

基于鄂尔多斯盆地西北缘地区烃源岩有机碳含量测井评价方法研究

郭聪聪¹, 何之龙¹, 徐志峰², 王子璇², 宋 姗²

¹中国石油大学(北京)地球物理学院, 北京

²中海油田服务股份有限公司油田技术研究院, 北京

收稿日期: 2025年4月28日; 录用日期: 2025年6月3日; 发布日期: 2025年6月16日

摘 要

烃源岩总有机碳(Total Organic Carbon, TOC)含量是非常规油气藏的重要参数之一, 准确评价TOC含量对非常规油气藏的勘探开发具有重要意义。本文研究了两类TOC评价方法, 一类方法基于物理模型评价烃源岩有机碳含量, 包括 ΔLogR 方法、变基线方法以及双差分方法(DD ΔLogR)。相较于 ΔLogR 方法, 变基线方法改进了基线确定方法, 提高了模型精度; 双差分方法不仅考虑基线问题, 还考虑到岩石矿物成分的变化和孔隙扩大的影响, 在三种基于物理模型的TOC评价方法中效果最优。另一类方法基于机器学习评价烃源岩有机碳含量, 本文对比了BP神经网络、LSTM神经网络、与物理模型结合的BP神经网络、与物理模型结合的LSTM神经网络, 在四种神经网络方法中, 与物理模型结合的LSTM神经网络评价TOC效果最优。本研究建立的综合评价方法体系, 为鄂尔多斯盆地西北缘复杂地质条件下烃源岩有机质含量预测提供了创新性的技术路径。

关键词

TOC评价, ΔLogR 方法, 机器学习

Integrated Logging Assessment of Organic Carbon Content in Source Rocks: Northwestern Ordos Basin Applications

Congcong Guo¹, Zhilong He¹, Zhifeng Xu², Zixuan Wang², Shan Song²

¹College of Geophysics, China University of Petroleum (Beijing), Beijing

²WellTech R & D, China Oilfield Services LTD, Beijing

Received: Apr. 28th, 2025; accepted: Jun. 3rd, 2025; published: Jun. 16th, 2025

文章引用: 郭聪聪, 何之龙, 徐志峰, 王子璇, 宋姗. 基于鄂尔多斯盆地西北缘地区烃源岩有机碳含量测井评价方法研究[J]. 石油天然气学报, 2025, 47(2): 211-222. DOI: 10.12677/jogt.2025.472025

Abstract

The total organic carbon (TOC) content of source rocks is one of the important parameters for unconventional oil and gas reservoirs. Accurate evaluation of the TOC content is of great significance for the exploration and development of unconventional oil and gas reservoirs. This study investigates two types of TOC evaluation methods. For physical-model-based methods, it includes the ΔLogR method, the variable baseline method, and the double difference method (DD ΔLogR). Compared with the ΔLogR method, the variable baseline method improves baseline determination and enhances model accuracy; the double-difference method (DD ΔLogR) further considers rock mineral composition changes and pore expansion, outperforming the other two physical-model-based methods. Regarding machine-learning-based methods, a comparison of BP neural network, LSTM neural network, and their hybrid versions integrated with physical models shows that the LSTM neural network combined with physical models yields the most accurate TOC evaluation. The comprehensive evaluation framework established here offers an innovative solution for predicting source-rock organic matter content in the complex geological environment of the northwestern Ordos Basin, potentially improving resource assessment efficiency and accuracy in similar settings.

Keywords

TOC Evaluation, ΔLogR Method, Machine Learning

Copyright © 2025 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

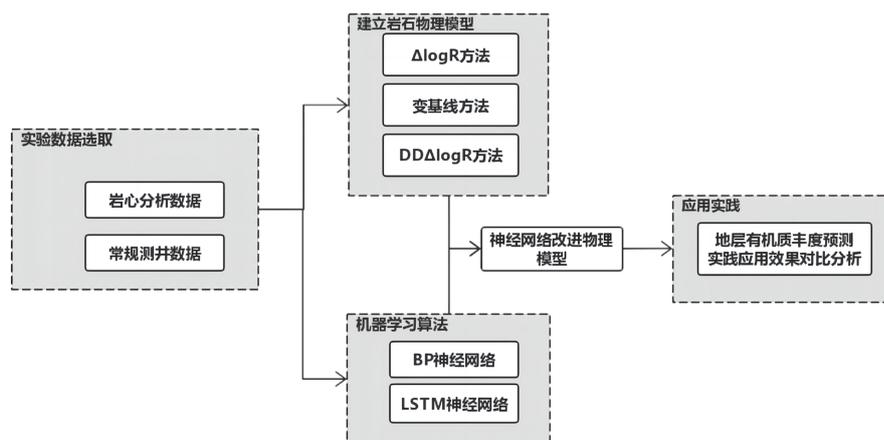


Figure 1. Technical route
图 1. 技术路线图

随着油气勘探开发的不断深入, 常规油气产量持续下降, 人类能源需求持续增加, 全球范围内正在开展页岩气储层、页岩油储层、致密油气储层等非常规储层研究。很多学者针对非常规油气藏进行了大量研究, 页岩储层的有机质含量(TOC)决定了其最终能产生的碳氢化合物数量, 反映了页岩地层的烃类生成能力, 是评估页岩储层质量的重要参数[1]-[5]。TOC 含量可通过地球化学实验确定。然而, 通过岩心

实验测量整个井的 TOC 是不现实的。因此，目前的主流方法是通过岩心实验获得地层 TOC 含量，然后利用测井曲线建立 TOC 评价模型。目前的 TOC 评价方法研究主要集中在两个方面：(1) $\Delta\log R$ 方法及改进 $\Delta\log R$ 方法[6]；(2) 使用机器学习方法来提高 TOC 预测精度[7]-[9]。本论文的研究内容是，利用岩石物理模型与机器学习方法评价烃源岩，提高页岩储层 TOC 含量的预测精度。技术路线如图 1，选取实验区域的岩心分析数据和常规测井数据，建立 $\Delta\log R$ 、变基线、DD $\Delta\log R$ 三种岩石物理模型和 BP 神经网络、LSTM 神经网络两种机器学习算法，分别对目标区域 TOC 进行评价。然后建立结合物理模型的神经网络[10][11]，以 $\Delta\log R$ 模型输出结果与测井曲线组合作为神经网络输入，基于 BP 神经网络与 LSTM 神经网络进行 TOC 评价。

2. 烃源岩测井响应特征

烃源岩定义为：富含有机质，可以生成油气并且排出油气的岩石，包括油源岩、气源岩和油气源岩，习惯上通常叫作生油岩。烃源岩通常为石灰泥岩和页岩，它们在地质发展和地层演化的过程中生成大量的烃类物质，包括石油和天然气。所以非烃源岩的主要组成可以划分为粘土矿物、非粘土矿物、内部流体，烃源岩的主要组成划分为粘土矿物、非粘土矿物、有机质、内部流体。

研究区域的测井响应见图 2，数据来源于鄂尔多斯盆地西北缘地区 XX 井，结合多条曲线特征分析。

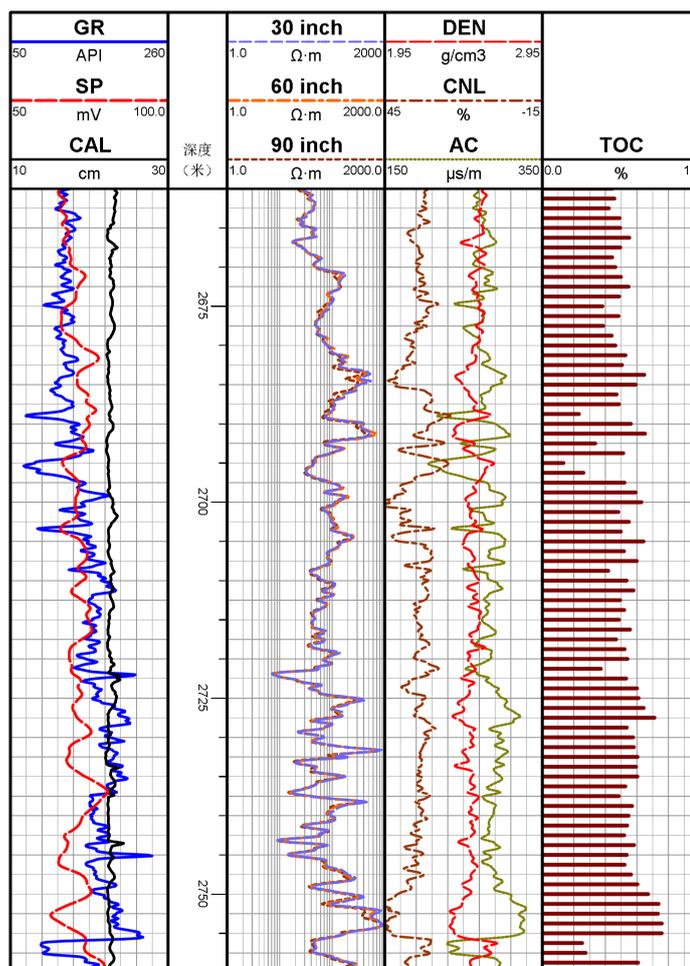


Figure 2. Well logging response in the study area
图 2. 研究区域测井响应

密度测井：相对于其他矿物成分，有机质的密度较低，在富含有机质的烃源岩地层，可以观察到密度测井曲线呈现低值。

声波时差测井：有机质的密度低，可以得到其声波时差较高的特性，有机质含量越高，声波时差就会越高，如果烃源岩含有气，则声波时差曲线会出现异常高值。

电阻率测井：随着烃源岩成熟阶段的提高，岩石中的油气含量不断升高，油气等烃类属于不导电介质，这些将会导致岩石电阻率升高，并且有机质含量越高，电阻率异常越大。

自然伽马测井：烃源岩中富含有机质，有机质吸附的铀元素将在伽马能谱曲线贡献高值，随着有机质含量增加，自然伽马测井曲线不断升高。

补偿中子测井：利用双探测器消除井眼对测量结果的影响的同位素源中子测井方法，主要反映地层对快中子的减速能力，显示出地层含氢量的变化。烃源岩中富含有机质，含氢很高，所以随有机质含量的增加会在中子曲线上呈现高值。

研究区域地层为富含有机质的页岩地层，存在大量的干酪根组分，在测井曲线上主要反映为低密度、高声波时差、高电阻率、高自然伽马和高补偿中子。

3. 基于物理模型的烃源岩有机质含量评价方法

3.1. 基于 ΔLogR 方法评价烃源岩有机质

ΔlogR 技术是埃克森于 1979 年开始开发和测试的，该技术已成功应用于生产实践[12]。尽管其他方法和新工具(如诱导伽马射线能谱测井)在直接测量有机碳含量方面可能具有特定的优势，但 ΔlogR 技术已被发现在碳酸盐岩和碎屑岩烃源岩中都能充分发挥作用，并且可以大范围准确预测 TOC， ΔlogR 的计算公式如下：

$$\Delta\text{logR} = \log_{10}\left(\frac{R}{R_{\text{baseline}}}\right) + 0.02(\Delta t - \Delta t_{\text{baseline}}) \quad (1)$$

公式 1 中， R 是电阻率取值， R_{baseline} 是电阻率基线， Δt 是声波时差， $\Delta t_{\text{baseline}}$ 是声波时差基线，式中 0.02 是指在地层中一定电阻率范围内声波时差的近线性变化系数，与地层因素有关。 ΔLogR 分离与总有机碳线性相关，是成熟度的函数， ΔLogR 的分离可以通过函数计算转换为总有机碳(TOC)，公式如下：

$$\text{TOC} = (\Delta\text{logR}) \times 10^{(2.297 - 0.1688 \times \text{LOM})} \quad (2)$$

ΔLogR 模型的应用效果如图 3 所示，图中横坐标为模型计算值，纵坐标为岩心实验分析值。利用均方误差(MSE)来评价方法的应用效果。均方误差是一种用于衡量预测模型误差的指标，在机器学习和统计学中 MSE 被广泛应用，计算公式如下：

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n} \quad (3)$$

公式中 n 表示样本数量， y_i 表示预测值， \bar{y}_i 表示真实值，MSE 的值越小，说明模型的预测能力越好。结合交汇图与 MSE 可以看出，交汇点分布在直线 $Y=X$ 两侧且分布散乱，MSE 为 3.413，说明使用 ΔLogR 模型评价研究区域井段的烃源岩有机碳含量效果不是很好，有待改进。

ΔLogR 方法自提出以来，应用一直十分广泛，其作者指出在有机质含量低的地层中 ΔLogR 分离也可能发生，但在大多数情况下，这种异常的分层主要发生在以下情况：(a) 油气藏，(b) 钻孔条件差，(c) 未压实的沉积物，(d) 致密层，(e) 火成岩，(f) 蒸发岩。并且该方法对声波曲线、密度曲线和中子曲线的计算中采用曲线均值代替基线或会对计算结果有影响。

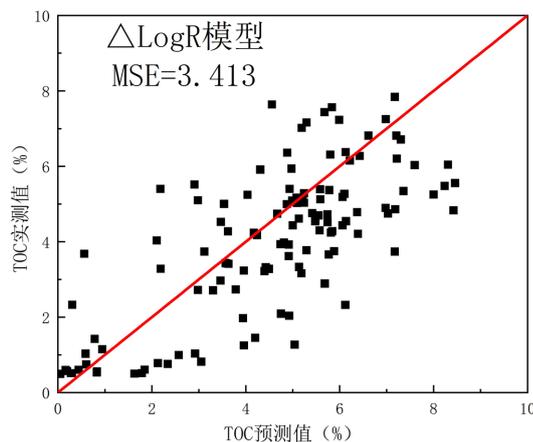


Figure 3. ΔLogR model results

图 3. ΔLogR 模型结果

3.2. 基于变基线方法评价烃源岩有机质

变基线方法(Zhao, 2017)是基于原始 ΔLogR 模型和修正模型(Wang, 2016)的改进,指出 ΔLogR 模型中手动确定基线存在的误差,提出用建立理论基线取代近线性基线,并且明确给出理论基线确定方法并验证其可行性,通过数学方法得到了不同的基线数值[13]。在贫岩数据区域建立理论基线,有机碳含量高的烃源岩声波时差和电阻率都比贫岩段的岩心点数据要高。进行 ΔLogR 模型有机碳含量计算时,不同点位的岩心都会有对应的基线点,反应为每个岩心点在理论基线上都有一距离的最小值点相对应。对每个岩心点计算得到相应的基线值后,使用改进后的公式计算 TOC:

$$TOC = aa\Delta \log R + bb(GR - GR_{baseline}) + cc \quad (4)$$

aa、bb、cc 为模型参数,使用多元回归的方法计算,参数必须根据油藏实际情况进行计算,不同井对应不同的模型参数。图 4 是变基线模型对于研究区域的 TOC 含量计算结果,可以看到变基线模型的预测数据与真实岩心数据交汇在 $Y = X$ 直线附近,并且 MSE 数值为 2.034,相较于 Passey 的 $\Delta\log R$ 模型有较为明显的改进,说明 Zhao 的改进模型有效,在改进 $\Delta\log R$ 模型基线手动标定取定值问题后,变基线模型更适合研究区域 TOC 含量的评价。

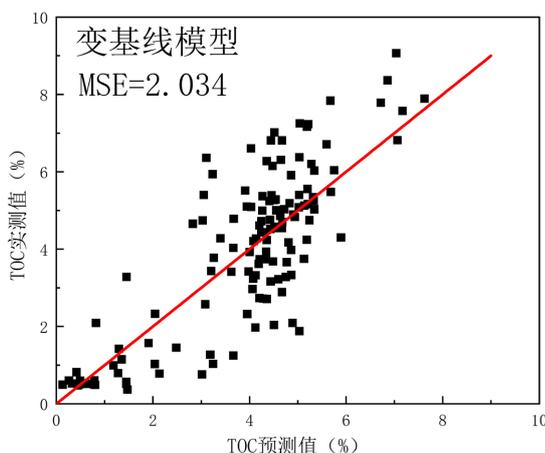


Figure 4. TOC prediction results of variable baseline model

图 4. 变基线模型预测 TOC

该模型建立理论基线取代近线性基线提高了 ΔLogR 模型计算精度,对于理论基线的建立,需要注意数据选取时需根据地层测井数据整体分析,选取合适的岩心点数据。而在整体烃源岩含量都不低的井段,选取到的数据可能会导致理论基线值偏高,计算 TOC 偏低。

3.3. 基于 DD ΔLogR 方法评价烃源岩有机质

DD ΔLogR 的方法(Zhu, 2019)是对 Passey 提出 ΔlogR (1990)模型的改进,此方法考虑到岩石矿物成分的变化和孔隙扩大的影响,有效地提高了 ΔlogR 的计算精度,极大地扩展了 ΔlogR 方法的适用性,可以有效地帮助页岩油藏的勘探和开发[14]。将基于各种矿物的体积和声波时差响应模型的体积模型演化为基于声波时差的体积模型,如图 5 所示。由于粘土和石英的 Δt 值差异较大,前者的声波时差值几乎接近后者的两倍,此外,页岩储层的粘土含量通常大于 40%,因此粘土的影响是建立页岩的 Δt 响应特性的体积模型时,必须要考虑的问题。方解石、白云石、黄铁矿等矿物的 Δt 响应值与石英无显著差异,可以在非粘土基质部分进行综合考虑。

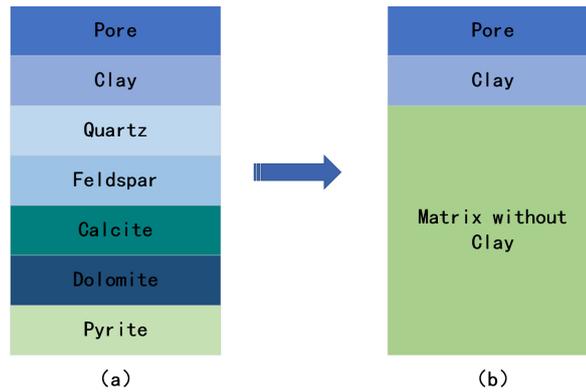


Figure 5. Acoustic volume model
图 5. 声学体积模型

图 5 中(a)图是理论的声学体积模型,(b)图是 Zhu 在文中新建立的声学体积模型,对于低有机质岩石 Δt 体积模型:

$$\Delta t = \phi t_w + V_{\text{clay}} \Delta t_{\text{clay}} + (1 - \phi - V_{\text{clay}}) \Delta t_m \quad (5)$$

Δt 是岩石声波时差, Δt_m 是不含黏土骨架声波时差, ϕ 是孔隙体积占比, V_{clay} 黏土体积占比。根据 Δt 公式计算页岩储层孔隙度:

$$\phi = \frac{\Delta t - \Delta t_m}{\Delta t_f - \Delta t_m} + V_{\text{clay}} \frac{\Delta t_{\text{clay}} - \Delta t_m}{\Delta t_f - \Delta t_m} \quad (6)$$

将 ϕ 式子带入阿尔奇公式 $R_o = \frac{aR_w}{\phi_m}$, 得到:

$$\log 10 R_o = \log \left(\frac{aR_w}{\left(\frac{(\Delta t - \Delta t_m) + V_{\text{clay}} (\Delta t_{\text{clay}} - \Delta t_m)}{\Delta t_f - \Delta t_m} \right)^m} \right) \quad (7)$$

新的理论关系考虑了矿物变化对 Δt 的影响,提取的 $\Delta \log R$ 结果不受矿物成分变化的影响。第二,理论关系曲线的形状发生了变化,曲线的形状与粘土含量有关,因此,该曲线将基线值的变化与矿物成分的变化联系起来。动态理论关系曲线考虑了矿物的影响,从而提高了 $\Delta \log R$ 计算精度。

$$\Delta t = V_{Quartz} \Delta t_{Quartz} + V_{Feldspar} \Delta t_{Feldspar} + V_{Calcite} \Delta t_{Calcite} + V_{Dolomite} \Delta t_{Dolomite} + V_{Clay} \Delta t_{Clay} + V_{Pyrite} \Delta t_{Pyrite} + V_{Por} \Delta t_{Por} + V_{OM} \Delta t_{OM} \quad (8)$$

$$\Delta t_{mapor} = V_{Quartz} \Delta t_{Quartz} + V_{Feldspar} \Delta t_{Feldspar} + V_{Clay} \Delta t_{Clay} + V_{Calcite} \Delta t_{Calcite} + V_{Dolomite} \Delta t_{Dolomite} + V_{Pyrite} \Delta t_{Pyrite} + V_{Por} \Delta t_{Por} \quad (9)$$

$$DD\Delta \log R = \Delta \log R - \Delta \log R_{mpor} \quad (10)$$

$\Delta \log R$ 反应有机岩, $\Delta \log R_{mapor}$ 反应有机贫岩, $DD\Delta \log R$ 可以看作作为烃源岩有机碳特征提取。 GR 曲线可以显著提高预测和测量的 TOC 含量之间的相关性,可作为岩石类型指标,降低有机贫、高电阻率非源岩(页岩层间、含烃、非常细粒砂岩-粉砂岩)对 TOC 含量的影响。在组成页岩的主要矿物中,具有较强放射性的矿物(OM 除外)是粘土,如果用放射性曲线来准确地表征 TOC 的含量,就必须排除粘土等放射性矿物的影响。由于动态理论关系曲线的精确测定随深度的变化,需要测定粘土含量;因此,文章直接使用粘土含量曲线来反映粘土的比例,最终推导出的公式如下:

$$\Delta D = \frac{GR - GR_{left}}{GR_{right} - GR_{left}} \frac{V_{clay} - V_{clay-left}}{V_{clay-right} - V_{clay-left}} \frac{V_{k-feldspar} - V_{k-feldspar-left}}{V_{k-feldspar-right} - V_{k-feldspar-left}} \quad (11)$$

在此公式中,不仅做了自然伽马校正,由于钾长石对自然伽马也作很大比例的贡献,所以又进行了钾长石的校正。最终 $DD\Delta \log R$ 模型表示如下:

$$TOC = aa(DD\Delta \log R) + bb(\Delta D) + cc \quad (12)$$

$DD\Delta \log R$ 模型评价 TOC 结果如图 6 所示,模型的预测数据与真实岩心数据交汇在 $Y = X$ 直线附近,并且 MSE 数值为 1.645,相较于 Passey 的 $\Delta \log R$ 模型有较为明显的改进,说明双差分的改进模型有效,并且双差分模型较为适合该研究区域 TOC 含量的研究。目标研究层段为页岩储层,方解石、白云石、黄铁矿等矿物含量不高,使用该模型进行非粘土基质等效对结果不会有很大影响,不过在碳酸盐岩和碎屑岩中,石英、长石、云母和重矿物含量升高会对模型结果产生很大影响, $DD\Delta \log R$ 模型不再适用。

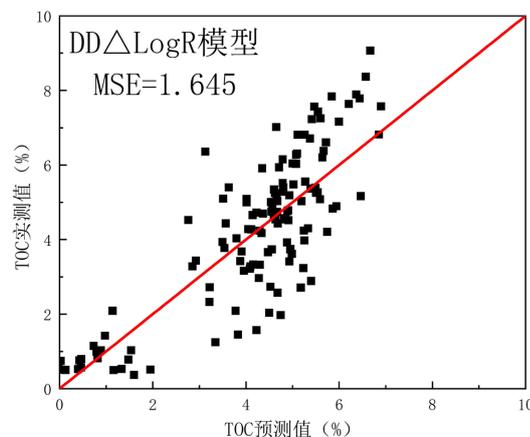


Figure 6. $DD\Delta \log R$ model results
图 6. $DD\Delta \log R$ 模型结果

4. 基于机器学习的烃源岩有机质含量评价方法

4.1. 基于 BP 神经网络评价烃源岩有机质

BP 神经网络预测 TOC 结构如图 7 所示：BP 神经网络包括有 1 个输入层、数个隐层(本实验中设置 5 个隐层)和 1 个输出层组成[15]。优选与 TOC 相关性高的测井曲线作为输入，以岩心 TOC 为标签，设置合适超参数，完成网络训练，之后用训练好的网络处理未经训练的数据集，完成 TOC 评价。

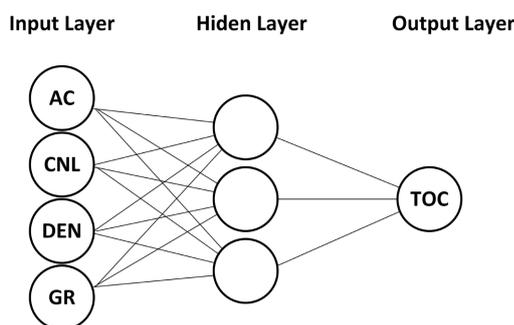


Figure 7. Structure diagram of TOC predicted by BP network
图 7. BP 神经网络预测 TOC 结构示意图

在进行数据训练之前，先讨论数据的输入问题，由测井响应与 TOC 的相关性分析得到岩心 TOC 与 GR、AC、CNL、Rt、SP 是正相关，相关系数是 0.667、0.752、0.639、0.501 和 0.439；TOC 与 DEN 为负相关，相关系数是 -0.622，根据相关性的绝对值大小比较来看，AC、GR、CNL、DEN 四条测井曲线与岩心 TOC 相关性较高，均大于 0.6，Rt 与 SP 曲线与岩心 TOC 的相关性较低，分别是 0.501 和 0.439。在关于输入数据的选择上，J Utans 等认为只输入与目标数据相关性好的数据可以得到较好的训练结果[16]，姚立忠等学者认为相关性好的和坏的数据合理进行输入，可以使得神经网络更加全面地了解输入数据的特征和规律，提高对数据的理解和学习能力[17]。基于此问题的讨论，针对本次研究，本文做了曲线的优选实验，分别选择不同的输入曲线组合，在参数设置相同的情况下，观测其结果，为了避免误差存在，对同一种曲线组合进行 10 次训练，取其验证集的相关性和均方误差 10 次测量的平均值作为比较样本，最终得到数据结果如表 1 所示：

Table 1. Curve optimization table
表 1. 曲线优选表格

| 曲线组合 | 高相关组合 | +RT | +SP | +RT SP |
|---------------|-------|-------|-------|--------|
| GR AC | 3.789 | 3.166 | 3.712 | 3.283 |
| CNL DEN | 4.238 | 3.364 | 3.272 | 3.475 |
| AC CNL | 4.354 | 3.123 | 7.989 | 3.699 |
| GR DEN | 3.689 | 2.926 | 3.054 | 3.664 |
| AC CNL GR | 4.205 | 3.554 | 3.931 | 3.748 |
| AC CNL DEN | 3.656 | 3.638 | 3.647 | 3.398 |
| AC DEN GR | 2.918 | 3.158 | 4.667 | 3.135 |
| CNL DEN GR | 3.419 | 3.364 | 3.292 | 2.868 |
| AC CNL GR DEN | 3.238 | 3.056 | 3.826 | 3.293 |

表格中分别讨论 9 种相关性好的测井曲线的不同组合，然后再分别添加一条和两条相关性差的测井

曲线,表中一共是 36 种不同的测井曲线组合进行比较其相关性和均方误差。从表格可以看出,只使用相关性差的数据作为输入结果是最差的,只使用相关性好的数据结果也不是最好的;在给相关性好的输入组添加相关性差曲线的 27 种组合中,只有 5 中组合数据结果变差,其余的 22 中组合结果均变好,也就是给相关性好的数据输入组合添加相关性差的数据会使结果更好的概率是 81.48%。最终得到的最优测井曲线输入组合为 CNL、DEN、GR、Rt、SP,该组合 MSE 最低,且相关性最高。说明,在进行神经网络输入时,输入相关性较好数据和相关性较差数据合理进行输入,可以使得神经网络更加全面地了解输入数据的特征和规律,提高对数据的理解和学习能力。同时,只输入相关性较好数据可能会导致过拟合的问题,即模型过度适应训练数据,而在实际使用中表现不佳。

使用最优测井曲线组合作为输入,为了更全面的评价 BP 神经网络方法,随机抓取 100 个岩心点位作为训练集,44 个岩心作为测试集,评价研究区域的 TOC 结果见图 8:可以看到,使用 BP 神经网络评价 TOC,十次训练 MSE 最高是 3.968,最低是 1.995,十次均值为 2.868。

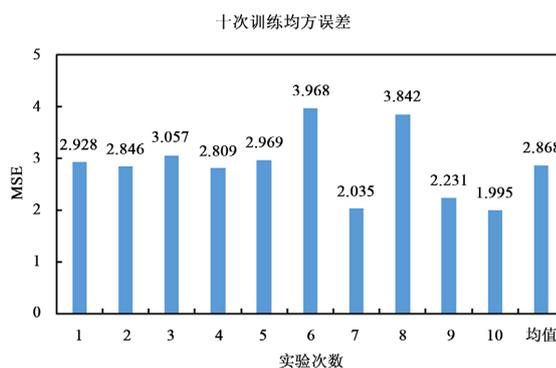


Figure 8. Prediction of TOC results by BP neural network
图 8. BP 神经网络预测 TOC 结果

4.2. 基于 LSTM 神经网络评价烃源岩有机质

LSTM 是 Long Short Term Memory 简称,又叫长短期记忆神经网络,是为了解决 RNN (循环神经网络)长期依赖问题设计出的神经网络,通过多个门函数将数据的重要特征进行保留,有效减缓了长序列问题中可能出现的梯度消失或爆炸问题,表现优于传统 RNN 神经网络[18] [19]。

选择 4.1 小节中优选的最佳测井曲线组合作为 LSTM 神经网络的输入,随机抓取 100 个岩心点位作为训练集,44 个岩心作为测试集,评价研究区域的 TOC 结果见图 9 所示,使用 LSTM 神经网络评价 TOC,十次训练 MSE 最高是 2.986,最低是 1.709,十次均值为 2.141。

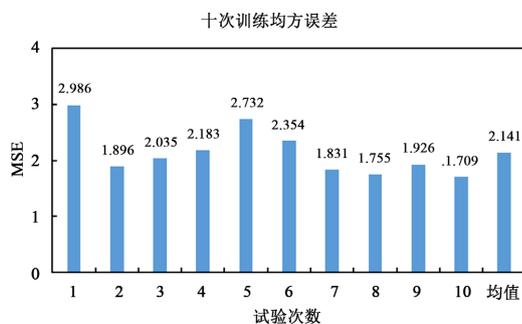


Figure 9. Prediction of TOC results by LSTM neural network
图 9. LSTM 神经网络预测 TOC 结果

4.3. 基于结合物理模型的神经网络评价烃源岩有机质

本次实验对神经网络与物理模型做结合,进行 TOC 含量的预测。由于变基线模型和双差分模型的计算结果都是进行了多元回归得到的,为防止过拟合问题,不作为选取,所以此次的结合选取 Passey 的 $\Delta\log R$ 模型,其计算结果全部来自公式计算而来,有数值上的误差。把本次实验中 Passey 的 $\Delta\log R$ 模型计算结果分别作为 BP 神经网络和 LSTM 神经网络的输入,又考虑到单一输入对神经网络训练效果影响较大,导致过拟合问题。最终选取最佳测井曲线组合一同作为神经网络输入,然后以真实岩心 TOC 作为标签,随机抓取训练集,进行训练。

图为四种神经网络方法评价 TOC: BP 神经网络、LSTM 神经网络、结合 $\Delta\log R$ 模型的 BP 神经网络、结合 $\Delta\log R$ 模型的 LSTM 神经网络,见图 10 四种神经网络方法进行对比:

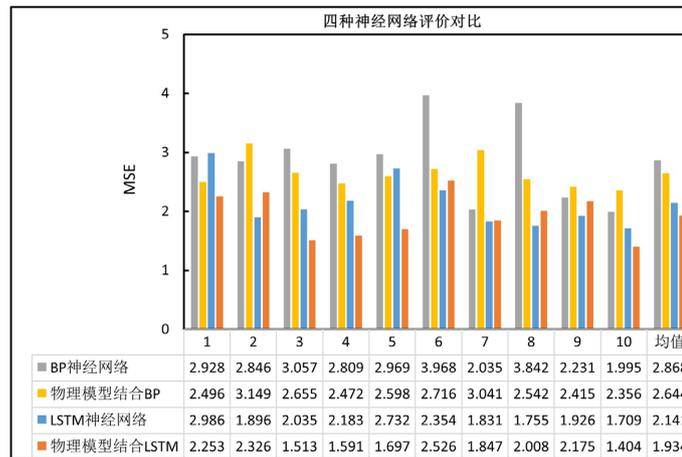
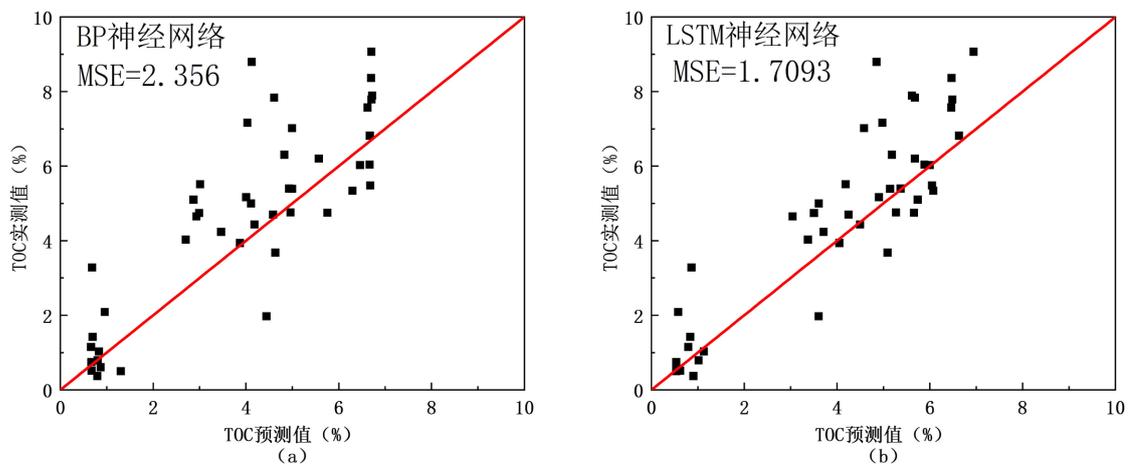


Figure 10. Comparison of TOC in neural network evaluation
图 10. 神经网络评价 TOC 对比

结合物理模型的神经网络评价 TOC 效果要优于不结合物理模型的神经网络, LSTM 神经网络整体评价 TOC 效果优于 BP 神经网络, LSTM 神经网络处理像 TOC 含量这种连续性数据具有更好的效果,四种神经网络中,结合 $\Delta\log R$ 物理模型的 LSTM 神经网络表现最好, MSE 均值最小为 1.934。图 11 为十次神经网络训练中随机一次训练效果展示:



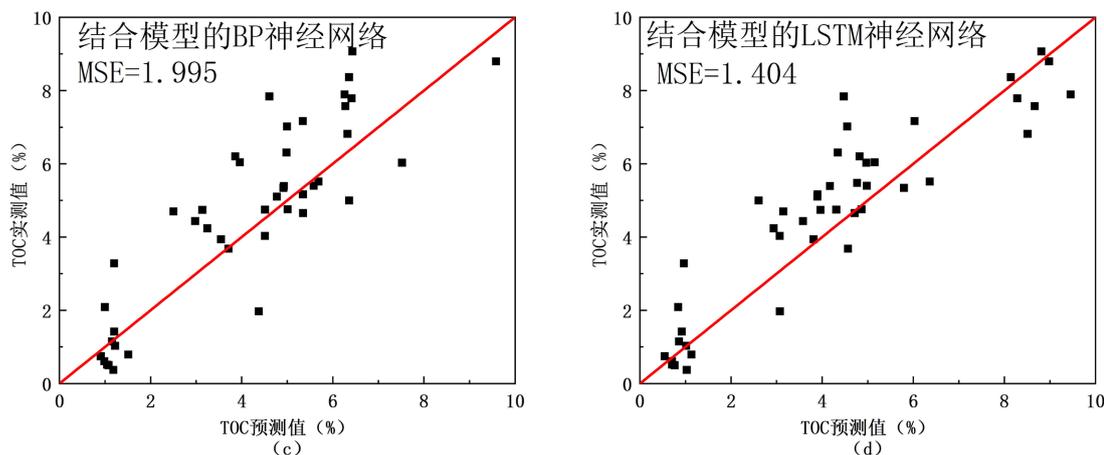


Figure 11. Comparison of TOC in neural network evaluation
图 11. 神经网络评价 TOC 对比

LSTM 神经网络评价效果优于 BP 神经网络, 因为 LSTM 神经网络对长序列数据处理更具优势, 它可以记忆先前的数据输入并将其作为后续预测的一部分考虑。结合 $\Delta\log R$ 模型的神经网络, 评价效果优于未与物理模型结合神经网络, 这说明结合物理模型后提升了神经网络的性能。综合对比四种神经网络模型, 结合 $\Delta\log R$ 模型的 LSTM 神经网络评价 TOC 效果最好。

5. 结论

本文基于常规测井数据, 对鄂尔多斯盆地中晚三叠世湖盆西北缘地区烃源岩 TOC 进行评价。共使用了三种物理模型($\Delta\log R$ 模型、变基线模型和 $DD\Delta\log R$ 模型), 以及四种神经网络(BP 神经网络和 LSTM 神经网络; 与 $\Delta\log R$ 结合的 BP 神经网络, 与 $\Delta\log R$ 结合的 LSTM 神经网络)评价 TOC, 通过实际数据处理, 得到以下结论:

(1) $\Delta\log R$ 模型在研究区域的效果不佳, 变基线方法改进了 $\Delta\log R$ 方法中基线确定的方法, 较好地提高了 $\Delta\log R$ 模型预测精度; 双差分方法考虑矿物成分变化与孔隙的影响, 同时也解决手动标定基线不稳定问题, 在变基线的基础上进一步改进了模型, 在三种物理模型中评价 TOC 效果最佳。

(2) 神经网络可以更准确地预测 TOC, LSTM 神经网络的效果优于 BP 神经网络。与 $\Delta\log R$ 模型结合的神经网络, 可以提高神经网络的预测能力, 在四种神经网络中, 与 $\Delta\log R$ 模型结合的 LSTM 神经网络评价 TOC 效果最优。

(3) 在进行神经网络训练时, 只输入相关性好的数据, 模型过度适应训练数据, 可能会导致模型过拟合的问题, 在实际使用中表现不佳。将与神经网络相关性较好与相关性较差的参数合理组合作为输入, 可以提高神经网络预测能力。

参考文献

- [1] 王宵宇, 谢然红, 毛治国, 等. 基于集成学习的烃源岩总有机碳含量测井评价方法研究[J]. 地球物理学进展, 2022, 37(2): 684-694.
- [2] 边雷博, 柳广弟, 孙明亮, 等. 优化的 $\Delta\log R$ 技术及其在中—深层烃源岩总有机碳含量预测中的应用[J]. 油气地质与采收率, 2018, 25(4): 40-45.
- [3] 陈海峰, 王凤启, 王民, 等. 基于变系数 $\Delta\log R$ 技术的烃源岩 TOC 精细评价[J]. 中国石油大学学报(自然科学版), 2017, 41(4): 54-64.
- [4] 张晗, 卢双舫, 李文浩, 等. $\Delta\log R$ 技术与 BP 神经网络在复杂岩性致密层有机质评价中的应用[J]. 地球物理学

- 进展, 2017, 32(3): 1308-1313.
- [5] 杨锋, 李生杰, 张小龙. 基于变基线的 $\Delta\log R$ 法研究[C]//中国地球物理学会, 中国地震学会, 全国岩石学与地球动力学研讨会组委会, 中国地球物理学会. 2016 中国地球科学联合学术年会论文集. 2016: 123-126.
- [6] Liu, C., Zhao, W., Sun, L., Zhang, Y., Chen, X. and Li, J. (2021) An Improved $\Delta\log R$ Model for Evaluating Organic Matter Abundance. *Journal of Petroleum Science and Engineering*, **206**, Article ID: 109016. <https://doi.org/10.1016/j.petrol.2021.109016>
- [7] Bai, Y. and Tan, M. (2021) Dynamic Committee Machine with Fuzzy-C-Means Clustering for Total Organic Carbon Content Prediction from Wireline Logs. *Computers & Geosciences*, **146**, Article 104626. <https://doi.org/10.1016/j.cageo.2020.104626>
- [8] Rui, J., Zhang, H., Zhang, D., Han, F. and Guo, Q. (2019) Total Organic Carbon Content Prediction Based on Support-Vector-Regression Machine with Particle Swarm Optimization. *Journal of Petroleum Science and Engineering*, **180**, 699-706. <https://doi.org/10.1016/j.petrol.2019.06.014>
- [9] Wang, P., Peng, S. and He, T. (2018) A Novel Approach to Total Organic Carbon Content Prediction in Shale Gas Reservoirs with Well Logs Data, Tonghua Basin, China. *Journal of Natural Gas Science and Engineering*, **55**, 1-15. <https://doi.org/10.1016/j.jngse.2018.03.029>
- [10] 魏红伟, 袁江, 田杰. 融合物理模型和神经网络的图像去雾方法与实现[J]. 现代电子技术, 2022, 45(9): 55-60.
- [11] Kim, T., Lee, H. and Lee, W. (2022) Physics Embedded Neural Network Vehicle Model and Applications in Risk-Aware Autonomous Driving Using Latent Features. arXiv: 2207.07920
- [12] Passey, Q.R., Creaney, S., Kulla, J.B., *et al.* (1990) A Practical Model for Organic Richness from Porosity and Resistivity Logs. *AAPG Bulletin*, **741**, 777-1794.
- [13] Zhao, P., Ma, H., Rasouli, V., Liu, W., Cai, J. and Huang, Z. (2017) An Improved Model for Estimating the TOC in Shale Formations. *Marine and Petroleum Geology*, **83**, 174-183. <https://doi.org/10.1016/j.marpetgeo.2017.03.018>
- [14] Zhu, L., Zhang, C., Zhang, Z., Zhou, X. and Liu, W. (2019) An Improved Method for Evaluating the TOC Content of a Shale Formation Using the Dual-Difference $\Delta\log R$ Method. *Marine and Petroleum Geology*, **102**, 800-816. <https://doi.org/10.1016/j.marpetgeo.2019.01.031>
- [15] 中文期刊网. 人工神经网络的实际应用范例 6 篇[Z/OL]. <https://www.baywatch.cn/>, 2024-03-28.
- [16] Utans, J., Moody, J., Ehfuss, S., *et al.* (1995) Input Variable Selection for Neural Networks: Application to Predicting the U.S. Business Cycle. *Proceedings of 1995 Conference on Computational Intelligence for Financial Engineering (CIFER)*, New York, 9-11 April 1995, 118-122
- [17] 姚立忠, 李太福, 易军, 等. 神经网络模型的透明化及输入变量约简[J]. 计算机科学, 2012, 39(9): 247-251+278.
- [18] 王维, 姜珊, 胡剑琇, 等. 基于特征选择的 CNN-LSTM 结构应变响应预测方法[J/OL]. 土木与环境工程学报(中英文), 1-11. <http://kns.cnki.net/kcms/detail/50.1218.TU.20250408.1700.006.html>, 2025-06-11.
- [19] Duan, J., Hu, L., Xiao, Q., Liu, M. and Yu, W. (2025) A Geographic Information Encryption System Based on Chaos-LSTM and Chaos Sequence Proliferation. *Frontiers of Information Technology & Electronic Engineering*, **26**, 427-440. <https://doi.org/10.1631/fitee.2300755>