基于物理的递减曲线分析与机器学习在增强型地热系统温度预测中的应用：犹他州 FORGE

Mina S. Khalaf

美国得克萨斯州休斯敦大学

摘要

增强型地热系统（EGS）的可靠温度预测对于储层设计和经济评估至关重要，但石油行业递减曲线以及许多机器学习（ML）代理模型并未对地热传热过程施加约束。此外，高保真热-水-力（THM）模拟的计算成本仍然很高。

本研究提出了一个统一且符合物理规律的框架，推进了地热温度预测中的递减曲线分析（DCA）和代理建模。通过引入受牛顿冷却定律启发的平衡温度项，将经典 Arps 递减族推广到地热应用中，从而确保温度在晚期达到有限极限；当平衡温度项设为零时，该模型可精确退化为传统 Arps 形式。扩展递减曲线首先通过犹他州 FORGE 的井下温度测量进行验证，随后用于在受控热-水-力数据集上构建学习代理模型，该数据集涵盖裂缝数量、井距、裂缝间距、围岩热导率和循环速率。方程信息神经网络将修正递减方程作为可微分的内部计算层嵌入，在保留可解释递减结构的同时，根据设计与运行输入生成完整的 0–60 个月温度轨迹。此外，还构建了概率高斯过程回归代理模型，用于直接进行多时域预测并提供校准不确定性；同时，以直接 XGBoost回归基线作为纯数据驱动的参照。在整个模拟数据集上，扩展递减模型以近乎完美的保真度重现了温度轨迹（决定系数中位数 =0.999；均方根误差中位数 = 0.071 °C），方程信息网络的典型留出误差为平均绝对误差= 3.06 °C 和均方根误差 = 4.49 °C。高斯过程代理模型在 3–60 个月预测时域内表现出最强的预测准确度（宏观决定系数 = 0.965；均方根误差 = 3.39 °C；平均绝对误差 = 2.34 °C），并具有良好校准的不确定性；相比之下，XGBoost 基线的误差更高，凸显了物理结构或概率建模对于可靠地热温度预测的价值。

关键词： 增强型地热系统（EGS）；物理信息机器学习（PIML）；递减曲线分析（DCA）；方程知情神经网络（EINN）；热-水-力（THM）建模；犹他州 FORGE。

1. 引言

地热能可持续提供来自地球内部的热量，从而实现低碳发电和直接供热。由于其运行不受天气和时间影响，能够提供可靠的基荷电力，因此有别于风能和太阳能资源。近几十年来，科研界和产业界一直致力于扩大地热能的部署，并将其应用范围拓展至天然具备有利水热系统以外的地区。全球评估显示，地热能正在逐步增长，勘探和钻井技术不断改进，其在发电和供热方面的应用也日益广泛。然而，地热能目前仍仅占全球发电量的一小部分，这表明如果能够克服技术、经济和环境方面的挑战，其潜力仍有很大一部分尚未开发 [1–4]。

增强型地热系统（EGS）将地热能的利用扩展到缺乏天然流体流动的干热岩（HDR）。增强型地热系统通过开采储存在干热岩中的热能来提升井的性能。这类系统依靠钻井和井刺激，形成流体流动的闭环。冷流体被泵入储层，流经受刺激的岩层并逐渐升温，随后返回地表。在地表，换热器捕获热能。图1展示了这一增强型地热系统的配置，包括循环回路和地表热能回收系统。关于增强型地热系统项目和技术的全球综述指出，场地条件至关重要，尤其是构造环境和应力状态；要维持生产，还需要谨慎开展井刺激和储层管理。这些综述还梳理了开发策略和研究方向，涵盖从勘探数据库建设到通过改进建模和维护方法降低成本等内容 [3,5]。

地热领域的产量预测通常会借鉴石油行业的递减曲线分析（DCA），采用传统的 Arps方程 [6]（指数型、调和型、双曲型）。然而，地热系统的行为可能有所不同，因为温度控制的热交换、注采耦合和裂缝网络演化，都会使其递减特征有别于以压力枯竭为主导的油气系统。案例研究呈现出不一的结果：The Geysers 的趋势接近调和型，且受井间距和连通效应影响；在 Geysers，采用另一种基于物理的 DCA 可提高后期预测的可靠性 [7,8]；而在 Ulubelu 等一些液体主导型地热田，指数型拟合也能奏效 [9],。这些结果凸显了针对地热数据和物理机制验证各种衰减形式的必要性。近年来，地热领域的代理模型和机器学习（ML）取得进展，与人工曲线拟合相比，它们能更快地给出通常也更准确的预测；若将方程结构嵌入模型，还能保留可解释性。相关示例包括用于优化、受方程约束的热递减代理模型，模拟水热时间序列的序列模型，以及在给出预测的同时量化不确定性的概率回归器 [10]。将递减分析与现代机器学习相结合，有助于决策者在可靠量化不确定性的基础上预测温度。典型曲线分析可提供兼顾物理机制的诊断依据和衰减曲线的约束，避免对后期趋势过度乐观；

机器学习可以通过学习不同情景中的模式并揭示不确定性来补充这一方法。此外，完全耦合的热-水-力耦合模拟计算负担沉重，因为需要同时求解非线性热传输、流体流动和岩石变形方程。这种耦合通常需要在标准工作站上运行数小时 [11–13]。

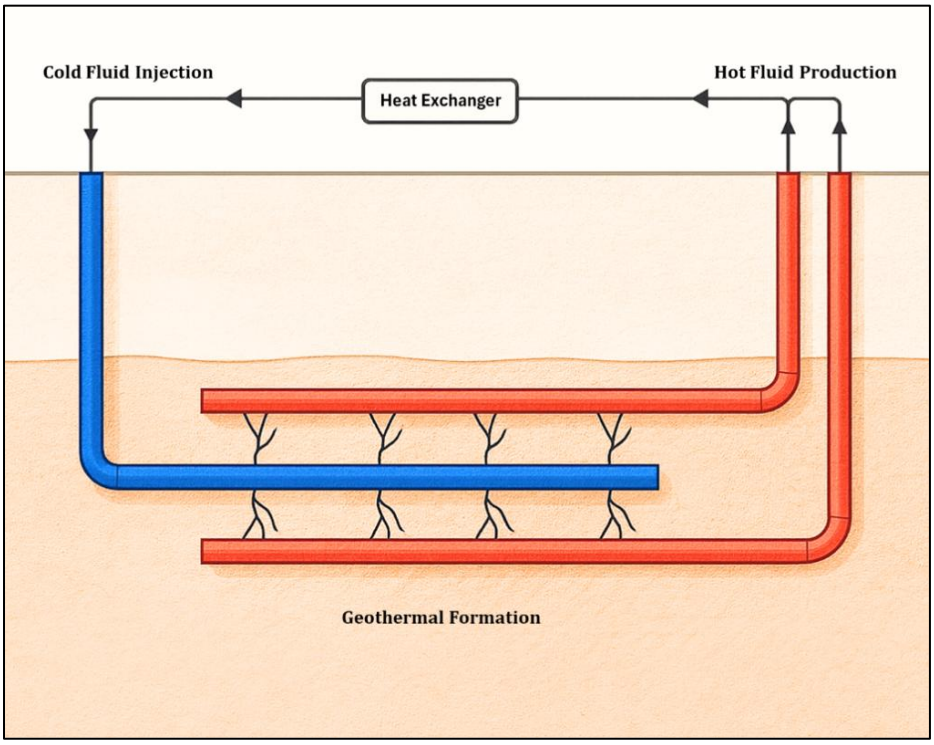


图*1*。 展示增强型地热系统的示意图，其中水平注入井和生产井通过水力裂缝相连。

本研究通过引入基于牛顿冷却定律的平衡温度扩展，改进了适用于地热应用的 Arps方程。研究使用来自犹他州 FORGE 项目的现场温度数据验证了地热扩展型递减曲线分析模型。此外，研究还将嵌入方程的神经网络用于地热温度递减分析，直接将物理定律嵌入模型，从而提高预测的可靠性。此外，研究采用高斯过程回归代理模型，实现快速且不确定性感知的预测。通过在单一热-水-力（THM）数据集上统一这些方法，本研究提出了一种连贯且符合物理规律的方法，从而减少对计算量大的模拟的需求。

2. 文献综述

增强型地热系统旨在通过在井间创建或改善流体通道，将高温但低渗透的岩石转变为可靠的换热器。近期综述指出，相关技术从

勘探到运行均稳步进展，同时也指出，结果在很大程度上取决于应力状态、增产策略和长期热管理 [1,3]。研究表明，增产通常涉及多种机制：天然裂缝发生滑移，同时伴随有限的张开和新通道的形成，这有助于解释流动局部化和压力受限行为等常见观测值 [5]。系统级评估认为，全球技术潜力巨大，但受场地适用性和经济运行条件下热可再生性限制的制约，这凸显了谨慎进行储层设计和井间距规划的必要性 [4]。井位布局和分段增产方面的进展，为在多井水平井配置中提高热提取量提供了切实可行的途径；在这种配置中，循环策略、裂缝特性以及生产井与注入井间距会显著影响温度衰减 [14]。除水之外，使用二氧化碳作为工作流体也是一种颇具吸引力的构想，因为浮力和膨胀效应能够增强自然循环、降低寄生运行负荷，并在压降相当的情况下提高净热提取量，不过化学相互作用仍是一个关键的不确定因素 [15]。

传统递减曲线分析广泛用于概括油气田的产量趋势并进行预测。Arps [6] 通过考察产量、时间、累积产量和递减百分比之间的数学关系，分析了指数型、调和型和双曲型递减模型。该研究还评估了损失比率法等统计工具，介绍了图解外推技术，并提出递减图表，以便进行直线预测 [6]。研究发现，递减曲线分析，尤其是双曲线形式，为石油工程师提供了一种实用且可靠的工具，可在产能生产条件下估算储量和未来单井产能 [6]。随后，Fetkovich 将这些经验形式与定压解和双对数型曲线联系起来，说明其具有坚实的物理基础，使递减分析成为一种还可用于诊断流动机理、估算渗透率、表皮系数和泄油面积，并在对产量进行适当归一化后处理回压变化的工具 [16,17]。随着生产转向致密气和页岩储层，针对水平多段压裂井的研究表明，在长期瞬态流期间应用传统的单段 Arps 拟合可能会误估储量；诊断优先工作流程（使用 β 导数图、归一化产率图等工具，并明确识别流态）应作为基于模型的生产分析的基础，再通过多种产量—时间递减关系（指数型、双曲型、拉伸指数型、幂律指数、Duong、逻辑斯蒂增长及相关形式）进行交叉验证，以界定随时间变化的储量范围 [18–21]。递减曲线分析在油气领域的现场应用包括鹰滩、巴肯、尼奥布拉拉和中东致密气 [22–25]。

在地热环境中，案例研究显示出多种行为特征。在一个大型蒸汽主导型地热田，归一化分析和型曲线拟合表明，许多井的趋势接近调和形式，而井间距和边缘效应则影响着各井的具体轨迹 [7]。人们提出了基于物理的改进，旨在

通过将产量率与累积产量的倒数关联起来，改善后期预测，从而得到线性趋势以及可解释的储层常数，这些常数与毛细力和重力有关 [8]。在液体主导型地热田中的应用表明，当产量递减平缓且回灌与补充钻井得到协调管理时，简单的指数形式有时能很好地拟合数据，并有助于实际规划；而对于其他地热储层，指数形式或双曲线形式也足以用于长期趋势分析 [7,9]。

机器学习和代理建模为模拟耦合热–水力行为以及加速设计迭代提供了高数据效率的方法。结合循环层和卷积层的时间序列代理模型已被用于预测压力–温度演化、支持井对布置，并降低寻找经济最优解的搜索成本 [10]。此外，近期研究提出了一种快速且高保真的方法，用于模拟裂缝型地热储层中的长期温度演化；该方法整合了复杂裂缝几何结构和注入参数，以支持实时和可持续运行 [26]。研究人员已应用机器学习算法，根据水文地球化学数据预测储层温度，并讨论机器学习模型和 Shapley 加性解释（SHAP）的可行性 [27,28]。基于降阶基且遵循物理规律的代理模型能够保留控制结构，并为敏感性分析和不确定性分析等多查询研究提供数量级的加速，同时也为纯黑箱网络提供了一种可解释的替代方案 [12]。更广泛的综述显示，地热领域的机器学习应用正在勘探、钻井、储层表征和生产等方面迅速增长；序列模型、卷积网络和概率学习器如今已普遍用于工作流程 [29–31]。相关研究还将机器学习扩展至诱发地震活动预测等运行风险问题；在此类研究中，更丰富的特征集和注意力机制改善了短期稳定性评估，但数据稀缺仍是主要限制因素 [32]。Yan 等人 [33]近期的研究展示了将热 DCA 用作增强型地热系统和热沉积盆地含水层系统代理模型的方法：他们引入 HyperReLU 递减模型，并训练神经网络，根据热–水–力模拟输入预测其参数。在其框架中，物理规律通过解析递减结构和损失函数中的基于方程的正则化项体现，而神经网络本身仍是一个标准的全连接参数回归模型。

此外，全物理场模拟器也是地热储层工程的核心工具，因为它们求解质量、动量和能量守恒方程的耦合系统，通常还会明确表示裂缝及其力学响应，从而能够在单一框架内，针对真实的储层几何结构和运行条件，预测流动、热传输和岩石变形 [14,15,34–36]。尽管这些模拟器能够稳健地预测生产温度、压力和长期热提取量，为循环速率、裂缝开度与渗透率、井距与井位布局以及运行寿命等设计选择提供指导，但其精细的物理过程以及较高的空间和时间分辨率使其

计算成本高昂，尤其是在纳入裂缝、复杂井位布局或较长生产周期时 [11,12,37]。

尽管这些方面都取得了进展，传统的（基于石油的）递减曲线方程并不符合地热温度变化，而用于地热预测的大多数机器学习代理模型也没有与热衰减的物理机制相结合。因此，单独采用任一方法都无法提供可靠且符合物理规律的温度预测，也没有统一的方法在不同增强型地热系统条件下对二者进行比较或结合。 consequently，目前尚不清楚何时传统递减分析已足够，何时机器学习能够带来决定性优势，以及混合方法能否结合二者的长处。

本研究通过在单一预测框架中统一基于物理的递减分析、数据驱动代理模型与高保真热-孔隙弹性模拟，弥合了三者之间长期存在的鸿沟。研究引入了完整 Arps方程族的平衡温度扩展，在保留精确退化为石油形式这一特性的同时，确保符合地热传热行为。此外，研究还采用方程嵌入神经网络，直接纳入这些物理定律，以实现可靠预测。该方法还包括一个高斯过程回归代理模型，用于快速生成不确定性感知预测。在单一热-水-力数据集中，研究评估了关键的地质、设计和运行变量。通过结合所有这些方法，研究提出了一个统一且符合物理规律的框架，从而减少了对高强度全物理场模拟的需求。

3. 增强型地热系统配置与数据集

本研究使用的数据集由 110 次高保真模拟构建而成，这些模拟代表了不同地质条件和运行条件下的增强型地热系统配置。用于生成数据的增强型地热系统模拟器集成了完全耦合的热-水-力（THM）框架，采用位移不连续法（DDM）公式模拟地热储层中的力学变形、流体流动和传热。该模型还包含裂缝内流动，其遵循泊肃叶定律。裂缝内热传输考虑了热传导、对流和岩石-流体相互作用 [13]。生成该数据集的首要目标，是捕捉裂缝几何形态、岩石热学性质和流体注入参数相互作用所形成的广泛温度递减曲线。通过系统地改变这些因素，该数据集为评估和比较传统递减曲线模型与现代机器学习代理模型在预测储层温度演化方面的表现，提供了全面的基础。

本研究采用双井式增强型地热系统配置，包括一口注入井和一口生产井。两口井通过水力裂缝相连，这些裂缝为井间流体循环提供流动通道；产出的高温流体则被输送至地表以回收热能。地层温度（T_{res}）为 146 °C，且

注入流体温度为 40 °C。模拟基于单相流体流动，并以水作为工作流体。在 146 °C、10 MPa 的储层条件下，温度为 40 °C 的注入水（T^{inj}）保持液态。地热地层性质列于表1。

表*1*。西部花岗岩的岩石性质 [38]。

ν	ν_u	G, GPa	k, md	ϕ, %	μ, cp	c_f, psi^{-1}	B	T_{res}, °C
0.25	0.33	15	0.0005	1	0.3	$2E-6$	0.85	146
σ_y, MPa	σ_x, MPa	σ_{xy}, MPa	p_r, MPa	α	B_f, $\frac{1}{°C}$	B_s, $\frac{1}{°C}$	C_T, $\frac{m^2}{s}$	T_{inj}, °C
20	13	0	10	0.44	$2.4E-4$	$2.1E-5$	$1.1E-6$	40

每个模拟案例都对应一种独特的增强型地热系统配置，该配置由五个设计变量定义：水力裂缝数量、井距、裂缝间距、围岩热导率和流体循环速率。选择这些变量，是因为它们代表了影响裂缝型地热储层热性能和长期采热的主要物理与运行参数。水力裂缝数量设为一条、两条或三条，以涵盖流动通道的不同复杂程度（图2）。井距范围为50至250米，反映了注入井与生产井之间的现场尺度间距（图3）。在多裂缝配置中，裂缝间距设为10至100米；单裂缝案例的间距则设为零（图4）。流体循环速率范围为30至300桶/米/天（图5）。热导率取值范围为1.5至15 W/(m·K)，涵盖结晶基底地层中常见的数值范围，并确保纳入低热导率和高热导率两类岩石（图6）。

为便于开展结构化分析，数据集按照水力裂缝数量进行分类（图2）。在单裂缝组中，井距相对较大，平均为143米；热导率均值为8.1 W/(m·K)，流体循环速率则相对较高，平均为152.4 桶/米/天（表2）。这种配置中没有裂缝间距，因此热交换几何结构较为简单，凸显了井距和热导率的影响。在双裂缝组中，平均井距降至约104米，同时引入了裂缝间距，均值为47米。该组的热导率均值略高，为9.2 W/m·K；流体循环速率平均为144 桶/米/天。三裂缝组的平均井距进一步缩小至86米，同时裂缝间距（均值

为46米）和热导率（均值为8.6 W/m·K）保持相近。流体循环速率仍相当，平均为143 桶/米/天，这表明一旦存在多条裂缝，裂缝数量对所需注入量的影响就较弱。

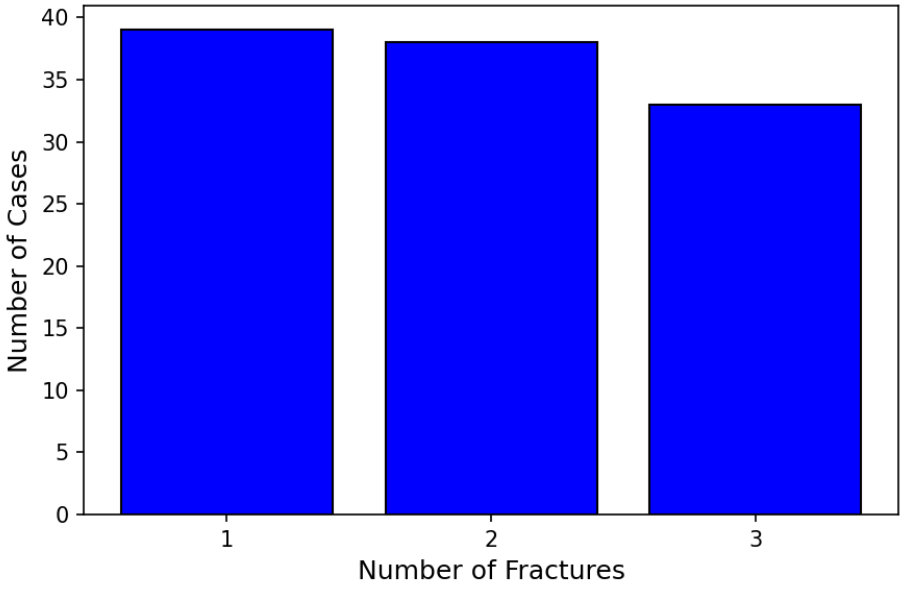


图*2.* 按水力裂缝数量划分的案例数。

表*2.* 按裂缝数量汇总井距、裂缝间距、热导率和流体循环速率的统计数据。

		均值	标准差	min	q25	q50	q75	max
单条裂缝	井距，m	143.1	58.5	50	100	130	200	250
	热导率，W/(m.K)	8.1	4.1	1.5	4.8	8.4	10.7	15
	流体循环速率，桶/天/米	152.4	74.6	37.5	100	147.4	209.3	299.4
两条裂缝	井距，米	103.9	32.3	50	82.5	100	120	200
	裂缝间距，米	47.3	25.3	10	30	39.2	64.4	99.8
	热导率，瓦/(米·开尔文)	9.2	3.9	1.8	6.6	10.7	10.9	15
	流体循环速率，桶/天/米	143.9	72.2	30.8	100	116.9	193.9	299.4
三条裂缝	井距，m	85.8	29.4	50	70	80	100	200
	裂缝间距，m	45.5	25.4	10	30	30	64.9	98.9
	热导率，W/(m.K)	8.6	4	1.5	5.1	10.7	10.7	15
	流体循环速率，bbl/day/m	142.7	74.3	35	100	117.7	200	296.8

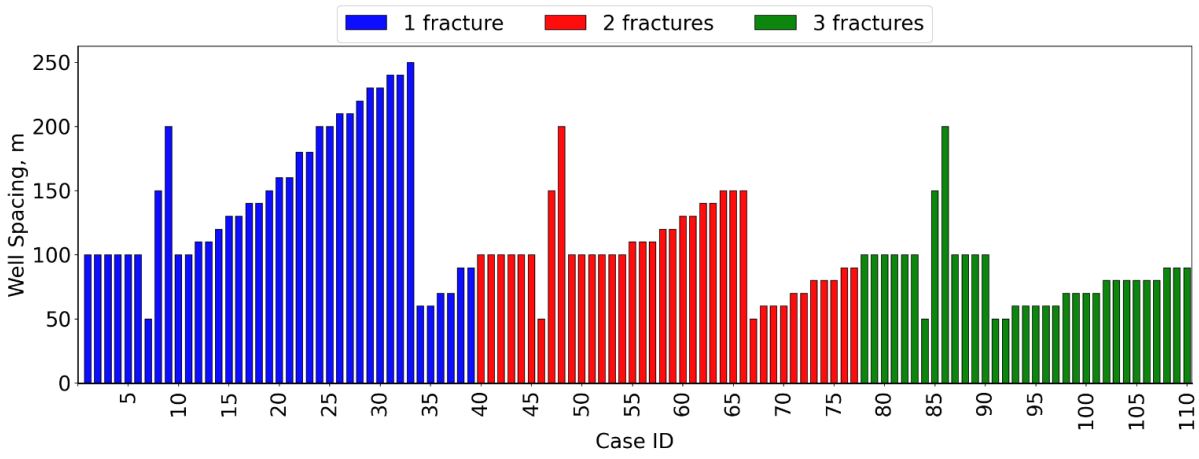


图 ***3***。*110* 个案例的井距分布（单位：米）。

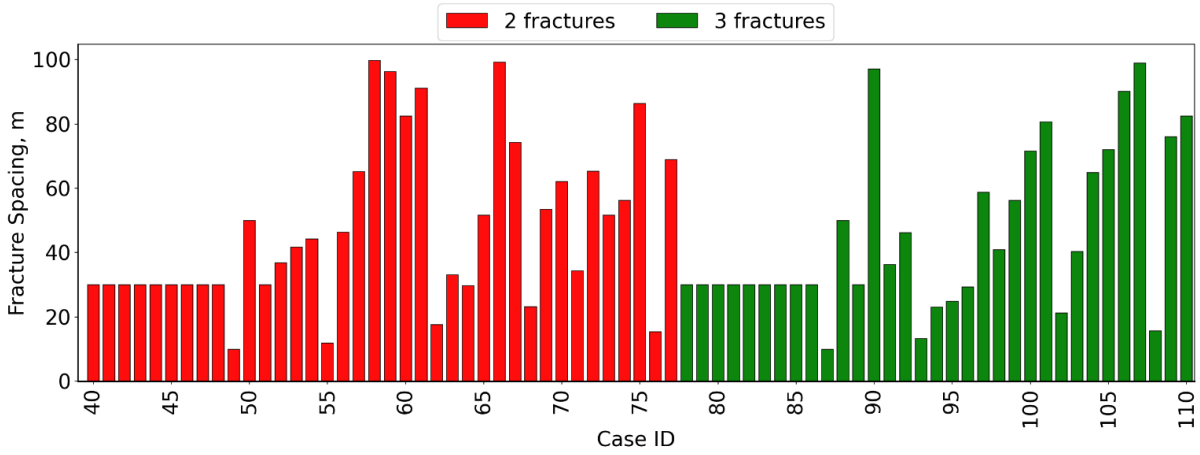


图 ***4***。*2* 条和 *3* 条裂缝案例的裂缝间距分布（单位：米）。

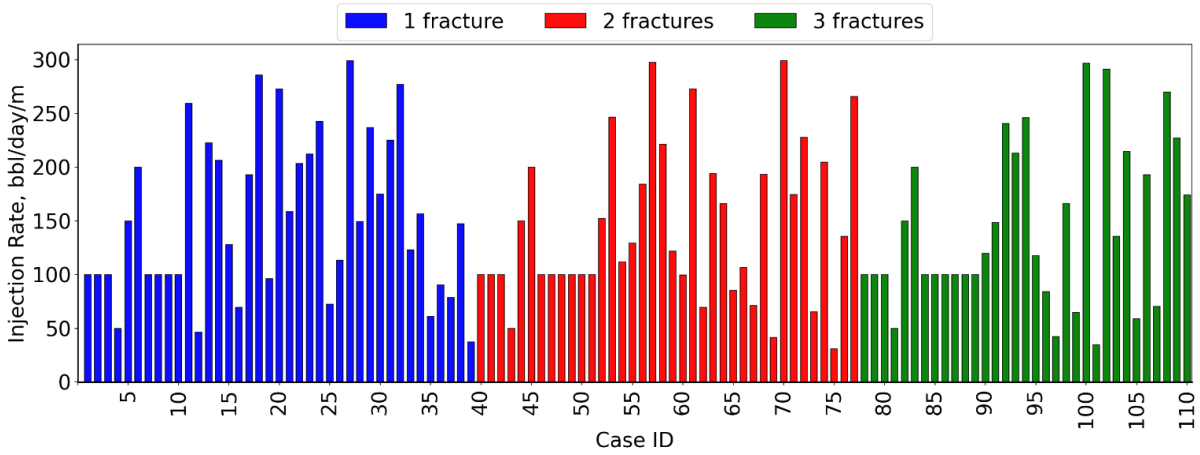


图 ***5***。*110* 个案例的流体循环速率分布，单位为桶/天/米厚度。

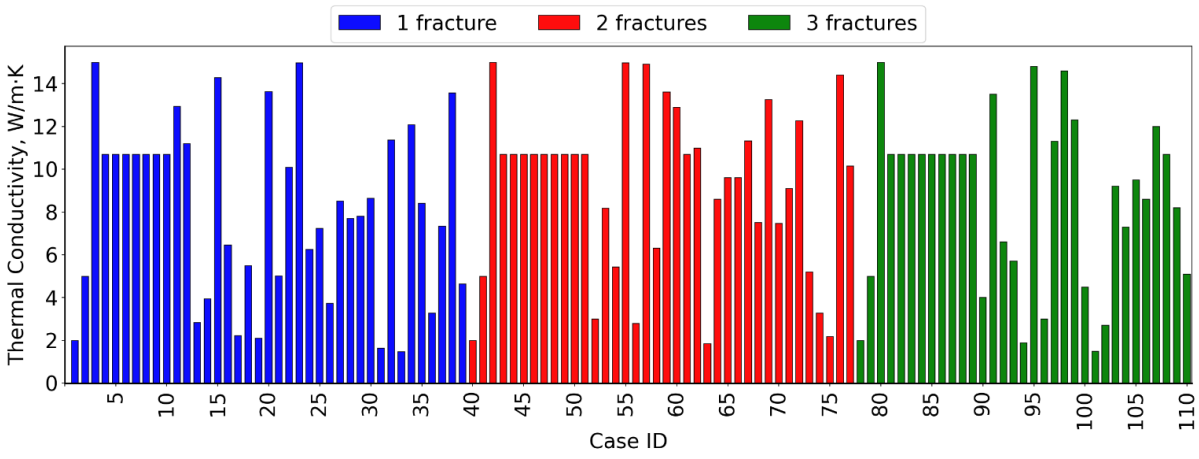


图 ***6***。 *110* 个案例的岩石热导率分布。

每个模拟案例都会生成一条温度-时间曲线，呈现储层温度的变化过程（图7）。时间序列按离散监测间隔（每隔3个月）选取，时间范围为3至60个月，为解析模型和机器学习模型提供因变量。通过将受控的地质与运行变量设计矩阵和一致的时间序列输出相结合，该数据集支持对不同衰减机制和建模框架的预测性能进行稳健的统计比较。最终形成的数据集为测试工程化地热储层中温度衰减模型的可解释性、准确性和泛化能力提供了全面的基准。

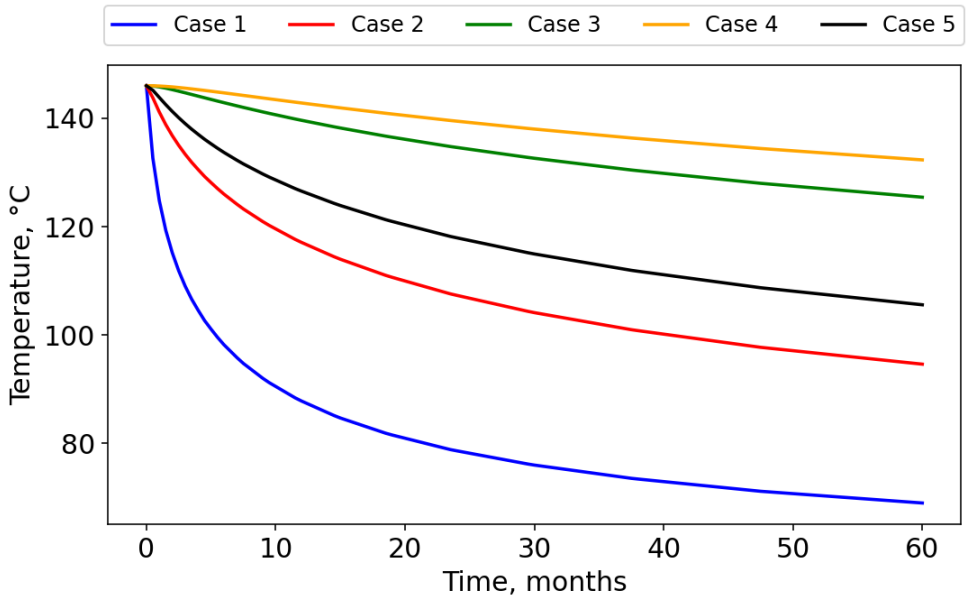


图 ***7***。 五种不同案例的温度随时间变化。

如相关矩阵（图8）所示，生产温度与岩石热导率始终呈强相关（在3至60个月期间为0.56–0.61）；其次，生产温度与井距的正相关性逐渐增强（3个月时为0.31，到60个月时升至≈0.42）。裂缝间距的影响较小且为正，并随时间略微减弱（从0.12降至0.10）。流体循环速率与温度呈中度负相关，且这种关系随时间增强（从-0.39至-0.44），表明在该数据集中，较高的循环速率与较低的温度相伴。裂缝数量的影响接近中性，并从微弱正相关逐渐转为轻微负相关（从0.05变为- 0.02）。总体而言，温度信号主要受热导率和井距（正向）的影响，并受到流体循环速率（负向）的抵消；裂缝间距和裂缝数量则起次要作用。

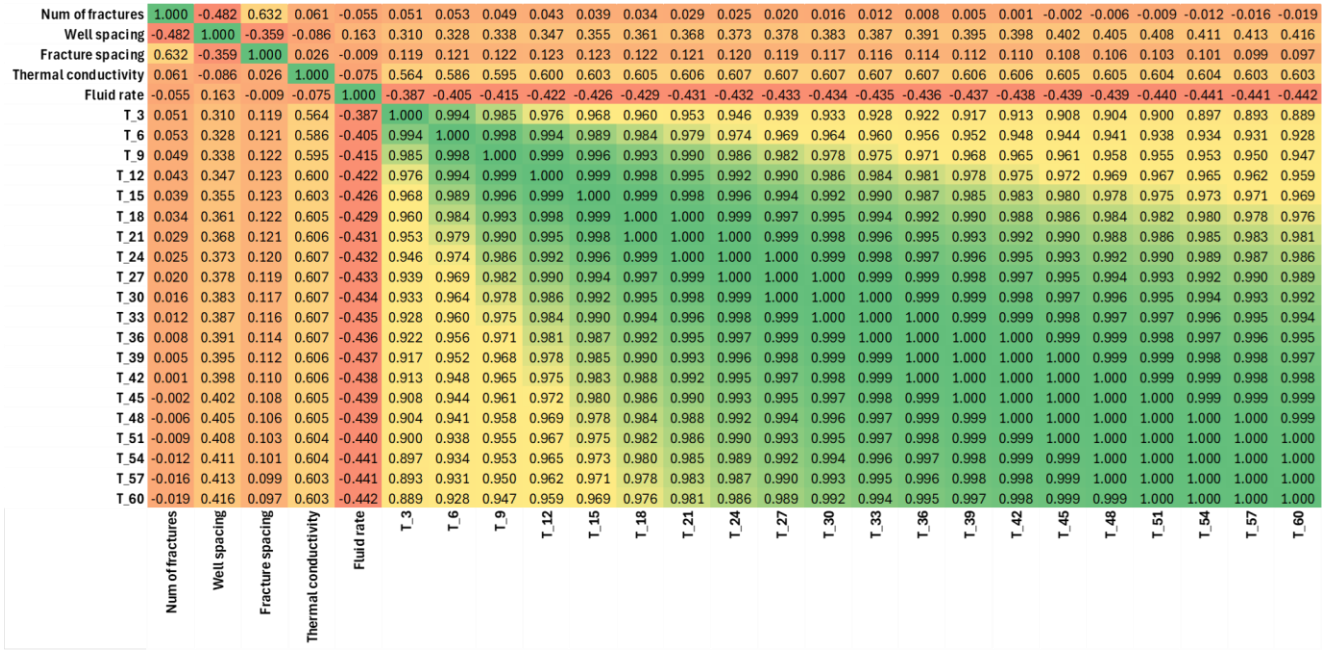


图*8*。 数据集变量的相关矩阵。

4. 方法

本节介绍用于分析和预测温度衰减的建模步骤。工作流程分为四个部分。首先，将改进的递减曲线方程应用于每个温度-时间序列，以获得简洁且符合物理规律的热行为概述。其次，训练方程信息神经网络，在保留 DCA 形式解析结构的同时，将储层和运行输入映射为完整的温度轨迹。第三，建立高斯过程回归模型，以直接生成多时域预测并量化不确定性。最后，XGBoost回归模型作为直接的数据驱动多时域基准。这些方法便于清晰比较基于物理的拟合、物理引导学习、灵活的统计预测以及纯数据驱动模型。

4.1. 地热扩展递减曲线模型

本研究扩展了 Arps 型递减方程（基于石油领域）[6,21,24,39,40]，引入平衡温度项（T_∞），使其适用于地热系统。该改进应用于所有衰减形式（指数型、调和型、双曲型和拉伸指数型），以表征围岩热基线。加入此项后，温度将趋近于有限的平衡值，而非零，从而准确反映地热系统的物理行为。若将平衡温度设为零，这些方程便直接简化为原始的 Arps 型递减曲线分析形式，常用于油气生产，其中关注变量渐近衰减至零。该公式遵循牛顿冷却定律 [41,42]，，该定律描述温度如何随时间逐渐趋近环境温度或平衡温度，确保其符合地热传热过程的物理规律。引入此项旨在克服 DCA 形式在表征地热系统发生热突破之前早期温度方面的局限。

采用改进递减曲线分析，通过以下四种改进模型刻画生产温度随时间的变化规律：指数型、调和型、双曲型和拉伸指数型。这些模型基于不同的物理假设描述温度随时间的演变，为解读递减趋势提供了定量框架。通过有界约束下的非线性最小二乘优化，确定各模型的参数，以确保解的合理性和稳定性。模型性能主要通过赤池信息准则（AIC）[43], 进行评估，该准则兼顾拟合质量与参数简洁性；此外，还采用 R^2 和均方根误差（RMSE）等指标进行诊断分析。

指数模型是一个三参数模型，假设分数递减率随时间保持恒定，因此是最简单且应用最广泛的 DCA 公式之一。该模型描述了这样一类系统：温度变化速率与初始状态和热平衡状态之间剩余的温差成正比。该模型适用于由传导冷却或指数递减等稳定、持续的衰减机制主导的过程。改进指数模型如下：

$$T(t) = T_\infty + (T^0 - T_\infty)\, e^{-D\, t} \qquad (1)$$

在该模型中，T(t)表示时间 t 时的温度；T0 是 t = 0 时的初始温度；T_∞ 是系统在较长时间后趋近的渐近温度或平衡温度；D 是控制温度衰减速度的指数递减率常数；t 表示时间。

调和模型是一个三参数模型，通过允许递减率随时间变化且变化较慢，扩展了指数形式。该模型假设温度变化的速率与时间成反比，从而描述递减逐渐减缓、而非遵循恒定指数趋势的情况。这种行为通常代表驱动力减弱或传热阻力增加的系统。修正调和模型可表示为：

$$T(t) = T_\infty + \frac{T^0 - T_\infty}{1 + D\,t} \qquad (2)$$

这里，T(t) 表示温度随时间 t 的函数；T0 为初始温度；T_∞ 为长期渐近温度；D 为决定温度随时间下降速度的调和衰减常数。与指数模型相比，该模型的递减率较低，反映了温度随时间逐渐下降的趋势。调和模型与指数模型采用相同的温度参数，并通过衰减常数控制衰减曲线的曲率。

双曲模型是一个四参数模型，它引入了额外的曲率参数，可对指数形式和调和形式进行一般化。这种灵活性使模型能够描述范围更广的衰减曲线，从快速衰减到逐渐趋缓的趋势。双曲指数决定递减率随时间变化的幅度。该模型常用于地下工程和储层研究，以捕捉复杂的非线性衰减模式。修正双曲模型可写为：

$$T(t) = T_\infty + (T^0 - T_\infty)\,(1 + b\,D\,t)^{-\frac{1}{b}} \qquad (3)$$

在此模型中，T(t) 表示时间 t 时的温度；T0 表示初始温度；T_∞ 表示最终温度或平衡温度；D 为衰减常数；b 为控制衰减曲线曲率的双曲指数；t 表示时间。当 $b \to 0$ 时，模型趋近于指数形式；当 $b = 1$ 时，模型则变为调和形式，因此能够灵活表示介于两者之间的衰减曲线。模型参数定义了初始温度和渐近温度、名义衰减常数，以及控制曲率变化的形状因子。

拉伸指数模型是一种四参数模型，代表了更广义的指数衰减形式；与单一特征速率不同，它能够考虑递减率的分布。该模型尤其适用于表现出非均匀弛豫或扩散过程的系统，因为这类系统的早期和晚期行为有显著差异。拉伸指数模型可写为：

$$T(t) = T_\infty + (T^0 - T_\infty)\,exp\left(-(D\,t)^\beta\right) \qquad (4)$$

这里，T(t) 表示温度随时间 t 的变化；T0 为初始温度；T_∞ 为渐近温度；D 为定义递减特征时间尺度的速率常数； β 为拉伸指数，用于控制曲率以及偏离简单指数行为的程度；t 表示时间。该模型能够描述递减率随时间变化、因而衰减较慢且不均匀的系统。模型参数包括初始温度、渐近温度、速率常数，以及用于描述偏离纯指数行为程度的拉伸指数。

所有模型均通过最小化观测温度与预测温度之间的残差平方和（RSS）进行拟合，确保在设定范围内获得尽可能好的拟合效果。

$$RSS = \Sigma (y_i - \hat{y}_i)^2 \quad (5)$$

在该表达式中，yi 是观测温度数据点，y i 是模型在相同时间 ti 对应的预测值。RSS 量化了观测值与模型预测值之间的总偏差，并作为曲线拟合过程中需要最小化的目标函数。

使用赤池信息准则 [43], 比较模型性能，该准则在奖励拟合优度的同时会惩罚过高的模型复杂度。

$$AIC = n\, ln\left(\frac{RSS}{n}\right) + 2k \quad (6)$$

在该方程中，n 是数据点的数量，RSS 是残差平方和，k 是拟合参数的数量。AIC 衡量模型的相对质量，兼顾拟合优度与复杂度。较低的 AIC 值表明模型在拟合质量与简洁性之间取得了最佳平衡，从而指导为每种情况选择最具代表性的衰减模型。

4.2. 方程知情神经网络（EINN）

方程嵌入模型采用神经网络（NN）实现，并将其应用于修正的DCA公式（第4.1节）生成的温度衰减输出。这样，NN便能从基于物理规律的衰减模式中学习，同时保留DCA定义的底层趋势和参数。开发这一方程嵌入监督模型的目的是在固定时间网格上预测衰减曲线参数及相应的温度轨迹。输入包含五个预测变量（裂缝数量、井距、裂缝间距、热导率和流体循环速率）。目标值遵循四种经典DCA方程（指数型、调和型、双曲型、拉伸指数），每种方程均由其标准符号参数化（例如，T∞、D、b、 β）。所有特征预处理（包括对数变换和标准化）均仅在每个数据划分或交叉验证折中的训练子集上拟合，随后不作任何更改地应用于验证子集和测试子集，以防止来自留出案例的信息泄漏

时间域覆盖0–60个月，间隔为0.5个月，初始温度T^0 = 146.0固定。

为防止情景级数据泄漏，方程信息神经网络的所有数据划分均以热–水–力模拟案例为单位进行。每个 THM 案例都会生成多个样本，分别对应不同的递减曲线公式；因此，源自同一模拟案例的所有样本均被完整地保留在同一个数据子集中。首先进行独立的分组留出划分，将 20% 的模拟案例留作测试集，其余 80% 的案例用于训练和验证。在训练验证集中，采用分组K折交叉验证（K = 5），确保同一个模拟案例不会同时出现在训练折和验证折中。

预测网络是一个前馈式多头神经网络，由包含两个全连接层的共享主干网络（每层 128 个神经元，使用ReLU激活函数）以及相互独立的单层线性输出头组成，每个输出头对应一个 DCA 模型（指数型、调和型、双曲型和拉伸指数型）。每个输出头预测其递减曲线方程特有的参数集：指数形式和调和形式各有两个参数，双曲型和拉伸指数形式则各有三个参数。储层初始温度 T^0, 不会在此机器学习模型中进行预测，因为它是已知值，代表储层的初始热状态。预测参数经过有界变换（使用 Softplus函数和缩放 Sigmoid函数），以确保其符合物理要求，例如递减率和渐近线为正值，并对 b 和 β 进行约束。给定某个输出头的参数后，方程约束计算层会使用相应的 DCA 方程，在 0–60 个月的时间网格上计算完整的温度–时间轨迹。

训练集优化了一个复合目标函数，该函数结合了时间加权曲线偏差和参数一致性项。曲线项是在时间网格上逐点计算的Huber损失（δ = 1.0），并乘以归一化的时间权重w(t) = 1 /（1 + t/τ），其中τ = 18 月，再对时间取均值。参数项是头部预测参数与参考递减曲线分析参数之间的均方误差。每个样本的总损失为:

$$L = \underbrace{Huber_{\delta}(T_{pred}(t), T_{ref}(t))\ averaged\ with\ w(t)}_{curve\ loss} + 0.2 \times \underbrace{||\theta_{pred} - \theta_{ref}||^2}_{parameter\ loss} \quad (7)$$

在该方程中，T_{pred} 和 T_{ref} 分别表示预测温度轨迹和参考温度轨迹，它们是时间 t 的函数；$Huber_{\delta}$ 表示Huber损失函数，其转折参数为 δ=1.0，用于对不同幅度的残差结合二次和线性区间；w(t) = 1 (/ 1 + /τ (t))是时间加权函数，其中τ= 18 月用于强调早期精度；θ_{pred} 和θ_{ref} 分别表示预测参数向量和

参考递减曲线分析参数向量；$||\theta_{pred} - \theta_{ref}||2$表示参数向量之间的均方差；系数 0.2 用于调整参数一致性项相对于曲线拟合项的权重，以平衡温度曲线保真度和参数准确性。

通过简单的按类别过采样，使各模型方程对应的小批次保持平衡。优化采用 Adam，学习率为 $10^{-}3$，权重衰减为 $10^{-}4$，批次大小为 64，训练 600 轮次。平台期衰减调度器（因子 0.5，耐心值 15）根据验证损失调整学习率。

使用均方根误差和平均绝对误差，将绘制出的机器学习曲线与参考 DCA 曲线进行比较。交叉验证后，使用完整的训练集+验证集训练最终模型，并在 20% 的测试集上采用相同指标进行评估；结果按方程分别报告，并给出总体结果（均值 ± 标准差）。输出包括所有案例的参数预测、逐案例曲线图（参考曲线与机器学习曲线），以及 15、30、45 和 60 个月时真实温度与预测温度的散点图。

4.3. 高斯过程回归（GPR）

对每个预测时域（从 3 个月到 60 个月，间隔为 3 个月）分别应用高斯过程回归（GPR），以五个预测变量为依据对温度演化进行建模：裂缝数量、井距、裂缝间距、热导率和流体循环速率。GPR 数据集中的每条观测对应一个模拟案例，每个案例由一组输入变量以及多个时间范围内的温度目标值向量表征。按案例标识独立进行留出划分，将 80% 的案例用于训练，剩余 20% 留作未见测试集（图 9）。这种按案例评估策略可防止相关温度预测时域之间发生信息泄漏，并确保报告的 GPR 性能能够反映模型对未见地热系统配置的泛化能力。

在模型开发过程中，仅针对训练案例执行了 K折交叉验证，以估计泛化性能并生成折外（OOF）预测。在每一折中，均使用仅根据该折训练数据计算的 Z分数变换对特征集进行标准化，使每个特征的均值为 0、标准差为 1。交叉验证完成后，在最终模型测试前，使用完整训练数据重新拟合了标准化器。

每个目标变量均使用高斯过程回归器进行建模，其复合协方差核由常数核与径向基函数（RBF）核相乘构成，并与线性（点积）分量及白噪声项相加。该结构使模型能够捕捉非线性平滑变化和线性全局趋势。模型使用归一化输出和数值正则化参数 $\alpha = 10^{-6}$。

采用决定系数（R^2）、均方根误差（RMSE）和平均绝对误差（MAE）评估预测性能。针对每一折计算各项指标，并汇总各折的均值和标准差，以概述训练性能。使用完整训练数据集完成训练后，采用相同指标评估模型在独立测试分区上的准确性。预测不确定性根据高斯过程的后验方差得出，95%预测区间计算为 $\mu \pm 1.96\,\sigma$，其中 μ 为均值，σ 为标准差。诊断输出包括散点图，用于比较选定预测时域（15、30、45 和 60个月）上的预测温度值与真实温度值。

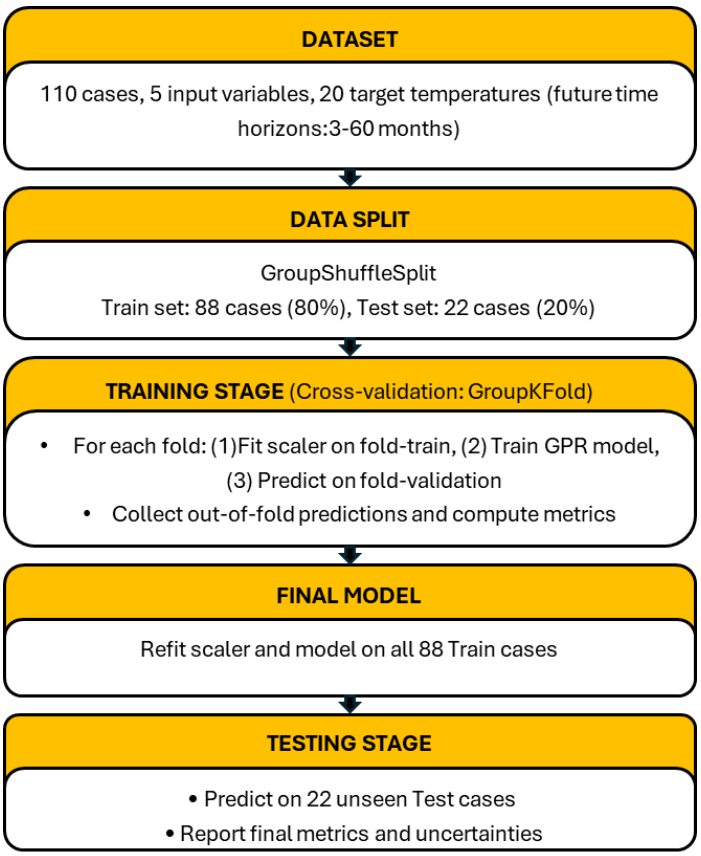


图 ***9***。 高斯过程回归示意图，展示了数据结构、划分为训练集和留出集的数据，以及训练集内部的交叉验证（用于生成折外预测）。

4.4. 直接 XGBoost

采用基于 XGBoost 的确定性梯度提升决策树模型，作为温度预测的非物理信息基线。该模型的实现旨在与其他代理模型采用相同的数据处理方式、数据划分策略和评估方案，以确保比较公平且一致。预测任务被表述为直接多时域回归。对固定时刻的储层温度进行预测，

预测时刻范围为 3 至 60个月，间隔为 3 个月，各时刻的温度分别独立预测。每个预测时域都作为一个独立的回归问题，使用相同的一组输入特征（包括水力裂缝数量、裂缝长度、裂缝间距、围岩热导率和注入速率），这些特征描述了裂缝特性、热学性质和运行条件。每个案例对应一次热-水-力模拟，并针对每个预测时域生成一个温度值。

为防止相关样本之间发生信息泄漏，所有数据划分均以模拟案例为单位，并使用组标识符进行。首先，使用分组感知随机划分构建一个由 20% 的案例组成的未见留出测试集。其余 80% 的案例构成训练池。在训练池中，使用分组感知K折交叉验证评估模型，确保同一模拟案例的所有样本只出现在训练折或验证折中的一方。这种方法既能保持各折之间的统计独立性，又能稳健地估计泛化性能。

XGBoost回归器采用平方误差目标函数和基于直方图的树构建方式进行训练，以提高计算效率。支持两种预定义配置：一种是容量适中的配置，可加快训练速度；另一种是容量更高的配置，采用较小的学习率和更深的树。固定随机种子以确保不同运行之间结果可复现。通过以分组感知的方式应用早停来控制过拟合。在每个训练折中，仅从训练部分创建一个基于分组的验证集划分。当验证集均方根误差在指定的提升轮数内未能改善时，模型训练即终止。使用兼容性封装器，确保早停在不同XGBoost版本中均能正常运行。此流程既用于交叉验证，也用于测试前在完整训练集上进行的最终训练。

模型性能分两个阶段进行评估。首先，使用 GroupKFold 交叉验证的折外预测计算无偏训练指标。其次，在完整训练集上重新训练的模型仅在未见过的留出测试集上评估一次。采用决定系数（R2）、均方根误差（RMSE）和平均绝对误差（MAE）量化性能，并按各预测时域分别报告，同时汇总所有预测时域的宏平均指标。

5. 地热扩展型递减曲线分析模型的验证

本节使用来自犹他州 FORGE 项目的现场案例，对地热扩展递减曲线分析模型进行验证。验证集由通过安装在 Utah FORGE 16B(78)-32井、测量井深为7,056.67英尺处的光纤传感器获取的井下压力-温度测量数据组成 [44]。这些数据包含2024年4月采集的压力和温度时间序列记录

（图10）。16B(78)-32 号井在 FORGE场址被用作储层构建、流体循环和采热示范的生产井。该井作为双井配置的一部分进行钻井，位于16A(78)-32注入井上方约300英尺处，并与其平行。尽管计划总井深为10,658英尺，但实际钻井深度超过了这一目标；作业于2023年4月26日开始，到2023年6月20日，测量井深达到10,947英尺，垂直深度为8,357英尺。钻井作业包括对由 Eavor Technologies 提供的隔热钻杆（IDP）进行现场试验 [45]。IDP 限制钻柱内较冷钻井液与较热环空返流之间的逆向传热，从而确保井底钻具组合始终浸没在相对较冷的流体中。钻井期间，分两次连续使用井底钻具组合，共部署了350根IDP钻杆 [45]。

图10以左侧纵轴（°F）表示井下温度，以右侧纵轴（psi）表示压力，并将二者绘制为当地日期和时间的函数。这些记录捕捉了整个流动期间的作业瞬态，包括生产压力保持相对稳定的时段。生产压力变化伴随着温度响应，而其他时段则显示，在压力近乎恒定的情况下，热演化更为缓慢。这些模式表明，井作业期间流动和运行条件发生了变化，并构成了后续分析中递减曲线拟合与验证集的实证基础。

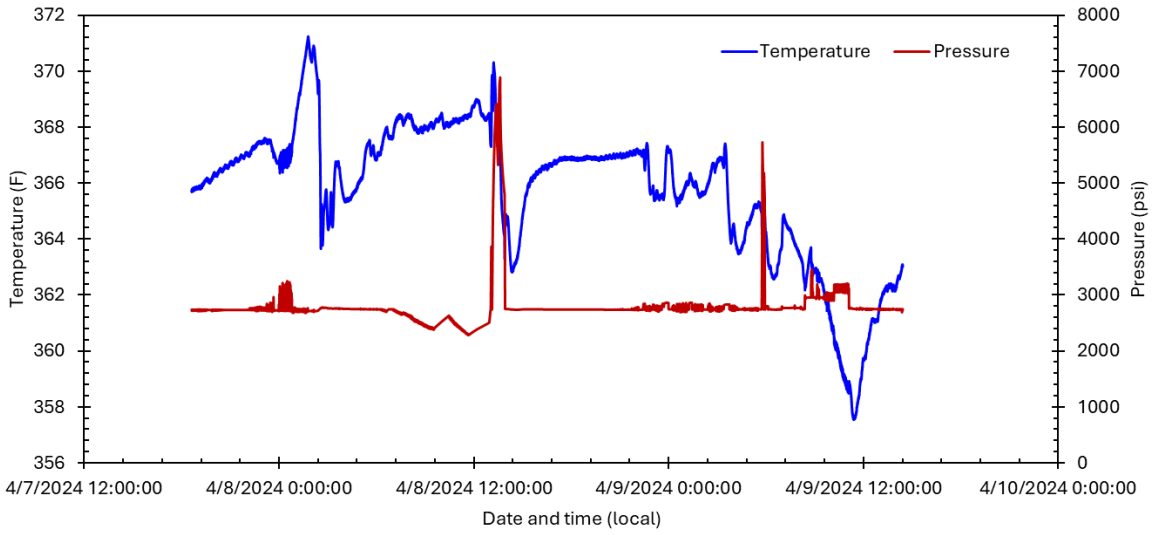


图***10***。 犹他州 *FORGE 16B(78)-32*井的井下压力和温度现场测量数据。

图11展示了实测井下温度数据与分别采用传统（基于石油行业）和修正递减曲线（扩展至地热领域）公式所得拟合结果的对比。温度观测值以蓝色数据点表示，传统模型和地热扩展模型则以连续曲线表示。每个子图对应一种

不同的递减曲线形式，包括指数型、调和型、双曲型和拉伸指数模型。所有面板均使用相同的温度时间序列数据进行模型拟合，从而能够直接且一致地比较各公式对整体温度递减行为的刻画效果。随着时间推移，传统公式与修正公式之间的差异日益明显；此时，曲率和递减率的差异会导致对观测温度趋势的呈现明显不同。

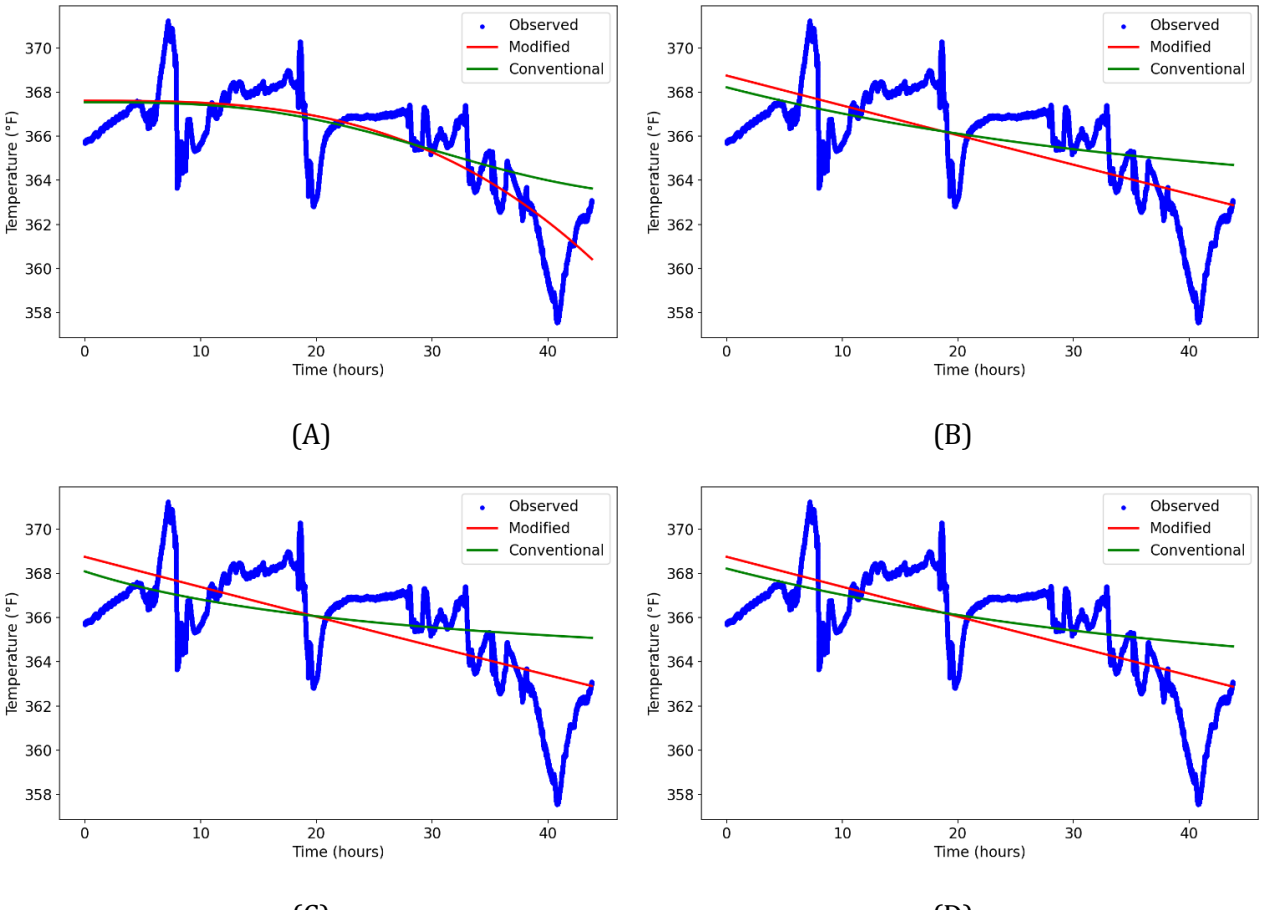
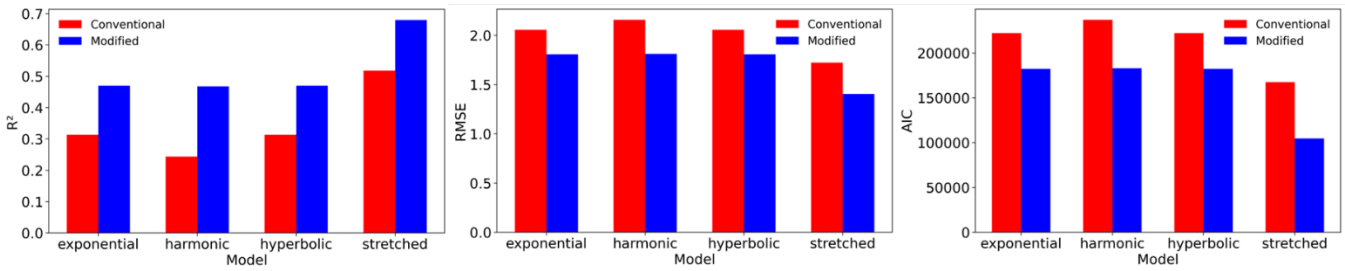


图*11*。采用传统和改进递减曲线模型拟合的温度数据：（*A*） 拉伸指数型，（*B*） 双曲型，（*C*） 调和型，以及（*D*）指数型。

图12汇总了应用于图11所示温度数据的传统和改进递减曲线模型的统计性能，并以决定系数（R^2）、均方根误差（RMSE）和赤池信息准则（AIC）作为拟合质量与模型简约性的定量指标。在所有递减公式（指数型、调和型、双曲型和拉伸指数型）中，改进模型始终优于传统模型，表现为更高的R^2值以及更低的RMSE和AIC值。例如，改进的指数模型将

R^2从约0.31提高至0.47，同时将RMSE从约2.06降至1.81，AIC也大幅降低。调和型和双曲线形式也取得了类似改进：解释方差增加、误差降低，同时AIC值更低，表明拟合效果有所改善且复杂度未过高。拉伸指数模型的性能提升最为显著，其R^2最高（约为0.68），RMSE最低（约为1.41），且AIC明显低于传统公式。这些指标表明，改进的递减曲线公式能够始终更优地表征观测到的温度递减行为，同时保持经复杂度调整后的模型性能优势。

尽管修正递减曲线公式引入了一个额外的拟合参数，但在评估模型性能时，并不会孤立地看待决定系数的提高和均方根误差的降低。因此，模型比较主要依据赤池信息准则，该准则通过对额外参数的引入施加惩罚，明确考虑了模型复杂度。所有修正公式所得的赤池信息准则值始终较低，这表明观察到的改进源自拟合优度与模型简洁性之间更有利的权衡，而非对温度数据的过拟合。此外，加入平衡温度项具有明确的物理依据，因为它能确保长期行为趋于有限值，与地热传热过程相一致。重要的是，当平衡温度设为零时，修正方程会精确地退化为常规 Arps递减形式，从而确保与经典的基于石油的递减分析向后兼容。



(A) (B) (C)

图***12***。 常规与修正递减曲线模型的统计性能：*(A)* 决定系数，*(B)* 均方根误差，以及 *(C)* 赤池信息准则。

6. 结果与分析

本节展示了完整热-水-力数据集的主要发现。结果的编排旨在展示每种建模方法如何捕捉产温行为，以及

这些方法在设计空间中的泛化能力。首先，评估扩展递减曲线模型对温度历史的拟合效果。随后，考察方程信息神经网络根据设计输入重现完整轨迹的能力。最后，分析高斯过程和XGBoost回归模型的准确性，以及它们在多个时间范围内提供不确定性感知预测的能力。

6.1. 地热递减曲线分析

分析采用四种经验递减形式，对110个时间-温度递减案例进行了评估，并使用赤池信息准则（AIC）为每个案例选出最佳表示形式。在所有最佳拟合案例中，模型拟合都极为紧密，决定系数的中位数为0.9999，均方根误差的中位数为0.071 °C，平均绝对误差的中位数为0.048 °C（表3）。根据AIC准则，110个案例中有66个（60.0%）选择了拉伸指数形式，41个（37.3%）选择了双曲线形式，3个（2.7%）选择了调和形式；简单指数替代形式并未被选为最合理的描述（图13）。这些结果表明，大多数温度下降轨迹要么轻微偏离纯指数响应，符合拉伸弛豫过程的特征，要么呈现出可由双曲线方程描述的曲率。

表***3***。*110*个案例中最佳拟合模型的统计数据：*R2*、*RMSE*和*MAE*的均值、中位数和标准差。

统计量	**R^2**	**RMSE, °C**	**MAE, °C**
mean	0.9998	0.116	0.084
中位数	0.9999	0.071	0.048
标准差	0.0003	0.124	0.094

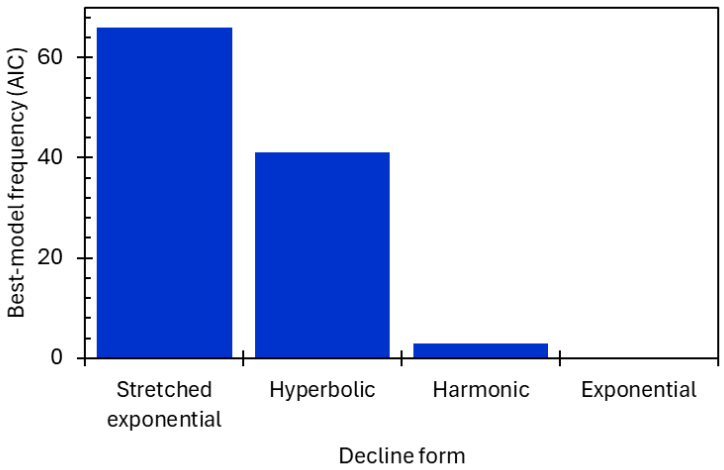


图***13***。 *AIC*最优 *fre* 在 *110* 个案例中的出现频率；拉伸指数型占主导，其次是双曲型。

选定模型的拟合优度指标集中在较低的误差水平，支持了所推断特征参数的可靠性。在选定的拟合结果中，拉伸指数情形的典型误差最低，简约性调整支持度也最强，其中位均方根误差为≈ 0.043，中位平均绝对误差为≈ 0.032（表4）。双曲线选择的表现也很出色，其中位均方根误差为≈ 0.089，中位平均绝对误差为≈ 0.069。三个调和选择的均方根误差中位数为≈ 0.068，平均绝对误差中位数为≈0.0385，但由于样本量较小，这些选择仅适用于特定的衰减几何形态。总的来说，误差分布表明，DCA模型能够以百分之几度的误差复现观测到的温度轨迹，且拉伸指数形式始终表现更优。图14展示了110个最佳拟合案例的均方根误差。

表***4***。*110*个最佳拟合情形的各模型性能统计量：*R2*、*RMSE*和*MAE*的均值、中位数和标准差。

模型	AIC最优案例	R^2			RMSE, °C			MAE, °C		
		mean	中位数	std	mean	中位数	std	mean	中位数	std
调和型	3	0.9999	0.9999	9.0E-06	0.0677	0.0676	0.0008	0.0386	0.0385	0.0002
双曲型	41	0.9998	0.9999	0.0003	0.1250	0.0890	0.0982	0.0971	0.0693	0.0883
拉伸指数	66	0.9998	0.9999	0.0003	0.1125	0.0429	0.1397	0.0786	0.0318	0.0983

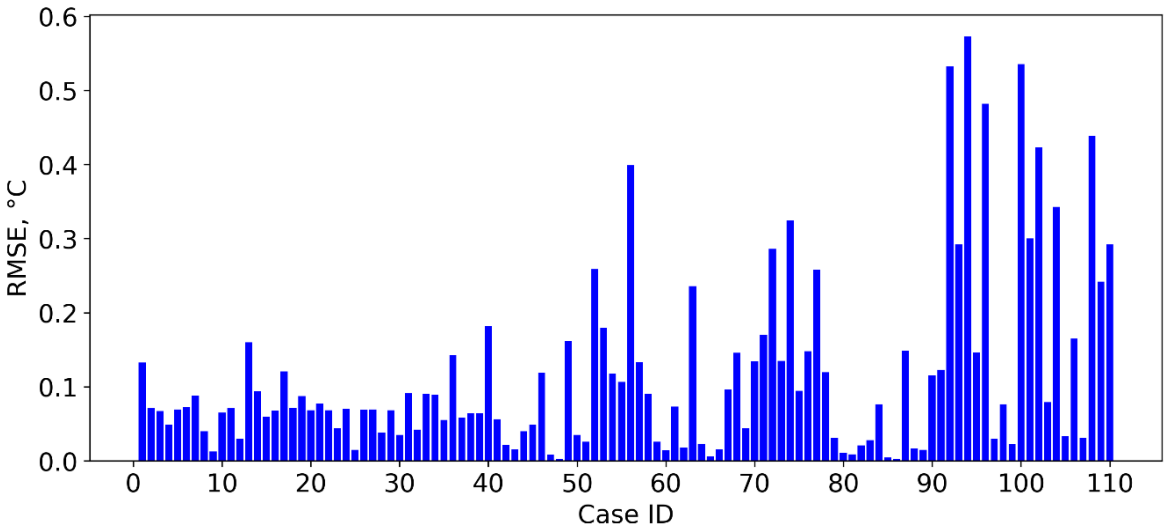


图***14.*** 使用 *DCA* 对 *110* 个 *AIC* 最优拟合案例进行的均方根误差。

总体而言，在数据集中的大多数情况下，温度下降曲线最适合用拉伸弛豫过程来描述（图15）。误差幅度普遍较小，参数估计稳定且易于解释，选择偏好与略微偏离单一

速率均衡的轨迹相一致。建议的呈现方式着重突出拟合残差的紧凑性，并展示观测曲线与选定曲线的代表性叠加图。

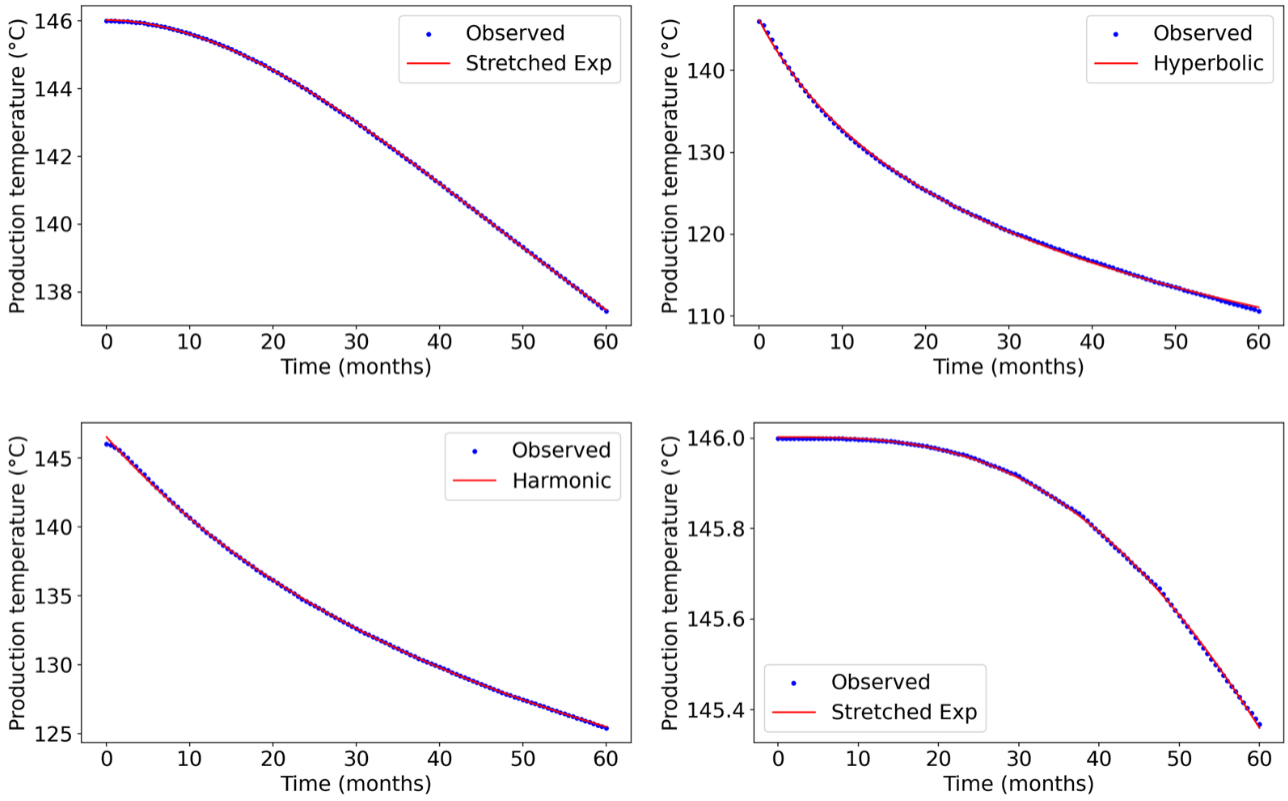


图*15*。 展示代表性 *DCA* 拟合的示例，说明拉伸指数型和双曲型分别能够捕捉早期曲率和灵活地描述衰减趋势。

6.2. 方程知情神经网络（EINN）

模型使用 440 个拟合后的递减案例进行训练，这些案例涵盖 110 种储层情景下的四种函数形式（指数型、调和型、双曲型和拉伸指数型），如第 6.1 节所述。这些案例的拟合质量始终很高，决定系数值介于 0.85 至 1.00 之间，平均值为 0.99，表明 DCA 拟合与参考温度历史高度吻合。分层抽取的20%案例（共 88 个，每种衰减形式 22 个）留作测试，其余 80% 用于模型开发，并采用 5 折交叉验证。

在训练验证集上进行的交叉验证进一步 corroborated 了这些发现。综合各折结果，整体验证误差为：平均绝对误差 5.3 °C，均方根误差 7.292 °C（表5）。按衰减形式汇总的验证均值如下：调和型的平均绝对误差为 5.15 °C，均方根误差为 7.066 °C；拉伸指数的平均绝对误差为 5.259 °C，均方根误差为 7.233 °C；指数型的平均绝对误差为 5.045 °C，均方根误差为 6.837 °C；双曲型的平均绝对误差为 5.745 °C，均方根误差为 7.923 °C。箱线图

展示了各衰减形式在所有交叉验证折中的平均绝对误差和均方根误差分布，见图16。

图17展示了在四个具有代表性的预测时域（15、30、45 和 60 个月）下，EINN 对训练集参考温度的预测一致性图。所有预测时域上的数据点都紧密聚集在 1:1 线附近，表明在整个温度范围内预测结果与参考值高度一致，系统偏差极小。在早期和中期预测时域（15–30 个月），预测准确度保持较高，决定系数为 = 0.98，均方根误差为 ≈ 2.82- 2.88 °C；随着变异性增加，后期预测时域的性能仅略有下降（45 个月时决定系数为 = 0.97，均方根误差为 = 3.89 °C；60 个月时决定系数为 = 0.96，均方根误差为 = 4.94 °C）。60 个月时散布略有增加，说明长时段预测误差有轻微累积；但整体校准仍然良好，且主导趋势得以保留。

表*5*。按衰减形式划分的 *EINN* 模型交叉验证折外性能。对于每种形式，该表报告所有折的平均绝对误差和均方根误差的均值 ± 标准差。

形式 / 模型	N	平均绝对误差 (°C)（均值 ± 标准差）	均方根误差 (°C)（均值 ± 标准差）	折数
指数型	88	5.045 ± 1.176	6.837 ± 1.465	5
调和型	88	5.150 ± 1.234	7.066 ± 1.505	5
双曲型	88	5.745 ± 1.299	7.923 ± 1.609	5
拉伸指数	88	5.259 ± 0.912	7.233 ± 1.076	5
All	352	5.300 ± 1.074	7.292 ± 1.322	5

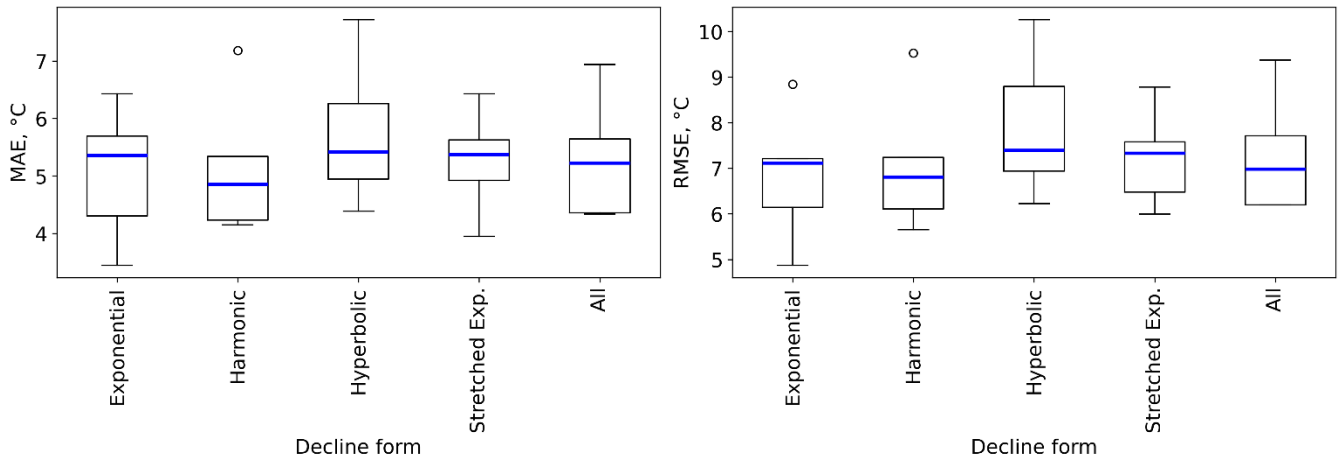


图*16*。 *EINN*模型在不同衰减形式下的预测性能。箱线图展示了每种形式在所有交叉验证折中的平均绝对误差和均方根误差分布。*All*类别表示所有衰减形式的汇总结果，

所有形式。

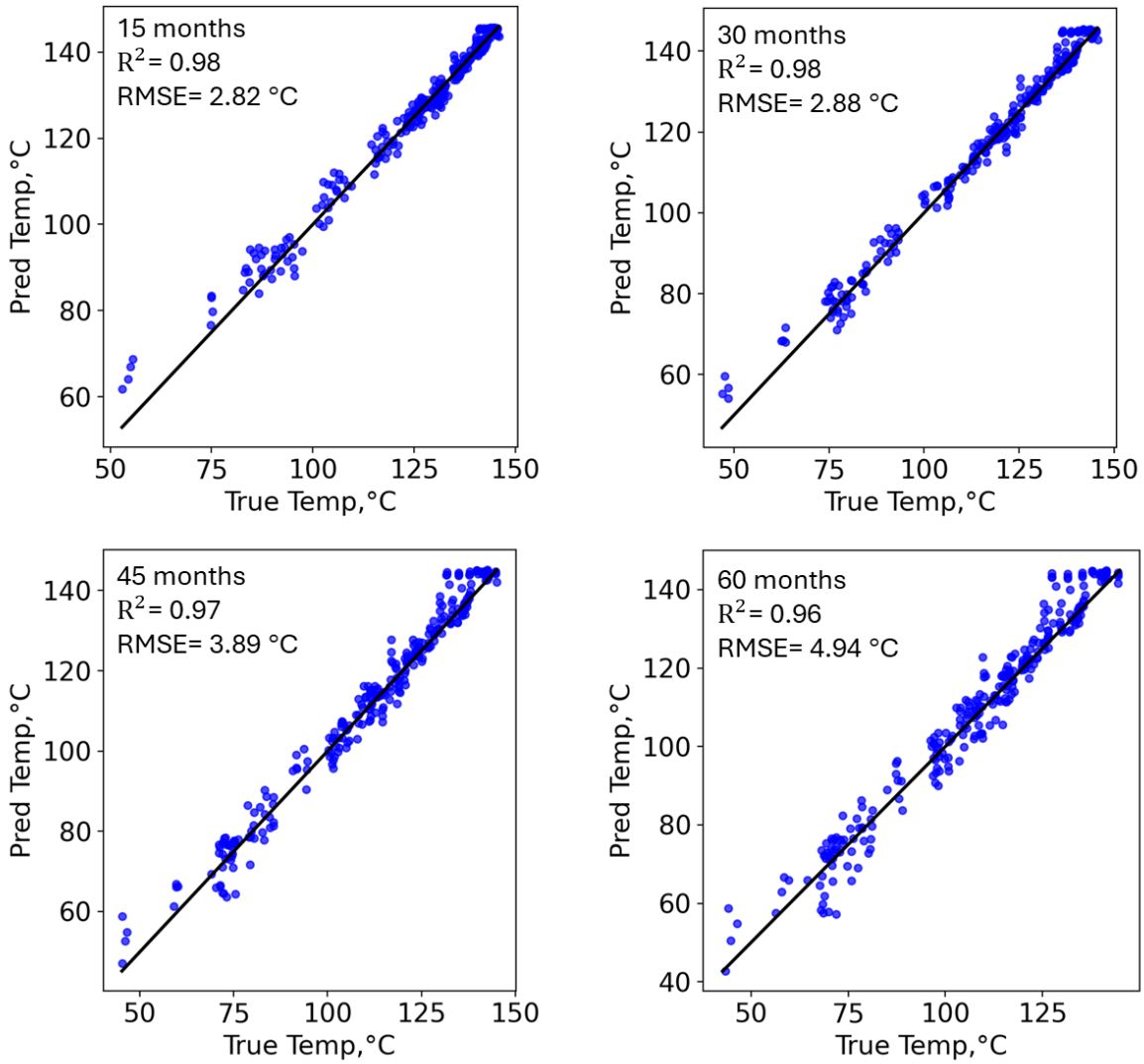


图***17***。训练集中四个代表性时间范围（*15*、*30*、*45* 和 *60* 个月）下 *EINN* 预测温度与真实温度的一致性图。每个面板均显示 *1:1* 参考线，以及相应预测时域的决定系数和均方根误差。

在独立测试集上，该方法的预测结果与参考曲线高度一致（表6）。总体平均绝对误差为 3.06 °C，相应的均方根误差为 4.49 °C。不同衰减形式下的表现稳定：拉伸指数型的典型误差最低（平均绝对误差为 2.586 °C，均方根误差为 3.423 °C），其次为调和型（平均绝对误差为 2.914 °C，均方根误差为 4.556 °C）。指数型和双曲型的离散度较高，但仍与参考值保持良好一致（指数型：平均绝对误差为 3.248 °C，均方根误差为 4.684 °C；双曲型：平均绝对误差为 3.489 °C，均方根误差为 5.117 °C）。这些结果表明，在 0–60 个月的预测时域内，该方法能够可靠地再现参考温度轨迹，其中拉伸指数型和调和型的吻合度尤其高。

表*6*。 按衰减形式划分的*EINN*模型留出测试性能。对于每种形式，该表报告所有测试案例的决定系数、平均绝对误差和均方根误差的均值± 标准差。

形式 / 模型	N	平均绝对误差（°C）（均值 ± 标准差）	均方根误差（°C）（均值 ± 标准差）
指数型	22	3.248 ± 2.689	4.684 ± 4.377
调和型	22	2.914 ± 2.932	4.556 ± 4.935
双曲型	22	3.489 ± 2.678	5.117 ± 3.902
拉伸指数	22	2.586 ± 1.726	3.423 ± 2.498
All	88	3.059 ± 2.571	4.489 ± 4.205

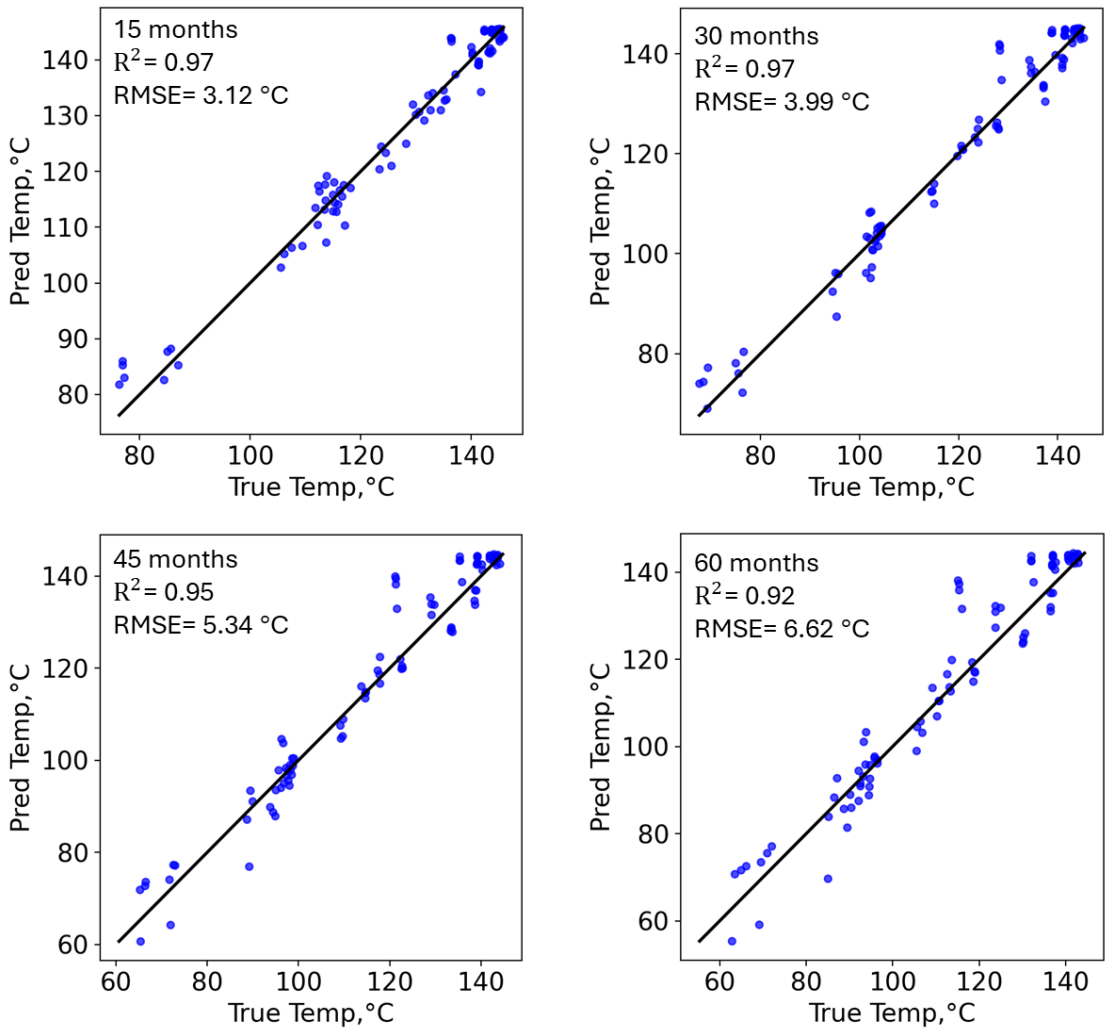


图*18*。 *EINN* 对*22*个留出案例在四个代表性

时间范围（*15*、*30*、*45*和*60*个月）内温度的预测值与真实温度的一致性图。每个面板均显示*1:1*参考线，以及针对相应预测时域计算的决定系数和均方根误差，表明模型在短期和长期预测中均具有较强的一致性

forecasts.

图18展示了一致性图，对比了四个预测时域（15、30、45和60个月）内，EINN对22个留出案例的温度预测与真实值。每个面板中的数据点都紧密聚集在1:1参考线周围，表明预测温度与观测温度之间校准良好。所有预测时域的拟合优度都很高，R^2值介于0.92至0.97之间，均方根误差介于3.12 °C至6.62 °C之间。尽管在最长预测时域下数据散布有所增加，但总体上的一致性表明，该模型从热衰减早期到后期均能保持可靠的预测准确度，证实其能够推广至训练数据之外。

图19对比了完整的3–60个月时间窗口内，若干典型留出案例的DCA温度衰减曲线与相应的EINN预测轨迹。在所有示例中，机器学习生成曲线（红色）与参考DCA曲线（蓝色）高度重合，既吻合早期陡峭的冷却阶段，也吻合后期逐渐平缓的变化趋势。仅观察到轻微偏差，这表明方程信息网络能够再现衰减曲线完整的时间形状，而非仅拟合个别数据点。不同衰减曲线均呈现出良好的视觉一致性，这证实模型在保持物理曲线结构的同时，能够提供准确的数据驱动预测，适用于代理预测。

总体而言，结果表明，所提出的方法能够高保真地复现参考温度衰减曲线。表现最佳的模型形式的典型绝对误差约为 3 °C，且在整个测试集上保持高度一致，说明该方法适用于需要快速、可靠地进行曲线表征的前瞻性应用。

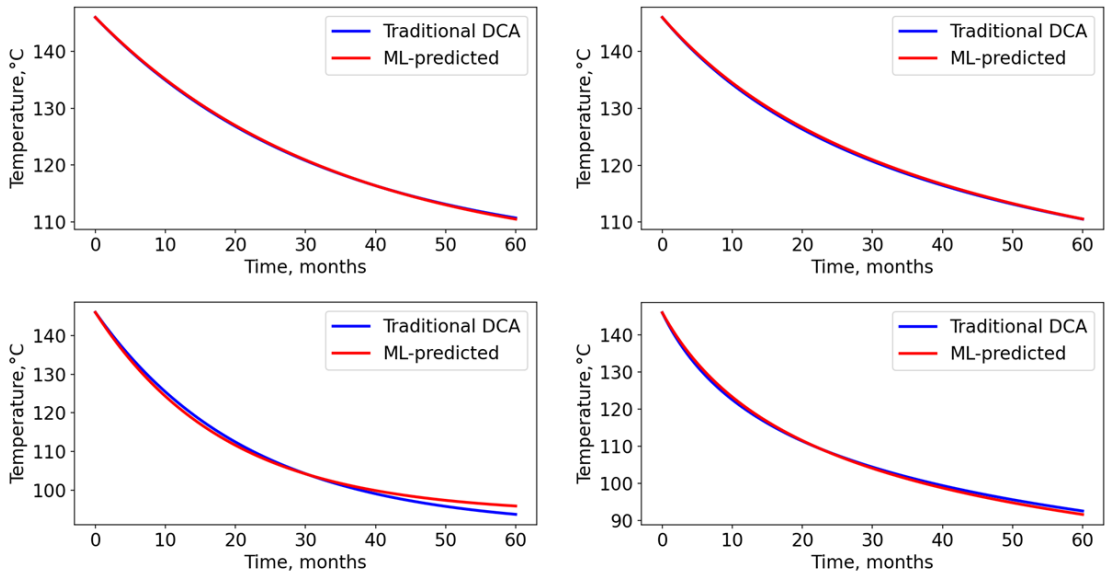


图***19***。 代表性留出测试案例在*3–60*个月期间观测到的（基于*DCA*）与*EINN*预测的温度轨迹。

6.3. 高斯过程回归

本分析采用组感知设计，评估模型在不同案例上的泛化能力。严格按照案例标识划分，预留了一个包含 22 个案例的未见留出集；其余案例则通过分组交叉验证进行折外评估。目标值为每隔 3 个月采集一次的月温度，覆盖 3 至 60 个月（共 20 个预测时域）。这种设计可减少相关观测值之间的信息泄漏，并从两个互补角度评估性能：交叉验证训练集的折外（OOF）指标，以及仅进行一次的留出集评估。

在交叉验证训练评估中，所有20个预测时域的宏观性能为决定系数 = 0.934（标准差 0.030）、均方根误差 = 3.67（标准差 0.26）°C 和平均绝对误差 = 2.57（标准差 0.22）°C，这些数值概括了对各目标值进行汇总后得到的折级均值（表7）。性能随预测时域延长而单调提升：决定系数从3个月时的 ≈ 0.819 升至60个月时的 ≈ 0.951，且各预测时域的误差幅度逐步减小。这些折外结果表明，在训练数据分区内拟合质量稳定，同时保留了组间分离。

表 *7*。 *GPR* 代理模型在抽样预测时距（*3–60*个月）上的交叉验证折外（*OOF*）性能。*R2*、*RMSE* 和 *MAE* 的数值以分组折上的均值 ± 标准差表示。宏观行汇总了所有预测时域上的平均性能。

预测时域（月）	决定系数（均值 ± 标准差）	RMSE（°C）（均值 ± 标准差）	MAE（°C）（均值 ± 标准差）
3	0.819 ± 0.290	2.679 ± 1.186	1.771 ± 0.615
9	0.931 ± 0.052	3.487 ± 1.232	2.363 ± 0.678
15	0.936 ± 0.063	3.765 ± 1.090	2.591 ± 0.516
21	0.940 ± 0.075	3.794 ± 1.171	2.676 ± 0.605
27	0.944 ± 0.075	3.739 ± 1.245	2.629 ± 0.674
33	0.946 ± 0.074	3.719 ± 1.298	2.597 ± 0.741
39	0.948 ± 0.074	3.691 ± 1.346	2.595 ± 0.811
45	0.946 ± 0.074	3.871 ± 1.516	2.763 ± 1.087
51	0.948 ± 0.072	3.815 ± 1.487	2.743 ± 1.048
57	0.950 ± 0.070	3.741 ± 1.447	2.697 ± 0.971
60	0.951 ± 0.069	3.712 ± 1.428	2.680 ± 0.932
宏观	0.934 ± 0.030	3.669 ± 0.262	2.573 ± 0.220

图20展示了四个具有代表性的预测时域（15、30、45和60个月）下，训练案例的高斯过程回归预测温度与真实温度的一致性图（基于折外结果）。所有预测时域的预测结果均与1:1参考线高度吻合。模型性能随时间保持一致，决定系数始终较高（决定系数 = 0.97-0.98），均方根误差则稳定在约3.9-4.2 °C的狭窄范围内。未观察到明显的、随预测时域变化而产生的性能下降；相反，从早期到后期的散布程度相近，表明高斯过程回归代理模型能够以稳定的精度捕捉短期和长期热行为。较低温度下的轻微离散反映的是具体案例的变异性，而非系统误差；总体趋势在所有预测时域内均得到良好保留。

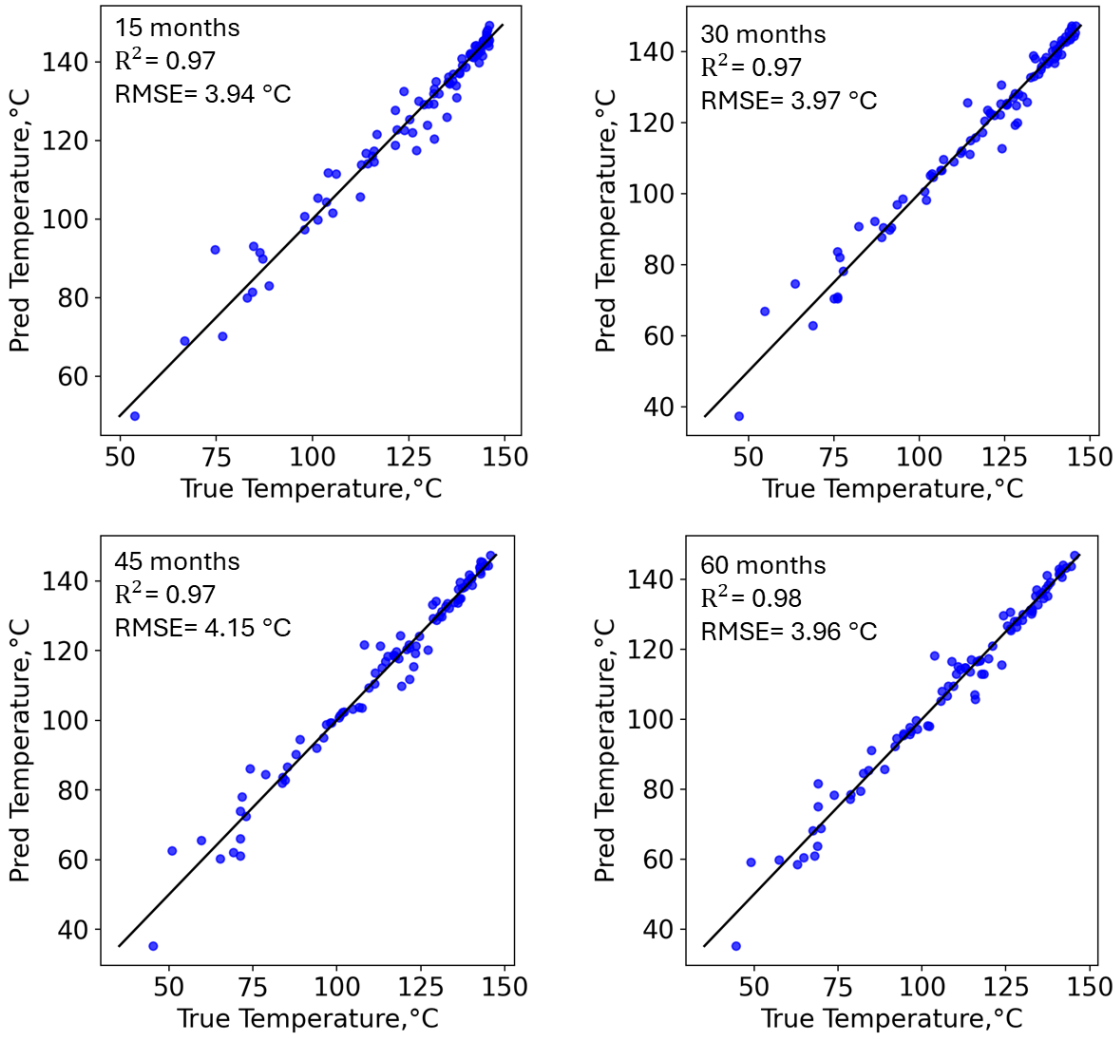


图***20***。 使用高斯过程回归代理模型，对*88*个训练案例在四个具有代表性的预测时域（*15*、*30*、*45*和*60*个月）的预测温度与真实温度进行比较的一致性图。每个面板均显示*1:1*参考线，以及对应预测时域的决定系数和均方根误差。

在留出评估中，20个预测时域的宏观性能共计440次目标评估，决定系数达到 = 0.965（跨预测时域标准差为0.013），均方根误差为 = 3.39（标准差为0.24）°C，平均绝对误差为 = 2.34（标准差为0.20）°C（表8）。各预测时域的表现均十分出色：最低决定系数出现在3个月时（决定系数 = 0.933），随后稳步上升，在60个月时达到最高值（决定系数 = 0.976）。误差始终处于较窄范围内，均方根误差从2.45 °C（3个月）升至21个月左右的峰值3.67 °C，之后趋于稳定，在较远预测时域维持在3.38 °C附近。平均绝对误差呈现类似趋势，各预测时域的平均值为2.34 °C。综合来看，留出评估指标略优于折外估计，这与稳定的泛化能力一致。

对留出残差中各目标值的离散度（绝对误差的标准差）进行检查后发现，不同预测时域的结果总体一致，说明预测可靠性随时间具有同质性。由于目标值采用相同尺度进行测量，RMSE 和 MAE 随预测时域变化的趋势近乎平行，这表明误差分布并非由离群值主导。此外，这一模式与较晚月份的 R^2值稳步上升相结合，表明预测信号会随预测时域延长而增强，同时绝对误差仍保持在合理范围内。图21展示了高斯过程回归代理模型在折外验证和留出测试集中的20个月度预测时域（3–60个月）内，随时间范围变化的误差概况（RMSE 和 MAE）。

表*8*。高斯过程回归代理模型在抽样预测时距（*3–60*个月）上的留出测试性能。对于每个时距的全部*22*个测试案例，指标均以均值 ± 标准差的形式报告，涵盖决定系数、均方根误差和平均绝对误差。宏观行总结了所有时距上的平均性能。

预测时域（月）	N	R^2（均值 ± 标准差）	RMSE（°C）（均值 ± 标准差）	MAE（°C）（均值 ± 标准差）
3	22	0.933	2.446 ± 2.327	1.803 ± 1.653
9	22	0.947	3.381 ± 3.376	2.397 ± 2.384
15	22	0.955	3.613 ± 3.605	2.729 ± 2.368
21	22	0.961	3.667 ± 3.650	2.584 ± 2.602
27	22	0.971	3.331 ± 3.325	2.345 ± 2.366
33	22	0.972	3.388 ± 3.367	2.344 ± 2.446
39	22	0.973	3.419 ± 3.376	2.311 ± 2.519
45	22	0.975	3.413 ± 3.349	2.266 ± 2.553
51	22	0.975	3.403 ± 3.325	2.234 ± 2.567
57	22	0.976	3.384 ± 3.298	2.193 ± 2.578
60	22	0.976	3.381 ± 3.290	2.172 ± 2.592
宏观	440	0.965 ± 0.013	3.386 ± 0.242	2.337 ± 0.202

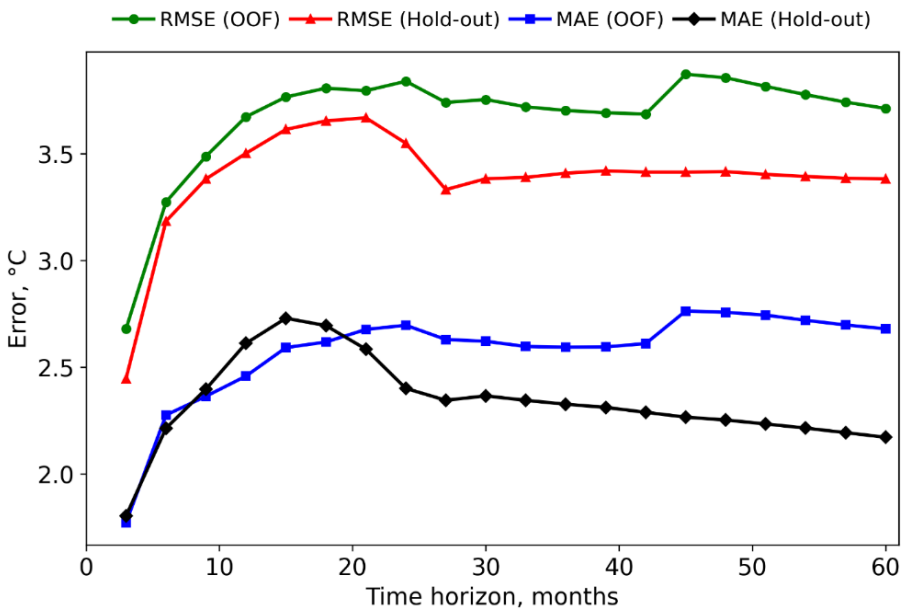


图***21***。高斯过程回归代理模型在不同时间范围内的误差概况。图中展示了*20*个月度预测时域（*3–60*个月）内的折外（*OOF*）验证集和留出测试集的均方根误差与平均绝对误差。误差幅度随时间保持稳定，表明预测性能始终如一。

图22比较了高斯过程回归在四个有代表性的预测时域（15、30、45和60个月）上预测的温度与真实留出值。所有时距上的数据点都紧密聚集在1:1线附近，表明校准效果良好且系统偏差较小。决定系数随预测时域略有增加（决定系数 = 0.955:0.976），而均方根误差稳定在约3.4 °C，证实预测准确度在较晚时域并未下降。这种一致性表明，即使向生产期更远的时段外推，代理模型仍具有可靠的泛化能力。图23展示了具有代表性的留出井的完整温度下降轨迹，对比了所有20个预测时距（3–60个月）内的真实观测值与高斯过程回归预测值。在整个生产期内，预测曲线都紧密跟随观测到的变化，而较窄的95 %预测区间表明认知不确定性较低。

这些结果表明，模型能够对所有目标月份做出准确且校准良好的预测，绝对误差保持稳定，且较长预测时域上的解释方差不断提高。OOF性能与留出测试集性能相吻合，再结合散点图与叠加诊断，说明模型对未见样本具有良好的泛化能力，同时保留了单条轨迹层面的可解释性。

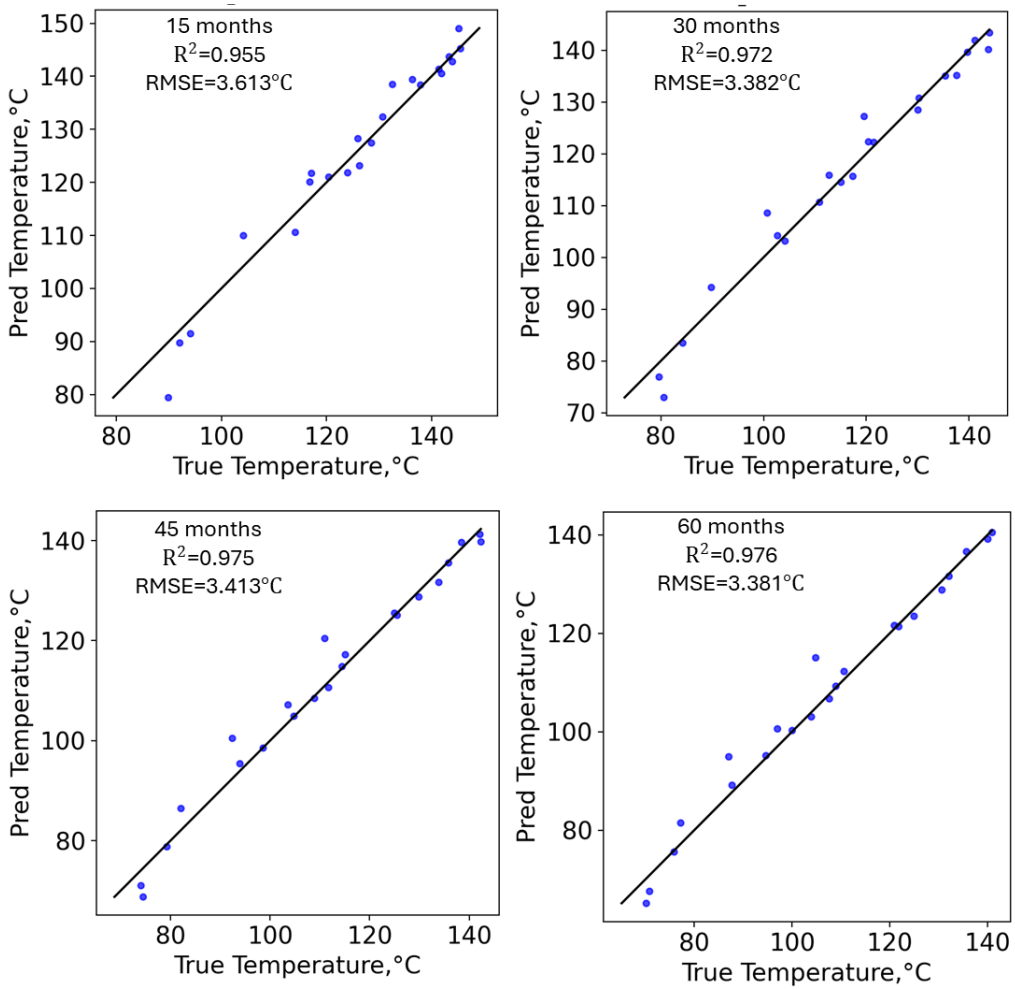


图*22*。 使用高斯过程回归代理模型，在四个具有代表性的预测时域（*15*、*30*、*45* 和 *60*个月）下，*22*个留出案例的预测温度与观测温度对比。每个面板均显示一一对应参考线，以及相应预测时域的决定系数和均方根误差。该模型在早期和后期预测时域均保持了良好的校准效果。

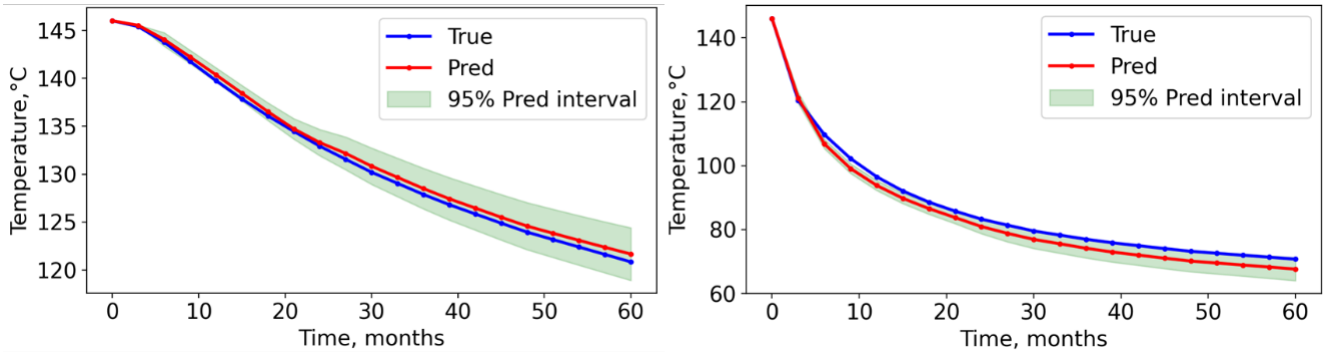


图*23*。 所选留出案例的观测温度演化与高斯过程回归预测的温度演化。实线表示跨越*20*个预测时域（*3–60*个月）的真实温度和预测温度，阴影带表示*95%*的预测区间/置信区间。

6.4. XGBoost代理模型

本节介绍 XGBoost 模型的性能。数据集按 THM 案例划分为两个互不重叠的子集，确保同一案例中的所有观测值仅被分配到训练集或测试集中的一方。20% 的案例（22 个案例）被保留为完全未见的测试集，其余 88 个案例则组成训练集。此划分确保所报告的结果反映模型在未见样本上的表现。

在所有时间点，误差指标（均方根误差和平均绝对误差）均在较早的预测时域上升，随后趋于稳定（图 24）。从最早的预测时域（3个月）到约 12–18 个月，训练集和测试集的均方根误差及平均绝对误差均有所上升，表明随着预测时域从短期延长至年度尺度，预测准确度会下降。约 20 个月后，曲线趋于平缓，说明进一步延长预测时域对误差的影响相对较小。

在每个预测时域上，训练误差（通过多个交叉验证折计算并取平均）始终高于测试误差。训练集均方根误差起初约为 7.2 °C，在 24 个月左右升至最高的 9.5 °C，随后缓慢下降，到 60个月时约为 9.0 °C。训练集平均绝对误差起初约为 4.7 °C，到 24 个月时升至 7.2 °C，并在最长的预测时域内保持在 7.1 °C 左右。这一趋势表明，预测时域达到约两年后，训练误差的典型幅度基本不再随时间变化，仅有轻微波动。

测试性能呈现出相同的先上升后趋于平台的特征，但误差水平更低。测试均方根误差从最短预测时域的约 5.5 °C 增加至 18 个月左右的约 7.2 °C，随后趋于稳定，并在 60 个月时小幅改善至约 6.9 °C。测试平均绝对误差从约 3.7 °C 上升至 21-33 个月左右的 5.8 °C，随后逐渐下降，在最长预测时域时约为 5.4 °C。测试平均绝对误差在较晚预测时域的小幅下降（以及测试均方根误差的轻微回落）表明，误差趋于饱和，甚至可能在多年预测时域内有所改善。

随着预测时域延长，训练集曲线与测试曲线之间的差距并未扩大，这反映的是稳定的偏移，而非随预测时域变化的发散。总体而言，结果表明有三个主要特征：（1）预测时域延长带来的大部分性能退化发生在早期（最多20个月）；（2）多年预测时域呈现明显的误差平台；（3）汇总后的训练集误差始终系统性高于测试误差，而测试误差则相对稳定，并在最长预测时域下略有改善。

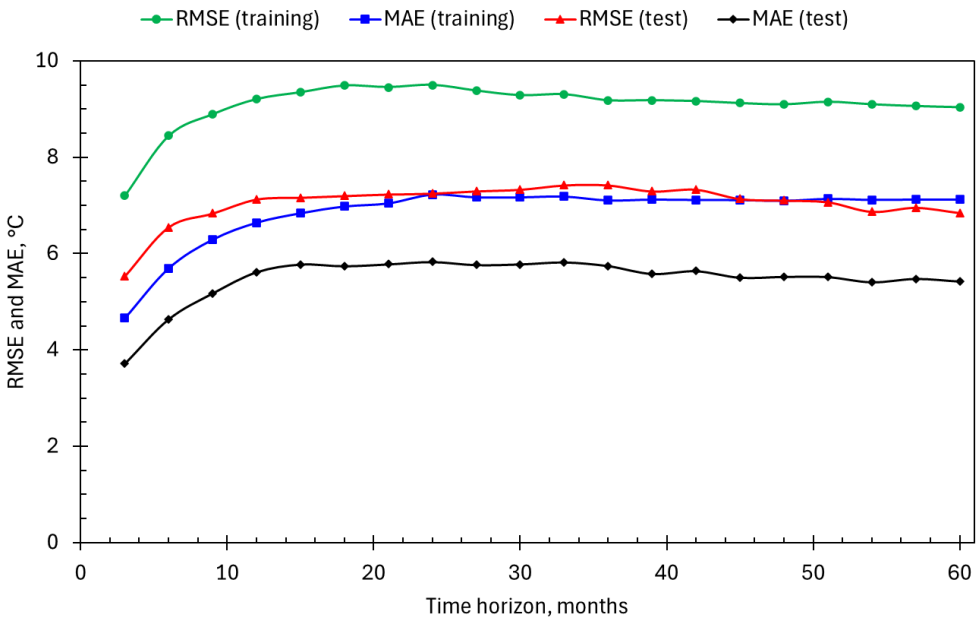


图24。 *XGBoost* 模型的均方根误差（*RMSE*）和平均绝对误差（*MAE*）随预测时域（月）的变化情况，展示汇总训练结果和测试集的结果。

训练数据一致性图见图25。图中显示，在所有预测时域（15、30、45和60个月）内，预测温度与真实温度始终保持很强的线性一致性。每个面板中的大多数点都聚集在1:1线附近，表明模型捕捉到了主要的温度信号，并生成了高度相关的估计值。报告的决定系数随时间范围延长而提高，从R2 = 0.80（15个月）升至0.84（30个月）、0.86（45个月）和0.87（60个月）。这一逐步上升表明，在训练集内，较长预测时域下的解释方差略有提升，其中45至60个月的预测时域整体吻合度最高。误差幅度也保持稳定，并随预测时域延长略有改善。报告的均方根误差从9.71 °C（15个月）降至9.55 °C（30个月）、9.37 °C（45个月）和9.31 °C（60个月）。这种变化是渐进的，而非显著的，说明无论预测时域如何，典型的训练集偏差仍在9–10 °C左右，但随着预测时域延长略有系统性降低。

数据点分布表明，在高温区间（约 120–150 °C）内，预测保真度最高，数据点云密集且紧贴 1:1线。相比之下，低温区域（约 50–90 °C）的散布更广，且有几处明显偏离一致性线的情况。较低真实温度下的准确度似乎更不稳定，而高温则能更一致地再现。

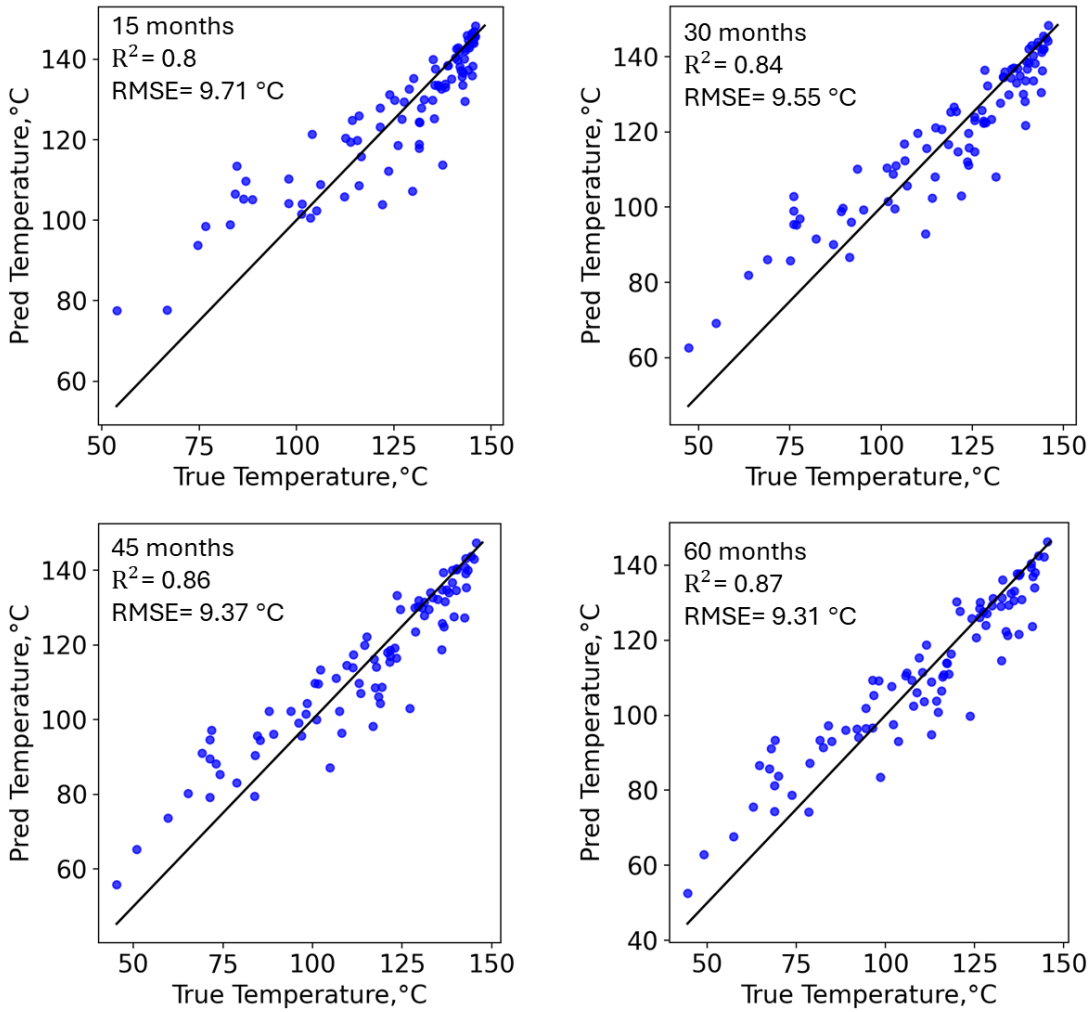


图***25***。*XGBoost*在*15*、*30*、*45*和*60*个月时的训练集预测值一致性图。图中展示了预测温度与真实温度的对比。每个面板均显示一一对应参考线，并标注了各预测时段对应的决定系数和均方根误差。

图26展示了四个预测时域（15、30、45、60个月）的留出测试散点图（预测温度与真实温度对比）。在15个月时，两者关系较强（决定系数 = 0.83，均方根误差 = 7.17 °C）。大多数数据点聚集在较高温度下的1:1线附近，表明高温范围内的准确度良好。最大差异出现在中低真实温度范围内。到30个月时，解释方差有所提高（决定系数 = 0.87），而误差幅度基本相近（均方根误差 = 7.33 °C）。决定系数提高而均方根误差相近，这意味着各样本的整体排序和线性关系更好。在45个月和60个月时，两者的一致性更加稳定。在45个月时（决定系数 = 0.89，均方根误差 = 7.14 °C）和60个月时（决定系数 = 0.90，均方根误差 = 6.85 °C），在大部分温度范围内，数据点散布与1:1线更加吻合，且较大的残差出现得较少。60个月面板显示出整体上最佳的平衡：决定系数最高，且

在四个预测时域中均方根误差最低，表明对较长预测时域的温度预测精度略高，且与真实值的对应关系更稳定，但仍有少数数据点与完全吻合存在明显偏差。总体而言，这些观测结果与图24一致：较短预测时域的预测通常较为可靠，但仍存在少数影响显著的误差。图27展示了具有代表性的测试集轨迹，对比了若干未见样本在完整的0-60个月预测时域内的预测温度演化与真实温度演化。在每个示例中，模型都能较好地再现总体降温趋势和时间形状，预测结果在整个时段内通常都能跟随真实曲线。

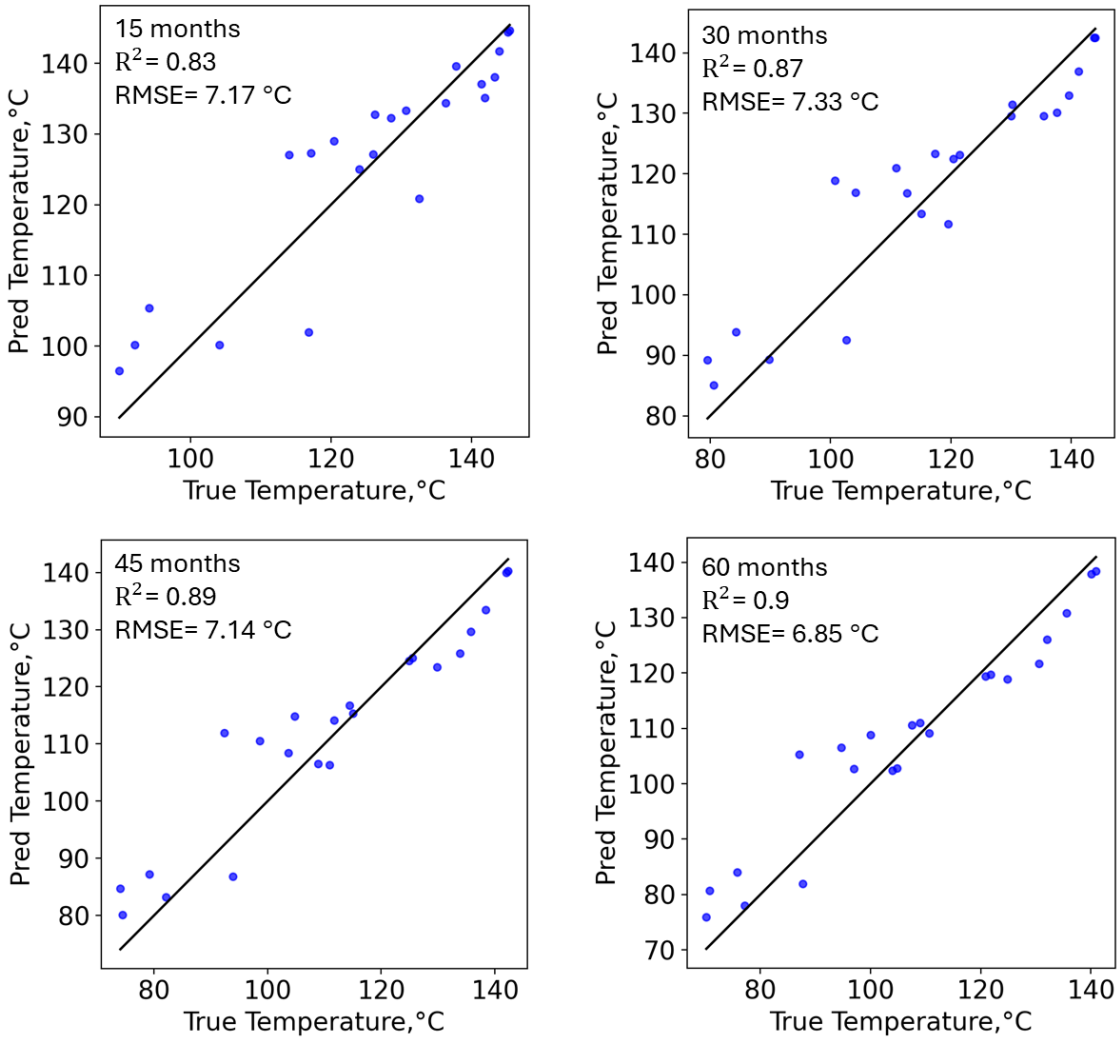


图***26.*** *XGBoost* 在 *15*、*30*、*45* 和 *60*个月时的测试集预测一致性图。图中展示了预测温度与真实温度相对于 *1:1*线 的关系。每个面板均显示一一对应参考线，以及各预测时距对应的决定系数和均方根误差。

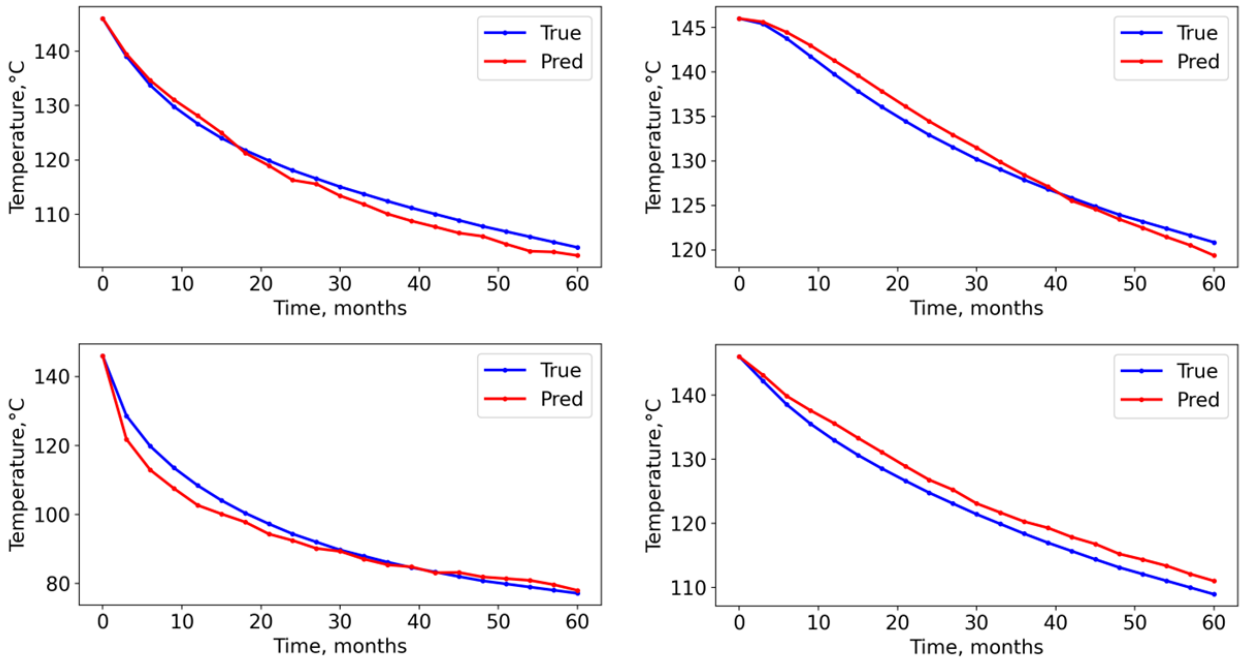


图*27. XGBoost*测试集图示例：对比预测时间范围为*0–60*个月的单个未见样本的真实温度轨迹与预测温度轨迹。

7. 讨论

由于数据集中的每次热-水-力学模拟都需要求解全耦合流动、热传输和力学变形，因此在标准个人工作站（Intel i7-10750H、16 GB RAM、Windows 11）上运行一次高保真模拟需要约6–8小时。这项计算开销源于热-孔弹性耦合，以及捕捉裂缝控制的热传输所需的精细空间和时间分辨率。相比之下，训练完成后，方程信息神经网络和高斯过程代理模型几乎可以即时返回完整的0–60个月温度预测，相对于底层模拟器所需的时间微不足道。这使用户能够评估新的配置、开展敏感性分析并探索设计方案，而无需承担反复进行热-水-力学模拟所带来的高昂计算开销。

递减曲线公式的平衡温度扩展通过强制满足由与周围地层热交换所决定的有限长期行为，为地热冷却提供了符合物理规律的表征。与犹他州 FORGE 项目的井下温度测量进行验证后发现，修正递减模型比传统的石油行业公式更可靠地捕捉到了观测到的地热温度趋势，尤其是在渐近行为变得重要的后期。尽管现有现场记录的持续时间有限，但它是目前可公开获取的、针对在产增强型地热系统生产井且适用于

递减曲线验证的最长连续井下温度数据集。在此背景下，修正递减曲线与现场观测值的一致性，为所提出的公式提供了重要的经验支持，并增强了人们对其作为符合物理规律的诊断和预测工具的信心。此外，当平衡温度设为零时，修正公式会精确退化为传统 Arps 递减方程，从而与经典的基于石油的递减分析完全向后兼容。此外，将修正模型应用于各个模拟案例时，其与温度数据高度吻合，误差极小。

本研究考察的机器学习模型各有互补作用，其主要区别在于如何融入物理结构和不确定性。方程信息神经网络（EINN）将地热扩展递减曲线方程直接嵌入学习架构，确保温度轨迹形状符合物理规律，并生成可靠的全曲线预测，典型测试误差为平均绝对误差 = 3.06 °C 和均方根误差 = 4.49 °C。高斯过程回归（GPR）作为直接多预测时域概率代理模型实现，在各预测时域上取得了整体最佳的预测准确度和稳定性（留出集上的宏观均方根误差为 = 3.39 °C，平均绝对误差为 = 2.34 °C），同时还能明确量化不确定性，因此非常适合快速筛查和风险感知决策。相比之下，未嵌入物理结构或不确定性的直接 XGBoost 模型虽能再现总体趋势，但误差较高（均方根误差约为 7 °C），凸显了融入物理规律或概率结构的价值。尽管模型开发所用的实现次数有限，所有模型在未见样本上的表现仍然稳定，这表明，支配地热温度递减的主要物理因素在数据集中表现显著，并能通过代理模型公式得到有效捕捉。

Yan 等人 [33] 近期提出了 HyperReLU 热递减模型，并训练了一个全连接神经网络，根据热-水-力模拟输入回归其三个递减参数；该方法通过解析递减形式以及损失函数中的基于方程的正则化项引入物理信息。本研究在两个根本方面有所不同。首先，通过引入与牛顿冷却定律一致的平衡温度项，将经典 Arps 递减族推广至地热温度预测；这一做法在保持后期热行为有限的同时，当平衡温度项设为零时，可精确还原为传统 Arps 方程。其次，本研究并非仅将物理信息作为针对回归参数的外部损失惩罚项，而是将修正后的递减方程嵌入方程信息神经网络内部，作为可微计算层，从而确保训练过程中学习信号能够通过解析递减物理传播。此外，扩展后的递减形式还通过犹他州 FORGE 井下温度测量进行了验证，并在同一个受控热-水-力数据集上，与概率（高斯

过程）代理模型和直接数据驱动（XGBoost）代理模型进行基准比较，从而能够客观比较物理引导结构与不确定性感知及黑箱（纯数据驱动）替代方案。

这一进展为递减分析与现代数据驱动方法搭建了实用的桥梁。这意味着用户既能利用快速、基于情景的机器学习预测，又能保留结果的物理意义。这些模型共同构成了一套灵活的工具：DCA 用于逐案例深入分析，高斯过程回归用于快速筛查并量化不确定性，而方程嵌入模型则用于提供符合物理规律的预测，以支持设计优化和敏感性分析。

8. 结论

从石油工程中移植而来的传统递减曲线分析往往无法准确描述地热温度变化，因为它忽略了与围岩之间的热交换以及有限的平衡温度；与此同时，高保真耦合模拟器在针对众多增强型地热系统设计进行快速筛查时，计算成本仍然高昂。

本研究通过为完整的Arps递减模型加入与牛顿冷却定律一致的平衡温度项，建立了物理引导预测框架。这一修正确保模型在晚期呈现适用于地热系统的有限值行为；当平衡温度设为零时，模型则精确退化为经典Arps方程，从而保持向后兼容性。

在受控的热-水-力数据集上，扩展后的递减模型能够以近乎完美的精度再现温度衰减（决定系数中位数 = 0.999；均方根误差中位数 = 0.071 °C）；其中，拉伸指数形式最常被选用（60%），其次是双曲线形式（37%）。与犹他州 FORGE 温度测量结果进行的现场验证表明，修正公式能够捕捉观测到的衰减，同时确保平衡行为有限。

基于这些基于物理的曲线，方程信息神经网络将修正递减方程嵌入内部计算层，根据五项设计与运行输入预测递减参数，并生成保留解析结构的完整 0–60 个月温度剖面。典型的留出误差为平均绝对误差 = 3.06 °C 和均方根误差 = 4.49 °C，表明该方法能够可靠地进行全轨迹预测，同时保持可解释性。

直接多时域高斯过程回归代理模型可快速提供从3个月到60个月、每3个月一次的不确定性感知预测，并在留出案例上展现出稳健的泛化能力（宏观决定系数 ≈ 0.965；均方根误差 ≈ 3.39 °C；平均绝对误差 ≈ 2.34 °C）；其预测区间可量化各预测时域的不确定性。

直接 XGBoost 模型作为确定性的非物理信息基线：它能够重现总体趋势，但误差较大（均方根误差约为 7 °C），这进一步凸显了嵌入物理结构或进行不确定性建模对于可靠地热预测的价值。

符号说明

B_f：流体热膨胀系数，1/°C B_s：固体热膨胀系数，1/°C

B：斯肯普顿孔隙压力系数 c_f：流体压缩性，psi^{-1}。（1 psi^{-1} = 1.450 $\times 10^{-4}$ Pa^{-1}）。

C_T：热扩散率，m^2/s D：递减率常数 G：剪切模量，Pa k：储层渗透率，md k_T：岩石热导率，W. m^{-1}. °C^{-1} p_r：储层孔隙压力，Pa T(t)：时间 t 的温度，°C T^{inj}：注入温度，°C T_{res}：地层温度，°C T0：初始温度，°C T_∞：渐近温度，°Cν_u：不排水泊松比ν：排水泊松比σ_x：x 方向的主应力，Paσ_{xy}：剪切应力，Paσ_y：y 方向的主应力，Pa

α：Biot系数μ：流体黏
度，cpϕ：储层孔隙度，
%

参考文献

[1] Soltani M, Moradi Kashkooli F, Dehghani-Sanij ARR, Nokhosteen A, Ahmadi-Joughi A, Gharali K 等。地热能演变与发展的全面综述。国际绿色能源期刊 2019年；16：971–1009。https://doi.org/10.1080/15435075.2019.1650047。[2] Nath F, Mahmood MN, Ofosu E, Khanal A。增强型地热系统：近期进展及清洁能源生产未来潜力的批判性综述。地热能源科学与工程 2024年；243：213370。https://doi.org/10.1016/j.geoen.2024.213370。[3] Lu SM。增强型地热系统（EGS）全球综述。可再生与可持续能源评论 2018年；81：2902–21。https://doi.org/10.1016/j.rser.2017.06.097。[4] Franzmann D, Heinrichs H, Stolten D。全球地热发电潜力：技术、经济及热可再生性评估。Renew Energy 2025年；250：123199。https://doi.org/10.1016/j.renene.2025.123199。[5] McClure MW, Horne RN。增强型地热系统增产机制研究。国际岩石力学与矿业科学期刊 2014年；72：242–60。https://doi.org/10.1016/j.ijrmms.2014.07.011。[6] Arps JJ。递减曲线分析。美国矿业、冶金及石油工程师学会会刊 1945年；160：228–47。https://doi.org/10.2118/945228- G。[7] Ripperda M, Bodvarsson GS。盖瑟斯地热田生产数据的递减曲线分析。第12届地热储层工程研讨会 1987年：23–7。[8] Reyes JLP, Horne RN, Li K。一种应用于盖瑟斯的新型递减曲线分析方法。第二十九届地热储层工程研讨会论文集，加利福尼亚州斯坦福斯坦福大学，2004年1月26日至28日，SGP-TR-175 2004年。[9] Hidayat I。液体主导型地热储层的产量预测与优化递减曲线分析。IOP地球与环境科学会议系列 2016年；42：012024。https://doi.org/10.1088/1755-1315/42/1/012024。[10] Li F, Guo X, Qi X, Feng B, Liu J, Xie Y 等。一种基于代理模型的优化方法，用于

采用正则化 LSTM-CNN 模型和灰狼优化器进行地热井双井组布置。Sustain 2025年；17。https://doi.org/10.3390/su17010266。[11] Duplyakin D, Beckers KF, Siler DL, Martin MJ, Johnston HE。利用机器学习对地热储层地下性能进行建模。Energies 2022年；15：967。https://doi.org/10.3390/en15030967。[12] Degen D, Cacace M, Wellmann F。探索基于物理的机器学习在地热应用中的应用。第49届地热储层工程研讨会 2024年：1–9。[13] Khalaf MS。增强型地热系统的全面热-水-力耦合框架：热刺激、能源回收与天然裂缝活化。Unconv Resour 2026年；9：100281。https://doi.org/10.1016/j.uncres.2025.100281。[14] Han S, Cheng Y, Gao Q, Yan C, Zhang J。多级压裂增强型地热系统热提取性能的数值研究。Renew Energy 2020年；149：1214–26。https://doi.org/10.1016/j.renene.2019.10.114。[15] Pruess K。以二氧化碳为工作流体的增强型地热系统（EGS）——一种通过同时封存碳来生产可再生能源的新方法。Geothermics 2006年；35：351–67。https://doi.org/10.1016/j.geothermics.2006.08.002。[16] Fetkovich MJ, Vienot ME, Bradley MD, Kiesow UG。利用类型曲线进行递减曲线分析——案例史。SPE Form Eval 1987年；2：637–56。https://doi.org/10.2118/13169-PA。[17] Fetkovich MJ。利用类型曲线进行递减曲线分析。J Pet Technol 1980年；32：1065–77。https://doi.org/10.2118/4629-PA。[18] Ilk D, Currie SM, Blasingame TA。致密气井和页岩气井的产量分析与井性能预测。SPE东部地区会议 2010年：451–65。https://doi.org/10.2118/139118-ms。[19] Okouma V, Symmons D, Hosseinpour-Zonoozi N, Ilk D, Blasingame TA。非常规储层递减曲线分析的实践考量——近期开发的时间-产率关系的应用。SPE油气经济评价研讨会 2012年：414–74。https://doi.org/10.2118/162910-ms。[20] Ilk D, Rushing JA, Blasingame TA。通过参数相关关系整合产量分析与产率-时间分析——理论考量与实际应用。Soc Pet Eng - SPE水力压裂技术会议 2011年 2011年：738–50。https://doi.org/10.2118/140556-ms。[21] Tan L, Zuo L, Wang B。页岩气储层递减曲线分析方法。Energies 2018年；11。

https://doi.org/10.3390/en11030552。

[22] Yuan H, Soar L, Packer R, Bhuiyan M, Xu J. 基于实际油井生产数据评估页岩油产能模型的案例研究。Unconv Resour Technol Conf 2013, URTC 2013 2013. https://doi.org/10.1190/urtec2013-254.[23] Wachtmeister H, Lund L, Aleklett K, Ho o k M. 鹰福特页岩致密油井的产量递减曲线。Nat Resour Res 2017;26:365–77. https://doi.org/10.1007/s11053-016-9323-2.[24] Konwar L, Hamad E, Barni S, Nemmwi N, Shamlooh M, Garadah A, et al. 通过整合多种递减曲线方法和模拟研究，评估巴林致密气藏的产能与采收率。Soc Pet Eng - SPE Conf Oman Pet Energy Show, OPES 2025 2025:1– 26. https://doi.org/10.2118/225094-MS.[25] Clark AJ, Lake LW, Patzek TW. 使用逻辑斯蒂增长模型进行产量预测。Proc - SPE Annu Tech Conf Exhib 2011;1:184–94. https://doi.org/10.2118/144790-ms.[26] Shi Y, Liu C, Song X, Yan S. 基于傅里叶神经算子的裂缝型地热储层温度场预测模型：应对不同裂缝形态和注采参数。Geothermics 2025;131:103404. https://doi.org/10.1016/j.geothermics.2025.103404.[27] Haffou FZ, Ouzzaouit LA, Qarbous A, Boudad L. 利用摩洛哥北部的水文地球化学数据预测地热储层温度：一种机器学习方法。Geothermics 2025;127:103259. https://doi.org/10.1016/j.geothermics.2025.103259.[28] Ibrahim B, Konduah JO, Ahenkorah I. 预测土耳其安纳托利亚西部地热系统的储层温度：聚焦机器学习模型的预测性能与可解释性。Geothermics 2023;112:102727. https://doi.org/10.1016/j.geothermics.2023.102727.[29] Chen G, Jiao JJ, Liu Q, Wang Z, Jin Y. 机器学习加速的裂缝型地热系统多目标设计。Nexus 2024;1:100044. https://doi.org/10.1016/j.ynexs.2024.100044.[30] Chen G, Jiao JJ, Jiang C, Luo X. 用于地热开采优化的代理辅助基于层级学习的进化搜索。Renew Sustain Energy Rev 2024;189:113860. https://doi.org/10.1016/j.rser.2023.113860.[31] Al-Fakih A, Abdulraheem A, Kaka S. 机器学习与深度学习在地热资源开发中的应用：趋势与展望。Deep Undergr Sci Eng 2024;3:286–301. https://doi.org/10.1002/dug2.12098.

[32] Karimpouli S, Kwiatek G, Martí nez-Garzo n P, Caus D, Wang L, Dresen G, et al. 使用机器学习预测增强型地热系统中的诱发地震活动：挑战与机遇。Geophys J Int 2025;242:1–13. https://doi.org/10.1093/gji/ggaf155.[33] Yan B, Gudala M, Sun S. 基于广义热衰减模型和深度学习的地热采收稳健优化。Energy Convers Manag 2023;286:117033. https://doi.org/10.1016/j.enconman.2023.117033.[34] Chen J, Jiang F. 设计增强型地热系统的多井布局，以更好地开发干热岩地热能。Renew Energy 2015;74:37–48. https://doi.org/10.1016/j.renene.2014.07.056.[35] Hofmann H, Babadagli T, Yoon JS, Blo cher G, Zimmermann G. 针对花岗岩基底中增强型地热系统（EGS）复杂水力压裂扩展的混合离散/有限元建模研究。Geothermics 2016;64:362–81. https://doi.org/10.1016/j.geothermics.2016.06.016.[36] Khalaf MS, Rezaei A, Soliman MY, Farouq Ali SM. 地热储层冷流体注入的热-孔弹性分析：取热可持续性。本文为 ARMA 22–0453，准备在2022年6月26日至29日于美国新墨西哥州圣菲举行的第56届美国岩石力学/地质力学研讨会上发表。https://doi.org/10.56952/ARMA-2022-0453.[37] Wang Y, Voskov D, Khait M, Bruhn D. 一种高效的地热模拟数值模拟器：基准研究。Appl Energy 2020;264:114693. https://doi.org/10.1016/j.apenergy.2020.114693.[38] McTigue DF. 流体饱和多孔岩石的热弹性响应。J Geophys Res Solid Earth 1986;91:9533–42. https://doi.org/10.1029/jb091ib09p09533.[39] Ravikumar A, Lee J. 使用正则化先验稳健估计ARPS衰减参数。SPE East Reg Meet 2019;2019-Octob:15–7. https://doi.org/10.2118/196576-ms.[40] Zhang H, Rietz D, Cagle A, Cocco M, Lee J. 扩展指数递减曲线分析。J Nat Gas Sci Eng 2016;36:402–13. https://doi.org/10.1016/j.jngse.2016.10.010.[41] Davidzon MI. 牛顿冷却定律及其解释。Int J Heat Mass Transf 2012;55:5397– 402. https://doi.org/10.1016/j.ijheatmasstransfer.2012.03.035.[42] da Silva SLEF. 广义统计力学中的牛顿冷却定律。Phys A Stat Mech Its Appl 2021;565:125539. https://doi.org/10.1016/j.physa.2020.125539.

[43] Shabib-Asl A、Plaksina T。基于赤池信息准则为非常规储层选择递减曲线分析模型。石油科学与工程杂志 2019年；182:106327。https://doi.org/10.1016/j.petrol.2019.106327。[44] McLennan J、Hughes B。犹他州 FORGE：16B(78)-32 号井压力-温度测井曲线。OpenEi 地热数据存储库 2024年。https://doi.org/10.15121/2368548。[45] McLennan J、Mock B、Swearingen L、Baldwin R、Hodder M、Vetsak A 等。犹他州 FORGE：16B(78)- 32 号井钻井数据。OpenEi 地热数据存储库 2023。https://doi.org/https://doi.org/10.15121/1998591。