Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

doi: 10.3850/978-981-18-5182-7_05-014-cd

ISGSR 2022

Geological Uncertainty Quantification Using Image Warping and Bayesian Machine Learning

Hui Wang[1] and Xingxing Wei[2]

[1]Department of Civil and Environmental Engineering, The University of Dayton, Dayton, OH, USA.
E-mail: hwang12@udayton.edu
[2]School of Civil Engineering, Central South University, Changsha, China.
E-mail: xingxingwei15@163.com

Abstract: Quantifying the uncertainty of stratigraphic condition is an essential task in geotechnical projects. However, delineating and simulating heterogeneous stratigraphic profiles (non-stationary field), such as tectonically distorted or irregularly deposited strata from limited borehole information is still an open question and challenging task in engineering geology. In this study, a novel approach that applies the image warping technique to non-stationary field and combines it with an advanced stratigraphic stochastic simulation model is proposed to address this challenge. The image warping technique is effective to transform non-stationary field into stationary field based on the thin plate splines warping algorithm. Subsequently, an in-house developed stratigraphic stochastic simulation model can be applied to the transferred stationary field. The developed stratigraphic simulation approach integrates the Markov random field (MRF) model and the discriminant adaptive nearest neighbor-based k-harmonic mean distance (DANN-KHMD) classifier into a Bayesian framework to efficiently estimate the stratigraphic uncertainty given sparse site exploration results. To demonstrate the effectiveness of the developed approach, a synthetic case is studied using the developed approach. We envision this approach can be further promoted in industry practices for an improved risk control in geotechnical engineering.

Keywords: Non-stationary field; image warping technique; Markov random field; discriminant adaptive nearest neighbor-based k-harmonic mean distance; stratigraphic uncertainty; Bayesian machine learning.

1 Introduction

Obtaining accurate site-specific stratigraphic interpretation with quantified uncertainty is a crucial and essential step for planning and designing geotechnical systems. However, the subsurface soil layer configuration at a specific project site can be difficult to infer due to the limited site-specific data, let it alone to quantify the associated uncertainty (Wang et al., 2016), especially for sites with extremely complex strata such as tectonically distorted or irregularly deposited strata.

Substantial works (Fenton, 1999; Miranda et al., 2009) have been conducted to delineate deterministic soil stratification for geotechnical design. Some stochastic modeling frameworks, such as Markov chain model (Qi et al., 2016), multiple-point geostatistics (Fadlelmula et al., 2014) and iterative convolutional XGBoost (Shi and Wang 2021), have been developed in recent years. These models are able to generate stratigraphic realizations with quantified uncertainty, however, they have certain strong assumptions (such as stationary transition probability matrices or predefined pattern templates) that may be violated by natural soil configurations.

Recently, Markov random field (MRF)-based stochastic simulation approaches have been applied in engineering geology (Wang et al., 2016), due to the MRF models can provide a flexible and intuitive way to reflect and reproduce the anisotropy and heterogeneity of subsurface geological structures. While the existing MRF models (Wang et al., 2016; Wang et al., 2017; Wang et al., 2019) have a certain drawback: (1) the model parameters are subjectively defined in priori, besides they are fixed during the entire inferential process without consideration of model bias/uncertainty; (2) the current approaches may generate unrealistic soil profiles if there is no proper regularization (which is non-trivial) on model parameters; (3) the current models are unable to cope with non-stationary field such as tectonically distorted or irregularly deposited strata.

In this study, a novel approach that applies the image warping technique to non-stationary field and combines it with an advanced stratigraphic stochastic simulation model is proposed. The image warping technique is effective to transform the complex and intractable non-stationary field into a much simpler and tractable stationary field based on the thin plate splines warping algorithm (Bookstein, 1989). Subsequently, an in-house developed stratigraphic stochastic simulation model, that integrates the Markov random field (MRF) model and the discriminant adaptive nearest neighbor-based *k*-harmonic mean distance (DANN-KHMD) classifier into a Bayesian framework, can be applied to the transferred stationary field to efficiently estimate the stratigraphic uncertainty given sparse site exploration results.

2 Method

The flowchart of the developed approach is shown in Figure 1. Detailed explanations regarding each component are provided blow.

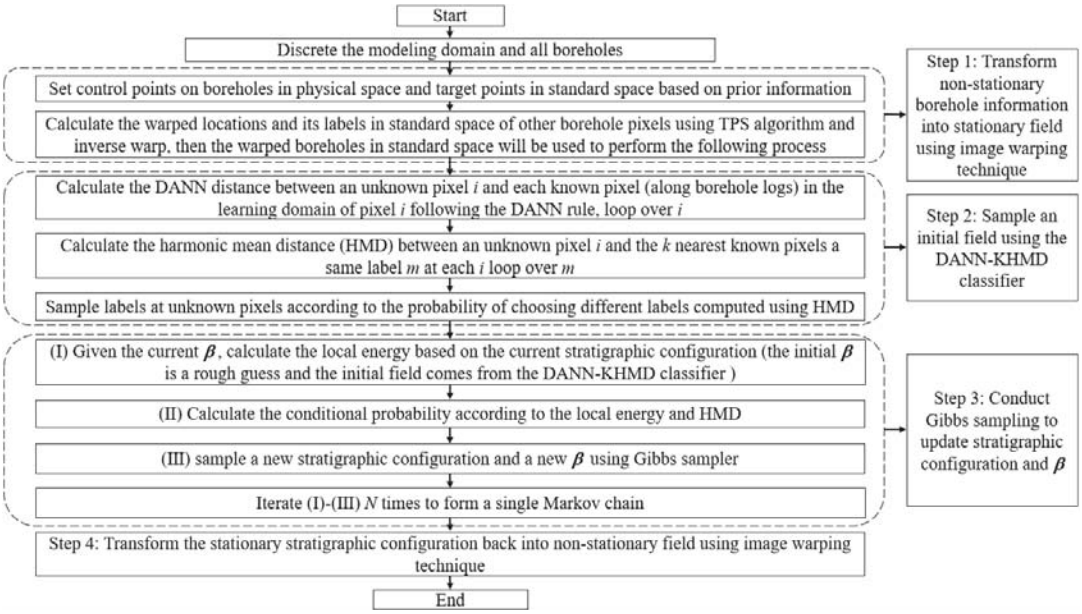


Figure 1. The flowchart of the developed approach.

2.1 Markov random field

An MRF model is a probabilistic graphical description of a spatial pattern defined on a pixel array. Pixels having the same label belong to the same soil type. Formally, the label field can be represented as a label configuration of all pixels $\boldsymbol{x} = \{x_i, i \in \mathbf{S}, x_i \in \mathbf{L}\}$, where $\mathbf{S} = \{1, 2, 3, \ldots, s\}$ indicates all pixels and $\mathbf{L} = \{1, 2, 3, \ldots, l\}$ is a set of all possible labels indicating different soil types.

A typical example of a graph model describing the spatial correlation of an MRF is the two-dimensional lattice with a second-order neighborhood system (Besag, 1986). The local neighborhood system ∂_i of pixel i consists of the nearest eight pixels around it. The local conditional probability $P(x_i \mid \boldsymbol{x}_{\partial_i})$ of a specific label given the labels of all neighbors can be calculated in the following form (Besag, 1986):

$$P(x_i \mid \boldsymbol{x}_{\partial_i}) = \frac{P(x_i, \boldsymbol{x}_{\partial_i})}{\sum_{x_i \in L} P(x_i, \boldsymbol{x}_{\partial_i})} = \frac{\exp[-U(x_i, \boldsymbol{x}_{\partial_i})]}{\sum_{x_i \in L} \exp[-U(x_i, \boldsymbol{x}_{\partial_i})]} \tag{1}$$

To compute the local energy $U(x_i, \boldsymbol{x}_{\partial_i})$, we adopt the widely used Potts model (Koller and Friedman, 2009) to characterize the local contextual interaction. Potts model has the following form:

$$U(x_i, \boldsymbol{x}_{\partial_i}) = V_i(x_i) + \sum_{j \in \partial_i} V_{i,j}(x_i, x_j) \tag{2}$$

where $V_i(x_i)$ is the potential function defined solely on pixel i indicating the preference of choosing different labels at pixel i and $V_{i,j}(x_i, x_j)$ is the potential function reflecting the local contextual interaction between neighboring pixels and defined as:

$$V_{i,j}(x_i, x_j) = \begin{cases} 0 & \text{if } x_i = x_j \\ \beta_d & \text{if } x_i \neq x_j \end{cases} \tag{3}$$

where $\beta_d \in \boldsymbol{\beta} = \{\beta_1, \beta_2, \beta_3, \beta_4\}$ indicates the contextual constraint, corresponding to one of the four independent directions (i.e., 0, π/2, π/4, 3π/4), within a two-dimensional lattice grid, and is referred to as a granularity coefficient. The behavior of an MRF model is intimately related to the granularity coefficients $\boldsymbol{\beta}$. More detailed information for expounding $V_{i,j}(x_i,x_j)$ and $V_i(x_i)$ can be found in the authors' previous work (Wei and Wang, 2022; Wang and Wei, 2022). For $V_i(x_i)$, we adopt the non-homogeneous version that depends on the pixel location, and characterize $V_i(x_i)$ via the DANN-KHMD classifier following Wei and Wang (2022).

2.2 Bayesian machine learning

All pixels in a two-dimensional array can be divided into two groups: a) pixels with known labels indicating sparse borehole information $\boldsymbol{x}_{BH}$, and b) pixels with unknown labels $\boldsymbol{x}_{unknown}$ elsewhere. Both $\boldsymbol{x}_{unknown}$ and the granularity coefficients $\boldsymbol{\beta}$ need to be inferred from $\boldsymbol{x}_{BH}$. A Markov chain Monte Carlo (MCMC) method is employed to sample ($\boldsymbol{x}_{unknown}$, $\boldsymbol{\beta}$) using two conditional a posteriori distributions $P(\boldsymbol{x}_{unknown} \mid \boldsymbol{x}_{BH}, \boldsymbol{\beta})$ and $P(\boldsymbol{\beta} \mid \boldsymbol{x}_{unknown}, \boldsymbol{x}_{BH})$ iteratively.

Given $\boldsymbol{x}_{BH}$ and $\boldsymbol{\beta}$, $P(\boldsymbol{x}_{unknown} \mid \boldsymbol{x}_{BH}, \boldsymbol{\beta})$ is a Gibbs distribution with fixed soil labels only at the borehole locations. The local energy at unknown pixels can be calculated using Eq. (2) and the corresponding conditional probability of choosing each label given the neighbors can be evaluated via Eq. (1). Then, given an initial filed sampled using the DANN-KHMD classifier, a realization of the conditional random field $P(\boldsymbol{x}_{unknown} \mid \boldsymbol{x}_{BH}, \boldsymbol{\beta})$ can be simulated via a parallel strategy named chromatic sampler (Wang et al., 2017).

During the iterative process, every time after a realization of $\boldsymbol{x}_{\text{unknown}}$ is simulated, $\boldsymbol{\beta}$ is sampled following the conditional posterior distribution:

$$\text{Post}(\boldsymbol{\beta}) \propto \text{Prior}(\boldsymbol{\beta}) L(\boldsymbol{x}_{BH} \mid \boldsymbol{\beta}, \boldsymbol{x}_{unknown}) \tag{4}$$

where $\text{Post}(\boldsymbol{\beta})$ and $\text{Prior}(\boldsymbol{\beta})$ are the posterior distribution and prior distribution of $\boldsymbol{\beta}$, respectively; $L(\boldsymbol{x}_{BH} \mid \boldsymbol{\beta}, \boldsymbol{x}_{unknown})$ is the likelihood function indicating the possibility of having the observed soil configuration given the simulated borehole information and $\boldsymbol{\beta}$. It can be evaluated via the following equation.

$$L(\boldsymbol{x}_{unknown}, \boldsymbol{x}_{BH} \mid \boldsymbol{\beta}) = \prod_{x_i \in \{\boldsymbol{x}_{unknown}, \boldsymbol{x}_{BH}\}} P(x_i \mid \boldsymbol{x}_{\partial_i}; \boldsymbol{\beta}) \tag{5}$$

In this work, *Prior*($\boldsymbol{\beta}$) is predefined by a multivariate Gaussian distribution with a mean vector $\boldsymbol{\mu}$ indicating the rough estimates of the granularity coefficients, and a diagonal covariate matrix $\sigma^2\mathbf{I}_4$, where $\boldsymbol{\sigma}$ is a predefined standard deviation of granularity coefficients and $\mathbf{I}_4$ is the identity matrix. $\boldsymbol{\mu}$ and $\sigma^2\mathbf{I}_4$.can be simply left as default or estimated roughly (less-informative prior) or customized if site-specific knowledge is available as *Prior*($\boldsymbol{\beta}$) is not sensitive to the final estimation results (Wei and Wang, 2022).

The Metropolis-Hasting algorithm is employed to update $\boldsymbol{\beta}$, and the log(target) function is expressed as

$$\log(\text{target}) = \log(\text{Prior}(\boldsymbol{\beta})) + \log(L(\boldsymbol{x}_{BH} \mid \boldsymbol{\beta}, \boldsymbol{x}_{unknown})) \tag{6}$$

The higher the log(target) is, the higher possibility that the boring logs can be observed and the corresponding granularity coefficients are compatible with the simulated field. In other words, Eq. (6) is being optimized during the Bayesian machine learning process via MCMC.

2.3 Discriminant adaptive nearest neighbor-based k-harmonic mean distance (DANN-KHMD) classifier

The DANN-KHMD classifier is essentially an approach to roughly "guess" possible labels of $\boldsymbol{x}_{unknown}$ given $\boldsymbol{x}_{BH}$ in a probabilistic manner. The local metric Ψ_i at unknown pixel *i* is evaluated over a rectangular learning domain $\aleph_i$, which is centered at *i* and includes all pixels in the rectangular region.

The DANN distance $D(j, i)$ between pixels *j* and the center unknown pixel *i* can be calculated using the local metric Ψ_i (Hastie and Tibshirani 1996):

$$D(j,i) = (\boldsymbol{v}_j - \boldsymbol{v}_i)^{\mathbf{T}} \Psi_i (\boldsymbol{v}_j - \boldsymbol{v}_i) \tag{7}$$

in which $\boldsymbol{v}_i = [c_i, r_i]^{\mathrm{T}}$ are the coordinates of pixel *i*, where *c* and *r* are the column and row indices, respectively. The metric Ψ_i is defined as follow (Hastie and Tibshirani 1996):

$$\Psi_i = \mathbf{W}^{-1/2}\left[\mathbf{W}^{-1/2}\mathbf{B}\mathbf{W}^{-1/2} + \varepsilon\mathbf{I}\right]\mathbf{W}^{-1/2} \tag{8}$$

where ε is a tuning parameter. **W** and **B** are the within-label covariance matrix and the between-label covariance matrix, respectively. They are calculated using known pixels *j* in $\aleph_i$. Detailed information for demonstrating the DANN distance can be found in the authors' previous work (Wei and Wang, 2022).

Then, the harmonic mean distance (HMD) is introduced to measure the mean distance between a subset of $\boldsymbol{x}_{BH}$ having a certain label in terms of $D(j, i)$ with the formulation below (Pan et al., 2017):

$$HMD_i^{(m)} = \frac{k}{\sum_{j=1}^{k} \frac{1}{D(j,i)}}, x_j = m \tag{9}$$

where $HMD_i^{(m)}$ is the HMD of the k nearest known pixels having label m to the unknown pixel i. $HMD_i^{(m)}$ indicates the local preference of label m. In Eq. (2), $V_i(x_i)$ is the potential function defined solely on pixel i and indicates the preference of choosing different labels at pixel i, it is reasonable to set $V_i(m) = HMD_i^{(m)}$. Accordingly, the local probabilities $Pro_i^{(m)}$ of choosing label m for pixel i can be calculated with the form below following Besag (1986) and Geman and Geman (1984):

$$Pro_i^{(m)} = \frac{\exp(-HMD_i^{(m)})}{\sum_{m' \in \mathbf{L}} \exp(-HMD_i^{(m')})} \tag{10}$$

The initial stratigraphic configuration can be sampled directly and independently via Eq. (10) at each unknown pixel. In order to hold the local label preference characterized by $HMD_i^{(m)}$ throughout the stochastic simulation, the HMD is integrated into the local energy and the conditional probability Eq. (1) for a local neighborhood system is updated as:

$$P(x_i \mid \boldsymbol{x}_{\partial_i}) = \frac{\exp[-(HMD_i^{(x_i)} + \sum_{j \in \partial_i} V_{i,j}(x_i, x_j))]}{\sum_{x_i' \in \mathbf{L}} \exp[-(HMD_i^{(x_i')} + \sum_{j \in \partial_i} V_{i,j}(x_i', x_j))]} \tag{11}$$

Accordingly, the likelihood function (i.e., Eq. (5)) is updated as:

$$L(\boldsymbol{x}_{unknown}, \boldsymbol{x}_{BH} \mid \boldsymbol{\beta}) = \prod_{x_i \in \{\boldsymbol{x}_{unknown}, \boldsymbol{x}_{BH}\}} \frac{\exp[-(HMD_i^{(x_i)} + \sum_{j \in \partial_i} V_{i,j}(x_i, x_j))]}{\sum_{x_i' \in \mathbf{L}} \exp[-(HMD_i^{(x_i')} + \sum_{j \in \partial_i} V_{i,j}(x_i', x_j))]} \tag{12}$$

The sensitivity of a few parameters including k and the parameters for learning domain $\aleph_i$ have been analyzed in the authors' previous work (Wei and Wang, 2022). The setting of these parameters follows Wei and Wang, 2022.

2.4 Image warping technique

Given an image with a sparse set of control points $\boldsymbol{v}_i = [c_i, r_i]^{\mathrm{T}}$ with corresponding displacements $\boldsymbol{d}_i = [\Delta c_i, \Delta r_i]^{\mathrm{T}}$, we can find a mapping $f: \boldsymbol{v}_i \to \boldsymbol{v}_i'$ from pixels in the input image to pixels in the warped/deformed image so that the corresponding warped target points $\boldsymbol{v}_i' = [c_i', r_i']^{\mathrm{T}}$ closely match its expected targets $[c_i + \Delta c_i, r_i + \Delta r]^{\mathrm{T}}$, and the surrounding pixels are deformed as smoothly as possible. Accordingly, the thin plate spline (TPS) warping algorithm is employed for image warping problem due to it can provide a smooth interpolation in both column and row directions using the following smooth function (Bookstein, 1989):

$$f\left(c_i', r_i'\right) = a_0 + a_1 c + a_2 r + \sum_{i=1}^{N} \omega_i G\left(\left\|\left(c_i, r_i\right) - \left(c, r\right)\right\|\right) \tag{13}$$

where $G(e) = e^2 \log(e)$. Besides, additional constraints are needed for the system following Donato and Belongie (2003):

$$\sum_{i=1}^{N} \omega_i = \sum_{i=1}^{N} \omega_i c_i = \sum_{i=1}^{N} \omega_i r_i = 0 \tag{14}$$

Having Eq. (13) and (14), a linear system can be yielded for the TPS coefficients (Donato and Belongie, 2003):

$$\begin{bmatrix} K & P \\ P^T & O \end{bmatrix} \times \begin{bmatrix} \boldsymbol{\omega} \\ \boldsymbol{a} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}' \\ \boldsymbol{o} \end{bmatrix} \tag{15}$$

where $K_{ij} = G\left(\left\|\left(c_i, r_i\right) - \left(c_j, r_j\right)\right\|\right)$, the ith row of P is $\left(1, c_i, r_i\right)$, O is a 3×3 matrix of zeros, $\boldsymbol{o}$ is a 3×1 column vector of zeros, $\boldsymbol{\omega}$ and $\boldsymbol{v}$ are column vectors formed from ω_i and v_i, respectively, and $\boldsymbol{a}$ is the column vector with elements $\left(a_0, a_1, a_2\right)$.

Given a few control points ($\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ are known) and the corresponding warped control points (i.e., target points, $\boldsymbol{v}'$ is known), the matrices K and P can be calculated directly. Therefore, the vector $\left[\boldsymbol{\omega} \quad \boldsymbol{a}\right]^T$ containing the weight ω_i for each control point *i* and coefficients $\left(a_0, a_1, a_2\right)$ can be solved according to Eq. (15) . Subsequently, the warped point $v' = \left(c', r'\right)$ of any point $v = \left(c, r\right)$ in the input image can be computed via the smooth equation Eq. (13) given a few control points $v_i = \left(c_i, r_i\right)$. The procedure is called forward warp. While there are some gaps during forward warp such as many integer pixels are not assigned labels due to the warped positions of these pixels computed via the smooth equation are often non-integer. Accordingly, the inverse warp method is employed in the warping of stratigraphic profile.

Unlike the forward warp, the matrices K and P are computed with the target points and the vector $\left[\boldsymbol{v}' \quad \boldsymbol{o}\right]^T$ is a vector with respect to the control points. Therefore, the vector $\left[\boldsymbol{\omega} \quad \boldsymbol{a}\right]^T$ is for the transformation of warped control points to control points. Accordingly, for each pixel in the profile to be obtained, its corresponding position in the original profile can be found using the inverse warp ideology. Hence, there is no gap in the warped profile as each pixel is labeled.

2.5 Uncertainty quantification

The maximum a posteriori (MAP) of soil profile can be determined using the following majority vote principle:

$$\mathrm{MAP}(i) = \arg\max_m \left(P_i^{(m)}; m \in \mathbf{L}\right) \tag{16}$$

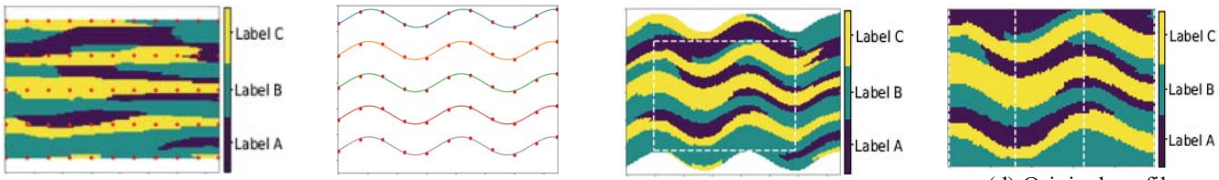
in which $P_i^{(m)}$ is the marginal probability of soil label *m* at pixel *i*. In order to quantify the uncertainty of the estimated soil labels at each pixel, the information entropy (IE) quantifying the uncertainty at a given pixel *i* is adopted and expressed as the formulation below (Li et al. 2016):

$$IE_i = -\sum_{m \in L}\left[P_i^{(m)} \log P_i^{(m)}\right] \tag{17}$$

The higher IE_i is, the higher uncertainty level it is at the given pixel *i*, and hence it will be more difficult to infer the actual soil label.

3 Case Study

The synthetic case is derived from an automatically generated "soil profile" that has size 100×100 pixels, using a Gibbs sampler with the granularity coefficients β =[4.50, 0.15, 0.15, 0.15]. First, put the generated soil profile in a larger space (see Figure 2a) so as to provide space for its deformation. Five series of control points (see red points in Figure 2a) are set on the profile and the corresponding target points lie on five sinusoids (see Figure 2b). Then the warped soil profile (see Figure 2c) can be acquired via image warping technique. The part (see Figure 2d) extracted from the white dotted box (see Figure 2c) is considered as the "ground truth" (original profile) throughout the sinusoidal case. Four evenly distributed virtual boreholes annotated by the dashed lines are extracted as shown in Figure 2d. Note that the soil profile deformation process can be used as prior information that combines with four extracted boreholes are the information used to infer the "unknown" pixels. In real practice project, the prior information of stratigraphic deformation can be caught according to the geological knowledge acquired by such as geological survey programs, experienced geologist, recorded geological data.

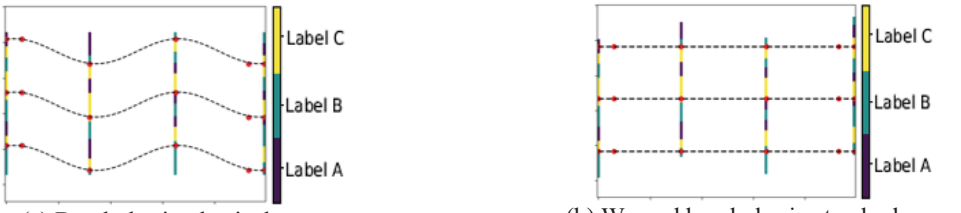


(a) Soil profile generated using Gibbs sampler (b) Five sinusoids (c) Warped soil profile (d) Original profile

Figure 2. The deformation process of soil profile.

3.1 Warping of borehole information

According to the prior information, the stratigraphic pattern in the area of modeling domain deforms into approximately sinusoidal in depositional history.



(a) Boreholes in physical space (b) Warped boreholes in standard space

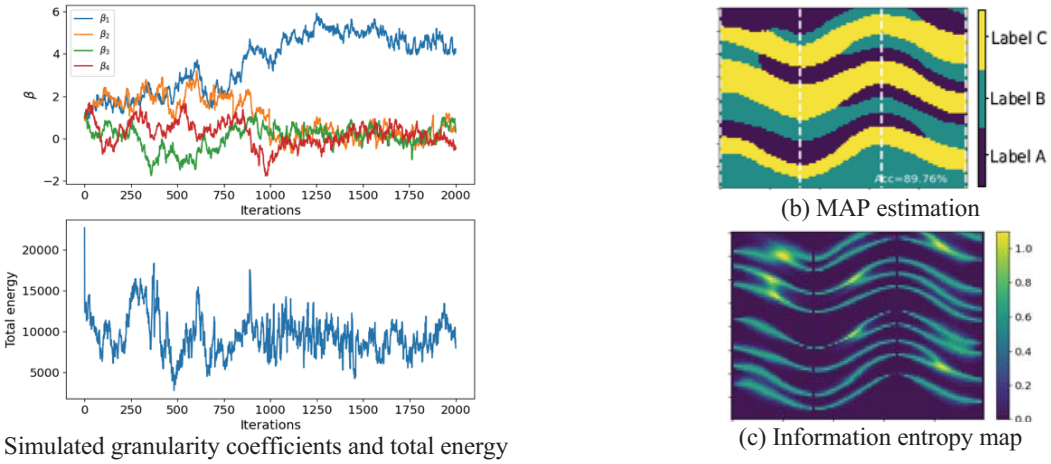
Figure 3. Boreholes and warped boreholes.

Then three series of control points are setup on borehole pixels and each series of control points is on the same sinusoidal, which can be seen in Figure 3a (red dots on boreholes). Moreover, a few additional control points (see the red dots not on boreholes in Figure 3a) used to adjust sinusoidal tangents at both sides of the modeling domain are set to be horizontal with and adjacent to the boundary control points as the sinusoidal tangents are horizontal on two boundaries that known from the prior information. Figure 3b shows the target points in standard space. The borehole information, in standard space, warped using the locations of control points and target points via image warping technique is also shown in Figure 3b.

3.2 Simulations results

Subsequently, the stochastic simulation is implemented using the warped boreholes in standard space (see Figure 3b). Two thousand MCMC iterations are performed. The simulated trace of $\boldsymbol{\beta}$ and total energy of each realization are shown in Figure 4a. As we can see, for this specific simulation, the chains of β_1, β_2, β_3, β_4 and total energy are convergent after 1200 iterations. The trace indicates that the convergence can be reached even with sparse boreholes and little or no prior information of $\boldsymbol{\beta}$.

The MAP estimation of the simulated soil profile is shown in Figure 4b, which is warped using image warping technique from the soil profile inferred via the warped borehole information (see Figure 3b). The accuracy of MAP estimation is 89.76%. The corresponding IE map in Figure 4b reflects the uncertainty level at a per-pixel basis. Generally speaking, higher information entropy can be found around the boundary of neighboring layers, while low information entropy exists inside different simulated layers. This result agrees with our intuitions. It shows that the proposed approach can well infer the soil profile and quantify the stratigraphic uncertainty in non-stationary field only use sparse boreholes (say, only 4% known pixels in this case).



(a) Simulated granularity coefficients and total energy (b) MAP estimation (c) Information entropy map

Figure 4. Simulation results.

4 Conclusion

In this study, a novel approach that implements the image warping technique to generate non-stationary field and combine it with an advanced stratigraphic stochastic simulation model is proposed. The proposed method aims at improving the performance of the developed MRF model for stratigraphic interpretation and uncertainty quantification, in non-stationary field such as tectonically distorted or irregularly deposited strata from, in the situation that only sparse boreholes are available. A synthetic example has been used to demonstrate the effectiveness and to validate the proposed method. It has been shown that the proposed approach can well interpret the soil profile and quantify the associated uncertainty with sparse borehole information in non-stationary field.

Acknowledgement

This work is partially supported by Central South University under the Project Number 1053320192341. This work is also partially supported by the Ohio Department of Transportation under the Agreement Number 31795, and by the STEM Catalyst grant from the University of Dayton. The financial supports from all sponsors are gratefully acknowledged.

References

Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society: Series B (Methodological)*, *48*(3), 259-279.

Bookstein, F. L. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence*, *11*(6), 567-585.

Donato, G., and Belongie, S. (2003). Approximation methods for thin plate spline mappings and principal warps.

Fadlelmula, M. M., Akin, S., and Duzgun, S. (2014). Parameterization of Channelized Training Images: A Novel Approach for Multiple-Point Simulations of Fluvial Reservoirs. *In Mathematics of Planet Earth (pp. 557-560). Springer, Berlin, Heidelberg*.

Fenton, G. A. (1999). Estimation for stochastic soil models. *Journal of Geotechnical and Geoenvironmental Engineering*, *125*(6), 470-485.

Hastie, T., and Tibshirani, R. (1996). Discriminant adaptive nearest neighbor classification. *IEEE transactions on pattern analysis and machine intelligence*, *18*(6), 607-616.

Koller, D., and Friedman, N. (2009). Probabilistic graphical models: principles and techniques. *MIT press*.

Li, Z., Wang, X., Wang, H. and Liang, R.Y. 2016c. Quantifying stratigraphic uncertainties by stochastic simulation techniques based on Markov random field. *Engineering Geology*, 201, 106-122.

Miranda, T., Correia, A. G., and e Sousa, L. R. (2009). Bayesian methodology for updating geomechanical parameters and uncertainty quantification. *International Journal of Rock Mechanics and Mining Sciences*, *46*(7), 1144-1153.

Qi, X. H., Li, D. Q., Phoon, K. K., Cao, Z. J., and Tang, X. S. (2016). Simulation of geologic uncertainty using coupled Markov chain. *Engineering geology*, *207*, 129-140.

Shi, C. and Wang, Y. 2021. Development of Subsurface Geological Cross-Section from Limited Site-Specific Boreholes and Prior Geological Knowledge Using Iterative Convolution XGBoost. *Journal of Geotechnical and Geoenvironmental Engineering*, **147**, 04021082.

Wang, H., and Wei, X. Stochastic Stratigraphic Simulation and Uncertainty Quantification Using Machine Learning. *In Geo-Congress 2022 (pp. 337-346)*.

Wang, H., Wellmann, J. F., Li, Z., Wang, X., and Liang, R. Y. (2017). A segmentation approach for stochastic geological modeling using hidden Markov random fields. *Mathematical Geosciences*, 49(2), 145-177.

Wang, X., Li, Z., Wang, H., Rong, Q., and Liang, R. Y. (2016). Probabilistic analysis of shield-driven tunnel in multiple strata considering stratigraphic uncertainty. *Structural safety*, 62, 88-100.

Wang, X., Wang, H., Liang, R. Y., and Liu, Y. (2019). A semi-supervised clustering-based approach for stratification identification using borehole and cone penetration test data. *Engineering Geology*, 248, 102-116.

Wei, X., and Wang, H. (2022). Stochastic stratigraphic modeling using Bayesian machine learning. *Engineering Geology*, 106789.