

DOI: 10.16285/j.rsm.2024.0500

深圳土工参数数据库及基于生成对抗网络的多元参数分布预测模型研究

潘秋景¹, 孙广灿¹, 蔡永敏², 苏 栋³, 李凤伟⁴

(1. 中南大学 土木工程学院, 湖南 长沙 410075; 2. 新加坡科技设计大学 建筑学与可持续设计系, 新加坡 487372;
3. 深圳大学 土木与交通工程学院, 广东 深圳 518060; 4. 中国铁建股份有限公司, 北京 100855)

摘 要: 借鉴大数据思想, 充分利用岩土工程勘察数据, 实现岩土参数精细化表征和建模, 是岩土工程数字孪生的重要组成部分。通过收集深圳市 75 个工程项目的岩土工程勘察报告, 建立了深圳黏性土及风化残积土 8 个土工试验参数数据库 SZ-SOIL/8/11369, 分析了深圳黏性土及风化残积土土工试验参数的分布特征和规律。进一步利用该数据库, 提出了基于生成对抗网络 (generative adversarial network, 简称 GAN) 的土工试验物理力学参数概率分布及预测模型, 并将提出的方法应用于深圳某项目, 针对单组土工试验样本利用物理参数成功预测了其力学参数, 并利用少量样本正确预测了该工程场地的土工试验参数的分布。结果表明, 所提方法能够对缺失参数样本进行合理预测, 并实现了通过大范围地区勘察数据降低局部工程场地岩土参数不确定性的目的, 可为深圳岩土与地下工程结构韧性设计和风险评价提供参数保障。

关 键 词: 土工参数分布; 数据库; 预测; 生成对抗网络

中图分类号: TU 413

文献标识码: A

文章编号: 1000-7598 (2025) 02-563-10

Shenzhen geotechnical parameter database and multivariate parameter distribution prediction model based on generative adversarial network

PAN Qiu-jing¹, SUN Guang-can¹, CAI Yong-min², SU Dong³, LI Feng-wei⁴

(1. School of Civil Engineering, Central South University, Changsha, Hunan 410075, China;

2. Department of Architecture and Sustainable Design, Singapore University of Technology and Design, Singapore 487372;

3. College of Civil and Transportation Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China;

4. China Railway Construction Corporation Limited, Beijing 100855, China)

Abstract: Inspired by big data, fully utilizing geotechnical data for precise characterization and modeling of geotechnical parameters is critical for the digitalization of geotechnical engineering. This study collected geotechnical investigation reports from 75 engineering projects in Shenzhen, established a database containing 8 geotechnical parameters of clay and weathered residual soil (SZ-SOIL/8/11369), and thoroughly analyzed the distribution characteristics of geotechnical parameter data in Shenzhen. Subsequently, a model for predicting geotechnical parameters was developed using this database and a generative adversarial network (GAN). The proposed method was applied to a project in Shenzhen, successfully predicting mechanical parameters from known physical parameters and accurately forecasting the geotechnical parameter distribution of the project site using small samples. The results indicate that the proposed method can make reasonable predictions for samples with missing parameters, achieving the goal of reducing the uncertainty in geotechnical parameters at local engineering sites through extensive regional survey data. This provides parameter assurance for the resilience design and risk assessment of geotechnical and underground engineering structures in Shenzhen.

Keywords: distribution of geotechnical parameters; database; prediction; generative adversarial network

收稿日期: 2024-04-22

录用日期: 2024-07-26

基金项目: 国家自然科学基金 (No. 52108388, No. 52378424); 国家重点研发计划 (No. 2023YFC3009300); 湖南省科技创新计划 (No. 2021RC3015); 深圳大学 2035 追求卓越研究计划 (No. 2022B007)。

This work was supported by the National Natural Science Foundation of China (52108388, 52378424), the National Key Research and Development Program of China (2023YFC3009300), the Science and Technology Innovation Program of Hunan Province (2021RC3015) and the Shenzhen University 2035 Pursuit of Excellence Research Program (2022B007).

第一作者简介: 潘秋景, 男, 1987 年生, 博士, 教授, 博士生导师, 主要从事盾构隧道掘进力学与智能决策等方面的研究。E-mail: qiuqing.pan@csu.edu.cn

通讯作者: 蔡永敏, 男, 1992 年生, 博士, 博士后, 主要从事岩土工程风险分析及数字岩土工程等方面的科研工作。E-mail: yongmin_cai@sutd.edu.sg

1 引言

中华人民共和国住房和城乡建设部在发布的《2016—2020 年建筑业信息化发展纲要》文件中,明确提出我国要全面提高建筑业的信息化水平^[1]。利用岩土工程勘察数据实现地层和岩土参数的精细化表征和建模,被视为实现岩土与地下工程全生命周期数字化与信息化不可或缺的组成部分^[2-4]。此外,岩土工程风险管理与可靠度设计是岩土与地下工程结构韧性评价和控制的核心要素^[3],其关键环节之一是利用现场勘察数据合理推断岩土参数的概率分布,其难点在于对于具体工程项目而言,现场勘察数据具有多维、稀疏、缺失、高变异性等特征。因此,如何高效地利用有限的土工试验数据来确定岩土参数的概率分布,以及如何对缺失的数据进行准确预测,对岩土工程的数字化和韧性设计至关重要,这些问题引起了国内外众多学者的广泛关注^[5-7]。

建立岩土物理力学参数的概率分布模型是数字岩土工程的研究热点问题,国内外已涌现出许多相关研究成果。在国内,一些学者借助不同地区的工程现场数据,采用 Kolmogorov-Smirnov 检验(K-S 检验)和贝叶斯更新等方法,对南宁冲积相黏土物理力学参数^[8]、上海软土抗剪强度参数^[9]、小浪底和关中灌区等大坝堆石料抗剪强度参数^[10]、重庆万州泥岩和砂岩强度参数^[11]、天津软土土性指标^[12]是否服从正态分布进行了深入分析和研究。在国外,Baecher 等^[13]认为岩石节理长度和节理间距分别服从对数正态分布和指数分布。然而,这些工作受到数据量或分析方法的限制,仅分析单个或极少数几个岩土参数的概率分布模型,而未对多个岩土物理力学参数之间的相关性进行充分分析,也未建立岩土参数的联合概率分布模型。这种忽视多元岩土参数之间相互关联性的方法不仅会导致信息浪费,还会限制对于参数的进一步分析和利用。Ching 等^[14-15]构建了岩石和黏性土物理力学参数的多元概率分布模型,采用 Johnson 分布来标准化各参数的边际分布,然后通过多元正态分布将这些标准化的分布耦合在一起。该方法充分利用了多元参数之间的关联性,并利用多元正态分布存在共轭分布的特性简化了预测过程。然而,以上的工作囿于固化的分布形式,对于分布较为复杂的数据的拟合和预测效果欠佳。

对于缺失参数预测问题,许多学者针对不同地区或者项目的区域性土工参数数据集,采用统计的方

法提出了相关的经验公式。典型的有渤海海洋土^[16]、杭州湾浅层软黏土^[17]、南京地铁地基黏土^[18]与武汉地铁黏土^[19]等的相关公式。但是现有的经验性成果都基于一些数学假设,并且研究对象具有鲜明的地区特征,因此泛化能力有待提升,且无法量化未知信息的不确定性。

以上的传统概率统计方法拟合能力有限,无法充分利用多维互相关信息建立准确的岩土参数分布模型。21 世纪以来,机器学习技术的发展带来了高性能的拟合方法。机器学习技术可以更好地挖掘自变量指标与被预测指标间复杂的非线性映射关系,从而建立精细化的多元参数分布模型。例如早期的研究利用 MATLAB 的 Box-Cox 方法研究土工参数之间的多元相关关系^[20-21]。更新的研究则是将人工神经网络与预测评估土体参数的工作结合起来^[22-23]。以上研究都证明,基于机器学习的方法能够弥补传统概率统计方法拟合能力的缺陷,得到的结果相比过去总结的各类回归公式拟合预测效果更佳。

然而,即使一般的机器学习方法解决了拟合能力的问题,其仍然存在一些实际应用方面的局限性。一方面,基于特定区域数据训练得到的机器学习模型缺乏泛化能力;另一方面,前文所述方法无法得到参数的概率分布模型,只能输出单一预测值,应用意义较低。

基于以上背景,本文建立了深圳土工参数数据库,并提出了基于生成对抗网络的岩土参数分布及预测模型。本文建立的深圳土工参数数据库包含 8 个常见的土工参数,提出的模型充分利用多元参数之间的相互关联性同时对这 8 个参数分布进行建模。针对模型泛化能力的问题,提出的模型融合大区域参数信息与局部小区域参数信息以保证模型的泛化能力。针对概率分布建模方面的局限性,则参考矿业^[24]、工业传感^[25]领域的相关研究,利用生成对抗网络(generative adversarial network,简称 GAN)的生成式特性,以样本分布的形式输出土工参数的概率分布。结合本文所建立的深圳土工参数数据库和预测模型可以降低局部工程场地岩土参数不确定性。

2 深圳土工参数数据库 SZ-SOIL/8/11369

本文选用深圳市 75 个工程项目的岩土勘察报告,整理并建立了 8 个土工参数的数据库。这 8 个土工参数分别为:① 4 个物理参数,孔隙比 e 、饱和度 S_r (%)、液限 ω_L (%)、塑限 ω_p (%);② 4 个力学参数,快剪黏聚力 c_q (kPa)、快剪摩擦角 φ_q (°)、压缩模量 E_s (MPa)、压缩系数 α (MPa⁻¹)。

数据库中的 75 个工程项目广泛分布于深圳市的 10 个区，总共包含 11 369 条数据。这 11 369 条数据按区位分布统计如图 1 所示。75 个工程项目的具体分布情况见表 1，南山区的数据量占比达到了 47.5%，近海区域的数据占比为 61%。

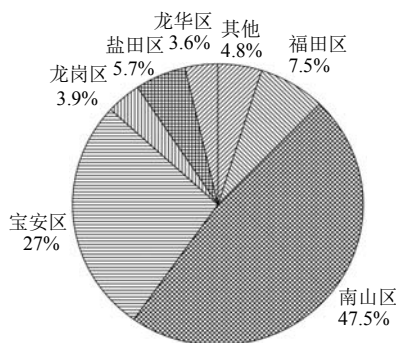


图 1 数据区位分布统计饼图

Fig.1 Pie chart of data location distribution

表 1 数据库中 75 个工程项目项目分布
Table 1 Distribution of 75 projects in the database

行政区划	项目数量 /个
福田区	5
南山区	37
宝安区	14
龙岗区	4
盐田区	4
龙华区	3
其他	8

2.1 类别差异性分析

数据库中的 11 369 条数据中有 5 285 条完整数据，主要有两大类，包括黏性土数据 3 700 条、风化残积土数据 951 条；除了这两大类，剩下的 634 条为砂、碎石、杂填土等其他种类。黏性土和风化残积土的详细组成见图 2。由图可以看到，每类土包含了许多子类，黏性土类包括一般黏土、粉质黏土、砂质黏土和砾质黏土等，风化残积土类包括强风化及全风化花岗岩和粉砂岩、强风化碎裂岩等。黏性土和风化残积土的统计规律见表 2。表中平均变异系数最大的为快剪黏聚力，最小的为饱和度。含水率、孔隙比和饱和度等物理性质参数的变异系数均小于黏聚力、摩擦角和压缩模量等力学参数。此外，数据库中各种土的饱和度均值都超过 80%，推测这是由于深圳市年平均降水量高达 1 863.9 mm^[26]，导致当地土的饱和度较高。黏性土的平均孔隙比较大，推测是因为深圳沿海地区黏性土多为海积土及

冲积土，孔隙比大^[27]。另外，黏性土的快剪黏聚力及快剪摩擦角的最大值和最小值差别非常大，如快剪黏聚力的最小值和最大值分别为 1.0 kPa 和 61.0 kPa，达到了 60 倍的差别，一方面黏性土子类之间性质差异大，另一方面获取数据的项目在深圳分布范围广，沿海区域与远海区域的岩土性质也有差异。

两类土的孔隙比 e 、塑限 ω_p 、压缩系数 α 与压缩模量 E_s 的边缘分布密度曲线见图 3。由图可以看到，两类土这 4 个土工参数的分布表现出明显的区别，黏性土的孔隙比、塑限以及压缩系数都相对较大，风化残积土则较小。

2.2 区域差异性分析

表 3 展示了数据库中深圳市不同区土工参数的均值，揭示了土工参数的区域性差异。南山区和宝安区的项目多位于近海区域，其孔隙比、饱和度和液限、塑限等物理参数普遍高于其他区域。力学参数的区域差异则不明显。与整体数据库的均值相比，南山区快剪摩擦角与整体数据的差距最小，仅为 0.4%，且南山区各参数与整体数据的平均差距也最小，仅为 6%。这主要是因为南山区数据在土工

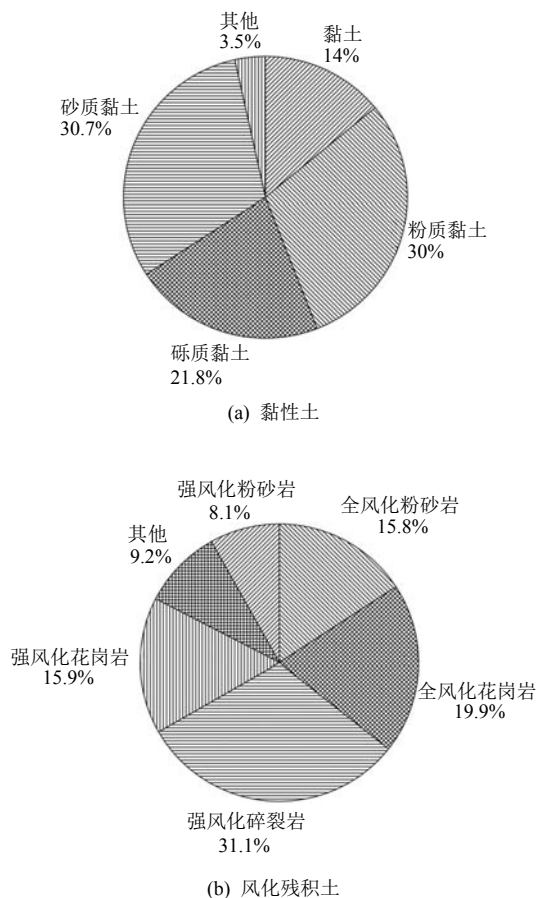


图 2 SZ-SOIL/8/11369 中两类土的组成图

Fig.2 Detailed composition diagram of two types of soil in SZ-SOIL/8/11369

表 2 SZ-SOIL/8/11369 数据的统计结果
Table 2 Statistics of SZ-SOIL/8/11369

土的种类	统计量	孔隙比 e	饱和度 S_r / %	液限 ω_L / %	塑限 ω_p / %	快剪黏聚力 c_q / kPa	快剪摩擦角 φ_q / (°)	压缩模量 E_s / MPa	压缩系数 α / MPa ⁻¹
黏性土	最大值	1.55	112.80	66.60	40.00	61.00	48.40	11.30	0.95
	最小值	0.44	64.10	21.80	14.80	1.00	0.10	1.05	0.13
	平均值	0.99	93.47	44.55	27.40	23.94	14.44	4.62	0.46
	变异系数/ %	19.82	6.44	9.41	10.14	45.96	29.85	32.72	39.38
风化残积土	最大值	1.44	103.20	62.90	38.30	66.20	34.40	14.20	0.84
	最小值	0.34	48.70	20.40	13.70	3.50	3.70	2.44	0.12
	平均值	0.75	89.79	35.52	22.61	24.53	22.52	7.25	0.28
	变异系数/ %	23.74	9.92	15.78	19.77	46.04	21.79	31.37	39.17

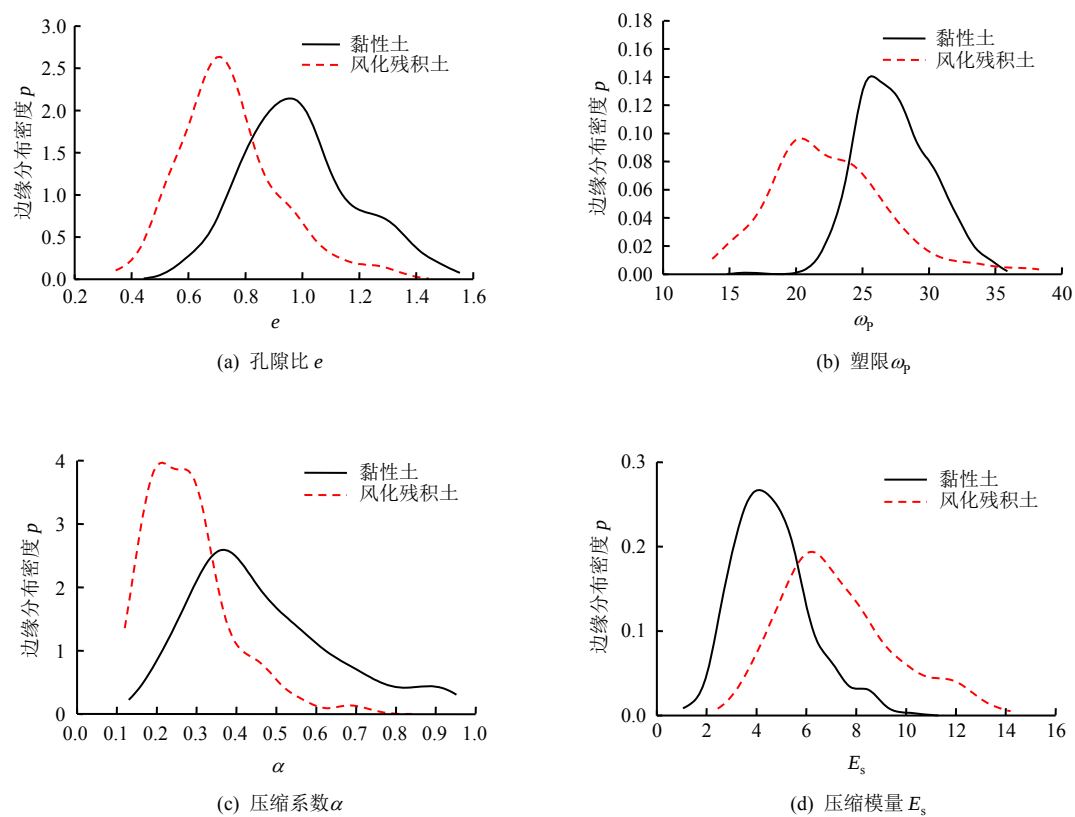


图 3 两类土孔隙比 e 、塑限 ω_p 、压缩系数 α 与压缩模量 E_s 的边缘分布密度曲线
Fig.3 Marginal distribution density curves of pore ratio e , plastic limit ω_p , compression coefficient α , and compression modulus E_s for two types of soil

表 3 SZ-SOIL/8/11369 中深圳各区参数均值
Table 3 Average parameter values of each district in Shenzhen in SZ-SOIL/8/11369

行政区划	孔隙比 e	饱和度 S_r / %	液限 ω_L / %	塑限 ω_p / %	快剪黏聚力 c_q / kPa	快剪摩擦角 φ_q / (°)	压缩模量 E_s / MPa	压缩系数 α / MPa ⁻¹
南山区	1.00	100.58	47.29	28.65	24.71	14.38	4.05	0.52
宝安区	0.98	89.69	43.82	26.95	24.11	16.73	6.02	0.35
龙岗区	0.95	87.38	41.33	25.32	29.18	8.00	4.60	0.38
盐田区	0.99	79.39	37.33	24.47	17.82	11.46	3.26	0.53
龙华区	0.98	80.62	36.00	23.67	21.62	4.66	3.49	0.50
福田区	0.97	83.82	40.05	25.12	20.66	18.52	5.16	0.38
其他	1.02	89.00	45.26	28.46	24.93	11.46	3.91	0.46

参数数据库中占比最大, 对整体统计参数的影响更为显著。

但这并不意味着本数据库无法反映其他区的土工参数分布。即使是完全内陆的龙华区和龙岗区, 其与整体数据库的参数平均数值差异也仅在 15% 左右, 差距并不明显。因此, 整理的深圳土工参数数据库仍能较为全面地反映深圳市各区的土工参数分布情况。

3 基于GAN的深圳土工参数分布预测模型

3.1 生成对抗网络

GAN 是由 Goodfellow 等^[28]于 2014 年提出的生成式机器学习模型。GAN 是一种隐式生成模型, 其独特之处在于不需要显式地假设生成数据的密度函数, 而是学习数据样本的生成过程。它通过生成器与判别器之间的博弈, 不断提升双方的性能, 进而学习数据样本的分布信息。GAN 在岩土与地质工程领域有许多应用, 如重建页岩图像^[29]、盾构掘进地层预测^[30]、生成地质模型^[31]等。

3.2 基于 GAN 的土工参数分布预测模型

基于 GAN 的土工参数分布预测模型的大致流程如下: 首先使用土工数据库训练 GAN 模型, 利用 GAN 的生成器拟合土工参数的概率分布, 并基于 GAN 反演技术进行参数预测。对于具体项目, 进一步使用项目数据对 GAN 模型进行微调, 从而得到针对具体项目的土工参数模型以用于后续的场地参数预测等问题。

3.2.1 训练

与原始的 GAN 对抗训练流程相同^[28], 在模型训练的每个循环中, 生成器接收一个 0~1 之间的均匀分布随机向量 $\mathbf{z}$, 并依此生成一个土工参数数据样本 $\mathbf{x}_g$ 。判别器则分别对真实的数据和生成器生成的数据进行评估, 输出它们为真实数据的概率。随后, 依据 GAN 的目标函数对这两个网络进行训练。训练完成后, 生成器将包含有关数据分布的信息, 进而可以用于解决土工参数的预测问题。

3.2.2 单组数据样本缺失值的预测

对部分缺失的样本 $\mathbf{x}_r$ 进行补全, 利用样本已知的部分参数 $\mathbf{x}_{r,k}$ 预测未知参数的值 $\mathbf{x}_{r,u}$, 从而提高数据完整性, 对支持工程决策、降低工程风险具有重要意义。作为一种生成式模型, GAN 生成器可以使用存储的参数分布信息生成符合数据分布的样本, 因此, 可以考虑将对参数样本 $\mathbf{x}_r$ 中缺失值 $\mathbf{x}_{r,u}$ 的预

测转化为使用生成器生成一个与 $\mathbf{x}_r$ 接近的样本 $\mathbf{x}_g$, 并使用该生成样本 $\mathbf{x}_g$ 中的对应部分 $\mathbf{x}_{g,u}$ 作为 $\mathbf{x}_{r,u}$ 的预测值。相似样本 $\mathbf{x}_g$ 的生成可以通过 GAN 反演技术实现。通过反演技术对生成器的输入 $\mathbf{z}$ 进行调整以得到与 $\mathbf{x}_r$ 接近的样本 $\mathbf{x}_g$, 其中 $\mathbf{x}_r$ 与 $\mathbf{x}_g$ 是否接近通过二者的已知部分 $\mathbf{x}_{r,k}$ 与 $\mathbf{x}_{g,k}$ 之间的差异来判断。调整 $\mathbf{z}$ 的优化目标函数如下:

$$\mathbf{z}^* = \arg \min_{\mathbf{z}} l(\mathbf{x}_{r,k}, \mathbf{x}_{g,k}) \quad (1)$$

式中: $\mathbf{x}_g = G(\mathbf{z}, \theta_g)$, θ_g 为第 3.2.1 节训练得到的生成器参数; $l(\cdot)$ 为差异度量函数。

综上所述, 模型具体的预测流程如下: 首先, 生成一个随机向量 $\mathbf{z}$, 将其传递给生成器以生成土工参数样本 $\mathbf{x}_g$ 。然后, 将生成的样本 $\mathbf{x}_g$ 中与已知参数对应的部分 $\mathbf{x}_{g,k}$ 与 $\mathbf{x}_{r,k}$ 进行比较, 将比较结果反向传播以更新生成器的输入 $\mathbf{z}$ 。这个过程循环进行, 直到 $\mathbf{x}_{g,k}$ 与已知值 $\mathbf{x}_{r,k}$ 的差异在限定的范围内, 例如控制二者之间的相对差距小于 5%。将生成器最后一次生成的参数数据中与未知参数对应的部分 $\mathbf{x}_{g,k}$ 作为预测值 $\mathbf{x}_{r,u}$ 。

3.2.3 基于小样本的局部场地土工参数概率分布

利用具体工程场地的少量土工试验数据得到该场地的岩土参数的概率分布模型, 对岩土工程的数字化和韧性设计至关重要^[3]。

微调是以预训练的模型为基础, 在特定任务上进行微调以提升性能的一种技术。本文使用 LST (ladder side-tuning) 技术使用具体项目数据对第 3.2.1 节得到的 GAN 生成器进行微调以得到具体工程项目岩土参数的分布^[32]。具体而言, 对于预训练得到的生成器 $G(\mathbf{z}, \theta_g)$, 在其基础上以连接辅助模型的方式添加少量的额外参数 θ_{LST} 以得到微调生成器模型 $G^*(\mathbf{z}, \theta_g, \theta_{LST})$ 。随后将具体工程项目的土工参数数据集 $\mathbf{X}^* = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_n^*\}$ 作为训练集, 采用第 3.2.1 节中的方法对微调生成器模型 $G^*(\mathbf{z}, \theta_g, \theta_{LST})$ 进行训练, 训练时固定预训练模型参数 θ_g 而只更新额外参数 θ_{LST} , 从而保证预训练的信息不丢失。得到的 $G^*(\mathbf{z}, \theta_g, \theta_{LST})$ 即可用于表征该工程项目的土工参数分布。

4 工程应用

4.1 深圳土工参数分布模拟

使用深圳土工数据库中的黏性土与风化残积土样本训练 GAN 模型, 得到的 GAN 模型生成器即为深圳黏性土与风化残积土参数分布模型。图 4 为

利用该模型生成的两类土的孔隙比 e 与快剪黏聚力 c_q 的分布结果。

从图 4 可以看到, 提出的 GAN 模型生成的土工数据分布与原数据分布十分接近, 准确地模拟了原始数据的分布。分别将本文方法与 Ching 方法^[15]

得到的分布与原数据分布进行 K-S 检验, Ching 方法^[15]得到的分布与原数据分布的 K-S 统计量为 0.22, 而本文方法得到的 K-S 统计量仅有 0.15, 说明本文提出方法得到的分布与原数据分布更接近, 进一步说明了本文方法的有效性。

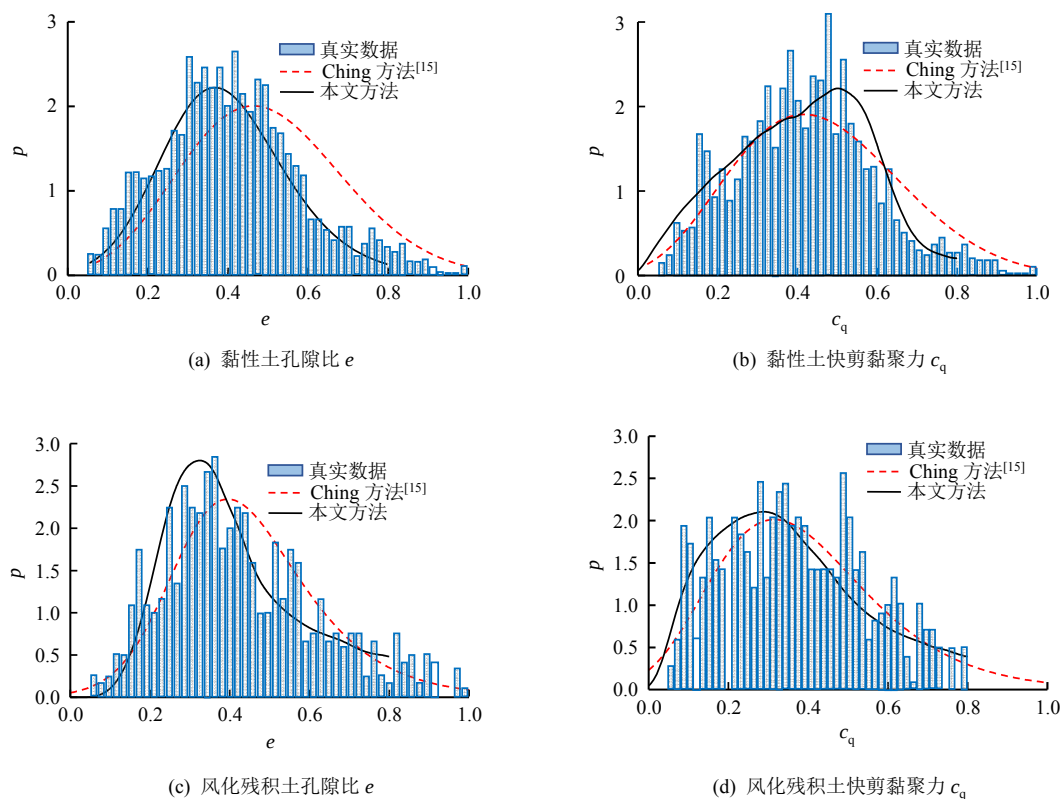


图 4 深圳黏性土与风化残积土参数数据分布直方图及 Ching 方法^[15]拟合的各参数边缘分布(点划线)与本文方法的各参数边缘分布(实线)

Fig.4 Histograms of parameters of clay and weathered residual soil, the marginal distributions of parameters fitted using the method proposed by Ching et al^[15] (dashed lines) and this paper (solid lines)

4.2 单组数据样本缺失值预测

深大城际铁路工程坪山站至聚龙站区间位于深圳市坪山区, 属于深圳土工数据库中土工参数数据量较稀疏的区域。该项目两个钻孔的土工试验数据见表 4。表 4 中待预测参数单元格有两行数据, 第 1 行是实际值, 第 2 行是使用提出的 GAN 模型预测的力学参数结果。在黏性土样本参数预测结果中, M33Z3-TBD-411-3 的快剪黏聚力误差最大, 达到了 43.5%, 而最小误差出现在该样本的压缩模量, 仅为 1.8%, 整体的预测平均准确率达到 83.3%。对于风化残积土, 最大误差出现于 M33Z3-TBD-151-2 的压缩系数, 达到 31.3%, 而最小误差则在 M33Z3-TBD-151-3 的快剪摩擦角, 为 11.7%, 预测的平均准确率为 81.9%。因此, 可以认为即使区域土工参数信息稀缺, 本文方法依旧有较高的可靠

度。

为了更直观地观察使用 GAN 的预测效果, 对样本 M33Z3-TBD-411-2 使用本文方法与 Ching 方法^[14]进行预测, 结果如图 5 所示, 图中等值线为深圳土工数据库的参数分布。从图中可以看到, 相比 Ching 等^[14]提出的传统概率统计方法, 本文方法预测结果更接近真实值。

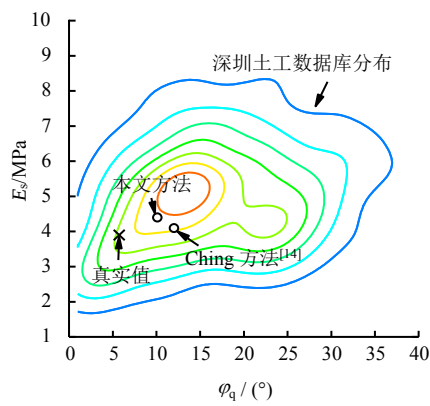
4.3 基于小样本的工程场地土工参数概率分布预测

深大城际铁路工程本段共有 198 条土工试验数据样本, 选取其中 16 条数据, 对深圳黏性土数据库土工参数 GAN 生成器模型进行微调。198 条土工试验数据样本的频率直方图作为该项目场地土工参数的真实概率分布。本文方法(即微调后的生成器)生成的土工参数分布曲线与深圳数据库的分布、该项目所有数据样本的岩土参数分布对比见图 6。

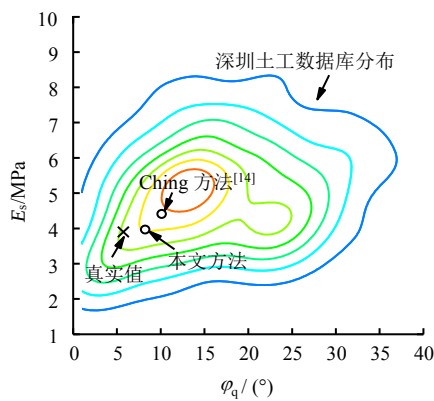
表 4 工程案例土工参数取值及预测值
Table 4 Values and predicted values of parameters in engineering case

钻孔编号	样本编号	孔隙比 e	饱和度 S_r /%	液限 ω_L /%	塑限 ω_p /%	快剪黏聚力 c_q /kPa	快剪摩擦角 φ_q /($^\circ$)	压缩模量 E_s /MPa	压缩系数 α /MPa $^{-1}$	土的种类
M33Z3-TBD-411	1	1.10	92.80	46.90	26.00	29.40	14.30	5.60 5.21	0.37 0.28	黏土
	2	1.22	92.30	53.70	29.00	17.20 20.53	7.80 8.13	5.30	0.42	黏土
	3	0.98	99.90	46.10	25.60	15.20	5.73 8.22	3.90 3.97	0.51	黏土
	4	0.81	85.40	40.70	24.70	— 31.52	— 8.93	— 7.20	— 0.26	砾质黏土
M33Z3-TBD-151	2	0.55	96.00	25.30	16.30	33.70	14.30	9.70 8.53	0.16 0.21	强风化粉砂岩
	3	0.66	71.40	29.90	18.00	54.40 48.03	21.50 25.27	7.20	0.23	强风化粉砂岩

注：表中“—”代表原始参数数据缺失。



(a) 已知 e 、 S_r 预测 φ_q 和 E_s



(b) 已知 e 、 S_r 、 ω_L 和 ω_p 预测 φ_q 和 E_s

图 5 本文方法和 Ching 方法^[14]预测 φ_q 和 E_s 结果对比
Fig.5 Comparison of prediction results of φ_q and E_s between this method and Ching method^[14]

从图 6 可以看到，本文方法利用小样本数据，结合深圳土工数据库，可以较好地拟合该场地土工数据的分布。将深圳土工数据库的样本分布、本文方法得到的分布分别与该场地 198 条土工试验数据

样本分布进行 K-S 检验，得到各参数的 K-S 统计量见表 5。结果表明，本文方法（即微调后的生成器）生成的所有参数样本分布的 K-S 统计量均小于深圳土工数据库分布的 K-S 统计量，其中差别最大的参数为孔隙比，相差约 4.7 倍，差别最小的为塑限，本文方法的平均 K-S 统计量为 0.09，深圳土工数据库分布的平均 K-S 统计量为 0.22，大于本文方法平均 K-S 统计量的 3 倍，这验证了微调后的分布更接近场地参数分布。

以上试验结果说明本文方法利用工程项目有限的勘察数据，借助大范围地区的勘察数据，可以降低局部工程场地岩土参数的不确定性，得到比较准确的场地土工参数的概率分布模型。

5 结 论

本文在 SZ-SOIL/8/11369 数据库的基础上，提出使用 GAN 拟合 8 个参数（孔隙比、饱和度、液限、塑限、快剪黏聚力、摩擦角、压缩模量和压缩系数）的多元概率分布，并利用 GAN 反演对未知参数进行预测，本文的主要结论如下：

(1) SZ-SOIL/8/11369 数据库收集深圳市 75 个工程项目勘察报告，包括孔隙比 e 、饱和度 S_r (%)、液限 ω_L (%)、塑限 ω_p (%)、快剪黏聚力 c_q (kPa)、快剪摩擦角 φ_q ($^\circ$)、压缩模量 E_s (MPa)、压缩系数 α (MPa $^{-1}$) 8 个物理力学参数，共计 11 369 组数据。

(2) 从数据库中抽取训练数据，训练得到 GAN 生成器。使用训练得到的生成器生成样本分布并与原数据分布进行比较，发现二者相当吻合，GAN 较为准确地捕捉了多维参数的分布信息。

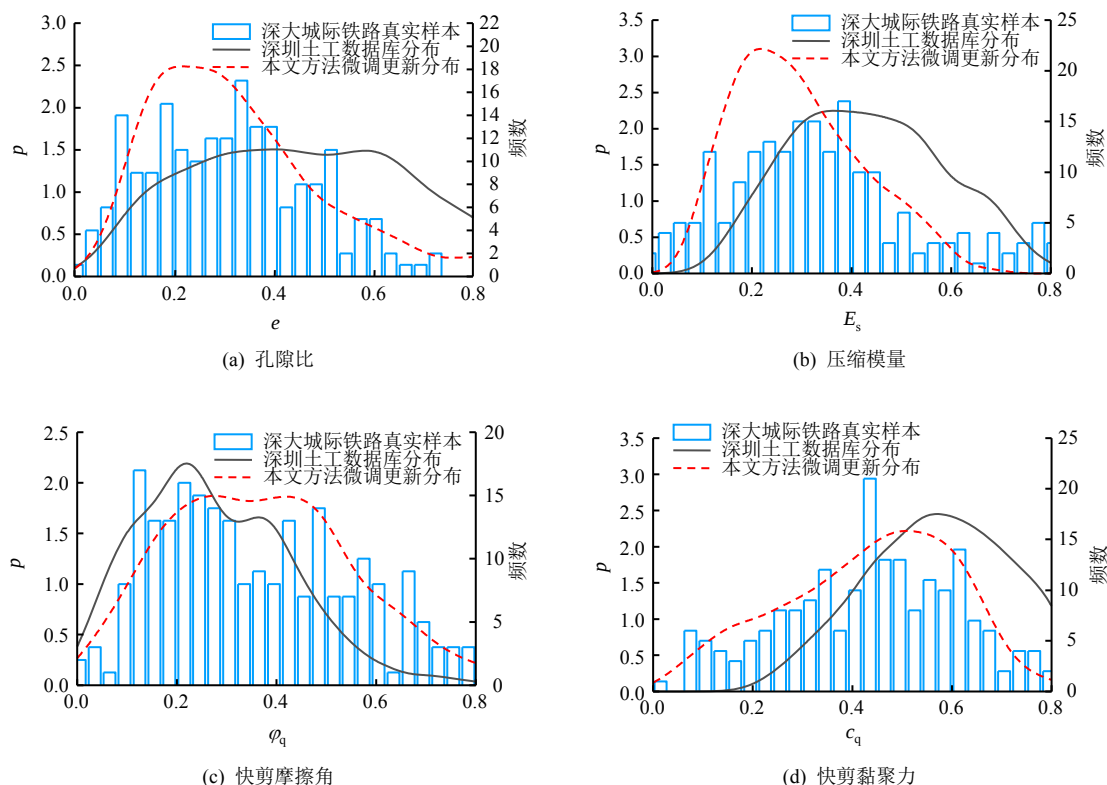


图6 微调生成器生成的部分岩土参数分布曲线、预训练所用深圳岩土数据库对应参数的分布曲线以及项目数据参数分布曲线

Fig.6 Distribution curves of geotechnical parameters generated by fine-tuned generator, corresponding parameters of database used for pre-training, and project data parameters

表5 场地数据分布与深圳土工数据库分布、本文方法分布的 K-S 统计量

Table 5 K-S statistics between the distribution of site data and the distribution of the Shenzhen geotechnical database, the distribution of the proposed method

对比对象	孔隙比 e	饱和度 S_r /%	液限 ω_L /%	塑限 ω_p /%	快剪黏聚力 c_q / MPa	快剪摩擦角 φ_q /(°)	压缩模量 E_s / MPa	压缩系数 α / MPa ⁻¹
深圳土工数据库样本分布	0.42	0.14	0.28	0.22	0.09	0.13	0.39	0.12
本文方法得到的分布	0.09	0.09	0.07	0.20	0.06	0.07	0.10	0.03

(3) 在训练得到 GAN 生成器的基础上, 借助 GAN 反演技术, 对验证数据中随机抽样得到的样本缺失值进行预测。结果表明, GAN 充分利用多个参数的信息, 有效提高了缺失值的预测效果。

(4) 将训练得到 GAN 生成器作为预训练模型, 结合项目数据, 使用微调的方法得到项目岩土参数的后验分布与实际参数分布相近, 说明微调得到的岩土参数分布可靠度较高。

岩土取样数据之间存在自相关和互相关的特性, 在数字岩土工程及场地表征中, 对岩土参数的空间变异性进行分析极其重要。但本文建立多元概率分布模型的过程中, 没有考虑各参数间的空间变异性, 这是本文存在的不足, 在后续的工作中将引入空间相关性的内容。

参考文献

- [1] 中华人民共和国住房和城乡建设部. 2016—2020 年建筑业信息化发展纲要[EB/OL]. (2016-08-23) [2024-12-21]. https://www.mohurd.gov.cn/gongkai/zc/wjk/art/2016/art_17339_228929.html.
Ministry of Housing and Urban-Rural Development of the People's Republic of China. 2016-2020 development outline for construction industry informatization[EB/OL]. (2016-08-23) [2024-12-21]. https://www.mohurd.gov.cn/gongkai/zc/wjk/art/2016/art_17339_228929.html.
- [2] 朱合华, 王长虹, 李晓军, 等. 数字地下空间与工程数据库模型建设[J]. 岩土工程学报, 2007, 29(7): 1098-1102.
ZHU He-hua, WANG Chang-hong, LI Xiao-jun, et al.

- Establishment of database model for geo-digital underground space and project[J]. **Chinese Journal of Geotechnical Engineering**, 2007, 29(7): 1098-1102.
- [3] 朱合华, 李晓军, 陈雪琴. 基础设施建养一体数字化技术(1)——理论与方法[J]. 土木工程学报, 2015, 48(4): 99-110, 123.
- ZHU He-hua, LI Xiao-jun, CHEN Xue-qin. Digitization techniques for integrated construction and maintenance of infrastructure—Part 1: theory and methodology[J]. **China Civil Engineering Journal**, 2015, 48(4): 99-110, 123.
- [4] 朱合华, 骆晓, 彭芳乐, 等. 我国城市地下空间规划发展战略研究[J]. 中国工程科学, 2017, 19(6): 12-17.
- ZHU He-hua, LUO Xiao, PENG Fang-le, et al. Development strategy on urban underground space planning in China[J]. **Strategic Study of CAE**, 2017, 19(6): 12-17.
- [5] 郑刚, 程雪松, 周海祚, 等. 岩土与地下工程结构韧性评价与控制[J]. 土木工程学报, 2022, 55(7): 1-38.
- ZHENG Gang, CHENG Xue-song, ZHOU Hai-zuo, et al. Resilient evaluation and control in geotechnical and underground engineering[J]. **China Civil Engineering Journal**, 2022, 55(7): 1-38.
- [6] CHING J, PHOON K K. Constructing a site-specific multivariate probability distribution using sparse, incomplete, and spatially variable (MUSIC-X) data[J]. **Journal of Engineering Mechanics**, 2020, 146(7): 04020061.
- [7] 倪红, 刘新宇, 秦玉. 土性参数概率特性对地基承载力可靠度的影响[J]. 解放军理工大学学报(自然科学版), 2004, 5(3): 67-69.
- NI Hong, LIU Xin-yu, QIN Yu. Reliability analysis of bearing capacity of soil[J]. **Journal of PLA University of Science and Technology (Natural Science Edition)**, 2004, 5(3): 67-69.
- [8] 周东, 韦海鑫, 杨涟, 等. 南宁市冲积相黏土物理力学参数概率统计分析[J]. 广西大学学报(自然科学版), 2017, 42(4): 1361-1368.
- ZHOU Dong, WEI Hai-xin, YANG Lian, et al. Probabilistic statistical analysis of physical and mechanical parameters of alluvial clay in Nanning city[J]. **Journal of Guangxi University (Natural Science Edition)**, 2017, 42(4): 1361-1368.
- [9] 苏卫卫, 黄宏伟, 张洁. 上海软黏土抗剪强度指标概率分布类型研究[J]. 地下空间与工程学报, 2012, 8(增刊 2): 1695-1699.
- SU Wei-wei, HUANG Hong-wei, ZHANG Jie. Risk analysis of the underground passage through rail transit construction scheme[J]. **Chinese Journal of Underground Space and Engineering**, 2012, 8(Suppl.2): 1695-1699.
- [10] 陈立宏, 陈祖煜, 刘金梅. 土体抗剪强度指标的概率分布类型研究[J]. 岩土力学, 2005, 26(1): 37-40, 45.
- CHEN Li-hong, CHEN Zu-yu, LIU Jin-mei. Probability distribution of soil strength[J]. **Rock and Soil Mechanics**, 2005, 26(1): 37-40, 45.
- [11] 吴越, 刘东升, 孙树国, 等. 岩土强度参数正态-逆伽马分布的最大后验估计[J]. 岩石力学与工程学报, 2019, 38(6): 1188-1196.
- WU Yue, LIU Dong-sheng, SUN Shu-guo, et al. Maximum posteriori estimation of strength parameters for geotechnical material obeying normal-inverse Gamma distribution[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2019, 38(6): 1188-1196.
- [12] 尹利华, 王晓谋, 张留俊. 天津软土土性指标概率分布统计分析[J]. 岩土力学, 2010, 31(增刊 2): 462-469.
- YIN Li-hua, WANG Xiao-mou, ZHANG Liu-jun. Probabilistical distribution statistical analysis of Tianjin soft soil indices[J]. **Rock and Soil Mechanics**, 2010, 31(Suppl.2): 462-469.
- [13] BAECHER G B, LANNY N A, EINSTEIN H H. Statistical description of rock properties and sampling[C]//The 18th US Symposium on Rock Mechanics (USRMS). Golden, Colorado: ARMA, 1977.
- [14] CHING J, PHOON K K, LI K H, et al. Multivariate probability distribution for some intact rock properties[J]. **Canadian Geotechnical Journal**, 2019, 56(8): 1080-1097.
- [15] CHING J, PHOON K K. Correlations among some clay parameters—the multivariate distribution[J]. **Canadian Geotechnical Journal**, 2014, 51(6): 686-704.
- [16] 顾小芸. 应用数理统计方法确定海洋土的工程特性[C]//中国土木工程学会第四届土力学及基础工程学术会议论文选集. 北京: 中国建筑工业出版社, 1983.
- GU Xiao-yun. Application of mathematical statistics to determine the engineering characteristics of marine soil[C]//Proceedings of the Fourth Academic Conference on Soil Mechanics and Basic Engineering of the Chinese Civil Engineering Society. Beijing: China Architecture & Building Press, 1983.
- [17] 吕小飞, 叶银灿, 潘国富. 杭州湾浅层黏性土物理力学

- 指标的统计分析[J]. 海洋学研究, 2005, 23(4): 8-14.
- LÜ Xiao-fei, YE Yin-can, PAN Guo-fu. Statistical analysis of clay property in the Hangzhouwan Bay[J]. **Journal of Marine Sciences**, 2005, 23(4): 8-14.
- [18] 蒋建平, 李晓昭, 高广运, 等. 南京地铁地基黏土物理力学参数相关性试验研究[J]. 中国铁道科学, 2007, 28(2): 17-24.
- JIANG Jian-ping, LI Xiao-zhao, GAO Guang-yun, et al. Experimental study on the correlativity between the physical and mechanical parameters of foundation clay in Nanjing subway engineering[J]. **China Railway Science**, 2007, 28(2): 17-24.
- [19] 郭淋, 王春艳, 张飞, 等. 标贯试验 N 值与土体物理力学参数的相关性分析[J]. 安全与环境工程, 2012, 19(4): 148-152.
- GUO Lin, WANG Chun-yan, ZHANG Fei, et al. Correlation analysis of SPT N -value and geotechnical physical mechanic parameters[J]. **Safety and Environmental Engineering**, 2012, 19(4): 148-152.
- [20] LIU S, ZOU H, CAI G, et al. Multivariate correlation among resilient modulus and cone penetration test parameters of cohesive subgrade soils[J]. **Engineering Geology**, 2016, 209: 128-142.
- [21] ZOU H, LIU S, CAI G, et al. Multivariate correlation analysis of seismic piezocone penetration (SCPTU) parameters and design properties of Jiangsu quaternary cohesive soils[J]. **Engineering Geology**, 2017, 228: 11-38.
- [22] HYUN I P, SEUNG R L. Evaluation of the compression index of soils using an artificial neural network[J]. **Computers and Geotechnics**, 2011, 38: 472-481.
- [23] ALADAG C H, KAYABASI A, GOKCEOGLU C. Estimation of pressuremeter modulus and limit pressure of clayey soils by various artificial neural network models[J]. **Neural Computing and Applications**, 2013, 23(2): 333-339.
- [24] 李泽政, 刘卫星, 李飞, 等. 基于数据增强的烧结矿转鼓强度预测研究[J]. 烧结球团, 2023, 48(6): 62-68.
- LI Ze-zheng, LIU Wei-xing, LI Fei, et al. Research on prediction of drum strength based on data enhancement sinter[J]. **Sintering and Pelletizing**, 2023, 48(6): 62-68.
- [25] 蒋栋年, 王仁杰. 基于 GAN 的软测量缺失数据生成方法研究[J]. 西北工业大学学报, 2024, 42(2): 344-352.
- JIANG Dong-nian, WANG Ren-jie. Research on the generation method of missing data for soft measurement based on GAN[J]. **Journal of Northwestern Polytechnical University**, 2024, 42(2): 344-352.
- [26] 刘业森, 刘舒, 刘媛媛, 等. 降雨移动对河道洪峰影响研究——以深圳为例[J]. 中国水利水电科学研究院学报(中英文), 2022, 20(4): 306-314.
- LIU Ye-sen, LIU Shu, LIU Yuan-yuan, et al. Influence of rainfall movement direction on flood peak in urban river with Hewan district in Shenzhen as a case study[J]. **Journal of China Institute of Water Resources and Hydropower Research**, 2022, 20(4): 306-314.
- [27] 张运标. 深圳地区第四系主要地层特征探讨[J]. 广州建筑, 2011, 39(5): 3-10.
- ZHANG Yun-biao. Shenzhen quaternary stratigraphy of the main layer[J]. **Guangzhou Architecture**, 2011, 39(5): 3-10.
- [28] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.
- [29] ZHA W, LI X, XING Y, et al. Reconstruction of shale image based on Wasserstein generative adversarial networks with gradient penalty[J]. **Advances in Geo-Energy Research**, 2020, 4(1): 107-114.
- [30] ZHANG C, LIANG M, SONG X, et al. Generative adversarial network for geological prediction based on TBM operational data[J]. **Mechanical Systems and Signal Processing**, 2022, 162: 108035.
- [31] ZHANG T F, 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.
- [32] SUNG Y L, CHO J, BANSAL M. LST: ladder side-tuning for parameter and memory efficient transfer learning[J]. **Advances in Neural Information Processing Systems**, 2022, 35: 12991-13005.