

doi:10.3799/dqkx.2026.013

基于人工神经网络优化算法的油气水三相平衡计算

沈安琪^{1,*}, 张博源¹, 任思扬², 魏建光¹, 张栋¹, 宋康¹

1. 东北石油大学多资源协同陆相页岩油绿色开采全国重点实验室, 黑龙江大庆 163318

2. 大庆油田有限责任公司第二采油厂, 黑龙江大庆 163000

摘要: 油气藏中的流体相态是评估储量、制定开发方案的基础; 相平衡计算可提供多组分流体在不同环境下相态预测, 为油气藏开发提供关键基础流体物性参数。本文基于传统热力学模型建立考虑 CO₂ 溶于水的油气水三相平衡计算模型。通过该模型获取数据样本, 结合遗传算法 (GA) 和模拟退火算法 (SA), 基于人工神经网络建立油水气 (含 CO₂) 三相相态预测模型, 可计算相分率及各组分在油气水三相间摩尔分数。研究结果表明, 基于人工神经网络优化算法的油气水三相平衡预测模型误差在 5% 以内, 其中 GA、SA 算法协同优化后的模型准确率最高。将传统热力学计算与神经网络遗传、模拟退火优化算法相结合, 可为复杂的三相平衡计算提供指导。

关键词: 相平衡计算; 遗传优化算法; 模拟退火优化算法; 亨利定律

中图分类号: TE312 收稿日期: 2025-09-28

A three-phase phase equilibrium calculation for oil, gas and water based on Artificial Neural Network and optimization algorithm

Abstract: The fluid phase state in oil and gas reservoirs is the basis for evaluating reserves and development plans; phase equilibrium calculations can provide phase state predictions of multi-component fluids under different conditions, providing basic fluid physical parameters for oil and gas reservoir. Due to the solubility of CO₂ in water, to study the phase fraction and component variation characteristics of the coexisting oil-water-gas (including CO₂) system, a three-phase phase equilibrium calculation model for oil-water-gas (including CO₂) is established based on the traditional thermodynamic model, considering the solubility of CO₂ in water. Data obtained through this model are used to form a sample set, and combined with genetic algorithm (GA) and simulated annealing algorithm (SA), an artificial neural network-based three-phase phase state prediction model for oil-water-gas (including CO₂) is established, obtaining the phase fraction and mole fraction of each component in the oil-water-gas three phases. The research results show that the error of the three-phase phase equilibrium prediction model based on the artificial neural network optimization algorithm is controlled within 5%, among which the model optimized by GASA together has the highest accuracy. The combination of traditional thermodynamic calculation and neural network genetic and simulated annealing optimization algorithms can provide guidance for complex three-phase phase equilibrium calculations.

Key words: Phase equilibrium calculation; Genetic algorithm; Simulated annealing algorithm; Henry's law

前言

相平衡分析是理解物质状态和行为的基础, 其重要性和普遍性存在于自然科学的各个角落和几乎所有的工业领域 (柳波等, 2014; Chen et al., 2025; 潘永帅等, 2025; 薛永安等, 2025) [3][14][18][26]。相平衡分析的核心是研究多组分系统在特定条件 (温度、压力、组成) 下能够稳定共存。化学热力学和统计热力学是

基金项目: 黑龙江省自然科学基金优秀青年项目 (YQ2024E008)

第一作者及通讯作者: 沈安琪, 女, 1986 年 3 月生, 2017 年获东北石油大学博士学位, 现为东北石油大学副教授、博士生导师, 主要从事渗流、提高油气采收率、流体相态等方面研究工作。Email: anqi1986@126.com, ORCID: <https://orcid.org/0000-0002-8171-7120>

相平衡理论的热力学核心。相图是吉布斯自由能最小原理最直观、最有力的体现。相平衡预测模型提供了预测物质在不同环境下（如加热、冷却、加压、混合）将发生何种变化（如熔化、凝固、沸腾、溶解、析出）的理论工具。

在石油工业中，油气藏中的流体相态是评估储量、制定开采方案的基础（白龙辉等，2021；Cai et al., 2025；朱光有等，2025）^{[1][2][30]}；尤其油藏气驱的驱油机理比传统注水开发更为复杂。传统热力学模型多以油气两相平衡计算为主，忽略了气水相互作用，这导致气驱和 CO₂ 埋存中的流体量估算中存在较大错误。以 CO₂ 驱油-埋存为例，由于 CO₂ 同样易溶于水，因此 CO₂、原油和地层水间存在复杂相互作用。CO₂ 驱油效果很大程度上取决于 CO₂ 与原油组分及其物理性质。受储层温度、压力非均一分布影响，CO₂ 与原油的相互作用可呈现混相驱替或非混相驱替模式。特别是在混相驱替过程中温度、压力变化和原油组分变化使流体相态发生改变（邸元等，2004）^[4]。通过热力学相平衡分析可定量表征多相体系相个数及组分摩尔浓度等关键参数，这些热力学参数的解析是提升驱油效率的关键问题，因此建立相平衡模型是 CO₂ 提高原油采收率的关键技术。

相平衡研究方面，实验是相平衡理论研究的基础，是获取多组分流体高温高压物性最直接、可靠的方法。许多学者开展高压相平衡实验直观、准确获取相平衡参数与流体物性（韩芳，2004；Kariznovi et al., 2013；翼胜合，2019；Liu et al., 2021）^{[7][13][15][27]}，但受限于实验室条件且时间成本高，理论计算方法可一定程度上弥补实验的不足。理论预测模型可利用较少实验数据推算出所需要的全部数据，并获得连续相变趋势的优点。多元流体的温度和压力范围覆盖广，常使用状态方程进行理论计算。状态方程是表征混合物体系压力、体积、温度和各组分组成的基本关系式，可计算物质的相变和多种热力学性质。状态方程的准确性和适用性直接决定了相态分析的可靠性（郭克群，2012）^[6]。水通常存在于所有的油气藏中。由于大多数储层岩石的水湿特性，水以小水滴或薄层的形式留在储层中，覆盖在颗粒上成为束缚水。除此之外，边水、底水等多种形式赋存水存在油气藏中，因此，在油气藏开采相平衡研究中需要考虑水相作用。于海洋等（2019）^[28]基于长庆油田 M 区致密储层，开展了碳化水驱，通过引入亨利定律和表面活性剂驱模型，能够较为准确地预测碳化水驱的开发效果，与实验结果相比误差低于 1.2%。Monroy-Loperena (Rosendo, 2023)^[22]提出一种三相闪蒸计算方法，即基于切线逼近各函数的 Rachford-Rice 方程组解的初始化方法，并证明该方法的有效性、鲁棒性和可靠性。

传统热力学相态计算方法理论与计算求解过程复杂，虽已有商业软件可对相态进行求解，但多限于经典理论方程，无法面对目前复杂应用条件，适用性受到一定挑战。虽然国内外学者在多元流体相平衡热力学计算领域已取得显著进展，但是基于状态方程的理论模型呈现出以下特点：（1）理论体系复杂，需运用 SRK、PR 等状态方程结合混合规则构建多组分模型。（2）算法实现难度大、计算时长，需牛顿-连续迭代、负闪蒸等计算方法解决传统 Rachford-Rice 方程组的收敛性问题。（3）计算 CO₂-油-水多相体系时，目前处于初步探索阶段，实验数据匮乏模型验证困难等。尽管如此，这些方法能揭示相变机理与关键参数影响规律，为优化油气开采工艺提供理论基础。未来需加强实验和理论的协同研究，发展新型热力学模型。

为解决计算复杂与时长问题，亟需一种快速且准确的相平衡计算方法。数据驱动方法是一种以数据为基础的决策和问题解决方法，通过对大量数据进行分析和挖掘，探索出潜在的数据关联和规律。与热力学状态方程不同，机器学习算法可以通过应用一些简单的调整来提供广泛条件下的精确结果。目前相态计算方面广泛应用数据驱动方法，主要是基于人工神经网络的机器学习和深度学习方法（Hui et al., 2025）^[11]。Huang 等（2022）^[10]开发了一种基于条件生成对抗网络方法的 CO₂-烃体系最小混相压力预测模型，选取 180 组数据进行实验，平均绝对误差仅为 6.81%，具有良好预测效果。Sun 等（2023）^[24]开发出一种随机森林方法预测气液平衡（VLE）的基本热力学特征，共构建了 210 个二元混合物的实验数据集，通过计算测试集的平均误差和对特定混合物的相图检验，证实该模型提供了准确的热力学性质预测结果。Moradkhani 等（2023）^[16]开发一种 CO₂-烃类混合物液体密度预测方法对于提高采收率（EOR）和 CO₂ 储存过程的建模、优化和设计至关重要；该径向基函数（Radial Basis Function, RBF）模型平均绝对相对误差为 1.61%，相对均方根误差（RRMSE）为 4.29%，该模型在精度、适用性和通用性方面优于早期的经验相关性。

人工神经网络方法与热力学理论的交叉融合为复杂系统建模提供了创新方法论。通过将热力学框架嵌入神经网络的架构中，可有效提升模型的泛化性能为解决复杂工程问题提供了兼具数据适应性与物理严谨性的解决方案。基于人工神经网络的相态模型有助于加速复杂的闪蒸计算，在未来还需进一步的提高准确性和有效性。本文针对 CO₂-油-水多相多组分体系的相平衡预测，建立考虑 CO₂ 溶解于水的相平衡预测方法，分别采用热力学方法和人工神经网络方法对流体油气、油气水相平衡进行计算。利用亨利定律描述 CO₂ 在水中的溶解特性，结合 PR-EoS (Equation of state, 状态方程) 构建三相相平衡模型。借助实验数据对比，完成对模型的验证。利用油气水三相相平衡热力学计算方法获得数据集，构建遗传算法与模拟退火算法相结合的反向传播神经网络模型 (GASA-BPNN)，优化网络权重和阈值，并验证了神经网络模型的准确与高效性。文中所涉及数学公式与算法均在 Matlab 软件中实现。

1 考虑 CO₂ 溶于水的油气水三相相平衡热力学模型

1.1 亨利定律

亨利定律（施文等，1992）^[23]是指在一定温度和平衡状态下，气体在液体内溶解度（摩尔分数表示）与其分压成正比。水中组分逸度可由公式（1）计算

$$f_i^{aq} = x_i^{aq} H_i, i \neq w \quad (1)$$

式中 f_i^{aq} 为组分 i 在水相中的逸度，Pa； x_i^{aq} 为组分 i 在水相中摩尔分数； H_i 是组分 i 在水相中的亨利常数，Pa。在不同温度压力下可以写为，

$$\ln H_i = \ln H_i^* + \frac{v_i(p - p^*)}{RT} \quad (2)$$

式中 p^* 为参考压力，Pa； H_i^* 为亨利常数在 p^* 下的值，Pa；T 为温度，K；R 为理想气体常数，J/(mol⁻¹.K⁻¹)； v_i 为组分 i 在水相中的摩尔体积，cm³/mol。。

通常采用 Harvey (1996)^[8] 提出的经验计算亨利常数

$$\ln H_i^s = \ln p_w^s + A(T_{r,w})^{-1} + B(1-T_{r,w})^{0.355} \times (T_{r,w})^{-1} + C[\exp(1-T_{r,w})](T_{r,w})^{-0.41} \quad (3)$$

式中， H_i^s 为组分 i 在饱和压力下的亨利系数； p_w^s 为水的饱和压力，MPa； $T_{c,w}$ 为水的临界温度，K； $T_{r,w}$ 为水的对比温度， $T_{r,w} = T / T_{c,w}$ ；系数 $A = -9.4234$ ， $B = 4.0087$ ， $C = 10.3199$ 。

Saul 和 Wagner (1987)^[25] 提出计算水在不同温度下的饱和压力的经验公式，

$$\ln \frac{P_w^s}{P_{c,w}} = \frac{T_c}{T} (a_1 \tau + a_2 \tau^{1.5} + a_3 \tau^{2.5} + a_4 \tau^{3.5} + a_5 \tau^4 + a_6 \tau^{7.5}) \quad (4)$$

式中， $P_{c,w}$ 为水的临界压力，Pa； T 为系统温度，K； T_c 为水的临界温度，K； $\tau = 1 - T / T_{c,w}$ ；常系数值 $a_1 = -7.85823$ ， $a_2 = 1.83991$ ， $a_3 = -11.7811$ ， $a_4 = 22.6705$ ， $a_5 = -15.9393$ ， $a_6 = 1.77516$ 。

因此，在任意温度、压力下的亨利常数可以表示为，

$$\ln H_i = \ln H_i^s + \frac{1}{RT} \int_{p_w^s}^p v_i dp \quad (5)$$

式中， v_i 为组分 i 在水相中的摩尔体积，cm³/mol。

采用 Garcia (2001)^[5] 提出的经验公式求解 CO₂ 摩尔体积，

$$v_{CO_2} = 37.51 - 9.585 \times 10^{-2} T + 8.74 \times 10^{-4} T^2 - 5.044 \times 10^{-7} T^3 \quad (6)$$

式中 T 为温度, $^{\circ}\text{C}$ 。

通过上述公式求解亨利系数, 进而求出油气水三相种各组分摩尔分数。

1.2 油气水三相平衡方程

水相的逸度由 Prausnitz (1998) [19] 提出的公式计算,

$$f^{aq} = x^{aq} \varphi^{aqs} P_{vp}^{aqs} \left(\int_{P_{vp}^{aqs}}^P \frac{v_m}{RT} dP \right) \quad (7)$$

式中 x^{aq} 为水相的摩尔分数; P_{vp}^{aqs} 为水的饱和蒸气压, Pa; v_m 为水的摩尔体积, m^3/mol ; φ^{aqs} 为纯水的逸度系数。逸度系数由 Rowe (1970) [21] 提出的公式计算,

$$\varphi^{aqs} = \begin{cases} 0.9958 + 9.6833 \times 10^{-5} T - 6.175 \times 10^{-7} T^2 - 3.08333 \times 10^{-10} T^3 & T > 32.22^{\circ}\text{C} \\ 1 & T < 32.22^{\circ}\text{C} \end{cases} \quad (8)$$

当流体处于稳定时应满足热力学平衡, 三相中的任意化学势应相等。

$$\mu_{i1} = \mu_{i2} = \mu_{i3} \quad (9)$$

根据化学势、逸度定义将逸度定义式 $\mu_i - \mu_i^0 = RT \frac{\ln f_i}{\ln f_i^0}$, 代入到式 (9) 中得到三相热力学平衡条件,

$$\begin{aligned} f_i^L &= f_i^V \\ f_i^L &= f_i^{aq} \end{aligned} \quad (10)$$

式中 μ_i^0 和 f_i^0 分别是参考状态下组分 i 的化学势和逸度, R 为理想气体常数, $\text{J}/(\text{mol}^{\cdot 1} \cdot \text{K}^{-1})$, f 为逸度。除了满足热力学平衡条件外, 对于三相系统也要满足质量守恒。

假设 1 摩尔流体系统总物料组分为 $\{c_i, i=1,2,\dots,n_c\}$ 。系统在温度、压力下分为三相其中液相 L 组分构成为 $\{x_i^L, i=1,2,\dots,n_c\}$ 、气相 V 组分构成为 $\{y_i^V, i=1,2,\dots,n_c\}$ 、水相 aq 组分构成为 $\{x_i^{aq}, i=1,2,\dots,n_c\}$ 。

$$n^L + n^V + n^{aq} = 1 \quad (11)$$

式中 n^L 、 n^V 、 n^{aq} 分别为液相、气相、水相的相分率。对于每个组分有,

$$n^L x_i^L + n^V y_i^V + n^{aq} x_i^{aq} = c_i \quad (12)$$

x_i^L 为组分 i 在液相中的摩尔分数, y_i^V 为组分 i 在气相中的摩尔分数, x_i^{aq} 分别为组分 i 在水相中的摩尔分数; c_i 为组分 i 的总摩尔分数。

三相平衡常数为,

$$K_i^V = \frac{y_i^V}{x_i^L} \quad i = 1, 2, \dots, n_c \quad (13)$$

$$K_i^L = \frac{y_i^V}{x_i^{aq}} \quad i = 1, 2, \dots, n_c \quad (14)$$

式中 K_i^V 为组分 i 的气-液相平衡常数; K_i^L 为组分 i 的气-水相平衡常数。

每个组分需满足质量守恒,

$$\sum_{i=1}^{nc} c_i = 1 \quad (15)$$

$$\sum_{i=1}^{nc} x_i^L = 1 \quad (16)$$

$$\sum_{i=1}^{nc} x_i^{aq} = 1 \quad (17)$$

$$\sum_{i=1}^{nc} y_i^V = 1 \quad (18)$$

结合式 (12) 和平衡常数组分关于相分率与平衡常数的表达式,

$$x_i^L = \frac{c_i}{1 + n^V(K_i^V - 1) + n^{aq}(K_i^L - 1)} \quad (19)$$

$$x_i^{aq} = \frac{c_i K_i^L}{1 + n^V(K_i^V - 1) + n^{aq}(K_i^L - 1)} \quad (20)$$

$$x_i^V = \frac{c_i K_i^V}{1 + n^V(K_i^V - 1) + n^{aq}(K_i^L - 1)} \quad (21)$$

结合 R-R 方程 $f(n^V) = \sum_{i=1}^{nc} \frac{c_i(K_i - 1)}{1 + n^V(K_i - 1)} = 0$, 可以得到,

$$\sum_{i=1}^{nc} (x_i^V - x_i^L) = \sum_{i=1}^{nc} \frac{(1 - K_i^V) z_i}{1 - (1 - K_i^V)n^V - (1 - K_i^L)n^{aq}} = 0 \quad (22)$$

$$\sum_{i=1}^{nc} (x_i^V - x_i^{aq}) = \sum_{i=1}^{nc} \frac{(1 - K_i^L) z_i}{1 - (1 - K_i^V)n^V - (1 - K_i^L)n^{aq}} = 0 \quad (23)$$

式中 z_i 为组分 i 的总摩尔分数。

式 (22)、(23) 为多相 R-R 方程。对于三相相平衡来说共有 $3nc+3$ 个未知数分别为: nc 个气相组分 x_i^V ; nc 个液相组分 x_i^L ; nc 个水相组分 x_i^{aq} ; 以及三个相分率 n^{aq} 、 n^L 及 n^V 。求解未知数需要公式 (10) ~ (12)、(22)、(23) 提供 $3nc+3$ 个方程。

1.3 方程求解

油气水三相进行相平衡计算整体流程具体步骤如下:

- (1) 给定初始参数, 利用经验公式 Wilson 方程获取初始平衡常数 K_v 、 K_w ;
- (2) 结合初始平衡常数, 利用多相 R-R 方程求出初值 n^V 、 n^L 、 n^{aq} 及各组分在不同相间的摩尔分数;
- (3) 基于亨利定律、PR-EoS 结合 VdW 混合规则, 计算水相、气相及液相的逸度系数, 利用逸度系数更新平衡常数, 为提高稳定性对 K 取对数处理, 见式 (24)、(25);

$$g_i^V = \ln(K_i^V) + \ln(\varphi_i^V) - \ln(\varphi_i^L) = 0, i = 1, 2, 3, \dots, nc \quad (24)$$

$$g_i^{aq} = \ln(K_i^{aq}) + \ln(\varphi_i^{aq}) - \ln(\varphi_i^L) = 0, i = 1, 2, 3, \dots, nc \quad (25)$$

φ_i^V 为组分 i 在气相中的逸度系数; φ_i^L 为组分 i 在液相中的逸度系数; φ_i^{aq} 为组分 i 在水相中的逸度系数。

(4) 根据公式 (24) (25) 求解 nc 个变量, 获得新的 K 值; 用牛顿法求解公式 (26), 根据雅可比矩阵计算结果更新 K 值,

$$\psi^{k+1} = \psi^k - J_g^{-1}(\psi^k) \times g(\psi^k) \quad (26)$$

式中 ψ 为平衡常数的对数变量, J_g 为 g 的雅可比矩阵。

(5) 利用方程 (22) (23) 更新 n^V 、 n^L 、 n^{aq} , 若 $\sum(|\psi^{k+1} - \psi^k|) < \varepsilon$ 则计算收敛输出结果, 否则跳回第三步继续迭代更新, 整体流程如图 1 所示。

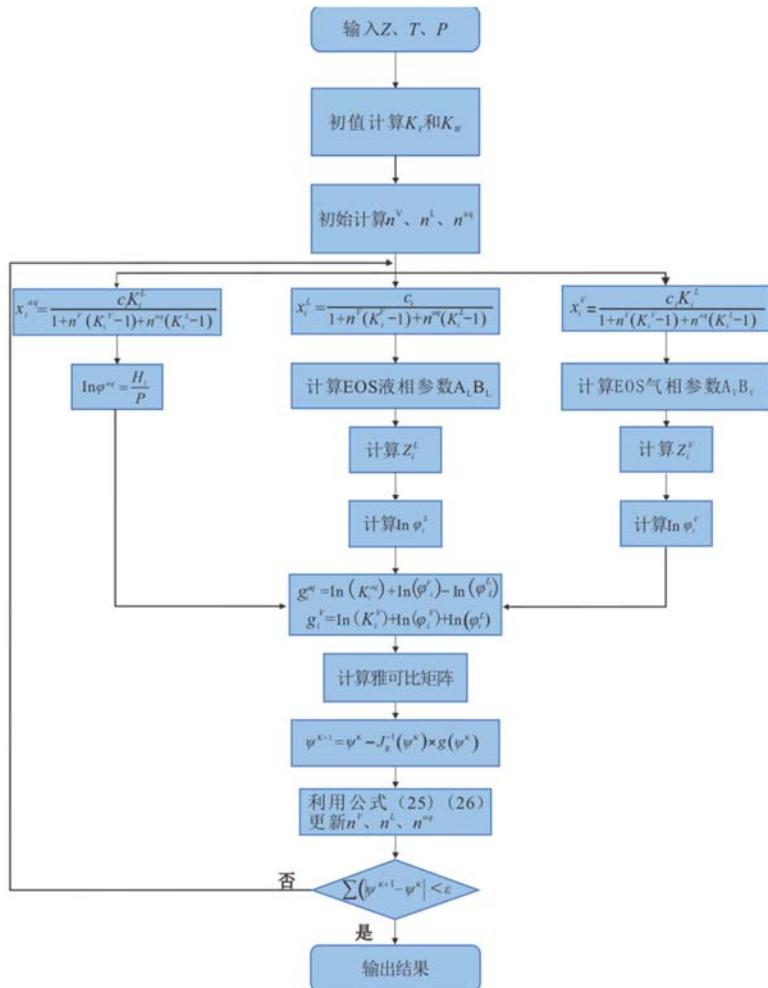


图 1 油气水三相相平衡计算流程图

Fig. 1 Flow chart of calculation of oil-gas-water three-phase phase balance

1.4 算例验证

本算例为一个多元体系油藏流体, 在温度 373.1K、10MPa 下的相平衡计算。模型基础参数见表 1, 二元交互作用参数 k_{ij} 见表 2, 表 3 为通过 Matlab 计算参数 Z^L 、 Z^V 、 Z^{aq} 、 n^V 、 n^L 、 n^{aq} 、相对误差验证模型的准确性。液相为烃类液体, 不包含水。图 2 为相平衡计算各组分在各相间的摩尔分数结果。

表 1 油藏流体组成及物性参数 (Nazari et al., 2019) [17]

Table 1 Reservoir fluid composition and physical property parameters

组分	Z (%)	Tc (K)	Pc (MPa)	w
----	-------	--------	----------	---

CH ₄	30	190.56	4.599	0.0115
nC ₅	15	469.70	3.370	0.2515
nC ₈	25	567.80	2.490	0.3996
CO ₂	10	304.19	7.382	0.2280
H ₂ S	10	373.53	8.963	0.0827
H ₂ O	10	647.13	22.055	0.3449

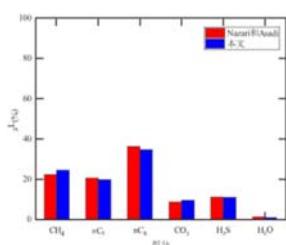
表 2 多元混合物二元相互作用系数 (K_{ij}) (Moradkhani et al., 2023) [16]Table 2 Binary Interaction Coefficient (K_{ij}) of multi-component mixture

	CH ₄	nC ₅	nC ₈	CO ₂	H ₂ S	H ₂ O
CH ₄	0	0.021	0.05	0.491	0.031	0.491
nC ₅	0.0206	0	0	0.125	0.095	0.500
nC ₈	0.0500	0	0	0.140	0.100	0.480
CO ₂	0.1030	0.125	0.14	0	0.096	0.200
H ₂ S	0.0310	0.095	0.10	0.096	0	0.120
H ₂ O	0.4907	0.500	0.48	0.200	0.120	0

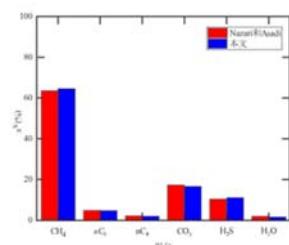
表 3 计算结果验证

Table 3 Verification of results

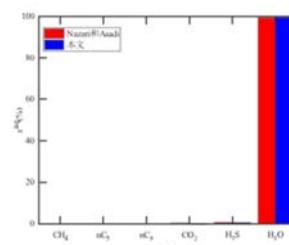
	Z^L		Z^V		Z^{aq}			
实验数据	计算结果	相对误差	实验数据	计算结果	相对误差	实验数据	计算结果	相对误差
0.3614	0.3936	0.089	0.8218	0.823	0.001	0.0776	0.0726	-0.064
	n^L						n^{aq}	
实验数据	计算结果	相对误差	实验数据	计算结果	相对误差	实验数据	计算结果	相对误差
67.61	71.28	0.054	23.6	19.59	-0.169	8.79	9.12	0.037



(a) 组分在液相(烃)间的摩尔分数占比



(b) 组分在气相间的摩尔分数占比



(c) 组分在水相间的摩尔分数占比

图 2 373.15K、10MPa 各组分摩尔分数结果

Fig. 2 Components molar fraction at 373.15K and 10MPa

从图 2、表 3 可以看出三相平衡模型计算的压缩因子、三相的相分率及各组分的摩尔分数与实验数据接近，计算精度高、误差较小说明模型具有稳定性。水的摩尔分数占据水相的 98%以上表明在水相中，水占据水相几乎全部的组成。为进一步分析整体气液(烃)水三相相分率变化情况，计算该流体 373.15K 温度不同压力下相平衡，图 3 为气液(烃)水三相相分率变化情况，该结果与文献中结果一致性较高，证实了该计算的可靠性。

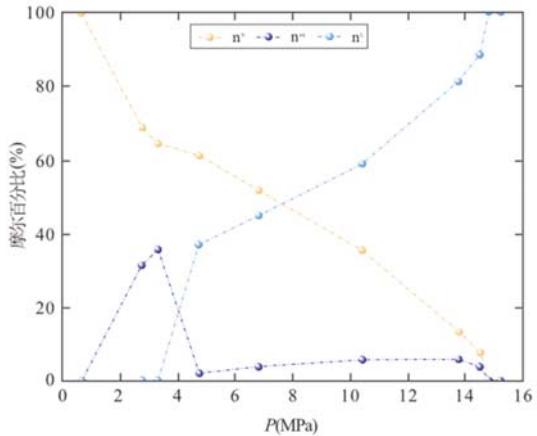


图 3 气液（烃）水三相相分率压力变化情况

Fig. 3 Pressure variation of phase fraction of gas-liquid (hydrocarbon) water

由图 3 可知在每一压力点气液（烃）水三相摩尔分数之和为 100% 符合质量守恒定律。随着压力的增加，气相摩尔分数逐渐减小，气相逐渐转变为液相，液相（烃）摩尔分数曲线呈上升趋势，水相摩尔分数曲线先增加后减小逐渐趋于零。水相摩尔分数在压力上升时先增加说明气体溶解于水中导致在低压时水相摩尔分数上升；但在进一步增加压力后，水相转变为液相的比例逐渐减少。随着压力的升高流体相态发生改变：压力在约 0.5MPa 时流体为单一气相，在 0.5~3MPa 下为气水两相，在 3~14.5MPa 下为气液（烃）水三相最终到 15MPa 下为单一液（烃）相。在约 3MPa 附近处于气水两相平衡，高于该压力点液（烃）相出现，且液（烃）相的比例急剧上升，气相部分转化为液（烃）相或水相。此时，水相的比例也达到极大值，随着压力增加分子间作用力增强使水相溶解度开始下降，水相最终趋于零。

因缺乏大量现场相平衡实验分析数据，因此利用该模型为人工神经网络算法提供基础数据。为模拟油藏中的真实环境，压力在 16~29MPa 下选取，温度在 343.15~413.15K 下选取，流体组分 CO₂、CH₄、nC₁₆、水的摩尔分数分别为 35%、30%、15%、20%，利用 Matlab 获得 92 组输入数据。

2 基于人工神经网络算法的油气水三相相平衡预测模型

利用实验与热力学模型进行多相多组分相平衡预测时，常存在成本高、耗时长与操作繁琐的问题。相平衡问题由于其数据类型单一，各参数均为数值型，使其成为一种非常适合利用数据驱动方法进行预测的问题。BPNN 结构清晰、可解释性强，便于与热力学框架对照，适合作为算法融合的稳定载体，也利于突出优化算法的贡献，同时其在工程领域应用成熟。BPNN 方法具有较强的非线性拟合能力和泛化性能，适用于油气水三相相平衡这类高维、非线性的回归问题。但传统 BPNN 在训练中容易陷入局部最优、产生收敛速度慢等问题。遗传优化算法（Genetic Algorithm, GA）作为全局优化算法，它具有较好的全局搜索能力，可有效探索解空间，避免陷入局部极小值；模拟退火优化算法（Simulated Annealing, SA）作为局部优化算法，通过概率突跳机制增强局部寻优能力，提高收敛稳定性。GA 与 SA 的引入并非替代传统梯度优化器，而是侧重于对网络初始权重、阈值及结构参数进行全局寻优，从而帮助模型摆脱局部最优、提升收敛稳定性。文章提出的 GASA-BPNN 融合了 GA 的全局搜索和 SA 的局部精细搜索能力，针对相平衡这类复杂问题具有更好的适应性，显著提升模型的收敛速度与预测精度，适用于复杂多相体系的相平衡计算，通过混合优化策略进一步提升模型的鲁棒性与泛化能力。因此，笔者利用人工神经方法中的 GA-BPNN 模型和 GASA-BPNN 模型构建油气水三相相平衡预测模型，对相分率进行计算与分析，神经网络模型算法逻辑图可在文献（邱锡鹏，2020）^[20]中查阅。

2.1 人工神经网络

在机器学习众多方法中 BPNN 作为一种多层前馈神经网络，有着其独特的反向传播算法和多层非线性结构，在人工智能领域展现出显著优势。神经网络在训练后具备较强的泛化能力，对未见数据做出合理预测，因此本文选用 BPNN 模型开展相平衡预测。BPNN 模型数据传输网络第一层是输入层，包含两个神经元 x_1 和 x_2 和截距项 b_1 。第二层是隐含层，包含两个神经元 h_1 和 h_2 和截距项 b_2 ；第三层是输出层，包含两个神经元 o_1 和 o_2 。每条线上标的 w_i 是不同层神经元之间连接的权重。 a 和 b 是用于调整截距项 b_1 、 b_2 的参数。BPNN 分为前向传播和反向传播。BPNN 工作流程分为迭代和优化两步：第一步，输入数据逐层前传，经加权计算和激活函数处理得到预测结果；第二步，计算预测值与真实值的误差，反向传播时利用链式法则逐层分析各权重对误差的影响程度，并用梯度下降法调整权重和阈值。每输入一个样本完成一次误差反传和参数更新，通过反复迭代逐步优化网络参数，最终使误差达到预设阈值，形成稳定且精准的输入输出映射模型。

完成网络训练后即可获得参数优化后的 BPNN 模型。为验证模型的泛化能力和可靠性，采用未训练的测试集对训练的模型进行泛化能力评估。通过预测值与实验值进行对比，构建误差评判体系讨论模型的预测性能。笔者选取四项统计学指标构建综合评价体系：决定系数 (R^2)、均方误差 (MSE)、均方根误差 (RMSE) 以及平均绝对偏差 (MAE)。上述误差评判能够表征模型预测与实验值之间的吻合度，其数学定义及计算方法见式 (27) ~ (30)。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (27)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (28)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (29)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (30)$$

公式中 y_i 和 \hat{y}_i 分别为第 i 个输入数据的真实值和预测值， \bar{y} 为真实值的平均值， n 为样本数。

2.2 遗传优化算法 (Genetic Algorithm, GA)

遗传算法最初源于利用计算机模拟生物系统而发展起来，它是一种随机全局搜索优化方法。该算法最早由美国密歇根大学 Holland 提出 (1992)^[9]，基于自然选择和遗传进化机制而建立。通过模拟生物界中复制、交叉和变异等过程，遗传算法从一个初始种群出发，利用随机选择、交叉和变异操作不断产生更适应环境的新个体，使得种群在搜索空间中逐步向最优区域演化，最终收敛到一个最优解的群体，从而实现对问题的高效优化。遗传算法的核心结构及其操作模块主要包括染色体编码方式、适应度函数、算法运行参数，以及三种基本遗传操作（选择、交叉和变异）。

(1) 染色体编码：染色体编码方式是将问题的解转换为遗传算法可操作的“染色体”形式，其核心在于如何用合适的方式对问题解进行数字化或符号串描述，从而使得遗传算法中的交叉、变异等操作能够在编码空间内有效地搜索最优解。目前染色体编码方式主要分为两种二进制编码和实数编码。与二进制编码相比，实数编码更直观又简便。在 GA 优化 BPNN 模型过程中，每个染色体代表一个初始的权值和阈值。(2) 适应度函数：该算法选用均方误差作为适应度函数，在回归问题中是一种有效的做法，能够量化模型的预

测误差，并为遗传算法提供清晰的优化目标。利用适应度函数，遗传算法通过迭代搜索最优解，提高回归模型的预测准确性。（3）选择操作：通过选择算子模拟“优胜劣汰”在遗传算法中占据着重要地位，即以一定的概率从当前种群中选择父代染色体来产生子代染色体。适应度高的算子被遗传到下一代的概率较大，适应度低的算子被遗传到下一代的概率较小。目前有多种选择方法如排名选择、精英选择、轮盘赌选择等；笔者选用轮盘赌选择法，该方法直观且操作简单。

$$p_i = \frac{f_i}{\sum_{i=1}^n f_i} \quad (31)$$

式中 p_i 为选择因子； f_i ($i=1,2,\dots,m$) 为种群中第 i 个个体的适应度函数。

（4）交叉操作：模仿自然界中基因的交换过程，将两个“父代”个体的部分遗传信息交换，生成一个或多个“子代”个体。交叉操作的目的是通过组合不同父代个体的优良特性来创造更优秀的后代个体。子代个体产生后，计算其适应度值，将具有最优适应度的子代个体重新放回原来的父代种群中。（5）变异操作：在自然界中，基因变异增加了个体的多样性，对于遗传算法来说增加了个体的多样性，可以增加找到最优解的概率。通过这一步骤可以使 GA 更好的搜索全局空间并有效防止过拟合。在 GA 中选择操作代表种群向适应度高的方向进化，交叉操作加速 GA 搜索过程，变异操作代表增加种群的多样性并防止过拟合的发生。通过这三种操作的有效结合，GA 更有效地逼近目标值。

2.3 模拟退火优化算法 (Simulated Annealing, SA)

模拟退火算法最早由 Kirkpatrick、Gelatt 和 Vecchi 三位学者提出 (Kirkpatrick et al., 1983) [13]。该算法从较高的初始温度开始，通过温度逐步降低，在解空间中结合概率突跳特性进行随机搜索，寻找目标函数的全局最优解。在搜索过程中，算法具有跳出局部最优解的能力，这种跳出是通过概率方式实现的，并且随着温度的下降，跳出局部最优的概率逐渐减小，最终使得算法趋向全局最优解。模拟退火算法有效避免陷入局部极小值，并确保最终收敛到全局最优解。

模拟退火算法由两个部分构成：Metropolis 算法和退火过程，分别对应内循环和外循环。内循环为 Metropolis 算法在每个温度保持不变的条件下对系统进行迭代，寻找该温度下的最小能量值作为最优解。外循环为退火过程，将系统加热至高温初始温度 T_{max} ，通过降温系数 α 使温度按一定比例逐渐下降。当温度降至终止温度 T_{min} 时，冷却过程结束，退火过程随之完成。在每一次迭代中，模拟退火算法会计算出新状态的能量与当前状态能量比较。如果新状态的能量比当前状态低，则接受新状态。如果新状态的能量更高，则根据公式 (32) 以一定的概率接受这个新状态并跳出循环。这个概率与温度和能量差有关。随着温度的降低，这个概率逐渐减小，使得算法更加稳定，趋向于选择能量更低的状态。该算法可与多种全局优化算法相融合形成新的算法，既保持全局搜索的广度，又提升局部优化的深度，尤其适用于组合优化等易陷入局部最优的问题，能有效提高解的质量和算法效率。

$$P = \begin{cases} 1 & E(n+1) < E(n) \\ e^{-\frac{E(n+1)-E(n)}{T}} & E(n+1) \geq E(n) \end{cases} \quad (32)$$

式中 P 为接受新状态概率； $E(n)$ 为当前状态（第 n 次迭代时）的能量； $E(n+1)$ 为新状态（第 $n+1$ 次迭代时）的能量； T 为当前温度； $e^{-\frac{E(n+1)-E(n)}{T}}$ 为 Metropolis 接受准则。

3 结果与讨论

将神经网络与遗传算法和模拟退火算法相结合，利用优化算法来优化反向传播神经网络的权重和偏置，从而提高神经网络的性能。利用 Matlab 软件编写程序，通过设置循环实现自动仿真。参照大庆油田 X 区块

地下原油组分构成（高含蜡原油），选取典型轻-重组分（CH₄、nC₁₆）模拟原油构成。模型输入参数为组分和温压条件（CO₂、CH₄、nC₁₆、水、温度、压力，共6个输入变量）；输出变量为相分率及各组分在油气水三相间摩尔分数，共15个输出项。将92组数据按照3:1进行随机划分，其中训练集69组、测试集23组。由于不同参数特征尺度差异会导致训练过程不稳定、收敛速度慢，因此，在训练前对数据进行归一化。同时，由于本研究样本容量共92个，为避免过拟合问题，采用交叉验证的方法提高模型的泛化能力。利用交叉验证的思想每次在训练集中选择1个样本轮空，剩余的68个数据作为训练样本，因此共有68个互不相容的训练样本，利用不相容的训练集创建68个BPNN，每个神经网络的结构相同，同取68个神经网络对测试集预测结果的平均值作为最终预测值。

3.1 不同优化算法预测对比

3.1.1 GASA-BPNN 模型结构

利用遗传-模拟退火混合算法优化神经网络（GASA-BPNN），通过融合模拟退火机制与变异算子构建混合优化策略，该模型在遗传算法全局搜索能力的基础上，引入退火概率函数增强局部寻优效率，改进后的GASA-BPNN在高维非线性油气藏数据建模中展现出更快的收敛速度与预测稳定性。GASA-BPNN整体流程如图4所示。

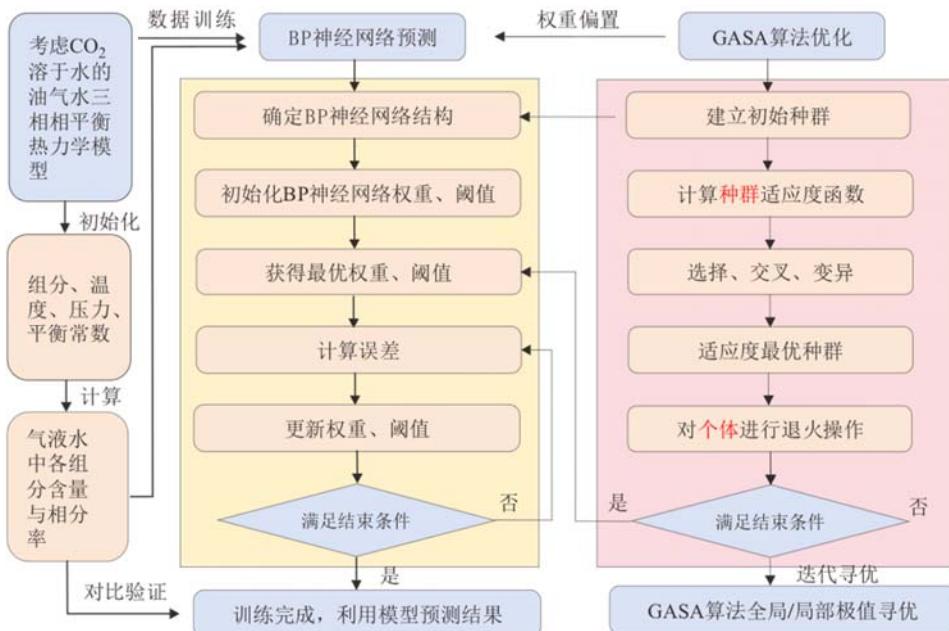


图4 GASA-BPNN 流程图

Fig.4 GASA-BPNN flow chart

表4列出GASA-BPNN模型的结构参数值、图5、6为模型参数与指标评判参数关系。

表4 模拟预测相平衡计算的最优GASA-BPNN模型结构参数

Table 4 Simulate and predict the optimal GASA-BPNN model structural parameters for phase equilibrium calculations

参数	值	参数	值
隐含层层数	1	最大迭代次数	1000
隐含层神经元个数	16	目标误差	0.0025
输入-隐含层的激活函数	Tanh	种群规模	50
隐含层-输出层的激活函数	Purelin	最大迭代次数	100
训练算法	Trainlm	交叉算子	0.3

学习速度	0.01	变异算子	0.1
动量因子	0.65	初始温度	100
		截止温度	30

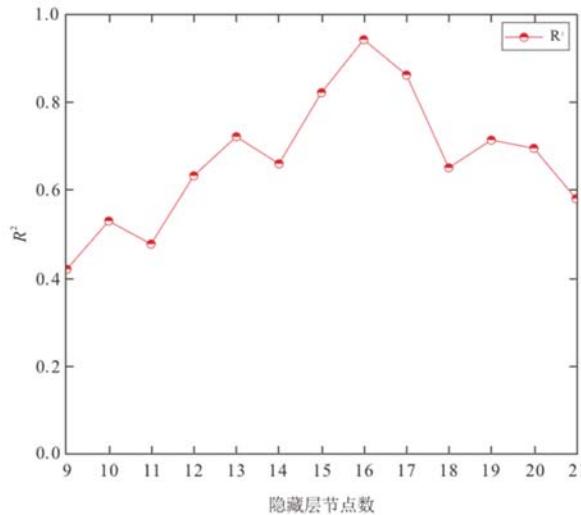


图 5 隐藏层神经元个数选取

Fig. 5 Selection of the number of neurons in the hidden layer

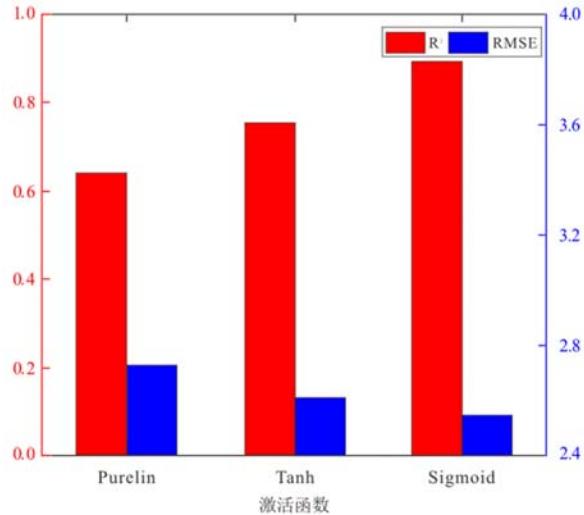


图 6 激活函数选取

Fig. 6 Selection of activation function

3.1.2 GASA-BPNN 与 GA-BPNN 模型训练结果

将 GA-BPNN 和 GASA-BPNN 两种模型中适应度计算最优的个体即权值和阈值代入到 BPNN 模型中训练。图 7 为两种算法适应度计算。将测试集代入模型中预测，通过误差指标评判（MAE、MSE、RMSE、 R^2 ）选出最优模型，图 8 为三种算法训练误差。

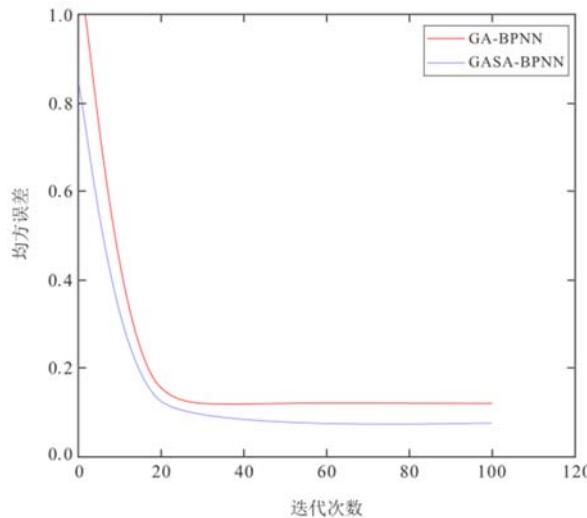


图 7 GA-BPNN 和 GASA-BPNN 计算适应度函数

Fig. 7 GA-BPNN and GASA-BPNN calculate fitness function

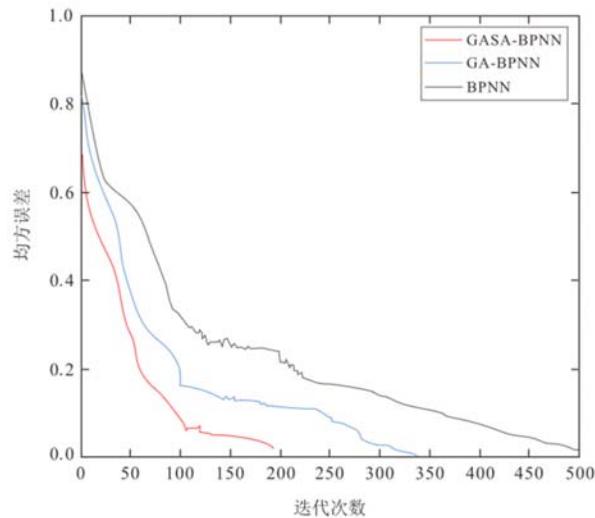


图 8 三种算法训练误差

Fig. 8 Training errors of the three algorithms

完成本文条件下的 BPNN 模型（含 GA 或 GASA 优化）的训练与测试，总耗时小于 2 分钟，优化算法的迭代次数不同时具体时长略有变化。将模型目标误差设置为 0.0025，当训练误差下降而验证集误差上升

即触发早停，防止因过度调参或训练导致的过拟合，确保了结果的稳定性。图9~图14为GA-BPNN和GASA-BPNN两种模型的训练结果及对应的误差值。由于仅考虑了温度和压力因素，模型中气相分率摩尔分数整体分布在0~50%，液（烃）相分率摩尔分数整体分布在30~80%，水相分率摩尔分数整体分布在17~20%。GA-BPNN和GASA-BPNN两种神经网络训练模型预测结果均在标准线附近，其中GASA-BPNN比GA-BPNN更靠近标准线表明GASA-BPNN训练模型预测性能更好。同时，三相分率的相对误差均在±5%以内，两种训练模型均能准确预测三相分率，而GASA-BPNN模型的相对误差集中在±1%以内表明该模型预测效果最好。

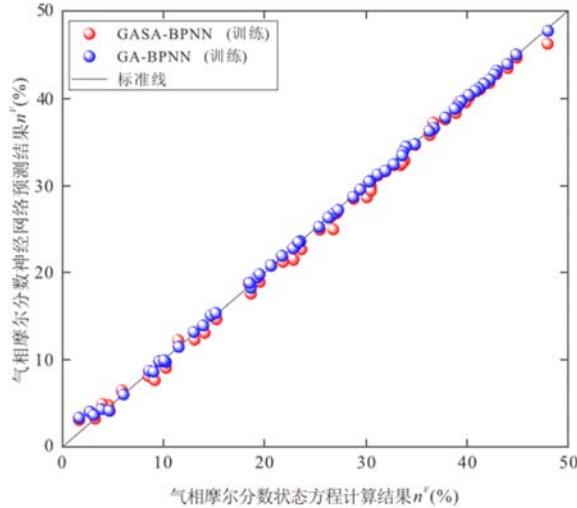


图9 GA-BPNN 和 GASA-BPNN 模型
 n^V 训练结果

Fig. 9 n^V training results of GA-BPNN and GASA-BPNN models

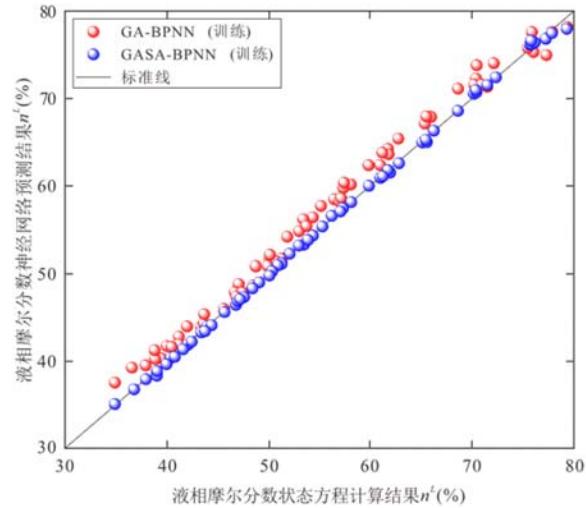


图10 GA-BPNN 和 GASA-BPNN 模型
 n^L 训练结果

Fig. 10 n^L training results of GA-BPNN and GASA-BPNN models

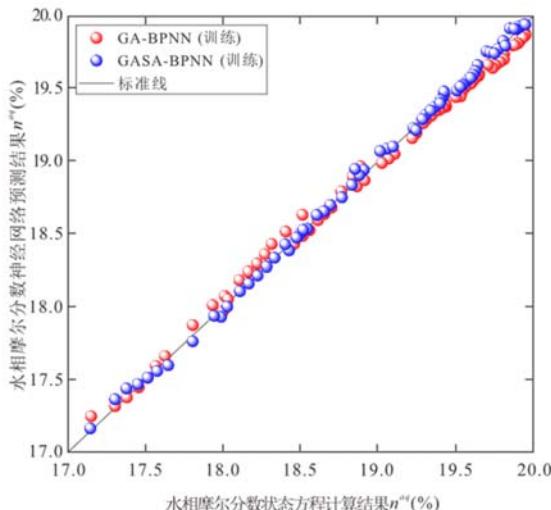


图11 GA-BPNN 和 GASA-BPNN 模型
 n^{aq} 训练结果

Fig. 11 n^{aq} training results of GA-BPNN and GASA-BPNN models

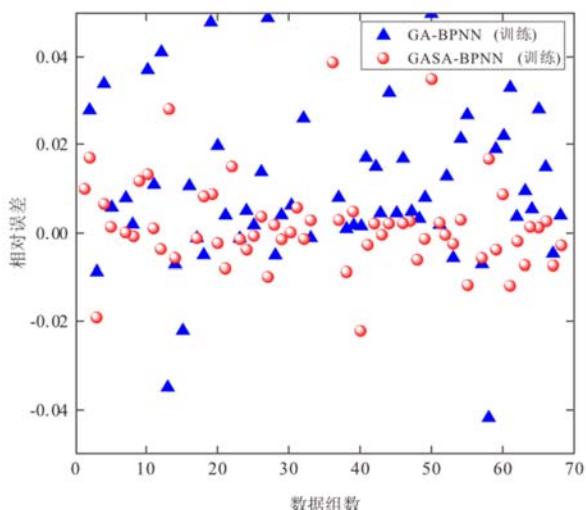


图12 GA-BPNN 和 GASA-BPNN 模型
 n^V 训练误差

Fig. 12 n^V training errors of GA-BPNN and GASA-BPNN models

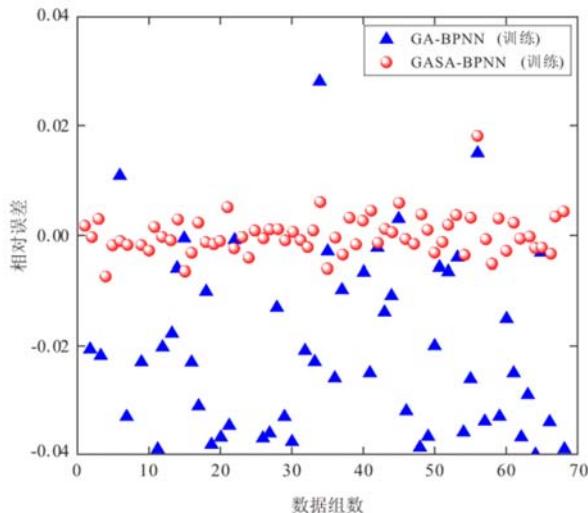


图 13 GA-BPNN 和 GASA-BPNN 模型

n^L 训练误差
Fig. 13 n^L training errors of GA-BPNN and GASA-BPNN models

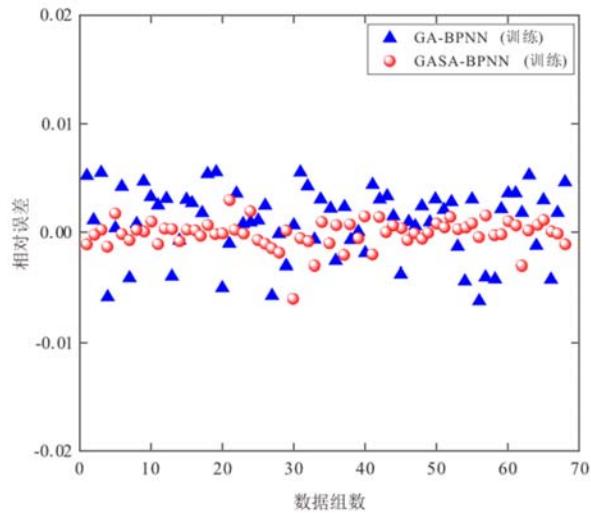


图 14 GA-BPNN 和 GASA-BPNN 模型

n^{aq} 训练误差
Fig. 14 n^{aq} training errors of GA-BPNN and GASA-BPNN models

3.1.3 GASA-BPNN 与 GA-BPNN 相平衡模型测试结果

从训练结果中看出 GA-BPNN 和 GASA-BPNN 两种模型能够预测相平衡，但还需进行测试将未处理的数据代入验证。表 5 为神经网络整体训练和测试的误差评判指标，可看出两种模型的训练效果均优于测试效果；通过两个模型 R^2 对比，GASA-BPNN 具有更高 R^2 ，说明该模型的泛化能力更强。图 15~图 20 为 GA-BPNN 和 GASA-BPNN 两种模型的测试结果及对应的误差值。GASA-BPNN 模型误差评判指标 MSE、RMSE、MAE 值最小，证明该模型准确率最高，能够准确预测相平衡。

表 5 神经网络模型测试集误差评判指标

Table 5 Error evaluation index of neural network model test set

预测模型		R^2	MSE	RMSE	MAE
GA-BPNN	训练值	0.9899	0.35	0.59	2.52
	测试值	0.9821	0.78	0.88	6.17
GASA-BPNN	训练值	0.9992	0.15	0.38	2.19
	测试值	0.9951	0.56	0.74	5.89

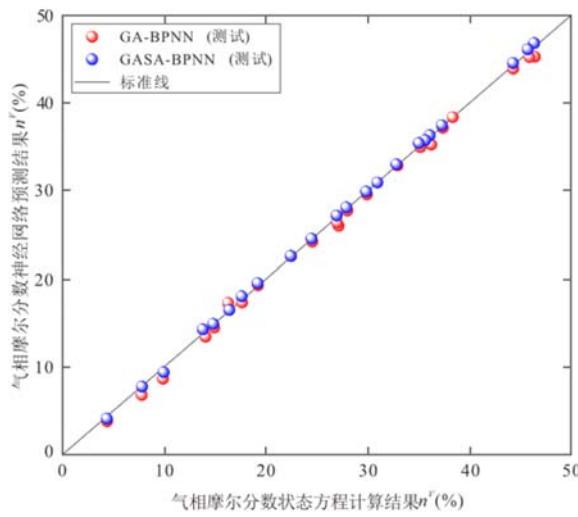


图 15 GA-BPNN 和 GASA-BPNN 模型
 n^V 测试结果

Fig. 15 n^V test results of GA-BPNN and GASA-BPNN models

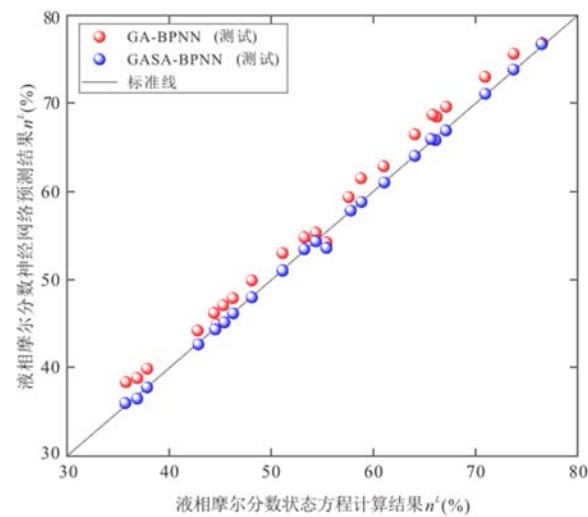


图 16 GA-BPNN 和 GASA-BPNN 模型
 n^L 测试结果

Fig. 16 n^L test results of GA-BPNN and GASA-BPNN models

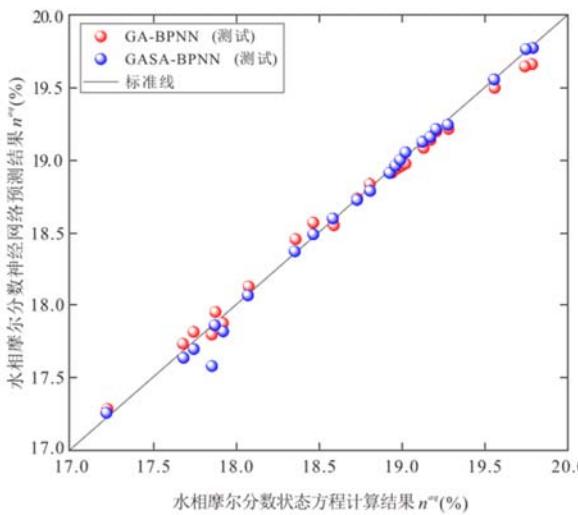


图 17 GA-BPNN 和 GASA-BPNN 模型
 n^{aq} 测试结果

Fig. 17 n^{aq} test results of GA-BPNN and GASA-BPNN models

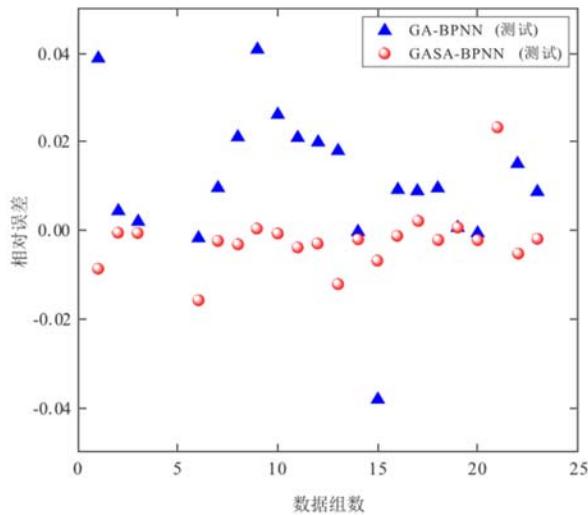


图 18 GA-BPNN 和 GASA-BPNN 模型
 n^V 测试误差

Fig. 18 n^V test error of GA-BPNN and GASA-BPNN models

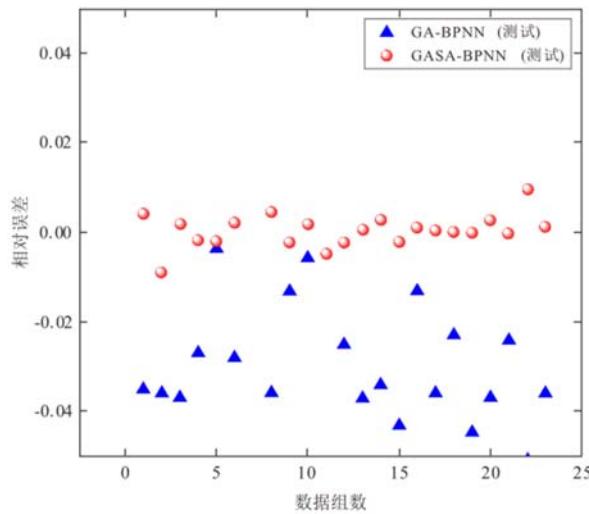


图 19 GA-BPNN 和 GASA-BPNN 模型

Fig. 19 n^L test error of GA-BPNN and GASA-BPNN models

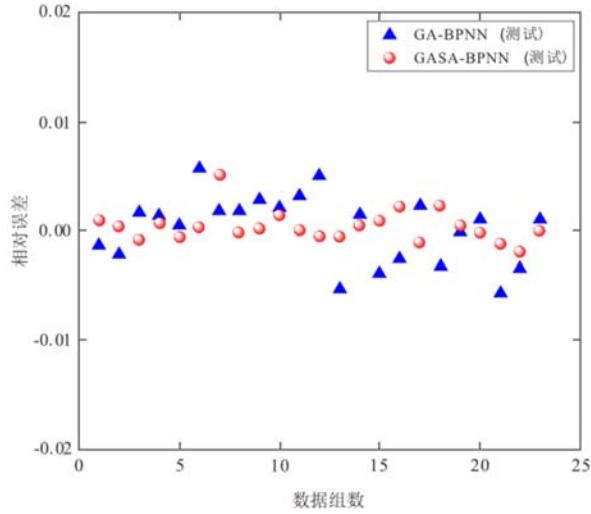


图 20 GA-BPNN 和 GASA-BPNN 模型

Fig. 20 n^{aq} test error of GA-BPNN and GASA-BPNN models

3.2 算例分析

依据华北油田油样在温度 323.15K、压力 1.02MPa 下的相平衡实验参数 (Zheng et al., 2023)^[29] 开展 GASA-BPNN 模型油气 (含 CO₂) 两相算例分析。实验结果数据见表 6; GASA-BNPP 模型验证结果见表 7, 同时表 7 中还对 GASA-BNPP 和热力学模型计算时间进行了对比; 图 21 为油气两相相分率变化情况。通过表 7 可知 GASA-BPNN 在预测 CO₂-原油体系相平衡上具有良好的精度, 组分及气液相摩尔分数预测结果与文献相比误差率均低于 5%。通过图 21 可知, 随着温度由 276.15K 增至 363.15K, 体系中的气相分率逐渐上升而液相分率逐渐下降。每组气相分率与液相分率摩尔分数之和为 100%, 满足质量守恒定律。

表 6 华北油田油样物性参数

Table 6 Physical property parameters of oil samples in Huabei Oilfield

组成	Mw (/g.mol ⁻¹)	T _c (K)	P _c (MPa)	w	C _i (mol)
CO ₂	44.01	304.2	7.37	0.225	0.25
C ₃ H ₈	44.01	369.8	4.25	0.008	0.25
C ₉₋₁₃	140.7	614.99	2.01	0.4998	0.0937
C ₁₄₋₁₈	224.3	719.85	1.43	0.755	0.09325
C ₁₉₋₂₄	295.5	785.66	1.17	0.9311	0.08575
C ₂₅₋₃₁	386.8	850.55	0.97	1.0909	0.0801
C ₃₂₋₄₈	546.1	929.86	0.82	1.2464	0.0788
C ₄₉₋₁₁₃ ⁺	1277.8	1090.76	0.79	1.3581	0.0684

表 7 两相相平衡计算结果对比表

Table 7 Comparison of phase equilibrium results for two phases

组成	x _i , mol%			y _i , mol%			时间,s	
	GASA-BPNN	实验	误差	GASA-BPNN	实验	误差	GASA-BPNN	PR-EoS
CO ₂	6.87	6.95	1.1%	70.59	71.11	0.7%		
C ₃ H ₈	14.12	13.68	3.2%	29.41	28.89	1.8%	10.27	21.78
C ₉₋₁₃	14.72	14.59	0.9%	0	0	0		

C ₁₄₋₁₈	15.89	15.91	0.1%	0	0	0
C ₁₉₋₂₄	14.38	14.77	2.6%	0	0	0
C ₂₅₋₃₁	13.73	13.75	0.1%	0	0	0
C ₃₂₋₁₁₃₊	20.29	20.35	0.3%	0	0	0
n ^V ,mol%						n ^L ,mol%
相分率	GASA-BPNN	实验	误差	GASA-BPNN	实验	误差
	27.68	28.11	-1.5%	72.32	71.89	0.5%

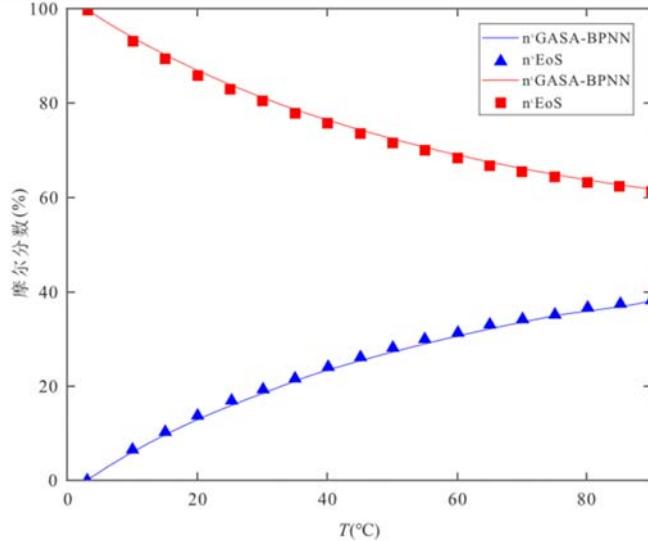


图 21 气液相分率变化情况

Fig. 21 Change of gas-liquid phase fraction

传统状态方程在分析多相多组分原油体系时，计算时间随组分复杂度呈指数级增长。相比之下，利用神经网络方法通过训练模型，可实现输入输出的快速预测，无需对中间变量逸度系数、平衡常数等进行计算，且仅需少量 PVT 实验数据即可构建全局代理模型，缓解了数据稀缺性问题。但需特别指出，人工神经网络方法的预测能力严格受限于训练集覆盖的原油物性范围，若应用于组分差异显著的其他原油体系，需补充实验数据并重新训练网络以保障预测精度。

利用该模型计算储层中某一点的温度和压力下油气水三相摩尔分数。由于未找到油气水三相多组分相平衡摩尔分数实验研究结果，利用本文 1 中方法计算得到的三相相平衡数据，开展温度 413.15K、压力 29.5MPa 下 GASA-BNPP 相平衡计算，如表 8 所示。图 22 为油气水三相中各组分摩尔分数分布情况，从图中可以看出 GASA-BPNN 模型结果与 EoS 模型结果几乎一致，表明该模型能够进行准确预测。随着压力的升高气相中 CO₂ 摩尔分数下降，由于分子间的相互作用力影响，液相中 CO₂ 摩尔分数呈上升趋势，与油互溶使油膨胀。随着压力增加 CO₂ 溶解于水，导致液相 CO₂ 摩尔分数增加，水相中的水摩尔分数减小，符合质量守恒定律。

表 8 三相相平衡计算结果对比表

Table 8 Comparison of phase equilibrium results for three phases

	x _i ,mol%		y _i ,mol%		x _i ^{ad} ,mol%	
	GASA-BPNN	EoS	GASA-BPNN	EoS	GASA-BPNN	EoS
CO ₂	42.074	42.183	49.169	48.533	1.503	1.488
H ₂ O	2.525	2.667	3.917	3.992	98.497	98.512
CH ₄	36.132	36.206	44.894	45.981	0	0
nC ₁₆	19.269	18.945	2.019	1.493	0	0

相分率	$n^V, \text{mol\%}$		$n^L, \text{mol\%}$		$n^{aq}, \text{mol\%}$	
	GASA-BPNN	EoS	GASA-BPNN	EoS	GASA-BPNN	EoS
	4.48	3.02	77.38	78.94	18.14	18.04

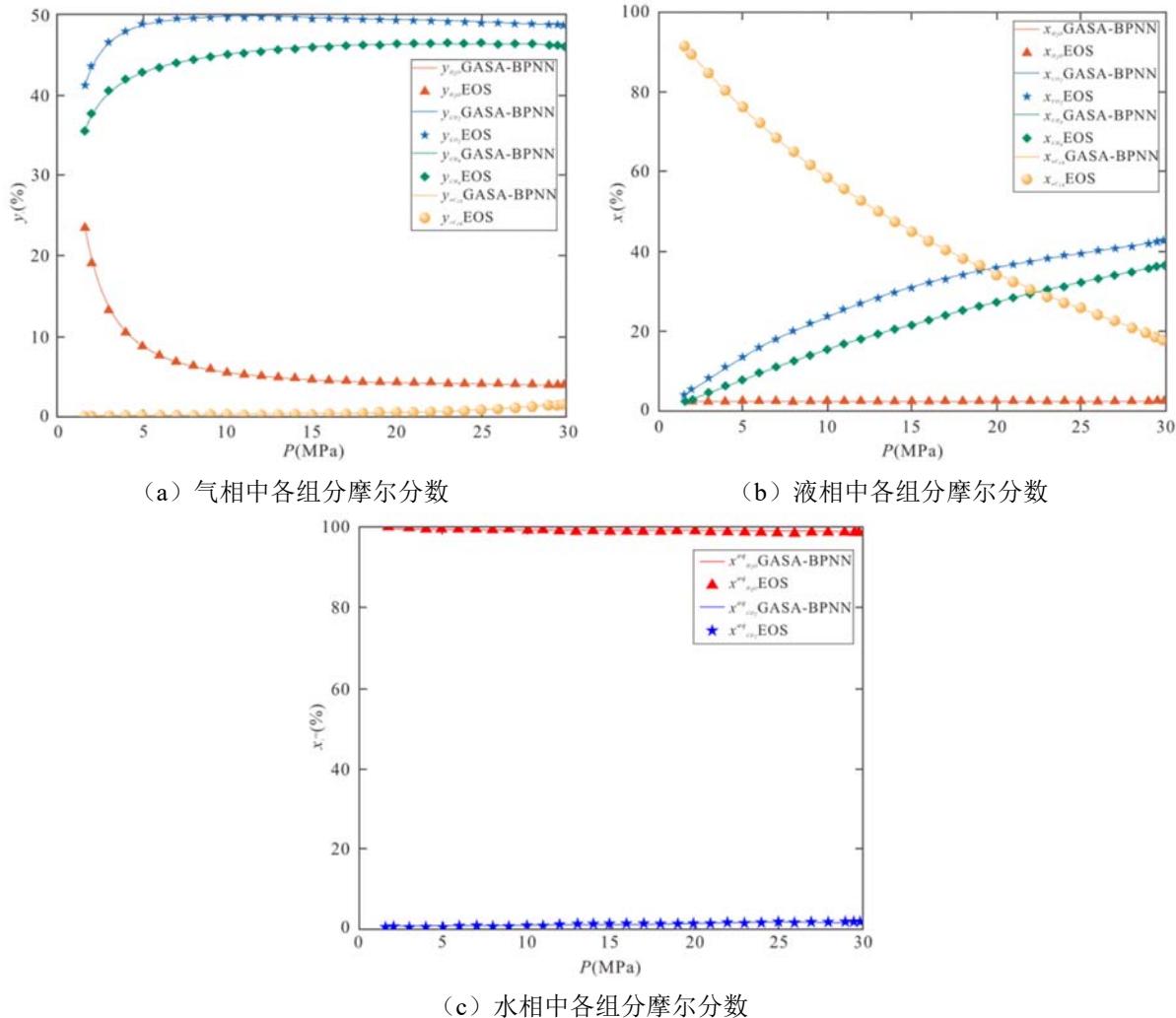


图 22 GASA-BPNN 模型与 EoS 模型摩尔分数计算结果

Fig. 22 Mole fraction calculation results of GASA-BPNN model and EoS model

4 结论

本文构建三种相平衡模型用于预测油气水三相相分率，分别为：利用传统热力学方法考虑 CO_2 溶解于水的油气水三相相平衡模型，利用人工神经网络方法构建 GA-BPNN 与 GASA-BPNN 模型。

(1) 通过公式推导得出水相的逸度系数与状态方程的油相水相逸度系数联立，相平衡方程求解结果与实验数据相比较误差较小，证明模型的可靠性。

(2) GASA-BPNN 通过全局搜索 (GA) 与局部寻优 (SA) 的协同机制，显著提升模型收敛速度迭代次数减小，测试集 R^2 值为 0.9951，优于单一优化模型。

(3) GA-BPNN 与 GASA-BPNN 两种模型均能准确预测油气水三相相分率, 相较而言 GASA-BPNN 精度更高, 泛化能力更强。可用于油气水三相相平衡预测, 与传统热力学模型相比, 可实现缺少实验数据时的准确计算, 该模型可用于实际工程中。

References

- [1] Bai,L.H.,Liu,B.,Chi,Y.A.,et al.,2021.2D NMR Studies of Fluids in Organic-Rich Shale from the Qingshankou Formation, Songliao Basin. *Oil & Gas Geology*, 42(06):1389-1400. doi: 10.11743/ogg20210613. (in Chinese with English abstract)
- [2] Cai,J.,Yang,X.,Yang,Z.,et al.,2025.Microscopic Flow and Reactive Transport in Geological Media: Recent Advances and Challenges. *Capillarity*, 17(3):77-80. doi: 10.46690/capi.2025.12.01.
- [3] Chen,Z.,Li,R.X.,Du,Y.F.,et al.,2025.Effect of Confinement on the Vapor-Liquid-Liquid Three-Phase Equilibrium During CO₂ Utilization and Sequestration in Shale Reservoirs. *Advances in Geo-Energy Research*, 16(3):199-210. doi: 10.46690/ager.2025.06.02.
- [4] Di,Y.,Zhang,Y.,W,Y.S.,2015.Phase Equilibrium Calculation of CO₂-Hydrocarbon-Water Systems Based on Gibbs Free Energy Minimization. *Acta Petrolei Sinica*, 36(5):593-599. doi: 10.7623/syxb201505008. (in Chinese with English abstract)
- [5] Garcia,J.E.2001.,Density of Aqueous Solutions of CO₂. *Office of Scientific & Technical Information Technical Reports*.
- [6] Guo,K.Q.2012. The Development of Phase Equilibrium Database Software.Qingdao University of Science and Technology, Qingdao. (in Chinese with English abstract)
- [7] Han,F.,2004. High Pressure Vapor-Liquid Equilibria of Two Binary Systems Containing Carbon Dioxide and Alkanes(Dissertation).Tianjin University, Tianjin. (in Chinese with English abstract)
- [8] Harvey,A.H.,1996.Semiempirical Correlation for Henry's Constants Over Large Temperature Ranges. *Aiche Journal*, 42(5):1491-1494. doi: 10.1002/AIC.690420531.
- [9] Holland,J.H.,1992.Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. The MIT Press, Cambridge. doi: 10.7551/mitpress/1090.001.0001.
- [10] Huang,C.,Tian,L.,Zhang,T.Y.,et al.,2022.Globally Optimized Machine-Learning Framework for CO₂-Hydrocarbon Minimum Miscibility Pressure Calculations. *Fuel*, 329:125312. doi: 10.1016/j.fuel.2022.125312.
- [11] Hui,G.,Ren,Y.L.,Bi,J.F.,et al.,2025.Artificial Intelligence Applications and Challenges in Oil and Gas Exploration and Development. *Advances in Geo-Energy Research*, 17(3):179-183. doi: 10.46690/ager.2025.09.01.
- [12] Kariznovi,M.,Nourozieh,H.,Abedi,J.,2013.Phase Composition and Saturated Liquid Properties in Binary and Ternary Systems Containing Carbon Dioxide, n-Decane, and n-Tetradecane. *The Journal of Chemical Thermodynamics*, 57:189-196. doi: 10.1016/j.jct.2012.08.019.
- [13] Kirkpatrick,S.,GelattJr,C.D.,Vecchi,M.P.,1983.Optimization by Simulated Annealing. *Science*, 220(13):671-680. doi: 10.1126/science.220.4598.671.
- [14] Liu,B.,Lü,Y.F.,Ran,Q.C.,et al.,2014.Geological Conditions and Exploration Potential of Shale Oil in Qingshankou Formation, Northern Songliao Basin. *Oil & Gas Geology*, 35(2):280-285. doi: 10.11743/ogg20140216. (in Chinese with English abstract)
- [15] Liu,G.F.,He,L.Y.,Fan,Z.Q.,et al.,2021.Investigation of Gas Solubility and Its Effects on Natural Gas Reserve and Production in Tight Formations. *Fuel*, 295:120507. doi: 10.1016/j.fuel.2021.120507.
- [16] Moradkhani,M.A.,Hosseini,S.H.,Ranjbar,K.,2023.Universal Intelligent Models for Liquid Density of CO₂+Hydrocarbon Mixtures. *Fuel*, 334(P2):126642. doi: 10.1016/j.fuel.2022.126642.
- [17] Nazari,M.,Asadi,M.B.,Zendehboudi,S.,2019.A New Efficient Algorithm to Determine Three-Phase Equilibrium Conditions in the Presence of Aqueous Phase: Phase Stability and Computational Cost. *Fluid Phase Equilibria*,

- 486(1):139-158. doi: 10.1016/j.fluid.2018.12.013.
- [18] Pan,Y.S.,Liu,B.,Zhao,Z.,et al.,2025. Diagenetic evolution and reservoir formation mechanisms of tuff-rich mixed shales: A case study of the Lucaogou Formation in the Santanghu Basin. *Oil & Gas Geology*, 1-20. (in Chinese with English abstract)
- [19] Prausnitz,J.M.,Lichtenthaler,R.N.,Azevedo,E.G.D.,1998. *Molecular Thermodynamics of Fluid-Phase Equilibria*. Pearson Education,London.
- [20] Qiu,X.P.,2020. *Neural Networks and Deep Learning*. China Machine Press,Beijing (in Chinese).
- [21] Rowe,A.M.,Chou,J.C.S.,1970. Pressure-Volume-Temperature-Concentration Relation of Aqueous Sodium Chloride Solutions. *Journal of Chemical and Engineering Data*, 15(1):61-66. doi: 10.1021/je60044a016.
- [22] Rosendo,M.,2023. An Efficient, Robust and Reliable Initialization of Phase Fractions in Rachford-Rice System of Equations for Three-Phase Split Calculation. *Fluid Phase Equilibria*, 566:113690. doi: 10.1016/j.fluid.2022.113690.
- [23] Shi,W.,Huan,G.R.,Guo,S.P.,1992. Study on Flash Calculation Method for Phase Equilibrium of CO₂-Hydrocarbon-Water System. *Petroleum Exploration and Development*, 19(03):48-56. (in Chinese)
- [24] Sun,G.L.,Zhao,Z.Y.,Sun,S.J.,et al.2023. Vapor-Liquid Phase Equilibria Behavior Prediction of Binary Mixtures Using Machine Learning. *Chemical Engineering Science*, 282:119358. doi: 10.1016/j.ces.2023.119358.
- [25] Wagner,W.,Pruss,A.,1987. International Equations for the Saturation Properties of Ordinary Water Substance. *Journal of Physical and Chemical Reference Data*, 16(4): 893-901. doi: 10.1063/1.555926.
- [26] Xue,Y.A.,Hu,A.W.,Wan,L.,et al.,2025. Accumulation Process and Phase Control Factors of Large Natural Gas Reservoirs in the Oil-Prone Bohai Bay Basin, East China. *Earth Science*, 50(2):377-387. doi: 10.3799/dqkx.2024.013. (in Chinese with English abstract)
- [27] Yi,S.H.,2019. The Study on CO₂+ crude oil+water phase equilibrium(Dissertation).China University of Petroleum, Beijing(in Chinese with English abstract).
- [28] Yu,H.Y.,Yang,Z.L.,Liu,J.H.,et al.,2019. EO R Method by the Carbonated Water Injection (CWI) for Tight Oil Reservoirs. *Petroleum Geology and Oilfield Development in Daqing*, 38(2):166-174. doi: 10.19597/J.ISSN.1000-3754.201807019. (in Chinese with English abstract)
- [29] Zheng,Z.X.,Di,Y.,Yu,E.Y.,2023. DL-KF Modeling for Acceleration of Flash Calculations in Phase Equilibrium Using Deep Learning Methods. *Frontiers in Earth Science*, 10:1041589.doi: 10.3389/feart.2022.1041589.
- [30] Zhu,G.Y.,Li,J.F.,Zhang,Z.Y.,2025. Origin of Deep Oil and Gas Phase State Diversity and Evaluation of Secondary Geochemical Intensity:A Case Study of Marine Oil and Gas in Tarim Basin. *Earth Science*, 50(6):2163-2178. doi: 10.3799/dqkx.2021.177. (in Chinese with English abstract)

中文参考文献

- [1] 白龙辉, 柳波, 迟亚恩, 等, 2021.二维核磁共振技术表征页岩所含流体特征的应用——以松辽盆地青山口组富有机质页岩为例.石油与天然气地质, 42(06):1389-1400.
- [4] 邸元, 张园, Wu Yu-Shu, 2015. Gibbs 自由能最小化法计算二氧化碳-烃-水系统相平衡.石油学报, 36(05): 593-599.
- [6] 郭克群, 2012.相平衡数据库软件开发(硕士学位论文).青岛: 青岛科技大学.
- [7] 韩芳, 2004.二氧化碳与饱和烷烃二元系统高压气液相平衡的研究(硕士学位论文).天津: 天津大学.
- [14] 柳波, 吕延防, 冉清昌, 等, 2014.松辽盆地北部青山口组页岩油形成地质条件及勘探潜力.石油与天然气地质, 35(02): 280-285.
- [18] 潘永帅, 柳波, 赵振, 等, 2025.富凝灰质混积型页岩成岩演化特征及成储机制——以三塘湖盆地二叠系芦草沟组为例.石油与天然气地质, 1-20.

- [20] 邱锡鹏, 2020.神经网络与深度学习.北京: 机械工业出版社.
- [23] 施文, 桓冠仁, 郭尚平, 1992.二氧化碳-烃-水系统相平衡闪蒸计算方法研究.石油勘探与开发, 19(03): 48-56.
- [26] 薛永安, 胡安文, 万琳, 等, 2025.渤海湾油型盆地深层大型气藏成藏过程与相态控制因素.地球科学, 50(2): 377-387.
- [27] 冀胜合, 2019. CO₂+原油+水相平衡研究(硕士学位论文).北京: 中国石油大学.
- [28] 于海洋, 杨中林, 刘俊辉, 等, 2019.致密油藏碳化水驱提高采收率方法[J].大庆石油地质与开发, 38(02): 166-174.
- [30] 朱光有, 李婧菲, 张志遥, 2025.深层油气相态多样性成因与次生地球化学作用强度评价: 以塔里木盆地海相油气为例.地球科学, 50(6): 2163-2178.