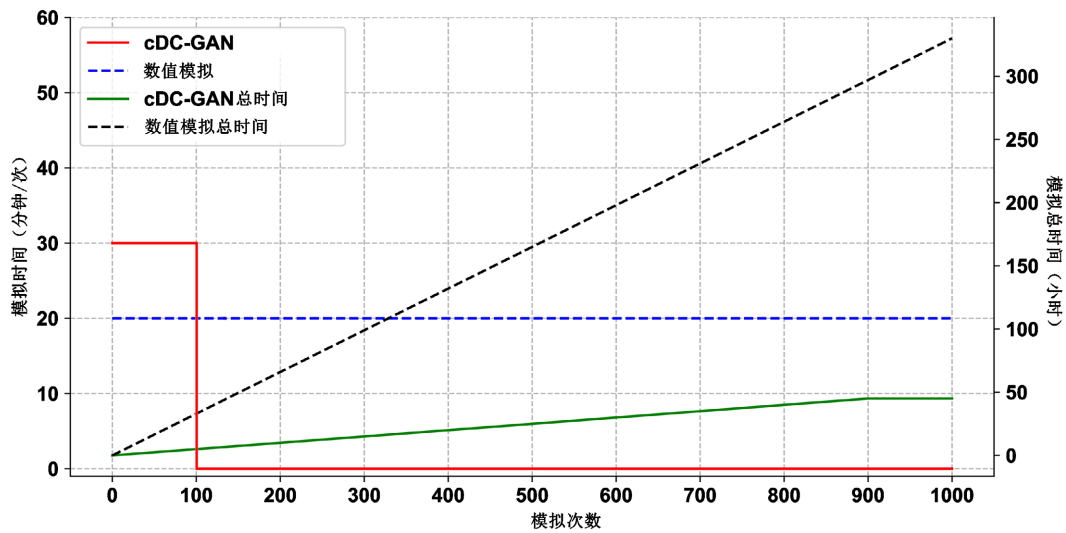


Figure 7. Computational time of proxy model versus numerical simulation time
图 7. 代理模型计算时间与数值模拟计算时间

5. 结论

基于 GAN，设计了三维孔隙型热储 CO₂ 循环采热过程的深度学习代理模型，代理模型预测结果与数值模拟结果的 SSIM 平均值为 97%，RMSE 平均值为 0.002，说明代理模型能够有效捕捉流体循环采热过程中储层温度场的动态演化，验证了三维代理模型在模拟 CO₂ 循环采热热储温度分布的可行性。

通过生成器和判别器的对抗学习，代理模型能够自动提取渗透率场和温度场之间的高维非线性映射关系，生成与数值模拟结果高度一致的三维温度分布。同时，代理模型的泛化能力得到验证，对于未参与训练的新渗透率场，模型依然表现出了优异的预测性能，这为代理模型在实际孔隙型热储 CO₂ 循环采热中的应用提供了可能。

与传统的数值模拟方法相比，训练完成的三维代理模型在预测效率上具有明显优势。数值模拟求解耗时 15~30 min，而代理模型的预测时间小于 1 s，速度呈数量级的提升。高效的代理模型使得实时热储动态分析、不确定性定量评估、参数优化等任务成为可能，为孔隙型热储开发决策提供强有力的工具。

综上所述，本研究提出了一种基于生成对抗网络的储层温度预测优化方法，旨在解决地热能开发中的储层温度预测问题，通过构建代理模型并进行训练，利用 GAN 学习渗透率场与储层温度分布之间的复杂关系实现了对储层温度的高效预测。实验结果表明，地热能开发领域具有广阔的应用前景，能够有效

地指导地热能资源的开发和利用, 提高地热能开发的效率和经济效益。

参考文献

- [1] 汪集昉, 胡圣标, 庞忠和, 等. 中国大陆干热岩地热资源潜力评估[J]. 科技导报, 2012, 30(32): 25-31.
- [2] 陈墨香, 邓孝. 中国地下热水分布之特点及属性[J]. 第四纪研究, 1996(2): 131-138.
- [3] 张薇, 王贵玲, 刘峰, 等. 中国沉积盆地型地热资源特征[J]. 中国地质, 2019, 46(2): 255-268.
- [4] 蔺文静, 王贵玲, 邵景力, 等. 我国干热岩资源分布及勘探: 进展与启示[J]. 地质学报, 2021, 95(5): 1366-1381.
- [5] 柯婷婷, 黄少鹏, 许威, 等. 关中盆地沔西地区地热对井采灌开发模式的数值模拟[J]. 第四纪研究, 2019, 39(5): 1252-1263.
- [6] Levy, E.K., Wang, X., Pan, C., *et al.* (2018) Use of Hot Supercritical CO₂ Produced from a Geothermal Reservoir to Generate Electric Power in a Gas Turbine Power Generation System. *Journal of CO₂ Utilization*, **23**, 20-28. <https://doi.org/10.1016/j.jcou.2017.11.001>
- [7] Gan, Q., Candela, T., Wassing, B., *et al.* (2021) The Use of Supercritical CO₂ in Deep Geothermal Reservoirs as a Working Fluid: Insights from Coupled THMC Modeling. *International Journal of Rock Mechanics and Mining Sciences*, **147**, Article ID: 104872. <https://doi.org/10.1016/j.ijrmmms.2021.104872>
- [8] Wang, Y., Wang, X., Xu, H., *et al.* (2022) Numerical Investigation of the Influences of Geological Controlling Factors on Heat Extraction from Hydrothermal Reservoirs by CO₂ Recycling. *Energy*, **252**, Article ID: 124026. <https://doi.org/10.1016/j.energy.2022.124026>
- [9] Wang, C.-L., Cheng, W.-L., Nian, Y.-L., *et al.* (2018) Simulation of Heat Extraction from CO₂-Based Enhanced Geothermal Systems Considering CO₂ Sequestration. *Energy*, **142**, 157-167. <https://doi.org/10.1016/j.energy.2017.09.139>
- [10] Bahrami, P., Sahari, M.F. and James, L.A. (2022) A Review of Proxy Modeling Highlighting Applications for Reservoir Engineering. *Energies, Multidisciplinary Digital Publishing Institute*, **15**, Article 5247. <https://doi.org/10.3390/en15145247>
- [11] Jaber, A.K., Al-Jawad, S.N. and Alhuraishawy, A.K. (2019) A Review of Proxy Modeling Applications in Numerical Reservoir Simulation. *Arabian Journal of Geosciences*, **12**, Article 701. <https://doi.org/10.1007/s12517-019-4891-1>
- [12] Goodfellow, I., Pouget-Abadie, J., Mirza, M., *et al.* (2014) Generative Adversarial Nets. https://proceedings.neurips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afcf3-Abstract.html
- [13] Holtz, M.H. (2002) Residual Gas Saturation to Aquifer Influx: A Calculation Method for 3-D Computer Reservoir Model Construction. *SPE Gas Technology Symposium*, Calgary, 30 April 2002, 1-10. <https://doi.org/10.2118/75502-MS>
- [14] Wang, Z., Bovik, A.C., Sheikh, H.R., *et al.* (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**, 600-612. <https://doi.org/10.1109/TIP.2003.819861>