



INRAe

Research internship on multi-scale geomechanics:
Machine learning from a mesoscale perspective

Author: Wenqing QU

Supervisor: Antoine WAUTIER

September 20, 2024

CONTENTS

| | |
|---|------------|
| List of Figures | vi |
| Abstract | vii |
| 1 Introduction | 1 |
| 1.1 Review and analysis of the H-model | 1 |
| 1.1.1 The H-model in brief | 2 |
| 1.1.2 Biaxial test at the material point scale | 9 |
| 1.2 Internship Topic: multi-scale geomechanics: Machine learning from a mesoscale perspective | 10 |
| 2 Methods | 13 |
| 2.1 Introduction to Artificial Neural Networks (ANNs) | 13 |
| 2.2 Motivation for Using ANNs in Granular Material Modeling | 14 |
| 2.3 Training Data Generation and Preprocessing | 14 |
| 2.4 Network Architecture and Design | 17 |
| 2.5 Model Validation and Evaluation | 18 |
| 3 Results and discussion | 23 |
| 3.1 Comparative Analysis with the original analytical H-model | 23 |
| 3.2 Future Work | 24 |
| Bibliography | 27 |

LIST OF FIGURES

| | | |
|-----|---|----|
| 1.1 | The directional distributions and the coordinates at the REV scale (e_1, e_2) and the mesoscale (n, t). | 2 |
| 1.2 | The symmetrical description of geometry (a) and forces (b) at the elementary hexagonal pattern of adjoining particles (H-cell). There are two types of contact c_1 and c_2 , the corresponding contact branches are d_1 and d_2 | 2 |
| 1.3 | General homogenization scheme relating both incremental stress and strain tensors. | 3 |
| 1.4 | Three options of the meso-volume: hexagonal domain (blue), rectangular domain going through the grain centers (green), bounding box domain (red) . . . | 7 |
| 1.5 | Two modes of pathological H-cell due to contact loss. | 8 |
| 1.6 | Two modes of over-compressed H-cell with the corresponding opening angle. . . | 9 |
| 1.7 | The stress ratio and volumetric strain curves based on the parameters from Veylon, 2017. Three confining pressures are considered. | 10 |
| 1.8 | The deviatoric stress and volumetric strain along the biaxial loading with $p_0 = 100$ kPa until a relative large axial strain. | 10 |
| 2.1 | Display the distribution of the total strains ($\delta\epsilon_1, \delta\epsilon_2$) variable in dataset 1 and incremental strains ($\delta\epsilon_1, \delta\epsilon_2$) in dataset 2 | 15 |
| 2.2 | The red-colored points represent the outliers that did not conform to the criteria and were subsequently removed. | 15 |
| 2.3 | Model 1 architecture | 17 |
| 2.4 | Model 1: Performance of 3 Hidden Layers Trained with MAE(a) and MSE(b) Loss Functions with 6k datasets | 19 |
| 2.5 | Model 8: Performance of 17 Hidden Layers Trained with MAE(a) and MSE(b) Loss Functions with 6k datasets | 20 |
| 2.6 | Model 1: Performance of 3 Hidden Layers Trained with MAE(a) and MSE(b) Loss Functions with 600k datasets | 20 |
| 2.7 | Test results for Model 1 with MAE(a) and MSE(b) Loss Functions with 60k datasets | 21 |

| | | |
|-----|--|----|
| 3.1 | Isotest results for Model trained with robust scaling with MAE Loss Functions with 600k datasets | 23 |
| 3.2 | Biaxial test results for Model trained with robust scaling with MAE Loss Functions with 600k datasets with lateral confining stress of 200 kPa | 24 |
| 3.3 | Undrained biaxial tests results for Model trained with robust scaling with MAE Loss Functions with 600k datasets | 24 |

ABSTRACT

Granular materials display complex behaviors that require extensive computational simulations to reveal their macroscopic properties. In this report, we embark on a transformative exploration of machine learning techniques to enhance the simulation of granular materials, with a focus on one particular multiscale model, namely the H-model. Such models propose to describe the overall behavior of granular materials as the average response of a collection of idealized structures of a few grains.

Our methodology revolves around the integration of machine learning models to replace the traditional 2D H-model structure(a regular hexagon formed of six grains of equal diameter), with the goal of improving computational efficiency and paving the way for the use of more complex structures . This endeavor unfolds through meticulous stages that involve data generation, data curing, data preprocessing, model construction, and comprehensive evaluation.

The results highlight the advantage of using machine learning to predict the behavior of the small stucture of grains instead of trying to predict directly the material behavior. First, considering a simple structure enable to construct easily a large database which cannot be acheived easily at the material scale. Then, the learning process preserve as much as possible the description of the physics of granular material by learning only the most difficul part of the derivation of the constitutive behavior. The culmination of our research underscores the significant advantages resulting from the integration of machine learning into geomechanics. These benefits extend beyond mere computational acceleration and encompass the acquisition of invaluable insights into the nuanced behaviors exhibited by granular materials.

1 INTRODUCTION

Granular materials have long fascinated researchers because of their intricate behaviors. Despite an apparent simplicity at the grain scale (solid bodies interacting through contacts), predicting the mechanical behavior of granular materials frequently require intensive computational simulations to reveal their macroscopic properties. Within this complex domain, the H-model stands out as a micromechanical-based constitutive model that provides insights into the behavior of granular materials. The following section will provide a brief introduction centered around the H model and the objectives of the internship project.

1.1 Review and analysis of the H-model

Proposed by Nicot and Darve in 2011, the H-model provides a valuable framework for describing the behavior of granular materials, with a particular focus on the distribution of mesostructures oriented in different directions.

The H-model introduces a novel approach to representing granular assemblies, utilizing a distribution of mesostructures that exhibit hexagonal patterns in two dimensions (2D) and bi-hexagonal patterns in three dimensions (3D). These mesostructures, known as H-cells, capture the essential characteristics of the granular material. Notably, the H-cells possess symmetric geometry and forces. For the sake of simplicity, we focus on the 2D H-model in the project.

One of the key advantages of the H-model lies in its ability to consider the collective rearrangement of granular materials by accounting for the deformation of the H-cells. This feature enables the model to capture the geometric effects that influence the macroscopic behavior of granular assemblies. By incorporating the deformations of mesostructures, the H-model provides insights into the intricate mechanisms underlying granular behavior, enhancing our understanding of complex granular systems.

1.1.1 The H-model in brief

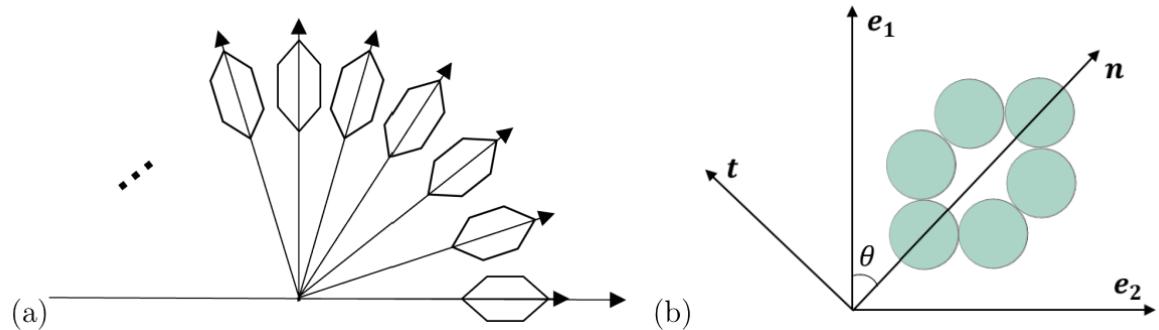


Figure 1.1. The directional distributions and the coordinates at the REV scale (e_1, e_2) and the mesoscale (n, t).

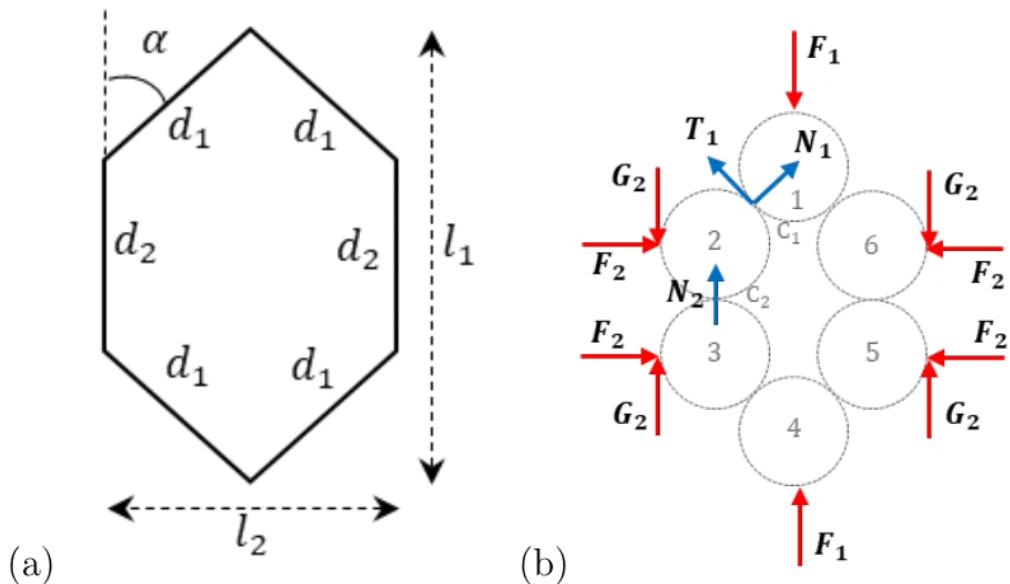


Figure 1.2. The symmetrical description of geometry (a) and forces (b) at the elementary hexagonal pattern of adjoining particles (H-cell). There are two types of contact c_1 and c_2 , the corresponding contact branches are d_1 and d_2 .

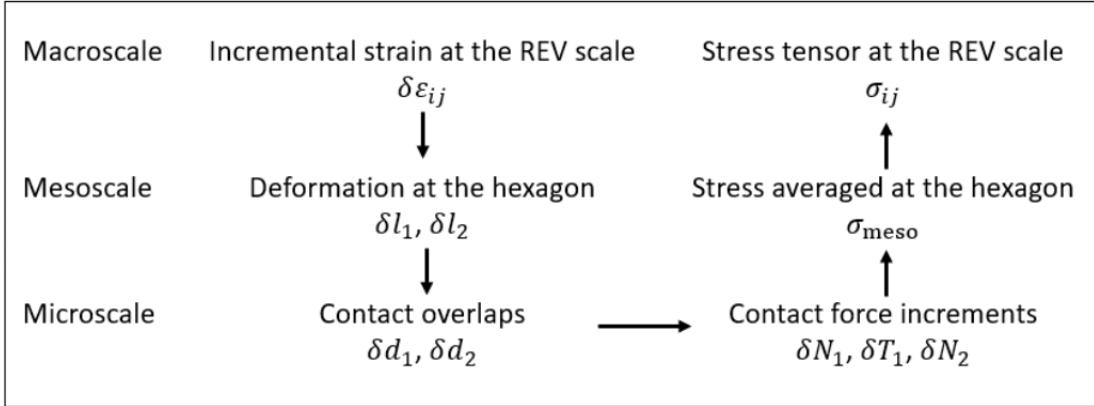


Figure 1.3. General homogenization scheme relating both incremental stress and strain tensors.

In the H-model, a granular assembly is described by a distribution of regular hexagons (the mesostructures) oriented in different directions, as shown in Figure 1.1 (a). The coordinate system for each hexagon is presented by (e_1, e_2) in the global frame and (n, t) in the local frame, as shown in Figure 1.1 (b). The collective rearrangement in granular materials is accounted in the H-model through the deformation of the H-cells. The geometry and forces at the H-cell are symmetric, as presented in Figure 1.2. Because of the symmetry of the H-cell, there are only two types of contact: c_1 and c_2 at each H-cell. Behaviors at different scales (the micro scale, the meso scale and the macro scale) are related based on the homogenization scheme in Figure 1.3. Thus the H-model accounts for geometric effects in the macroscopic behavior of granular material. In the following we detail each of the steps illustrated in Fig 1.3.

Strain localization hypothesis

The global incremental strain $\delta\varepsilon_{ij}$ is passed to each hexagon according to the equations as follows:

$$\begin{cases} \delta l_1 = -l_1 \delta\varepsilon_{ij} n_i n_j \\ \delta l_2 = -l_2 \delta\varepsilon_{ij} t_i t_j \end{cases} \quad (1.1)$$

These formula corresponds to the relative length variation in direction n and t . Note that no distortion is taken into account in the model for the sake of preserving the symmetry of the H-cell.

Directional behavior

Each H-cell representing one direction can be as illustrated in Figure 1.2. The state of a H-cell can be defined by three geometrical parameters: the opening angle α , and the inter-

granular distances d_1 and d_2 . These parameters relate to the cell dimensions l_1 and l_2 as

$$\begin{cases} l_1 = d_2 + 2d_1 \cos\alpha \\ l_2 = 2d_1 \sin\alpha, \end{cases} \quad (1.2)$$

The mechanical behavior of each hexagon depends on both the contact law between adjoining grains and the mechanical equilibrium with external forces. An elastic-frictional contact law is adopted. Such a contact law relies on three parameters: a normal stiffness k_n , a tangential stiffness k_t , and an inter-granular friction angle ϕ_g . The normal and tangential contact forces can be expressed in an incremental form as

$$\begin{cases} \delta N_i = k_n \delta u_n^i \\ \delta T_i = \min[|T_i + k_t u_t^i|, \tan\phi_g(N_i + \delta N_i)] \xi - T_i, \end{cases} \quad (1.3)$$

where ξ is the sign of the quantity $T_i + k_t u_t^i$, u_n^i and u_t^i are the normal and tangential relative displacements of the grains at contact i . Given the symmetries in the H cell, we only need to consider the contact forces called N_1 , N_2 and T_1 as shown in Figure 1.2, The corresponding incremental displacements are obtained by differentiating the expression of the branch vector joining the grains in contact:

$$\begin{aligned} \delta u_n^1 &= -\delta d_1 \\ \delta u_n^2 &= -\delta d_2 \\ \delta u_t^1 &= -d_1 \delta \alpha, \end{aligned} \quad (1.4)$$

Mechanical balance of grain 2 reads, along directions n tells that:

$$N_2 = N_1 \cos\alpha + T_1 \sin\alpha + G_2 = N_1 \cos\alpha + T_1 (\sin\alpha + 1) \quad (1.5)$$

Equation 1.5 differentiates into

$$\cos\alpha \delta N_1 - (G_2 + \sin\alpha) \delta T_1 - \delta N_2 + (N_1 \sin\alpha - T_1 \cos\alpha) \delta \alpha = 0 \quad (1.6)$$

Combining these equations at the H-cell scale, the incremental evolution δl_1 and δl_2 can be related to δd_1 , δd_2 and $\delta \alpha$ in matrix form as

$$A \begin{bmatrix} \delta d_1 \\ \delta d_2 \\ \delta \alpha \end{bmatrix} = \begin{bmatrix} \delta l_1 \\ \delta l_2 \\ \lambda \end{bmatrix}, A = \begin{bmatrix} 2\cos\alpha & 1 & -2d_1 \sin\alpha \\ 2\sin\alpha & 0 & 2d_1 \cos\alpha \\ A_{31} & -1 & A_{33} \end{bmatrix} \quad (1.7)$$

where

$$A_{31} = \begin{cases} \cos\alpha & \text{if no sliding occurs} \\ \cos\alpha + \xi \tan\phi_g (i_{G_2} + \sin\alpha) & \text{if sliding occurs} \end{cases} \quad (1.8)$$

$$A_{33} = \begin{cases} [N_1 \sin\alpha - T_1 \cos\alpha - k_t d_1 (i_{G_2} + \sin\alpha)]/k_n & \text{if no sliding occurs} \\ (N_1 \sin\alpha - T_1 \cos\alpha)/k_n & \text{if sliding occurs} \end{cases} \quad (1.9)$$

$$\lambda = \begin{cases} 0 & \text{if no sliding occurs} \\ \xi(1 + \sin\alpha)(\tan\phi_g N_1 - T_1) & \text{if sliding occurs} \end{cases} \quad (1.10)$$

Based on the updated variables δd_1 , δd_2 and $\delta\alpha$, the contact evolution (δN_1 , δN_2 and δT_1) can be computed according to Equation 1.7. This complex relationship between the H-cell geometry and the contact forces is the main focus of the internship, as precised in section.

Stress averaging

The macroscopic stress tensor σ at a material point is calculated by a directional averaging of all contact forces acting in a collection of unit H-cells (belonging to a virtual REV corresponding to this material point). This is done in two steps by first defining a meso stress σ_{meso} for each H-cell and then averaging all the meso-stresses of the H-cell collection.

The meso-stress tensor is defined at the scale of the unit H-cell using the Love-Weber formula:

$$\sigma_{\text{meso}} = \frac{1}{V_{\text{meso}}} \sum_{c \in V_{\text{meso}}} f_c \otimes l_c \quad (1.11)$$

where f_c is the contact force at contact c , l_c the branch vector joining the centers of the two grains in contact and V_{meso} the volume of the unit H-cell.

Expressed in the local frame (n, t) , equation 1.11 gives with soil mechanics conventions (positive compression).

$$\begin{cases} V_{\text{meso}} \sigma_{nn}^{\text{meso}} = 4N_1 d_1 \cos^2\alpha + 4T_1 d_1 \cos\alpha \sin\alpha + 2N_2 d_2 \\ V_{\text{meso}} \sigma_{tt}^{\text{meso}} = 4N_1 d_1 \sin^2\alpha - 4T_1 d_1 \cos\alpha \sin\alpha \\ V_{\text{meso}} \sigma_{nt}^{\text{meso}} = V_{\text{meso}} \sigma_{tn}^{\text{meso}} = 0 \end{cases} \quad (1.12)$$

Finally, the macroscopic stress tensor for the REV can be determined as an integration of all meso-stresses of individual unit H-cells, knowing their statistical distribution via the probability density function $\omega(\theta)$. Thus, for the entire collection of unit H-cells,

$$\boldsymbol{\sigma} = \frac{1}{V} \int_{\theta=0}^{\pi} \omega(\theta) V_{\text{meso}}(\theta) \boldsymbol{\sigma}_{\text{meso}}(\theta) d\theta \quad (1.13)$$

such that

$$V = \int_{\theta=0}^{\pi} \omega(\theta) V_{\text{meso}}(\theta) d\theta \text{ and } \int_{\theta=0}^{\pi} \omega(\theta) d\theta = 1 \quad (1.14)$$

If $\boldsymbol{\sigma}_{\text{meso}}$ is expressed as a matrix $\boldsymbol{\sigma}_{\text{meso}}^{n,t}$ in the local frame n, t , the stress matrix $\boldsymbol{\sigma}$ in the global frame e_1, e_2 is readily obtained by conversion from local to global configuration, i.e.

$$\boldsymbol{\sigma}^{e_1, e_2} = \frac{1}{V} \int_{\theta=0}^{\pi} \omega(\theta) V_{\text{meso}}(\theta) \mathbf{P}^{-1} \boldsymbol{\sigma}_{\text{meso}}^{n,t}(\theta) \mathbf{P} d\theta \quad (1.15)$$

where \mathbf{P} is the transformation matrix,

$$\mathbf{P} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (1.16)$$

The probability density function $\omega(\theta)$ describes the statistical distribution of unit H-cells and accounts for possible anisotropy. It is assumed that the initial distribution of unit H-cells can be approximated for instance as

$$\omega(\theta) = \frac{1}{\pi} [1 + \alpha_\omega \cos 2(\theta - \beta_\omega)] \quad (1.17)$$

where α_ω is a parameter ranging from 0 to 1 describing anisotropy of the hexagon cell distribution, and β_ω its major principal direction. In practice, the distribution is discretized in n_θ directions.

The global volume in Equ.1.15 reads as $V = \int_{\theta=0}^{\pi} \omega(\theta) V_{\text{meso}}(\theta)$. Note that the choice of V_{meso} is not unique as long as the strain localization equation 1.1 is updated accordingly to ensure that $\frac{\delta V_{\text{meso}}}{V_{\text{meso}}} = Tr(\delta \boldsymbol{\sigma})$. There are three main options for the volume of H-cell as shown in Fig.1.4. From the smallest to the largest, they are labelled as V_{meso}^1 , V_{meso}^2 and V_{meso}^3 . With these different meso-domains being considered, there are some changes in the porosity and the strain localization equation 1.1. In the project, V_{meso}^1 is used in the following.

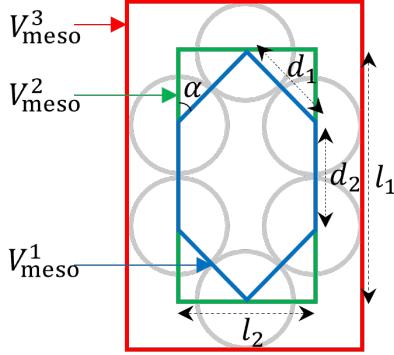


Figure 1.4. Three options of the meso-volume: hexagonal domain (blue), rectangular domain going through the grain centers (green), bounding box domain (red)

- To consistent with Equation 1.1, it is the rectangular domain (green in Fig.1.4) should be considered. The volume reads:

$$V_{\text{meso}}^2 = l_1 l_2 = 2d_1 \sin\alpha(d_2 + 2d_1 \cos\alpha) \quad (1.18)$$

- When the bounding box domain (red in Fig.1.4) is considered, the volume reads:

$$V_{\text{meso}}^3 = (l_1 + d_0)(l_2 + d_0) \quad (1.19)$$

In this case, Equ.1.1 should be updated with $l_1 + d_0$ and $l_2 + d_0$ instead of l_1 and l_2 . The geometric compatibility (Equation 1.2) matrix A (Equ.1.7) hold.

- When the hexagonal domain (blue in Fig.1.4) is considered, the meso-volume V_{meso}^h reads,

$$V_{\text{meso}}^1 = 2d_1 \sin\alpha(d_2 + d_1 \cos\alpha) \quad (1.20)$$

A virtual rectangle with the same area of the hexagon can be considered to make the same volume adopted in both strain localization and stress averaging processes. A rectangle of $l_1^* \times l_2^*$ can be proposed. With this hypothesis, Equ.1.2 should be replaced by:

$$\begin{cases} l_1^* = d_2 + d_1 \cos\alpha \\ l_2^* = 2d_1 \sin\alpha \end{cases} \quad (1.21)$$

The matrix A needs to be updated as A^* with two elements replaced: $A_{11}^* = \cos\alpha$ and $A_{13}^* = -d_1 \sin\alpha$.

In the original version of the H-model, V_{meso} is considered as the area of the hexagon, even though this volume is not consistent with the volume $l_1 l_2$ for strain localization in Equation 1.1. The option is inspired by the fact that granular materials in 2D can be tessellated by contact branches.

The general scheme of the standard 2D H-model can be summarized as follows:

1. Based on given $(\delta\varepsilon_1, \delta\varepsilon_2)$, update the H-cell dimensions l_1, l_2 of each H-cell according to a kinematic localization hypothesis.
2. Update d_1, d_2, α through Equation 1.7.
3. Update N_1, N_2, T_2 based on the incremental evolution of the H-cell geometry.
4. A meso-stress is defined based on the application of Love-Weber formula for each H-cell.
5. The macroscopic stress is eventually obtained by statistical averaging of all the meso-stresses.

In conclusion, the 2D H-model relies on three contact parameters (k_n, k_t and ϕ_g), on one geometrical parameter (the initial opening angle α_0), and on α_ω and β_ω for tuning the microstructure anisotropy.

Pathological cases

Pathological types of H-cell may be obtained owing to contact loss or gain during external loading. There are two modes of abnormal H-cell due to contact open as shown in Fig.1.5. The meso-stress on the broken H-cell is 0. The H-cell can be over-compressed when $\alpha \leq 30^\circ$ or $\alpha \geq 90^\circ$, as shown in Fig.1.6. In this case, additional contacts are generated with the associated contact forces. They contribute to the mesoscopic stress through additionnal terms in equation 1.11.

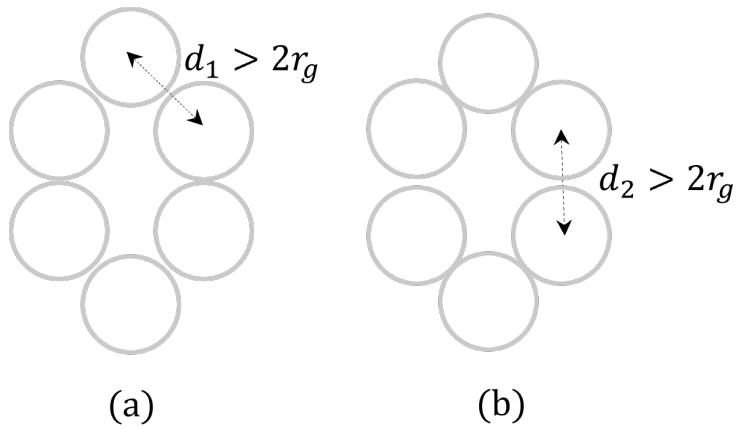


Figure 1.5. Two modes of pathological H-cell due to contact loss.

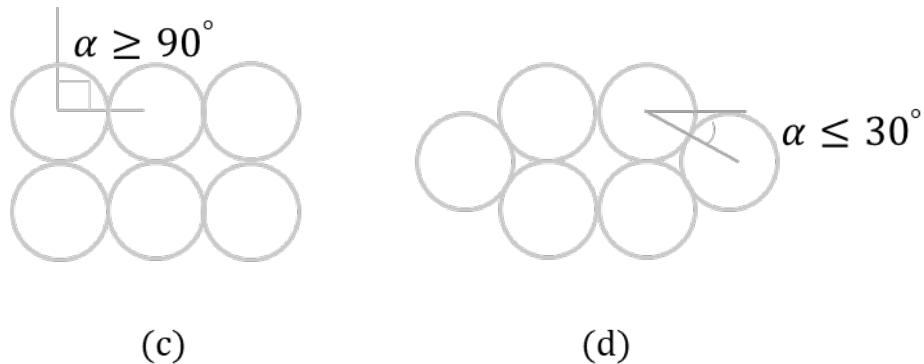


Figure 1.6. Two modes of over-compressed H-cell with the corresponding opening angle.

1.1.2 Biaxial test at the material point scale

Biaxial tests at the material point scale (homogeneous sample) are standard mechanical tests in the field of geomechanics. They can be simulated using the H-model based on the codes implemented in Python. Under this type of loading, a constant lateral stress condition is imposed together with a constant strain rate in the axial direction: $\delta\varepsilon_1 = \text{const}$ and $\delta\sigma_2 = 0$. In practice, as the H-model formulation is strain controlled, an iterative calculation process is run to find incremental strains ($\delta\varepsilon_1, \delta\varepsilon_2$) that gives the target confining pressure. The parameters (Veylon, 2017) are shown in Table 1.1. Three confining pressures 100 kPa, 200 kPa and 400 kPa are considered. The number of directions θ_i is set to 360.

Stress and strain curves are presented in Fig.1.7. It can be observed that, with the increase in p_0 , there are larger maximum stress ratios and stronger dilatancy. The trend is consistent to the results reported in [Veylon, 2017](#). If we pursue the loading up to a large level of axial strain (ε_{11} was 40%), the behavior is obviously not physically relevant. As shown in Fig.1.8, the deviatoric stress decreases to a very low value and the volumetric strain turns to contractancy at a large level of axial strain. This is due to the fact that a large number hexagons are broken and do not participate to stress transmission. Details are presented in the following section.

Table 1.1. Parameters used in biaxial tests Veylon, 2017

| k_n (N/m) | k_t/k_n | ϕ_g | α_0 | α_ω | β_ω | p_0 (kPa) |
|-------------|-----------|----------|------------|-----------------|----------------|---------------|
| 2e8 | 0.5 | 30° | 45° | 0 | 0 | 100, 200, 400 |

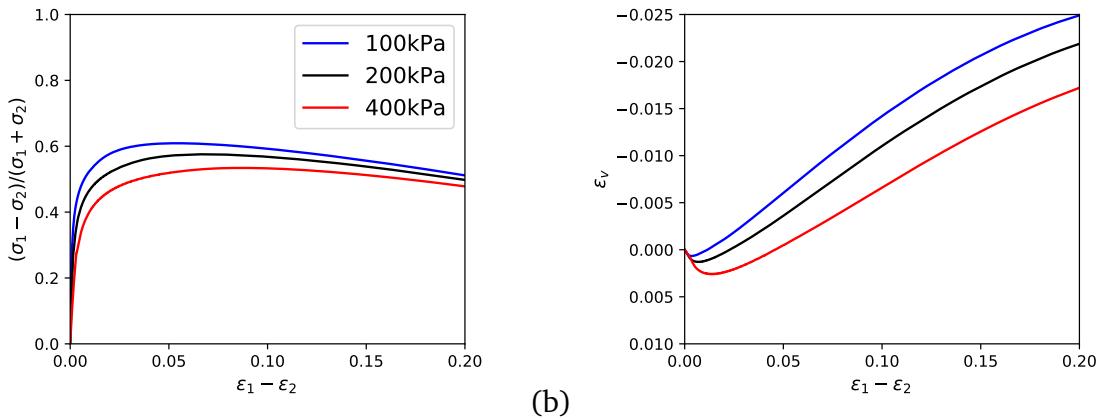


Figure 1.7. The stress ratio and volumetric strain curves based on the parameters from [Veylon, 2017](#). Three confining pressures are considered.

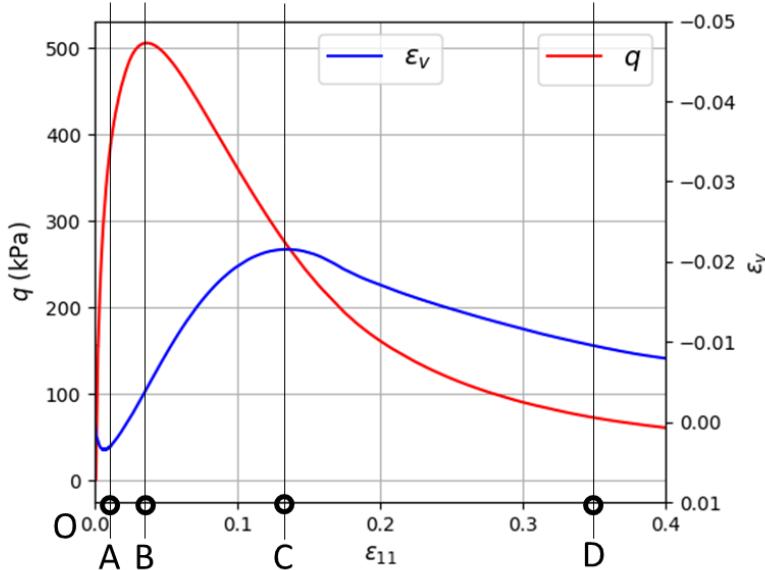


Figure 1.8. The deviatoric stress and volumetric strain along the biaxial loading with $p_0 = 100$ kPa until a relative large axial strain.

1.2 Internship Topic: multi-scale geomechanics: Machine learning from a mesoscale perspective

Through this internship, we delve into the realm of machine learning, specifically leveraging Artificial Neural Networks (ANNs) to simulate the behavior of mesostructures within granular materials. The overarching objective is to develop a new class of multi-scale constitutive models that can accurately capture the intricate behaviors exhibited by granular materials across

different scales. By replacing the traditional 2D H-model framework with machine learning models, we aspire to expedite computational efficiency while gaining valuable insights into the complexities of granular material behaviors.

To reach this objective the main tasks of the internship project are listed below:

1. Generate a training data base for the standard H-cell with use of the analytical description of it;

2. Train a deep neural network to replace analytical relationships for the H-cell and compare the computation efficiency at macroscale on different loading paths;

And offering the following two points as perspectives:

1. Apply the same training procedure for more complex mesostructures solved through discrete element simulations.

2. Explore the feasibility to set up an automatic calibration procedure.

In the following sections, we delve into the methodology, data preprocessing, model development, and results analysis, leading to a comprehensive evaluation of the machine learning approach to simulating the H-model. By doing so, we aim to provide valuable insights and recommendations for future applications of machine learning in granular material simulations.

2 METHODS

In this section, we will present the machine learning models and architectures adopted for our study, accompanied by a display of training outcomes.

2.1 Introduction to Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) represent a fascinating emulation of the human brain's neural structure within the realm of computational modeling. Comprising interconnected nodes or "neurons," ANNs are designed to process and learn patterns from data, making them particularly adept at capturing intricate relationships that might elude traditional methods.

At the core of an ANN lies its architecture, characterized by layers of interconnected neurons. These layers include an input layer to receive data, one or more hidden layers that process information, and an output layer that produces the network's final predictions. The connections between neurons, governed by weights and biases, allow the network to adapt and adjust its internal representations as it learns from data.

Key Components of ANNs:

1.Neurons and Activation Functions:

Neurons in an ANN mimic biological neurons, processing incoming data and generating an output signal. Activation functions determine the neuron's output based on the weighted sum of inputs. Common activation functions include the sigmoid, ReLU (Rectified Linear Unit), and tanh (Hyperbolic Tangent).

2.Weights and Biases:

Weights and biases are parameters that govern the strength of connections between neurons. During training, the network adjusts these parameters to minimize the difference between its predictions and the actual outcomes in the training data.

3.Forward and Backward Propagation:

The process of predicting an output from input is called forward propagation. The calculated output is then compared to the ground truth to compute an error. Backward propagation

involves propagating this error backward through the network to adjust weights and biases, iteratively fine-tuning the network's parameters.

2.2 Motivation for Using ANNs in Granular Material Modeling

The utilization of ANNs within the context of granular material modeling arises from the need to enhance the computational efficiency and to consider more complex mesostructures for which no analytical relationship (such as eq 1.7) exist.. While traditional analytical descriptions provide valuable insights, they can be labor-intensive and impractical for complex mesostructures. ANNs offer an innovative approach by directly learning from data, enabling the prediction of granular material behavior under varying conditions.

2.3 Training Data Generation and Preprocessing

In this section, we establish the form variables of the H-model in both the horizontal and vertical directions. These form variables are represented as $\delta\epsilon_1$ and $\delta\epsilon_2$, defined as $\delta\epsilon_1 = \frac{\delta l_1}{l_{10}}$ and $\delta\epsilon_2 = \frac{\Delta l_2}{l_{20}}$. Here, δl_1 and δl_2 signify the incremental length in the horizontal and vertical orientations of H cells, while l_{10} and l_{20} represent the original lengths in the horizontal and vertical orientations of H cells under zero stress conditions. Consequently, we can equate any stress in the e_1, e_2 direction in Figure 1.1 to stress in the n, t direction.

The horizontal and vertical strain can be equivalently presented in polar coordinates with a magnitude, denoted as ϵ_{tot} , and an angle θ . These transformations are defined as $\delta\epsilon_{\text{tot}} = \sqrt{\delta\epsilon_1^2 + \delta\epsilon_2^2}$, $\delta\epsilon_1 = \delta\epsilon_{\text{tot}} \cos \theta$, and $\delta\epsilon_2 = \delta\epsilon_{\text{tot}} \sin \theta$.

Within our model, the goal is to input information about the H cell and deformation to derive insights about the updated H cell. Specifically, we consider the H cell's attributes, including its length l_1 , width l_2 , inter-cellular angle α , and the forces N_1, N_2, T_1 , as the model inputs. For deformation, we incorporate ϵ_1 and $\delta\epsilon_2$ as additional model inputs. The desired outputs from the model encompass $l'_1, l'_2, \alpha', N'_1, N'_2, T'_1$ to characterize the updated H cell. With $l'_1 = l_1 + \delta\epsilon_1 \cdot l_1$ and $l'_2 = l_2 + \delta\epsilon_2 \cdot l_2$ can be get easily and do not require specific learning. Consequently, our model accepts inputs $\delta\epsilon_1, \delta\epsilon_2, \alpha, l_1, l_2, N_1, N_2, T_1$, and generates outputs $\alpha', N'_1, N'_2, T'_1$.

We've generated two datasets for our model training. In the first dataset, all inputs share the characteristic that we load the H-cell from the initial state $l_1 = l_{10}, l_2 = l_{20}, N_1 = 0, N_2 = 0, T_1 = 0$ in various directions θ between 0 and 90°. The total strain value ϵ_{tot} varies from 1% to 30% (shown in Fig2.1(a)). Dataset 1's H cell information and status have been preserved

and utilized as input data for H cells in dataset 2. In dataset 2, we start from states reached in data set 1 and apply strain increments of magnitude 1%, while the inter-cellular angle θ varies from 0 to 2π . (shown in Fig2.1(b))

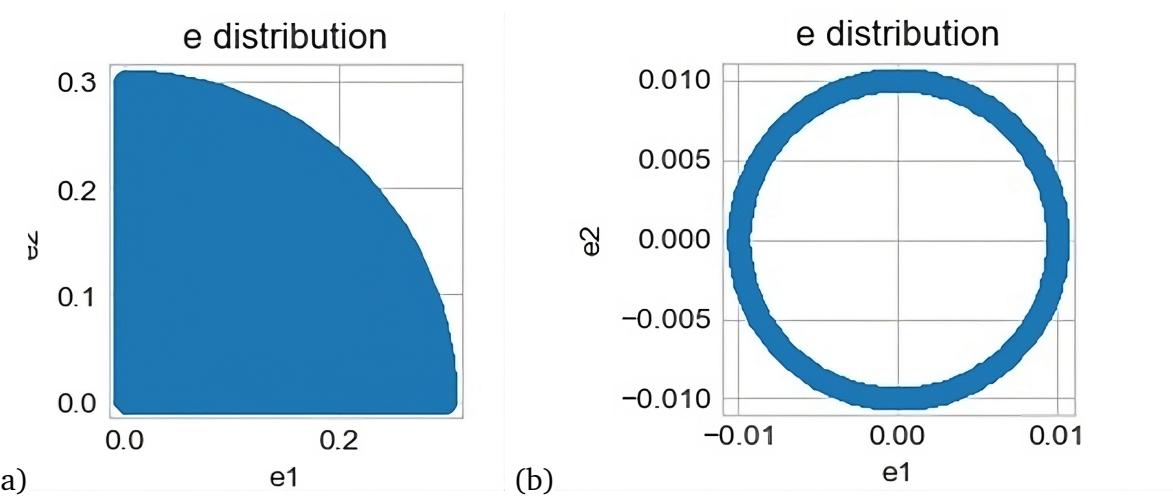


Figure 2.1. Display the distribution of the total strains ($\delta\epsilon_1, \delta\epsilon_2$) variable in dataset 1 and incremental strains ($\delta\epsilon_1, \delta\epsilon_2$) in dataset 2

It is worth noting that, due to the non-linear nature of the deformation simulated by the H model (eq 1.7), it is advisable to keep the deformation length at each step as small as possible during data generation to maintain simulation accuracy. A step size of 0.1% was employed during data generation, resulting in a total of 10 consecutive deformations using the H model for a 1% deformation in the dataset.

After generating all the data, certain outliers that did not conform to sensible characteristics of the H model were removed. For instance in Fig 2.2, cases where the length l_1 was less than $2r$ or where the length l_2 was less than $2r$ after deformation were eliminated.

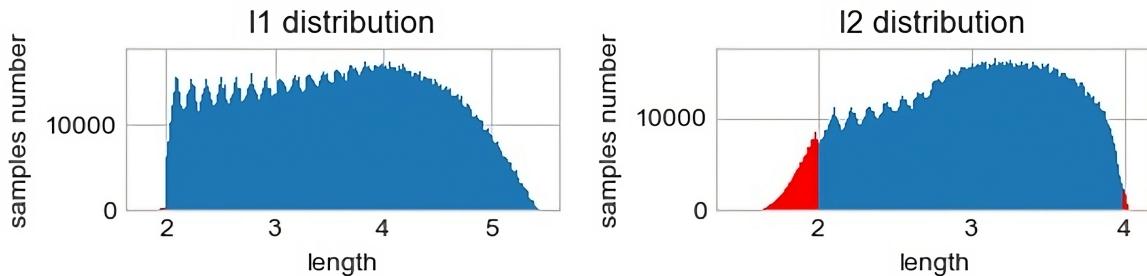


Figure 2.2. The red-colored points represent the outliers that did not conform to the criteria and were subsequently removed.

After the processing was completed, Dataset 1 had a size of $7 * 10^5$, and Dataset 2 had a size of $1.2 * 10^7$. To control the size of the training dataset with an homogeneous sampling of the

training domain, we employed random, grid and LHS (Latin Hypercube Sampling) to extract training and testing subsets from each of them for model training and evaluation purposes. Because input and output data do not always commensurate (we have strain, lengths, angle and forces of different units), normalization is needed. Five different normalization methods were applied to the input data to generate corresponding training datasets. These methods include:

1. **Min-Max Scaling:** Min-Max scaling, also known as feature scaling, rescales data to a specified range, typically between 0 and 1. It is achieved through the following formula for each feature:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

Here, X is the original feature value, X_{\min} is the minimum value of that feature in the dataset, and X_{\max} is the maximum value.

2. **Z-Score Standardization:** Z-score standardization, also known as standard scaling, transforms data to have a mean μ of 0 and a standard deviation σ of 1. It is done using the formula:

$$X' = \frac{X - \mu}{\sigma}$$

3. **Robust Scaling:** Robust scaling is a normalization method that is less sensitive to outliers. It uses the median (M) and the interquartile range (IQR) to scale the data. The formula is as follows:

$$X' = \frac{X - M}{IQR}$$

The median and IQR are more robust statistics compared to the mean and standard deviation, making this method suitable for datasets with outliers.

4. **Unit Vector Scaling:** Unit Vector Scaling, also known as vector normalization, scales data such that each data point becomes a vector with a magnitude of 1. It doesn't change the direction of the data vectors, only their magnitude.

5. **Max Absolute Scaling:** Max Absolute Scaling scales data by dividing each data point by the maximum absolute value in the dataset. It brings the data within the range of -1 to 1.

2.4 Network Architecture and Design

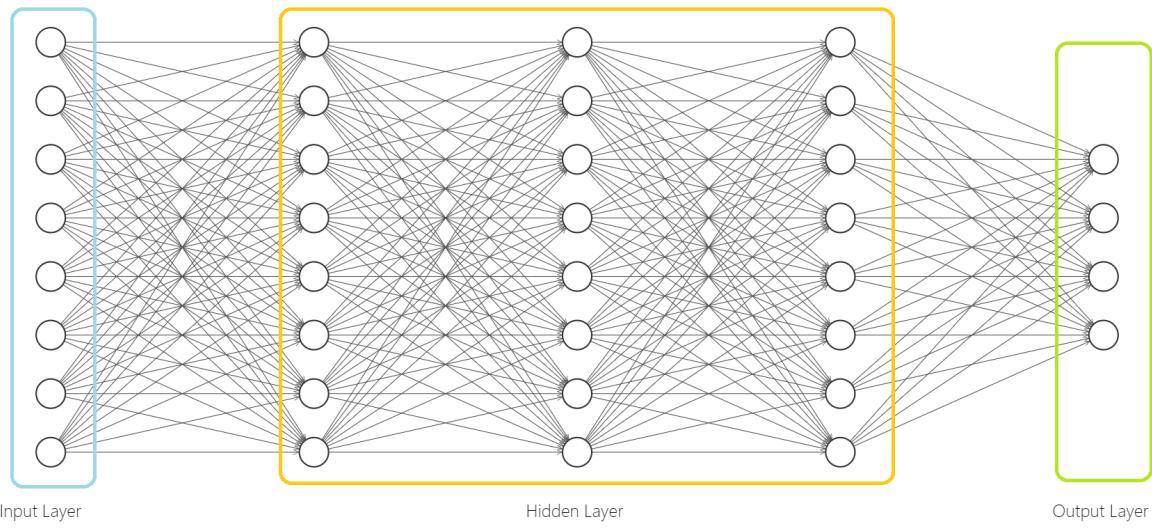


Figure 2.3. Model 1 architecture

In the context of our neural network model architecture, there are input $\delta\epsilon_1, \delta\epsilon_2, \alpha, l_1, l_2, N_1, N_2, T_1$ and output $\alpha', N'_1, N'_2, T'_1$. I have opted for 8 nodes within each hidden layer. Throughout our research, we have consistently employed the hyperbolic tangent (tanh) function as the activation function across all layers. Our investigative efforts encompassed a comprehensive examination of eight distinct model configurations, systematically varying the number of hidden layers.

Table 2.1. 8 models with their characteristics

| Model name | number of nodes per hidden layer | number of hidden layers | loss function |
|------------|----------------------------------|-------------------------|---------------|
| Model 1(a) | 8 | 3 | MAE |
| Model 1(b) | 8 | 3 | MSE |
| Model 2(a) | 8 | 5 | MAE |
| Model 2(b) | 8 | 5 | MSE |
| Model 3(a) | 8 | 7 | MAE |
| Model 3(b) | 8 | 7 | MSE |
| Model 4(a) | 8 | 9 | MAE |
| Model 4(b) | 8 | 9 | MSE |
| Model 5(a) | 8 | 11 | MAE |
| Model 5(b) | 8 | 11 | MSE |
| Model 6(a) | 8 | 13 | MAE |
| Model 6(b) | 8 | 13 | MSE |
| Model 7(a) | 8 | 15 | MAE |
| Model 7(b) | 8 | 15 | MSE |
| Model 8(a) | 8 | 17 | MAE |
| Model 8(b) | 8 | 17 | MSE |

Beginning with Model 1, we established a baseline with three hidden layers, each consisting of 8 nodes. In the case of Model 1(a), we chose the Mean Absolute Error (MAE) as the primary loss function. Subsequently, in Model 1(b), we adhered to the same three-hidden-layer architecture with 8 nodes each, but this time, we adopted the Mean Squared Error (MSE) as the loss function.

We proceeded in this manner, incrementally increasing the number of hidden layers by two for each subsequent model, until we explored the 8th model with 17 layers. This systematic approach enabled us to thoroughly investigate the performance of each neural network model variant in terms of varying model depth and different loss functions.

During model training, 20% of the data is randomly allocated as a validation set. This is done to monitor the training of the model and to prevent potential issues related to overfitting.

2.5 Model Validation and Evaluation

Due to the fact that the data in Dataset 2 better aligns with the application scenario of the H model (continuing deformation from a non-initial state), all the models were trained using a subset of Dataset 2 as the training set. After training the model, it was observed that having

three hidden layers was sufficient for addressing the problem at hand. Furthermore, it can be discerned from the graphical representation that increasing the number of hidden layers from 3(Fig 2.3) to 17(Fig 2.4) did not significantly improve the model's performance.

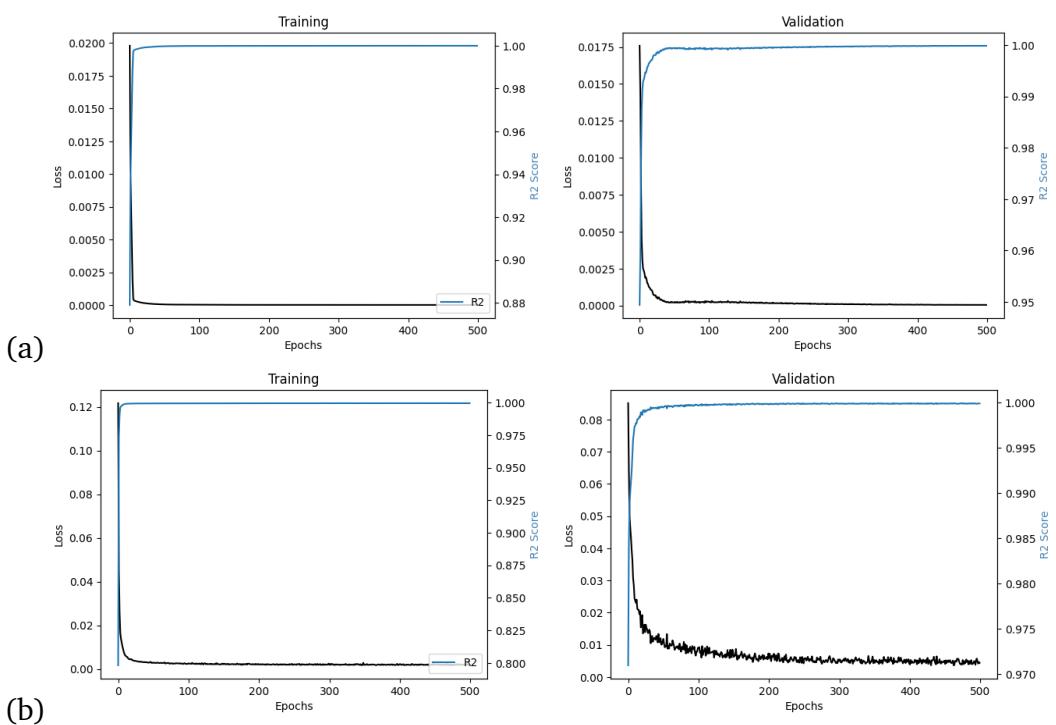


Figure 2.4. Model 1: Performance of 3 Hidden Layers Trained with MAE(a) and MSE(b) Loss Functions with 6k datasets

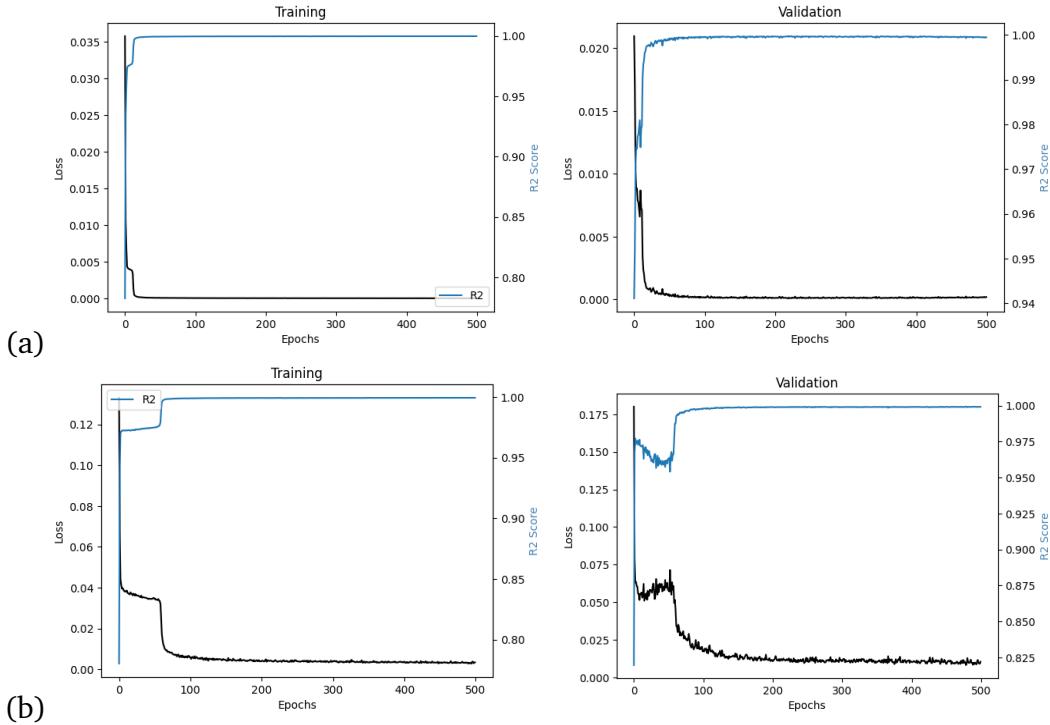


Figure 2.5. Model 8: Performance of 17 Hidden Layers Trained with MAE(a) and MSE(b) Loss Functions with 6k datasets

Similarly, when trained on a larger dataset, the performance of the same model remains consistent. Take the results from Figure 2.3 and Figure 2.5 as examples:

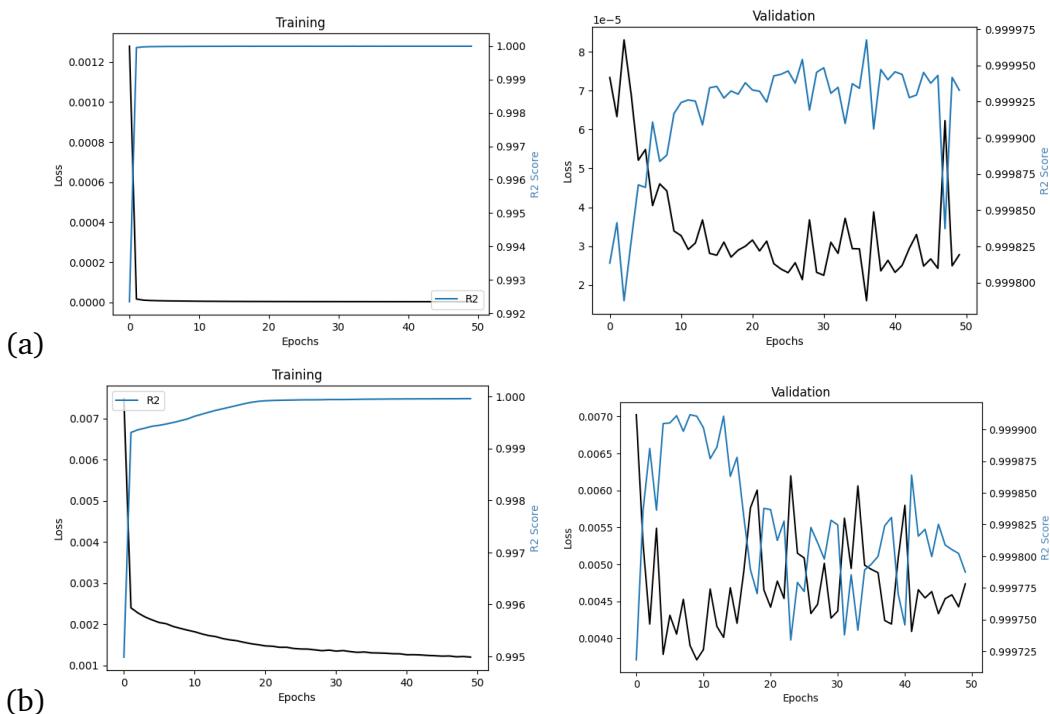


Figure 2.6. Model 1: Performance of 3 Hidden Layers Trained with MAE(a) and MSE(b) Loss Functions with 600k datasets

During the model performance validation phase, we evaluated the models by computing the MAE (Mean Absolute Error), RMSE (Root Mean Square Error), and R-squared (R²) values on both Dataset 1 and Dataset 2. Additionally, we considered the maximum values of output force and angles as simple indicators for assessment.

We observed that despite the models are trained on Dataset 2, they performed well on both datasets (Figure 2.6). Therefore, for subsequent training on the normalized data, we continued to use the training set derived from Dataset 2 and tested it on both model1(a) and model1(b).

```

test on dataset1
force max: [1.76252620e+08 1.18703511e+08 3.46426572e+07] angle max: 1.5815281059447595
true force max: [2.62044649e+08 1.14114918e+08 3.58215381e+07] true angle max: 1.5533418615598973
R2 force: 0.9936103 R2 angle: 0.9978329
RMSE force: 2053983.4027934917 RMSE angle: 0.012878974260744575
MAE force: 1027287.5057290787 MAE angle: 0.008618506962517676
test on dataset2
force max: [2.36022614e+08 1.18439468e+08 3.64819889e+07] angle max: 1.5631415415492529
true force max: [6.20017763e+08 1.17792602e+08 3.72855729e+07] true angle max: 1.5707600376421085
R2 force: 0.9982584 R2 angle: 0.99747574
RMSE force: 1099740.6071886998 RMSE angle: 0.013810898698387737
(a) MAE force: 269311.5010334855 MAE angle: 0.0025588723417527006
test on dataset1
force max: [2.54787792e+08 1.12053471e+08 3.44567386e+07] angle max: 1.5899907085333709
true force max: [2.62044649e+08 1.14114918e+08 3.58215381e+07] true angle max: 1.5533418615598973
R2 force: 0.99906445 R2 angle: 0.99820596
RMSE force: 785949.4258878548 RMSE angle: 0.011717967977173588
MAE force: 547470.3972178464 MAE angle: 0.008653581617502928
test on dataset2
force max: [2.84731289e+08 1.15859959e+08 3.61699697e+07] angle max: 2.054645643857
true force max: [6.20017763e+08 1.17792602e+08 3.72855729e+07] true angle max: 1.5707600376421085
R2 force: 0.9996026 R2 angle: 0.99910367
RMSE force: 525325.7164393435 RMSE angle: 0.008229814778000335
(b) MAE force: 364731.42883928766 MAE angle: 0.004217218558574834

```

Figure 2.7. Test results for Model 1 with MAE(a) and MSE(b) Loss Functions with 60k datasets

After testing all the normalized data, it was observed that the performance of the models trained on min-max scaling, z-score standardization, unit vector scaling, and max absolute scaling was subpar in terms of both MAE and MSE. However, the robust scaling method exhibited poor performance in MSE but performed exceptionally well in terms of MAE. The subpar performance of the underperforming models is primarily manifested in significant losses in angle prediction, with the maximum angle values (around 0.98) falling far below the expected maximum values (around 1.57).

3 RESULTS AND DISCUSSION

3.1 Comparative Analysis with the original analytical H-model

After conducting all the tests, two models were selected for testing in place of the H model: the robust scaling model with MAE loss and model 5 with MAE loss trained on 600k dataset. They are used to replace equation 1.7 in the use of the H-model to create a ANN based H-model. Then this surrogate model was compared with the original analytical H-model at the macroscopic scale (with a collection of H-cells) on standard mechanical tests : isotropic compression tests, drained biaxial tests and undrained biaxial tests. The following are the results of their testing.

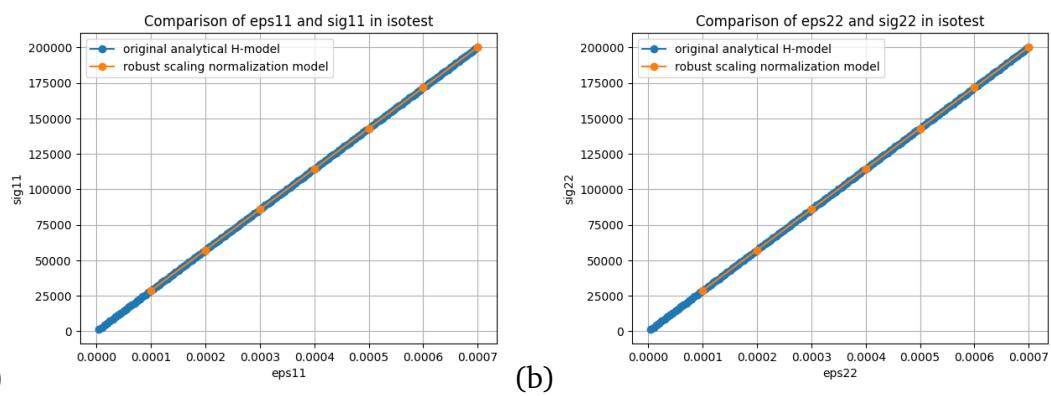


Figure 3.1. Isotest results for Model trained with robust scaling with MAE Loss Functions with 600k datasets

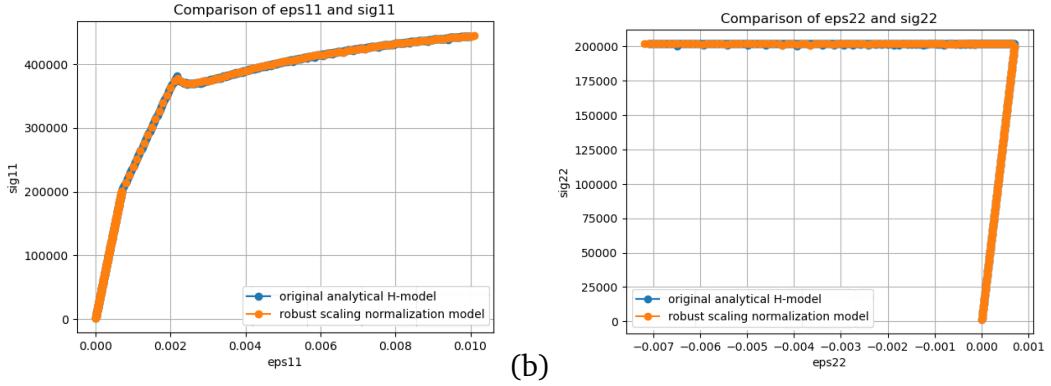


Figure 3.2. Biaxial test results for Model trained with robust scaling with MAE Loss Functions with 600k datasets with lateral confining stress of 200 kPa

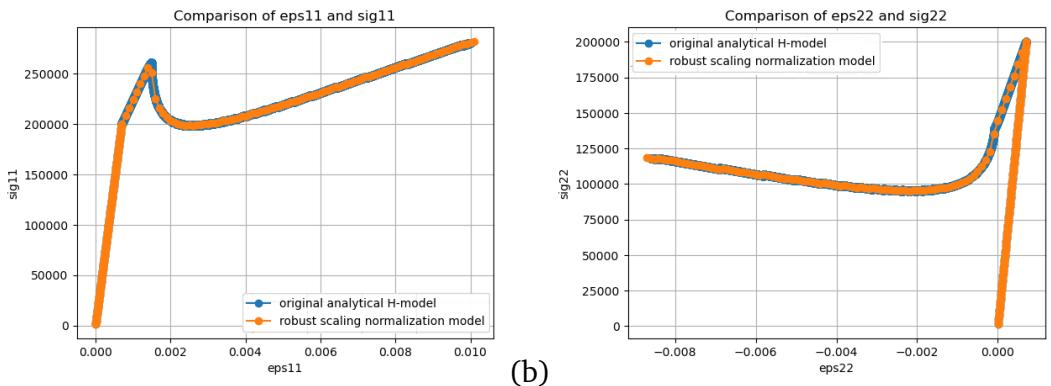


Figure 3.3. Undrained biaxial tests results for Model trained with robust scaling with MAE Loss Functions with 600k datasets

In the above Figure 3.1, Figure 3.2, Figure 3.3, blue dots represent the results of the H model, while yellow dots represent the outcomes of the neural network trained with robust scaling normalization. It is evident that the neural network model exhibits a perfect fit to the curve generated by the H model. And it is noteworthy that other non-normalized models have also exhibited similar results.

3.2 Future Work

In this section, we outline potential avenues for future research and development based on the findings and limitations of our current study. These directions may include:

1. Model Enhancement: Investigating ways to further improve the performance of our machine learning models, such as exploring advanced architectures, feature engineering, or hyperparameter tuning.
2. More complex mesostructures: Expanding our dataset to encompass a broader range of granular materials, such as Adding smaller particles within the original hexagonal structure

of the H model and simulating them using Discrete Element Method (DEM) generated a new dataset for training.

BIBLIOGRAPHY

Veylon, Guillaume (May 2017). “Modélisation numérique du mécanisme de liquéfaction des sols : application aux ouvrages hydrauliques”. Theses. Université Grenoble Alpes. URL: <https://theses.hal.science/tel-01689855>.