

沉积过程模拟驱动下的深度学习地质建模方法

刘彦锋,段太忠,黄 渊,张文彪,李 蒙

(中国石化石油勘探开发研究院,北京 102206)

摘要:油气藏勘探开发逐步向深层化、复杂化方向发展,观测数据不足、分辨率低等资料难题突显,传统的地质建模方法无法适应技术需求。以深度学习为代表的智能化地质建模方法可以充分整合多尺度、多维度的数据信息以及专家认识,是地质建模技术发展的重要方向。在综合分析地层沉积模拟和深度学习地质建模技术优缺点的基础上,探索形成了沉积过程模拟驱动的深度学习地质建模方法。首先,基于综合地质分析开展沉积正演模拟,分析参数不确定性,通过参数扰动形成大规模地质模型作为训练样本库;其次,利用条件化生成对抗网络学习样本库中蕴含的地质模式和规律,其中生成网络以井-震等条件数据作为输入、地质模型作为输出;最后,利用训练后生成网络在实际条件数据上的应用,得到目标区块的地质模型。通过在四川盆地普光气藏主力区块典型地质剖面的测试应用,该方法的可行性得到了验证,并分析了训练样本库大小对模拟结果的影响。沉积模拟和深度学习相结合,弥补了训练样本不足的缺陷,间接实现了知识驱动的深度学习地质建模,具有重要的推广意义。

关键词:人工智能;大数据;知识驱动;生成对抗网络;沉积过程模拟;深度学习地质建模;普光气藏;四川盆地

中图分类号:TE132.1

文献标识码:A

Deep learning-based geological modeling driven by sedimentary process simulation

LIU Yanfeng, DUAN Taizhong, HUANG Yuan, ZHANG Wenbiao, LI Meng

(Petroleum Exploration and Production Research Institute, SINOPEC, Beijing 102206, China)

Abstract: Problems including insufficient quantity and low resolution of data face exploration and development of deep and complex reservoir targets, and the traditional geological modeling methods have been inadequate in terms of technical needs. The intelligent geological modeling represented by deep learning capable of fully integrating multi-scale and multi-dimensional data as well as expert knowledge, is a key research and development direction of geological modeling technology. The study discusses the deep learning-based geological modeling driven by sedimentary process simulation following the comprehensive analysis of the advantages and disadvantages of stratigraphic forward modeling and deep learning-based geological modeling technology. First, forward modeling of sedimentation is carried out based on comprehensive geological analysis, parameter uncertainty is analyzed, and a large amount of geological models are established through parameter disturbance as a training dataset; Second, the geological patterns contained in the learning dataset are learned with the conditional Generative Adversarial Nets (cGAN), in which the Generative Adversarial Networks (GAN) takes the conditional data such as well and seismic data as the input, and the geological model as the output; Finally, the trained GAN is applied to the real conditional data to obtain the geological model of the target block. The feasibility of this method is verified through testing on the typical geological profiles of the main block of Puguang gas reservoir, and the impact of the training dataset scale on simulation results is analyzed. The combination of sedimentary simulation and deep learning could make up for the shortage of training data and indirectly realize knowledge-driven deep learning-based geological modeling. The method is therefore of great significance to popularization.

收稿日期:2022-06-01;修订日期:2022-11-03。

第一作者简介:刘彦锋(1986—),男,博士、副研究员,油气田开发地质。E-mail: liuyf_syky@sinopec.com。

通讯作者简介:段太忠(1961—),男,博士、教授,油气田开发地质。E-mail: duantz_syky@sinopec.com。

基金项目:中国科学院A类战略性先导科技专项(XDA14010204);中国石化科技部项目(P21038-3)。

Key words: artificial intelligence (AI), big data, knowledge-driven, Generative Adversarial Networks (GAN), sedimentary process simulation, deep learning-based geological modeling, Puguang gas reservoir, Sichuan Basin

引用格式:刘彦锋,段太忠,黄渊,等. 沉积过程模拟驱动下的深度学习地质建模方法[J]. 石油与天然气地质,2023,44(1):226-237. DOI:10.11743/ogg20230119.

LIU Yanfeng, DUAN Taizhong, HUANG Yuan, et al. Deep learning-based geological modeling driven by sedimentary process simulation[J]. Oil & Gas Geology, 2023, 44(1): 226-237. DOI: 10.11743/ogg20230119.

三维地质模型是勘探开发决策的重要依据。地质建模的基本原则是综合应用地质、测井、地震和生产动态等各类数据和知识,但深层油气藏可用的资料相对较少、分辨率低,对地质建模的方法要求高^[1-2],地质知识需要发挥更大的作用。40年来,地质建模大致经历了两点地质统计、多点地质统计建模阶段^[3-4],随着大数据和人工智能技术的发展,目前进入了以深度学习为代表的人工智能地质建模发展阶段。从地质建模技术发展历程看,知识在地质建模中的作用越来越大。

在传统的两点地质统计学中,知识通过变差函数的方式参与地质建模,以概率统计的形式体现地质体的长、宽、厚度和方向等参数^[5]。基于目标的建模方法可以给出更多的关于几何形态的地质认识,比如弯曲度、长宽比、厚度范围等参数,用来模拟更具象的地质模型^[6]。多点地质统计学建模通过训练图像表达概念性地质知识,表达的内容比变差函数更丰富,表达的方式比单纯的几何参数更灵活^[7]。基于地质过程的建模是量化地质知识驱动的建模方法,地质知识以微分方程的形式内嵌于地质建模算法之中,是纯粹的地质知识驱动。以深度学习为代表的人工智能地质建模方法以大规模训练样本为基础,地质知识蕴含在大规模训练样本中,相比于多点地质统计学,可以从训练样本中挖掘更丰富的地质知识,建立更符合地质规律的地质模型,并且在同类地质问题中可以一次训练多次应用。

生成对抗网络(GAN)是目前深度学习地质建模中采用的主要神经网络^[8],分为非条件和条件化生成对抗网络。非条件生成对抗网络与非条件地质建模类似,条件化生成对抗网络与条件化地质建模类似。非条件地质建模通过变差函数、训练图像或目标体长、宽、高等少量参数输入得到一系列三维模型,这些模型不考虑与条件数据的吻合情况,但能够体现较为真实的地质认识。生成对抗网络深度学习体现了类似过程。标准的生成对抗网络^[9],即非条件生成对抗网络,其目标是训练出一个性能良好的生成网络,输入一个低维的随机数序列生成到高维的网络层(地质模型),通过训练学习地质模式,利用训练后的生成网络可以得

到满足地质认识的非条件模拟结果^[10]。在生成对抗网络的地质模式重现能力得到广泛认可之后,众多学者将研究重点转移到对条件化生成对抗网络(cGAN)的地质建模上,以期在实用性方面取得更大突破。

目前生成对抗网络地质建模的条件化主要有两种途径:一是先训练非条件生成对抗网络,再二次条件化,比如预训练的生成器通过启发式迭代优化进行条件化,或者训练过程中通过正则化方法进行条件化,前人用卷积生成对抗网络建立了条件化的河流相地质模型^[11-12];另一种是通过条件化生成对抗网络,直接在网络的输入参数中加入条件化数据,比如井数据、地震数据以及砂地比等参数^[13-14],这种方法通用性更强,更受青睐。

目前,深度学习地质建模方法整体上仍处于探索阶段,实际工业落地应用难,面临的主要问题之一是缺少样本^[15],采用人工合成样本是目前比较可行的方式。尽管多点地质统计建模中训练图像获取的方法也适用于深度学习地质建模,包括手绘地质模式图、野外露头、卫星图像分析、地质条件类似的密井网区块、基于目标的模拟方法和基于沉积过程的模拟方法等。考虑到深度学习对样本的需求量大,基于目标模拟和沉积过程模拟是可选的方法。

沉积过程模拟与深度学习建模相结合可以实现优势互补。基于沉积过程模拟的建模方法可以更好地体现地质知识的约束,与地质统计学建模方法相比,井间预测更符合地质规律。但是,由于沉积模拟的输入数据通常都难以直接测量,不确定性较大,模拟结果与井数据、地震数据等条件数据吻合难度大,需要大量的手动调整工作,通常需耗时1~2个月才能获得与观测数据吻合度较高的模型,效率较低。深度学习建模对样本的需求量大,目前的技术还无法直接得到大规模地下实际模型用于深度神经网络训练。但沉积模拟恰好具备大规模样本生成能力,有效弥补了训练样本不足问题。本文试图将两种方法结合起来,建立沉积过程数值模拟驱动的深度学习地质建模方法,并利用普光气藏主力区块地质剖面对其进行验证。

1 沉积过程模拟驱动的深度学习地质建模方法

1.1 沉积过程驱动的深度学习地质建模基本思想

将沉积过程模拟技术和深度学习相结合,可以将沉积模拟中蕴含的地质知识以训练样本的形式赋予深度学习的神经网络参数,间接实现知识驱动的地质建模。量化领域知识与深度学习结合的方式有多种^[16-17],最直接的方法是将量化可微分的公式直接嵌入到神经网络的隐藏层,或嵌入到神经网络的目标函数;还有一种是为神经网络提供大规模训练图像。后者属于间接的知识驱动方法,解耦性好,不同类型的领域知识和神经网络可以灵活组合。

地层沉积过程正演数值模拟通过数值计算的方式,考虑各种地质因素的影响,再现地层形成过程,最终得到完整的二维或三维沉积模型^[18-19]。前人对地层沉积模拟进行了大量研究^[20-21],提出了多种沉积模拟方法,涵盖了大部分的沉积相类型。代表性的沉积模拟软件有法国石油研究院的Dinoisos,斯坦福大学的SedSim,斯伦贝谢公司的GPM等。地层沉积过程正演模型描述了可容空间、沉积物剥蚀、供给、生产、搬运、堆积以及压实作用,输出结果是模拟地层和一系列古环境条件,比如整个演化历程的古地貌、沉积间断、沉积相等。地层沉积数值模拟模型可以表示为:

$$M = F(s_1, s_2, s_3, \dots, s_n) \quad (1)$$

式中: M 表示模拟结果,即地质模型; F 表示正演模拟方程,通常是描述沉积物沉积搬运和沉淀的数学物理方程; $s_1, s_2, s_3, \dots, s_n$ 表示 n 个模拟的输入参数。由于沉积模拟输入参数多,且大部分难以直接测量,如地质历史时期的海平面曲线、碳酸盐产率、搬运速度、基底沉降速度等,导致模拟结果不确定性大。为了得到与实际的钻井或地震数据吻合度较高的模型,往往需要通过试错法或参数优化方法反复调整输入参数^[22-25]。

沉积过程模拟驱动的深度学习地质建模基本思路如下(图1)。针对某个区块,综合资料分析,获取初始地形、海平面曲线、碳酸盐产率、沉积物剥蚀搬运等相关参数,开展初步沉积模拟研究。以观测数据为基础,不断修改沉积模拟的输入,缩小其可能的取值区间,降低参数的不确定性,获得每个参数可能的最小值和最大值。根据超立方采样的参数采样方法,获取一定规模的等概率的参数组合,把这些参数的模拟结果(沉积模拟地质模型)作为训练样本提供给深度学习网络。

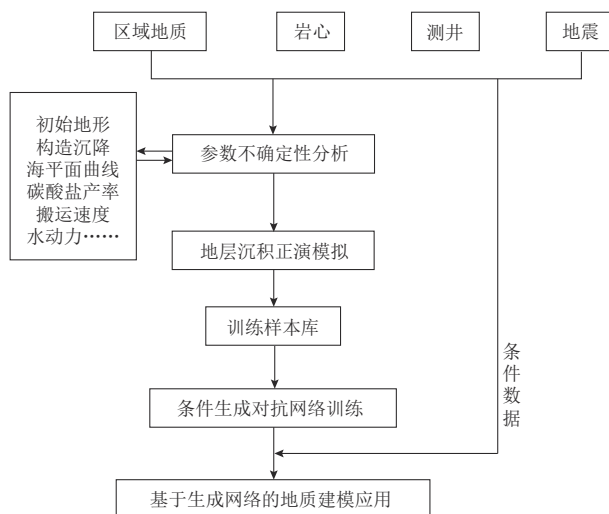


图1 沉积模拟驱动的深度学习地质建模技术流程

Fig. 1 Workflow showing deep learning-based geological modeling driven by sedimentary process simulation

利用训练后的生成网络,直接输入条件数据,即可得到该区块的地质模型。搭建训练条件化生成对抗网络时需要考虑实际沉积条件数据的情况,以地质模型上抽取的虚拟井数据,甚至基于地质模型合成地震数据作为训练输入,地质模型作为输出,具体方法见1.2节的描述。

1.2 基于生成对抗神经网络的地质建模方法

1.2.1 生成对抗网络地质建模原理

在深度学习框架下,地质建模是一个生成式问题。按照训练目标的特点,神经网络分为判别式网络和生成式网络:前者由高维神经网络层通过一系列的隐藏层得到低维输出层,比如分类、聚类等;后者输入低维神经网络层得到高维神经网络层,如输入低维的随机数据得到清晰的语音、图像、文字等^[26]。油气藏地质建模通常是输入稀疏的、高精度的井点数据,在稠密的、相对低精度的地震数据约束条件下建立高精度三维网格模型,可视为生成式神经网络问题。相对于判别式问题,生成式问题的机器学习难度大,直到生成式对抗网络的出现^[9,27],才大大加快了生成式问题的解决。该技术迅速在图像和语音合成领域得到普及,并应用到了地质建模领域^[28-29]。

生成对抗网络(GAN)由生成网络(G)和判别网络(D)组成,采用对抗学习的策略进行训练,目标是获取生成表达能力强的生成网络 G 。该训练框架中,生成网络(G)将从先验分布 $p(z)$ 采样的随机向量 z 映射到地质模型空间,鉴别网络(D)将输入地质模型映射到

似然概率(判断真假的概率)。 G 的作用是生成大量让 D 认为“真实”的模型,而 D 起着对抗的作用,即判断由 G 生成的模型为假,而来自样本库的模型为真, G 和 D 在相互竞争中学习,因此称为生成对抗网络。标准生成对抗网络的目标函数如下:

$$\min_G \max_D (D, G) = E_{x \sim p_{\text{data}}(x)} [\lg D(x)] + E_{z \sim p_z(z)} [\lg (1 - D(G(z)))] \quad (2)$$

式中: x 是来自 p_{data} 分布的样本; p_{data} 是全部的训练样本; $D(x)$ 是判别网络对样本的判别; z 是潜在空间上的随机向量; $G(z)$ 是生成的地质模型; $D(G(z))$ 是判别网络对生成的地质模型的判别; E 是一组样本已判别结果的期望。

通过大规模的训练可以得到生成网络 G ,它体现了随机向量 z 到地质模型的映射。变换不同的 z 可以得到相应的地质模型,但是 z 没有明确的物理意义。往往需要二次的条件化才能建立具有实际地质意义的模型^[11]。条件化生成网络输入有明确物理意义的建模参数^[30],输出结果更容易被解释。

1.2.2 条件化生成对抗网络地质建模

条件化生成对抗网络地质建模直接把井、震等条件数据作为网络的输入得到二维或三维地质模型。标准GAN提出之后,为了约束生成的结果,使其更加可控,很快就出现了条件化GAN,称为cGAN^[31]。cGAN是GAN的扩展,它的生成网络和判别网络都利用条件数据 y ,条件数据可以是任意类型,比如类型标签(相类型、是否为陡坡沉积、是否发育生物等),或者储层分布的条件约束(波阻抗、砂体发育概率等),同时参与到生成网络和判别网络的训练。在生成网络中,通常把条件数据作为 y 和随机向量 z 一起作为网络的输入层;在判别网络中,样本标签数据 x 和条件数据 y 同时作为输入层(图2)。cGAN的目标函数可以表示为如下的形式:

$$\min_G \max_D (D, G) = E_{x \sim p_{\text{data}}(x)} [\lg D(x, y)] + E_{z \sim p_z(z)} [\lg (1 - D(G(z, y)))] \quad (3)$$

在条件化生成对抗网络的基本框架下,在不同领域的应用,产生了不同的网络结构。地质模型具有多尺度复杂特征,需要生成网络具备更强的生成能力,U-Net神经网络^[32]具有多尺度卷积和跨层连接的特点,在刻画多尺度结构特征方面具有较强能力,把它作为cGAN的生成网络,可以用于生成二维和三维高精度地质模型。通常,在判别网络中不只是对整个输出结果做单一的判别,而是考虑地质模型的多尺度结构

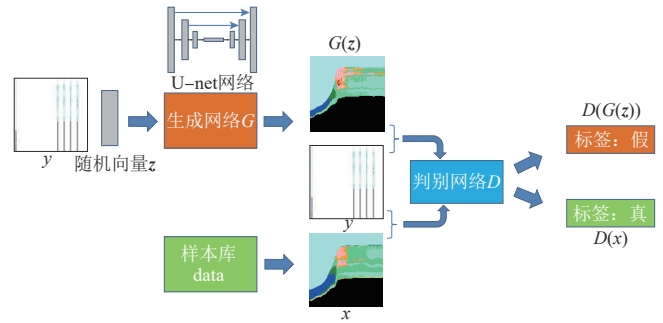


图2 条件化生成对抗网络基本结构

Fig. 2 Basic architecture of a cGAN

化特征,进行多尺度判别。

2 深度学习地质建模方法在普光气藏的应用

2.1 基于沉积过程模拟的样本生成

2.1.1 碳酸盐岩地层沉积正演模拟器

地层沉积模拟可以生成沉积规律约束的二维和三维地质模型。碳酸盐沉积体系的形成、演化和消亡除了受构造活动、海平面变化、气候条件、海洋环境和水动力条件等因素的直接或间接控制外,还尤其受到体系中生物与生态因素控制^[33-35]。我们采用了自主研发的产碳酸盐地层沉积正演模拟方法^[36]和CarbSIMS软件,考虑了最新的碳酸盐岩工厂、生态可容空间和层序地层学等基本原理。把地层形成过程视为生物能、势能和水体动能这3种能量场相互耦合的过程,生物能体现为碳酸盐岩生物的生长,动能表现为风能和风能引起的波浪能,势能主要以地形的形式体现。模型的具体细节参考文献^[35]。

CarbSIMS软件中的地层沉积正演模拟技术,包含碳酸盐岩原位生长模型、水体能量分布模型和沉积物搬运模型,可以模拟碳酸盐岩沉积物剥蚀、搬运和沉积作用,形成并展示碳酸盐岩体系的几何形态、沉积层序展布,还可以通过沉积反演模拟技术自动匹配沉积过程模拟结果与观测数据^[21-22]。该软件可以模拟台缘带、缓坡、孤立台地以及复杂地形下的碳酸盐岩沉积体系。

在该模型中,输入参数包括4大类12小类。这4大类包括可容空间类、沉积物供给类、沉积物搬运剥蚀类和水体动能分布类,其涉及的每一小类又包括多个具体参数。下表列出了主要参数类型和推荐的分析方法(表1)。

表1 碳酸盐岩沉积模拟主要参数类型及分析方法

Table 1 Main parameter types and analytical methods of sedimentary process simulation of carbonates

参数类型	具体参数	分析方法
可容空间类	构造沉降曲线	地层回剥法,地震资料解释
	海平面曲线	Haq曲线,岩石的水深指示曲线
	初始地形	标志层厚度,沉积微相对水深指示意义
沉积物供给类	沉积物供给速率,最大产率	地层厚度,井上地层厚度序列
	沉积物供给集中度,透光带厚度	物源方向初始地层厚度变化程度,高能相带厚度
	沉积物供给成分比例,产率下降系数	观测数据中岩性比例,岩性变化频率
沉积物搬运剥蚀类	势能扩散系数	沿沉积物搬运方向的地层厚度变化
	扩散系数变化因子	岩性纵向变化程度
	动能对流系数	沿沉积物搬运方向的岩性变化程度
流体动能分布类	波浪能	水体流速相对大小
	风能	高能 and 低能相带分布的突变程度
	地形消浪能	沉积相对水体能量的指示意义

2.1.2 普光区块沉积正演模型基本框架

前人对该地区的层序地层、沉积特征、地质建模方法进行了大量研究^[37-39],为沉积过程模拟参数获取和深度学习地质建模方法分析奠定了基础。普光气藏位于川东断褶带双石庙—普光NE向背斜构造带上,是川东北地区已探明的大型碳酸盐岩构造-岩性气藏^[40]。普光气藏主要含气层系是上二叠统长兴组和下三叠统飞仙关组,处于有利的台地边缘沉积相带。其中,长兴组台地边缘生物礁及飞仙关组台地边缘颗粒滩为该区最有利的沉积相,是模拟的主要层位。

碳酸盐岩结合地质分析和前人研究成果,初步获得该地区海平面变化、构造沉降、碳酸盐产率等参数,对主产区垂直台地边缘的7口井开展了沉积模拟框架下的沉积相解释,指导正演模拟参数的调整。搭建了

该地区的地层沉积过程模拟基本框架,构建了样本库的沉积模拟基础模型(图3)。沉积模拟范围为19.0 km×8.8 km,时间跨度为3.5 Ma,平面网格数为129×129,总时间步数为200。

根据深度学习对泛化能力的要求,需要训练样本库涵盖尽可能多的地质模式,因此对基础模型要求不高,不要求与井数据完全吻合,甚至允许有些井有较大差别,只要模拟出整体的沉积规律即可。

2.1.3 样本库构建

为了保证神经网络模型的泛化能力,在手动调整沉积模型的基础上,应使样本库涵盖更多的沉积特征。在样本库构建过程中,针对不确定性参数,在计算量允许的情况下,以基础模型参数为基础尽量扩大参数扰动范围,获取尽量多的样本。

根据该地区前期的资料分析,我们认为初始地形

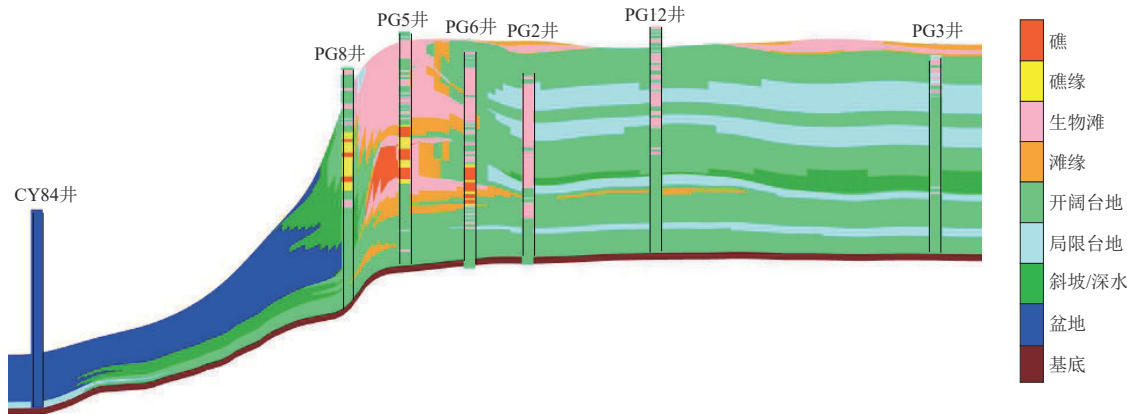


图3 手动参数调整的沉积正演模拟剖面与井资料对比

Fig. 3 Comparison of stratigraphic forward modeling by manual parameter adjustment with well data

的不确定性较低,采用确定值。海平面曲线、碳酸盐产率、沉积物搬运等相关参数共计41个,不确定性较大,以手动调整的参数值为基础,分析不同参数变化区间中样本地质模式的多样性,最终确定50%参数变化区间(表2)。这41个参数中,参数名字的最后—一个数字指示这类参数的第一个参数,有些参数考虑了时间的变化,比如根据时间段再划分为3个,其中长兴组2个,飞仙关1个。具体参数意义如下:参数1—6与海平面曲线有关(分别表示三级和四级海平面旋回正弦函数的振幅、周期和相位,以及线性约束),第7个参数与构造沉降有关(给定构造沉降面数据的系数),参数8—11与沉积物搬运有关(分别是势能和动能在 X 和 Y 方向引起的沉积物搬运系数),参数12—27与碳酸盐产率有关(分别是势能产率最大幅度、势能产率递减系数、透光带厚度、动能产率系数、动能产率基准值,以及势能相关产率的权重),参数28—41与水体动能有关(即动能幅度、波浪能下降系数、地形消浪能系数、浪基面和风能系数)。

考虑上述41个不确定性较大的参数,在确定的参数区间内,通过超立方随机采样的方式生成大量的参

数组组合,构建深度学习所需的样本库。本实例中,样本库为地质模型二维剖面。模型以二维矩阵表示,矩阵尺寸为 128×128 ,通过人工和自动筛选,去除模拟结果明显不合理的少量样本后,该样本库共计样本11 503个,训练集和测试集随机划分,通常80%用于训练,20%用于测试。从训练集中随机抽取几个样本,如图4所示(不同颜色表示不同的沉积相带),可以看出它们具备不同的沉积相结构样式,体现了样本包含的沉积模式的多样性。

2.2 神经网络搭建与训练

在条件化生成对抗网络的框架下,搭建并数据约束的深度学习地质模型网络,以U-net网络作为生成网络的基本结构,多尺度卷积网络作为判别网络,目标函数为判别损失加上带权重的第一范式损失(L1损失),具体的网络结构见图5。其中L1损失的权重是超参,需要人为给定,对模拟结果有一定影响。上采样和下采样作为基本的网络层单元,由卷积(转置卷积)层、丢弃层、归一化层和非线性化层组成。生成网络由一系列的下采样和上采样层组成,通过不断下采样捕捉

表 2 沉积反演模拟参数区间

Table 2 Parameter ranges of stratigraphic inverse modeling

序号	参数	最小值	最大值	序号	参数	最小值	最大值
1	AmpSeaL1	40.000	120.000	22	KEng2	1.000	3.000
2	PerdSeaL1	500.000	1500.000	23	KEng3	0.500	1.500
3	AmpSeaL2	20.000	60.000	24	bEng1	0.500	1.500
4	PerdSeaL2	80.000	150.000	25	bEng2	0.040	0.120
5	PerdSeaL3	20.000	60.000	26	bEng3	10.000	30.000
6	LinEuElev3	-55.000	-10.000	27	pPr1	0.350	1.000
7	SubsidenceScale1	0.500	1.500	28	KV1	5.000	15.000
8	PDX2	40.000	120.000	29	KV2	17.000	52.000
9	PDX1	800.000	2200.000	30	KV3	12.000	35.000
10	KDX1	5.000	15.000	31	KW1	0.007	0.022
11	KDX2	25.000	750.000	32	KW2	0.012	0.040
12	Apord1	2.500	7.500	33	KW3	0.007	0.022
13	Apord2	15.000	45.000	34	kR1	0.025	0.000
14	Apord3	7.500	22.000	35	kR2	0.020	0.060
15	Kpord3	0.010	0.030	36	kR3	0.007	0.020
16	Kpord1	0.010	0.030	37	kH1	0.030	0.090
17	Kpord2	0.020	0.070	38	kH2	0.100	0.300
18	Wpord1	30.000	90.000	39	kH3	0.030	0.090
19	Wpord2	15.000	45.000	40	WaveBase1	20.000	60.000
20	Wpord3	35.000	100.000	41	KL1	0.002	0.007
21	KEng1	0.020	0.060				

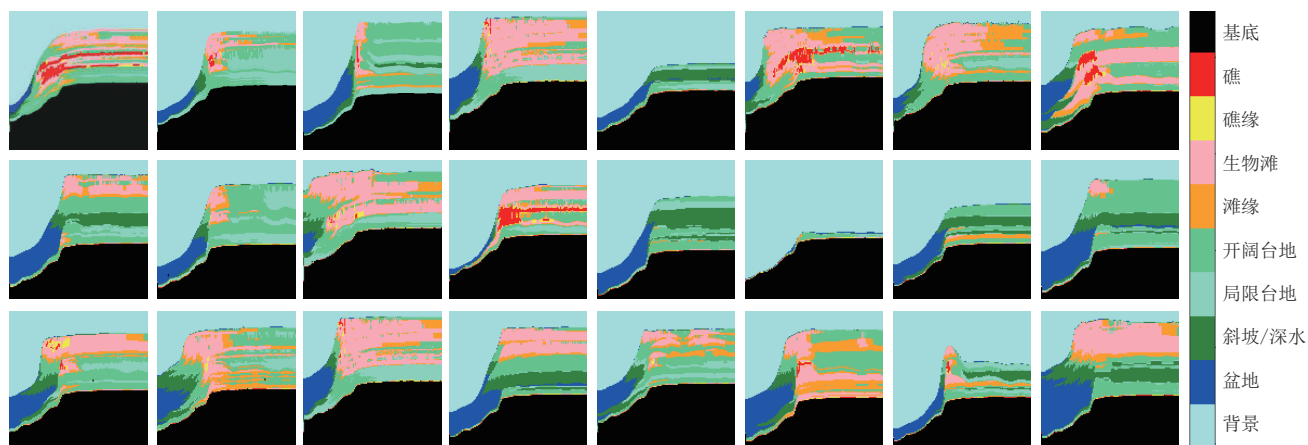


图4 从训练集随机抽取的一组训练样本

Fig. 4 A group of training samples randomly selected from the training dataset

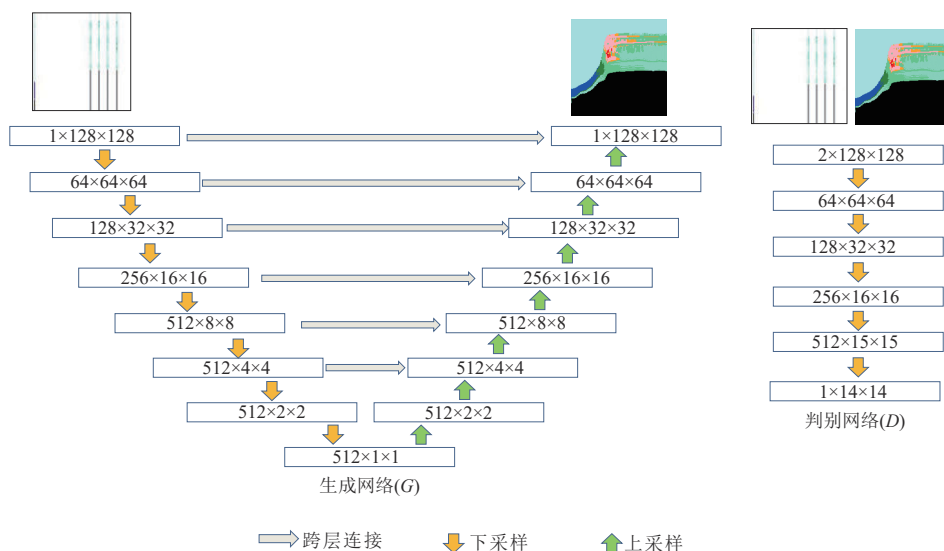


图5 条件化生成对抗网络中生成网络G和判别网络D的结构

Fig. 5 Architectures of generative (G) and discriminative (D) networks of cGAN

更大尺度的结构化特征,然后再通过上采样获取更小尺度的地质特征,同时通过跨层连接实现信息跨尺度融合,提高生成模型的分辨率。判别网络由一系列的下采样组成,最后得到 14×14 的判别结果,即对模拟结果给出196个判别指标,分别指示不同区域,并配合L1损失进行综合判别。

该生成网络以从沉积模拟样本中抽取的虚拟井数据为输入,二维地质模型为输出。设置最大训练轮次为200,通常迭代30轮次能得到比较理想的效果。从生成网络基本结构可以知道,按照多尺度结构化特征,以井数据为基础,先通过卷积运算刻画不同尺度的宏观特征,然后通过反卷积刻画小尺度特征,并通过跨尺度连接实现大尺度对小尺度的约束,生成新的地质模型,生成的地质模型整体上符合沉积过程模拟的地质

规律。图6是训练的生成网络在测试集上的表现,图中 a_1 到 a_9 , a_1 到 a_9 ,共计9组图形,分别为正演模拟器和生成网络模拟结果,两者基本一致。

2.3 模型效果分析

把实际观测的井上解释的沉积相数据作输入训练后的生成网络,得到该区块的模拟结果。分析了不同情况下的模拟结果,可以看出基于生成对抗网络的地质建模在井数据处与井上观测数据完全吻合,井间的沉积相分布符合地质规律。同时分析了深度学习地质建模与多点地质统计学建模效果的差异,在此基础上测试了改变训练样本数量及损失函数中L1损失权重时的模拟结果。

图7对比了深度学习地质建模与多点地质统计学

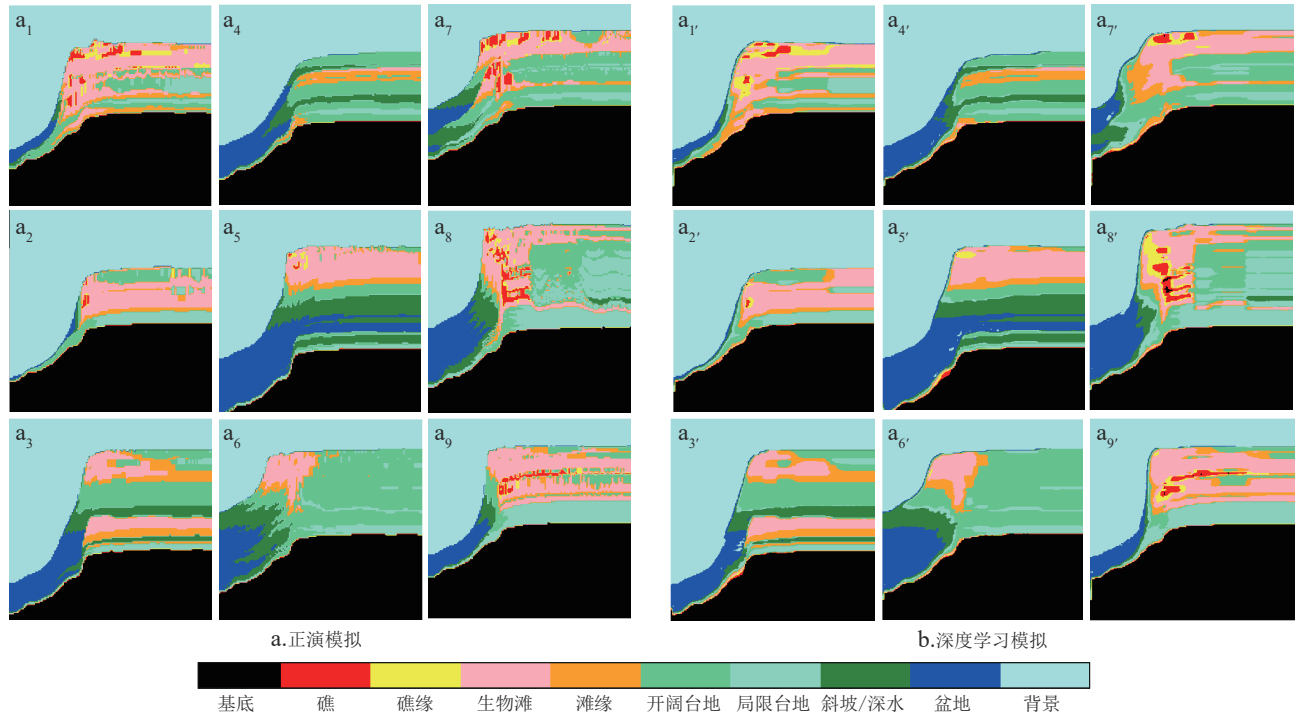


图6 训练后的生成网络模型在随机挑选的测试集上的模拟结果

Fig. 6 Results of applying the trained generative network model on randomly selected test dataset

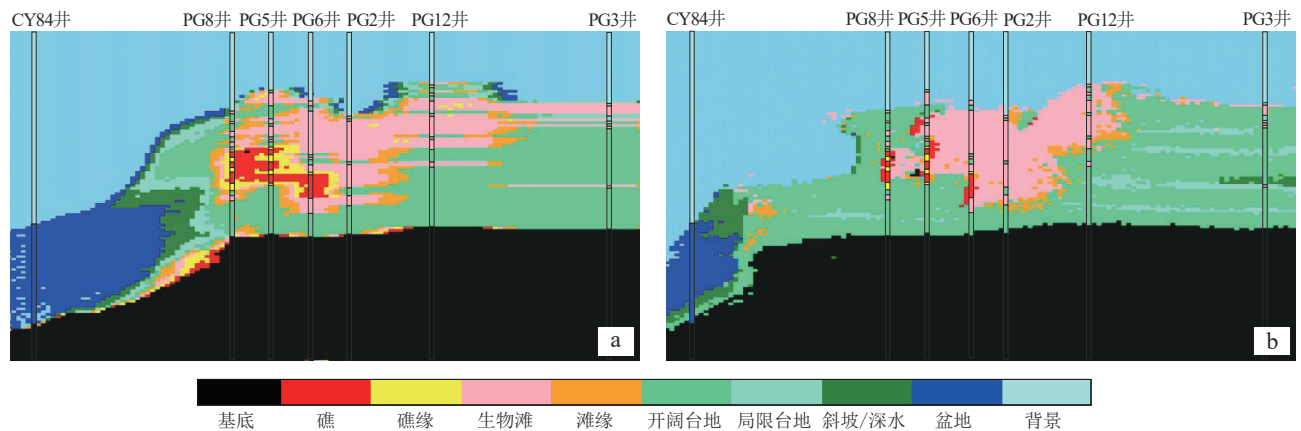


图7 深度学习地质建模(a)与多点地质统计学建模(b)比较

Fig. 7 Comparison of deep learning-based geologic modeling (a) and multi-point geostatistics-based modeling (b)

建模的效果,两者均以Cy84和PG8等8口井数据为硬数据。图7a是本文提出的深度学习地质建模,图7b是基于直接采样法的多点地质统计学建模,以上述沉积模拟基础模型为训练图像(图3),可以看出两种方法得到的相带整体分布规律一致,自左向右均表现出深水、斜坡、台缘和开阔台地的沉积相带特征。分析认为多点地质统计学地质统计学和深度学习地质建模各有优势。在相带分布的细节刻画方面,通常深度学习建模方法优于多点建模方法。多点地质统计学建模只需一个训练图像,深度学习方法需要大规模样本,在实际应用中前者更容易实施,另外对抗生成网络训练不容

易收敛,也进一步限制了该方法的应用潜力。

图8和图9对比分析了训练样本量和不同L1损失权重情况下的模拟结果,为了不遮挡模拟结果,4个子图只有一个显示了井上条件数据。

实际应用时应该增加样本,尽量减少过拟合现象,同时降低不合理样本对训练的影响。图8中4个子图采用的样本量分别占总样本的80%,60%,40%和20%,可以看出:训练集占总样本的60%和80%时模拟效果差别不大;40%时出现明显差异,特别是少井区域;20%时模拟结果明显不合理,完全不具备地质规律,预测失败,说明样本量过少,存在严重的过拟合

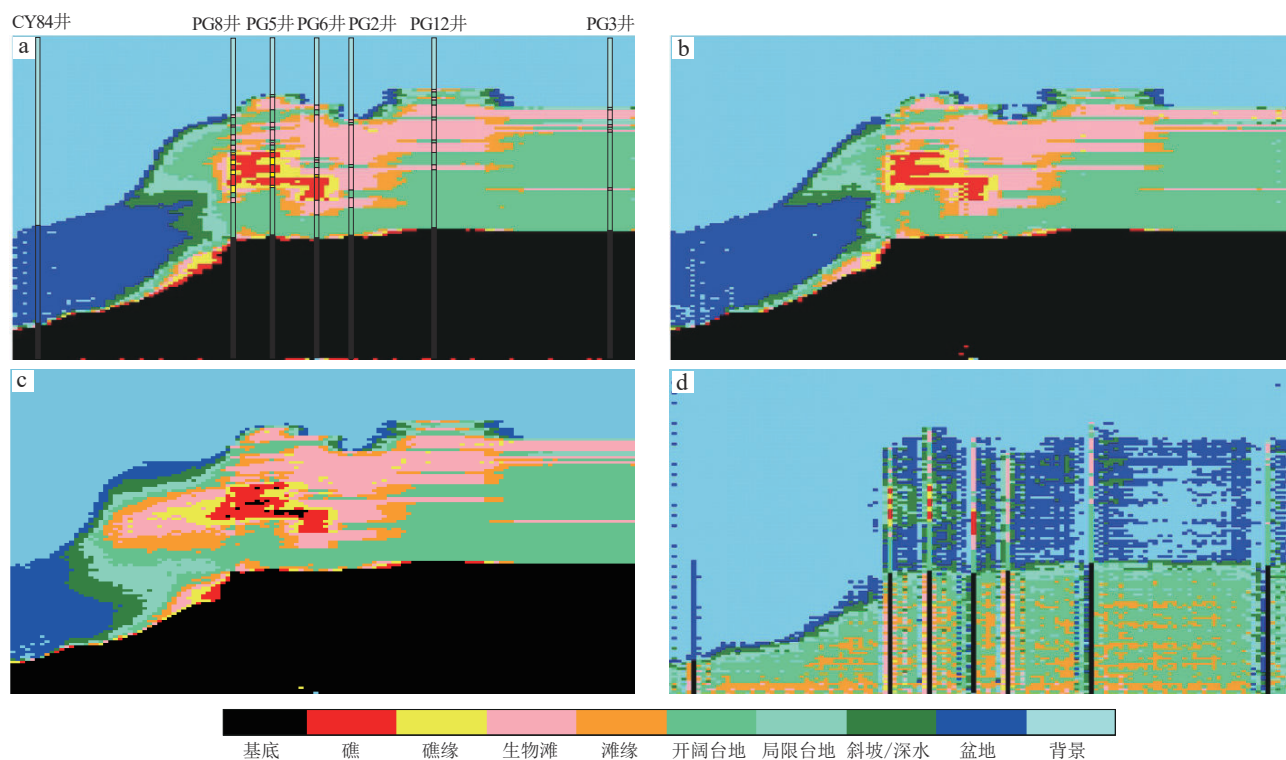


图8 训练样本量不同时生成网络模拟结果

Fig. 8 Results of generative network modeling with training datasets of different scales

a. 训练样本占总样本 80 % 时的模拟结果; b. 训练样本占总样本 60 % 时的模拟结果; c. 训练样本占总样本 40 % 时的模拟结果; d. 训练样本占总样本 20 % 时的模拟结果

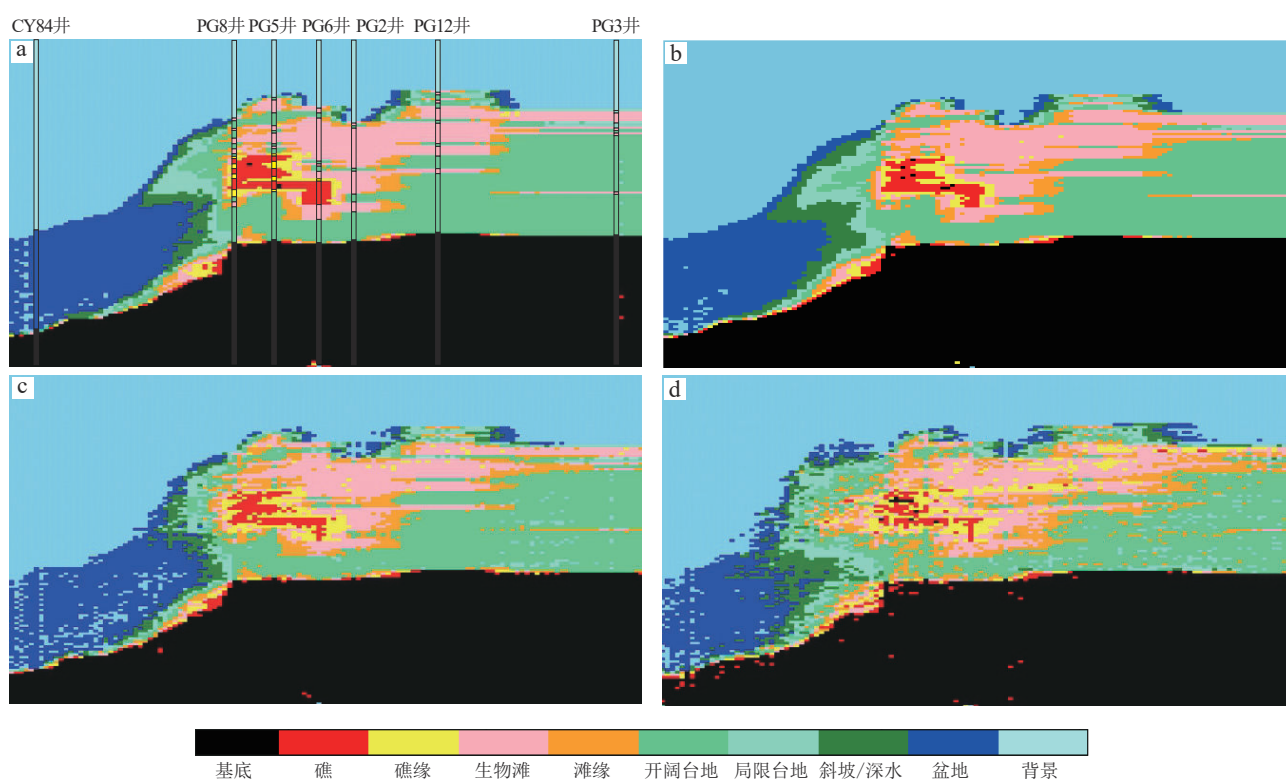


图9 损失函数L1权重不同时生成网络模拟结果

Fig. 9 Results of generative network modeling with different weights of L1 loss

a. L1 损失权重为 100 时的模拟结果; b. L1 损失权重为 50 时的模拟结果; c. L1 损失权重为 10 时的模拟结果; d. L1 损失权重为 1 时的模拟结果

现象。需要说明的是,图8d中没有叠加显示井数据,井所在位置显示的是模拟结果,可以看出即使在神经网络过拟合,井间预测不合理的情况下,井点处也能保证和条件数据吻合。

面向地质建模的条件化生成对抗网络中总的损失函数包含判别损失和L1损失。添加L1损失的目的是增加模型的分辨率,图9中4个子图分别是L1损失权重取值为100,50,10和1时的模拟结果,可以看出降低L1损失的权重会降低生成模型的分辨率,导致预测结果中散点增多。但是,过度增加L1损失的权重会导致判别器对模型的真假识别能力降低,权重值的确定需要根据实际的样本量和地质模式的复杂度,通过反复测试后确定。

3 讨论

深度学习地质建模有广阔的应用前景。目前阶段面临的主要问题是样本需求量大而实际地质模型样本不足,地层沉积过程模拟技术可以有效缓解这一难题,但不可忽视采用沉积模拟方法生成大规模样本时存在的不确定性。

首先应明确选用的沉积正演模拟器是否能够刻画目标区块的主要沉积过程,错误的或过度简化的正演模拟器不可能给出符合沉积规律的地质模型作为样本。其次是输入数据的不确定性,假设两个极端情况,即所有参数均不确定和所有参数均确定。前者将产生无限量样本的训练集,后者只产生一个模型用来训练。但实际情况是总能通过有限的观测数据和区域地质分析获得一些比较确定性的认识,降低参数的不确定性。但是,我们不可能获得足够数据以确定全部参数,因此需要设置一定的参数扰动区间,基本原则是在计算量允许的范围内越大越好,以降低人为影响。还有一部分不确定性由测井数据的人工解释引起,这部分反映了地质认识的综合不确定性,属于可接受的范围。

基于对上述3种不确定性的理解,可以进一步分析该方法的使用范围。基于沉积模拟的深度学习地质建模方法依赖两个基本假设:①沉积正演模拟采用的数学物理方程可以比较准确地刻画地层沉积形成的基本过程;②沉积正演模拟输入参数的不确定性在合理区间内。适合用该方法建立合理地质模型的区块需要满足这两个基本假设。很明显,地质过程正演模拟方法并不是要精细刻画每一个地质过程,而是模拟主要地质特征,忽略不必要的细节。因为一般的地质体都

经历了百万年甚至上亿年的演化,历经非常复杂的物理、化学作用,正演模拟器只要能刻画预期的主要特征,就满足了第一个基本假设。在充分理解所选沉积正演模拟器的基础上,结合地质分析,获取输入参数的可能取值范围。有些参数不确定性较低,有些参数不确定性较高,需要利用钻井数据和地震数据进行标定;在参数不确定性分析的基础上构建针对目标区块的训练样本库。

深度学习在各个行业落地应用的过程中,与专业领域知识的结合是必然要求。数据密集场景率先进行了智能化升级,比如测井相识别、地震数据解释、产量递减曲线分析等。地质领域传统上以概念和定性的知识为主,量化的数据较少,智能化的发展相对缓慢。但是智能化离不开量化,为了使地质知识更好地服务于深层油气藏勘探开发的实践,有必要对沉积过程、成岩过程、溶蚀过程、构造过程及地应力过程等地质过程进行量化模拟。同时探索小样本和基于符号的机器学习方法也是发展的重要方向,可以直接把地质领域的知识融入到智能化地质建模之中。

4 结论

1) 智能化地质建模是地质建模技术的重要发展方向,训练样本获取难度大是该领域面临的主要问题之一。沉积过程模拟与深度学习建模的结合可以实现优势互补,解决沉积正演模拟与观测数据吻合难度大和深度学习地质建模缺少样本的问题。

2) 通过在普光气藏主力区块典型地质剖面上的测试应用,该方法的可靠性得到了验证,构建了10 000多个地质模型组成的样本库,并以样本上抽取的虚拟井数据作为条件输入,以地质模型作为输出,进行网络训练和测试。利用训练后的生成网络,输入7口实际的观测井数据作为条件数据,建立了同时满足井数据和碳酸盐岩沉积成因模式的地质模型。

3) 深度学习在油气藏地质领域的落地应用需要充分利用专业领域知识,有必要加强沉积模拟、成岩模拟、溶蚀模拟、构造演化以及应力模拟等地质知识的量化研究,促进地质知识直接驱动的智能地质建模。

参 考 文 献

- [1] 吴胜和,李宇鹏. 储层地质建模的现状与展望[J]. 海相油气地质, 2007, 12(3): 53-60.
WU Shenghe, LI Yupeng. Reservoir modeling: Current situation and development prospect[J]. Marine Origin Petroleum Geology,

- 2007, 12(3): 53-60.
- [2] 张文彪, 段太忠, 刘彦锋, 等. 定量地质建模技术应用现状与发展趋势[J]. 地质科技情报, 2019, 38(3): 264-275.
ZHANG Wenbiao, DUAN Taizhong, LIU Yanfeng, et al. Application status and development trend of quantitative geological modeling [J]. Geological Science and Technology Information, 2019, 38(3): 264-275.
- [3] 尹艳树, 张昌民, 李玖勇, 等. 多点地质统计学研究进展与展望[J]. 古地理学报, 2011, 13(2): 245-252.
YIN Yanshu, ZHANG Changmin, LI Jiuyong, et al. Progress and prospect of multiple-point geostatistics [J]. Journal of Palaeogeography, 2011, 13(2): 245-252.
- [4] 吴胜和, 李文克. 多点地质统计学——理论、应用与展望[J]. 古地理学报, 2005, 7(1): 137-144.
WU Shenghe, LI Wenke. Multiple-point geostatistics: Theory, application and perspective [J]. Journal of Palaeogeography, 2005, 7(1): 137-144.
- [5] 尹艳树, 吴胜和. 储层随机建模研究进展[J]. 天然气地球科学, 2006, 17(2): 210-216.
YIN Yanshu, WU Shenghe. The progress of reservoir stochastic modeling[J]. Natural Gas Geoscience, 2006, 17(2): 210-216.
- [6] 胡向阳, 郑文波, 游瑜春, 等. 四川盆地元坝长兴组礁滩相气藏概率体约束地质建模[J]. 石油与天然气地质, 2020, 41(1): 157-163.
HU Xiangyang, ZHENG Wenbo, YOU Yuchun, et al. Probability body-constrained geomodeling of reef-shoal reservoir in Changxing Formation, Yuanba area, Sichuan Basin[J]. Oil & Gas Geology, 2020, 41(1): 157-163.
- [7] 张文彪, 段太忠, 郑磊, 等. 基于浅层地震的三维训练图像获取及应用[J]. 石油与天然气地质, 2015, 36(6): 1030-1037.
ZHANG Wenbiao, DUAN Taizhong, ZHENG Lei, et al. Generation and application of three-dimensional MPS training images based on shallow seismic data [J]. Oil & Gas Geology, 2015, 36(6): 1030-1037.
- [8] 宋随宏, 史燕青, 侯加根. 基于生成对抗网络的储层地质建模方法研究进展[J]. 石油科学通报, 2022, 7(1): 34-49.
SONG Suihong, SHI Yanqing, HOU Jiagen. Review of a generative adversarial networks (GANs) -based geomodelling method[J]. Petroleum Science Bulletin, 2022, 7(1): 34-49.
- [9] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]//Ghahramani Z, Welling M, Cortes C, eds. Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, 2014. Cambridge: MIT press, 2014: 2672-2680.
- [10] 刘彦锋, 张文彪, 段太忠, 等. 深度学习油气藏地质建模研究进展[J]. 地质科技通报, 2021, 40(4): 235-241.
LIU Yanfeng, ZHANG Wenbiao, DUAN Taizhong, et al. Progress of deep learning in oil and gas reservoirgeological modeling [J]. Bulletin of Geological Science and Technology, 2021, 40(4): 235-241.
- [11] ZHENG Qiang, ZENG Lingzao, KARNIADAKIS G E. Physics-informed semantic inpainting: Application to geostatistical modeling [J]. Journal of Computational Physics, 2020, 419: 109676.
- [12] MOSSER L, DUBRULE O, BLUNT M J. Conditioning of three-dimensional generative adversarial networks for pore and reservoir-scale models [EB/OL]. (2018-02-15) [2022-01-03]. <https://arxiv.org/abs/1802.05622>.
- [13] SONG Suihong, MUKERJI T, HOU Jiagen. Geological facies modeling based on progressive growing of generative adversarial networks (GANs) [J]. Computational Geosciences, 2021, 25(3): 1251-1273.
- [14] SONG Suihong, MUKERJI T, HOU Jiagen. GANSim: conditional facies simulation using an improved progressive growing of generative adversarial networks (GANs) [J]. Mathematical Geosciences, 2021, 53(7): 1413-1444.
- [15] NESVOLD E, MUKERJI T. Geomodeling using generative adversarial networks and a database of satellite imagery of modern river deltas [C]//Petroleum Geostatistics 2019, Florence, 2019. Bunnik: European Association of Geoscientists & Engineers, 2019: 1-5.
- [16] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations [J]. Journal of Computational Physics, 2019, 378: 686-707.
- [17] KARNIADAKIS G E, KEVREKIDIS I G, LU Lu, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3(6): 422-440.
- [18] SLOSS L L. Stratigraphic models in exploration [J]. AAPG Bulletin, 1962, 46(7): 1050-1057.
- [19] HARBAUGH J W. Mathematical simulation of marine sedimentation with IBM 7090/7094 computers[M]//Merriam D F. Computer Contributions 1. Lawrence: Kansas Geological Survey, 1966: 1-52.
- [20] PAOLA C. Quantitative models of sedimentary basin filling [J]. Sedimentology, 2000, 47(S1): 121-178.
- [21] BURGESS P M. 14-A brief review of developments in stratigraphic forward modelling, 2000-2009 [M]//Roberts D G, Bally A W. Regional Geology and Tectonics: Principles of Geologic Analysis. Boston: Elsevier Science, 2012: 378-404.
- [22] HUANG X, GRIFFITHS C M, LIU J. Recent development in stratigraphic forward modelling and its application in petroleum exploration [J]. Australian Journal of Earth Sciences, 2015, 62(8): 903-919.
- [23] ROSLI R, POPPELREITER M C, JAMALUDIN S N F. A review of stratigraphic forward models (Sfm) for carbonate platform [J]. International Journal of Engineering & Technology, 2018, 7(4.35): 143-147.
- [24] DUAN Taizhong. Similarity measure of sedimentary successions

- and its application in inverse stratigraphic modeling [J]. *Petroleum Science*, 2017, 14(3): 484–492.
- [25] 刘彦锋, 段太忠, 龚伟, 等. 基于深度学习的沉积模拟代理模型构建与应用[J/OL]. *沉积学报*: 1–16[2022-01-03]. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=CJXB20211217002&DbName=CAPJ2021>. DOI: 10.14027/j.issn.1000-0550.2021.152.
- LIU Yanfeng, DUAN Taizhong, GONG Wei, et al. Construction and application of a proxy model for stratigraphic forward modeling based on deep learning [J/OL]. *Acta Sedimentologica Sinica*: 1–16[2022-01-03]. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=CJXB20211217002&DbName=CAPJ2021>. DOI: 10.14027/j.issn.1000-0550.2021.152.
- [26] GOODFELLOW I, BENGIO Y, COURVILLE A. *Deep learning* [M]. Cambridge: MIT Press, 2016.
- [27] HARSHVARDHAN G M, GOURISARIA M K, PANDEY M, et al. A comprehensive survey and analysis of generative models in machine learning [J]. *Computer Science Review*, 2020, 38: 100285.
- [28] ZHANG Tuanfeng, TILKE P, DUPONT E, et al. Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks[J]. *Petroleum Science*, 2019, 16(3): 541–549.
- [29] CHAN S, ELSHEIKH A H. Parametric generation of conditional geological realizations using generative neural networks [J]. *Computational Geosciences*, 2019, 23(5): 925–952.
- [30] GAO Xiaoyang, HE Wenxiang, HU Yong. Modeling of meandering river deltas based on the conditional generative adversarial network [J]. *Journal of Petroleum Science and Engineering*, 2020, 193: 107352.
- [31] MIRZA M, OSINDERO S. Conditional generative adversarial nets [EB/OL]. (2014-11-06)[2022-01-03]. <https://arxiv.org/abs/1411.1784>.
- [32] HUANG Huimin, LIN Lanfen, TONG Ruofeng, et al. UNet 3+: A full-scale connected UNet for medical image segmentation[C]// *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, 2020. Piscataway: IEEE, 2020: 1055–1059.
- [33] POMAR L. Ecological control of sedimentary accommodation: Evolution from a carbonate ramp to rimmed shelf, Upper Miocene, Balearic Islands [J]. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 2001, 175(1/4): 249–272.
- [34] POMAR L, HAQ B U. Decoding depositional sequences in carbonate systems: Concepts vs experience [J]. *Global and Planetary Change*, 2016, 146: 190–225.
- [35] POMAR L. Chapter 12-Carbonate systems [M]//Scarselli N, Adam J, Chiarella D, et al. *Regional Geology and Tectonics: Volume 1: Principles of Geologic Analysis*. 2nd ed. Amsterdam: Elsevier, 2020: 235–311.
- [36] 段太忠, 王光付, 廉培庆, 等. 油气藏定量地质建模方法与应用[M]. 北京: 石油工业出版社, 2019.
- DUAN Taizhong, WANG Guangfu, LIAN Peiqing, et al. Quantitative geological modeling method and application of oil and gas reservoir[M]. Beijing: Petroleum Industry Press, 2019.
- [37] 马永生, 蔡勋育. 四川盆地川东北二叠系–三叠系天然气勘探成果与前景展望[J]. *石油与天然气地质*, 2006, 27(6): 741–750.
- MA Yongsheng, CAI Xunyu. Exploration achievements and prospects of the Permian-Triassic natural gas in northeastern Sichuan Basin[J]. *Oil & Gas Geology*, 2006, 27(6): 741–750.
- [38] 姜贻伟, 刘红磊, 杨福涛, 等. 震控储层建模方法及其在普光气田的应用[J]. *天然气工业*, 2011, 31(3): 14–17, 106.
- JIANG Yiwei, LIU Honglei, YANG Futao, et al. Seismic-constrained reservoir modeling and its application in the Puguang gas field[J]. *Natural Gas Industry*, 2011, 31(3): 14–17, 106.
- [39] 马永生, 储昭宏. 普光气田台地建造过程及其礁滩储层高精度层序地层学研究[J]. *石油与天然气地质*, 2008, 29(5): 548–556.
- MA Yongsheng, CHU Zhaozhong. Building-up process of carbonate platform and high-resolution sequence stratigraphy of reservoirs of reef and oolitic shoal facies in Puguang gas field[J]. *Oil & Gas Geology*, 2008, 29(5): 548–556.
- [40] 马永生, 蔡勋育, 郭旭升, 等. 普光气田的发现[J]. *中国工程科学*, 2010, 12(10): 14–23.
- MA Yongsheng, CAI Xunyu, GUO Xusheng, et al. The discovery of Puguang gas field[J]. *Strategic Study of CAE*, 2010, 12(10): 14–23.

(编辑 张 晟)