SchemaGAN

A conditional Generative Adversarial Network for geotechnical subsurface schematisation

Campos Montero, F. A.; Zuada Coelho, B.; Smyrniou, E.; Taormina, R.; Vardon, P. J.

DOI
10.1016/j.compgeo.2025.107177

Publication date
2025

Document Version
Final published version

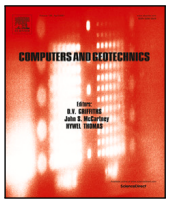
Published in
Computers and Geotechnics

Computers and Geotechnics 183 (2025) 107177

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo



Research paper

SchemaGAN: A conditional Generative Adversarial Network for geotechnical subsurface schematisation

F.A. Campos Montero [a,b], B. Zuada Coelho [a], E. Smyrniou [a], R. Taormina [b], P.J. Vardon [b,*]

[a] Deltares, 2600 MH, Delft, P.O. Box 177, The Netherlands
[b] Delft University of Technology, Faculty of Civil Engineering and Geosciences, 2600 GA Delft, Delft, P.O. Box 5048, The Netherlands

ARTICLE INFO

Dataset link: https://github.com/fabcamo/schemaGAN, https://zenodo.org/records/13143431

Keywords:
Generative adversarial network
AI
Machine learning
Subsurface schematisation
Cone penetration test

ABSTRACT

Subsurface schematisations are a crucial geotechnical problem which generally consists of filling substantial gaps in subsurface information from the limited site investigation data available and relying heavily on the engineer's experience and occasionally geostatistical tools. To address this, schemaGAN, a conditional Generative Adversarial Network (GAN) to generate geotechnical subsurface schematisations from site investigation data is introduced. This novel method can learn complex underlying rules that govern the subsurface geometries and anisotropy from a big database of training cross-sections, and can produce subsurface schematisations from Cone Penetration Tests (CPT) in an insignificant timeframe. To test and demonstrate the performance of schemaGAN, a database of 24,000 synthetic geotechnical cross-sections with their corresponding CPT data was created, including spatial variability and gradually spatially varying layers. After training, the effectiveness of schemaGAN was compared against several interpolation methods, and it is seen that schemaGAN outperforms all other methods, with results characterised by clear layer boundaries and an accurate representation of anisotropy within the layers. SchemaGAN's superior performance was confirmed through a blind survey, and in two real case studies in the Netherlands, where the model demonstrates better predictive accuracy for known CPT data.

1. Introduction

Site investigation typically comprises penetration tests and boreholes, which gather one-dimensional data at discrete locations, that are used to inform schematisations of the subsurface between these locations and to estimate material parameters or inform sample collection (Clayton et al., 1995; Wang and Tian, 2023). Occasionally, such data is also used to investigate the statistical properties of layers, including both more standard statistical measures (mean and standard deviation) and less well known (e.g. spatial correlation of properties) (Jaksa et al., 2005; de Gast et al., 2021), and it is unsurprisingly observed that the more data that is collected, the better the outcome (Ching et al., 2022). However, geotechnical exploration is still often underfunded (Jaksa, 2000; Oluwatuyi et al., 2023) and sparsely measured (Wang and Tian, 2023), which results in substantial gaps in subsurface models for which no information is available, yet assumptions need to be made to complete the subsurface models (Shi and Wang, 2021).

Traditionally, experienced engineers perform the schematisation of layer geometry and variability using results from in-situ tests and their expert knowledge (Phoon et al., 2019, 2022), and occasionally interpolation and geostatistical tools (McBratney et al., 2003; Grunwald, 2009; Heuvelink and Webster, 2022). This process is labour-intensive, time-consuming, and heavily reliant on the engineer's expert judgment (Shi and Wang, 2021; Phoon et al., 2022; Wang et al., 2022; Hudson et al., 2023), leading to non-unique and non-traceable solutions. As projects grow in size and complexity, maintaining consistency and accuracy in subsurface schematisation becomes increasingly challenging (Clayton et al., 1995; Shi and Wang, 2021). Moreover, valuable knowledge and information about the subsurface from previous projects are often not utilised in the new schematisations, leading to a waste of resources.

To address this, Machine Learning (ML), particularly Deep Learning (DL) emerges as an attractive option that has been increasingly utilised in geotechnical engineering (Shi and Wang, 2021; Wang and Tian, 2023; Phoon et al., 2023) due to its ability to model complex and nonlinear processes without requiring initial assumptions about the relationships between input and output variables (Goh et al., 2018; Nguyen et al., 2019; Zhang et al., 2021). Recent reviews on the state-of-the-art Artificial Intelligence (AI) in geotechnics are offered by Zhang et al. (2021), Baghbani et al. (2022) and Phoon and Zhang (2022).

* Corresponding author.
E-mail address: P.J.Vardon@tudelft.nl (P.J. Vardon).

https://doi.org/10.1016/j.compgeo.2025.107177
Received 10 June 2024; Received in revised form 15 February 2025; Accepted 26 February 2025
Available online 11 March 2025

Past studies on subsurface schematisation and ML techniques offer various approaches to tackle this problem. Hudson et al. (2023) used unsupervised learning with K-means clustering and a Gaussian model to detect layers in CPT data, but manual parameter calibration and layer identification review were needed. Wang et al. (2019) proposed a Bayesian unsupervised learning method for layer detection from CPT data, providing uncertainty quantification with low computation cost, yet the approach struggled with thin layering. Wu et al. (2021) proposed an approach using random forests for 1D layer detection from CPTs and used neural networks for 3D interpolation with high accuracy with limitations of only working with sand, clay and gravel layers. Shi and Wang (2021) developed a convolutional method for learning stratigraphic patterns and interpolating subsurface cross-sections from boreholes, but this was limited to specific sites. Wang et al. (2020) combined Bayesian compressive sensing and edge detection techniques for layer delimitation in CPT data, which was able to interpolate consistently between CPT soundings in a multi-layer system, while it was observed to lose resolution between layers, which was improved when using significantly more data. Li et al. (2016b) and Wang et al. (2017) and Li et al. (2016a), Qi et al. (2016) explored Markov random fields and hidden Markov chain models, respectively, facing challenges with complex stratigraphy. Zhou et al. (2019) utilised recurrent neural networks for building 3D geostratigraphic series based on borehole data but also required a substantial amount of drilling data for building a reliable model. Finally, Wang et al. (2022) adopted a multi-point statistics method for automatic generation of subsurface cross-sections from sparse borehole data and geological knowledge, with the disadvantage that the model applies only to the geological context in which it was trained.

This study approaches the problem by leveraging generative AI, particularly Generative Adversarial Networks (GANs). GANs are a type of DL model consisting of two neural networks, which are trained simultaneously in a competitive setting to learn the underlying distribution of the training data (Goodfellow et al., 2014). Unlike previous methods, GANs can generate new data points that closely mirror the existing data (Creswell et al., 2018; Goodfellow et al., 2020), a critical aspect for accurately capturing the complex relationships in subsurface schematisation, where data is sparse and heterogeneous. State-of-the-art research consistently presents GAN applications as an improvement over traditional methods (Gao et al., 2020; Ravuri et al., 2021; Yan et al., 2021; Heuvelink and Webster, 2022), including scenarios in which complex geometries need to be solved (Sun et al., 2023).

In geotechnical engineering, GANs have been applied in various areas such as porous media structure analysis (Mosser et al., 2017; Valsecchi et al., 2020), geological modelling (Azevedo et al., 2020), tomography image processing (Chen et al., 2020), flow characteristics assessment (Janssens et al., 2020), and seismic data enhancement (Oliveira et al., 2019). They have shown capabilities in interpolating signals from sparse monitoring data for environmental mapping (Gao et al., 2020), and spatial interpolation of elevation data (Zhu et al., 2020).

Recently, Lyu et al. (2024) presented a GAN developed to build 3D subsurface geological models from borehole data. This model showcased the potential of GANs, with the practical drawback of being dependent on a prior geological knowledge model during training, making it a site-specific model.

Building upon the ideas of Smyrniou and Zuada Coelho (2023) and the groundwork from Isola et al. (2017), this research introduces schemaGAN, a novel model that can learn complex underlying rules that govern the subsurface layer geometry and anisotropy and can generate accurate cross-sections from sparse CPT data. The system is built as a general tool that is trained once and learns from a big database of synthetic geotechnical cross-sections that encompass all levels of complexity from simple layered models to highly complex geometries with missing layers, indentations and lenses among others. It is designed to incorporate practical considerations such as using less than 1% of the data as input and rules for the separation and depth of the CPT data, all aimed at enhancing the realism and applicability of the generated outputs.

Section 2 details the design of the GAN, including the objective function, the architectural decisions behind the GAN, and the training algorithm. As part of a DL framework, the model requires a robust dataset for training. In this context, a synthetic training dataset is used, which is produced based on several realistic assumptions of a geotechnical system. This is presented in Section 3.

To evaluate its effectiveness, several experiments were conducted to compare schemaGAN with traditional interpolation methods, primarily using synthetic data but also using a real case study. These experiments are detailed in Section 4, with the results presented in Section 5. The paper concludes by summarising the findings and outlining potential future directions in Section 6.

2. SchemaGAN

At their core, GANs operate on the principles of a minimax game between two distinct networks: the Generator (G) and the Discriminator (D). They engage in adversarial training, where the Discriminator aims to maximise its ability to classify between real and fake/generated data, while the Generator aims to minimise the Discriminator's ability to correctly classify generated data (Goodfellow et al., 2014; Goodfellow, 2017; Tomczak, 2022). As they train, the Generator improves its generative capability until it produces samples that are ideally indistinguishable from real data. This process enables the Generator to effectively learn and replicate the underlying distribution even in complex data (Goodfellow, 2017; Goodfellow et al., 2020).

An extension of traditional GANs is the conditional GAN (cGAN). In addition to the adversarial training process, cGANs introduce conditioning, which provides additional information to the Generator to guide the generation process (Mirza and Osindero, 2014; Tomczak, 2022). In a cGAN, the model learns the relationship between a conditional input c and the real data y. After training, the Generator can then be utilised to generate data $G(c, z)$ that closely resembles the original data based on both the conditional input c and a random noise vector z.

In this study, the network architecture from Isola et al. (2017) (i.e. the pi×2pi× method) is modified in order to perform translation tasks from one-dimensional sparse CPT data to complete subsurface schematisations. For the training of schemaGAN, a synthetic database of geotechnical schematisations was generated. For each schematisation, an accompanying synthetic CPT image was generated as described in Section 3.4 and will be referred to as CPT-like data. Illustrated in Fig. 1, the training framework comprises two independent networks: the Generator and the Discriminator. The Generator takes as input the CPT-like data that acts as a condition c, along with a noise vector z, and produces a schematisation of the subsurface. Meanwhile, the Discriminator is presented with a balanced mix of real schematisations and schematisations generated by the Generator, each paired with its corresponding conditional CPT-like data, and it is tasked with distinguishing between real and fake schematisations. Depending on the Discriminator's assessment, feedback is provided to either the Generator, encouraging it to refine its generation of fake schematisations, or to the Discriminator, prompting it to improve in differentiating between real and fake schematisations. This feedback mechanism is integrated into the networks through the utilisation of loss functions. Through multiple training cycles, the Generator progressively learns to deceive the Discriminator, ultimately becoming a proper tool to translate CPT-like data into accurate subsurface schematisations.

2.1. Objective function

The objective function of SchemaGAN is a composite loss function that captures the training process illustrated in Fig. 1, and it is defined as:

$$\min_G \max_D \mathcal{L}_{cGAN}(D, G) + \lambda \mathcal{L}_{L1}(G) \tag{1}$$

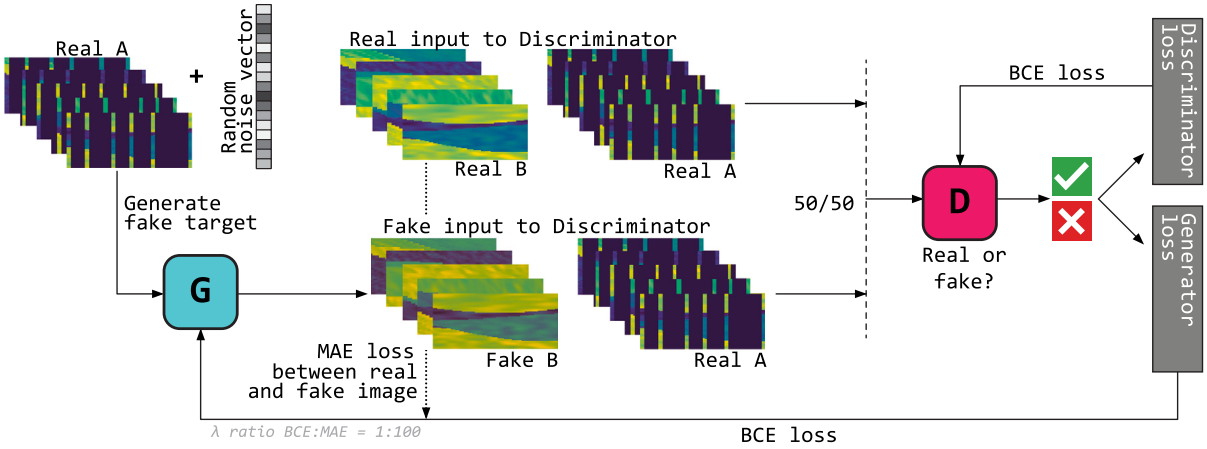


Fig. 1. Adversarial behaviour in the SchemaGAN training for the Generator and Discriminator.

where the adversarial loss ($\mathcal{L}_{cGAN}$) measures the Generator's ability to deceive the Discriminator, with the Discriminator striving to maximise this loss and the Generator aiming to minimise it. On the other hand, the $L1$ loss evaluates the similarity between the generated and real samples on a pixel-wise level, encouraging the Generator to produce outputs resembling the ground truth. The coefficient λ adjusts the relative importance of the $L1$ loss compared to the adversarial loss.

The cGAN loss function $\mathcal{L}_{cGAN}(D,G)$ employs a Binary Cross-Entropy (BCE) loss, a common choice for binary classification tasks, including adversarial training. In schemaGAN, the BCE loss quantifies the difference between the Discriminator's predictions for real and generated schematisations, penalising deviations from the true distribution and promoting adversarial learning (Goodfellow et al., 2016):

$$\mathcal{L}_{cGAN}(D,G) = \mathbb{E}[y \log D(y,c)] + \mathbb{E}[(1-y)\log(1-D(c,G(z,c)))] \tag{2}$$

In the first term, $\mathbb{E}[y \log D(c,y)]$, $\mathbb{E}$ represents the expectation, D is the Discriminator network, c denotes the CPT-like data serving as conditional information, and y represents real subsurface schematisations. The first term penalises the Discriminator when it misclassifies real samples ($y = 1$), encouraging it to assign higher probabilities to real data and increasing the ability to discern between real and fake schematisations. In the second term, $\mathbb{E}[(1-y)\log(1-D(c,G(c,z)))]$, the Generator G produces fake subsurface schematisations $G(c,z)$ based on both the conditional data c and a noise vector z. The goal is for these generated schematisations to be indistinguishable from real ones. Thus, the Discriminator is penalised for incorrectly classifying generated samples ($y = 0$), prompting it to assign lower probabilities to generated data and pushing the Generator to produce more realistic schematisations over time, as it learns to create outputs that deceive the Discriminator into classifying them as real.

By combining these terms, the cGAN loss function facilitates adversarial training, where the Discriminator learns to differentiate between real and fake subsurface schematisations while the Generator aims to produce convincing representations of subsurface conditions guided by CPT-like data. This iterative process drives the Generator to capture the intricate patterns and characteristics inherent in real subsurface scenarios.

The second loss term $\lambda\mathcal{L}_{L1}(G)$ from Eq. (1) represents the mean absolute error (MAE) between the generated subsurface schematisations $G(c,z)$ and the actual subsurface schematisations y. It is defined as:

$$\mathcal{L}_{L1}(G) = \mathbb{E}(|y - G(c,z)|). \tag{3}$$

This loss term addresses mode collapse, a common challenge in cGANs (Li et al., 2017; Bau et al., 2019), by guiding the Generator network G to produce more accurate subsurface schematisations. By computing the MAE between the generated outputs and the ground truth, the Generator aims to create outputs that closely resemble the real data at a pixel level, thereby enhancing fidelity.

The λ parameter in this Eq. (1) acts as a balancing factor, allowing adjustments in the trade-off between image diversity and accuracy. A smaller λ encourages the Generator to generate more diverse outputs, while a larger λ prioritises fidelity to the real data. This flexibility enables fine-tuning of the Generator performance (Isola et al., 2017).

2.2. The generator architecture

The Generator employs convolutional neural networks (CNNs), specifically a modified U-Net architecture, due to their efficacy in handling high-dimensional data (LeCun et al., 2015). U-Net is a neural network architecture commonly used for image segmentation tasks, characterised by its encoder–decoder structure with skip connections in a symmetric structure (Ronneberger et al., 2015), facilitating the conveyance of both local and global structural information for generating high-quality images.

Fig. 2 illustrates the general architecture of the Generator, which comprises regular and irregular encoder and decoder blocks arranged in a U-Net framework (Ronneberger et al., 2015).

Beginning with CPT-like images of dimensions 512 × 32 the encoder progressively downsamples the images through eight sequential blocks. The initial four blocks are regular encoder blocks (RE), each with increasing numbers of filters (64, 128, 256, and 512), alongside standard batch normalisation to improve the stability of the network by normalising the activations of each layer (Santurkar et al., 2018). Subsequently, irregular encoder blocks (IE) are employed from the fifth to the eighth block, with 512 filters each and standard batch normalisation. These blocks adjust the stride of the convolutional layer to 1 × 2, which determines the step size of the filter's movement across the input data, accommodating the schematisation 1:16 aspect ratio. Following findings from Goodfellow et al. (2014), Isola et al. (2017) layer weights are initialised with a standard deviation of 0.02 for stable training, and a leakyReLU activation function with an alpha of 0.2, where alpha defines the slope of the function for negative input values. This helps to improve the flow of gradients and prevents neurons from becoming saturated. Following the last encoder block, a bottleneck is formed by a convolutional 2D layer (Conv2D) with 512 filters, a kernel size of 4 × 4, and a stride of 2 × 2, compressing the image into a lower-dimensional representation with a ReLU activation function. The decoder process mirrors the encoder's structure symmetrically, starting with a regular decoder block (RD), followed by four irregular decoder blocks (ID), and concluding with three additional RD blocks. Dropout regularisation of 50% is used to prevent overfitting by randomly dropping out a fraction of the units during training (Srivastava et al., 2014).

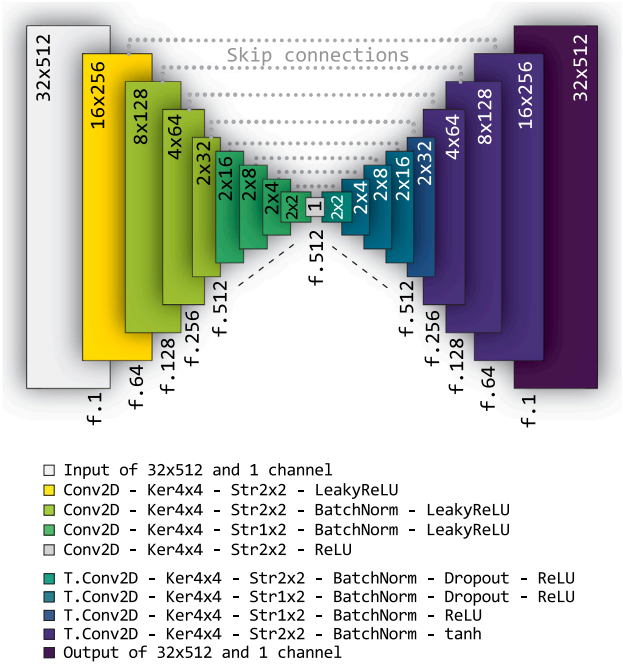


Fig. 2. SchemaGAN Generator architecture based on a U-Net framework.

Each decoder block receives a skip connection from the corresponding encoder block's output, aiding in localising and reconstructing details in the output (Ronneberger et al., 2015). The final layer is a transposed Conv2D layer using a 4 × 4 kernel and a stride of 2 × 2, which upscales the feature maps back to the original image size of 512 × 32 pixels with a *tanh* activation function. The output is an image of size 512 × 32 pixels with pixel values ranging from −1 to 1.

Throughout training, the weight parameters within the Generator are iteratively adjusted based on the results from the objective function, optimising its ability to generate subsurface schematisations that closely align with the desired output.

2.3. The discriminator architecture

The Discriminator functions as a combination of a CNN and a patch-based discriminator, as proposed by Demir and Unal (2018). Initially, the input images are passed through the CNN which extracts features at multiple levels of abstraction, capturing spatial patterns and textures from the input images. Following the feature extraction, the discriminator evaluates the realism of these features at a patch level rather than considering the entire image at once. This patch-based discrimination allows the Discriminator to focus both on local and finer details (Isola et al., 2017; Demir and Unal, 2018). By combining the feature extraction from the CNN and the patch-based discrimination, the Discriminator learns to distinguish between real and fake schematisations, ultimately guiding the Generator towards producing the desired results.

Fig. 3 provides an overview of the Discriminator architecture used in schemaGAN. The process initiates with two input images, each sized 512 × 32 × 1, which are merged and passed through a Conv2D layer featuring 64 filters, a 4 × 4 kernel size, and strides of 2 × 2, with a LeakyReLU activation function. Sequentially, the model uses four additional Conv2D layers, each employing a 4 × 4 kernel size with an escalating number of filters: 128, 256, 512, 512. These layers apply a stride of 1 × 2 to change the aspect ratio from 1:16 to 1:1 for the patch-based discrimination. Batch normalisation is used in all but the first layer, followed by a LeakyReLU activation function. Subsequently, the model channels the output through another Conv2D layer, maintaining

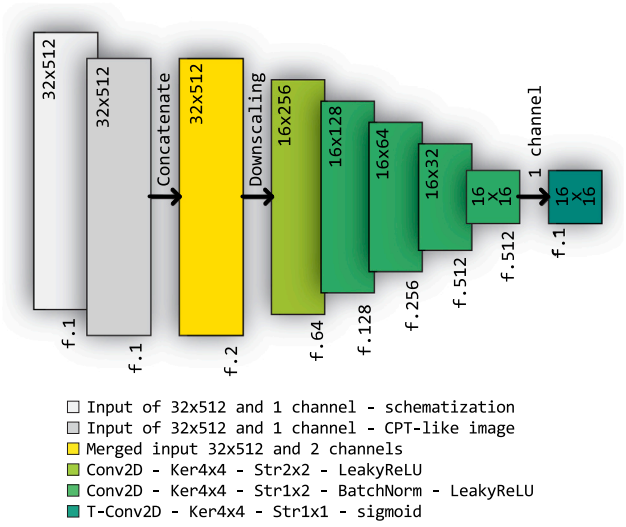


Fig. 3. SchemaGAN Discriminator architecture based on a PatchGAN framework as a binary classifier.

spatial dimensions at 16 × 16 with a stride of 1 × 1, and including batch normalisation and LeakyReLU activation. The last step involves another Conv2D layer with a single filter and a 4 × 4 kernel size, producing patch-based output scaled to a range of 0 to 1 using a sigmoid activation function, signifying the probability of each 16 × 16 patch in the image being real or fake.

For training, the Discriminator is compiled with the Adam optimiser (Kingma and Ba, 2014), employing a learning rate of 0.0002 and a beta coefficient of 0.5. The learning rate controls how much to change the model in response to the estimated error each time the model weights are updated, while the beta coefficient, specifically β_1 in the context of the Adam optimiser, refers to the exponential decay rate for the first-moment estimate (mean), which helps in smoothing out the gradient updates. These values were chosen based on experimentation and draw upon typical GAN hyperparameters established in previous research (Goodfellow et al., 2014; Salimans et al., 2016), adhering to best practices in the field of GAN training. During training, the Discriminator's parameters are adjusted based on the results from the loss function just like in the Generator.

2.4. The combined cGAN architecture

For the compilation of the combined Generator and Discriminator networks into the schemaGAN model, the Adam optimiser was also used with a learning rate of 0.0002 and a beta value of 0.5 (Goodfellow et al., 2014; Salimans et al., 2016).

The complete schemaGAN model boasts a total of 78,172,226 parameters representing the weights and biases of the network's layers. From these, a total of 67,003,137 are trainable parameters adjusted during the training process to optimise the model's performance. The remaining 11,169,089 parameters are non-trainable and contribute to the structure and functionality of the network without being influenced by training.

2.5. Training and validation of the model

The final model undergoes training for 200 epochs, where each epoch represents one complete pass through the entire dataset. During training, a batch size of 1 is utilised, meaning that each training iteration involves processing a single sample from the dataset. This approach mitigates mode collapse, allows for precise gradient updates, conserves memory, and is known for faster convergence in the models (Goodfellow et al., 2014; Salimans et al., 2016).

After experimentation with different loss functions and weighting factors, the combined approach outlined in Eq. (1) with a λ value of 100 was selected as the optimal architecture. This choice implies a weight ratio of 1:100 between the BCE conditional GAN loss and the MAE loss. By prioritising the MAE loss, the generated schematisations maintain fidelity to the input data while also exhibiting realistic features and structures (Isola et al., 2017).

A database of 16,000/4,000/4,000 cross-sections was created for the training, testing, and validation of schemaGAN, and for each sample, the corresponding CPT image was generated. All the intensive computation tasks were performed on DelftBlue supercomputer (DHPC, 2022) as detailed in the Section 4.

3. Synthetic data generation

Training a cGAN capable of capturing all the complex relationships underlying a subsurface schematisation, related to geometry in the layer boundaries and anisotropy inside each layer, requires a database of thousands of geotechnical cross-sections (Bowles et al., 2018; Karras et al., 2020). Given the impracticality of collecting and digitising a vast number of geotechnical models for training, a synthetic database of subsurface schematisations was developed.

These datasets are generated to train schemaGAN in the form of cross-sections of the subsurface (i.e. 2D images) populated with Soil Behaviour Type index (Ic) values (Robertson, 2010; Robertson and Cabal, 2022). This was chosen in order to mimic values which can be directly calculated from CPT data. It is noted that the cGAN technique is not limited to Ic values and can be extended to any parameter represented as a single measurement, offering a versatile approach in geotechnical practice. For detailed information on the synthetic data generation, readers are referred to Campos Montero (2023).

The creation of synthetic images involves three primary steps. Firstly, the computation of layer boundary geometry is conducted. Subsequently, materials are assigned with intrinsic variability and anisotropy. Lastly, accompanying CPT-like images are generated. The images are formatted to a size of 512 × 32 pixels, maintaining a ratio of 1:16 for length over depth. This balance ensures computational efficiency while preserving visual fidelity to capture the intricacies inherent in geotechnical schematisations.

3.1. Geotechnical assumptions

The synthetic subsurface schematisation process incorporates several assumptions in order to generate what are thought to be reasonably realistic cross-sections, with a general reference to the Dutch subsurface:

- Layered system: multiple layers with distinct properties.
- Irregular layer shapes: variations in thickness, shape, and orientation.
- Lenses and old channels: presence of lenses and old channels.
- Heterogeneity and anisotropy: anisotropic spatial variations in soil properties within each layer.

The site investigation considered, follows similar constraints to those of real site investigations:

- Incomplete site investigation penetration: limited depth of geotechnical surveys.
- Minimum distance between CPTs: CPTs cannot be directly adjacent to another.

To limit the range of possibilities to test the model, the following limitations are used:

- Limited Layers: Maximum of five layers for practical analysis while capturing essential variations.
- Horizontal surface: Assumes the top layer is horizontal.

Table 1
Soil type Ic values for the generation of the random synthetic models (Robertson and Cabal, 2022).

Material	Mean Ic	std. dev. Ic	Distribution
Sand	1.68	0.12	Gaussian
Sand mixtures	2.33	0.09	Gaussian
Silt mixtures	2.78	0.06	Gaussian
Clays	3.28	0.11	Gaussian
Organic soils	3.90	0.10	Gaussian

3.2. Layer boundary variability

The first step in creating a synthetic cross-section is computing the geometry of the layer boundaries. This is achieved using a periodic function, either a sine or cosine, with randomly determined parameters that vary in their characteristics for each layer boundary and cross-section. These functions range from largely horizontal to extremely sinuous. The phase and vertical location are also randomly determined (Campos Montero, 2023).

The process is conducted four times for each cross-section, aiming to generate a maximum of five layers. Interactions between periodic wave functions may yield models with fewer layers (if the function falls outside the domain) and introduce complexity, leading to phenomena like indentation, lens formation, and the occurrence of old channels, as seen the series of example layers shown in Fig. 4.

3.3. Simulating the subsurface heterogeneity

For each layer, a material is randomly assigned from a list of five soil types with different Ic values, as detailed in Table 1. This random allocation process permits repeated occurrences of the same material within a single model, including adjacent regions.

2D Gaussian random fields (Fenton, 1999) are generated using the GStools framework (Müller et al., 2022). These fields are used to populate Ic values for each layer. The random fields exhibit anisotropic variability in both the horizontal and vertical directions, with scales of fluctuation based on Phoon et al. (2000) for sand and clay materials. A triangular distribution was employed to represent horizontal fluctuations and the angle factor, centring around the mean value and prioritising sub-horizontal anisotropy, while the vertical fluctuations were constructed using a uniform distribution. Fig. 5 illustrates the application of the random fields on the property values in the synthetic cross-sections.

3.4. CPT-like data

After generating the subsurface schematisation, a corresponding CPT-like data image is created by randomly removing columns until only 1% of the original data remains. The deletion process follows specific rules. To simulate the distances used in actual surveys, when removing columns it is not possible to have CPT data that is closely spaced. Furthermore, to account for the variability in depths reached during the tests, a random number of rows are deleted from the bottom to mimic incomplete penetration depths. The depth of the CPT data is controlled by a triangular distribution with min and max values of zero and half of the model thickness, and centred around zero to encourage complete depth of the CPT data.

The transformation process from the complete synthetic model to the CPT-like image is visually represented in Fig. 6. Here the thickness of the CPT data has been increased to ease the understanding of the process, as the real CPT-like data image would have 1 pixel wide CPT-like columns.

Fig. 4. Synthetic models simulating subsurface layering variability with simple sub-horizontal layering, steep angles, irregular layers, indentations, and complex geometries with lenses and old channels.

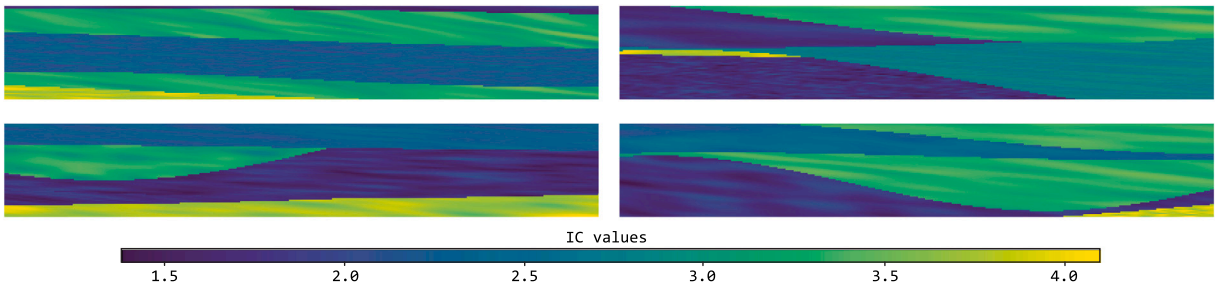


Fig. 5. Synthetic models simulating the anisotropic spatially variable property values within each material layer.

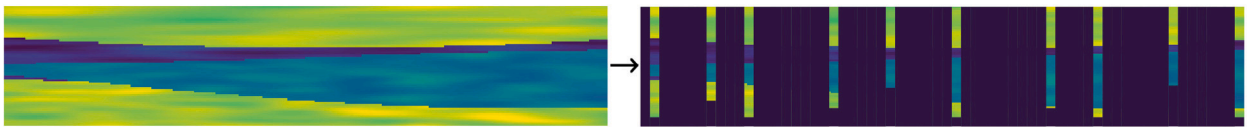


Fig. 6. Transformation from a synthetic schematisation to a CPT-like image (Exaggerated CPT-like data width for illustration purposes).

3.5. Training database limitations

An important effort was made to design a schematisations database that covered an extensive range of layer patterns and internal anisotropy, and included both realistic geotechnical assumptions and practical logic in the quantity placement and depth of the CPT-like data. However, due to the nature of ML models, schemaGAN (or any other DL model to date) will not perform well for sites that present subsurface layering patterns or complexity that the training database does not capture.

4. Experiments

In order to evaluate the effectiveness of schemaGAN in generating geotechnical subsurface schematisations, several experiments were performed, comparing results against a set of traditional interpolations and one advanced interpolation technique. These experiments can be categorised into two main groups based on the origin of the input data used for generating schematisations: synthetic and real data experiments.

The synthetic data experiments involved evaluating schemaGAN against benchmark methods in a large database of variable complexity, investigating the impact of CPT location, and conducting an expert criteria survey. On the other hand, the experiments using real CPT data focused on a rail embankment stretch between Delft and Schiedam in South Holland, leveraging the high density of CPT tests available in a linear array.

All computations were performed using a single NVIDIA Tesla V100S 32 GB on the DelftBlue supercomputer (DHPC, 2022), which allowed for optimisation of the training and testing. The final model was also trained in a single GPU for 95 h. After training, the final model does not require further fine-tuning and can be used on any personal computer to generate schematisation in fractions of a second.

4.1. Benchmark methods

For comparison with schemaGAN, several widely used interpolation methods were selected: nearest neighbour (NeaNe), ordinary Kriging, inverse distance weighting (IDW), and natural neighbour (NatNe) interpolation. These methods were chosen due to their established effectiveness in spatial interpolation tasks (Lam, 1983; Li, 2008; Ledoux et al., 2022).

NeaNe assigns the value of the nearest data point to the location being estimated. Ordinary Kriging is an advanced geostatistical method that estimates values at unmeasured locations based on the spatial correlation of the data. A Radial Basis Function (RBF) kernel combined with a White Noise kernel was used for the Kriging implementation. The RBF kernel was initialised with anisotropic length scales of 50 in the x-direction and 0.5 in the y-direction to account for spatial sampling constraints and variation in different directions. The hyperparameters are optimised using maximum likelihood estimation (MLE). IDW estimates values at unmeasured locations by averaging the values of nearby data points, with closer points weighted more heavily. NatNe interpolation utilises the Delaunay triangulation of the data points to determine the influence of neighbouring points in estimating values at unmeasured locations (Li, 2008; Ledoux et al., 2022).

Additionally, Bayesian Compressive Sampling (BCS) was used via the graphical interface from Lyu et al. (2023) as an advanced data-driven interpolation technique in the comparison. BCS is a probabilistic method that interpolates data from limited measurements by assuming it is sparse in a transformed domain (e.g., wavelets). It uses Bayesian inference for generating a reconstructed signal and estimating the uncertainty (Wang and Zhao, 2017; Zhao et al., 2018). This makes BCS useful for handling incomplete or undersampled data and a good advanced comparison method.

4.2. Evaluation

In order to evaluate the performance of schemaGAN against the benchmark methods, a quantitative piexel-wise evaluation was conducted using the Mean Absolute Error (MAE) and Mean Squared Error (MSE) to measure accuracy. While MAE provides robustness against outliers and MSE highlights larger discrepancies, these metrics have limitations in evaluating image generation tasks (Salimans et al., 2016; Salehi et al., 2020; Benny et al., 2021). These metrics fail to account for the intricacies of image generation regarding subsurface schematisation, in which complex geometries and hard boundaries are present. Therefore, the reliance solely on these metrics may overlook the performance of the methods. As a result, a qualitative evaluation through visual inspection by domain experts was considered to be a valuable tool for assessing the fidelity and practical utility of the schematisations (Borji, 2018, 2022). Additionally, computational efficiency was considered for practicality in geotechnical engineering projects.

4.3. Synthetic data

The evaluation of schemaGAN involved an analysis using 4,000 synthetic cross-sections and their corresponding CPT-like image from the validation database (Fig. 9). These samples represented a diverse spectrum of subsurface complexities, with a slight emphasis on subhorizontal layering models. The performance of schemaGAN and benchmark methods was assessed against the ground truth.

In another experiment, the influence of the CPT data location on the generation of the subsurface schematisations was investigated (Fig. 11) using two distinct cross-sections: one with simple geometry and another with complex geometry. CPT-like data was generated 1,000 times for each cross-section, with randomised locations following the already stated rules about CPT location and depth. Using the CPT-like images, schemaGAN and the benchmark methods were used to generate subsurface schematisations. Analysing the variability and average Mean Absolute Error (MAE) across these experiments provided insights into how the spatial distribution of CPT data affects the accuracy and consistency of schemaGAN and traditional interpolation methods.

Additionally, the realism and geotechnical relevance of the generated schematisations were evaluated by expert judgement. A blind survey was conducted, where experts ranked 10 synthetic schematisations, varying from simple to complex geometries, based on their perceived realism and practical utility (Fig. 12). Readers are referred to the appendix of Campos Montero (2023) to observe all the schematisations used in the survey.

4.4. Real data

The last experiment to evaluate schemaGAN pertained to the use of real CPT data from the Netherlands open database (https://www.dinoloket.nl/). Two field case studies are presented. First a rail embankment between Delft and Schiedam in South Holland, and second the Eemskanaal along a canal between Groningen and Delfzijl (Fig. 13).

The Delft-Schiedam embankment provided 51 Cone Penetration Tests (CPTs) with semi-ordered spacing and depths up to −36 m NAP. For the Groningen-Delfzijl canal, a section of 101 CPTs reaching up to −30 m NAP was selected.

The CPT data was pre-processed for schemaGAN, with IC values calculated as group averages to fit a 32 px height resolution. Both test sites were divided into two 512 px sections, and the CPT positions were aligned with the surface elevation (NAP). The pre-processing included applying a top cutoff to create a flat surface.

The absence of a 2D schematisation as ground truth prompted a different approach for evaluation, where 6 CPTs were randomly chosen to generate a subsurface model, with the remaining CPTs serving as ground truth for comparison at specific locations. This process was repeated 1,000 times in each cross-section, each iteration with a selection of a different set of random 6 CPTs to be used as input to generate schematisation, and the remaining locations used to evaluate the performance.

5. Results and discussion

5.1. SchemaGAN applied to synthetic data

The performance of schemaGAN is assessed by comparing the accuracy of the 4,000 samples validation dataset against the traditional and one advanced interpolation method. Figs. 7 and 8 are shown as two examples from the 4,000 results generated, corresponding to a more simple and more complex schematisation, respectively. The results show that the schemaGAN has the lowest MAE error in both examples. SchemaGAN is able to capture the internal variability, the overall structure and layering of the original cross-section, including complex layer structures and variable properties. NeaNe achieves the lowest MAE error among the alternative methods when a simple geometry is considered, and as the complexity increases, Kriging performs the second best.

Visually, SchemaGAN produces the best results from all the methods, with clear layer boundaries even in scenarios where data was limited at the bottom of the schemas, and capturing the variability inside the layers. NeaNe and Kriging closely match SchemaGAN and the original schematisation in simple sub-horizontal models (Fig. 7) but face challenges as complexity increases (Fig. 8). NeaNe introduces discontinuities and sharp vertical boundaries, distorting the natural subsurface layer behaviour. Conversely, Kriging exhibits increased blurriness in layer boundaries with rising complexity, making it challenging to visualise features like indentations and old channels. The loss of spatial variability further hampers subsurface structure understanding in Kriging interpolations. IDW and NatNe significantly underperform in accurately modelling simple subsurface schematisations compared to SchemaGAN. While NatNe slightly outperforms IDW, it cannot extrapolate beyond available CPT data. For very simple schematisations, both IDW and NatNe initially capture the structure but rapidly deteriorate with increasing complexity. IDW exhibits overly blurry layer boundaries and sharp transitions between provided data, leading to a complete loss of the layering structure. NatNe faces the same issue and additionally lacks the capacity to extrapolate, resulting in consistently incomplete edges in the schematisation.

These findings are supported by state-of-the-art research (Gao et al., 2020; Ravuri et al., 2021; Yan et al., 2021; Heuvelink and Webster, 2022; Sun et al., 2023), which also show GANs outperforming traditional methods. Furthermore, Zhu et al. (2020) showed how for interpolation tasks when using digital elevation models, conditional GANs create clearer representations when compared with IDW and Kriging which have the tendency to generate blurry outputs, just like the comparisons from schemaGAN showed.

For its part, BCS does not perform well under the testing conditions. The generated schematisations show very blurry layer boundaries and struggle to correctly interpolate the bottom of the cross-sections where very little data is present. Additionally, it cannot capture the variability inside the soil layers.

Although BCS has been successful in interpolating points within a single CPT test and creating 3D data-driven models (Lyu et al., 2023), in this evaluation, the model struggles with the sharp layer transitions present in the dataset. BCS assumes sparsity in the wavelet domain, a condition not met by our data. This can be caused by the abrupt changes in layer properties and sharp transitions, which are a characteristic of the training dataset. For this reason and the complexities of evaluating the 4,000 samples manually, the BCS method is not evaluated further.

Fig. 9 illustrates a statistical summary of the results from the evaluation metrics for the whole 4,000 samples validation dataset. The results show that SchemaGAN outperforms all traditional methods, with the lowest mean error values as well as the narrower interquartile ranges, meaning that schemaGAN performs well regardless of the schematisation considered.

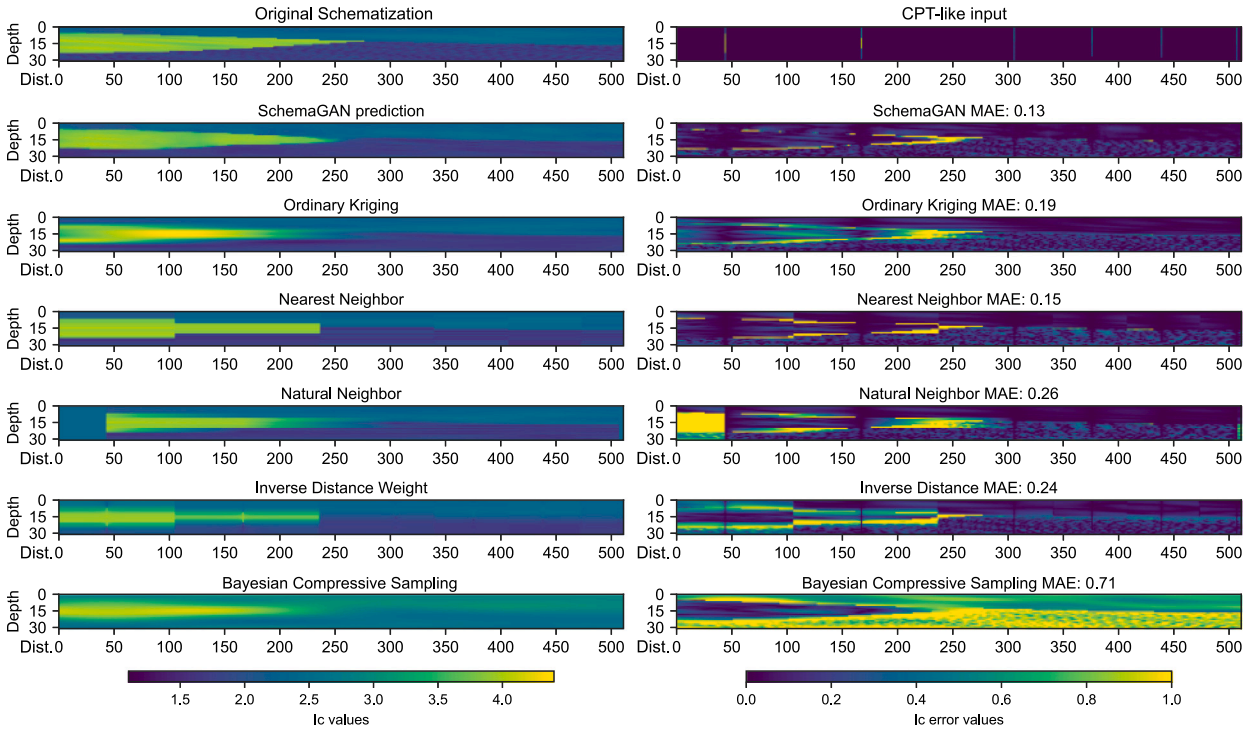


Fig. 7. Example 1: Left: An example subsurface schematisation with irregular layers, including the original (truth), the schematisation generated by SchemaGAN, the traditional interpolation methods and the BCS interpolation. Right: the CPT-like conditional input, and the MAE error map for each method.

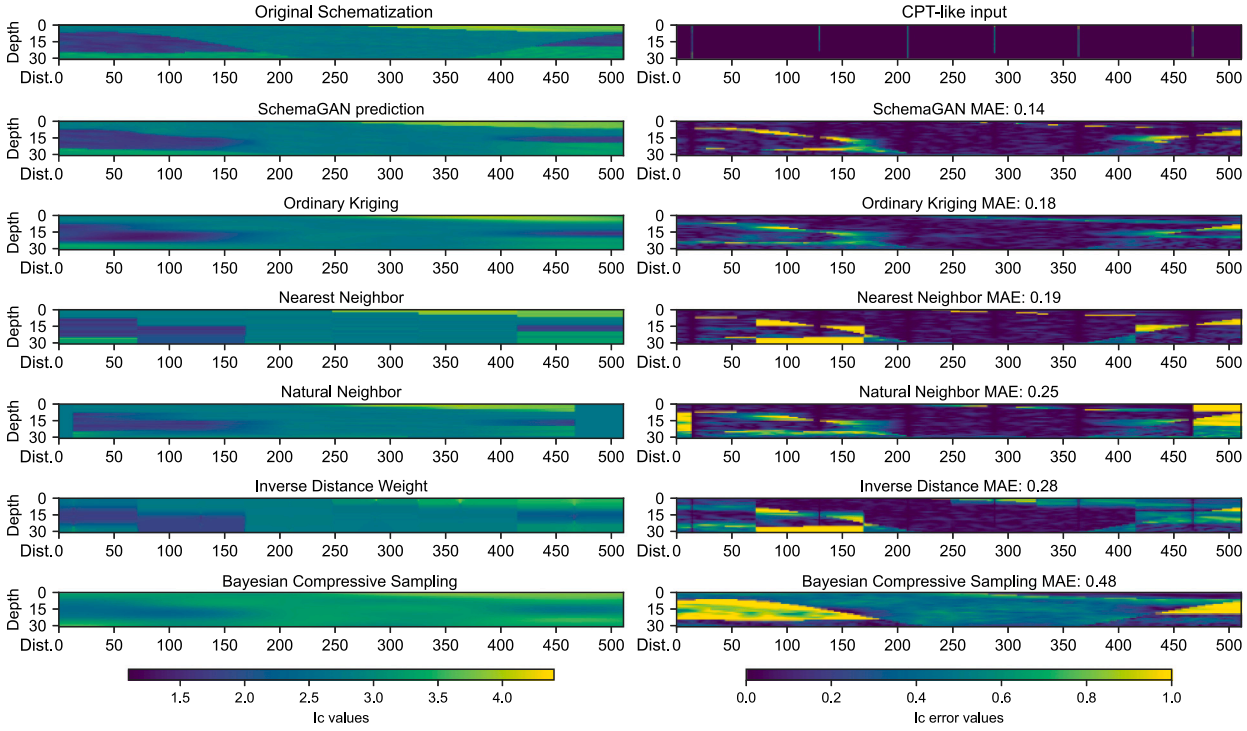


Fig. 8. Example 2: Left: An example subsurface schematisation with irregular layers, including the original (truth), the schematisation generated by SchemaGAN, the traditional interpolation methods and the BCS interpolation. Right: the CPT-like conditional input, and the MAE error map for each method.

All methods show a spread of results. In general, it is seen that simpler schematisations exhibit both lower MAE and MSE values across all methods, attributed to their ease of interpretation. Conversely, more complex structures, with indentations, irregular layering and buried channels, result in higher errors for the traditional methods due to the challenges in accurately interpolating intricate geometries.

The MAE box plots identify NeaNe and Kriging as the second and third-best interpolation methods, respectively. NeaNe excels with simpler geometries but struggles with complexity, as indicated by its larger interquartile range. In contrast, Kriging performs consistently across complexities, evident in its closer alignment with SchemaGAN in both MAE and MSE results.

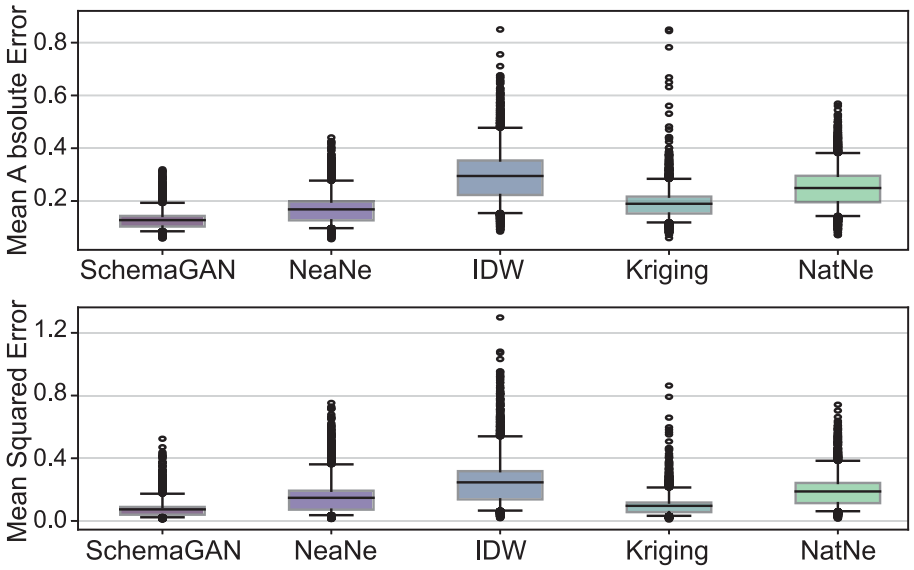


Fig. 9. MAE and MSE error distribution comparison for the traditional interpolation methods against the SchemaGAN model.

Table 2
Inference time for the interpolation of the 4000 schematisations in the test database.

Method	SchemaGAN	NeaNe	IDW	Kriging	NatNe
Time (m)	7.0	1.1	14.8	102.2	18.7

5.2. Computational efficiency

Considering the computational efficiency of each method is a crucial aspect of their practical applicability. Table 2 provides an overview of the inference times required for interpolating 4,000 schematisations in the validation database. NeaNe emerged as the fastest method, completing the task in slightly over a minute due to its inherent simplicity. In comparison, schemaGAN, while more computationally demanding than NeaNe, exhibited reasonable efficiency with a total execution time just exceeding seven minutes. This places it in a pragmatic position, striking a balance between the sophistication of deep learning models and computational speed, especially considering the quality of its output.

Conversely, traditional interpolation methods, including kriging, IDW and NatNe demonstrated considerably lengthier computation times. Kriging, in particular, stood out as the most time-intensive, requiring over 100 min for image generation due to its iterative optimisation step. It is crucial to approach these computational results with caution, considering that traditional interpolation methods are not fully optimised.

The inference times for the schemaGAN exclude the substantial training time requisite for schemaGAN, which amounted to 95 h on the DelftBlue supercomputer (DHPC, 2022). While the training phase is a one-time commitment, it necessitates a significant allocation of computational resources and time.

Similar results have been found by Laloy et al. (2017) and Zhu et al. (2020), indicating that once the GAN model is trained, inference times are fast in comparison to geostatistical alternatives

5.3. Influence of the CPT location

To explore the influence of the location of the CPT tests on the interpolated schematisations two synthetic models; a simple layered system and a complex irregular model, were chosen from the validation database (Fig. 10).

Results in Fig. 11 clearly favour SchemaGAN over other methods, indicated by its narrow interquartile range and lowest MAE values

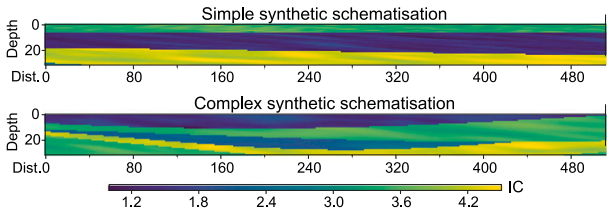


Fig. 10. Simple (top) and complex (bottom) geometry synthetic subsurface schematisation for investigation of the influence of the CPT location.

for both simple and complex models. NeaNe and Kriging follow as the next best methods. NeaNe struggles in simpler models with a wider interquartile range when CPT data misses layer changes at the bottom. Kriging, while maintaining comparable performance on average for both schematisations, exhibits a slight decrease in effectiveness compared to the preceding methods. IDW and NatNe exhibit the largest interquartile range, suggesting heightened sensitivity to CPT data-position across complexity levels. This illustrates the robustness of the schemaGAN in consistently providing good schematisations independently of the location of the CPTs.

5.4. Expert criteria blind survey

To address the limitations in quantitative assessments of performance, a blind survey of 13 expert geotechnical engineers was carried out on 10 schematisations based on the approach presented by Ravuri et al. (2021). The assessment involved evaluating the accuracy in reflecting CPT-like data, quality in representing subsurface layers, realism, and ease of understanding. Each series comprised six images: CPT-like data on top, followed by randomly ordered schematisations labelled A to E from the five methods in a random order. The experts ranked each image from 1 to 5 (worst to best).

The violin plot in Fig. 12 presents the results of the expert survey. SchemaGAN stands out with the widest section at a score of 5 (being the preferred method in 55% of cases), gradually narrowing down, followed by Kriging, which peaks at a score of 4. NeaNe maintains a constant width across all scores, reflecting mixed expert opinions. In contrast, IDW and NatNe consistently occupy the lower scores.

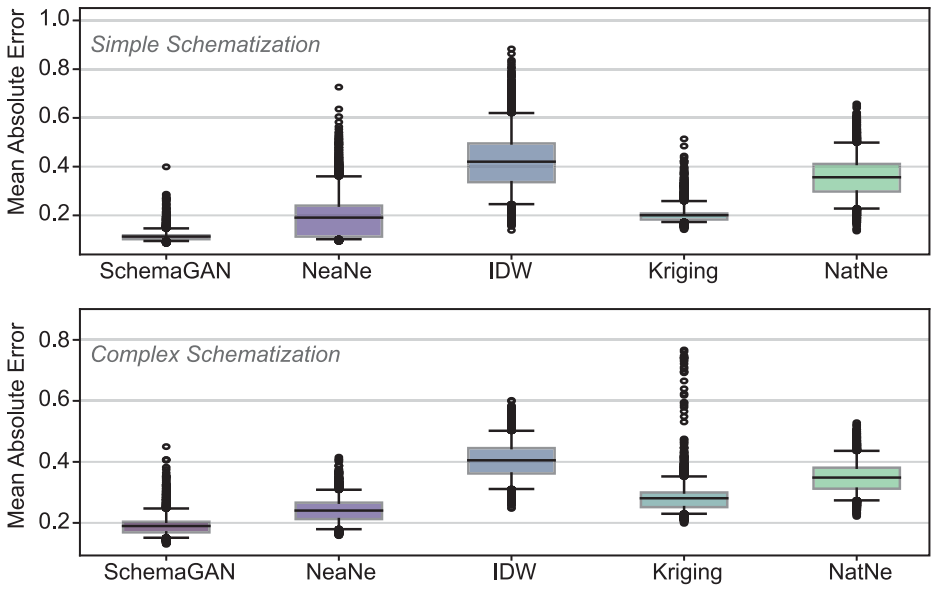


Fig. 11. Influence of the CPT location for the generation of schematisations with SchemaGAN and the traditional interpolation methods.

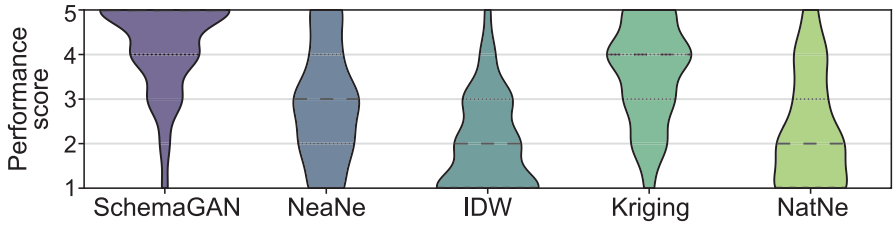


Fig. 12. Percentage of times that each method was chosen as the best subsurface schematisation by expert criteria.

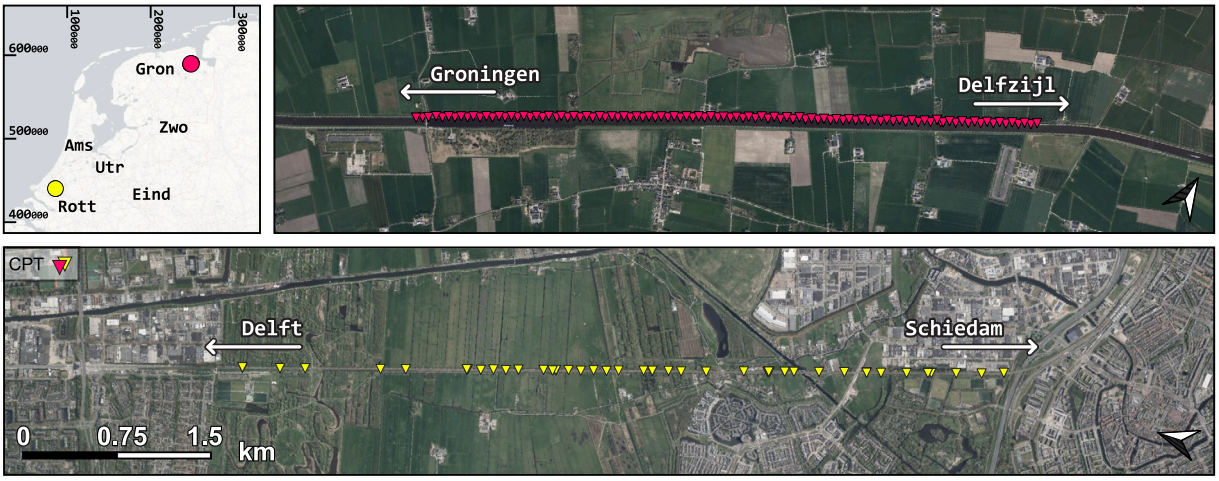


Fig. 13. Overview of the CPT locations in the field for the Delft-Schiedam rail embankment (top) and the Eemskanaal between Groningen-Delfzijl (bottom).

5.5. SchemaGAN applied to real data

In Fig. 13, the Delft-Schiedam rail embankment and the Eemskanaal between Groningen-Delfzijl are shown with 51 and 101 CPT tests respectively with a semi-ordered spacing and depths up to −36 m NAP.

The evaluation procedure is explained in Section 4.4. The results for both sites are shown as boxplots for MAE and MSE in Fig. 14 comparing schemaGAN, NeaNe and Kriging, as they were identified as the most effective methods through synthetic data analysis (see Section 5). SchemaGAN's results show a lower mean and narrower interquartile range compared to other methods, illustrating its effectiveness. Nearest Neighbour interpolation achieved a lower minimum MAE value than SchemaGAN, attributed to the simplification of the CPT data, while Kriging interpolation closely aligned with schemaGAN, with distinctions in the upper interquartile range, in agreement with findings by Rahman et al. (2021) and Section 5.

Fig. 15 shows one instance of schematisations generated by schemaGAN using 6 random real CPT tests in both test sites. The visual representation shows how schemaGAN is able to provide a realistic schematisation of the subsurface with clear boundaries and complex

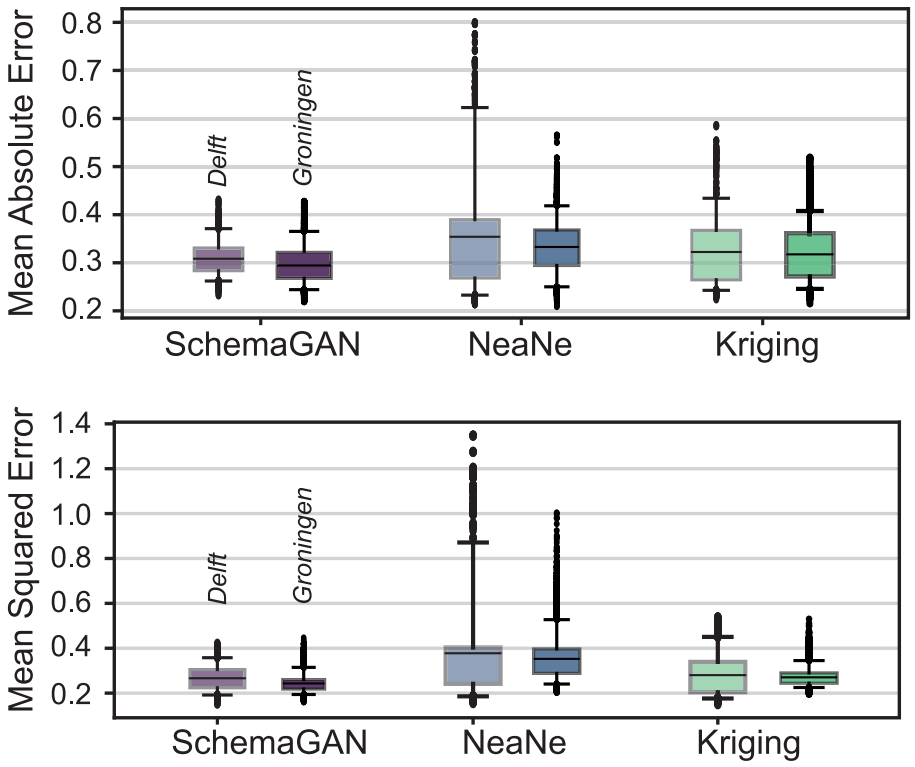


Fig. 14. Evaluation metrics from the best-performing methods on the real CPT data for the case studies in Delft-Schiedam (left) and Groningen-Delfzijl (right).

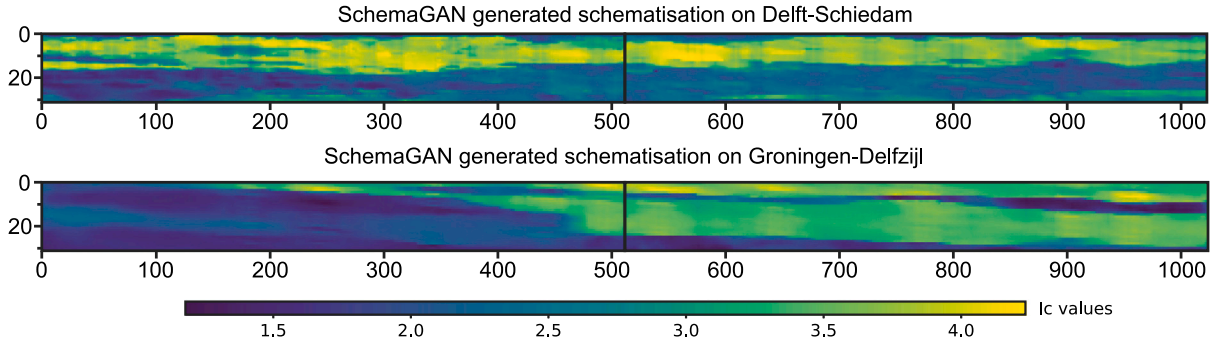


Fig. 15. SchemaGAN schematisation with 6 random CPT tests for Delft-Schiedam (top) and the Eemskanaal (bottom).

geometries like lenses and sharp changes in materials, with variability within the subsurface units.

6. Conclusions

The paper introduces schemaGAN, a novel method designed to transform discrete CPT data, into a continuous subsurface schematisation, with a focus on the soil behaviour index (Ic) while remaining adaptable to other geotechnical parameters. This model is designed to effectively extract complex spatial patterns in the subsurface by learning from nearby and distant features and following practical site investigation decisions like the spacing and depth of the CPT soundings and less than 1% of the total cross-section used as input, resulting in a system that is trained only once and can generate realistic and accurate schematisations with minimal input information.

It has been shown that cGANs are effective tools for learning the complex obscure patterns that govern the construction of geotechnical subsurface schematisations. The adversarial learning process is a simple, yet powerful concept that extracts huge amounts of information from the training data through deep neural networks. This is evident in the generated cross-sections where schemaGAN excels in both the layer boundary shape and definition, and the internal layer anisotropy, while the traditional interpolation methods struggle with blurry boundaries and continuity. Furthermore, cGANs were shown to be highly reliable across a wide range of different conditions, including varying levels of complexity in subsurface structures and different positions of CPT locations. Nevertheless, a significant obstacle in DL applications, lies in the availability of training data. Given the diverse and disparate nature of geotechnical investigations, consolidating disparate datasets into a cohesive training body remains a formidable challenge. Presently, reliance on synthetic data is the only way forward, although future research should strive to integrate real CPT data to refine the schemaGAN model, especially in real-world scenarios in order to better account for the extensive averaging of CPT tests needed for schemaGAN.

Future developments in this approach could entail the incorporation of additional conditional inputs, such as geological and geomorphological maps, or geophysical findings. These additional conditions can guide the model in generating results particular to specific spatial contexts, thereby enriching the characterisation of complex layering systems and broadening the applicability of schemaGAN in geotechnical contexts.

Additionally, to gain insights into the effectiveness of the new ML frameworks for subsurface site characterisation, it is recommended to conduct comparative research using a common benchmark. Although this might be challenging due to the nature and mostly closed access of the new ML methods, it is crucial to evaluate the benefits, drawbacks, and differences of these tools. SchemaGAN and the models presented in this paper are openly available to facilitate this process (Campos Montero, 2024).

CRediT authorship contribution statement

F.A. Campos Montero: Writing – original draft, Visualization, Software, Investigation, Formal analysis, Conceptualization, Methodology. **B. Zuada Coelho:** Writing – review & editing, Supervision, Resources, Conceptualization, Software. **E. Smyrniou:** Software, Conceptualization. **R. Taormina:** Writing – review & editing, Supervision, Conceptualization. **P.J. Vardon:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The code for the generation of the synthetic subsurface schematisations and the architecture of schemaGAN's components can be accessed in the repository https://github.com/fabcamo/schemaGAN. The training, testing and validation data, as well as the model weights, are open access and available at https://zenodo.org/records/13143431 (Campos Montero, 2024).

References

Azevedo, L., Paneiro, G., Santos, A., Soares, A., 2020. Generative adversarial network as a stochastic subsurface model reconstruction. Comput. Geosci. 24, 1673–1692. http://dx.doi.org/10.1007/s10596-020-09978-x.

Baghbani, A., Choudhury, T., Costa, S., Reiner, J., 2022. Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. Earth-Sci. Rev. 228, 103991. http://dx.doi.org/10.1016/j.earscirev.2022.103991.

Bau, D., Zhu, J.-Y., Wulff, J., Peebles, W., Strobelt, H., Zhou, B., Torralba, A., 2019. Seeing What a GAN Cannot Generate. http://dx.doi.org/10.48550/ARXIV.1910.11626.

Benny, Y., Galanti, T., Benaim, S., Wolf, L., 2021. Evaluation metrics for conditional image generation. Int. J. Comput. Vis. 129 (5), 1712–1731. http://dx.doi.org/10.1007/s11263-020-01424-w.

Borji, A., 2018. Pros and Cons of GAN Evaluation Measures. arXiv:1802.03446.

Borji, A., 2022. Pros and cons of GAN evaluation measures: New developments. Comput. Vis. Image Underst. 215, 103329. http://dx.doi.org/10.1016/j.cviu.2021.103329.

Bowles, C., Chen, L., Guerrero, R., Bentley, P., Gunn, R., Hammers, A., Dickie, D.A., Hernández, M.V., Wardlaw, J., Rueckert, D., 2018. GAN Augmentation: augmenting training data using generative adversarial networks. arXiv:1810.10863.

Campos Montero, F.A., 2023. Deep Learning for Geotechnical Engineering, The Effectiveness of Generative Adversarial Networks in Subsoil Schematization (Master's thesis). Delft University of Technology, Delft.

Campos Montero, F., 2024. SchemaGAN: Dataset and pre-trained model for geotechnical schematisation. http://dx.doi.org/10.5281/ZENODO.13143431.

Chen, H., He, X., Teng, Q., Sheriff, R.E., Feng, J., Xiong, S., 2020. Super-resolution of real-world rock microcomputed tomography images using cycle-consistent generative adversarial networks. Phys. Rev. E 101, 023305. http://dx.doi.org/10.1103/PhysRevE.101.023305.

Ching, J., Phoon, K.-K., Yang, Z., Stuedlein, A.W., 2022. Quasi-site-specific multivariate probability distribution model for sparse, incomplete, and three-dimensional spatially varying soil data. Georisk: Assess. Manag. Risk Eng. Syst. Geohazards 16 (1), 53–76. http://dx.doi.org/10.1080/17499518.2021.1971256.

Clayton, C.R.I., Matthews, M.C., Simons, N.E., 1995. Site Investigation, second ed. Blackwell Science, Oxford [England] ; Cambridge, Mass., USA.

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A., 2018. Generative adversarial networks: An overview. IEEE Signal Process. Mag. 35, 53–65. http://dx.doi.org/10.1109/MSP.2017.2765202.

de Gast, T., Vardon, P.J., Hicks, M.A., 2021. Assessment of soil spatial variability for linear infrastructure using cone penetration tests. Géotechnique 71 (11), 999–1013. http://dx.doi.org/10.1680/jgeot.19.SiP.002.

Demir, U., Unal, G., 2018. Patch-based image inpainting with generative adversarial networks. arXiv:1803.07422.

DHPC, 2022. DelftBlue supercomputer (phase 1). URL: https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase1.

Fenton, G.A., 1999. Random field modeling of CPT Data. J. Geotech. Geoenviron. Eng. 125 (6), 486–498. http://dx.doi.org/10.1061/(ASCE)1090-0241(1999)125:6(486).

Gao, Y., Liu, L., Zhang, C., Wang, X., Ma, H., 2020. SI-AGAN: spatial interpolation with attentional generative adversarial networks for environment monitoring. In: 24th European Conference on Artificial Intelligence. Santiago de Compostela, Spain.

Goh, A.T.C., Zhang, W., Zhang, Y., Xiao, Y., Xiang, Y., 2018. Determination of earth pressure balance tunnel-related maximum surface settlement: A multivariate adaptive regression splines approach. Bull. Eng. Geol. Environ. 77 (2), 489–500. http://dx.doi.org/10.1007/s10064-016-0937-8.

Goodfellow, I.J., 2017. NIPS 2016 tutorial: Generative adversarial networks. CoRR abs/1701.00160.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. In: Adaptive Computation and Machine Learning, The MIT Press, Cambridge, Massachusetts.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial networks.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2020. Generative adversarial networks. Commun. ACM 63, 139–144. http://dx.doi.org/10.1145/3422622.

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152, 195–207. http://dx.doi.org/10.1016/j.geoderma.2009.06.003.

Heuvelink, G.B., Webster, R., 2022. Spatial statistics and soil mapping: A blossoming partnership under pressure. Spat. Stat. 50, 100639. http://dx.doi.org/10.1016/j.spasta.2022.100639.

Hudson, K.S., Ulmer, K.J., Zimmaro, P., Kramer, S.L., Stewart, J.P., Brandenberg, S.J., 2023. Unsupervised machine learning for detecting soil layer boundaries from cone penetration test data. Earthq. Eng. Struct. Dyn. 52 (11), 3201–3215. http://dx.doi.org/10.1002/eqe.3961.

Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A., 2017. Image-to-Image Translation with Conditional Adversarial Networks. IEEE, pp. 5967–5976. http://dx.doi.org/10.1109/CVPR.2017.632.

Jaksa, M., 2000. Geotechnical risk and inadequate site investigations: A case study. Aust. Geomech. 35, 39–46.

Jaksa, M.B., Goldsworthy, J.S., Fenton, G.A., Kaggwa, W.S., Griffiths, D.V., Kuo, Y.L., Poulos, H.G., 2005. Towards reliable and effective site investigations. Géotechnique 55 (2), 109–121. http://dx.doi.org/10.1680/geot.2005.55.2.109.

Janssens, N., Huysmans, M., Swennen, R., 2020. Computed tomography 3D super-resolution with generative adversarial neural networks: Implications on unsaturated and two-phase fluid flow. Materials 13, 1397. http://dx.doi.org/10.3390/ma13061397.

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., Aila, T., 2020. Training generative adversarial networks with limited data. arXiv:2006.06676.

Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. http://dx.doi.org/10.48550/ARXIV.1412.6980.

Laloy, E., Hérault, R., Jacques, D., Linde, N., 2017. Training-image based geostatistical inversion using a spatial generative adversarial neural network. http://dx.doi.org/10.1002/2017WR022148.

Lam, N.S.-N., 1983. Spatial interpolation methods: a review. Am. Cartogr. 10 (2), 129–150. http://dx.doi.org/10.1559/152304083783914958.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. http://dx.doi.org/10.1038/nature14539.

Ledoux, H., Arroyo Ohori, K., Peters, R., Pronk, M., 2022. Computational modelling of terrains. http://dx.doi.org/10.5281/ZENODO.3992107.

Li, J., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Canberra.

Li, J., Madry, A., Peebles, J., Schmidt, L., 2017. On the limitations of first-order approximation in GAN dynamics. http://dx.doi.org/10.48550/ARXIV.1706.09884.

Li, D.-Q., Qi, X.-H., Cao, Z.-J., Tang, X.-S., Phoon, K.-K., Zhou, C.-B., 2016a. Evaluating slope stability uncertainty using coupled Markov chain. Comput. Geotech. 73, 72–82. http://dx.doi.org/10.1016/j.compgeo.2015.11.021.

Li, Z., Wang, X., Wang, H., Liang, R.Y., 2016b. Quantifying stratigraphic uncertainties by stochastic simulation techniques based on Markov random field. Eng. Geol. 201, 106–122. http://dx.doi.org/10.1016/j.enggeo.2015.12.017.

Lyu, B., Hu, Y., Wang, Y., 2023. Data-driven development of three-dimensional subsurface models from sparse measurements using Bayesian compressive sampling: A benchmarking study. ASCE- ASME J. Risk Uncertain. Eng. Syst. Part A: Civ. Eng. 9 (2), 04023010. http://dx.doi.org/10.1061/AJRUA6.RUENG-935, URL: https://ascelibrary.org/doi/10.1061/AJRUA6.RUENG-935.

Lyu, B., Wang, Y., Shi, C., 2024. Multi-scale generative adversarial networks (GAN) for generation of three-dimensional subsurface geological models from limited boreholes and prior geological knowledge. Comput. Geotech. 170, 106336. http://dx.doi.org/10.1016/j.compgeo.2024.106336.

McBratney, A., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. http://dx.doi.org/10.1016/S0016-7061(03)00223-4.

Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. CoRR abs/1411.1784.

Mosser, L., Dubrule, O., Blunt, M.J., 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys. Rev. E 96, 043309. http://dx.doi.org/10.1103/PhysRevE.96.043309.

Müller, S., Schüler, L., Zech, A., Heße, F., 2022. GSTools v1.3: a toolbox for geostatistical modelling in python. Geosci. Model. Dev. 15, 3161–3182. http://dx.doi.org/10.5194/gmd-15-3161-2022.

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., García, Á.L., Heredia, I., Malík, P., Hluchý, L., 2019. Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. Artif. Intell. Rev. 52, 77–124. http://dx.doi.org/10.1007/s10462-018-09679-z.

Oliveira, D.A.B., Ferreira, R.S., Silva, R., Brazil, E.V., 2019. Improving seismic data resolution with deep generative networks. IEEE Geosci. Remote. Sens. Lett. 16, 1929–1933. http://dx.doi.org/10.1109/LGRS.2019.2913593.

Oluwatuyi, O.E., Ng, K., Wulff, S.S., 2023. Improved resistance prediction and reliability for bridge pile foundation in shales through optimal site investigation plans. Reliab. Eng. Syst. Saf. 239, 109476. http://dx.doi.org/10.1016/j.ress.2023.109476.

Phoon, K.-K., Cao, Z.-J., Liu, Z., Ching, J., 2023. Report for ISSMGE TC309/TC304/TC222 Third ML dialogue on "Data-Driven Site Characterization (DDSC)": 3 December 2021, Norwegian Geotechnical Institute, Oslo, Norway (Online). Georisk: Assess. Manag. Risk Eng. Syst. Geohazards 17 (1), 227–238. http://dx.doi.org/10.1080/17499518.2022.2105366.

Phoon, K.-K., Ching, J., Shuku, T., 2022. Challenges in data-driven site characterization. Georisk: Assess. Manag. Risk Eng. Syst. Geohazards 16 (1), 114–126. http://dx.doi.org/10.1080/17499518.2021.1896005.

Phoon, K.-K., Ching, J., Wang, Y., 2019. Managing risk in geotechnical engineering — from data to digitalization. In: Proceedings of the 7th International Symposium on Geotechnical Safety and Risk. ISGSR 2019, Research Publishing Services, pp. 13–34. http://dx.doi.org/10.3850/978-981-11-2725-0-SL-cd.

Phoon, K.-K., Kulhawy, F.H., Grigoriu, M.D., 2000. Reliability-based design for transmission line structure foundations. Comput. Geotech. 26 (3–4), 169–185. http://dx.doi.org/10.1016/S0266-352X(99)00037-3.

Phoon, K.-K., Zhang, W., 2022. Future of machine learning in geotechnics. Georisk: Assess. Manag. Risk Eng. Syst. Geohazards 1–16. http://dx.doi.org/10.1080/17499518.2022.2087884.

Qi, X.-H., Li, D.-Q., Phoon, K.-K., Cao, Z.-J., Tang, X.-S., 2016. Simulation of geologic uncertainty using coupled Markov chain. Eng. Geol. 207, 129–140. http://dx.doi.org/10.1016/j.enggeo.2016.04.017.

Rahman, M.H., Abu-Farsakh, M.Y., Jafari, N., 2021. Generation and evaluation of synthetic cone penetration test (CPT) data using various spatial interpolation techniques. Can. Geotech. J. 58, 224–237. http://dx.doi.org/10.1139/cgj-2019-0745.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., Mohamed, S., 2021. Skilful precipitation nowcasting using deep generative models of radar. Nature 597, 672–677. http://dx.doi.org/10.1038/s41586-021-03854-z.

Robertson, P.K., 2010. Soil behaviour type from the CPT: An update. pp. 575–583.

Robertson, P.K., Cabal, K., 2022. Guide to Cone Penetration Testing, seventh ed. Gregg Drilling LLC, URL: https://www.cpt-robertson.com/PublicationsPDF/CPT-Guide-7th-Final-SMALL.pdf.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Springer International Publishing, Switzerland, pp. 234–241. http://dx.doi.org/10.1007/978-3-319-24574-4_28.

Salehi, P., Chalechale, A., Taghizadeh, M., 2020. Generative adversarial networks (GANs): An overview of theoretical model, evaluation metrics, and recent developments. arXiv preprint arXiv:2005.13178.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., 2016. Improved techniques for training GANs. arXiv:1606.03498.

Santurkar, S., Tsipras, D., Ilyas, A., Madry, A., 2018. How does batch normalization help optimization? http://dx.doi.org/10.48550/ARXIV.1805.11604.

Shi, C., Wang, Y., 2021. Development of subsurface geological cross-section from limited site-specific boreholes and prior geological knowledge using iterative Convolution XGBoost. J. Geotech. Geoenviron. Eng. 147 (9), 04021082. http://dx.doi.org/10.1061/(ASCE)GT.1943-5606.0002583.

Smyrniou, E., Zuada Coelho, B., 2023. Using generative adversarial networks to create a 2D subsoil schematization. In: 4th International Symposium of Machine Learning and Big Data in Geoscience. ISSMGE, University College Cork.

Srivastava, N., Hinton, G., Krizhevsky, A., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Sun, C., Demyanov, V., Arnold, D., 2023. Geological realism in fluvial facies modelling with GAN under variable depositional conditions. Comput. Geosci. http://dx.doi.org/10.1007/s10596-023-10190-w.

Tomczak, J.M., 2022. Deep Generative Modeling. Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-030-93158-2.

Valsecchi, A., Damas, S., Tubilleja, C., Arechalde, J., 2020. Stochastic reconstruction of 3D porous media from 2D images using generative adversarial networks. Neurocomputing 399, 227–236. http://dx.doi.org/10.1016/j.neucom.2019.12.040.

Wang, Y., Hu, Y., Zhao, T., 2020. Cone penetration test (CPT)-based subsurface soil classification and zonation in two-dimensional vertical cross section using Bayesian compressive sampling. Can. Geotech. J. 57 (7), 947–958. http://dx.doi.org/10.1139/cgj-2019-0131.

Wang, Y., Shi, C., Li, X., 2022. Machine learning of geological details from borehole logs for development of high-resolution subsurface geological cross-section and geotechnical analysis. Georisk: Assess. Manag. Risk Eng. Syst. Geohazards 16 (1), 2–20. http://dx.doi.org/10.1080/17499518.2021.1971254.

Wang, Y., Tian, H.-M., 2023. Digital geotechnics: From data-driven site characterisation towards digital transformation and intelligence in geotechnical engineering. Georisk: Assess. Manag. Risk Eng. Syst. Geohazards 1–25. http://dx.doi.org/10.1080/17499518.2023.2278136.

Wang, H., Wang, X., Wellmann, J.F., Liang, R.Y., 2019. A Bayesian unsupervised learning approach for identifying soil stratification using cone penetration data. Can. Geotech. J. 56 (8), 1184–1205. http://dx.doi.org/10.1139/cgj-2017-0709.

Wang, H., Wellmann, J.F., Li, Z., Wang, X., Liang, R.Y., 2017. A segmentation approach for stochastic geological modeling using hidden Markov random fields. Math. Geosci. 49 (2), 145–177. http://dx.doi.org/10.1007/s11004-016-9663-9.

Wang, Y., Zhao, T., 2017. Statistical interpretation of soil property profiles from sparse data using Bayesian compressive sampling. Géotechnique 67 (6), 523–536. http://dx.doi.org/10.1680/jgeot.16.P.143, URL: https://www.icevirtuallibrary.com/doi/10.1680/jgeot.16.P.143.

Wu, S., Zhang, J.-M., Wang, R., 2021. Machine learning method for CPTu based 3D stratification of New Zealand geotechnical database sites. Adv. Eng. Inform. 50, 101397. http://dx.doi.org/10.1016/j.aei.2021.101397.

Yan, L., Tang, X., Zhang, Y., 2021. High accuracy interpolation of DEM using generative adversarial network. Remote. Sens. 13 (4), 676. http://dx.doi.org/10.3390/rs13040676.

Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y., Ding, X., 2021. Application of deep learning algorithms in geotechnical engineering: a short critical review. Artif. Intell. Rev. 54, 5633–5673. http://dx.doi.org/10.1007/s10462-021-09967-1, URL: https://link.springer.com/10.1007/s10462-021-09967-1.

Zhao, T., Hu, Y., Wang, Y., 2018. Statistical interpretation of spatially varying 2D geo-data from sparse measurements using Bayesian compressive sampling. Eng. Geol. 246, 162–175. http://dx.doi.org/10.1016/j.enggeo.2018.09.022, URL: https://linkinghub.elsevier.com/retrieve/pii/S0013795218306495.

Zhou, C., Ouyang, J., Ming, W., Zhang, G., Du, Z., Liu, Z., 2019. A stratigraphic prediction method based on machine learning. Appl. Sci. 9 (17), 3553. http://dx.doi.org/10.3390/app9173553.

Zhu, D., Cheng, X., Zhang, F., Yao, X., Gao, Y., Liu, Y., 2020. Spatial interpolation using conditional generative adversarial neural networks. Int. J. Geogr. Inf. Sci. 34 (4), 735–758. http://dx.doi.org/10.1080/13658816.2019.1599122.