A data driven CPTU-PINNs method to predict soil settlement

Haoxi Li[1], *Zhongkun* Ouyang[2#]

[1&2]*Tsinghua Shenzhen International Graduatate School,*
University Town of Shenzhen, Nanshan District, Shenzhen 518055 P.R. China
[#]*Corresponding author: ouyangzk@sz.tsinghua.edu.cn*

ABSTRACT

Obtaining soil parameters through laboratory tests and solving the governing equations that describe soil settlement can be time-consuming, making immediate on-site predictions of soil settlement challenging. In-situ testing provides a more efficient approach to obtain soil parameters than laboratory tests. Data from the Piezocone penetration test (CPTu) can be used for on-the-spot interpretation of soil mechanical parameters, which can then be incorporated into the governing equations for soil settlement calculation. Physics Informed Neural Networks (PINNs) algorithm uses automatic differentiation method to directly embed partial differential equations (PDEs) into a deep learning neural network and provides solution for these PDEs in a cost-effective manner compared to traditional numerical methods. In this paper, a framework integrating data from CPTu and PINNs to predict soil settlement is proposed and evaluated through comparison with numerical simulations from Finite Element Methods (FEMs). Results show that the framework gave a reasonably good agreement with the FEMs benchmark while substantially reduced the computation time. This method allows for immediate on-site prediction of soil settlement during site investigations, thus better guiding surveying and construction activities.

Keywords: CPTU; SCPTU; PINNs; Settlement; 2D settlement; Biot' s theory.

1. Introduction

The occurrence of unexpected consolidation and settlement in soil can pose risks and hazards to engineering practices. Like most geotechnical engineering problems, theories of soil consolidation involving complex nonlinear relationships often described by partial differential equations (PDEs), such as the one-dimensional consolidation theory (Terzaghi 1925), the Terzaghi-Rendulic three-dimensional consolidation theory (Rendulic 1936), and Biot's three-dimensional consolidation theory (Biot 1941). Solving PDEs for most engineering problems analytically under various initial and boundary conditions is challenging. Therefore, common methods for solving PDEs are numerical, like finite difference methods (FDMs) (Shwetank et al. 2023) and finite element methods (FEMs) (Sandhu and Wilson 1969; Ferronato 2010; Rodríguez et al. 2023). These numerical approaches often require discretization of calculations, and the best balance between the efficiency and accuracy of calculation depend on adjusting grid density and computational step size.

Neural networks serve as universal function approximators (Hornik et al. 1989) and, as a mesh-free algorithm, its computation accuracy is independent of the step size. Theoretically, neural networks can accurately solve all types of PDEs with sufficient training data. Furthermore, the recently developed Physics-Informed Neural Networks (PINNs) algorithm informs physical knowledge to introduce physics-driven constraints into neural networks (Raissi et al. 2019). It leverages automatic differentiation (Baydin et al. 2018), an underutilized yet powerful technique in scientific computing to formulate loss functions according to the governing PDEs. This dual-drive approach reduces the demand for extensive training data. Raissi et al. (2019) demonstrated the research value and application capabilities of PINNs by solving the Schrödinger equation and the Korteweg-de Vries equation.

However, certain physical parameters in engineering problems PDEs need to be determined before conducting PINNs algorithm on PDEs. Commonly, corresponding laboratory tests are conducted to obtain these parameters, but the in-situ test, such as the piezocone penetration test (CPTu), provides another efficient, cost-effective, and convenient method. Geotechnical practitioners can stratify subsurface layers or interpret geological parameters through analysing the measurements of CPTu test (Robertson 2009). Some of the detailed specifications for the equipment, test procedures, and methods for interpreting CPTu results are provided by Lunne et al. (1997), Mayne (2007), Schnaid (2008), and Mayne (2023). A standard CPTu involves the static penetration of the cone into the soil at a constant rate of 20 mm/s, accessing three parameters that vary continually with depth, including: measured cone tip resistance (q_c), sleeve friction (f_s), and porewater pressure at the shoulder (u_2) schematically shown in Fig. 1. In the case of clayey soils, the parameter q_c is converted to total cone tip resistance (q_t) using q_t=q_c+ (1-a_{net}) *u_2, where a_{net} is defined as net area ratio (Campanella and Robertson 1988). Furthermore, some other CPTu includes additional equipment that enable the measurement of other soil properties. For example, the Seismic Cone Penetration Test with pore pressure measurements (SCPTu) can determine the seismic shear

wave velocity (V_s) of the soil (Schneider et al. 2000). This can be utilized to interpret additional soil parameters.

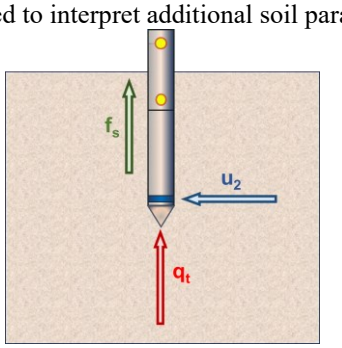

Figure 1. CPTU test and soundings.

This study proposes a framework that integrates CPTu and PINNs to predict the muti-dimensional settlement of soil without prior investigation and laboratory tests. This framework initially conducts a CPTu test on the soil, and by interpreting the measurements of the CPTu test, it determines the relevant physical parameters involved in the theory describing the settlement of the soil. Subsequently, the governing equations, the initial and the boundary conditions of the soil with known physical parameters will be inputted into PINNs for mesh-free solving. This method combines the advantages of CPTu, which provides a rapid and cost-effective on-site continuous depth testing of soil, with the fast and convenient PDEs solving capabilities of PINNs. As a result, it enables quick on-site predictions of settlement for soil over large scale space after fully consolidation without the time-consuming and complex steps of obtaining soil parameters through laboratory tests and numerical methods in traditional approaches.

2. Physical theory base

In soil mechanics, Biot's theory is a three-dimensional consolidation theory for soils based on the principle of effective stress (Biot 1941). This theory, assuming small strains in the soil, with the pore fluid being saturated and incompressible, establishes a system of equations that combines the equilibrium equation within the framework of effective stress principle, stress-strain relationships of the soil skeleton, deformation compatibility conditions, and continuity conditions. By solving these equations, it provides insights into the evolution of stress, strain, and pore pressure in the soil under loading conditions. It is also capable of predicting and describing the Mandel-Cryer effect (Mandel 1953; Cryer 1963) and is theoretically rigorous. In this study, it is assumed that the soil behaves as a linear elastic material, whereas the volume of soil particles remains constant, the volume of voids changes and the pore fluid follows Darcy's law. Under these assumptions, the settlement of the soil and the pore water pressure can be described by the Biot consolidation theory as follows:

$$-G\nabla^2 u_S - \frac{G}{1-2v}\left(\frac{\partial^2 u_S}{\partial x^2} + \frac{\partial^2 v_S}{\partial x \partial y}\right) + \frac{\partial u}{\partial x} = 0$$

$$-G\nabla^2 v_S - \frac{G}{1-2v}\left(\frac{\partial^2 u_S}{\partial x \partial y} + \frac{\partial^2 v_S}{\partial y^2}\right) + \frac{\partial u}{\partial y} = \gamma$$

$$\frac{k}{\gamma_w}\nabla^2 u + \frac{\partial}{\partial t}\left(\frac{\partial u_S}{\partial x} + \frac{\partial v_S}{\partial y}\right) = 0 \qquad (1.1\text{-}1.3)$$

where, G is the small-strain shear modulus of the linear elastic material, v is the Poisson's ratio, u_S and v_S are displacements in the 2 directions, u is the pore water pressure, γ is the saturated unit weight of the soil, k is the coefficient of permeability, and γ_w is the unit weight of the fluid (water in this study) in the porous medium.

3. PINNs algorithm

In the context of the PINNs algorithm, a neural network is needed to be constructed as an approximator for the solutions to PDEs. The chosen neural network model in this study is the Multilayer Perceptron (MLP), known for its simplicity and clarity (Bishop 1995). The MLP structure comprises an input layer, an output layer, and hidden layers, with no connections within each layer, but fully connected between adjacent layers in a unidirectional forward pass, as in Fig. 2. The computation process of MLP proceeds from the input layer through the hidden layers to the output layer and the computation for each neuron is defined as (Pinkus 1999):

$$x_{i,j} = \sigma_{i,j}\left(\boldsymbol{w}_{i-1,j} * \boldsymbol{x}_{i-1} + b_{i,j}\right) \qquad (2)$$

where, $x_{i,j}$ represents the output of the j-th neuron in the i-th layer, $\boldsymbol{w}_{i-1,j}$ is the weight matrix linking all the neurons in the $(i-1)$-st layer to the j-th neuron in the i-th layer, $\boldsymbol{x}_{i-1}$ is the output matrix of the neurons in the $(i-1)$-st layer, $b_{i,j}$ is the bias of the j-th neuron in the i-th layer, and $\sigma_{i,j}$ is the activation function of the j-th neuron in the i-th layer. These parameters collectively form the neural network internal parameter set $\boldsymbol{\theta} = \{\boldsymbol{w}, \boldsymbol{b}\}$. One of the sources of nonlinear ability of the neural network is the multiple hidden layers, the other one is the nonlinear activation functions, which makes the fitting ability of neural network stronger. The common activation functions include the Sigmoid function, the Hyperbolic Tangent (Tanh) function, the ReLU function and its improvements, and the ELU function (Ding et al. 2018). A well-constructed neural network should include an appropriate number of the hidden layers, with the careful selection of activation functions and the quantity of neurons within each layer. The output of the output layer is denoted as $\mathbf{u}(\boldsymbol{\theta}, \mathbf{x})$, where x represents the input to the input layer.

The necessary steps of training a neural network involve feeding it a suitable database comprising N sets of inputs $\mathbf{x} = \{x_1, x_2, x_3 \ldots x_n\}$ (also referred to as "feature" in deep learning) and the corresponding "correct" outputs $\mathbf{u}(\mathbf{x}) = \{u_1, u_2, u_3 \ldots u_m\}$ (also referred to as "label"). For an untrained neural network which means its internal parameters set θ has not adjusted based on this database, the output of the neural network result $\mathbf{u_{nn}}(\boldsymbol{\theta}, \mathbf{x})$ will exhibit residuals compared to the labels from the database. The residuals are quantitatively described by a loss function, commonly formulated using the Mean Squared Error (MSE) function in mathematics:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{u}_{\boldsymbol{nn}}^{\boldsymbol{i}} - \boldsymbol{u}^{\boldsymbol{i}}\right|^2 \quad (3)$$

Building upon this foundation, the loss function will be minimized and converged using a gradient descent optimizer. This process, referred to as the training of the network, is achieved by adjusting the parameters θ of the network through the optimizer. From the dataset being passed forward along the neural network until the optimizer finishes adjusting θ, it is called an epoch for training. There will be many epochs until the optimizer finishes a reduction in the loss function as people want.

Common gradient descent optimizers include the Adam algorithm (Kingma and Ba 2014), and the L-BFGS algorithm (Byrd et al. 1995). The Adam algorithm terminates when the neural network training reaches a specified number of iterations, impacting its computational precision but offering a fast convergence rate. On the other hand, the L-BFGS algorithm terminates the iteration when the relative change in the loss function is below a set threshold, demonstrating superior performance for smaller training datasets.

In general, the mathematical form of PDEs with physical information (governing equations, initial conditions, and boundary conditions) can be unified as:

$$\left(\boldsymbol{u}; \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x_1}} \cdots \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x_n}}; \frac{\partial^2 \boldsymbol{u}}{\partial \boldsymbol{x_1}\partial \boldsymbol{x_1}} \cdots \frac{\partial^2 \boldsymbol{u}}{\partial \boldsymbol{x_1}\partial \boldsymbol{x_n}}; \cdots; \boldsymbol{\lambda}\right) = 0 \quad (4)$$

where, $\mathbf{x}$ (=$\{x_1, x_2, x_3 \ldots x_n\}$) are the variables of the PDEs, $\boldsymbol{\lambda}$ (= $\{\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_p\}$) are the parameters, and $\mathbf{u}$ = ($\{u_1, u_2, u_3 \ldots u_m\}$) are the solutions. For each PDE, a certain number of points are selected on each computational domain where each PDE is defined, referred as residual points. $\mathbf{x}$ (as features), $\mathbf{u}$ (as labels), and $\boldsymbol{\lambda}$ of those residual points are used as the database for constructing the residual for that corresponding PDE. By leveraging automatic differentiation, the residual of the particular PDE with physical information can be individually formulated using the mathematical form of MSE with the parameters, features and labels from residual points. For each PDE, the above steps are performed. And the entire database for training PINNs is the sum of the data provided by residual points for each PDE. Then, sum those MSEs up with timing suitable weights τ_j, the loss function that is informed all the physical information is constructed:

$$\sum_{j=1}^{M}\frac{\tau_j}{N}\sum_{i=1}^{N}\left|\boldsymbol{u}_{\boldsymbol{nn}}^{\boldsymbol{ij}}; \frac{\partial \boldsymbol{u}_{\boldsymbol{nn}}^{\boldsymbol{ij}}}{\partial \boldsymbol{x}^{\boldsymbol{ij}}}; \frac{\partial^2 \boldsymbol{u}_{\boldsymbol{nn}}^{\boldsymbol{ij}}}{\partial \boldsymbol{x}^{\boldsymbol{ij}}\partial \boldsymbol{x}^{\boldsymbol{ij}}}; \cdots; \boldsymbol{\lambda}^{\boldsymbol{ij}}\right|^2 \quad (5)$$

Unlike the original loss function only simply reflecting residuals between $\mathbf{u_{nn}}$ and $\mathbf{u}$ as in Eq. (3), the loss function informed physics information goes beyond the conventional purely data-driven neural network training. The MSEs, incorporating physical information, achieves the regularization of embedding physics information into the neural network. It constrains the neural network training with known physical information. This makes PINNs satisfy the corresponding physical regulations during the training process, which enhances the training speed of the model and reduces the demands on neural network complexity.

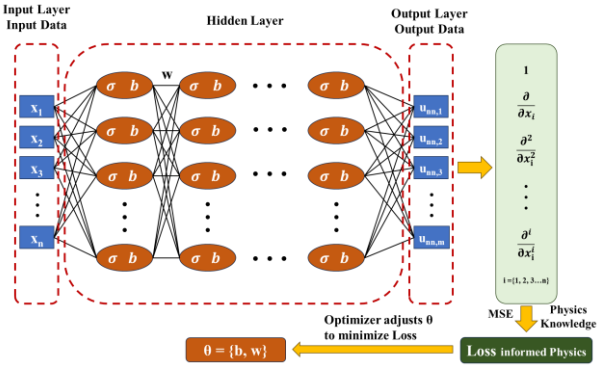

Figure 2. Structural diagram of PINNs.

4. CPTU interpretation theories

The soil parameters the elastic modulus E (or the small-strain shear modulus G_{max}), Poisson's ratio v, the coefficient of permeability k, and the soil unit weight γ contained in the Eq. (1) applied in this study to characterize soil settlement, can all be interpreted with the measurements (q_t, f_s, u_2, and V_s) of the CPTU (SCPTU) test. CPTU tests is a fast and inexpensive method of acquiring the soil parameters, serving as a key step for achieving in situ prediction of soil settlement.

The small-strain shear modulus G_{max} can be calculated from elasticity theory using the total soil mass density ρ and shear wave velocity V_s (Mayne 2023):

$$G_{max} = \rho \cdot V_s^2 \quad (6)$$

where $\rho = \gamma/g$, where gravitational acceleration constant $g = 9.8m/s^2$. According to the mechanics of materials, E and G_{max} can be interconverted incorporating Poisson's ratio v, while Poisson's ratio v can be calculated through the relationship with the at-rest lateral stress coefficient (K_0) (Federico and Elia 2009):

$$K_0 = \frac{v}{1-v} \quad (7)$$

K_0 can be estimated as $K_0 = (1 - sin\,\varphi') \cdot OCR^{sin\,\varphi'}$ (Kulhawy and Mayne 1990), where φ' is effective friction angle, and OCR ($= \sigma_p'/\sigma_{v_0}'$) is the overconsolidation ratio.

The well recognized CPTu interpretation theory of the effective friction angle φ' is proposed and calibrated by Ouyang and Mayne (2018, 2019):

$$\varphi' = 29.5B_q^{0.121} \cdot \left[0.256 + 0.336 \cdot B_q + logQ'\right] \quad (8)$$

where $B_q(= \frac{u_2 - u_0}{q_t - \sigma_{v0}})$ is normalized pore-water pressure parameter and the modified cone resistance number Q' was defined as:

$$Q' = \frac{q_t - \sigma_{v_0}}{\sigma_{v_0}' \cdot OCR^{\Lambda}}, \text{ when OCR} \geq 2.5$$
$$Q' = \frac{q_t - \sigma_{v_0}}{\sigma_{v_0}'}, \text{ when } 0 > \text{OCR} > 2.5 \quad (9)$$

The effective vertical stress σ_{v_0}' is determined by the total unit weight and the ground water level, and the total vertical stress σ_{v0} is only determined by the soil unit weight. Λ is plastic volumetric strain potential. The effective preconsolidation stress σ_p' can be estimated by a general estimation theory proposed by Mayne (2017) and Agaiby and Mayne (2019):

$$\sigma_p' = 0.33(q_t - \sigma_{v0})^m \quad (10)$$

where the exponent parameter m values can be assigned as: m = 0.72 in clean quartz sands, 0.8 in silty sands, 0.85 in silts, 0.9 in organic clays, and 1.0 in intact clays of low sensitivity (Agaiby and Mayne 2019).

A quick estimation theory of soil unit weight γ is proposed by the studies that used data from offshore sites (Mayne and Peuchen 2013) and onshore location (Mayne 2014):

$$\gamma = \gamma_w \cdot [1.22 + 0.345 \cdot log_{10}(10^2(f_s/\sigma_{atm}) + 10^{-2})] \quad (11)$$

where, σ_{atm} is atmospheric pressure (kPa), approximately equals to 100kPa in common.

The normalized CPT soil behavior type chart (SBTn) method including three normalized CPTu measurements: Q, F, and B_q, was suggested by Robertson 1990, where $F(=\frac{f_s}{q_t-\sigma_{v_0}} \cdot 100\%)$ was defined as the normalized parameter of the sleeve friction f_s.

Jefferies and Davies 1993 identified a Soil Behavior Type Index (I_c) that can represent the soil behavior type zones in the SBTn. Regarding the calculation of I_c, Robertson 2009 defined the updated normalized cone resistance $Q_{tn}(=\frac{q_t-\sigma_{v_0}}{\sigma_{atm}}\left(\frac{\sigma_{atm}}{\sigma_{v_0}'}\right)^m)$ and presented the now commonly used formula for I_c based on the relationship with F:

$$I_c = [(3.47 - logQ)^2 + (logF + 1.22)^2]^{0.5} \quad (12)$$

where the parameter n varies with soil type and is calculated as $n = 0.381 \cdot I_c + 0.05\left(\frac{\sigma_{v_0}'}{\sigma_{atm}}\right) - 0.15 \leq 1$.

In the above equations, it can be observed that there is a mutual dependence between I_c and n. Therefore, it is necessary to iterate from n=1, calculate Q_{tn}, I_c and n twice, and determine the convergence of the iterative calculation of them. Then, the calculated Q_{tn} and F can be used for soil classification based on the SBTn.

Lunne et al. 1997 provided estimated values and ranges of soil permeability for each type of soil in SBTn. On this basis, Robertson and Cabal (2010) proposed an algorithm that relates the coefficient of permeability k in terms of I_c for soil zones 2 to zone 7 in SBTn:

$$\begin{aligned} k(m/s) &= 10^{0.952-3.04 \cdot I_c};\ 1.0 < I_c \leq 3.27 \\ k(m/s) &= 10^{-4.52-1.37 \cdot I_c};\ 3.27 < I_c \leq 4.0 \end{aligned} \quad (13)$$

5. Case study

5.1. CPTU data and interpretation

To illustrate the CPTU-PINNs method for predicting soil settlement, a field case study is presented using SCPTU data from Mount Pleasant in Chelston, USA, as reported by Mayne and Woeller (2008). At a depth of approximately 15-20m at this location, there is a thick marine deposit consisting of calcareous clay named as Cooper marl, which is typically the bearing layer for local engineering structures (Camp 2004). Therefore, in this case study, the multi-settlement prediction for the soil is focused on the Cooper marl.

A series of in-situ tests and laboratory tests were conducted to investigate the site for constructing a bridge spanning the Cooper River. Specifically, a total of 55 SCPTU tests were conducted. The representative piezocone data (q_t, f_s, and u_2) as shown in Fig. 3 are from No. 31 SCPTU test, while V_s is averaged from 5 nearby SCPTs. The Cooper marl layer has a high calcite content ranging from 60 to 80%, an average plasticity index (PI) of 38%, a typical water content (w_n) of 48%, and a clay content of 78% (Camp et al. 2002).

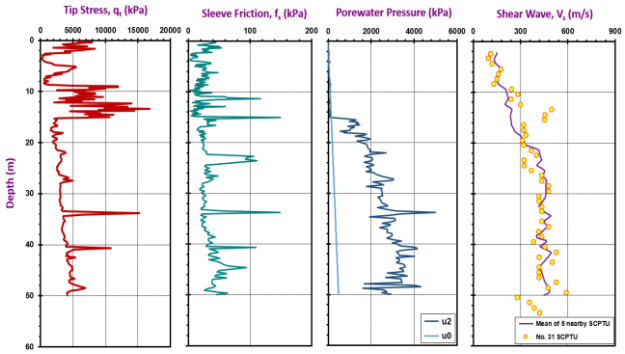

Figure 3. No. 31 SCPTU data at the Mount Pleasant Shaft MP-1 (data from Mayne and Woeller 2008).

Its depth ranges from 20m to 50m where the porewater pressures sounding of SCPTU are high. Above the Cooper marl layer, there is a fill layer with low porewater pressure soundings, ranging from 0m to 20m in depth. Other essential geotechnical parameters for the test site, such as unit weight, are interpreted from SCPTU test data using the aforementioned interpretation theories, the average values of them are as shown in Table 1.

Table 1. Parameters interpreted from SCPTU data

γ	OCR	φ'	K_0
17.2kN/m^3	2.21	44.3°	0.89
E	ν	I_c	**k**
118Mpa	0.47	2.68	1.84*10^{-7} m/s

Typically, a Poisson's ratio ν of 0.47 is not entirely accurate and may not be reflective of engineering reality. In certain depths, significant interpreted values may arise due to disturbances or errors in CPTu testing, leading to an elevated mean of interpreted ν. However, as Poisson's ratio is not a primary determinant in this case study and other parameters exhibit no significant issues with the use of mean values, the interpretation criterion for the Poisson's ratio remains consistent with other parameters, employing a mean value of 0.47.

While still focusing on the performance of the CPTU-PINNs framework on prediction for the settlement during the multi-dimensional consolidation of soil, in this case study the consolidation problem is simplified to the free drainage (except the loading area) consolidation process of a semi-infinite homogeneous, isotropic, linearly elastic strip of Cooper marl with a thickness of 10m and a width of 10m under a semi-infinite strip instant loading with uniform load that does not vary with time as illustrated in Fig. 4. Therefore, the consolidation problem is simplified to a two-dimensional stress-strain problem coupled with pore pressure, and the soil parameters interpreted from

SCPTU data will be averaged. The boundary conditions are detailed as:

$$u\,(0,y,t) = u\,(10,y,t) = 0$$

$$u_S\,(0,y,t) = u_S\,(10,y,t) = 0$$

$$u\,(x,10,t) = u_S\,(x,10,t) = v_S\,(x,10,t) = 0$$

$$u\,(0{\sim}4,0,t) = u\,(6{\sim}10,0,t) = 0 \qquad (14.1\text{-}14.4)$$

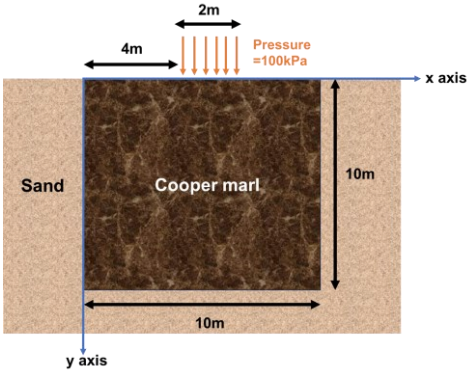

Figure 4. Simplified Cooper marl two-dimensional consolidation problem under free drainage conditions.

5.2. Finite element methods

To evaluate the accuracy of PINNs in solving the two-dimensional consolidation problem of Cooper marl, the results are compared with numerical simulation results from FDMs.

The symmetrically loaded 2D consolidation problem of Cooper marl in the case study is modeled using a 1/2 model in finite element software. The Cooper marl model in the simulation analysis is 10m thick and 5m wide, simplified as an isotropic, uniformly elastic, two-dimensional porous material with constant parameters. The material has a linear elastic modulus (E) of 118 MPa, a Poisson's ratio (v) of 0.47, a dry density of 804.35 kg/m³ (dry unit weight of 7400 N/m³), a cofficient of permeability (k) of $1.84*10^{-7}$ m/s ($6.624*10^{-4}$m/h). The pore fluid is assumed to be pure water, resulting in a fluid density of 9800 N/m³. Symmetric boundary conditions required for the 1/2 model, drainage and displacement boundary conditions described by Eq. (14), instantaneous uniformly distributed loads, and the gravitational field of the Earth's stress are all applied to the model. Predefined fields of pore pressure are also applied according to the gravity direction to ensure that the model represents a saturated soil model.

When seeding for meshing in the load region ($0m \leq x \leq 1m$) along the x-axis, 15 seeds are set every meter, and in the non-load region ($0m < x \leq 1m$), 10 seeds are set every meter. Along the y-axis, in the region $0m \leq y \leq 4m$, 15 seeds are set every meter, and in the region $4m < y \leq 10m$, 10 seeds are placed every meter. A total of 6600 four-node plane strain quadrilateral elements, with bilinear displacement and bilinear pore pressure (CPE4P), are used to mesh the soil, as shown in Fig. 5. With this model, a simulation of the consolidation process for 100 hours is conducted, outputing the 2D distribution of displacement and pore pressure after consolidation, as shown in Fig. 6. These results are used for comparison with the outcomes from PINNs.

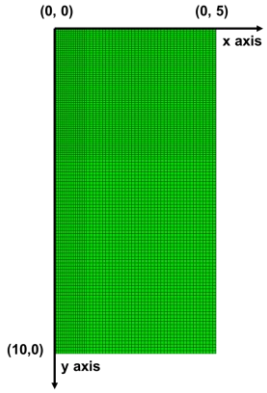

Figure 5. Finite element model of the Cooper marl 2D consolidation problem.

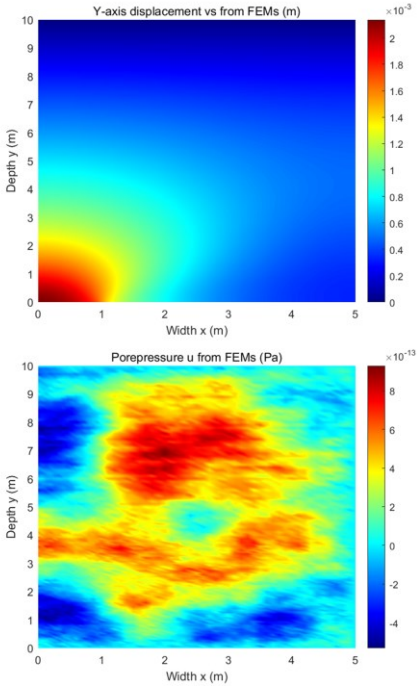

Figure 6. Spatial distribution cloud plots of vertical displacement (v_S) and pore pressure (u) after 100 hours of consolidation from FEMs (y-axis is positive upwards).

5.3. PINNs methods

The governing equations for the 2D consolidation problem of Cooper marl are given by Eq. (1), and the boundary conditions are described by Eq. (14). Based on these equations, the PDEs are nondimensionalized and normalized using the interpreted parameters as aforementioned as follows:

$$-1.8\left(\frac{\partial^2 \overline{us}}{\partial \bar{x}^2}+\frac{\partial^2 \overline{us}}{\partial \bar{y}^2}\right)-30\frac{\partial}{\partial \bar{x}}\left(\frac{\partial \overline{us}}{\partial \bar{x}}+\frac{\partial \overline{vs}}{\partial \bar{y}}\right)+\frac{\partial \bar{u}}{\partial \bar{x}}=0$$

$$-1.8\left(\frac{\partial^2 \overline{vs}}{\partial \bar{x}^2}+\frac{\partial^2 \overline{vs}}{\partial \bar{y}^2}\right)-30\frac{\partial}{\partial \bar{y}}\left(\frac{\partial \overline{us}}{\partial \bar{x}}+\frac{\partial \overline{vs}}{\partial \bar{y}}\right)+\frac{\partial \bar{u}}{\partial \bar{y}}=\gamma$$

$$\frac{1.84*1.18*3.6*5}{9.8*5.3}\nabla^2\bar{u}+\frac{\partial}{\partial \bar{t}}\left(\frac{\partial \overline{us}}{\partial \bar{x}}+\frac{\partial \overline{vs}}{\partial \bar{y}}\right)=0 \qquad (15.1{\sim}15.3)$$

$$\bar{u}\,(0,\bar{y},\bar{t}) = \bar{u}\,(1,\bar{y},\bar{t}) = 0$$

$$u_S\,(0,\bar{y},\bar{t}) = u_S\,(1,\bar{y},\bar{t}) = 0$$

$$\bar{u}\,(\bar{x},1,\bar{t}) = u_S\,(\bar{x},10,\bar{t}) = v_S\,(\bar{x},1,\bar{t}) = 0$$

$$\bar{u}\,(0{\sim}0.4,0,\bar{t}) = \bar{u}\,(0.6{\sim}1,0,\bar{t}) = 0 \qquad (16.1{\sim}16.4)$$

where, $\bar{x} = \frac{x}{10m}$, $\bar{y} = \frac{y}{10m}$, $\bar{t} = \frac{t}{3600s*50}$ are the dimensionless and normalized space and time coordinates in the consolidation process, $\bar{u} = \frac{u}{10000Pa}$, $\overline{us} = \frac{us}{\frac{5.3}{1180}m}$, $\overline{vs} = \frac{vs}{\frac{5.3}{1180}m}$ are the dimensionless and normalized pore water pressure and displacements. Eq. (15) & (16) together constitute the nondimensionalized and normalized PDEs describing the consolidation problem in the study case.

As shown in Fig. 2, a neural network is initially constructed with $\bar{\boldsymbol{x}}$, $\bar{\boldsymbol{y}}$, $\bar{\boldsymbol{t}}$ as inputs and $\overline{\boldsymbol{u}_{\boldsymbol{s}_{\boldsymbol{nn}}}}$, $\overline{\boldsymbol{v}_{\boldsymbol{s}_{\boldsymbol{nn}}}}$, $\overline{\boldsymbol{u}_{\boldsymbol{nn}}}$ as outputs, with internal parameters as $\boldsymbol{\theta}$. This neural network serves as a numerical approximation for Eq. (15). Residual points are selected within the computational domain for Eq. (15) & (16), serving as training database to create specific assigned loss functions. Finally, a gradient descent optimizer is employed to adjust the neural network parameters $\boldsymbol{\theta}$, minimizing the loss function until the neural network is appropriately trained.

For the neural network in this case, there are 12 hidden layers, each consisting of 360 neurons. Each neuron includes a Tanh activation function and bias. The Tanh function is chosen for its infinite differentiability, and the outputs are within the range of -1 to 1:

$$Tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \qquad (17)$$

For the output layer, no activation function is set. In the choice of the gradient descent optimizer, the commonly used Adam optimizer is selected with an initial learning rate of 10^{-5}, decayed by 20% every 100 steps, and a total of 15,000 epochs. For the MSE of Eq. (15), there are 1000 residual points, and for the MSE of Eq. (16), there are 100 residual points, sampled using Latin hypercube sampling. The total loss function of the neural network can be summed up after giving proper weights τ to each MSE according to its magnitude and importance.

The numerical solutions obtained based on the PINNs algorithm are shown in Fig. 7, while the absolute errors compared to the FEMs solutions are illustrated in Fig. 8. After 100 hours consolidation is completed, the pore pressure is almost zero, and its small magnitude renders the spatial distribution analysis less meaningful. Nevertheless, PINNs still successfully achieve a magnitude of pore pressure comparable to FEMs. However, the key distinction between the Biot's theory employed in this study and the commonly used Terzaghi one-dimensional consolidation theory or Terzaghi-Rendulic multi-dimensional consolidation theory lies in its incorporation of the stress-strain relationships of the soil skeleton, enabling the prediction of soil deformation. Therefore, the predictive performance of PINNs in terms of the vertical displacement v_S of the soil is worthy of evaluation. In the PINNs results, v_S in the loaded region of the soil is much greater than in the unloaded region, with specific values close to the FEMs results, both at magnitude of 10^{-3}m while the errors are maintained at the smaller magnitude. This indicates that the neural network accurately captures the influence of the load, which is only present in a specific region in this case, and effectively handles the abrupt change in the load distribution along the x-axis. The results above indicate that PINNs, by solving Biot's theory, not only effectively captured the dissipation of pore pressure under fully drainage conditions but also demonstrated proficiency in directly predicting soil settlement. This capability allows for the assessment of potential soil deformations, providing insights into the likelihood of ground movements at the site.

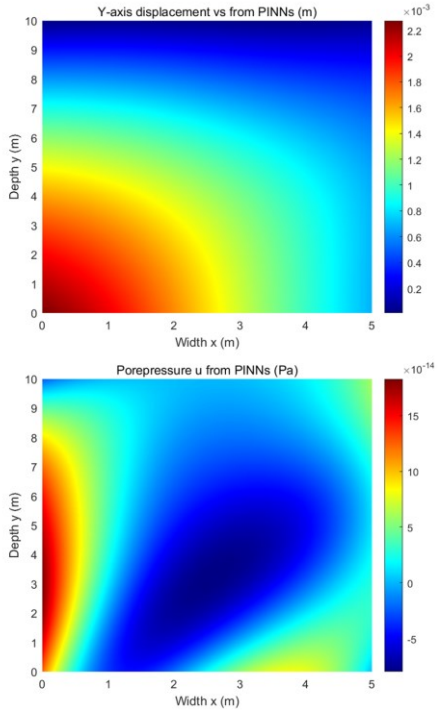

Figure 7. Spatial distribution cloud plots of vertical displacement (v_S) and pore pressure$v_S$after 100 hours of consolidation from PINNs(y-axis is positive upwards).

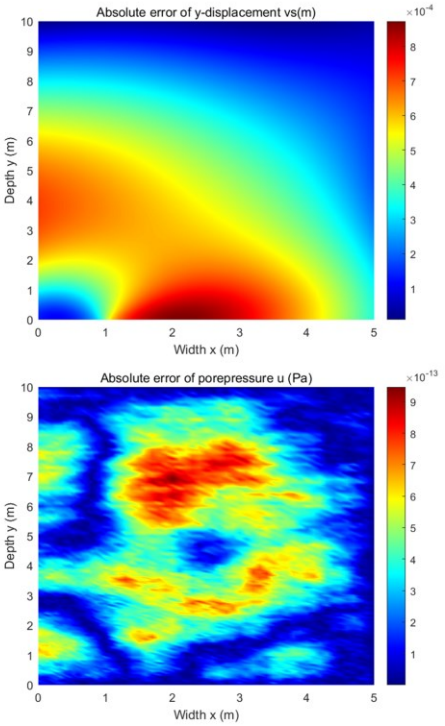

Figure 8. Spatial distribution cloud plots of the absolute error of vertical displacement (v_S) and pore pressure (u) between results from PINNs and FEMs after 100 hours of consolidation (y-axis is positive upwards).

6. Discussion

In the case study presented in this paper, to simplify the problem, the research object is assumed to be a homogeneous soil, and mean values interpreted from CPTU data are used in both FEMs and PINNs. However, one significant advantage of CPTU is it can provide continuous responses of soil properties along the depth, allowing for the interpretation of soil parameters that vary continuously with depth. The simplification in the case study weakens the advantage of CPTU. In future work, incorporating depth-varying soil parameters interpreted from CPTU into the PINNs framework should enhance the precision and reliability of the predictions made by the CPTU-PINNs framework.

Since this study is using the Biot‘s theory to describe the two-dimensional consolidation problems, the advantages are evident. Not only can it predict variations in pore water pressure, but it can also forecast soil strain. However, the drawback lies in the increased complexity of the PDEs describing this problem, including the governing equations and boundary conditions, compared to previous works, like Raissi et al. 2019, Lu et al. 2019, and Tartakovsky et al. 2020. This also requires the assignment of weights to more loss functions embedded with PDEs when forming the overall loss function. It is a more difficult Multi-Task Learning (MTL) problem. Properly assigning weights to those MSEs is not a trivial or insignificant task. Loss functions assigned with too small weights may fail to make the neural network adequately reflect the physical relationships inherent in the corresponding PDEs.

In this study, the evolution of each loss function corresponding to PDEs is monitored during training. Adjustments are made to the weights of those loss functions whose magnitudes differ significantly from others during the training process. In the field of deep learning, a common approach for MTL problems is to employ random weighting (Lin et al. 2022). However, since the magnitudes of loss functions for different PDEs can vary significantly, simply using random weighting might disproportionately diminish the weights of loss functions with inherently larger magnitudes. Therefore, there is much room for improvement in using PINNs to solve problems described by a greater number of PDEs, such as higher-dimensional consolidation problems.

7. Conclusions

This paper proposes a data-driven CPTU-PINNs method, incorporating Biot's theory to predict multi-dimensional soil settlement. The analysis is based on field data from an SCPTU test, and the main conclusions are as follows:

PINNs as a new deep learning algorithm that informs physical information into regularization, it has low requirements for the complexity of the neural networks, which also makes training faster. Additionally, the introduction of physical constraints in PINNs means that the training is not purely data-driven. This not only improves the fitting capability of the neural network to relevant problems but also reduces the demand for the size of training database. These advantages, compared to traditional data-driven neural networks, make PINNs less demanding on training equipment and more suitable for training on portable devices. Therefore, the collaboration between PINNs and CPTU is more efficient, enabling a swift process from in-situ testing to settlement prediction.

As an attempt to address the multi-task learning problem in PINNs, the results of the CPTU-PINNs framework in solving multi-dimensional consolidation problems under continuous drainage boundary conditions, as presented in the article, have successfully produced numerical results for soil strain and pore water pressure that have reference value. It means the CPTU-PINNs framework can be considered as a part of engineering reconnaissance that can rapidly and cost-effectively identify high-risk areas in engineering projects, thereby reducing the need for time-cost and expensive investigations such as drilling.

Although this study utilized Biot's theory for PINNs to solve consolidation problems, there were significant simplifications in the case study. Better and more general solutions for addressing the challenges of multi-dimensional consolidation problems in PINNs were not explored extensively. Additionally, the advantages of CPTU in interpreting continuous soil parameters in depth were not fully token in the presented case. These suggest that the CPTU-PINNs framework is a method with further research and practical value.

Acknowledgements

The authors are grateful for the support provided by Tsinghua Shenzhen International Graduate School.

References

Agaiby, S. S., and Mayne, P. W. 2019. “CPT Evaluation of Yield Stress Profiles in Soils.” *Journal of Geotechnical and Geoenvironmental Engineering*, *145*(12). https://doi.org/10.1061/(asce)gt.1943-5606.0002164

Biot, M. A. 1941. “General theory of three-dimensional consolidation.” *Journal of Applied Physics*, *12*(2). https://doi.org/10.1063/1.1712886

Bishop, C. M. 1995. *Neural networks for pattern recognition*. Oxford university press.

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. 1995. “A Limited Memory Algorithm for Bound Constrained Optimization.” *SIAM Journal on Scientific Computing*, *16*(5). https://doi.org/10.1137/0916069

Camp, W. M. 2004. *Drilled and Driven Foundation Behavior in a Calcareous Clay*. https://doi.org/10.1061/40713(2004)1

Camp, W. M., Mayne, P. W., and Brown, D. A. 2002. *Drilled Shaft Axial Design Values: Predicted Versus Measured Response in a Calcareous Clay*. https://doi.org/10.1061/40601(256)108

Campanella, R.G., and Robertson, P.K. 1988. “Current status of the piezocone test.” *Penetration Testing 1988.* Vol. 1, (Proc. ISOPT, Orlando), A.A. Balkema Publishing, Rotterdam: 93-116.

Cryer, C. W. 1963. “A comparison of the three-dimensional consolidation theories of biot and terzaghi.” *Quarterly Journal of Mechanics and Applied Mathematics*, *16*(4). https://doi.org/10.1093/qjmam/16.4.401

Ding, B., Qian, H., and Zhou, J. 2018. “Activation functions and their characteristics in deep neural networks.” *Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018*. https://doi.org/10.1109/CCDC.2018.8407425

Federico, A., and Elia, G. 2009. “At-rest earth pressure coefficient and Poisson’s ratio in normally consolidated soils.” *Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering: The Academia and Practice of Geotechnical Engineering*, *1*. https://doi.org/10.3233/978-1-60750-031-5-7

Ferronato, M., Castelletto, N., and Gambolati, G. 2010. “A fully coupled 3-D mixed finite element model of Biot consolidation.” *Journal of Computational Physics*, *229*(12). https://doi.org/10.1016/j.jcp.2010.03.018

Güneş Baydin, A., Pearlmutter, B. A., Andreyevich Radul, A., and Mark Siskind, J. 2018. “Automatic differentiation in machine learning: A survey.” In *Journal of Machine Learning Research* (Vol. 18).

Hornik, K., Stinchcombe, M., and White, H. 1989. “Multilayer feedforward networks are universal approximators.” *Neural Networks*, *2*(5). https://doi.org/10.1016/0893-6080(89)90020-8

Kingma, D. P., and Ba, J. L. 2015. “Adam: A method for stochastic optimization.” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.

Kulhawy, F. H., and Mayne, P. W. 1990. “Manual on Estimating Soil Properties for Foundation Design.” In *Ostigov*.

Lin, B., Ye, F., Zhang, Y., and Tsang, I. W. 2021. “Reasonable effectiveness of random weighting: A litmus test for multi-task learning. ” arXiv preprint arXiv:2111.10603.

Lu, L., Meng, X., Mao, Z., and Karniadakis, G. E. 2021. “DeepXDE: A deep learning library for solving differential equations.” *SIAM Review*, *63*(1). https://doi.org/10.1137/19M1274067

Lunne, T., Robertson, P., and Powell, J. 1997. “Cone Penetration Testing in Geotechnical Practice., 1997.” *Blackie Academic and Professional, ….*

Mandel, J. 1953. “Consolidation des sols (étude mathématique).” *Geotechnique*, *3*(7). https://doi.org/10.1680/geot.1953.3.7.287

Mayne, P. W. 2007. “National Cooperative Highway Research Program (NCHRP Synthesis 368) -- Cone Penetration Testing.” In *Nchrp*.

Mayne, P. W., and Woeller, D. 2008. “Evaluating O-cell response from seismic cone tests” *Proc. GeoEdmonton 2008*, (61th Canadian Geotechnical Conference), Canadian Geotechnical Society: www.cgs.ca

Mayne, P. W. 2014. “Interpretation of geotechnical parameters from seismic piezocone tests.” *3rd International Symposium on Cone Penetration Testing (CPT`14)*.

Mayne, P. W. 2017. “Stress history of soils from cone penetration tests.” *Soils and Rocks*, *40*(3). https://doi.org/10.28927/SR.403203

Mayne, P. W., and Peuchen, J. 2013. “Unit weight trends with cone resistance in soft to firm clays.” *Geotechnical and Geophysical Site Characterization 4 - Proceedings of the 4th International Conference on Site Characterization 4, ISC-4*, *1*.

Mayne, P. W. 2023. *The Cone Penetration Test: A CPT Design Parameter Manual.* ConeTec

Ouyang, Z., and Mayne, P. W. 2018. “Effective friction angle of clays and silts from piezocone penetration tests.” *Canadian Geotechnical Journal*, *55*(9). https://doi.org/10.1139/cgj-2017-0451

Ouyang, Z., and Mayne, P. W. 2019. “Modified NTH Method for Assessing Effective Friction Angle of Normally Consolidated and Overconsolidated Clays from Piezocone Tests.” *Journal of Geotechnical and Geoenvironmental Engineering*, *145*(10). https://doi.org/10.1061/(asce)gt.1943-5606.0002112

Pincus, H., Jefferies, M., and Davies, M. 1993. “Use of CPTu to Estimate Equivalent SPT N60.” *Geotechnical Testing Journal*, *16*(4). https://doi.org/10.1520/gtj10286j

Pinkus, A. 1999. “Approximation theory of the MLP model in neural networks.” *Acta Numerica*, *8*. https://doi.org/10.1017/S0962492900002919

Raissi, M., Perdikaris, P., and Karniadakis, G. E. 2019. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations.” *Journal of Computational Physics*, *378*. https://doi.org/10.1016/j.jcp.2018.10.045

Rendulic, L. 1936. “Porenziffer und porenwasserdruck in Tonen. ” *Der Bauingenieur*, 17, 559-564.

Robertson, P. K. 1990. “Soil classification using the cone penetration test.” *Canadian Geotechnical Journal*, *27*(1). https://doi.org/10.1139/t90-014

Robertson, P. K. 2009. “Interpretation of cone penetration tests - A unified approach.” *Canadian Geotechnical Journal*, *46*(11). https://doi.org/10.1139/T09-065

Robertson, P. K., and Cabal, K. L. 2010. “Guide to Cone Penetration Testing for Geotechnical Engineering.” *Gregg Drilling & Testing, Inc.*

Sandhu, R. S., and Wilson, E. L. 1969. “Finite-Element Analysis of Seepage in Elastic Media.” *Journal of the Engineering Mechanics Division*, *95*(3). https://doi.org/10.1061/jmcea3.0001124

Schnaid, F. (2008). In situ testing in geomechanics: The main tests. In *In Situ Testing in Geomechanics: The Main Tests*. https://doi.org/10.1201/9781482266054

Schneider, J. A., Mayne, P. W., and Rix, G. J. 2000. “Ground improvement assessment using SCPTu and crosshole data.” *Proceedings of Sessions of Geo-Denver 2000 - Innovations and Applications in Geotechnical Site Characterization, GSP 97*, *285*. https://doi.org/10.1061/40505(285)13

Shwetank, K., Deb, D., and Pramanik, R. 2023. “Coupled meshfree (SPH) and grid based (FDM) procedures for modeling fluid flow through deformable porous media.” *International Journal of Rock Mechanics and Mining Sciences*, *170*. https://doi.org/10.1016/j.ijrmms.2023.105494

Tartakovsky, A. M., Marrero, C. O., Perdikaris, P., Tartakovsky, G. D., and Barajas-Solano, D. 2020. “Physics-Informed Deep Neural Networks for Learning Parameters and Constitutive Relationships in Subsurface Flow Problems.” *Water Resources Research*, *56*(5). https://doi.org/10.1029/2019WR026731

Terzaghi, K. 1925, "Principles of soil mechanics," Engineering News-Record. Vol. 95.