

Article

Not peer-reviewed version

A Mathematical Model of the Generalized Finite Strain Consolidation Process and Its Deep Galerkin Solution

[Guang Yih Sheu](#) *

Posted Date: 1 September 2025

doi: 10.20944/preprints202509.0037.v1

Keywords: finite strain consolidation; Reynolds transport theorem; deep Galerkin method; mass and moment balance equations; self-weight



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Mathematical Model of the Generalized Finite Strain Consolidation Process and Its Deep Galerkin Solution

Guang Yih Sheu 

Department of Accounting and Information Systems, Chang-Jung Christian University, No.1, Changda Rd., Gueiren District, Tainan City 711301, Taiwan; xsheu@mail.cjcu.edu.tw; Tel.: +886-6-2785123-2225

Abstract

Developing classical three-dimensional consolidation theories considers the small-strain assumption. This small-strain assumption is inappropriate when studying the consolidation process of soft or very soft clay layers. Instead, this study derives a generalized mathematical model describing a three-dimensional finite-strain consolidation process and applies the deep Galerkin method to deduce its numerical solution. Developing this mathematical model uses the Reynolds transport theorem to describe mass and momentum balances for clay grain and pore water phases. The governing equation is the sum of the resulting mass and momentum balance equations. Next, the deep Galerkin method is applied to train a deep neural network to minimize the loss function defined by the governing equation and available initial and boundary conditions. The unknowns are the average velocity, effective stress, and pore water pressure. Predicting consolidation settlements is implemented by updating the problem domain using the resulting average velocity. Beneficial from the deep Galerkin method, two real-world examples demonstrate that the current mathematical model provides accurate predictions of consolidation settlements caused by the self-weight of two very soft clay layers. The deep Galerkin method helps eliminate the need to specify complete boundary conditions, which is difficult to achieve in modeling a finite strain consolidation process.

Keywords: finite strain consolidation; Reynolds transport theorem; deep Galerkin method; mass and moment balance equations; self-weight

1. Introduction

Consolidation represents the time-dependent process by which clay decreases its volume due to the expulsion of pore water under an external loading or its self-weight. A consolidation theory or a mathematical model describing the consolidation process is essential in many actual problems such as estimating the consolidation settlements due to a construction loading (e.g., [1]), evaluating the results of a sand drain scheme (e.g., [2]), and scheduling a land reclamation project (e.g., [3]). Unexpected consolidation settlements may damage a building and delay a land reclamation project.

Nevertheless, developing classical mathematical models describing consolidation processes considers the small-strain assumption (e.g., [4]) in which consolidation settlements do not substantially change the thickness of a clay layer. It is invalid for a soft or very soft clay layer since it may be significantly thinner due to its consolidation settlements. These consolidation settlements may be caused only by their self-weight. A real example may be the Kansai airport, which may gradually sink into the sea due to excessive settlements of undersea clay layers. However, we only had one-dimensional [5,6] or quasi multi-dimensional mathematical models [7] to describe the consolidation process of a soft or very soft clay layer. Developing these mathematical models assumes that a soft or very soft clay layer deforms only in the vertical direction. This assumption is inconsistent with the actual condition.

Therefore, the author's thesis [8] developed a generalized mathematical model describing a generalized finite-strain consolidation process. The Reynolds transport theorem [9] is employed to

represent mass and momentum balances. The previous thesis [8] further combined the resulting mass and momentum equations into one single governing equation. Nevertheless, the resulting governing equation is long and complex. Applying it to model a real problem is difficult since we have few numerical methods to solve it. To solve this drawback, this study applies the deep Galerkin method (DGM) [10] to minimize the loss function defined by the sum of resulting momentum equations, initial, and boundary conditions. The unknowns are the average velocity field, effective stress, and pore water pressure. Calculating consolidation settlements is implemented by updating the problem domain according to the resulting average velocity field.

This DGM (e.g., [10]) is a meshless deep algorithm for solving high-dimensional partial differential equations. Choosing it considers the difficulties of collecting necessary data to define the boundary conditions (for example, a deep boundary) of a soft or very soft clay layer. By using the DGM, we train a deep neural network to provide a solution to a partial differential equation at random points in a problem domain. Defining complete boundary conditions is unnecessary.

The objectives of creating this study are

1. Compensate for the absence of a useful mathematical model for describing a generalized finite-strain consolidation process. It advances the simulation of the finite-strain consolidation process of soft or very soft clay layers. Compared to traditional small-strain consolidation theories (e.g., [4]), the proposed mathematical model has more applications, such as estimating settlements of dredge fill deposits.
2. Extend the application of the DGM to a partial differential equation with a problem domain subjected to large strains.

The remainder of this article has four sections. Section 2 presents a review of published articles relevant to this study. Section 3 presents the development of a generalized finite strain consolidation theory and its DGM solution. Section 4 presents the numerical results generated using the current mathematical model. Comparing these numerical results to observed settlements of the phosphatic waste clay and Osaka Bay mud is done in this section. Based on this comparison, Section 5 presents a discussion. Section 6 presents conclusions and the concluding remark of this research.

2. Related Works

This study intends to develop a mathematical model to describe a generalized finite-strain consolidation process and apply the DGM to solve the unknowns. Therefore, relevant references are those articles presenting a two- or three-dimensional finite strain consolidation theory and applications of the DGM to solve various partial differential equations.

2.1. Finite Strain Consolidation Theory

When we intend to develop a generalized finite strain consolidation theory, the origin is usually the one-dimensional finite strain consolidation theory proposed by Gibson et al. [5,6]. Different from Biot [4], they used the void ratio as the main unknown of their mathematical model. Considering that consolidation settlements may substantially thin a soft clay layer, they defined a convective coordinate system to describe the changes in a clay layer's thickness. Besides, Gibson et al. [5,6] accounted for the effects of a thick clay layer's self-weight on its consolidation process.

After the one-dimensional finite strain consolidation theory [5,6], few articles presented a truly generalized finite strain consolidation theory. However, quasi-two- or three-dimensional theories are available. For example, Jeeravipoolvarn et al. [7] developed a quasi-three-dimensional finite strain consolidation theory, which assumed that pore water flows in any direction and pore pressure dissipation causes only vertical deformations. Huerta and Rodriguez (1992) [11] calculated finite-strain consolidation settlements of soft sediment fillings at high water levels. They considered one-dimensional deformations and two-dimensional fluxes. Liu et al. (2021) [12] assumed vertical strains. They evaluated the vertical drains combined with vacuum pressure using a quasi-axisymmetric finite strain consolidation theory.

The author's Ph.D. thesis [8] presented a truly three-dimensional finite strain consolidation theory. This generalized finite strain consolidation theory extended the one-dimensional works provided by Gibson et al. [5,6]. Creating this extension used the Reynolds transport theory to model the mass and momentum balances for the clay grain and pore water phases. Limiting the deformation and pore water dissipation in the vertical direction is not required. Besides, the effect of the self-weight of a thick clay layer on its consolidation process is studied. However, the resulting mathematical model may be complex.

2.2. Deep Galerkin Method

Since this study chooses the DGM (e.g., [10]) to solve the proposed mathematical model, it is necessary to review the representative partial differential equations, which were solved using the DGM:

Kumar et al. (2022) [13] used the DGM to solve the one-dimensional Burger-Huxley and Huxley equations. They are second-order partial differential equations with initial and boundary conditions. In the application of DGM, estimating the time derivatives using a finite difference scheme was not required. Approximating a time derivative by a finite difference scheme is frequently seen in the application of the finite element method in solving a time-dependent partial differential equation.

Sirignano and Spiliopoulos (2018) [10] applied the DGM to a free boundary equation. It is a one-dimensional and second-order partial differential equation. Its theoretical background is the stock price dynamics. One of the boundaries is a time-dependent function. It enters into the derivation of the loss function defined to implement the DGM.

Masaharu (2021) [14] employed the DGM to solve a compressible Navier-Stokes equation. The Navier-Stokes equation contains convective accelerations. They are highly nonlinear differential terms. We may notice that Masaharu (2021) [14] studied supersonic flows around a blunt body without specifying full boundary conditions. This example is quite different from the application of a finite element method.

3. Generalized Finite Strain Consolidation

3.1. Mathematical Model

Suppose a homogeneous and saturated soil layer. Besides, clay grains and pore water are incompressible. The convective coordinate $\mathbf{x}$ is defined to identify a point in a problem domain Ω .

If Θ denotes the total amount of an extensive property within Ω_0 , the Reynolds transport theorem states that the conservation of Θ can be represented by [9]

$$\frac{d\Theta}{dt} = \frac{d}{dt} \left(\int_{\Omega} \theta dV \right) + \int_S \theta (\mathbf{v}_S \cdot \mathbf{n}_S) dS = 0 \quad (1)$$

where $\frac{d}{dt}$ is the total derivative with respect to the time t , θ is the density of Θ , V is the volume of Ω , S is the boundary of Ω , the subscript s denotes the boundary, $\cdot$ is the dot operator, $\mathbf{v}$ is the velocity field, and $\mathbf{n}_S$ is a unit vector normal to S .

3.1.1. Mass Balance for the Clay Grain Phase

Substituting $\theta = (1 - n)\rho_s$ into Equation (1) results in

$$\frac{d}{dt} \left[\int_{\Omega} (1 - n)\rho_s dV \right] + \int_S (1 - n)\rho_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS = 0 \quad (2)$$

where f is a function, the subscript s denotes the clay grain phase, n is the porosity, ρ is the density. Applying the definition of total derivative $\frac{df}{dt} = \frac{\partial f}{\partial t} + \nabla f \cdot \mathbf{v}$ and coordinate transform to simplify Equation (2) yields

$$\begin{aligned} \frac{d}{dt} \left[\int_{\Omega} (1-n) \rho_s dV \right] + \int_S (1-n) \rho_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS &= \frac{d}{dt} \left[\int_{\Omega_0} (1-n_0) \rho_{s0} J dV_0 \right] \\ + \int_S (1-n) \rho_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS &= \int_{\Omega_0} \frac{d[(1-n_0) \rho_{s0} J]}{dt} dV_0 + \int_S (1-n) \rho_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS \\ &= \int_{\Omega_0} \left\{ \frac{\partial[(1-n_0) \rho_{s0} J]}{\partial t} + \mathbf{v}_{s0} \cdot \nabla[(1-n_0) \rho_{s0} J] \right\} dV_0 + \int_S (1-n) \rho_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS = 0 \end{aligned} \quad (3)$$

in which the subscript 0 denotes the Lagrangian coordinate $\mathbf{a}$, ∇ is the gradient vector, $J = \left| \frac{\partial(x_1, x_2, x_3)}{\partial(a_1, a_2, a_3)} \right|$ is the Jacobian, $x_i (i = 1, 2, 3)$ are the components of convective coordinate $\mathbf{x}$, and a_i are the components of Lagrangian coordinate $\mathbf{a}$.

Applying the Gauss theorem and coordinate transform to simplify the final term of Equation (3) yields

$$\int_S (1-n) \rho_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS = \int_{\Omega_0} (1-n) \rho_s \nabla \cdot \mathbf{v}_s dV = \int_{\Omega_0} (1-n_0) \rho_{s0} \nabla \cdot \mathbf{v}_{s0} J dV_0 \quad (4)$$

Combining the final term of Equation (4) with the eight term $\left(\int_{\Omega_0} \mathbf{v}_{s0} \cdot \nabla[(1-n_0) \rho_{s0} J] dV_0 \right)$ in Equation (3) yields $\int_{\Omega_0} \nabla \cdot [(1-n_0) \rho_{s0} J \mathbf{v}_{s0}] dV_0$. Substituting the resulting expression into Equation (3) yields

$$\int_{\Omega_0} \left\{ \frac{\partial[(1-n_0) \rho_{s0} J]}{\partial t} + \nabla \cdot [(1-n_0) \rho_{s0} J \mathbf{v}_{s0}] \right\} dV_0 = 0 \quad (5)$$

Or we can consider the localization and write

$$\frac{\partial[(1-n_0) \rho_{s0} J]}{\partial t} + \nabla \cdot [(1-n_0) \rho_{s0} J \mathbf{v}_{s0}] = 0 \quad (6)$$

Equation (6) is the mass balance equation for the clay grain phase. Furthermore, simplifying Equation (5) by considering that only the pore water dissipation from the dV causes the consolidation settlements yields

$$\int_{\Omega_0} \nabla \cdot [(1-n_0) \rho_{s0} J \mathbf{v}_{s0}] dV_0 = 0 \quad \text{and} \quad \int_{\Omega_0} \frac{\partial[(1-n_0) \rho_{s0} J]}{\partial t} dV_0 = 0 \quad (7)$$

Further simplifying the final expression of Equation (7) leads to

$$(1-n_0) \rho_{s0} J = C = [1-n_0(t=0)] \rho_{s0}(t=0) \quad (8)$$

where C is independent upon the time t . Since this study has assumed that clay grains are incompressible, we can obtain: $\rho_{s0} = \rho_{s0}(t=0)$ and simplify Equation (8) to

$$J = \frac{1-n_0(t=0)}{1-n_0} \quad (9)$$

3.1.2. Mass Balance for the Pore Water Phase

Substituting $\theta = n \rho_w$ into Equation (1) yields

$$\frac{d}{dt} \left(\int_{\Omega} n \rho_w dV \right) + \int_S n \rho_w (\mathbf{v}_w \cdot \mathbf{n}_S) dS = 0 \quad (10)$$

where the subscript w represents the pore water phase.

Referring to Section 3.1.1 to simplify Equation (10) results in

$$\begin{aligned} \frac{d}{dt} \left(\int_{\Omega} n \rho_w dV \right) + \int_S n \rho_w (\mathbf{v}_w \cdot \mathbf{n}_S) dS &= \int_{\Omega_0} \left[\frac{\partial(n_0 \rho_{w0} J)}{\partial t} + \mathbf{v}_{w0} \cdot \nabla(n_0 \rho_{w0} J) \right] dV_0 \\ + \int_{\Omega_0} n_0 \rho_{w0} (\nabla \cdot \mathbf{v}_{w0}) J dV_0 &= \int_{\Omega_0} \left[\frac{\partial(n_0 \rho_{w0} J)}{\partial t} + \nabla \cdot (n_0 \rho_{w0} \mathbf{v}_{w0} J) \right] dV_0 = 0 \end{aligned} \quad (11)$$

Considering the localization to simplify Equation (11) yields

$$\frac{\partial(n_0 \rho_{w0} J)}{\partial t} + \nabla \cdot (n_0 \rho_{w0} \mathbf{v}_{w0} J) = 0 \quad (12)$$

Equation (12) is the mass balance equation for the pore water phase.

3.1.3. Momentum Balance for the Clay Grain Phase

Substituting $\theta = (1 - n) \rho_s \mathbf{v}_s$ into Equation (1) yields

$$\frac{d}{dt} \left[\int_{\Omega} (1 - n) \rho_s \mathbf{v}_s dV \right] + \int_S (1 - n) \rho_s \mathbf{v}_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS = 0 \quad (13)$$

Similar to the derivation of Equation (3), one can simplify Equation (13) to

$$\begin{aligned} \frac{d}{dt} \left[\int_{\Omega} (1 - n) \rho_s \mathbf{v}_s dV \right] + \int_S (1 - n) \rho_s \mathbf{v}_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS &= \frac{d}{dt} \left[\int_{\Omega_0} (1 - n_0) \rho_{s0} \mathbf{v}_{s0} J dV_0 \right] \\ + \int_S (1 - n) \rho_s \mathbf{v}_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS &= \int_{\Omega_0} \frac{d[(1 - n_0) \rho_{s0} \mathbf{v}_{s0} J]}{dt} dV_0 + \int_S (1 - n) \rho_s \mathbf{v}_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS \\ &= \int_{\Omega_0} \rho_{s0} (1 - n_0) J \frac{d\mathbf{v}_{s0}}{dt} dV_0 + \int_{\Omega_0} \mathbf{v}_{s0} \frac{d[\rho_{s0} (1 - n_0) J]}{dt} dV_0 + \int_S (1 - n) \rho_s \mathbf{v}_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS \end{aligned} \quad (14)$$

By simplifying Equation (14) using the Gauss theorem and transforming the final term of the same equation using the coordinate transformation, we can reduce the equation to

$$\begin{aligned} \int_{\Omega_0} \rho_{s0} (1 - n_0) J \frac{d\mathbf{v}_{s0}}{dt} dV_0 + \int_{\Omega_0} \mathbf{v}_{s0} \frac{d[\rho_{s0} (1 - n_0) J]}{dt} dV_0 + \int_S (1 - n) \rho_s \mathbf{v}_s (\mathbf{v}_s \cdot \mathbf{n}_S) dS \\ = \int_{\Omega_0} \rho_{s0} (1 - n_0) J \frac{d\mathbf{v}_{s0}}{dt} dV_0 + \int_{\Omega_0} \mathbf{v}_{s0} \frac{d[\rho_{s0} (1 - n_0) J]}{dt} dV_0 + \int_{\Omega_0} (1 - n_0) \rho_{s0} \mathbf{v}_{s0} J (\nabla \cdot \mathbf{v}_{s0}) dV_0 \end{aligned} \quad (15)$$

Combining the final two terms of Equation (15) yields

$$\begin{aligned} \int_{\Omega_0} \mathbf{v}_{s0} \frac{d[\rho_{s0} (1 - n_0) J]}{dt} dV_0 + \int_{\Omega_0} (1 - n_0) \rho_{s0} \mathbf{v}_{s0} J (\nabla \cdot \mathbf{v}_{s0}) dV_0 \\ = \int_{\Omega_0} \mathbf{v}_{s0} \left\{ \frac{\partial[\rho_{s0} (1 - n_0) J]}{\partial t} + \nabla[\rho_{s0} (1 - n_0) J] \cdot \mathbf{v}_{s0} + (1 - n_0) \rho_{s0} J (\nabla \cdot \mathbf{v}_{s0}) \right\} dV_0 \\ = \int_{\Omega_0} \mathbf{v}_{s0} \left\{ \frac{\partial[\rho_{s0} (1 - n_0) J]}{\partial t} + \nabla \cdot [\rho_{s0} (1 - n_0) J \mathbf{v}_{s0}] \right\} dV_0 \end{aligned} \quad (16)$$

Based on Equation (6), Equation (16) is equal to 0. We then simplify Equation (13) to

$$\int_{\Omega_0} \rho_{s0} (1 - n_0) J \frac{d\mathbf{v}_{s0}}{dt} dV_0 = 0 \quad (17)$$

where the term $\frac{d\mathbf{v}}{dt}$ is the acceleration field and $\rho_{s0} (1 - n_0)$ denotes the mass of clay grains (in $J dV_0$). Therefore, this equation is equal to the forces exerted on clay grains. Thus, we can define:

$$\int_{\Omega_0} \rho_{s0} (1 - n_0) J \frac{d\mathbf{v}_{s0}}{dt} dV_0 = \int_{\Omega_0} \mathbf{f}_s (1 - n_0) dV_0 + \int_{\Omega_0} \mathbf{h}^s dV_0 + \int_{S_0} (\boldsymbol{\sigma}'_0 \cdot \mathbf{n}_{S_0}) dS_0 \quad (18)$$

in which σ denotes the stress tensor, the prime represents the effective stress, $\mathbf{h}^s$ is the seepage force vector per unit volume arising from the frictional drag of pore water, and $\mathbf{f}_s$ is the body force vector. This study considers that the $\mathbf{f}_s$ comes from the weights of clay grains. It and σ'_0 can be further equated by

$$\sigma'_0 = \begin{bmatrix} \sigma'_{11,0} & \sigma'_{12,0} & \sigma'_{13,0} \\ \sigma'_{21,0} & \sigma'_{22,0} & \sigma'_{23,0} \\ \sigma'_{31,0} & \sigma'_{32,0} & \sigma'_{33,0} \end{bmatrix} \quad \text{and} \quad \mathbf{f}_s = \pm \rho_{s0} \mathbf{g} \quad (19)$$

where $\mathbf{g} = (0, 0, g)^T$ and g is the gravity acceleration. The sign $\pm$ depends upon the direction of a_3 . If the direction of a_3 is opposite to the direction of the gravitational force, the sign $+$ is adopted. Meanwhile, applying the Gauss theorem to simplify the final term of Equation (18) yields $\int_{S_0} (\sigma'_0 \cdot \mathbf{n}_{S_0}) dS_0 = \int_{\Omega_0} (\nabla \cdot \sigma'_0) dV_0$. Substituting the resulting expression into Equation (18). Considering Equation (17) and the localization, we can deduce:

$$\rho_{s0}(1 - n_0)J \frac{d\mathbf{v}_{s0}}{dt} = \pm \rho_{s0}(1 - n_0)\mathbf{g} + \mathbf{h}^s + \nabla \cdot \sigma'_0 \quad (20)$$

Equation (20) is the momentum balance equation for the clay grain phase.

3.1.4. Momentum Balance for the Pore Water Phase

Substituting $\theta = n\rho_w \mathbf{v}_s$ into Equation (1) yields

$$\frac{d}{dt} \left(\int_{\Omega} n\rho_w \mathbf{v}_w dV \right) + \int_S n\rho_w \mathbf{v}_w (\mathbf{v}_w \cdot \mathbf{n}_S) dS = 0 \quad (21)$$

Referring to the derivation of Equation (14) to simplify the first term of Equation (21) and applying the Gauss theorem and coordinate transform to the final term of the same equation, the resulting equation is

$$\begin{aligned} \frac{d}{dt} \left(\int_{\Omega} n\rho_w \mathbf{v}_w dV \right) + \int_S n\rho_w \mathbf{v}_w (\mathbf{v}_w \cdot \mathbf{n}_S) dS &= \frac{d}{dt} \left(\int_{\Omega_0} n_0 \rho_{w0} \mathbf{v}_{w0} J dV_0 \right) \\ + \int_S n\rho_w \mathbf{v}_w (\mathbf{v}_w \cdot \mathbf{n}_S) dS &= \int_{\Omega_0} \rho_{w0} n_0 J \frac{d\mathbf{v}_{w0}}{dt} dV_0 + \int_{\Omega_0} \mathbf{v}_{w0} \frac{d(\rho_{w0} n_0 J)}{dt} dV_0 \\ + \int_{S_0} n_0 \rho_{w0} \mathbf{v}_{w0} J (\nabla \cdot \mathbf{v}_{w0}) dV_0 \end{aligned} \quad (22)$$

The final two terms of Equation (22) are combined and reduced to

$$\begin{aligned} \int_{\Omega_0} \mathbf{v}_{w0} \frac{d(\rho_{w0} n_0 J)}{dt} dV_0 + \int_{\Omega_0} n_0 \rho_{w0} \mathbf{v}_{w0} J (\nabla \cdot \mathbf{v}_{w0}) dV_0 \\ = \int_{\Omega_0} \mathbf{v}_{w0} \left[\frac{\partial(\rho_{w0} n_0 J)}{\partial t} + \nabla \cdot (\rho_{w0} n_0 J \mathbf{v}_{w0}) \right] dV_0 \end{aligned} \quad (23)$$

Based on Equation (12), Equation (23) is equal to 0. Thus, we can reduce Equation (21) to

$$\int_{\Omega_0} \rho_{w0} n_0 J \frac{d\mathbf{v}_{w0}}{dt} dV_0 = 0 \quad (24)$$

Similar to Equation (17), Equation (24) is equal to the forces exerted on the pore water phase. Thus [9]

$$\int_{\Omega_0} \rho_{w0} n_0 J \frac{d\mathbf{v}_{w0}}{dt} dV_0 = \int_{\Omega_0} \mathbf{f}_w n_0 dV_0 + \int_{\Omega_0} \mathbf{h}^w dV_0 + \int_{S_0} (\mathbf{p}_0 \cdot \mathbf{n}_{S_0}) dS_0 \quad (25)$$

where $\mathbf{p}$ denotes the pore water pressure tensor, $\mathbf{h}^w$ is the reactive force vector per unit volume exerted by clay grains on the pore water phase as the pore water seeps through the clay, and $\mathbf{f}_w$ is the body force vector. The $\mathbf{p}_0$ and $\mathbf{f}_w$ are further defined by

$$\mathbf{p}_0 = \begin{bmatrix} p_0 & 0 & 0 \\ 0 & p_0 & 0 \\ 0 & 0 & p_0 \end{bmatrix} \quad \text{and} \quad \mathbf{f}_w = \pm \rho_{w0} \mathbf{g} \quad (26)$$

in which p is the pore water pressure and setting the sign $\pm$ still depends upon the direction of a_3 . On the other hand, applying the Gauss theorem to the final term of Equation (25) yields $\int_{S_0} (\mathbf{p}_0 \cdot \mathbf{n}_{S_0}) dS_0 = \int_{\Omega_0} (\nabla \cdot \mathbf{p}_0) dV_0$. Substituting the resulting expression into Equation (25) and considering the localization, we can obtain:

$$\rho_{w0} n_0 J \frac{d\mathbf{v}_{w0}}{dt} = \pm n_0 \rho_{w0} \mathbf{g} + \mathbf{h}^w + \nabla \cdot \mathbf{p}_0 \quad (27)$$

Equation (27) is the momentum balance equation for the pore water phase.

3.2. DGM Formulation

Summing Equations (20) and (27) results in

$$\rho_{total} J \frac{d\mathbf{v}_{avg}}{dt} = \pm \rho_{total} \mathbf{g} + \mathbf{h}^w + \mathbf{h}^s + \nabla \cdot \mathbf{p}_0 + \nabla \cdot \boldsymbol{\sigma}'_0 \quad (28)$$

where $\mathbf{v}_{avg} = \frac{\rho_{s0}(1-n_0)\mathbf{v}_{s0} + \rho_{w0}n_0\mathbf{v}_{w0}}{\rho_{total}}$ denotes the average velocity field, $\frac{d\mathbf{v}_{avg}}{dt}$ represents the average acceleration field, and $\rho_{total} = \rho_{s0}(1-n_0) + \rho_{w0}n_0$ is the total mass density. Different from the author's thesis [8], this study does not try to combine Equations (9), (12), (20), and (27) into a governing equation. Thus, deriving a DGM solution of the current mathematical model is simpler.

Furthermore, assume $\mathbf{h}^w + \mathbf{h}^s = 0$ [8] or neglects the interaction between pore water and clay grains. Simplifying Equation (28) with this consideration yields

$$\rho_{total} J \frac{d\mathbf{v}_{avg}}{dt} = \pm \rho_{total} \mathbf{g} + \nabla \cdot \mathbf{p}_0 + \nabla \cdot \boldsymbol{\sigma}'_0 \quad (29)$$

A DGM solution of Equation (29) is next derived: Suppose the time t is between 0 and T . The initial and boundary conditions are

$$\frac{d\mathbf{v}_{avg}(t=0)}{dt} = 0, \quad \mathbf{p}(t=0) = \mathbf{p}_{ini}, \quad \text{and} \quad \boldsymbol{\sigma}'(t=0) = \boldsymbol{\sigma}'_{ini} \quad (30)$$

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma}_d \quad \text{on} \quad \Gamma_f \quad (31)$$

$$p = 0 \quad \text{on} \quad \Gamma_D \quad (32)$$

$$\mathbf{v}_{avg} = 0 \quad \text{on} \quad \Gamma_F \quad (33)$$

$$\frac{\partial p_0}{\partial n_{Uj}} = 0 \quad \text{on} \quad \Gamma_U \quad (34)$$

where Γ_f , Γ_D , Γ_F , and Γ_U are free, drained, fixed, and undrained boundaries, respectively, $\mathbf{p}_{ini}$ is the initial pore water pressure tensor, $\boldsymbol{\sigma}'_{ini}$ represents the initial effective stress tensor, $\boldsymbol{\sigma}_d$ denotes an imposed loading, and n_{Uj} ($j = 1, 2, 3$) is the component of a unit vector $\mathbf{n}_U$ normal to the undrained boundary Γ_U . Note that Γ_f , Γ_D , Γ_F , and Γ_U may not exist simultaneously in a particular problem domain Ω_0 . Also the computation of $\boldsymbol{\sigma}'_{ini}$ and $\mathbf{p}_{ini}$ depends upon a particular problem domain Ω_0 .

Based on Equations (29)-(34), this study applied the DGM algorithm to approximate the $\mathbf{v}_{avg}$ using $\boldsymbol{\alpha}(\mathbf{a}, t, \phi)$, $\mathbf{p}_0$ using $\mathbf{P}(\mathbf{a}, t, \phi)$, and $\boldsymbol{\sigma}'_0$ using $\mathbf{S}(\mathbf{a}, t, \phi)$ in which ϕ is the neural network's parameter. The DGM algorithm defines the loss function by (e.g., [10])

$$\begin{aligned} Loss = & \left\| \rho_{total} J \frac{d\boldsymbol{\alpha}}{dt} - \nabla \cdot \mathbf{P} - \nabla \cdot \mathbf{S} \right\|_{(0,T) \times \Omega_0, \nu_1}^2 + \|\boldsymbol{\alpha}(t=0)\|_{\Omega_0, \nu_2}^2 \\ & + \|\mathbf{P}(t=0) - \mathbf{p}_{ini}\|_{\Omega_0, \nu_2}^2 + \|\boldsymbol{\sigma}'(t=0) - \boldsymbol{\sigma}'_{ini}\|_{\Omega_0, \nu_2}^2 \\ & + \|\mathbf{P}\|_{(0,T) \times \Gamma_D, \nu_3}^2 + \|\mathbf{S} - \boldsymbol{\sigma}_d\|_{(0,T) \times \Gamma_f, \nu_4}^2 + \left\| \frac{\partial P_{ij}}{\partial n_{Uj}} \right\|_{(0,T) \times \Gamma_U, \nu_5}^2 \end{aligned} \quad (35)$$

where P_{ij} ($i, j = 1, 2, 3$) is the component of the $\mathbf{P}$ tensor, the symbol $\|\cdot\|$ represents the L^2 norm function in which $Loss$ measures how well the functions $\boldsymbol{\alpha}$, $\mathbf{P}$, and $\mathbf{S}$ satisfy Equations (29)-(34). If we can obtain $Loss = 0$, the $\boldsymbol{\alpha}$, $\mathbf{P}$, and $\mathbf{S}$ are the solutions of Equation (29).

The goal of defining Equation (35) is to find the ϕ parameter with which the corresponding $\boldsymbol{\alpha}(\mathbf{a}, t, \phi)$, $\mathbf{P}(\mathbf{a}, t, \phi)$, and $\mathbf{S}(\mathbf{a}, t, \phi)$ minimize the $Loss$ function. Then, they are solutions of the $\mathbf{v}_{avg}$, $\mathbf{p}$, and $\boldsymbol{\sigma}'$. The best ϕ value is estimated using an existing optimization method (for example, the Adam optimizer). Discretizing the problem domain Ω into some elements is unnecessary.

Algorithm 1 [14] illustrates the proposed steps of implementing the DGM algorithm and three post-processing steps (6th-8th steps) in which k denotes an iteration number, E denotes an error, β represents the learning rate, and $\mathbf{C}$ is the compliance matrix, R denotes the void ratio-effective stress relationship, and $\boldsymbol{\epsilon}$ is the strain tensor. The sixth step provides the strain tensor $\boldsymbol{\epsilon}$ or void ratio e . The seventh step outputs the numerical results for updating the Jacobian J and ρ_{total} . Since the initial displacement and velocity fields are zero, the eighth step is employed to update the coordinates x_j ($j = 1, 2, 3$). Updating the coordinates x_j also provides the consolidation settlements. Note that this study does not calculate consolidation settlements from the strain tensor $\boldsymbol{\epsilon}$ since collecting undisturbed samples of soft or very soft clay to construct its stress-strain and void ratio-stress relationships is difficult.

Algorithm 1 DGM algorithm

Require: A problem domain Ω , a time interval $(0, T)$, initial and boundary conditions, the neural network's parameter ϕ , a learning rate β , a maximum iteration number, and a compliance matrix $\mathbf{C}$ or a void ratio-effective stress relationship R .

Ensure: Solutions of the $\mathbf{v}_{avg}$, $\mathbf{p}$ and $\boldsymbol{\sigma}'$.

- 1: Generate random points from $\Omega \times [0, T]$ and existing boundaries
 - 2: **while** a maximum iteration number is not exceeded **do**
 - 3: Calculate the $\boldsymbol{\sigma}'_{ini}$ and $\mathbf{p}_{ini}$ according to the initial problem domain Ω .

$$E \leftarrow \left[\rho_{total} J \frac{d\boldsymbol{\alpha}(\mathbf{a}, t, \phi_k)}{dt} - \nabla \cdot \mathbf{P}(\mathbf{a}, t, \phi_k) - \nabla \cdot \mathbf{S}(\mathbf{a}, t, \phi_k) \right]_{\Omega}^2 + [\boldsymbol{\alpha}(t=0, \phi_k)]_{\Omega}^2$$
 - 4: Calculate: $+ [\mathbf{P}(t=0, \phi_k) - \mathbf{p}_{ini}]_{\Omega}^2 + [\boldsymbol{\sigma}'(t=0, \phi_k) - \boldsymbol{\sigma}'_{ini}]_{\Omega}^2 + [\mathbf{P}(\mathbf{a}, t, \phi_k)]_{\Gamma_D}^2$
 $+ [\mathbf{S}(\mathbf{a}, t, \phi_k) - \boldsymbol{\sigma}_d]_{\Gamma_f}^2 + \left[\frac{\partial P_{ij}}{\partial n_{Uj}} \right]_{\Gamma_U}^2$
 - 5: Update $\phi_{k+1} \leftarrow \phi_k - \beta_k \nabla_{\phi} E$.
 - 6: Calculate $\boldsymbol{\epsilon} \leftarrow \mathbf{C}\boldsymbol{\sigma}'$ or $e \leftarrow R(\boldsymbol{\sigma}')$.
 - 7: Update J and ρ_{total} .
 - 8: Update $\mathbf{x} = \mathbf{x} + \int \mathbf{v}_{avg} dt$.
 - 9: $k \leftarrow (k+1)$.
 - 10: **end while**
-

The learning rate β_k in Algorithm 1 decreases with the iteration number k . The calculation $\nabla_{\phi} E$ is unbiased estimation of $\nabla_{\phi}(Loss)$:

$$\mathbb{E}[\nabla_{\phi} E(\cdot | \phi_k)] = \nabla_{\phi} [Loss(\cdot : \phi_k)] \quad (36)$$

where $\mathbb{E}$ denotes the mathematical expectation. Under technical conditions, the ϕ parameter will converge to a critical point of the *Loss* function as $k \rightarrow \infty$:

$$\lim_{k \rightarrow \infty} \|\nabla_{\phi} \text{Loss}(\cdot : \phi_k)\| = 0 \quad (37)$$

Figure 1 (e.g., [13,14]) shows an example of the neural network's architecture for implementing the DGM. It approximates the $\mathbf{v}_{avg}$ by $\alpha(\mathbf{a}, t, \phi)$, $\mathbf{p}$ by $\mathbf{P}(\mathbf{a}, t, \phi)$, and σ by $\mathbf{S}(\mathbf{a}, t, \phi)$ in which ϕ is the neural network's parameter. In mathematical expressions, the deep neural network contains three types of layers:

- Input layer: The neural network calculates:

$$\mathbf{Z}^{(0)} = \psi(\mathbf{W}_0 \mathbf{X} + \mathbf{b}_0) \quad (38)$$

in which $\mathbf{X} = (\mathbf{a}, t)^T$, $\mathbf{Z}$ is the intermediate hidden feature vector, ψ represents the activation function, $\mathbf{W}_0$ denotes the input weight matrix, and $\mathbf{b}_0$ represents the input bias vector.

- Hidden layer: Suppose L hidden layers are generated. For each hidden layer, the neural network computes:

$$\begin{aligned} \mathbf{F}^{(k)} &= \psi \left[\mathbf{W}_f^{(k)} \mathbf{Z}^{(k-1)} + \mathbf{b}_f^{(k)} \right] \\ \mathbf{r}^{(k)} &= \psi \left[\mathbf{W}_r^{(k)} \mathbf{Z}^{(k-1)} + \mathbf{b}_r^{(k)} \right] \\ \tilde{\mathbf{Z}}^{(k)} &= \tanh \left\{ \mathbf{W}_h^{(k)} \left[\mathbf{r}^{(k)} \odot \mathbf{Z}^{(k-1)} \right] + \mathbf{b}_h^{(k)} \right\} \\ \mathbf{Z}^{(k)} &= \left[1 - \mathbf{F}^{(k)} \right] \odot \mathbf{Z}^{(k-1)} + \mathbf{F}^{(k)} \odot \tilde{\mathbf{Z}}^{(k)} \end{aligned} \quad (39)$$

where $k = 1, 2, \dots, L$ denotes the L -th hidden layer, $\mathbf{F}$ is the gate vector, the subscripts f , r , and h denote the update gate, reset gate, and candidate hidden gate; respectively, $\mathbf{r}$ represents the reset gate. $\tilde{\mathbf{Z}}$ represents the candidate hidden state, $\odot$ is the Hadamard product, and $\tanh$ is the hyperbolic tangent function. This $\tanh$ function serves as a nonlinear activation function.

- Output layer:

$$\mathbf{U} = \mathbf{W}_{out} \mathbf{Z}^{(L)} + \mathbf{b}_{out} \quad (40)$$

where $\mathbf{U} = (\mathbf{v}_{avg}, \mathbf{p}, \sigma')^T$ represents the output of the neural network, $\mathbf{W}_{out}$ is the weight matrix of the output layer, and $\mathbf{b}_{out}$ is the bias vector of the output layer.

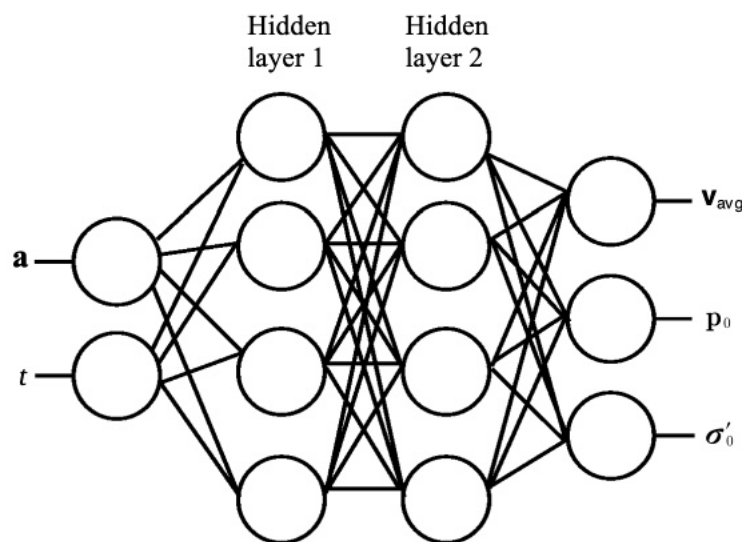


Figure 1. An example for the neural network's architecture for implementing the DGM (e.g., [10])

3.3. Implementation of the DGM

This study uses PyTorch in coding a Python program for implementing Algorithm 1. It has an automatic differentiation function that can easily output the derivatives with respect to independent variables. Moreover, back propagation is adopted to train a deep neural network. Besides, optimizing the training process adopts the Adam optimizer based on the stochastic gradient descent method. Furthermore, a Mac Pro is employed to generate numerical results. Since it has 16 GPUs, the total time of implementing Algorithm 1 can be limited to about 10 minutes.

4. Application

Considering the acquisition of experimental data, this study introduces the recorded consolidation settlements of the phosphatic waste clay [15] and Osaka Bay [16] mud to test Algorithm 1.

4.1. Phosphatic Waste Clay

The phosphate waste clay is the by-product of the beneficiation process of the phosphate ore. This phosphate ore occurs in a gravelly, clayey sand and contains $\frac{1}{3}$ phosphate, $\frac{1}{3}$ granular materials (sand), and $\frac{1}{3}$ clays. Merely the phosphate is collected as the primary source of phosphorus in inorganic fertilizers. After extracting the phosphate, the phosphate waste clay is pumped into large retention ponds and allowed to consolidate without any imposed loading. For designing the storage capacity of the retention ponds, we must predict well the consolidation settlements of the phosphate waste clay caused by its self-weight.

The phosphate waste clay has a liquid limit between 100 and 200, a plastic index between 70 and 150, and a specific yield G_s equal to 2.71 [15]. Some previous studies have constructed empirical relationships between the void ratio and effective stress. Equation (41) provides an example [17]:

$$e = (e_b - e_\infty) \exp(-\lambda \sigma'_{33}) + e_\infty \quad (41)$$

where e_b and e_∞ are the anticipated void ratios before and after the consolidation process, and λ is an empirical parameter. This word 'anticipated' denotes the difficulty of measuring the void ratio within the whole phosphate waste clay layer.

Testing the performance of the current mathematical model is through the results of a field test [17]. In this field test, a retention pond settled due to the self-weight of poured phosphate waste clay. Cargill [3] derived a one-dimensional mathematical model to predict the consolidation settlements. Instead, this study applied the two-dimensional version of the proposed mathematical model to predict the consolidation settlements and compare the measured results. The initial problem domain is $0 \leq a_j \leq H$ ($j = 1, 3$) in which H is the depth of the retention pond, the direction of a_3 is identical to the direction of the gravitational force, and the a_1 is from right to left. A small figure inside the subsequent Figure 2 also illustrates the initial problem domain. The two-dimensional version of Equation (29) is the governing equation:

$$\begin{aligned} \rho_{total} J \frac{dv_{avg,1}}{dt} &= \frac{\partial p_0}{\partial a_1} + \frac{\partial \sigma'_{11,0}}{\partial a_1} + \frac{\partial \sigma'_{13,0}}{\partial a_3} \\ \rho_{total} J \frac{dv_{avg,3}}{dt} &= -\rho_{total} g + \frac{\partial p_0}{\partial a_3} + \frac{\partial \sigma'_{13,0}}{\partial a_1} + \frac{\partial \sigma'_{33,0}}{\partial a_3} \end{aligned} \quad (42)$$

The initial and boundary conditions are

$$\begin{aligned} \mathbf{v}_{avg}(t=0) &= 0, \quad p(t=0) = \gamma_w a_3, \\ \sigma'_{11}(t=0) &= \sigma'_{13}(t=0) = 0, \\ \sigma'_{33}(t=0) &= \frac{\gamma_w (G_s - 1) a_3}{1 + e(t=0)}, \\ \sigma'(a_3=0) &= 0, \quad p(a_3=0) = 0 \end{aligned} \quad (43)$$

where γ_w is the unit weight of pore water. Moreover, the previous study [17] provided the following data:

$$T = 633 \text{ days}, \quad \lambda = 3.53 \quad H = 6.33 \text{ m}, \quad e_b = 6.936, \quad \text{and} \quad e_\infty = 3.64 \quad (44)$$

in which the high ratio e_b value implies that the phosphate waste clay layer is very soft. Meanwhile, training a deep neural network uses the Adam optimizer, learning rate β equal to 10^{-3} , and 10^5 epochs. This deep neural network has two hidden layers, and each hidden layer has 128 neurons. Discretizing the problem domain Ω_0 and Γ_f uses 2^{22} interior nodes and 2^9 boundary nodes.

After subtracting predicted consolidation settlements from the elevation of the retention pond, Figure 2 compares the remaining elevation. Figures 3(a)-3(d) selectively show the predicted $\sigma'_{11,0}$, $\sigma'_{33,0}$, $\sigma'_{13,0}$, and p_0 at the time $t = 30$ days. The unit of predicted $\sigma'_{11,0}$, $\sigma'_{33,0}$, $\sigma'_{13,0}$, and p_0 is kg/cm^2 .

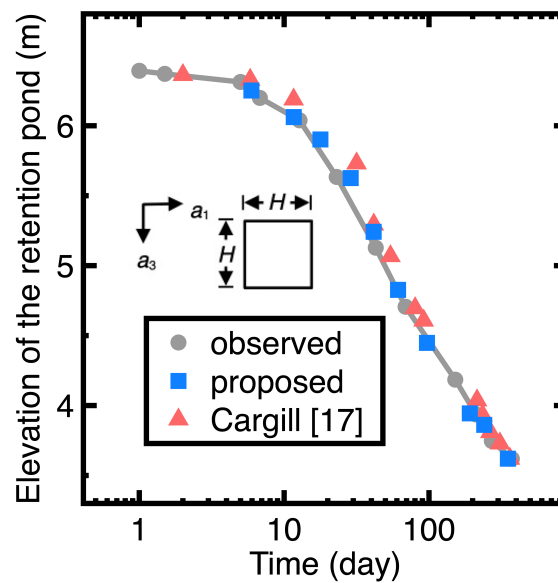


Figure 2. Comparison of the predicted and observed elevations of the retention pond

Figure 2 shows the necessity of developing a general mathematical model for studying a finite strain consolidation process. Considering the horizontal pore water dissipation, the current mathematical model provides more accurate predicted elevations of the retention pond, especially during 10-100 days. The corresponding maximum error $\left(= \frac{|\text{predicted value} - \text{true value}|}{\text{true value}} \right)$ is about 2.3%, whereas the previous one-dimensional model [17] outputs the maximum error about 4.2 %. Figures 3(a) and 3(b) show that the $\sigma'_{11,0}$ and $\sigma'_{33,0}$ vary symmetrically to two skewed vertical axes. Besides, Figure 3(c) demonstrates that a nonuniform $\sigma'_{13,0}$ exerts on the retention pond. It is impossible to use a one-dimensional mathematical model to output variations of the $\sigma'_{11,0}$, $\sigma'_{33,0}$, and $\sigma'_{13,0}$ shown in Figures 3(a)-3(c). Meanwhile, Figure 3(d) demonstrates that the pore water pressure p_0 varies uniformly across the depth of the retention pond.

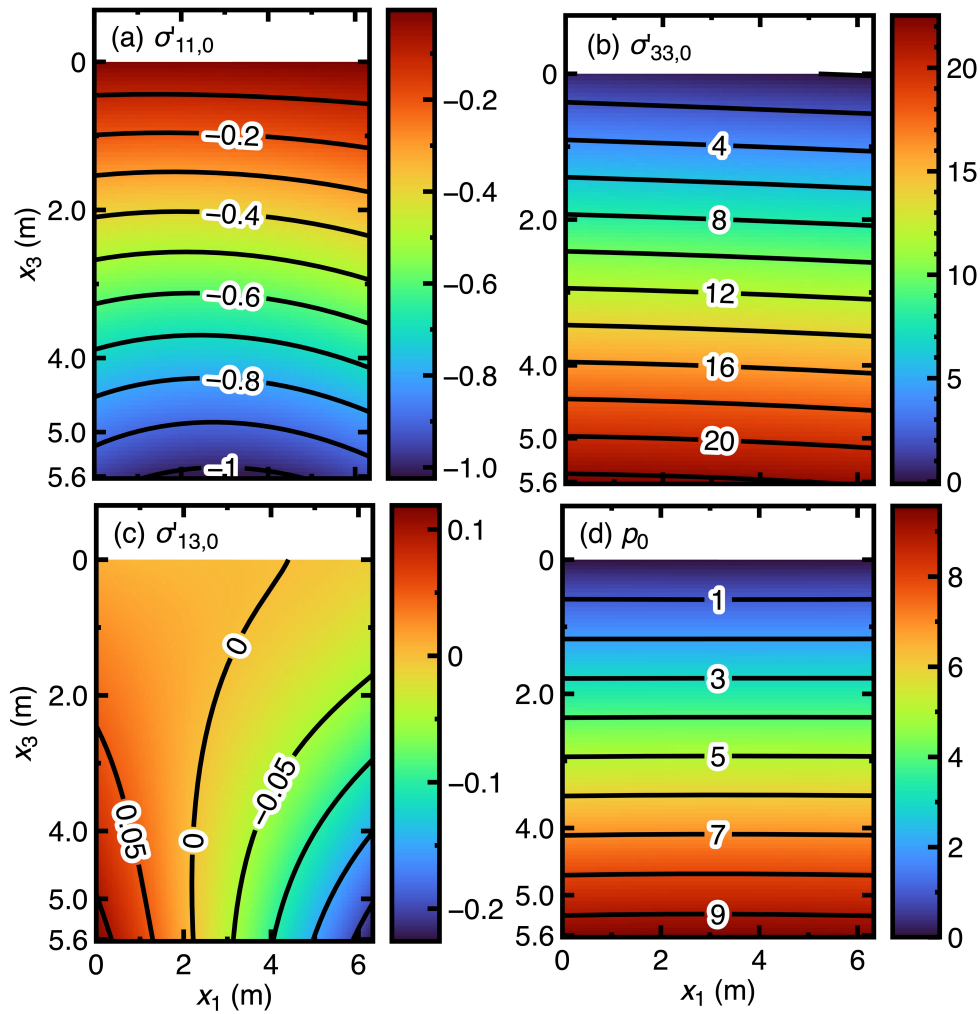


Figure 3. Predicted $\sigma'_{11,0}$, $\sigma'_{33,0}$, $\sigma'_{13,0}$, and p_0 of a phosphate waste layer consolidated by its self-weight at the time $t = 30$ days

Figure 4 demonstrates the convergence of the *Loss* value with respect to the number of epochs. This figure indicates that the *Loss* value decreases to a low value after running Algorithm 1 sufficiently.

If we desire the *Loss* values below 10^{-4} , Figure 4 indicates that we can adopt merely 4000 to 5000 epochs to generate such predictions. However, it may be interesting to study how to investigate the variation of *Loss* values with respect to the learning rate β and the number of interior or boundary nodes

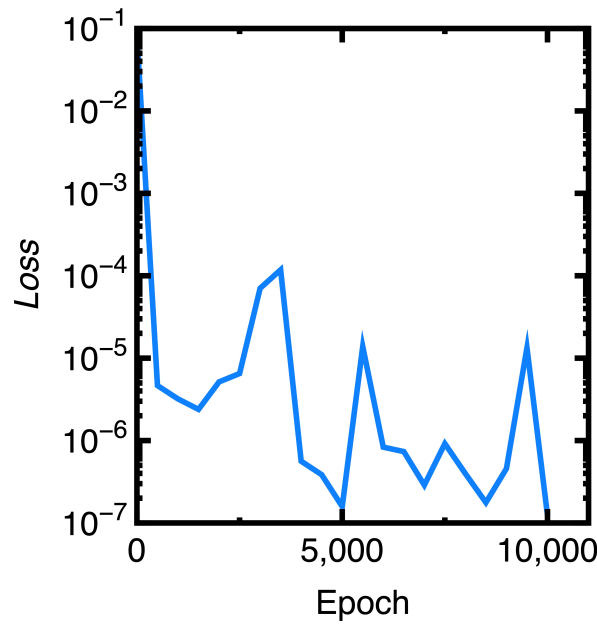


Figure 4. Convergence of the *Loss* value in predicting consolidation settlements of a phosphate waste clay layer

4.2. Osaka Bay Mud

Since Japan has limited land and a dense population, it has no choice but to build some offshore public facilities. Therefore, some published studies (e.g., [16]) investigated the consolidation behaviors of offshore clay layers near big cities such as Tokyo and Osaka. This study uses the results of a model test [16] whose goal is to monitor the consolidation process induced by the self-weight of an Osaka Bay mud layer to test the current mathematical model.

The previous study [16] reported that Osaka Bay mud has a liquid limit equal to 102.8, a plastic index identical to 45.8, and a specific yield G_s equal to 2.59. Another published study [18] provided an empirical relationship between the void ratio and vertical effective stress σ'_{33} :

$$e = 1.35 - 0.45 \log \left(\frac{\sigma'_{33}}{25 \text{ kPa}} \right) \quad (45)$$

Similar to Section 4.1, Equation (42) is still the governing equation for predicting consolidation settlements of the Osaka Bay mud layer with setting the initial problem domain identical to $0 \leq a_j \leq H$ ($j = 1, 3$) in which H is the thickness of the Osaka Bay mud layer. The direction of a_3 is identical to the direction of the gravitational force, whereas the direction of a_1 is from left to right. The small figure inside the subsequent Figure 5 further illustrates this problem domain Ω_0 . The initial and boundary conditions are

$$\begin{aligned} \mathbf{v}_{avg}(t=0) &= 0, \quad p(t=0) = \gamma_w a_3, \quad \sigma'_{33}(t=0) = \frac{\gamma_w(G_s-1)a_3}{1+e(t=0)} \\ \sigma'_{11}(t=0) &= \sigma'_{13}(t=0) = 0 \\ \sigma'(a_3=0) &= 0, \quad p(a_3=0) = 0 \\ \frac{\partial p(a_1=0)}{\partial a_1} &= \frac{\partial p(a_1=H)}{\partial a_1} = v_{avg,1}(a_1=0) = v_{avg,1}(a_1=H) = 0 \\ \frac{\partial p(a_3=H)}{\partial a_3} &= v_{avg,3}(a_3=H) = 0 \end{aligned} \quad (46)$$

Other necessary data are [16]:

$$T = 112.5 \text{ days}, \quad H = 80 \text{ cm}, \quad e(t=0) = 7.849 \quad (47)$$

in which the initial void ratio $e(t=0)$ indicates that the Osaka mud layer is softer than the phosphate waste clay layer in Section 4.1. Meanwhile, training a deep neural network uses the Adam optimizer, learning rate β equal to 10^{-3} , and 10^5 epochs. This deep neural network has two hidden layers, and each hidden layer has 128 neurons. Discretizing the problem domain Ω_0 and Γ_f uses 2^{22} interior nodes

and 2^9 boundary nodes. Figure 5 compares the predicted and measured vertical settlements. Figures 6(a)-6(d) selectively show variations of the predicted $\sigma'_{11,0}$, $\sigma'_{33,0}$, $\sigma'_{13,0}$, and p_0 at the time $t = 1.55 \times 10^4$ mins.

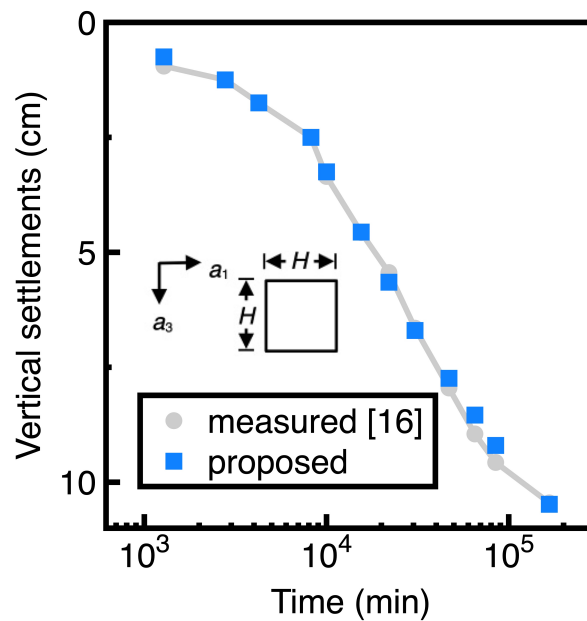


Figure 5. Comparison of predicted and measured vertical settlements of an Osaka Bay mud layer at the time $t = 1.55 \times 10^4$ mins

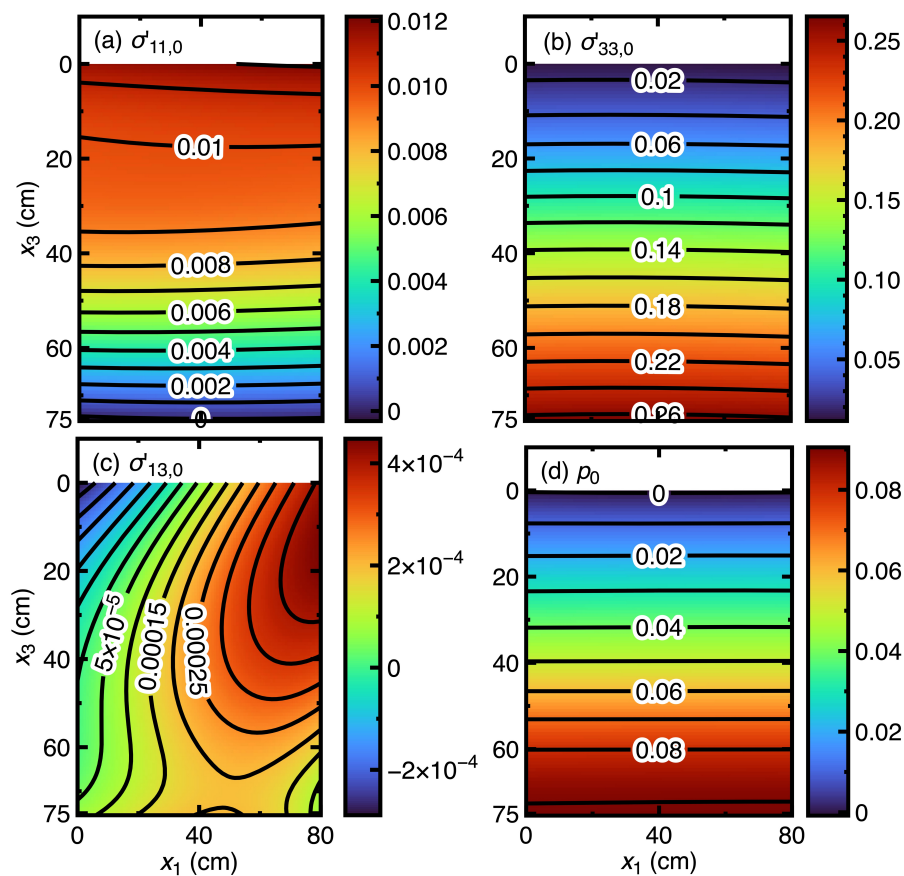


Figure 6. Predicted $\sigma'_{11,0}$, $\sigma'_{33,0}$, $\sigma'_{13,0}$, and p_0 of an Osaka Bay mud layer consolidated by its self-weights at the time $t = 1.55 \times 10^4$ mins

Figure 5 indicates that the current mathematical model provides accurate predictions of vertical settlements. Further inspection of the data used to plot this figure finds that the maximum error is below 5%. Although the Osaka Bay mud layer in this section is thinner than the phosphate waste clay layer in Section 4.1, Figure 6(a) still shows that nonuniform $\sigma'_{11,0}$ imposes the Osaka Bay mud layer. Larger stresses $\sigma'_{11,0}$ are imposed on the middle part of the Osaka Bay mud layer. Similarly, Figure 6(c) demonstrates that the shear stress $\sigma'_{13,0}$ varies non-uniformly. In contrast, Figure 6(b) and 6(d) shows that variations of the $\sigma'_{33,0}$ and p_0 are uniform.

Figure 7 visualizes the convergence of the *Loss* value with respect to the number of epochs. Similar to Section 4.1, if the desired *Loss* value is below 10^{-4} , this figure shows that we can adopt merely 2000 epochs to obtain such *Loss* values. Convergence of the *Loss* value is even faster in this section than in Section 4.1, although more complex boundary conditions are in this section.

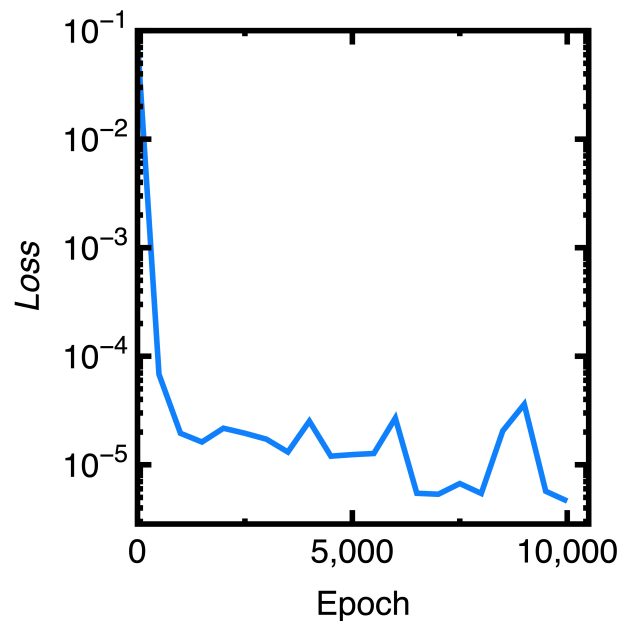


Figure 7. Convergence of the *Loss* value in predicting consolidation settlements of an Osaka Bay mud layer

4.3. Ablation Study

In Section 3.2, Equation (32) defines the *Loss* value. Several parameters can affect its convergence. Compared to the error analysis in the application of other numerical methods (for example, the finite element method) to solve a partial differential equation, this study chooses to inspect the *Loss* value with respect to different learning rates β and numbers of interior and boundary nodes. Obviously, increasing the number of hidden layers or neurons can provide lower *Loss* values. Discussing the influence of the number of hidden layers or neurons on the *Loss* value can't deliver new results.

Figures 8(a)-8(b) compare the *Loss* values with respect to the learning rates $\beta = 10^{-2}$ and 10^{-4} . Except for the learning rate β , creating Figure 8(a) uses the required data or settings for plotting Figures 2, 3(a)-3(d), and 4, whereas generating Figure 8(b) employs the required data or settings for drawing Figures 5, 6(a)-6(d), and 7.

Observing Figures 8(a)-8(b), one may find that training the deep neural network using the faster learning rate $\beta = 10^{-2}$ is not appropriate. The corresponding *Loss* value does not converge to an acceptable value (below 10^{-4}) in Figure 8(a). In contrast, the *Loss* value is below 10^{-4} when training the deep neural network over 4000 to 5000 epochs with the learning rate $\beta = 10^{-4}$.

Figures 9(a)-9(b) compare the *Loss* values versus different numbers of interior and boundary nodes. Except for the number of interior or boundary nodes, editing Figure 9(a) uses the data and settings required to prepare Section 4.1m while creating Figure 9(b) is based on the necessary data and settings employed to generate Figures 5, 6(a)-6(d), and 7. For keeping a particular nodal spacing, the total number of interior and boundary nodes is simultaneously increased in creating Figure 9(a)-9(b).

Increasing the total number of interior and boundary nodes means a finer discretization of the problem domain. Since this study creates random interior and boundary nodes, Figures 9(a)-9(b) indicate that generating a finer discretization of the problem domain does not apparently improve the *Loss* value. These two figures may show a distinguishing characteristic, which does not exist in other existing numerical methods (for example, the finite element method). A finer discretization of the problem domain can apparently improve the accuracy of numerical results.

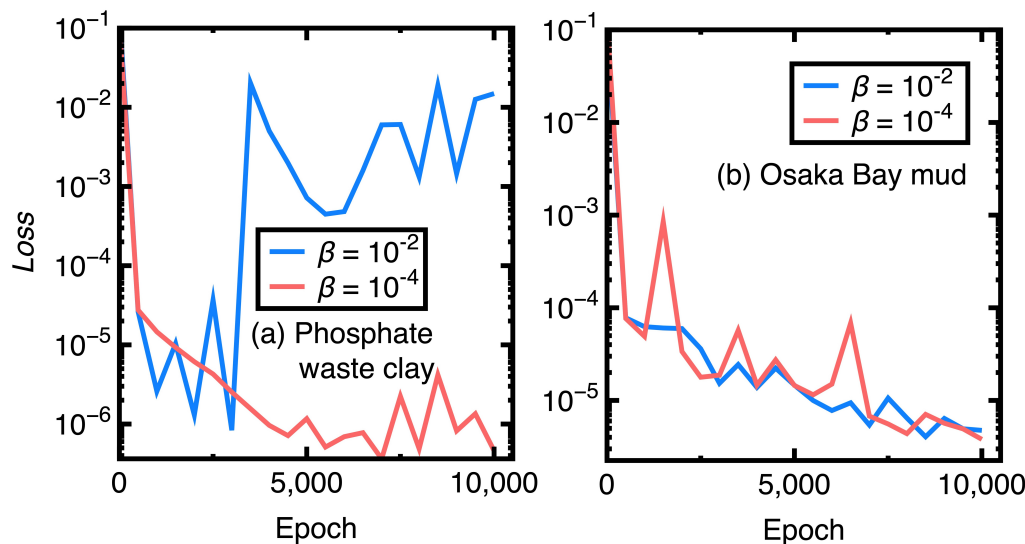


Figure 8. Comparison of the *Loss* values with respect to different learning rates β : (a) the phosphate waste clay (b) Osaka Bay mud

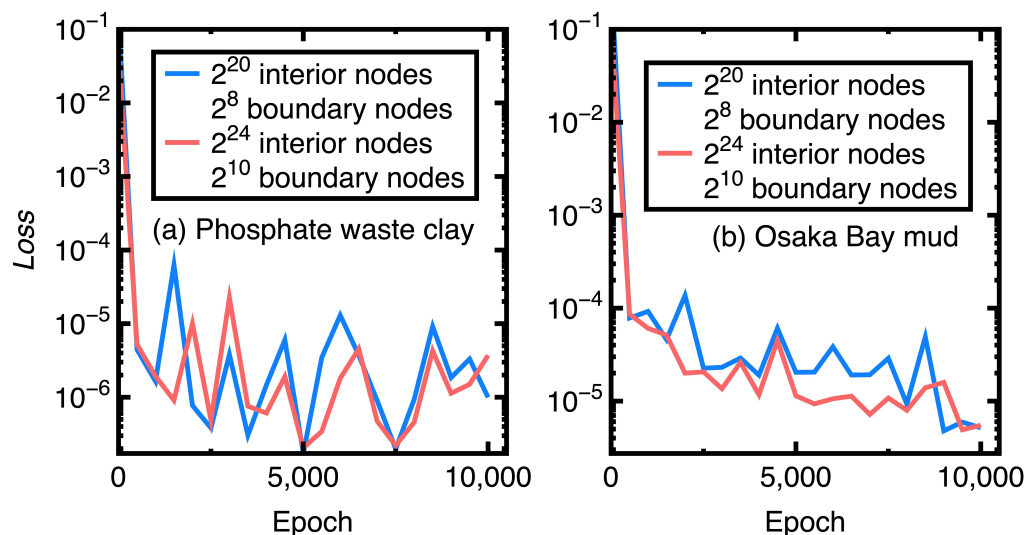


Figure 9. Comparison of the *Loss* values with respect to different numbers of interior and boundary nodes: (a) the phosphate waste clay (b) Osaka Bay mud

5. Discussion

Sections 4.1 and 4.2 reveal four difficulties of developing a mathematical model for an engineering problem; however, this study resolves these difficulties to a certain extent:

1. The first difficulty is the challenge of balancing the number of assumptions and the simplicity of the corresponding governing equations. This study eliminates the assumption that a clay layer consolidates only in the vertical direction in modeling a finite strain consolidation process; however, the current governing equation (Equation (29)) is not complex.

2. The second difficulty is the harassment of choosing a suitable numerical method for solving a real-world problem. For this study, it is difficult to define some boundary conditions of a natural soft clay layer; nevertheless, boundary conditions are prerequisites for implementing existing numerical methods (for example, the finite element method). Section 4.1 provides an example. Probably due to the lack of field measurements, some boundary conditions were unavailable in the previous study [17]. However, the DGM can resolve this difficulty since its goal is to minimize the *Loss* value at random nodes.
3. The third difficulty is the requirement that the total number of governing equations equals the total number of unknowns. The DGM helps eliminate this requirement. In Sections 4.1 and 4.2, two governing equations are available, but six unknowns (pore water pressure, two average velocity components, and three effective stress components) exist. If this study does not adopt the DGM, reducing the total number of unknowns must be implemented using available material properties (for example, a void ratio-stress relationship). The author's Ph.D. thesis provided an example in which the void ratio is the unknown of a single and complex governing equation.
4. The fourth difficulty is the non-homogeneity of clay's properties. This difficulty represents the limitation of this study. Natural clay's properties are non-homogeneous. Although we can create a particular probability model to regress the distribution of a clay's property, there must be enough clay samples to provide accurate regression points. However, gathering clay samples of a natural soft clay layer is not easy. Probably due to this reason, it is unavoidable that the accuracy of predicted consolidation settlements is limited.

Meanwhile, Section 4.3 demonstrates that the DGM may be a promising numerical method for studying engineering problems. Even in a finite-strain engineering problem, we only need to adjust the learning rate β in implementing a particular deep neural network.

6. Conclusions and Concluding Remark

This study develops a mathematical model for describing a generalized finite strain consolidation process and its DGM solution. Developing this mathematical model adopts the Reynolds transport theorem to describe mass and momentum balances for clay grain and pore water phases. The governing equation is the sum of the resulting mass and momentum balance equations. Applying the DGM to train a deep neural network to minimize a loss function defined using the resulting governing equation, available initial, and boundary conditions, is next implemented. Two real-world problems show that the current mathematical model outperforms previous one-dimensional models in predicting the consolidation settlements caused by the self-weights of two natural soft clay layers. Besides, an additional ablation study finds that the learning rate has an apparent effect on the value of the loss function.

Based on Sections 4.1-4.3, the conclusions of this study are

1. Deriving the current mathematical model advances the modeling of a finite-strain engineering problem. The current governing equation is simple but adapts to the changes of the problem domain. Based on it, we can predict more accurate settlements.
2. The DGM is a promising numerical method for resolving the difficulties of deducing sufficient governing equations or specifying full initial and boundary conditions in solving a partial differential equation.
3. To obtain the desired accuracy of numerical results provided by the DGM, adopting a lower learning rate in training a deep neural network is preferred.

Furthermore, the limitation of this study is the non-homogeneity of clay properties. Since implementing the DGM uses random interior and boundary nodes, constructing particular probability models to regress the distributions of clay properties and combining the resulting probability model with a DGM solution may be the main interest of future research.

Funding: This research is the author's private work and received no external funding.

Data Availability Statement: Data used to implement Sections 4.1 and 4.2 are digitized from two previous studies [16,17]. They are available on <https://github.com/xsheu/consolidation>.

Acknowledgments: The author completed this article without any financial support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DGM Deep Galerkin Method

References

1. Loktev, K.A.; Ulanov, I.; Shishkina, I.; Savulidi, M.; Klekovkina, N.; Kuznetsov, A. Determination of settlement parameters of highway embankment and base consolidation time depending on soil characteristics. *Transp. Res. Proc.* **2022**, *63*, 946–955.
2. Kamash, W.E.; Hafez, K.; Zakaria, M.; Moubarak, A. Improvement of soft organic clay soil using vertical drains. *KSCE. J. Civ. Eng.* **2021**, *25*(2), 429–441.
3. Cao, L.F.; Chang, M.-F.; Teh, C.I. and Na, Y.M. Back-calculation of consolidation parameters from field measurements at a reclamation site. *Can. Geotech. J.* **2001**, *38*(4), 755–769.
4. Biot, M.A. General theory of three-dimensional consolidation. *J. Appl. Phys.* **1941**, *12*(2), 155–164.
5. Gibson, R.E.; England, G.L.; Hussey, M.J.L. The theory of one-dimensional consolidation of saturated clays. *Géotechnique* **1967**, *17*(3), 261–273.
6. Gibson, R.E.; Schiffman, R.L.; Cargill, K.W. Back-calculation of consolidation parameters from field measurements at a reclamation site. *Can. Geotech. J.* **2001**, *38*(4), 755–769.
7. Jeeravipoolvarn, S.; Scott J.D.; Chalaturnyk R. Multi-dimensional finite strain consolidation theory: Modeling study. In Proceedings of the 61st Canadian Geotechnical Conference and the 9th Joint CGS/IAH-CNC Groundwater Conference, Edmonton, AB, Canada, September 21–24, 2008; 167–175.
8. Sheu, G.Y. A general finite strain consolidation theory. Ph.D. Thesis, National Cheung Kung University, Tainan, Taiwan, 1997.6.30.
9. Malvern, L.E. *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall: Englewood Cliffs, New Jersey, USA, 1977.
10. Justin S.; Konstantinos S.: DGM: A deep learning algorithm for solving partial differential equations. *J. Comput. Phys.* **2018**, *375*, 1339–1364.
11. Huerta, A.; Rodriguez, A.: Numerical analysis of nonlinear large-strain consolidation and filling. *Degree-Comput. Struct.* **1992**, *44*(1), 357–365.
12. Liu, S.J.; Sun, H.L.; Pan, X.D.; Shi, L.; Cai, Y.Q.; Geng, X.Y.: Analytical solutions and simplified design method for large-strain radial consolidation. *Comput. Geotech.* **2021**, *134*, 103987.
13. Kumar, H.; Yadav, N.; Nagar, A.K.: Numerical solution of Generalized Burger-Huxley & Huxley's equation using deep Galerkin neural network method. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105289.
14. Matsumoto, M.: Application of deep Galerkin method to solve compressible Navier-Stokes equations. *Trans. Jpn. Soc. Aeronaut. Space Sci.* **2021**, *64*(6), 348–357.
15. McVay, M.; Townsend, F.; Bloomquist, D.: Quiescent consolidation of phosphatic waste clays. *J. Geotech. Eng.* **1986**, *112*(11), 1033–1049.
16. Zen, K.; Umehara, Y. (2019). A new consolidation testing procedure and technique for very soft soils. In *Consolidation of Soils: Testing and Evaluation*; R. N. Yong; F. C. Townsend, Eds.; American Society for Testing and Materials: Philadelphia, USA; pp. 405–432.
17. Cargill, K. W.: Prediction of consolidation of very soft soil. *J. Geotech. Eng.* **1984**, *110*(6), 775–795.
18. Hiroyuki T.; Jacques L.: A microstructural investigation of Osaka Bay clay: the impact of microfossils on its mechanical behaviour. *Can. Geotech. J.* **1999**, *36*(3), 493–508.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.