

基于多分支回归生成对抗网络的稀土萃取分离流程模拟方法

朱建勇¹ 王佳浩¹ 韩泽润¹ 廖良芳² 杨辉¹

摘 要 针对基于萃取机理的稀土萃取工艺流程模拟难以吻合实际萃取生产工况的问题, 提出一种基于多分支回归生成对抗网络的稀土萃取分离流程模拟方法, 实现稀土萃取分离各级萃取槽稀土元素组分含量的精准计算. 首先, 针对稀土萃取生产现场有效样本数量较少的问题, 采用生成对抗网络 (GAN) 构造生成对抗模型, 依据稀土萃取工艺串级分离特点, 使用多分支深层网络构造 GAN 的生成器, 逐级学习萃取级间数据的深层特征; 提出判别器与回归器的浅层特征共享机制, 回归器复用判别器首层卷积特征以提升预测性能, 并通过回归一致性约束以生成更真实的样本; 根据多分支网络结构特点, 设计一种递归渐进式对抗训练策略, 固定前一支子 GAN 模型学习到的网络参数并作为下一分支子 GAN 的共同特征, 各分支子 GAN 内部生成器、判别器、回归器三者循环对抗训练, 在保证模型稳定收敛的同时, 精准捕捉级间耦合特征. 仿真结果表明了本文所提方法的有效性.

关键词 生成对抗网络; 稀土元素; 多分支网络; 组分含量

引用格式 朱建勇, 王佳浩, 韩泽润, 廖良芳, 杨辉. 基于多分支回归生成对抗网络的稀土萃取分离流程模拟方法. 自动化学报, 2026, 52(8): 1791–1806

DOI 10.16383/j.aas.c250450 **CSTR** 32138.14.j.aas.c250450

Simulation Method for Rare-earth Extraction and Separation Processes Based on Multi-branch Regression Generative Adversarial Networks

ZHU Jian-Yong¹ WANG Jia-Hao¹ HAN Ze-Run¹ LIAO Liang-Fang² YANG Hui¹

Abstract To address the problem that the flow simulation of rare-earth extraction process based on extraction mechanism hardly conforms to actual extraction production conditions, a simulation method for rare-earth extraction and separation processes based on multi-branch regression generative adversarial networks is proposed to achieve accurate calculation of rare-earth element component concentrations in each stage of rare-earth extraction and separation. First, in view of the limited number of effective samples in the rare-earth extraction production site, a generative adversarial networks (GAN) is adopted to construct a generative adversarial model. According to the cascade separation characteristics of the rare-earth extraction process, a multi-branch deep network is used to construct the generator of the GAN, learning deep features of inter-stage extraction data level by level; A shallow feature sharing mechanism between discriminator and regressor is proposed, the regressor reuses the first-layer convolutional features of the discriminator to improve prediction performance, and a regression consistency constraint is introduced to generate more realistic samples; In view of the structural characteristics of multi-branch networks, a recursive progressive adversarial training strategy is designed. The network parameters learned by the previous branch sub-GAN model are fixed and serve as common features for the next branch sub-GAN. Within each branch sub-GAN, the generator, discriminator, and regressor undergo cyclic adversarial training, ensuring stable model convergence while accurately capturing inter-stage coupling features. Simulation results demonstrate the effectiveness of the proposed method.

Keywords generative adversarial networks; rare-earth elements; multi-branch networks; component concentration

Citation Zhu Jian-Yong, Wang Jia-Hao, Han Ze-Run, Liao Liang-Fang, Yang Hui. Simulation method for rare-earth extraction and separation processes based on multi-branch regression generative adversarial networks. *Acta Automatica Sinica*, 2026, 52(8): 1791–1806

收稿日期 2025-09-04 录用日期 2026-03-26

Manuscript received September 4, 2025; accepted March 26, 2026

国家重大专项 (2026ZD16010302), 国家自然科学基金 (62363010), 江西省自然科学基金 (20252BAC250019), 江西省双千计划 (SSQ2023018) 资助

Supported by National Major Special Project (2026ZD16010302), National Natural Science Foundation of China (62363010), Natural Science Foundation of Jiangxi Province (20252BAC250019), and Jiangxi Province Double Thousand Plan (SSQ2023018)

本文责任编辑 刘毅

稀土包含元素周期表中镧系元素、钪 (Sc) 及钇 (Y) 共计 17 种元素, 享有“工业维生素”的美誉^[1]. 自 20 世纪 70 年代徐光宪教授提出“串级萃取理论”

Recommended by Associate Editor LIU Yi

1. 华东交通大学电气与自动化工程学院 南昌 310013 2. 中稀江西稀土有限公司 赣州 341000

1. School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 310013 2. China Rare Earth Jiangxi Co., Ltd., Ganzhou 341000

以来,我国稀土萃取工艺逐渐处于国际领先水平^[2].其团队后续提出的串级萃取静态计算和动态模拟方法进一步为稀土萃取工艺提供优化指导^[3-4].为提高工业智能化水平,根据生产工艺参数对稀土萃取过程进行流程模拟,预测可能的萃取分离工况,为后续的流程重组和工艺调整提供决策依据,对于避免人工经验的主观性和提升生产效率具有重要意义^[5].

稀土萃取过程存在机理复杂、级数多且级间强耦合等特点^[6].文献[7]指出,在连续逆流过程中,进料端微小的浓度扰动会破坏全流程的酸度平衡,在级间产生动态滞后与累积偏差,导致稳态模型失效.文献[8]针对多组分体系的模拟发现,受限于萃取剂的负载容量,前级易萃组分的分配系数变化会通过流体 pH 值的波动,呈级联式放大并抑制后级难萃组分的传质效率.文献[9-10]将简化后的热力学平衡模型应用于多组分稀土萃取工艺的流程模拟,但忽视了诸多因素影响,依此建立的机理模型难以准确描述实际生产过程并预测萃取工艺稀土元素组分含量.文献[11]设计稀土萃取多级串联的混合澄清槽仿真模拟系统,通过对槽体体积与流量平衡的模拟,展现两相体积和流量变化的非线性和滞后性,但其堰流系数等核心参数依赖经验确定,其结果与实际存在一定偏差.

目前数据驱动方法开始被应用于稀土萃取分离流程建模^[12].文献[13]利用人工神经网络对稀土萃取传递过程中稀土元素含量变化进行预测,相比传统机理模型结果更加准确.但是对于工况多变的复杂工业场景,工业原始数据存在样本稀缺的问题,导致神经网络因样本不足难以处理实际复杂的非线性关系、学习能力不足、泛化能力较弱^[14].此外,人工神经网络的预测结果在网络最终输出层进行单一维度或多维度输出,无法准确刻画稀土萃取工艺组分含量逐级多层次化输出的特性^[15].

针对多层次化输出的建模需求,文献[16]设计一种多分支神经网络结构,通过在传统卷积神经网络架构中增加额外的侧支分类器,使得网络能够从不同分支上输出具有不同置信度的预测结果,适用于多分支逐级输出的场景.此外,生成对抗网络(generative adversarial networks, GAN)作为一种强大的生成模型^[17].当生成样本缺乏多样性时,生成器(generator, G)可用梯度信息不足以覆盖数据样本的分布,易导致训练过程中出现模式崩溃现象^[18].为此,文献[19]提出条件生成对抗网络(conditional generative adversarial networks, CGAN),将辅助信息作为额外的输入,馈送到 CGAN 模型的生成器和判别器(discriminator, D),根据特定条件生

成具有不同特征的样本. CGAN 已广泛应用于传感器数据生成^[20]、图像生成^[21]、临床诊断^[22]等领域.文献[23]使用 CGAN 作为模型框架,训练生成器用于回归任务.文献[24]提出一种多判别器的 GAN 模型,各子 GAN 模型间共享单个生成器,对不同的数据模式进行训练.文献[25]进一步提出回归生成对抗网络(regression GAN, RGAN),通过判别器和回归器(regressor, R)共享浅层捕获一般低级特征,为模型用于回归任务提供正则化优势.目前基于 GAN 的多输出回归研究文献较少.

针对稀土萃取工艺的串级特性及标签样本数量有限的问题,本文提出一种基于多分支回归生成对抗网络(multi-branch regression generative adversarial networks, MB-RGAN)的稀土萃取分离流程模拟方法,该网络包含一个公用的生成器、多个独立的判别器和回归器. MB-RGAN 逐级引出包含当前层信息的分支层,每一分支判别器和回归器共享单个生成器分支组成一个子 GAN 模型.为解决传统 GAN 训练不稳定的问题,引入带梯度惩罚的 Wasserstein 距离优化目标函数;同时,设计一种判别器和回归器共享首层卷积特征的机制,利用判别器提取的潜在特征辅助回归预测,并引入回归一致性约束以指导生成器生成更符合生产实际的样本.在训练阶段,设计一种递归渐进式对抗训练(recursive progressive adversarial training, RPAT)策略,先采取冻结前序网络权重参数的方式逐级预训练分支子模型,再通过全局微调阶段解冻所有网络参数进行联合训练,从而在保证收敛的同时精准捕捉复杂的级间耦合关系.

1 工艺描述及生成对抗网络

1.1 稀土萃取工艺

稀土萃取工艺采用多级混合澄清槽级联操作,根据不同溶解度差异将稀土元素分为易萃和难萃组分.图1描述了一段 $n+m$ 级(n 级萃取段和 m 级洗涤段组成)的稀土串级萃取分离工艺流程,通过电机搅拌将每一萃取级对应萃取槽的料液分为易萃组分 A (有机相) 和难萃组分 B (水相),假设为 $Y_j = [y_{j1}, y_{j2}]$, $j \in [1, n+m]$, y_{j1} 和 y_{j2} 分别表示第 j 级萃取槽中有机相和水相稀土组分含量值.

由于稀土萃取工艺由数十个甚至上百个萃取槽串联而成,采用神经网络建立组分含量计算模型时,浅层神经网络适合模拟前几级萃取槽组分含量,而深层神经网络模拟前几级萃取槽组分含量易出现过拟合现象,适合计算后级萃取槽组分含量.因此,若

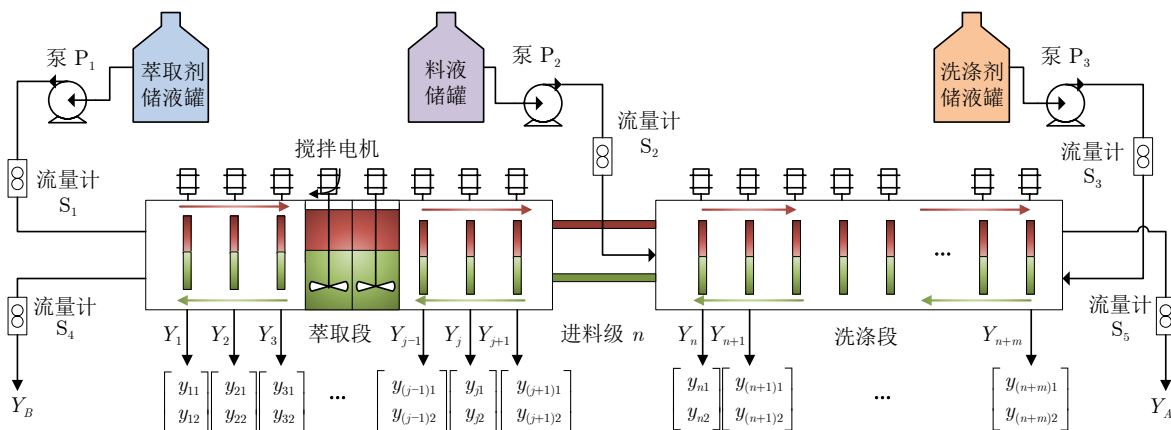


图1 稀土串级萃取分离工艺流程图

Fig.1 Flowchart of rare-earth cascade extraction and separation process

需要完全模拟稀土萃取工艺每级萃取槽的组分含量, 显然需要建立几十个神经网络模型, 然而这几十个神经网络模型难以体现萃取槽与萃取槽之间的级联耦合关系. 不同于传统神经网络在同一层进行多输出的方式, 本文将多分支网络引入 CGAN 模型, 设计一种多分支回归生成对抗网络 MB-RGAN, 串联在公共生成器的分支子 GAN 模型, 对不同级数萃取槽进行联合建模, 通过对抗训练能够捕捉级间耦合关系.

1.2 CGAN 的基本理论

生成对抗网络提供了一种新的生成模型训练优化架构. 整个网络由生成器 G 和判别器 D 两部分组成, G 的目标是捕捉真实样本的概率分布, 学习后生成能够尽可能“欺骗”D 的数据样本. D 的目标是尽可能准确地将真实样本与生成样本区分开来. CGAN 是在 GAN 基础上, 将条件变量 c 同时加入 G 与 D 中, 使得样本数据的生成取决于包含不同特征的条件变量. 相较于传统回归模型, CGAN 具有根据特征条件将随机噪声 z 映射到潜在空间中的特定分布的能力, 通过对抗训练使模型能更充分学习和表达输入-输出间的复杂映射关系. CGAN 结构示意图如图 2 所示.

CGAN 的网络模型优化可以归结为一个二元最小-最大对抗过程, 目标函数 $V(D, G)$ 可表示为:

$$\min_G \max_D V(D, G) = E_{y_r \sim P_r} [\log D(y_r, c)] + E_{z \sim P_z} [\log(1 - D(G(z, c), c))] \quad (1)$$

式中, y_r 为服从真实分布 P_r 的稀土元素组分含量数据; c 为辅助生成过程的工艺条件变量; z 为服从先验分布 P_z 的随机噪声向量; $G(z, c)$ 表示生成器生成的稀土元素组分含量样本; $D(\cdot)$ 表示判别器输

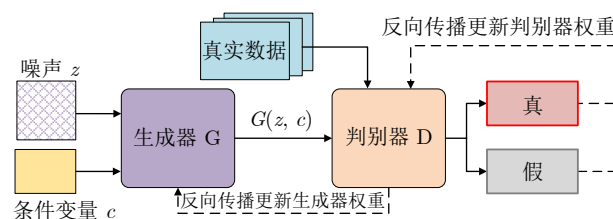


图2 CGAN 原理图

Fig.2 CGAN principle diagram

出样本为真的概率.

2 多分支回归生成对抗网络结构设计

2.1 MB-RGAN 模型总体架构

本文所提的多分支回归生成对抗网络结构设计如图 3 所示, 包括一个公共多分支生成器 G、多个判别器 D 和多个回归器 R. G 的主网络使用多隐层逐级学习原始特征, 并在主网络下引出 $n+m$ 级独立分支, 各分支首先引入全连接层逐级映射到潜在空间生成样本, 使 G 的当前分支提取前一支对应隐含层学习到的特征, 该过程准确刻画了稀土萃取生产中稀土元素组分含量逐级多层次化输出的特性. G 的每一个分支将噪声 z 和条件变量 c 通过潜在空间映射到一个单独的样本 y , 由于生成器可用的梯度信息不足以覆盖所有数据样本的分布, 多输出情况下可能会导致生成样本仅围绕唯一样本的模式崩塌. 因此, 在每个分支下继续引入一个独立判别器, 每个 D 通过各自损失函数约束, 强制生成器 G 学习数据集的不同模式, 防止 G 总是接受 D 的一致固定反馈.

在 MB-RGAN 模型中, 判别器 D 通过对抗训练能够学习到有效区分真实样本和生成样本的特征

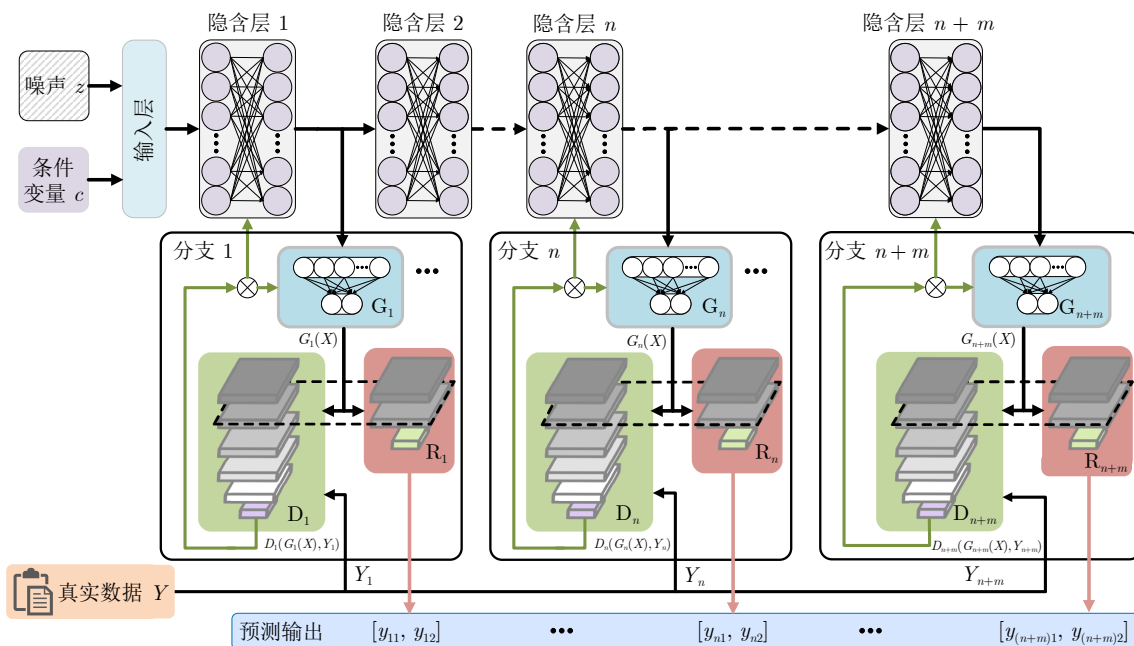


图3 多分支回归生成对抗网络结构图

Fig.3 MB-RGAN structure diagram

表示. 考虑到 D 的首层潜在特征层是对各分支真实样本的浅层表达, 有利于回归预测, 因此在每一分支 D 后并联一个回归器 R, D 和 R 共享首层一维卷积神经网络 (1D convolutional neural network, 1D-CNN), 各分支低级特征能够被浅层网络捕获并共享, 提高回归器性能的同时减少模型冗余度. R 的目的是回归预测, 为了关注各分支下的数据生成和变量之间的回归关系, 将目标函数设计为回归损失的同时, 引入回归一致性损失作为约束项施加于判别器 D 和生成器 G, 这种反馈作用约束 G 生成更真实的样本; 同时, 基于 D 和 R 的首层特征共享机制, 促使浅层网络捕捉到兼具判别性与回归价值的特征, 进一步通过 G 和 D 的对抗学习, 促进 R 的回归能力, 由此在分支内部形成一个正反馈循环, 提升模型的整体性能.

2.2 生成器 (G) 网络结构

MB-RGAN 的生成器网络为多分支深层网络, 包括主网络和分支层, 其详细结构如图 4 所示.

由于企业在工艺流程中仅设有 5 个关键组分含量监测级 (即第 15、25、30、35、45 级), 为使模型输出与实际观测数据的维度保持一致, 本文将生成器网络确定为 5 个隐含层及 5 个分支结构, 使每个分支对应一个具体的萃取监测级. 每个隐含层由包含 64 个神经元的全连接层构成, 全连接层后采用 Tanh 激活函数以标准化特征分布, 每个隐含层后分出一个包含 32 个神经元的独立分支, 每个分支末端增

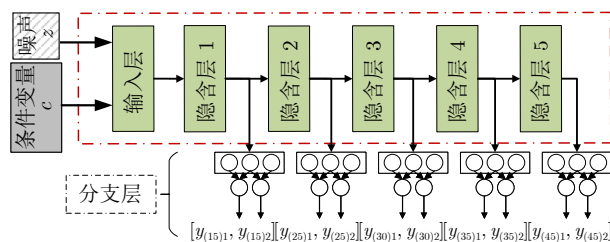


图4 生成器 (G) 网络结构图

Fig.4 Network structure diagram of generator (G)

加一个回归层用于得到相应级数下的组分含量值 y_{j1} 和 y_{j2} (分别对应有机相与水相组分), 为了严格满足组分含量非负的物理约束, 输出层激活函数选用 Sigmoid.

2.3 判别器 (D) 与回归器 (R) 网络结构

考虑到卷积神经网络具有出色的特征提取能力, 因此选择一维卷积神经网络 (1D-CNN) 构建判别器 D 和回归器 R, 其详细结构如图 5 所示.

针对输入数据的一维序列, 模型首先通过输入层 (Input) 接收批次大小为 Bs 的数据, 随后通过一个共享卷积层 (Conv1d) 提取基础浅层特征. 该共享机制不仅能够有效减少模型参数冗余, 更能使回归器 R 利用判别器 D 在对抗训练中学到的强鲁棒性特征, 从而提升模型的泛化性能. 对于判别器 D, 共享特征随后依次经过两个独立的一维卷积层 (分别包含 64 个和 128 个卷积核) 逐步提取高阶抽

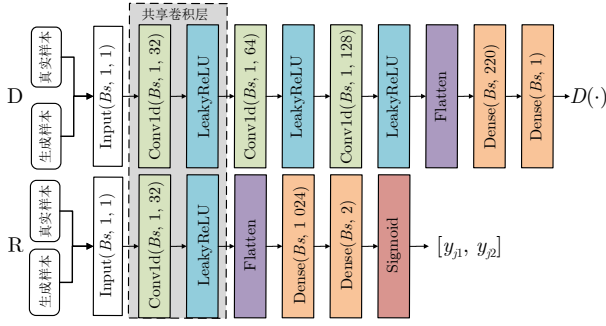


图 5 判别器 (D) 和回归器 (R) 网络结构图

Fig. 5 Network structure diagram of discriminator (D) and regressor (R)

象特征, 各卷积层后均使用 LeakyReLU 激活函数, 然后经过 Flatten 层进行展平操作, 最终在最后一层全连接层 (Dense) 输出一个无界实数 $D(\cdot)$ 作为 Wasserstein 距离的评估. 回归器 R 复用判别器 D 的首层共享卷积层, 经展平操作后通过全连接层进行回归拟合, 末端输出层采用 Sigmoid 激活函数, 以确保预测的有机相与水相组分含量严格满足非负的物理约束.

3 多分支回归生成对抗网络训练

3.1 多分支回归生成对抗网络的目标函数

为了解决传统 GAN 训练中存在的梯度消失和模式崩溃问题^[26], 引入带梯度惩罚的 Wasserstein 距离^[27]作为度量指标来衡量生成数据分布与真实数据分布的差异. 由于直接计算 Wasserstein 距离较为困难, 使用 Kantorovich-Rubinstein 对偶性, 将计算 Wasserstein 距离转化为一个最优化问题, 对于任意两个分布 P_r (真实分布) 和 P_g (生成分布), 二者间的 Wasserstein 距离定义如下:

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x, y) \sim \gamma} [\|x - y\|] = \sup_{\|f\|_L \leq 1} E_{y_r \sim P_r} [f(y_r)] - E_{y_g \sim P_g} [f(y_g)] \quad (2)$$

式中, $\Pi(P_r, P_g)$ 是所有以 P_r 和 P_g 为边缘分布的联合分布 γ 的集合; $\inf$ 和 $\sup$ 分别表示取最大下界与最小上界; $\|f\|_L \leq 1$ 表示函数 f 满足 1-Lipschitz 连续性约束. 直接寻找满足 1-Lipschitz 约束的函数 f 较为困难, 在 MB-RGAN 中, 使用判别器 D 来近似该函数 f . 为了使 D 满足 1-Lipschitz 连续性条件并提高模型训练稳定性, 在损失函数中引入梯度惩罚项 (gradient penalty, GP), 定义如下:

$$L_{GP} = E_{\tilde{y} \sim P_{\tilde{y}}} [(\|\nabla_{\tilde{y}} D(\tilde{y})\|_2 - 1)^2] \quad (3)$$

式中, $P_{\tilde{y}}$ 为插值样本 $\tilde{y}$ 的分布; $\tilde{y}$ 表示在真实样本 $y_r \sim P_r$ 和生成样本 $y_g = G(z, c)$ 之间进行线性插值采样得到的样本点, 计算方式为 $\tilde{y} = \epsilon y_r + (1 - \epsilon) y_g$, 其中 $\epsilon \sim U(0, 1)$ 服从标准均匀分布; $\|\nabla_{\tilde{y}} D(\tilde{y})\|_2$ 表示判别器在 $\tilde{y}$ 处的梯度二范数. 结合 Wasserstein 距离与多分支结构特性, MB-RGAN 第 j 分支的生成器 G_j 与判别器 D_j 的目标函数定义为:

$$\min_G \max_D V(D_j, G_j) = E_{y_r \sim P_r} [D_j(y_r)] - E_{z \sim P_z} [D_j(G_j(z, c))] - \lambda_{GP} L_{GP} \quad (4)$$

式中, λ_{GP} 为梯度惩罚系数; $G_j(z, c)$ 为生成器第 j 分支输出组分含量样本.

考虑到 G 的数据生成和 D 的判别反馈应受到稀土萃取工艺回归关系的约束, 本文在生成对抗网络中引入回归器 R, 基于 D 通过对抗学习能提取深层潜在特征的能力, 且回归器与判别器在浅层特征提取阶段具有任务相似性, 设计一种浅层特征共享机制. 如图 5 所示, R 与 D 共享首层卷积特征, 回归器 R 利用共享特征准确预测输入样本的组分含量值, 其损失函数定义如下:

$$L_R = \frac{1}{N} \sum_{i=1}^N [(\hat{y}_{g,i} - y_{g,i})^2 + (\hat{y}_{r,i} - y_{r,i})^2] \quad (5)$$

式中, N 为样本数量; $y_{r,i}$ 表示真实样本的观测值, $\hat{y}_{r,i}$ 表示回归器对真实样本特征的预测值; $y_{g,i}$ 表示生成样本的观测值; $\hat{y}_{g,i}$ 表示回归器对生成样本特征的预测值.

完整的 MB-RGAN 内部结构如图 6 所示, 为了约束生成器 G 生成符合物理规律的样本, 模型在传统判别器 D 到生成器 G 反馈通路的基础上, 还构建了从回归器 R 到生成器 G 的反馈通路.

针对 MB-RGAN 的分支结构不同于传统 GAN, 对 $n + m$ 级分支分别建立独立的优化目标函数, 以分支 j 为例, 对应的 D_j 和 G_j 的目标函数表达式如式 (6) 和 (7) 所示.

$$L_{D_j} = \frac{1}{N} \sum_{i=1}^N \left\{ D_j(G_j(z, c)) - D_j(y_{r,i}) + \lambda_{GP} (\|\nabla_{\tilde{y}} D_j(\tilde{y})\|_2 - 1)^2 + \lambda_1 ((\hat{y}_{g,i} - y_{g,i})^2 + (\hat{y}_{r,i} - y_{r,i})^2) \right\} \quad (6)$$

$$L_{G_j} = \frac{1}{N} \sum_{i=1}^N \left\{ -D_j(G_j(z, c)) + \lambda_2 ((\hat{y}_{g,i} - y_{g,i})^2 + (\hat{y}_{r,i} - y_{r,i})^2) \right\} \quad (7)$$

最终, 各网络参数依据计算出的梯度进行更新:

$$\theta \leftarrow \theta - \eta \text{Adam}(\nabla_{\theta} L) \quad (16)$$

式中, η 为学习率; $\text{Adam}(\cdot)$ 表示 Adam 优化器根据一阶矩和二阶矩估计对梯度进行的修正。

3.3 多分支回归生成对抗网络训练策略

针对 MB-RGAN 网络的多分支深层结构特性, 为了避免深层网络训练中常见的梯度消失及收敛困难问题, 本文提出一种递归渐进式对抗训练 (RP-AT) 策略, 其训练原理如图 7 所示, 该策略将复杂的整体网络优化分解为逐级递进的子任务。

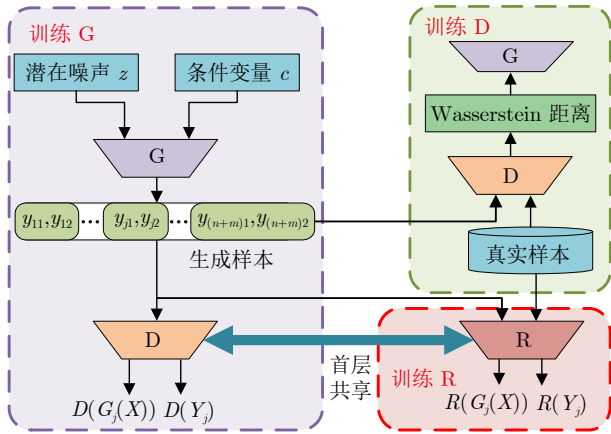


图 7 MB-RGAN 训练原理图

Fig.7 MB-RGAN training principle diagram

训练过程从第 1 级隐含层及对应的分支子 GAN 开始, 逐级向后推进。当训练进行到第 j 分支时, 前序 1 至 $j-1$ 级的主干网络 (生成器隐含层) 及对应分支的判别器 D 与回归器 R 参数已被视为训练完毕, 此时对参数空间 $\theta^{(1:j-1)}$ 执行冻结 (freeze) 操作, 即在当前分支的反向传播过程中, 不计算也不更新这些参数的梯度。前序分支生成器隐含层学习到的高阶特征 h_{j-1} 被传递给当前第 j 层分支网络作为输入, 这种机制不仅显著降低了单次训练的计算复杂度, 还通过保留级间耦合特征的稳定性, 使得模型能够更专注于当前萃取级数据的分布拟合。

对于每个子 GAN 模型 (由第 j 级生成器分支、判别器 D_j 和回归器 R_j 组成) 都在前一轮生成特征的基础上进行优化。此外, 为了解决训练中生成器与判别器收敛速度不匹配的问题, 在每一轮迭代中, 固定生成器参数, 循环训练判别器 D_j 和回归器 R_j 共 n_c 次 (本文设为 5 次), 随后再更新一次生成器参数。

在模型所有分支均完成逐级迭代预训练之后, 对网络中全部可训练参数执行解冻 (unfreeze) 操作。随后采用全量数据集对包含所有分支的完整网

络进行联合训练。在此阶段, 梯度能够在网络所有隐含层间传播, 网络参数能够根据后续所有分支的反馈梯度信息进行微调, 从而进一步增强模型对稀土萃取级间强耦合关系的全局捕捉能力。

MB-RGAN 模型的训练细节如算法 1 所示。分别将 $\theta_G^{(j)}$, $\theta_D^{(j)}$ 和 $\theta_R^{(j)}$ 定义为第 j 分支中生成器 G、判别器 D 和回归器 R 的参数空间, 并采用梯度下降法迭代训练该网络。该训练策略采用 Adam 优化器更新网络参数, 其中生成器 G 和判别器 D 的初始学习率都为 0.000 1, 学习率衰减系数设为 0.1, 批次大小为 64, 梯度惩罚超参数 λ_{GP} 设为 10。

算法 1. MB-RGAN 递归渐进式训练算法

输入. 条件变量 c , 噪声 z , 标签数据 y_r , 分支数 L , 批次大小 Bs , 梯度惩罚系数 λ_{GP} , 回归权重 λ_1, λ_2 , 判别器和回归器循环次数 n_c , 预训练迭代次数 S_{pre} , 微调迭代次数 S_{fine} , 学习率 η 。

输出. 优化后的网络参数 $\theta_G, \theta_D, \theta_R$ 。

- 1) 初始化所有网络参数;
- 2) 阶段一: 递归渐进式预训练
- 3) **for** $j = 1$ **to** L **do** $\triangleright$ 逐级遍历每一分支
- 4) **if** $j > 1$ **then**
- 5) 冻结 $\theta_G^{(1:j-1)}, \theta_D^{(1:j-1)}, \theta_R^{(1:j-1)}$ $\triangleright$ 冻结前序分支及主干参数;
- 6) **end if**
- 7) **for** $s = 1$ **to** S_{pre} **do**
- 8) **for** $t = 1$ **to** n_c $\triangleright$ 训练判别器与回归器
- 9) 采样真实数据 y_r 和噪声 z , 生成插值样本 $\tilde{y}$;
- 10) 计算回归损失 L_R :

$$L_R = \frac{1}{N} \sum_{i=1}^N ((\hat{y}_{g,i} - y_{g,i})^2 + (\hat{y}_{r,i} - y_{r,i})^2)$$
- 11) 更新 $\theta_D^{(j)}, \theta_R^{(j)}$ 以最小化 L_D :

$$\nabla_{\theta_D, \theta_R} \frac{1}{N} \sum_{i=1}^N \{D_j(G_j(X)) - D_j(y_{r,i}) + \lambda_{GP} (\|\nabla_{\tilde{y}} D_j(\tilde{y})\|_2 - 1)^2 + \lambda_1 L_R\}$$
- 12) **end for**
- 13) 采样噪声 z 和条件 c ; $\triangleright$ 训练生成器
- 14) 更新 $\theta_G^{(j)}$ 以最小化 L_G :

$$\nabla_{\theta_G} \frac{1}{N} \sum_{i=1}^N [-D_j(G_j(X)) + \lambda_2 L_R]$$
- 15) **end for**
- 16) **end for**
- 17) 阶段二: 全局微调
- 18) 解冻所有参数 θ ; $\triangleright$ 解冻整个网络
- 19) **for** $s = 1$ **to** S_{fine} **do**
- 20) 从全量数据集中采样批次样本;
- 21) 基于 G、D、R 目标函数计算全局损失;
- 22) 利用全局梯度反向传播同步微调所有参数。
- 23) **end for**

4 仿真结果及分析

4.1 数据集与实验设置

实验数据采集自江西某稀土企业实际生产线, 现场技术人员主要依据少数关键监测级的组分变化趋势来进行工艺调整. 考虑到取样化验成本, 仅选取了对工艺波动较为敏感的 5 个关键监测级 (第 15、25、30、35、45 级) 作为数据采集点. 操作人员从上述 5 个萃取槽手工提取水相/有机相溶液并送实验室进行离线化学分析, 获得稀土元素组分含量, 并从分布式控制系统 (distributed control system, DCS) 历史数据库中提取对应的工艺输入变量. 原始数据经过清洗 (剔除含缺失值、负值和明显离群点的样本) 后将其划分为训练集 (80%)、验证集 (10%) 和测试集 (10%). 数据集中包含了 14 个工艺输入变量以及相应萃取槽的稀土元素组分含量, 各输入变量物理意义描述如表 1 所示.

表 1 工艺输入变量
Table 1 Process input variables

输入变量	物理意义描述
c_1	La 元素组分含量
c_2	Ce 元素组分含量
c_3	Pr 元素组分含量
c_4	Nd 元素组分含量
c_5	Ce/La 元素间分离系数
c_6	Pr/Ce 元素间分离系数
c_7	Nd/Pr 元素间分离系数
c_8	主产品出口模式
c_9	进料模式
c_{10}	萃取剂流量
c_{11}	洗涤剂流量
c_{12}	进料级数
c_{13}	有机相出口分数
c_{14}	水相出口分数

变量 $c_1 \sim c_4$ 表示原料液中 La、Ce、Pr、Nd 四种待分离元素的摩尔浓度; 变量 $c_5 \sim c_7$ 为不同元素间的分离系数, 由现场技术人员在萃取槽两相取样后化验计算得到; c_8 为主产品出口模式, 数值 0 表示易萃组分为主产品, 数值 0.5 表示难萃组分为主产品, 数值 1.0 表示两者均为主产品; c_9 为进料模式, 数值 0 表示采用水相进料方式, 数值 1 表示有机相进料方式; c_{10} 和 c_{11} 分别表示萃取剂与洗涤剂的入口流量; c_{12} 表示原料液进料级数, 设置不同的进料级会形成不同的萃取段和洗涤段; c_{13} 和 c_{14} 表示萃取流程两端出口生产指标, 由设计人员预先拟

定, 其中 c_{13} 表示有机相产品出口分数 (即输出中易萃组分在料液中的占比), c_{14} 表示水相产品出口分数 (即输出中难萃组分在料液中的占比).

在数据预处理阶段, 为了消除不同量纲对模型训练的影响, 确保所有特征处于相同的尺度, 采用最小-最大归一化方法将输入特征变量映射至 $[0, 1]$, 归一化值 c' 计算公式如下:

$$c' = a + \frac{(c - c_{\min}) \times (b - a)}{c_{\max} - c_{\min}} \quad (17)$$

式中, $c_{\min}$ 和 $c_{\max}$ 分别表示原始特征数据中的最小值和最大值; a 和 b 表示归一化目标区间的上下边界.

图 8 展示了 200 个测试样本归一化后的输入变量和第 25 级易萃组分含量的数据分布趋势, 各个特征的分布差异显著, 没有明显的规律性.

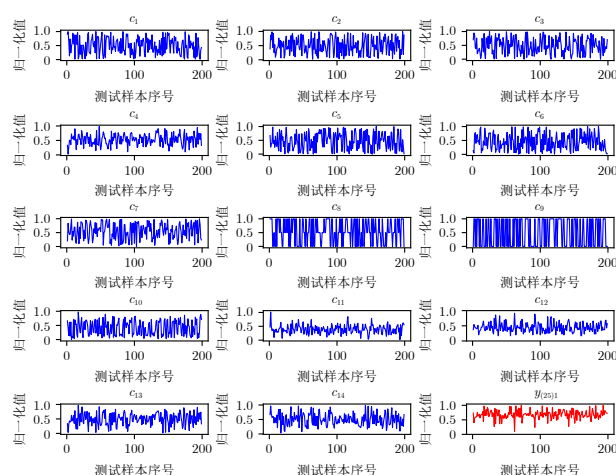


图 8 输入变量 c 与输出数据 $y_{(25)1}$ 分布趋势图

Fig.8 Distribution trend chart of input variable c and output data $y_{(25)1}$

仿真实验基于 Tensorflow 深度学习框架和 Python 3.8 编程环境实现, 硬件平台采用一台配备 NVIDIA GeForce RTX 4050 GPU 的计算机.

为了全面评估所提 MB-RGAN 模型在稀土萃取流程模拟任务上的有效性及优越性, 本文选择了 8 种不同机制的代表性模型作为对比基准, 具体包括: 传统机器学习模型 SVR^[28] 和 XGBoost^[29]、时序深度学习模型 LSTM^[30] 与 GRU^[31]、反向传播神经网络 BPNN^[32]、条件生成对抗网络 CGAN 以及多分支神经网络 MB-DNN 与 MB-RDN^[33].

为确保对比实验的严谨性与公平性, 本文对所有对比基线模型及本文提出的 MB-RGAN 均采用了网格搜索 (grid search) 策略进行超参数寻优. 各模型经调参后的最佳设置如表 2 所示.

经调参优化后, MB-RGAN 的最终参数设置如

表 2 各模型最佳超参数设置
Table 2 Optimal hyperparameter settings for each model

模型	关键参数
SVR	惩罚系数 $C = 1$, 核系数 $\gamma = 0.1$, 不敏感区宽度 $\epsilon = 0.01$
XGBoost	学习率 $lr = 0.05$, 子采样率 $S = 0.8$, 最大深度: 5
BPNN	隐藏层单元数: 64, 学习率 $lr = 0.001$
LSTM	隐层神经元数: 128, 学习率 $lr = 0.01$, Dropout: 0.2
GRU	隐层神经元数: 128, 学习率 $lr = 0.01$, Dropout: 0.2
CGAN	生成器学习率 $lr_G = 0.000\ 2$, 判别器学习率 $lr_D = 0.000\ 2$, L_1 范数权重 $\lambda_{L1} = 10$
MB-DNN	隐层神经元数: 64, 学习率 $lr = 0.001$, Dropout: 0.2
MB-RDN	隐层神经元数: 64, 学习率 $lr = 0.001$, Dropout: 0.2
MB-RGAN	生成器学习率 $\eta_G = 0.000\ 1$, 判别器学习率 $\eta_D = 0.000\ 1$, 回归器学习率 $\eta_R = 0.001$, $\lambda_{GP} = 10$, $\lambda_1 = \lambda_2 = 1$

下: 采用 Adam 优化算法更新网络参数, 生成器 G 和判别器 D 的初始学习率均设为 0.000 1, 学习率衰减系数设为 0.1, 批次大小设为 64. 为保证训练的稳定性, 梯度惩罚超参数 λ_{GP} 设为 10, 判别器与回归器的联合训练循环次数设为 5.

4.2 MB-RGAN 模型实验结果与分析

为了评估 MB-RGAN 模型多分支训练的稳定性, 实验记录了模型训练过程中各分支生成数据分布与真实数据分布之间的 Wasserstein 距离随迭代轮次的变化曲线, 结果如图 9 所示.

从图中可以清晰地观察到, 在模型训练初期 (0 ~ 5 000 轮迭代), 各分支的 Wasserstein 距离均呈现出不同程度的振荡. 在迭代超过 5 000 轮后, 所有分支的 Wasserstein 距离均趋于收敛, 8 000 次迭代后各分支 Wasserstein 距离收敛至一个较低且稳定的数值区间. 这表明, MB-RGAN 训练过程各分支都能达到收敛, 具备在多级级联工况下持续生成高保真样本的鲁棒性.

为了深入评估 MB-RGAN 模型对稀土萃取过程随机性特征的捕获能力, 以及验证模型是否出现了生成对抗网络常见的“模式崩溃”问题, 本文采用

核密度估计 (kernel density estimation, KDE) 方法, 对测试集中 5 个关键监测级的有机相与水相稀土元素组分含量的概率密度分布进行了可视化对比, 结果如图 10 所示. 图中绿色填充区域代表真实工业数据的分布 (Real), 蓝色填充区域代表模型模拟生成的数据分布 (Simulated).

从整体上看, 模型生成的蓝色概率密度曲线与真实的绿色曲线在所有监测级上均表现出较高的重叠度, 这表明 MB-RGAN 成功学习到了数据的高阶统计特性 (如方差、偏度和峰度). 观察不同萃取级的分布演变趋势, 可以发现模型生成的分布变化完全符合稀土串级萃取的传质机理: 随着萃取级数增加, 有机相组分含量的概率密度峰值逐渐从低值区间 (约 0.10) 向高值区间 (约 0.95) 移动, 这反映了稀土离子逐渐从水相转移至有机相并不断富集的过程, 而水相中的易萃组分逐渐被贫化, 说明模型有效地捕捉到了萃取槽之间的耦合效应.

为了进一步定量评估 MB-RGAN 模型在各个萃取级上的点对点回归预测精度, 本文绘制了测试集上各分支 (对应第 15、25、30、35、45 级) 真实值与预测值的散点拟合图, 如图 11 所示. 图中横坐标代表稀土元素组分含量的真实值, 纵坐标代表模型的预测值, 红色虚线表示理想的完全拟合线. 观察散点分布可知, 无论是在有机相还是水相, 绝大多数测试样本点都紧密聚集在红色对角线附近, 表明 MB-RGAN 在各分支都能实现良好拟合. 克服了传统单一输出网络在处理多级级联输出时容易出现顾此失彼现象 (即难以同时兼顾低浓度级与高浓度级的精度), 从而实现了全流程的高精度模拟.

为了验证模型预测是否学习到稀土萃取过程中物理传质规律, 选取部分测试样本绘制其在 5 个关键监测级上的预测值与真实值变化趋势, 如图 12 所示. 从图 12 可以观察到, 随着萃取级数的增加, 有机相中的稀土组分含量呈现显著的单调上升趋势, 而水相中的组分含量则呈现单调下降趋势. 这

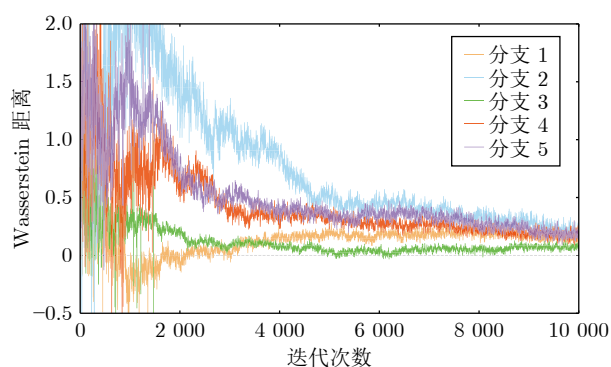


图 9 各分支 Wasserstein 距离曲线

Fig.9 Wasserstein distance curve of each branch

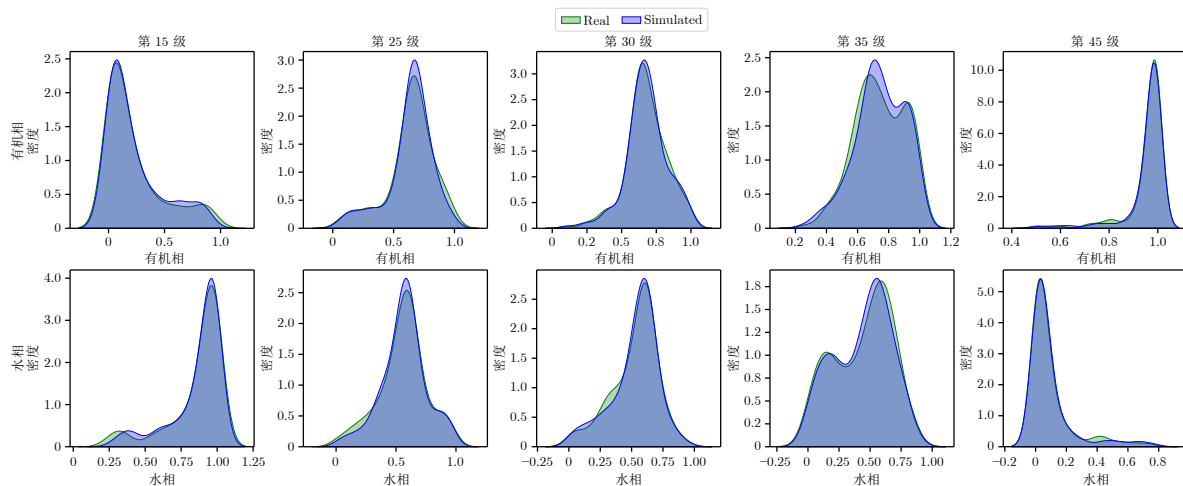


图 10 测试集上各萃取级真实值与模拟值的概率密度分布对比

Fig.10 Comparison of probability density distributions between real values and simulated values for each extraction level on the test set

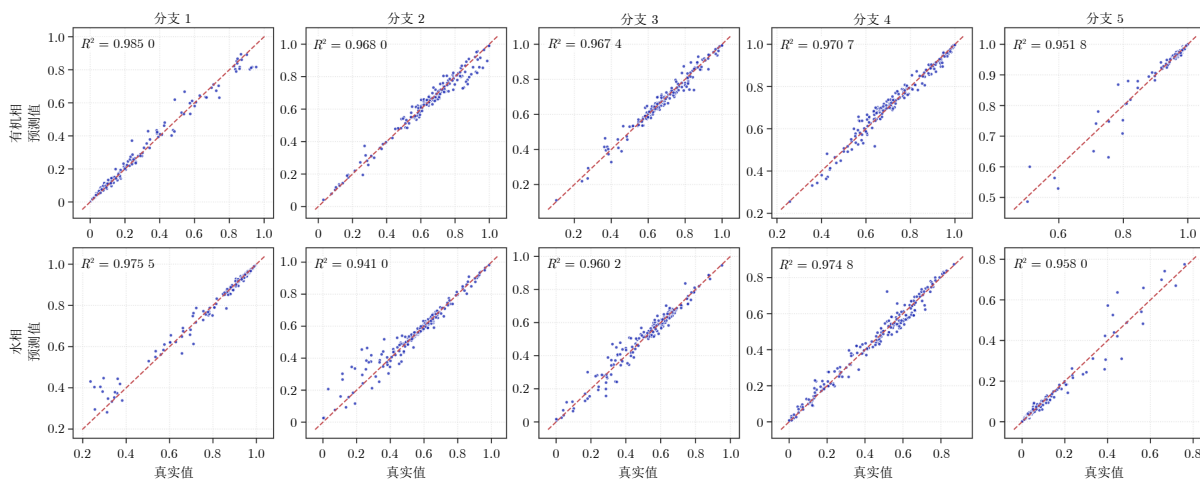


图 11 各分支的真实值与预测值散点拟合图

Fig.11 Scatter fitting plot of real values and predicted values for each branch

与稀土串级萃取工艺中稀土元素易萃组分在有机相富集、难萃组分保留在水相的传质机理完全一致。

4.3 不同模型的性能对比分析

为了全面衡量模型在稀土组分含量预测任务中的精度与拟合能力, 本文采用平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean squared error, RMSE) 和决定系数 (coefficient of determination, R^2) 作为评价指标. 其计算公式为:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (19)$$

式中, y_i 为数据真实值; $\hat{y}_i$ 为模型预测值; $\bar{y}$ 为所有数据样本真实值总和的平均值. 对于 MAE 和 RMSE 两个指标, 其数值越接近于 0 则代表模型的预测精度越高, 而指标 R^2 越接近于 1 则代表模型的预测值和原始数据样本的真实值误差越小.

表 3 详细展示了本文提出的 MB-RGAN 模型与其他 8 种基线对比模型分别在有机相与水相上的性能对比结果. 为了更直观展示各模型在各项指标上的性能差异, 绘制各指标的对比柱状图如图 13 所示.

综合分析表 3 中的评价指标数据以及图 13 的

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (20)$$

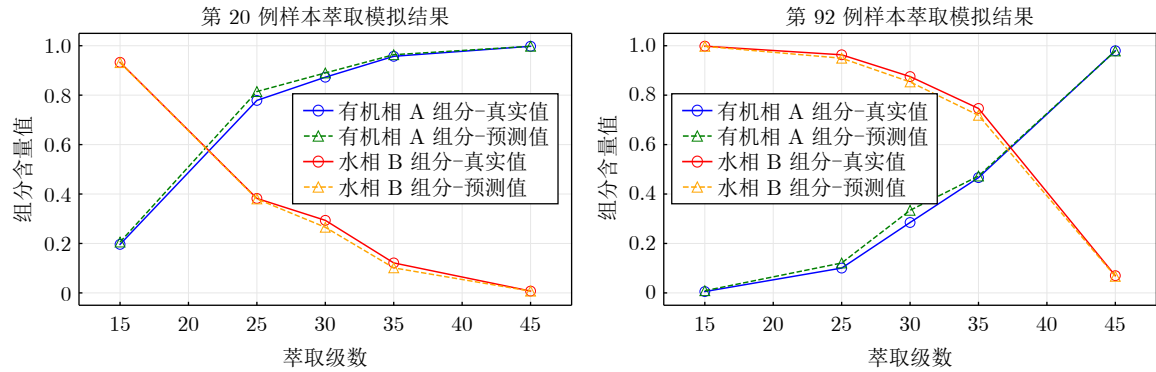


图 12 MB-RGAN 流程模拟结果

Fig.12 Process simulation results of MB-RGAN

表 3 不同模型在测试集上的性能指标对比
Table 3 Comparison of performance metrics of different models on the test set

模型	组分	平均 MAE	平均 RMSE	平均 R^2
SVR	有机相	0.050 0	0.062 9	0.833 1
	水相	0.057 7	0.075 1	0.841 2
XGBoost	有机相	0.036 9	0.051 4	0.879 0
	水相	0.048 4	0.069 4	0.858 9
LSTM	有机相	0.035 9	0.049 8	0.902 9
	水相	0.043 1	0.062 7	0.886 5
GRU	有机相	0.030 0	0.042 3	0.921 1
	水相	0.036 3	0.053 7	0.916 0
BPNN	有机相	0.033 1	0.045 4	0.907 9
	水相	0.041 5	0.058 1	0.902 0
CGAN	有机相	0.028 8	0.042 0	0.926 4
	水相	0.033 9	0.052 6	0.918 0
MB-DNN	有机相	0.032 8	0.045 3	0.915 6
	水相	0.039 3	0.057 5	0.904 8
MB-RDN	有机相	0.030 3	0.042 3	0.921 1
	水相	0.038 3	0.055 3	0.909 7
MB-RGAN	有机相	0.017 6	0.027 7	0.968 5
	水相	0.021 5	0.036 2	0.961 9

可视化结果, 可以看出 MB-RGAN 模型表现最优, 预测效果最佳, 模型在有机相及水相的决定系数 (R^2) 均高于 0.960 0, 有机相的 R^2 更是接近 0.970 0, 显著优于其他对比模型. SVR 和 XGBoost 的表现相对较差, 这是因为稀土萃取过程具有高度的非线性和强耦合特征, SVR 和 XGBoost 机器学习模型虽然计算开销小, 但在处理此类高维复杂映射关系时, 其浅层结构限制了对深层特征的挖掘能力, 无法充分捕捉组分变化的动态规律; BPNN、LSTM 和 GRU 等神经网络模型虽然取得了优于机器学习模型的效果, 但针对稀土萃取分离这种多输出问题, 单一输出层神经网络通过一组参数同时兼顾所有监测级的预测, 网络训练目标是最小化预测值与真实值的整体误差, 在反向传播更新权重参数时, 为了降低部分监测级上的误差, 往往会牺牲其他级的精度, 以实现整体误差最小化; MB-DNN、MB-RDN 以及本文提出的 MB-RGAN 这三种具有多分支结构的方法更适合多输出工艺流程模拟任务, 此类模型各分支具有独立的回归层参数, 能对各级输出进行微调, 降低不同级之间的训练冲突, 并且每个分支都显式输出对应级的预测结果, 模型具备更强的工程可解释性. MB-RDN 模型在多分支结构中引入残差连接, 有效缓解了深层网络中的梯度消失问

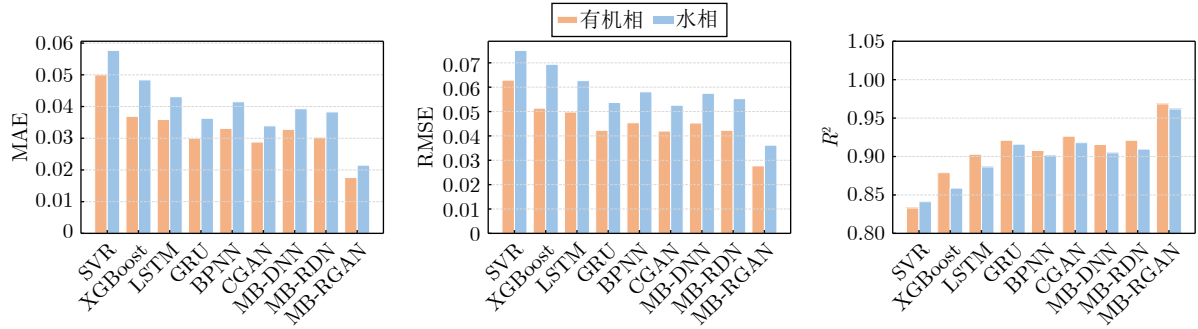


图 13 不同模型性能指标对比

Fig.13 Comparison of performance metrics for different models

题,使网络能够学习到更深层的特征表示,从而提高精度,但由于其本质上仍主要依赖均方误差 (MSE) 作为损失函数,模型倾向于生成平滑的预测值,在一定程度上限制了模型对数据高频波动和细节分布的表达能力,因此其预测精度低于 MB-RGAN. 而生成对抗机制很好地解决了这个问题,MB-RGAN 预测性能优于 CGAN 和 MB-RDN,说明 MB-RGAN 模型能捕捉数据的深层分布规律,预测精度表现最优.

为了进一步直观对比不同模型在关键监测级上的预测精度与稳定性,绘制第 25 级有机相和水相的绝对预测误差曲线与误差分布统计箱线图,如图 14 所示. 通过对比可以明显看出,MB-RGAN 模型 (红色实线) 的绝对误差曲线在大多数样本点上均位于所有对比模型的最下方,且波动幅度最小. 观察箱线图可以看到 MB-RGAN 的误差箱体位置最低且最为扁平,表明其预测误差的中位数最小,且四分位距 (interquartile range, IQR) 最窄. 这意味着 MB-RGAN 不仅整体预测精度最高,而且预测结果的离散度极低,表现出最优的稳定性.

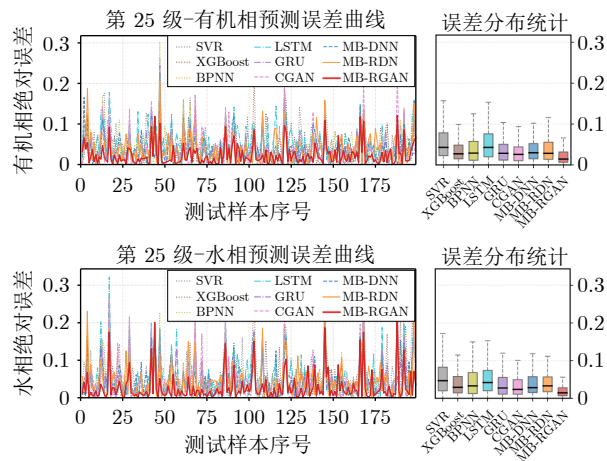


图 14 第 25 级各模型预测绝对误差及误差分布统计对比
Fig.14 Comparison of prediction absolute error and error distribution statistics for each model at the 25th stage

为了评估 MB-RGAN 在实际工业部署中的可行性,将模型参数量作为衡量空间复杂度的指标,将单样本推理耗时作为衡量时间复杂度的指标,与其他模型进行对比,结果如表 4 所示. MB-RGAN (Total) 的参数量为 461 233,单样本推理耗时为 26.10 ms,高于其他对比模型,这是由于模型引入了多分支结构,每个分支有独立的判别器与回归器,而 MB-RGAN (G) 仅包含模型生成器 G 的参数,不包括分支判别器 D 和回归器 R 的参数,其参数量仅为 18 890. MB-RGAN 虽然在参数量和计算耗时上有所增加,但其换来的高精度预测回报远大于计算成本的投入,完全满足稀土萃取工业现场对实时性的严格要求.

表 4 模型计算复杂度与推理效率对比 Table 4 Comparison of model computational complexity and inference efficiency		
模型	参数量	单样本推理耗时 (ms)
SVR	N/A	1.01
XGBoost	N/A	15.08
LSTM	148 682	18.99
GRU	114 634	19.64
BPNN	18 762	5.37
CGAN	57 354	4.99
MB-DNN	104 426	5.80
MB-RDN	172 682	7.79
MB-RGAN (G)	18 890	N/A
MB-RGAN (Total)	461 233	26.10

4.4 消融实验

为了验证 MB-RGAN 模型中多分支结构、浅层特征共享机制、独立回归器设计、递归冻结训练策略以及关键超参数对模型性能的具体贡献,本节设计了 5 组变体模型与完整的 MB-RGAN 进行对比实验. 各变体模型的具体设置如表 5 所示.

实验结果如表 6 与图 15 所示, MB-RGAN 在所有指标上均表现最优,其中 Model B 和 Model

表 5 消融实验变体模型设置 Table 5 Setting of the ablation experiment variant models		
模型编号	核心变动描述	损失函数/训练策略
Model A	移除多分支结构, 仅在网络末端输出预测值	带梯度惩罚项的 Wasserstein 距离损失 + 回归一致性/递归训练
Model B	移除回归器 R, 移除回归一致性约束	带梯度惩罚项的 Wasserstein 距离损失/递归训练
Model C	移除首层特征共享, D 与 R 独立提取特征	带梯度惩罚项的 Wasserstein 距离损失 + 回归一致性/递归训练
Model D	移除递归训练, 所有层同时更新	全局端到端 (end-to-end) 训练
Model E	移除梯度惩罚项, 设定 $\lambda_{GP} = 0$	无约束 Wasserstein 距离损失/递归训练
MB-RGAN	完整模型, 保留所有改进模块	带梯度惩罚项的 Wasserstein 距离损失 + 回归一致性/递归训练

表 6 不同变体模型的消融实验结果对比
Table 6 Comparison of ablation experiment results for different variant models

模型变体	组分	平均 MAE	平均 RMSE	平均 R^2
Model A	有机相	0.025 9	0.039 6	0.938 0
	水相	0.032 2	0.050 5	0.928 4
Model B	有机相	0.036 6	0.059 2	0.688 9
	水相	0.039 7	0.059 5	0.902 1
Model C	有机相	0.024 6	0.038 1	0.934 0
	水相	0.031 8	0.049 3	0.929 1
Model D	有机相	0.021 3	0.032 6	0.957 5
	水相	0.028 8	0.045 2	0.943 9
Model E	有机相	0.032 5	0.046 8	0.895 0
	水相	0.039 2	0.059 4	0.882 0
MB-RGAN	有机相	0.017 6	0.027 7	0.968 5
	水相	0.021 5	0.036 2	0.961 9

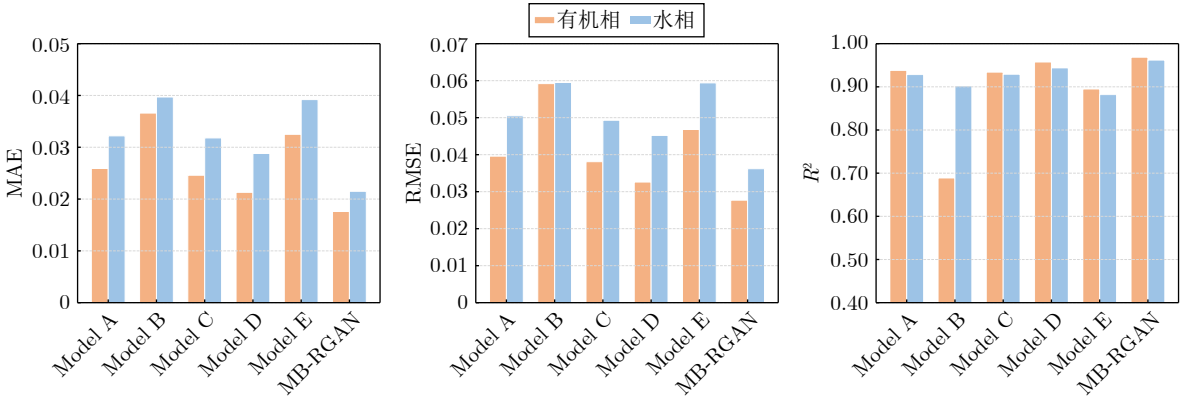


图 15 不同模型性能指标对比

Fig.15 Comparison of performance metrics for different models

E 性能表现最差. Model B 移除了回归器与回归一致性约束, 其有机相和水相平均 MAE 与平均 RMSE 误差均增大, 且有机相 R^2 跌至 0.688 9, 说明 Model B 虽然生成的样本在统计分布上接近真值, 但在具体的数值对应关系上存在巨大偏差, 这表明回归一致性约束对于建立输入特征与输出组分之间精确的映射关系十分重要. 它不仅约束生成器模拟数据的统计分布, 更通过均方误差反馈强制其符合物理规律的数值精度, 从而避免了生成对抗网络常见的模式坍塌问题. Model E 移除了梯度惩罚项, 其两相的平均 R^2 均降至 0.900 0 以下, 说明移除梯度惩罚后, 判别器无法满足 1-Lipschitz 连续性约束, 导致 Wasserstein 距离估计偏差较大, 无法为生成器提供稳定的梯度指引, 导致训练过程不稳定, 从而影响了最终的生成质量与预测精度. Model A (无多分支结构)、Model C (R、D 无特征共享层) 和 Model D (无递归训练) 的性能均有不同程度的下降. 由于稀

土萃取是一个由数十级萃取槽串联组成的复杂物理系统, 各关键监测级 (如第 15、25、30、35、45 级) 所处工序位置不同, 其组分分布规律存在差异. Model A 因缺乏多分支结构, 单一输出层结构网络在训练时旨在最小化所有输出节点的全局整体误差. 在反向传播过程中, 为了降低整体损失, 不同监测级的梯度方向可能存在冲突, 而多分支结构各分支有自己独立的参数, 每个分支输出都能达到较高精度.

为了更直观地展示各模型在不同指标上的综合性能表现, 图 16 给出了各变体模型在归一化性能指标下的雷达图对比结果.

为了进一步探究梯度惩罚系数 λ_{GP} 对模型预测性能的影响. 将模型在 $\lambda_{GP} = 0$ (Model E) 与 $\lambda_{GP} = 10$ (MB-RGAN) 两种设置下的预测结果进行了可视化对比分析. 图 17 展示了两个模型在不同梯度惩罚系数下各级有机相和水相的 KDE 分布曲线. 可以看出在第 30 级、第 45 级有机相和第

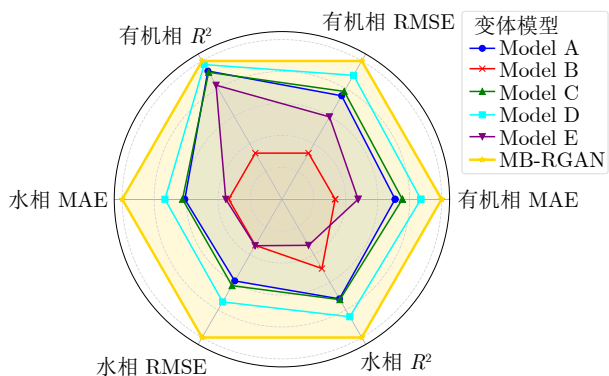


图 16 各模型性能指标雷达图

Fig.16 Performance metrics radar chart of each model

25 级、第 45 级水相中, Model E 的预测分布呈现出极度尖锐的峰值, 并且在第 25 级水相中和第 30 级有机相中, 蓝色虚线的峰值位置明显向右偏移 (偏离真实均值). 这种均值与方差的双重偏差表明, 在缺乏 1-Lipschitz 连续性约束的情况下, 模型对 Wasserstein 距离估计偏差较大, 导致生成器无法正确学习到数据的真实分布. 结果证明了梯度惩罚项在维持生成对抗网络训练稳定性方面的关键作用, 有效地约束了判别器的梯度范数, 避免模型在训练过程中出现模式坍塌问题.

为了探究回归一致性约束在 MB-RGAN 模型中的作用机制, 进行不同权重系数 $\lambda \in \{0, 0.1, 1.0, 5.0\}$ 的对比实验 (实验中令 $\lambda_1 = \lambda_2 = \lambda$), 将服从高斯分布的随机噪声 z 及条件变量 c 分别采样作为 MB-RGAN 的生成器 G 的输入, G 的输出以第 25 萃取级水相的 200 个样本点组分含量数据表示. 如图 18

所示, 图中蓝色曲线为真实样本数据, 橙色曲线为生成样本数据.

当回归一致性约束权重 λ 为 0 时, 回归器 R 对 G 无回归约束作用, 可以看到生成样本剧烈振荡, 当 λ 分别为 0.1、1.0 时, 可以观察到随着 MB-RGAN 中回归器 R 对生成器 G 的反馈作用增强, 生成的组分含量值上下波动较小, 抑制了离群点和假峰的产生, 更能贴合真实的稀土萃取样本点数据; 当 λ 为 5.0 时, 约束太过强烈, 导致生成器 G 难以充分学习数据的特征, 抑制了 G 的生成能力. 结果表明 MB-RGAN 各分支通过回归器 R 对 G 的约束作用, 使得 MB-RGAN 生成器 G 的样本生成性能得到提高.

5 结束语

针对基于萃取机理的稀土萃取分离流程模拟方法难以吻合实际萃取生产工况, 提出了一种基于多分支回归生成对抗网络的流程模拟方法. 依据稀土串级萃取工艺特点, 设计了多分支生成器网络, 逐级捕捉萃取槽间的深层耦合特征; 提出判别器与回归器浅层特征共享机制以提升预测性能, 并通过回归一致性损失进一步约束生成样本的真实性; 引入带梯度惩罚的 Wasserstein 距离并设计一种递归渐进式对抗训练策略解决网络训练不稳定及模式崩溃问题. 仿真结果表明, MB-RGAN 性能优于多种对比模型, 表明了本文方法的有效性.

未来的研究将重点探索机理与数据混合驱动的建模模式, 尝试将物料平衡方程等物理化学机理嵌入神经网络结构中, 进一步提升模型预测结果的物

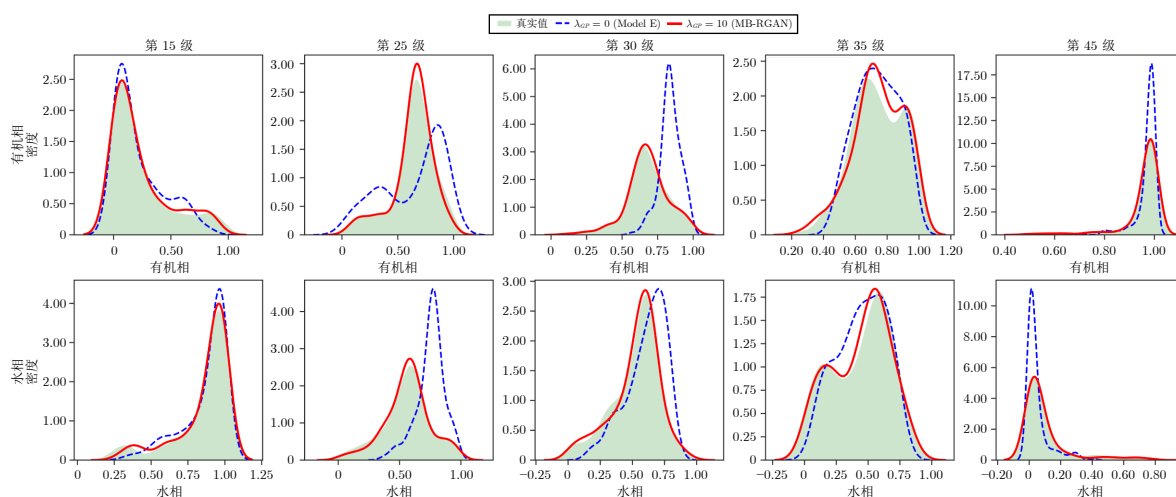
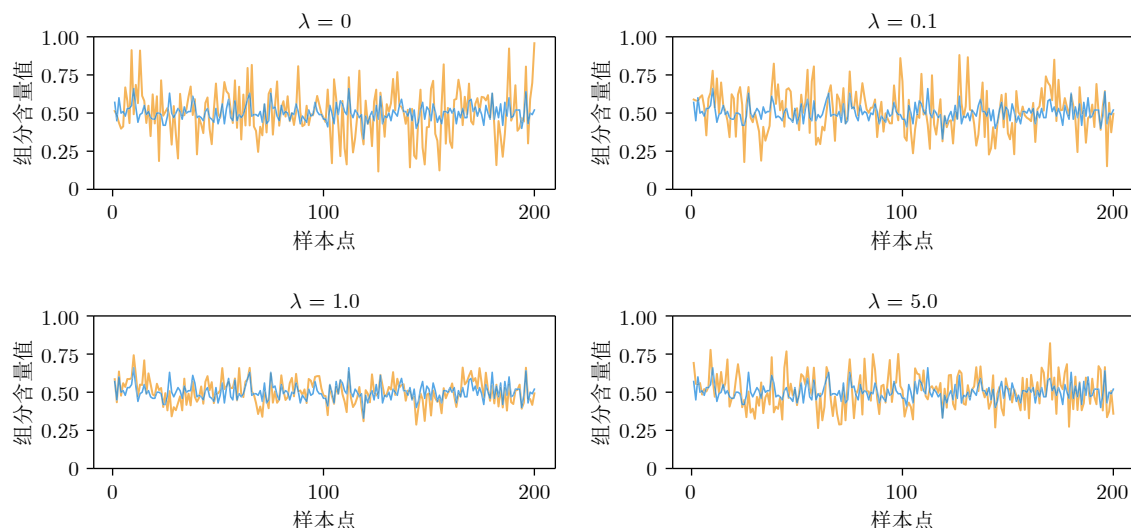


图 17 不同梯度惩罚系数下的组分分布核密度估计曲线对比

Fig.17 Comparison of kernel density estimation curves for component distributions under different gradient penalty coefficients

图 18 超参数 λ 对生成组分含量样本质量的影响Fig. 18 The influence of the hyperparameter λ on the quality of the generated component concentration samples

理一致性与可解释性, 为稀土萃取过程的智能化优化控制提供更加坚实的决策依据。

参考文献

- Feng Zong-Yu, Wang Meng, Zhao Long-Sheng, Xu Yang, Zhang Yong-Qi, Huang Xiao-Wei. Development status and prospect of rare earth extraction and separation technology. *Journal of the Chinese Society of Rare Earths*, 2021, **39**(3): 469–478 (冯宗玉, 王猛, 赵龙胜, 徐阳, 张永奇, 黄小卫. 稀土元素萃取分离提纯技术发展现状与展望. 中国稀土学报, 2021, **39**(3): 469–478)
- Xu Guang-Xian, Li Biao-Guo, Yan Chun-Hua. Progress on theory of counter current extraction and its applications in rare earth separation processes. *Chinese Rare Earths*, 1985(1): 56–67 (徐光宪, 李标国, 严纯华. 串级萃取理论的进展及其在稀土工艺中的应用. 稀土, 1985(1): 56–67)
- Wu Sheng, Liao Chun-Sheng, Jia Jiang-Tao, Yan Chun-Hua. Static design for multiple components and multiple outlets rare earth countercurrent extraction (I): Algorithm of static design. *Journal of the Chinese Society of Rare Earths*, 2004, **22**(1): 17–21 (吴声, 廖春生, 贾江涛, 严纯华. 多组分多出口稀土串级萃取静态优化设计研究 (I) 静态设计算法. 中国稀土学报, 2004, **22**(1): 17–21)
- Wu Sheng, Liao Chun-Sheng, Jia Jiang-Tao, Yan Chun-Hua. Static design for multi-component and multi-outlet rare earth counter-current extraction (II): Static design and its verification. *Journal of the Chinese Society of Rare Earths*, 2004, **22**(2): 171–176 (吴声, 廖春生, 贾江涛, 严纯华. 多组份多出口稀土串级萃取静态优化设计研究 (II) 静态程序设计及动态仿真验证. 中国稀土学报, 2004, **22**(2): 171–176)
- Zhu Jian-Yong, Yang Hui, Lu Rong-Xiu, Xu Fang-Ping, Yu Yun-Jun. Static setting and dynamic compensation based optimal control for the flow rate of the reagent in CePr/Nd extraction process. *Acta Automatica Sinica*, 2019, **45**(6): 1186–1197 (朱建勇, 杨辉, 陆荣秀, 徐芳萍, 余运俊. 基于静态设定和动态补偿的铈镨/钕萃取过程药剂剂量优化控制. 自动化学报, 2019, **45**(6): 1186–1197)
- Yang Hui, Chai Tian-You. Concatenation extraction separation constituent content soft measurement and application. *Nonferrous Metallurgical Design and Research*, 2003(S1): 129–134 (杨辉, 柴天佑. 串级萃取分离组分含量软测量及应用. 有色冶金设计与研究, 2003(S1): 129–134)
- Lyon K L, Utgikar V P, Greenhalgh M R. Dynamic modeling for the separation of rare earth elements using solvent extraction: Predicting separation performance using laboratory equilibrium data. *Industrial & Engineering Chemistry Research*, 2017, **56**(4): 1048–1056
- Srivastava V, Werner J, Honaker R. Design of multi-stage solvent extraction process for separation of rare earth elements. *Mining*, 2023, **3**(3): 552–578
- Ryu K H, Lee C, Lee G G, Jo S, Sung S W. Modeling and simulation of solvent extraction processes for purifying rare earth metals with PC88A. *Korean Journal of Chemical Engineering*, 2013, **30**(10): 1946–1953
- Yun C Y, Lee C, Lee G G, Jo S, Sung S W. Modeling and simulation of multicomponent solvent extraction processes to purify rare earth metals. *Hydrometallurgy*, 2016, **159**: 40–45
- Chang Zhi-Duo, Wu Sheng, Cheng Fu-Xiang, Liao Chun-Sheng, Yan Chun-Hua. Simulation system of mixer-settlers in rare earth extraction separation. *Journal of the Chinese Society of Rare Earths*, 2023, **41**(3): 623–630 (常智舵, 吴声, 程福祥, 廖春生, 严纯华. 稀土萃取分离中混合澄清槽的仿真模拟系统. 中国稀土学报, 2023, **41**(3): 623–630)
- Lu Rong-Xiu, Rao Yun-Chun, Yang Hui, Zhu Jian-Yong, Yang Gang. Prediction of Pr/Nd component content based on improved just-in-time learning algorithm. *Control Theory & Applications*, 2020, **37**(8): 1846–1854 (陆荣秀, 饶运春, 杨辉, 朱建勇, 杨刚. 基于改进即时学习算法的铈/钕元素组分含量预测. 控制理论与应用, 2020, **37**(8): 1846–1854)
- Giles A E, Aldrich C, van J S J. Modelling of rare earth solvent extraction with artificial neural nets. *Hydrometallurgy*, 1996, **43**(1): 241–255
- Ma L, Wang M, Peng K. A two-phase soft sensor modeling framework for quality prediction in industrial processes with missing data. *Journal of Process Control*, 2023, **129**: Article No. 103061
- Anitha M, Singh H. Artificial neural network simulation of rare earths solvent extraction equilibrium data. *Desalination*, 2008, **232**(1–3): 59–70
- Teerapittayanon S, McDanel B, Kung H T. BranchyNet: Fast inference via early exiting from deep neural networks. arXiv preprint arXiv: 1709.01686, 2017.
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2014: 2672–2680
- Zhang K. On mode collapse in generative adversarial networks. In: Proceedings of International Conference on Artificial Neural

- Networks. Virtual Event: Springer International Publishing, 2021. 563–574
- 19 Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv: 1411.1784, 2014.
 - 20 Mohammadzadeh M, Ghadami A, Taheri A, Behzadipour S. cGAN-based high dimensional IMU sensor data generation for enhanced human activity recognition in therapeutic activities. *Biomedical Signal Processing and Control*, 2025, **103**: Article No. 107476
 - 21 Isola P, Zhu J Y, Zhou T, Efros A A. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017. 1125–1134
 - 22 Tavalara T E, Niaz M K K, Arole V, Chen W, Frankel W, Gurcan M N. A modular cGAN classification framework: Application to colorectal tumor detection. *Scientific Reports*, 2019, **9**(1): Article No. 18969
 - 23 Aggarwal K, Kirchmeyer M, Yadav P, Keerthi S S, Gallinari P. Benchmarking regression methods: A comparison with CGAN. arXiv preprint arXiv: 1905.12868, 2019.
 - 24 Megahed M, Mohammed A. Multi-GANs with shared generator: An approach for handling mode collapse issue. In: Proceedings of the 6th International Conference on Computing and Informatics (ICCI). Cairo, Egypt: IEEE, 2024. 483–489
 - 25 Zheng Nian-Zu, Ding Jin-Liang. Regression GAN based prediction for physical properties of total hydrogen in crude oil. *Acta Automatica Sinica*, 2018, **44**(5): 915–921 (郑念祖, 丁进良. 基于 Regression GAN 的原油总氢物性预测方法. *自动化学报*, 2018, **44**(5): 915–921)
 - 26 Barnett S A. Convergence problems with generative adversarial networks (GANs). arXiv preprint arXiv: 1806.11382, 2018.
 - 27 Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A C. Improved training of Wasserstein GANs. arXiv preprint arXiv: 1704.00028, 2017.
 - 28 Basak D, Pal S, Patranabis D C. Support vector regression. *Neural Information Processing Letters and Reviews*, 2007, **11**(10): 203–224
 - 29 Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA: ACM, 2016. 785–794
 - 30 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, **9**(8): 1735–1780
 - 31 Chung J, Gulcehre C, Cho K H, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv: 1412.3555, 2014.
 - 32 Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, **323**(6088): 533–536
 - 33 Zhu Jian-Yong, Wang Wei, Yang Hui, Xu Fang-Ping, Lu Rong-Xiu. Process simulation of rare earth extraction based on multi-branch residual deep network. *Control Theory & Applications*, 2022, **39**(12): 2242–2253 (朱建勇, 王伟, 杨辉, 徐芳萍, 陆荣秀. 基于多分支残差深度网络的稀土萃取流程模拟. *控制理论与应用*, 2022, **39**(12): 2242–2253)

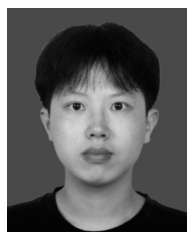


朱建勇 华东交通大学电气与自动化工程学院教授. 主要研究方向为复杂工业过程建模与优化控制, 预测控制, 智能控制.

E-mail: zhujiyemail@163.com

(**ZHU Jian-Yong** Professor at the School of Electrical and Automation Engineering, East China Jiaotong University. His

research interests include modeling and optimal control of complex industrial processes, predictive control, and intelligent control.)



王佳浩 华东交通大学电气与自动化工程学院硕士研究生. 主要研究方向为复杂工业过程建模与优化控制.

E-mail: wangjiahao2628@163.com

(**WANG Jia-Hao** Master student at the School of Electrical and Automation Engineering, East

China Jiaotong University. His research interests include modeling and optimal control of complex industrial processes.)



韩泽润 华东交通大学电气与自动化工程学院硕士研究生. 主要研究方向为复杂工业过程建模与优化控制.

E-mail: 13955656869@163.com

(**HAN Ze-Run** Master student at the School of Electrical and Automation Engineering, East China

Jiaotong University. His research interests include modeling and optimal control of complex industrial processes.)



廖良芳 中稀江西稀土有限公司主办. 2007 年获得江西理工大学学士学位. 主要研究方向为流程自动化控制, 工厂数字化.

E-mail: Liaoliangfang@regcc.cn

(**LIAO Liang-Fang** Coordinator at China Rare Earth Jiangxi Co., Ltd.

He received his bachelor degree from Jiangxi University of Science and Technology in 2007. His research interests include process automation control and factory digitalization.)



杨 辉 华东交通大学电气与自动化工程学院教授. 主要研究方向为复杂工业过程建模、控制与优化, 轨道交通自动化与运行优化. 本文通信作者.

E-mail: yhshuo@163.com

(**YANG Hui** Professor at the School of Electrical and Automation Engineering, East China Jiaotong University. His

research interests include modeling, control and optimization of complex industrial processes, rail transit automation, and operation optimization. Corresponding author of this paper.)