Modeling Water Flow and Contaminant Transport in Unsaturated Soil by Artificial Intelligence

Thesis by

Xun Ding

A thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

School of Civil Engineering

University of Sydney

NSW, Australia

https://orcid.org/0000-0001-5496-4700

Originality Statement

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Any contribution by other people, including co-authors of papers, has been clearly stated in the Authorship Contribution Statements.

Candidate Signature:

Statement of AI usage

I acknowledge the use of Grammarly with Overleaf to detect spelling and grammar errors and to improve the language quality when preparing this thesis with LaTeX.

Candidate Signature:

Author Attribution Statement

This thesis contains materials published in the following peer-reviewed articles:

- Ding, X., & El-Zein, A. (2024). Predicting soil water retention curves using machine learning: A study of model architecture and input variables. Engineering Applications of Artificial Intelligence, 133(Part B), 108122. https://doi.org/10.1016/j.engappai.2024.108122
- Ding, X., El-Zein, A. (2023). Water Retention Curves of Clayey Soils by Artificial Neural Networks with Uneven Datasets. In: Barla, M., Di Donna, A., Sterpi, D., Insana, A. (eds) Challenges and Innovations in Geomechanics. IACMAG 2022. Lecture Notes in Civil Engineering, vol 288. Springer, Cham. https://doi.org/10.1007/978-3-031-12851-6_6

This thesis also includes materials from manuscripts that have been submitted for publication and are currently under peer review:

- Ding, X., & El-Zein, A. (Under review). Solving High-Peclet Advection-Diffusion Equations Using Hard-Constraint Physics-Informed Neural Network with Peclet-Adjusted Distance Functions. Engineering with Computers. Submitted on 16th May 2025.

The author contributions of above-mentioned papers are specified according to the CRediT taxonomy as follows:

- Xun Ding (Candidate): Conceptualization; Methodology; Software; Formal Analysis; Validation; Data curation; Writing – original draft
- Abbas El-Zein (Supervisor): Supervision; Conceptualization; Formal Analysis; Writing – review & editing

Candidate Signature: **Supervisor Signature:**

This page was intentionally left blank.

ABSTRACT

Modeling Water Flow and Contaminant Transport in Unsaturated Soil by Artificial Intelligence

Xun Ding

Background. The flow of water in unsaturated soil is typically represented by the Richardson-Richards Equation (RRE), while solute transport is often described by the Advection-Diffusion Equation (ADE). These two equations are fundamental for addressing problems in water and waste resources management but are challenging to solve because of their input-data requirements and inherent non-linearities. The RRE's nonlinearity arises from the strong dependence of material properties (hydraulic conductivity and gradient of water retention curve) on state variables (soil water content and pore pressure), which can cause convergence issues in numerical solutions. Additionally, solving the RRE requires knowledge of the water retention curve (WRC), which is time-consuming, expensive and sometimes difficult to measure. While pedo-transfer functions (PTF) for estimating WRC from more easily measurable soil properties have been developed, uncertainties remain about the most relevant soil properties required for the task. The ADE, for its part, depends on output from the RRE (water content and seepage velocities) and is challenging to solve under advection-dominated scenarios (high Peclet number) in which spurious numerical oscillations may occur.

Recent advances in artificial intelligence (AI) offer potentially innovative solutions to these challenges. The universal approximation capabilities of AI provide a new approach to estimating complex relationships. Furthermore, machine learning methods can incorporate physics-based constraints, enabling models to perform well even with limited data. This new class of machine learning algorithm is called Physics-Informed Neural Networks (PINN) and has been shown to

be effective in solving complex partial differential equations (PDE). Very recently, new hard-constraint PINN approaches based on strict enforcement of boundary conditions, have been proposed but are yet to be explored in the context of water flow and contaminant migration.

Goal and Objectives. The goal of this thesis is to advance the modeling of water flow and solute transport in unsaturated soils through the application of the above AI techniques. Four objectives are pursued to achieve this goal: a) investigate the ability of shallow and deep ANNs to generate accurate estimates of WRC, b) develop a new hard-constraint PINN approach to the solution of the time-dependent RRE, c) develop and explore a new hard-constraint PINN based on novel distance functions to solve the steady-state and time-dependent ADE under low and high Peclet numbers and d) assess the ability of PINN to solve the coupled, time-dependent RRE-ADE problem. All of these developments are conducted in one dimension in space.

Methods. Two major sets of developments were conducted to achieve the above objectives: a) neural networks for predicting WRC and b) PINN solvers for ADE and RRE. ANN predictors of WRC were trained and tested using the Unsaturated Soil Database 2.0 (UNSODA 2.0), including 790 samples collected worldwide. Thirteen different models were constructed with varying combinations of input parameters. Each of the thirteen models was developed in two different versions of neural network structure: SNN and DNN. In addition, a clay-specific dataset was extracted from the National Cooperative Soil Survey (NCSS) Soil Characterisation Database and used to explore whether improvements in prediction accuracy for clayey soils could be achieved. For PINN solvers, soft-constraint solvers of ADE and RRE were first developed and then modified to introduce hard constraints for boundary and initial conditions. A set of distance functions – linear, exponential and logarithmic – were investigated to explore their effects on the accuracy of the hard-constraint approach. Performances of the PINN models were validated by comparing their predictions to analytical and numerical so-

lutions. The RRE-ADE solvers were then coupled and validated. Finally, the combined usage of the WRC-predictor and ADE-RRE solver was illustrated by application to a problem of contaminant transport in unsaturated soil. All developments were made in Python and associated libraries, and all ADE and RRE equations were in one dimension in space.

Key Findings. Usage of DNN did not always outperform SNN in predicting WRC and the latter sometimes provided higher levels of accuracy. Soil texture, augmented by either dry density or soil porosity, was found to be essential for high accuracy in WRC prediction, whereas the contribution of specific gravity, soil depth, and organic content appeared minimal. The two best-performing models, out of 13 models developed and tested, compared favourably with WRC AI models reported in the literature. For clayey soils, training an ANN on a clay-soil-specific dataset, improved prediction accuracy by 20%.

A hard-constraint PINN algorithm was found to be able to generate accurate solutions of both the RRE and ADE. In particular, in developing a hard-constraint PINN solver for the ADE, it was found that using an adjustable exponential distance function achieved remarkably high levels of accuracy even at very high Peclet numbers (up to 1 million and 1000 under steady-state and transient conditions, respectively). Finally, a hard-constraint PINN algorithm was found to be able to solve the coupled RRE-ADE problem and, when used in conjunction with the WRC-ANN predictor developed earlier, offered an AI suite of solvers for simulating the coupled time-dependent water flow and contaminant migration problem in soils.

Hence, this thesis indicates a strong potential for deploying AI tools to model time-dependent problems of unsaturated flow and contaminant transport in geoenvironmental engineering. Extension of the developments reported here to two and three spatial dimensions and comparison of computational economies of AI to more conventional solvers would consolidate these findings and allow researchers to identify conditions under which AI solvers are advantageous in this field.

This page was intentionally left blank.

ACKNOWLEDGEMENTS

I would like to begin by expressing my deepest gratitude to my supervisor, Professor Abbas El-Zein, for his continuous support, insightful guidance, and constant encouragement throughout my doctoral candidature. His patience and generous mentorship have not only guided my research but also deeply influenced how I approach problems and challenges. It is my great honour to be under his supervision.

I would also like to thank all the members of the Geoenvironmental Research Group for sharing ideas, engaging in valuable discussions, and making this doctoral journey both enjoyable and meaningful. I am also thankful to all my friends in Office 101 and 360 for making this journey more enjoyable.

Finally, I want to show my deepest appreciation to my parents. Thank you for always respecting my choices and encouraging me to pursue my dreams. You gave me love, education, and the foundation of who I am.

This page was intentionally left blank.

List of Publications

The following publications have been produced during this PhD.

Journal Articles

1. Ding, X., & El-Zein, A. (2024). Predicting soil water retention curves using machine learning: A study of model architecture and input variables. *Engineering Applications of Artificial Intelligence*, **133**(Part B), 108122. https://doi.org/10.1016/j.engappai.2024.108122

Conference Papers

1. Ding, X., & El-Zein, A. (2023). Water retention curves of clayey soils by artificial neural networks with uneven datasets. In Barla, M., Di Donna, A., Sterpi, D., & Insana, A. (Eds.), *Challenges and Innovations in Geomechanics. IACMAG 2022. Lecture Notes in Civil Engineering* (Vol. 288). Springer, Cham. https://doi.org/10.1007/978-3-031-12851-6_6

Manuscripts Under Review

1. Ding, X., & El-Zein, A. (2025). Solving High-Peclet Advection-Diffusion Equations Using Hard-Constraint Physics-Informed Neural Network with Peclet-Adjusted Distance Functions. *Engineering with Computers*, Under Review.

Manuscripts Under Preparation

1. Ding, X., & El-Zein, A. (2025). Water Flow in Unsaturated Soil: A Physics-informed Neural Network with Hard Constraints and Adaptive Resampling.

This page was intentionally left blank.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page was intentionally left blank.

LIST OF SYMBOLS

Symbol	Description
a	Left boundary coordinate [L]
a_i	i_{th} actual value
b	Right boundary coordinate [L]
b_i	Biases matrix of the i^{th} hidden layer
b_j	Bias associated with the j^{th} neuron
D	Dispersion coefficient [L^2 T^{-1}]
$D(\theta)$	Soil water diffusivity [L^2 T^{-1}]
D_{eff}	The effective dispersion coefficient
D_L	Longitudinal dispersivity coefficient
D_m	Molecular diffusion coefficient
$g(x,t)$	Boundary condition values
h	Pressure head (matric potential) [L]
h^*	Reduced pressure head
$h(x,0)$	Initial condition values
H	Total hydraulic head [L]
i_a	Dirichlet BC switch at $x = a$
i_b	Dirichlet BC switch at $x = b$
k_1	A resampling scaling factor for RAR
k_2	A resampling scaling factor for RAD
$K(\theta)$	Unsaturated hydraulic conductivity [L T^{-1}]
K_s	Saturated hydraulic conductivity [L T^{-1}]
K^*	Reduced hydraulic conductivity
$l(t)$	Distance function for initial condition

$l^{\partial\Omega}(x)$	Distance function for boundary condition
L_d	The domain length, which is $L_d = b - a$
L_{total}	Total loss function
L_B	Loss term for boundary condition
L_I	Loss term for initial condition
L_R	Loss term for PDE residual
m_{fx}	Fitting parameter (Fredlund-Xing model)
m_{vg}	Fitting Parameter (van Genuchten model)
n	number of inputs for this neuron
n_a	Total number of data points to be analysed
n_{fx}	Fitting parameter (Fredlund-Xing model)
n_{vg}	Fitting Parameter (van Genuchten model)
nl	Total number of hidden layers
N_c	Number of collocation points
$N_x[u]$	The spatial differential operator
p_i	i_{th} predicted value
PWAE_i	Point-wise absolute error [L]
PWAPE_i	Point-wise absolute percentage error
q	Water flux [L T^{-1}]
R^2	Coefficient of determination
S	Effective saturation
t	Time [T]
t_{lim}	Upper bound of time [T]
t_{min}	Minimum time [T]
t_{max}	Maximum time [T]
T	Reduced time
u	Contaminant concentration [M.L^{-3}]
$u^{\partial t}(x, 0)$	Initial condition function
$u^{\partial\Omega}(x, t)$	Boundary condition function

v	Seegage velocity [L T^{-1}]
$V1_i$	A given input variable
$\bar{V1}$	The mean of $V1_i$ in the dataset
$V2_i$	A given input variable
$\bar{V2}$	The mean of $V2_i$ in the dataset
w_B	Weight for boundary condition loss
w_I	Weight for initial condition loss
w_R	Weight for PDE residual loss
W	Reduced volumetric water content
W_{ij}	Weight connecting the i^{th} input to the j^{th} neuron
$\mathbf{W_i}$	Weights matrix of the i^{th} hidden layer
x	Spatial coordinate [L]
x_{pred}	The depth predicted [L]
x_{true}	The target value [L]
X	Reduced spatial coordinate
α_{bc}	Fitting parameter (Brooks-Corey model) [L^{-1}]
α_{fx}	Fitting parameter (Fredlund-Xing model) [L^{-1}]
α_t	Distance function parameter for time
α_{vg}	Fitting Parameter (van Genuchten model) [L^{-1}]
α_x	Distance function parameter for space
β_{fx}	Fitting parameter (Fredlund-Xing model)
$\chi(W)$	Coefficient in generalised solution
δ	Parameter that controls the boundary type
$\varepsilon(x,t)$	The point-wise PDE residual
η	The parameter collection of the NN
$\hat{\eta}$	The optimised parameter collection of the NN
γ	Parameter that controls the boundary type
$\lambda(W)$	Coefficient in generalised solution
λ_{bc}	Fitting parameter (Brooks-Corey model)

λ_{fx}	Fitting parameter (Fredlund-Xing model)
μ	Scaling factor for reduced coordinates $[L^{-1}]$
$\psi(\theta)$	Suction head function [L]
$\psi(W)$	Coefficient in generalised solution
σ	The activation function
θ	Volumetric water content $[L^3\ L^{-3}]$
$\theta(h)$	Volumetric water content to h $[L^3\ L^{-3}]$
θ_r	Residual water content $[L^3\ L^{-3}]$
θ_s	Saturated water content $[L^3\ L^{-3}]$
$\tau(\theta)$	The tortuosity factor
Ω	A spatial domain
$\partial\Omega$	The boundary of Ω

LIST OF ABBREVIATIONS

Abbreviation	Description
ADE	Advection-Diffusion Equation
AE	Absolute Error
AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
BC	Boundary Condition
BEM	Boundary Element Methods
BNN	Biological Neural Network
CFD	Computational Fluid Dynamics
DD	Dry density
DIS	Displacement
DS	Depth of sample
DT	Decision Tree
DNN	Deep Neural Networks
FDM	Finite Difference Method
FEM	Finite Element Method
FBPINN	Finite Basis Physics-Informed Neural Network
FVM	Finite Volume
HDC	Hydrodynamic Dispersion Coefficient
HC	Hard-Constraint
HFC	Hydraulic Conductivity Function
HPD-DNN	Hybrid physical data-informed DNN model
IC	Initial Condition

kNN	k-Nearest Neighbors
Leaky ReLU	Leaky Rectified Linear Unit
LHS	Latin hypercube sampling
MAE	Mean Absolute Error
MSE	Mean Squared Error
NCSS	National Cooperative Soil Survey
NN	Neural Network
OM	Organic matter content
PDE	Partial Differential Equation
PINN	Physics-Informed Neural Network
PO	Porosity
PTF	Pedotransfer Function
PWAE	Point-Wise Absolute Error
PWAPE	Point-Wise Absolute Percentage Error
RAD	Residual-based adaptive distribution
RAR	Residual-based adaptive refinement
RRE	Richardson–Richards Equation
ReLU	Rectified Linear Unit
RMSE	Root Mean Squared Error
RF	Random Forest
SHAP	Shapley Additive explanations
SG	Specific gravity
SNN	Shallow Neural Networks
SVM	Support Vector Machine
SWCC	Soil Water Characteristic Curve
UNSODA	Unsaturated Soil Database
VG	Van-Genuchten
WBC	Weakly Imposed Boundary Conditions
WRC	Water Retention Curve

Chapter 1

Introduction

1.1 Background and Overview

Understanding water movement and solute transport in unsaturated soils is of utmost importance for water resources, waste management and protection of groundwater, geotechnical and structural design. Contaminant transport is commonly simulated by the advection-diffusion equation (ADE), while water movement in unsaturated soil is described by the Richardson–Richards Equation (RRE) [3]. However, analytical solutions for these equations are only available for a limited number of simple scenarios with very specific geometries, boundary conditions and usually homogeneous material properties [4, 5]. Consequently, numerical methods such as the Finite Difference Method (FDM), the Finite Element Method (FEM), and Finite Volume Method (FVM) have conventionally been employed to solve these PDEs.

Nonetheless, certain obstacles persist in effectively and accurately modeling water movement and solute transport in unsaturated soils. Firstly, experimental determination of the water retention curve (WRC), a critical soil property, is time-consuming, technically complex and expensive [6]. This is especially the case for soils with significant swelling clay content, with each point on the curve requiring up to two weeks, particularly in the low- and mid-suction ranges [7]. Secondly, the accuracy and computational efficiency of FDM and FEM solvers decline when used to solve highly nonlinear problems such as the ADE and RRE. For example, when advection dominates transport in the ADE, numerical dispersion is known to distort the FEM solutions and yield erroneous results [8, 9]. The

traditional FEM approach also struggles to effectively solve the RRE because of the dependence of water content and hydraulic conductivity on pore pressure and the highly nonlinear and hysteretic nature of this relationship [4]. Hence, new approaches to solving these equations remain an active research pursuit.

In recent years, remarkable advances have been achieved in artificial intelligence (AI) and deep learning, driven by the rapid evolution of computing power and data science. These advances have led to significant breakthroughs in various scientific disciplines, including computer visualisation, data analysis, classification and regression tasks. Machine learning techniques are especially powerful in finding hidden patterns and relationships within data due to their universal approximation capabilities. Recent studies have applied AI techniques to develop pedotransfer functions (PTFs) that estimate difficult-to-measure soil properties from more readily available ones. One application is the use of neural networks to predict the WRC. However, several important research gaps remain unsolved. First, model performance is highly sensitive to AI model's architecture. Most existing work relies on shallow neural networks, with limited exploration of deeper models. Second, the influence of individual input variables on model accuracy is not well understood or inconclusive. Finally, current models tend to perform poorly for certain soil textures, especially clay-rich soils, and strategies to improve predictions for these cases are lacking. This research aims to address these gaps.

Until recently, machine learning was largely limited to problems in which large datasets exist, was entirely data-driven and could not incorporate fundamental physical laws that govern relationships between input and output variables. In 2017, a new category of machine learning techniques, physics-informed neural network (PINN), emerged to address the above limitations [10, 2, 11]. This approach addresses the challenge of training with limited data by incorporating available physics knowledge as additional training resources when building machine learning models. It also offers insights into solving challenging PDEs, particularly in scenarios involving significant nonlinearities or shocks where standard numerical

approaches often face difficulties [10]. Hence, there is potential for PINN to contribute to the solution of ADE and RRE equations and help address the issues faced by conventional numerical techniques discussed above. However, several challenges remain in the current PINN framework. In advection-dominated scenarios with high Peclet numbers, standard PINNs often experience convergence issues due to the competition between multiple loss terms. The hard-constraint approach, which enforces boundary and initial conditions exactly, has the potential to address this by eliminating such competitions. However, its ability to solve advection-dominated ADEs remains untested. In addition, hard-constraint PINNs typically rely on simple linear distance functions to apply constraints, and there has been no systematic investigation into whether alternative forms, such as exponential or logarithmic, could improve model performance. For the Richards equation, the use of PINNs is still in its early stages, with only five known studies attempting to solve it [12, 13, 14, 15, 16]. Among these, only one presented a forward solution, and the results reported were highly sensitive to model configurations and unstable during training. This study will explore whether the hard-constraint formulation can improve both the stability and accuracy of RRE predictions. Furthermore, an AI-inspired model will be developed to couple the RRE and ADE equations for modeling water and solute transport in unsaturated soil.

In summary, three problems in modeling water flow and contaminant transport in unsaturated soils remain, in particular are addressed in this thesis through machine learning: WRC determination, advection-dominated solute transport, and accurate solutions of the RRE.

1.2 Research Aims

This project aims to advance the modeling of water flow and contaminant transport in unsaturated soils using AI techniques. The study focuses on addressing key challenges and gaps in current approaches. The specific aims are:

1. Investigate the potential of deep neural networks to enhance the accuracy of PTFs for estimating WRC by exploring various combinations of input parameters.

2. Explore the use of ANN with clay-specific datasets to improve the predictive accuracy of PTFs for WRC in clayey soils.

3. Evaluate the effectiveness of hard-constraint approaches in improving the ability of PINN to solve the ADE under high Peclet number conditions and assess different distance function formulations.

4. Assess the capability of hard-constraint approaches to enhance the stability and accuracy of PINN solvers for RRE and identify strategies for further improving solver performance.

5. Develop an integrated ANN-PINN-based solver for coupled water flow and solute transport in saturated and unsaturated soils.

1.3 Research Questions

This research aims to explore the application of ANN and PINN in modeling water flow and contaminant transport in unsaturated soils. The study is structured around the following key research questions:

1. **Can ANN improve predictions for WRC?**

   (a) Can deep ANN improve the predictive accuracy of PTFs for WRC?

   (b) Which input variables are most critical for achieving high predictive accuracy in ANN models?

   (c) Can datasets enriched with clay-specific data enhance the accuracy of WRC predictions for clay-dominated soils?

2. **Can PINNs provide better solutions for the ADE?**

(a) Can a hard-constraint approach improve the ability of PINN to solve the ADE for high Peclet numbers?

(b) What is the most effective formulation for a distance function when applying a hard-constraint approach in PINNs?

3. **Can PINNs enhance the modeling of RRE?**

   (a) Can a hard-constraint approach improve the stability and accuracy of PINN solvers for the RRE?

   (b) How can the performance of PINN-based RRE solvers be further improved?

4. **Can a robust and accurate ANN-PINN-based solver be developed for the coupled ADE-RRE equations?**

   (a) How can ADE and RRE solvers be effectively coupled to simulate the complete coupled system?

   (b) Can such a hybrid solver effectively integrate contaminant transport and water flow dynamics?

1.4 Thesis Structure

The remainder of this thesis is composed of nine chapters:

Chapter 2 reviews current knowledge and techniques relevant to this research. It examines advancements in AI and its application in predicting WRC. The chapter also explores the development and application of physics-informed machine learning, with a focus on its use in solving advection-dominated ADE and modeling the RRE. Additionally, studies on coupling water movement and solute transport in unsaturated soils are reviewed.

Chapter 3 outlines the methodologies employed in this thesis. It provides a detailed description of the database utilised, including its structure and relevance to the study. The chapter discusses theories of solute and water movement,

along with their governing equations: ADE and RRE. Furthermore, it discusses the different methodologies and algorithms, either used or developed, and the evaluation metrics used to assess model performance and accuracy.

Chapters 4 to 8 presents the results of various developments detailed in Chapter 3. Chapter 4 presents the results of using machine learning to predict WRC. This chapter examines whether transitioning from shallow neural networks (SNN) to deep neural networks (DNN) improves WRC prediction accuracy. It also investigates the most effective input parameters for enhancing prediction performance.

Chapter 5 focuses on the development of a point-based PTF model specifically for clayey soils, addressing the challenge of lower prediction accuracy for this soil texture compared to others. The model is designed to predict gravimetric water content at predefined matric potentials and to estimate the parameters of the Van Genuchten model, providing a more accurate and reliable solution for clayey soils.

Chapter 6 explores the use of physics-informed machine learning to solve the ADE, particularly under advection-dominated conditions. The chapter addresses optimisation challenges arising from conflicts among loss terms. A new approach is introduced, incorporating a novel exponential distance function formulation within the hard constraint framework, along with a recommended shape factor. This approach successfully resolves the ADE across a wide range of Peclet numbers (P_e). The results highlight the model's effectiveness in solving both steady-state and transient problems.

Chapter 7 discusses modeling the RRE using PINNs. Traditional numerical approaches face significant challenges due to the nonlinearity and dependency on complex parameters. This chapter details the development of a robust model that effectively converges to accurate solutions using a hard constraint method. This chapter also tested two resampling strategies to find out if further improvement can be achieved.

Chapter 8 illustrates the coupling of the RRE and the ADE equations to

simulate water and solute transport in unsaturated soils. The chapter evaluates the performance of the coupled machine learning PDE model in solving a reference problem.

Chapter 9 concludes the thesis by summarising findings of this research, discussing its limitations and highlighting future research directions, including suggestions for advancing machine learning applications in modeling unsaturated soil systems.

This page was intentionally left blank.

Chapter 2

Literature Review

2.1 Data-Driven Machine Learning

Data is a collection of recorded facts and observations. It allows humans to document knowledge, understand the world, and make better decisions [17]. In the fields of scientific research and professional practice, data is analysed to uncover the underlying relationships between variables, understanding complex patterns, and supporting prediction and decision-making. This process is referred to in the literature as statistical learning [18, 19]. Building on these principles, data-driven machine learning has emerged over the past decade as a powerful technique to learn and elicit hidden patterns in data, make predictions and/or aid decisions based on learned knowledge. Commonly used algorithms in data-driven machine learning, include Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF) and Artificial Neural Network (ANN). Each algorithm has its own strengths and limitations, and their areas of application vary according to their characteristics and modelling capabilities. These commonly used AI algorithms are described and justified in the following paragraphs.

kNN is a classical AI algorithm that has been widely used for classification tasks. The name "kNN" was formally introduced and named by Cover and Hart in 1967 [20]. Its theoretical foundation is the nonparametric discrimination method proposed by Fix and Hodges in 1951 [21]. The prediction for a new sample is determined by label aggregation of its k nearest neighbours in the dataset. The similarity between the new sample and every single record in the dataset is quantified using a distance metric, which measures how far apart the new sample is from each

existing record within the feature space. The simplicity and ease of implementation of kNN make it suitable for simple and low-dimensional classification and regression tasks [22]. However, the drawbacks of kNN are also straightforward. Because the algorithm must compute the distance between the new sample and all existing records, its computational cost increases linearly with the dataset size, making kNN less suitable for applications with large datasets [23]. Additionally, the data similarity evaluation of kNN relies purely on distance metrics in feature space. Therefore, irrelevant features can impact the model's performance, and the distance metrics become less meaningful in high-dimensional spaces [23, 24].

SVM is another popular AI algorithm introduced by Cortes and Vapnik in 1995 [25]. This algorithm learns the relationships and differences among samples to identify an optimal boundary that could effectively separate data from different groups. Once trained, the model can classify new samples into their corresponding groups. Since its introduction, SVM algorithm has been widely applied to classification tasks such as face recognition [26, 27] and text categorisation [28, 29, 30]. Despite the successful applications of SVM in various types of tasks, a common limitation of this algorithm is its computational complexity. The training cost of the SVM model is typically between $O(N^2)$ to $O(N^3)$ (for a definition of big O notation, the reader is referred to [31]) [32]. Therefore, the training cost increases exponentially as the size of the training dataset grows. The model can become computationally prohibitive when the training dataset is too large. So, SVM is more suitable for tasks with small to medium-sized training datasets [33, 34].

Another widely used AI algorithm is DT, whose development can be traced back to the mid-1980s [35, 36]. A DT model has a hierarchical structure (Figure 2.1) starting from a root node and branching through several levels of internal nodes until reaching the leaf nodes. From root to leaf, each hierarchical level selects one feature and uses a splitting threshold to divide the data into groups. This splitting process continues level by level until all available features

have been used or the samples within each leaf node already belong to the same category. As described above, the structure of DT is very straightforward and can be easily applied in practice. Its rule-based modeling strategy also made it highly interpretable [37, 38]. However, DT is prone to overfitting, especially when the tree grows too deep, leading to a reduced generalisation capability of the model[39, 40, 41]. This limitation occurs because in each DT level, the model makes splitting decisions based on local information at each node. This strategy makes the algorithm lose the global view of the data, which can let the tree grow to a level of detail that captures noise in the training data rather than meaningful patterns. To address the overfitting issues of DT algorithm, Breiman introduced the concept of randomness into the decision tree framework and proposed the Random Forest algorithm in 2001 [42]. In RF, a large number of modified decision trees are constructed and combined to form an ensemble model. Two types of randomness are introduced to each single DT during the RF development. The first type is data randomness, where each tree is trained on a slightly different subset of the original dataset by randomly excluding a small portion of the records. The second type is feature randomness, where each tree considers only a fraction of randomly selected features from the feature field rather than using all of them. These two types of randomness create variations among DTs. By aggregating their predictions through a majority voting process, RF could significantly increase its generalisation and reduce the risk of overfitting [42, 43]. RF has shown excellent performance in both classification and regression tasks across various science disciplines [44, 45, 46, 43, 47, 48]. However, RF has a significant limitation that is widely recognised by researchers when this tree-based algorithm is applied in regression tasks [49, 50, 51], it produces intrinsically discontinuous predictions. When visualised, the resulting prediction has a serrated appearance (Figure 2.2 provides an example). This behaviour arises from the piecewise constant function (i.e., output has the same value in the same leaf region) of tree-based algorithms, which makes the prediction lack smoothness.

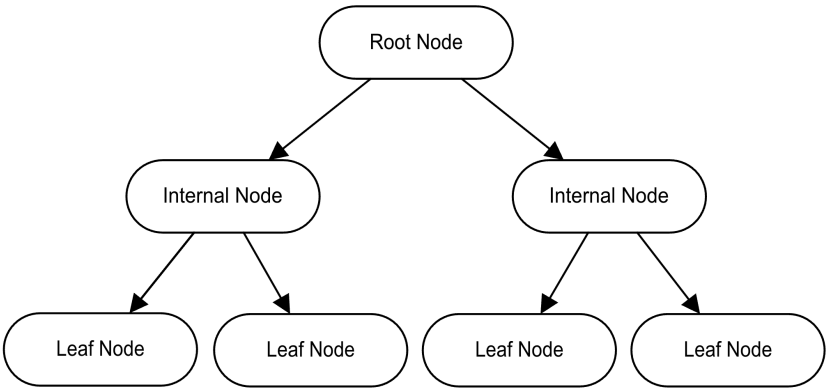


Figure 2.1: Structure of a decision tree model

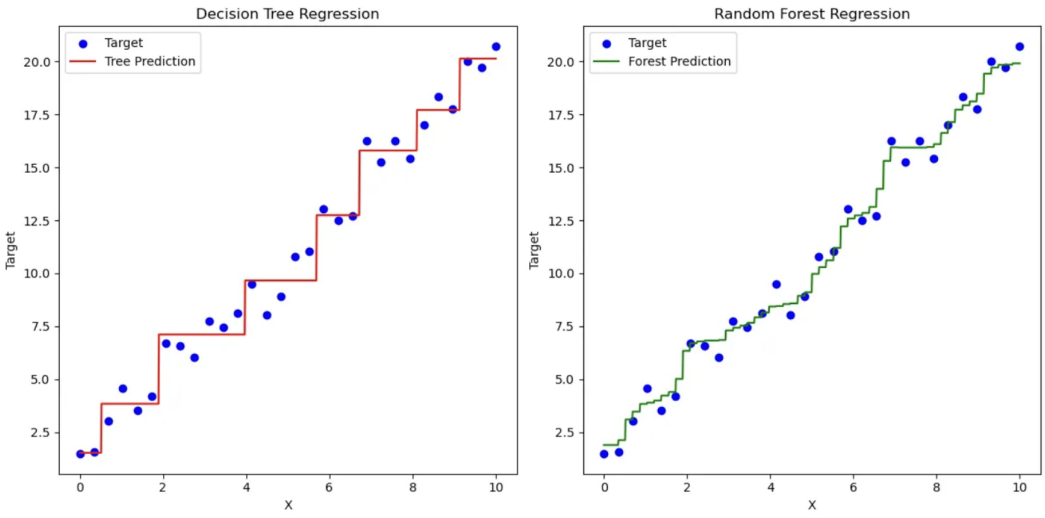


Figure 2.2: Sample regression results by DT and RF [1].

The last AI algorithm to be reviewed in this section is ANN, this algorithm has been found to be successful in classification and regression tasks in a wide range of disciplines and sectors, especially when working with large data sets. It is particularly powerful in capturing intricate patterns and relationships in data, especially high-dimensional data (i.e., where a large number of input parameters is present), and data exhibiting high levels of nonlinearity. For example, ANN has been extensively applied in the fields of computer vision, speech recognition and processing, medical imaging, and aerospace [52, 53, 54, 55, 56, 57, 58]. ANN has also been successfully applied in the geotechnical field, especially in problems in which high levels of heterogeneity, nonlinearity and uncertainty in soil behaviour are present, such as ground excavation [59, 60], estimation of soil

hydraulic properties [61, 62, 63], earth retaining structures [64, 65], site characterisation [66, 67, 68, 69, 70, 71], pile-bearing capacity [72, 73], settlement of structures [74], liquefaction [75, 76, 77], slope stability [78, 79, 80], mapping landslide susceptibility [81, 82], and soil classification [83, 84, 85].

While ANNs have demonstrated remarkable success across various fields, their effectiveness largely depends on how they are structured and trained, particularly in the design of hidden layers and selection of activation functions. In terms of the structure of hidden layers, the number of neurons selected for each layer significantly influences the accuracy of the ANN model [86]. Unlike input and output neurons, which are equal to the number of input and output variables, respectively, the number of hidden layers and the number of hidden layer neurons must be chosen depending on the problem's environment.

The structure of hidden layers in neural networks is also greatly influenced by the choice of activation function. These functions are used to introduce non-linearity between outputs of consecutive layers, a critical element that enables the network to learn and accurately represent data patterns. Different activation functions bring distinct properties and condition the learning process differently. Their selection can significantly influence various aspects of neural network training, including the rate at which the model converges, and its overall performance when tackling tasks [87]. Thus, choosing the right activation function is a key consideration in structuring the neural network. This is a complex and time-consuming step because there is no rule of thumb to identify an optimum hidden layer's structure.

ANN models can be classified as SNN and DNN, depending on whether they contain one or more hidden layers, reflecting increased model complexity [88]. While adding a hidden layer might enable the model to capture more complex data patterns, it can also cause the training process to be slower. Thomas et al. [89] analysed the effects of additional hidden layers for 10 datasets from different scientific disciplines. They found that, while shifting from SNN to DNN improved

accuracy for some datasets, no improvements were detected in others, and, in one case, a 3% decline was found. This was consistent with the findings reported in earlier work by Hayashi, Sakata, and Gallant [90].

As discussed earlier in the introduction, this thesis uses the predictive capacity of AI for modeling water flow and solute transport. So, the selection of a suitable AI algorithm is key to the thesis. In this thesis, the selected AI algorithm serves two primary purposes (1) WRC prediction and (2) as the function approximator in a PINN framework for modeling water flow and solute transport in unsaturated soils. To achieve accurate WRC estimation, the algorithm must demonstrate a powerful expressive capability to capture complex and highly nonlinear relationships within the soil hydraulic dataset, and produce smooth, continuous regression predictions. For an effective integration of PINN modeling framework, the selected algorithm should not only have strong function approximation capabilities but also be fully differentiable throughout the computational domain. Based on the review of AI algorithms above, ANN is selected in this thesis because it best satisfies the required criteria.

2.2 Classical Numerical Methods for Solving PDEs

At the macroscopic scale, PDE is a parametric model that provides a quantitative description of how physical quantities evolve in space and time [91]. Accurate solutions of PDEs are the basis of the meaningful interpretation of the natural system under the given initial and boundary conditions. However, analytical solutions are typically limited to simple PDEs [4, 5]. For problems with high-order equations, complex boundary and initial conditions, high levels of nonlinearity, or solutions with sharp fronts, such closed-form analytical solutions are not usually available [92, 93, 94].

In the history of numerical modeling, the first widely recognised attempt to use numerical methods for solving PDEs is often attributed to von Neumann's research on Lagrangian hydrodynamic schemes in the mid-1940s [91, 95, 96]. Since

then, a family of classical numerical methods built on the space and time discretisation principle has been developed over the past 8 decades, including FDM, FEM, and FVM [97, 98, 99, 100]. These techniques allow researchers to solve PDEs and address a wide range of tasks, including spatial prediction, temporal prediction, data assimilation between observations and deterministic/stochastic model outputs, and model parameter inference [101].

Classical numerical methods also led to the development of a number of numerical software, such as COMSOL Multiphysics, ANSYS, and ABAQUS, which greatly improve the research efficiency and allow the industrial applications [102, 103]. For example, in structural engineering, numerical methods were used to identify weak points by testing the digital twin of the real structure, and improve the reliability and safety of the design [104, 105, 106, 107, 108]. In aerodynamics, numerical analysis can efficiently model flow behaviour under various conditions. One noticeable example is the use of computational fluid dynamics (CFD) simulations, a numerical approach for solving the governing equations of fluid flow, to analyse the vortex around aircraft wings [109, 110, 111]. The application of CFD provides valuable information for adjusting wing design and improving aerodynamic performance and structural stability [112, 113, 114]. Furthermore, classical numerical analysis has also contributed to fields such as biomechanics [115, 116, 117, 118], energy and thermal systems [119, 120, 121], and many other branches of science and engineering disciplines [122, 123, 124, 125].

Although classical numerical methods have achieved significant success across many scientific disciplines, this family of discretisation-based approaches still faces difficulties in certain types of problems due to their conceptual limitations. For problems whose solution fields have strong spatial and temporal variations, fine spatial meshes and small time-step sizes are always required. As a result, computational cost, including memory requirements, increases rapidly. The situation is even worse when the problem is also multidimensional [126, 127, 128, 129, 130, 131]. When dealing with highly nonlinear PDEs, numerical instability and

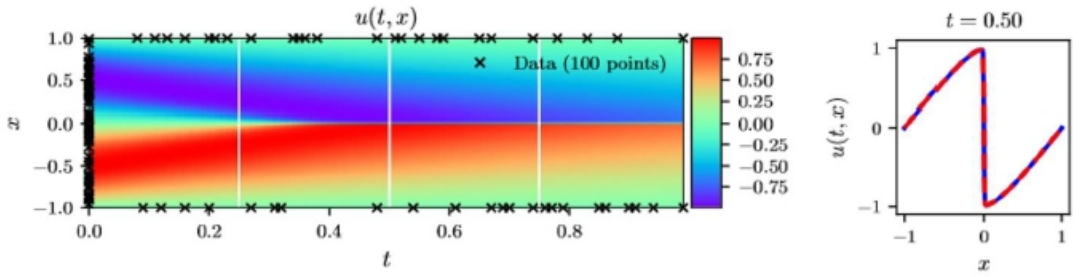
potential convergence issues continue to be major concerns and challenges despite decades of development by researchers around the globe [132, 133, 134]. In this thesis, an alternative modeling framework that combines statistical data-driven techniques with physical principles is investigated, aiming to test if the new hybrid approach can address the conceptual limitations of classical numerical methods. The new framework is introduced in the following subsection.

2.3 Physics-Informed Machine Learning

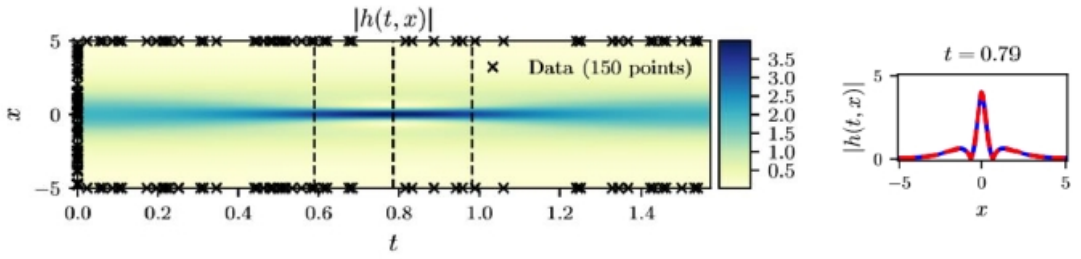
A new class of machine learning models, PINN, has been proposed as a viable alternative for solving complex partial differential equations (PDEs) [2, 11]. PINN is a class of neural networks that incorporates physical laws as constraints, along with available data, to address scientific problems effectively. PINN is a mesh-free approach that treats the process of solving PDE as an optimisation problem. The neural network for a given PDE undergoes unsupervised training to generate a functional model that minimises an aggregate of residuals of PDE at collocation points inside the problem's domain, as well as residuals of initial and boundary conditions. Once training is completed, the PINN is capable of predicting the PDE solution at any location within the domain.

PINN was first proposed in 2017 in a paper that demonstrated the method's potential by solving three representative PDEs [2]. As shown in Figure 2.3, the solutions of Burger's equation, Schrödinger equation, and Allen-Cahn show excellent agreement between PINN predictions and exact analytical solutions.

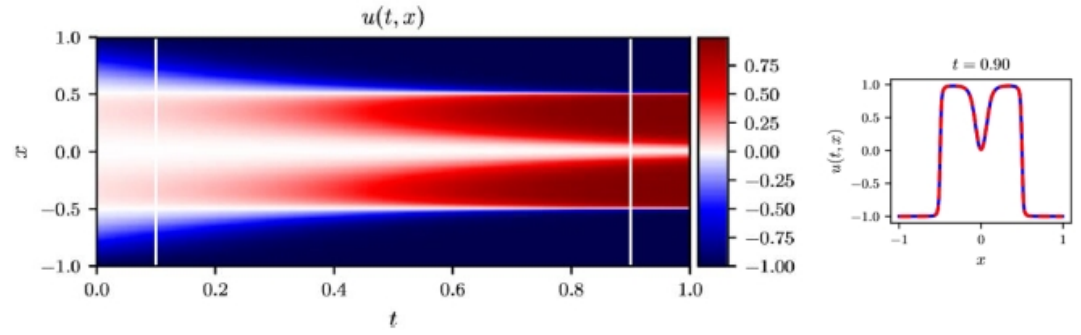
Since its introduction in 2017, PINN has been successfully applied to a range of challenging PDEs, including the Navier-Stokes equations [135, 136, 137], Maxwell's equations [138, 139], and the Schrödinger equation [140, 141, 142]. PINN has also been used in industries across different domains [10, 143]. For instance, in aerospace engineering, PINNs have been used to optimise aerofoil shapes for improved aerodynamic performance [144]. Their study found minimal discrepancies between PINN simulations and traditional Computational Fluid Dynamics



(a)



(b)



(c)

Figure 2.3: PINN predictions contours, and their comparison to analytical solutions at a given time station in graphs: (a) Burger's equation PINN solution. (b) Schrödinger equation PINN solution. (c) Allen-Cahn equation PINN solution. The figure is adopted from Raissi et al. [2] (solid and dashed lines in the graph represent analytical solutions and PINN, respectively)

(CFD). Aside from the success of PINN in solving forward problems, another important value of this framework lies in its strong potential for inverse problems. In the inverse PINN algorithm, unknown PDE parameters and the underlying solution field are estimated simultaneously from sparse and noisy observation data. The procedure for implementing an inverse PINN model is very similar to the implementation of a forward PINN model. In a forward PINN solver, the ANN parameters are optimised to approximate the PDE solution field by enforcing initial and boundary conditions together with the governing PDE as penalty terms in

the loss function. In an inverse PINN model, the ANN is trained to approximate both the solution field and the unknown PDE parameters simultaneously, with available measurements incorporated as additional constraints on top of the forward model setup during optimisation. Compared with inverse solvers developed based on classical numerical methods, inverse PINN does not require iterative forward solvers and can operate more effectively [10, 145]. Inverse PINN has been successfully applied to diverse problems, including estimating the thermal conductivity in heat transfer [146] and predicting lift and drag coefficients in fluid dynamics [11]. The inverse PINN model has also attracted significant attention in large-scale infrastructure projects. A notable example is its application in the settlement estimation of the immersed tunnel of Hong-Kong-Zhuhai-Macao Bridge, a 55 km bridge-tunnel system [147]. The customised PINN model, named the hybrid physical data-informed DNN model (HPD-DNN), was employed to predict immersion joints' axial displacement (DIS). In that study, HPD-DNN utilised a limited set of DIS monitoring data combined with fundamental relationships between DIS and volume change of adjacent tunnel segments caused by temperature changes in the concrete tunnel. The average error of prediction by the HPD-DNN model was 30% lower than that of a model trained by purely data-driven machine learning.

It is worth noting that, as a relatively new technique, PINN has not yet outperformed traditional numerical methods across all scenarios. Nevertheless, researchers continue to actively investigate PINN due to its potential to address certain classes of problems that remain challenging for conventional approaches. Cuomo et al. reviewed and summarised several drawbacks and issues that require additional attention in the development of PINN [10]. The biggest concern is the loss imbalance, as the different components of the loss function do not have the same contributions to the model training. Without appropriate balancing strategies, the optimisation process may converge poorly or lead to overfitting. The other concern is the spectral bias, particularly for problems involving high-

frequency or sharp gradient solutions. Recent studies have demonstrated the feasibility of PINN in these contexts, although further methodological advances are still required to fully boost its capabilities and broaden its practical applicability.

In this study, the PINN approach will be incorporated to solve both the ADE and RRE. This chapter will later present the limitations of existing PINN methods in addressing each equation, followed by the novel contributions introduced in this study and the specific implementation strategies employed to solve the ADE and RRE.

2.4 Prediction of Water Retention Curves

Knowledge of the WRC – sometimes referred to as soil water characteristic curve (SWCC) – is imperative for predicting water movement, mechanical behaviour and solute transport in unsaturated soil [148, 149]. WRC depicts the relationship between soil's water content (gravimetric or volumetric) and soil matric suction (referred to as 'suction' henceforth for brevity). The soil matric potential for a given water content can be measured in the laboratory by the dew-point potentiometer method, axis translation technique, or vapour adsorption and desorption methods [62, 150, 151]. However, direct measurement of WRC is time-consuming and expensive [6, 152, 153, 154, 155]. Hence, developing faster and more economical ways of estimating WRC has been of interest to researchers for decades [156].

The term PTF was introduced by Bouma in 1989 to describe mathematical tools that estimate complex soil properties from readily accessible ones, such as soil texture, porosity, and dry density [157]. PTF can be established by traditional regression or data mining methods [158]. The advantage of using machine learning and data-mining techniques, including ANN, to generate PTF, is that no prior model concept is needed [159]. The predictive power of the model depends on the quality of the dataset used to train the neural network.

The first attempt to predict WRC with neural networks was made by Pachepsky et al. in 1996 [160]. They found that an ANN model performed better than traditional regression analysis. Three research questions have been of interest to scholars since this early development [160], but especially in the past decade which saw accelerating interest in ANN. What is the best PTF form for predicting WRC? What are the most critical soil properties to be used as input variables in ANN for achieving high prediction accuracy? What is the most suitable ANN structure for WRC estimation? These three questions are discussed next.

WRC-PTF can be classified into three categories, namely point, parametric, and continuous PTFs. Point PTFs estimate soil water content at several predefined pressure heads (e.g. -1, -10, -100, -1500 kPa, etc) [159, 161]. A drawback of point PTF is that soil water content at a matric potential other than the predefined ones cannot be directly determined from the model.

Parametric PTF is an alternative approach that directly estimates the fitting parameters of analytical WRC equations [61, 148, 149, 159, 161, 162, 163]. The advantage of using parametric PTF is, in addition to its ability to estimate water content at any value of matric potential, the reduction of the WRC to a small number of parameters (usually 2 or 3 depending on the specific analytical equation adopted). These parameters can then be readily used in computational modeling of unsaturated soil behaviour. However, a key limitation of the parametric PTF is that its accuracy depends on how well the predefined analytical WRC equation captures the real shape of the WRC.

Finally, continuous PTF overcome the drawbacks of point and parametric PTF by introducing the matric potential as an additional input parameter, with the model yielding the water content at this specific matric potential as the only output [164]. Hence, continuous PTF are not limited to the predefined matric potential and do not depend on the choice of the WRC analytical equation. On the other hand, it is possible to fit, a posteriori, the predictions of a continuous PTF for a given set of soil parameters to an analytical function, hence combining

the advantages of parametric and continuous PTF. Nevertheless, the accuracy of the continuous approach still depends on suitable coverage by the training dataset of a wide range of suction values.

As mentioned above, determining the best combination of input variables when developing PTF to estimate WRC is another research pursuit. It has already been proven that soil texture data is a fundamental soil property required for a good WRC prediction [148, 149, 160, 165]. Some additional input variables may lead to further improvements in predictive power, but others may have the opposite effect, and research on which additional variables are beneficial remains inconclusive. For example, Minasny and McBratney [165] compared the performances of PTF with three (% sand, % silt, % clay) and four (% sand, %silt, % clay, dry density) input parameters. They found that introducing soil dry density as input variable reduced the mean root mean squared error (RMSE) from 0.049 m^3/m^3 to 0.043 m^3/m^3, a 12.3% improvement in accuracy. This conclusion was reinforced in later work by Bayat et al. [162]. On the other hand, results reported by Haghverdi et al. [164] did not show a significant improvement in WRC prediction with the addition of either soil organic carbon content or dry density for all three types of PTF models considered (point, parametric, and continuous).

As discussed earlier, the size and structure of the ANN model are another critical factor that influences the accuracy of PTF predictions [166]. The feedforward neural network has been widely adopted in previous research on estimating WRC by machine learning and data mining techniques, and shows considerable accuracy [148, 149, 152, 158, 159, 164, 165]. In this thesis, the impact of network architecture and model configurations on the prediction performance of the WRC will be systematically investigated.

Despite the wide adoption of ANN models, an important question remains insufficiently explored: can DNN, with increased model complexity, offer a significant performance advantage over SNN in predicting the WRC? Achieng [167] compared SNN and DNN performances in regressing suction measurements onto

a WRC equation for four loamy sand samples but did not attempt to build a PTF model, since the model can only be used if experimental measurements for the soil in question already exist. To the best of the author's knowledge, studies for building a PTF model based on DNN or comparing the performances of SNN and DNN for the prediction of WRC from easy-to-measure soil properties are lacking.

In addition to model complexity, another important factor affecting WRC prediction accuracy is the representation of soil texture classes in the training dataset. Previous studies have shown that the estimation accuracy of PTFs for predicting WRC is highly sensitive to the prevalence of specific soil texture classes in the dataset [63, 168]. When certain textures are underrepresented, the resulting PTFs tend to perform poorly for those classes. This issue is especially pronounced for clayey soils, which are often sparsely represented in global datasets [168, 169]. A notable exception is the study by Singh et al. [170], which employed a dataset where clay samples accounted for 48.1% of the total, leading to improved prediction performance for that class. This thesis will investigate the accuracy of WRC predictions for clayey soils and assess how their representation in the training data influences the performance of PTF models in predicting clayey soil's WRC.

2.5 Solute Transport: Advection-Diffusion Equation

ADE is commonly used to describe a wide range of problems including contaminant transport in soil, water and air [171, 172, 173], drug delivery through the circulatory system [174, 175], sediment transport [176, 177] and financial forecasting [178]. Of particular importance is the use of the equation to model the degradation of groundwater quality caused by pollutants leaching from surface sources [179, 180]. The ADE has been shown to capture key processes governing contaminant transport in the subsurface and has therefore been widely used in water quality modelling [181, 182, 183, 184].

The ADE carries two key components, namely a) molecular diffusion or mechanical dispersion and b) advection. Molecular diffusion refers to the redistribution of solute in soil water along concentration gradients caused by Brownian motion. Mechanical dispersion, also proportional to concentration gradients, is driven by fluctuations in pore-scale seepage around the mean velocity. Both diffusion and dispersion are often represented by a Fickian diffusion equation and quantified through the hydrodynamic dispersion coefficient (HDC) which incorporates both processes in a single coefficient. Advection, on the other hand, is a process through which solutes are carried by the movement of water in soil. The importance of advection relative to diffusion/dispersion is captured by the Peclet number: $P_e = vL/D$, where v is the seepage velocity, L is a representative length, and D is the HDC.

At low P_e, diffusion/dispersion is dominant over advection, the ADE is primarily hyperbolic in nature and common discretisation techniques such as the FEM, FDM, finite volume (FVM) or boundary element methods (BEM) yield stable numerical solutions. However, when P_e is high, advection is dominant, the ADE becomes primarily parabolic, and results generated by the above conventional numerical methods suffer from spurious numerical oscillations. Several remedies have been proposed in the literature to tackle this problem including discretisation mesh refinement, special shape functions and stabilisation techniques [8, 185, 186, 187, 188]. However, these approaches do not always work, can be computationally expensive and may require complex modifications to numerical algorithms.

As previously discussed, PINN provides an AI-driven framework for solving PDEs and has been successfully applied to several challenging problems. In recent years, significant developments have been made in using PINN to solve the ADE. He and Tartakovsky [189] showed that PINN can solve problems with P_e up to approximately 60 in one-dimensional steady-state cases and up to 200 in two-dimensional cases. However, they observed a decline in performance as

P_e increased, particularly in scenarios involving complex boundary conditions and time-dependent dynamics. Hou et al. [190] and Sukumar and Srivastava [191] reported a good fit between PINN predictions and analytical solutions for steady-state problems with P_e of 50 and 100, respectively. Similar outcomes under low to moderate P_e regimes have been reported by several other studies [192, 193, 194, 195, 196, 197, 198].

Early attempts to tackle high P_e problems using PINN include the work of Gomes et al. [199], who explored several steady-state ADE cases with P_e values up to 1000. They tested two PINN structures to solve the problem. The first structure used the standard PINN models with only the coordinates as inputs, but it resulted in large discrepancies between PINN predictions and analytical solutions. However, they reported that if an additional diffusion/dispersion coefficient is added as an input feature, the model yielded good results at P_e values up to 1000, though with some errors at low P_e values. Sikora et al. [200] and Hsieh and Huang [201] also attempted to address high P_e problems. Similar to Gomes et al., they observed poor performance when using standard PINN models without modification. However, Hsieh and Huang [201] achieved better outcomes by incorporating a weakly imposed boundary conditions (WBC) approach combined with transfer learning, in which knowledge from a pre-trained model is reused to improve the training performance of a new model. Although their results showed significant improvements, there was still a sharp fluctuation near the boundary that was absent in the analytical solution. Currently, no PINN approach has accurately solved steady-state ADE problems across a wide range of P_e. In contrast, accurate PINN solutions for transient ADE with higher Peclet numbers have been reported. He and Tartakovsky [189], Niu et al. [202] and Bo et al. [203] all reported accurate predictions of transient ADE solutions at Peclet numbers of 200, 1000 and 6000, respectively.

Most of the above-mentioned studies have considered only Dirichlet-type boundary conditions, with only a few exploring the ability of PINN to simulate Neumann-

type boundary conditions. The latter are more complex because they require the enforcement of gradient variables at or near the boundary. However, they are critical for real-life applications of the ADE in science and engineering.

Hence, several gaps remain in the application of PINNs to ADEs: (1) Solving steady-state ADE with PINNs under high P_e is still challenging, (2) existing methods often struggle with accuracy and stability as advection dominance increases and (3) little consideration has been given to Neumann-type boundary conditions.

In standard PINN algorithms, the loss function typically includes three components, namely the residuals of the PDE, boundary conditions and initial conditions. In other words, the boundary condition (BC) and initial condition (IC) are enforced in a 'soft' fashion, by ensuring that deviations from the correct BCs and ICs incur a penalty during the optimisation process. This approach maintains a useful flexibility in the optimisation process which enables the model to search a broader solution space for a more globally optimal solution. However, multiple residual terms in the loss function may also compete during the optimisation process and their contributions may not always be effectively balanced, hence leading to poor convergence [195, 204, 205, 206].

To address this problem, some researchers have applied relative weights to each loss term, resulting in some improvements [137, 144, 207]. However, adjusting these weights is highly case-specific and determining the appropriate weighting for the boundary condition loss term can be challenging [208]. It has also been shown that soft boundary constraints often lead to suboptimal solutions [209].

Recognising these challenges, an alternative approach in which the BCs and ICs are enforced in a 'hard' fashion has been proposed by Krishnapriyan et al. [205]. This method was first applied by Lu et al. [210] who demonstrated the effectiveness of a hard-constraint (HC) PINN in solving a holography problem in optics and a fluid problem of Stokes flow. HC-PINN has since been successfully applied to problems of wave propagation [208, 211, 212].

To the best of the authors' knowledge, only two studies have applied HC-PINN to ADE problems. Mojgani et al. [213] proposed a PINN within a Lagrangian framework to solve advection-dominated ADE problems, up to $P_e = 6000$, albeit with periodic boundary conditions. Li et al. [195] developed an HC-PINN that uses Fourier-induced activation functions to address ADE problems with high-frequency components but low Peclet numbers.

When using HC method, a distance function needs to be defined to enforce constraints at spatial boundaries or initial time while restricting the spread of its effect. A distance function helps demarcate boundary/initial conditions points from the rest of the spatial/time domains [195, 208, 210, 211, 212]. It should be smooth and maintain the differentiability of the neural network output [195]. While the possibility of using non-linear distance functions has been raised by some authors (e.g., [195, 214]), to the best of the authors' knowledge, no attempt has been made to explore the effects of different types of distance functions on the performance of hard-constraint PINN. Two questions hence arise as to what the most effective form for a distance function is and whether non-linear distance functions can improve the accuracy and convergence speed of a hard-constraint PINN.

Hence, in summary, existing literature reveals several unresolved challenges in the application of PINNs to the ADE, particularly under conditions of high P_e. In this thesis, the following gaps will be addressed, namely a) improving the accuracy and convergence speed of PINN solutions to steady-state and transient ADE under high P_e values, including both Dirichlet and Neumann boundary conditions, and b) assessing the ability of a hard-constraint approach to improve PINN predictions while exploring the effects of different types of distance functions associated with it. Following this, a new PINN is proposed and evaluated, based on a hard constraints approach with a P_e-adjusted exponential distance function. The proposed PINN is shown to be able to tackle significantly higher Peclet numbers than the current PINN solutions found in the literature.

2.6 Water Flow in Unsaturated Soils: Richards' Equation

Understanding and modeling water movement in the vadose zone is important for a range of engineering and environmental problems including agricultural irrigation management [215, 216, 217], slope and foundation stability [218, 219, 220, 221], and contaminant transport in soil and groundwater protection [222, 223, 224]. Over the past century or so, soil water movement in unsaturated soil has been modelled by the Richards Equation [225]. Richards applied the principle of fluid mass conservation and incorporated capillary forces into Darcy's law, deriving a governing equation for water movement in unsaturated soil. The Richards equation involves two constitutive relationships that are highly non-linear, namely between a) pressure head and water content, a relationship commonly referred to as WRC, and b) hydraulic conductivity and pressure head (known as hydraulic conductivity function, HFC). The few available analytical solutions of the Richards equation typically rely on highly simplifying assumptions for boundary conditions and material properties (e.g., using Gardner's equation for WRC) [226, 227, 228, 229]. Hence, the equation is much more commonly solved numerically [4]. However, the highly non-linear nature of the equation makes it difficult to solve numerically as well, and despite significant advances in the past decades, research continues on developing more accurate, stable and computationally efficient solutions to this equation [230, 231].

Recent applications of PINN to nonlinear PDEs suggest it may offer a promising framework for addressing such problems. However, there have been relatively few attempts to use PINN to solve the Richards equation [12, 13, 14, 15, 16]. The first attempts were made by Bandai and Ghezzehei [12, 13] who reported a PINN inverse solver of the Richards equation to estimate fitting parameters of the Van Genuchten equation. Employing three concurrent neural networks and a monotonicity constraint applied to water content and hydraulic conductivity, they showed that PINN can effectively solve the inverse problem but did not provide a solution to the forward problem. Depina et al. [15], on the other

hand, used a single network in conjunction with closed-form WRC and hydraulic conductivity functions and successfully solved the inverse problem. However, exploring the ability of PINN to solve the forward solver, they reported difficulty in achieving consistent convergence to sufficiently accurate solutions, owing to the non-linearity of the equation, and hence did not present quantitative results. To the best of the authors' knowledge, Chen et al. [14] and Haruzi et al. [16] are the only two papers to present a PINN forward solver of the Richards equation. Both studies employed a method that dynamically adjusts the weights of the loss terms in the PINN residuals, and their results demonstrate that this approach can achieve high accuracy. However, Chen et al. [14] mentioned that a drawback of their approach is that it requires problem-specific adjustment of weights of various contributions to the loss term (PDE and boundary and initial conditions) in order to achieve the desired accuracy.

Based on the above, there remains significant room for improvement of PINN approaches to the solution of the Richards equation. Enhancing the predictive performance of AI models can be achieved by either improving the algorithm's expressive capacity or increasing the quality of the training data [232, 233, 234]. In this study, both strategies will be explored. First, as noted in the ADE section, the soft enforcement of loss terms in traditional PINNs can lead to convergence challenges when solving complex PDEs [195, 204, 205, 206]. In this study, the hard-constraint approach is also evaluated to assess whether reducing conflicts among loss terms can enhance model performance and prediction stability.

Second, this study also investigates methods to enhance the quality of training data. As an unsupervised learning approach, the quality and distribution of these collocation points play a critical role in determining the training performance and accuracy of the model. In the PINN framework, a series of collocation points are selected from the physical domain to enforce the underlying PDE. In more conventional PINN, the location of these points is predetermined and remains fixed throughout the entire training process. Lu et al. [210] first suggested adap-

tive resampling in which, after a few training epochs, the PDE residuals at each point are evaluated, and additional collocation points are adaptively deployed in regions with higher residual values. This refinement allows the model to focus more on regions with higher errors, and the results show a significant improvement compared to fixed sampling. Following this work, adaptive resampling has shown the capacity to improve the performance of PINN in several applications [235, 236, 237, 238]. Wu et al. [239] compared several resampling strategies and found that residual-based adaptive refinement yielded the most promise.

Based on the above, in this thesis, the following aims will be pursued:

1. To develop a reliable forward solver for Richard's equation using the PINN framework;

2. To investigate the effect of improving training data quality by adaptively resampling collocation points to strengthen the model training in regions with high residual errors.

2.7 Coupled Water Flow and Solute Transport in Unsaturated Soils

So far, the use of PINN to solve the ADE and RRE has been reviewed separately. However, coupling the two equations is of great interest in a range of fields in water resources, geoenvironmental engineering and waste management. Only one prior study has attempted to couple these two equations within a unified PINN framework to model the complete water and solute transport in unsaturated soils [16]. Haruzi and Moreno [16] proposed an innovative approach by employing PINN in combination with time-lapse geoelectrical data to reconstruct the spatiotemporal distribution of water content and solute concentration in unsaturated soils. Their method integrates established physical models with indirect observational data derived from geoelectrical measurements, along with available initial and boundary conditions, to estimate the coupled processes governing water movement and solute transport. The geoelectrical data used in their study

are apparent resistivity measurements, which reflect the bulk electrical resistivity of the subsurface. This bulk resistivity is influenced by both the water content and the electrical conductivity of the pore fluid, making it a coupled function of these two variables [240, 241]. Conversely, changes in water content and solute concentration at a given spatiotemporal location will result in corresponding changes in the apparent resistivity. When integrated with PINNs, this indirect information is incorporated as a soft observational constraint, helping guide the network to learn physically consistent representations of the underlying water and solute distributions.

Integrating indirect observational data with physics-based learning significantly enhances predictive accuracy and robustness, particularly in scenarios where direct measurements are limited. However, as Haruzi and Moreno also acknowledge, there are certain limitations to this approach [16]. First, the inversion of geoelectrical signals is an ill-posed problem, which often leads to smoothing artefacts and structural inaccuracies [242, 243]. Second, the traditional interpretation of bulk electrical conductivity cannot unambiguously separate the effects of water content and solute concentration, introducing uncertainty in the recovered distribution of [244, 245]. Furthermore, geoelectrical data are not always available in practice. In the absence of such data, the PINN model alone may not be able to accurately predict the coupled dynamics of water and solute transport.

In this thesis, the individual studies conducted on PINN solvers for ADE and RRE serve as foundational components for the development of a coupled PINN framework. The knowledge gained from studies in previous sections, including their structure, optimisation strategies, and treatment of boundary and initial conditions, will be utilised to construct a coupled model that simultaneously simulates water flow and solute transport in unsaturated soils. This coupling aims to capture the dynamic interactions between moisture movement and solute distribution.

Chapter 3

Methods

3.1 Overall Research Design

This thesis aims to bring about advances in AI-based modeling of water flow and contaminant transport in unsaturated soils. Figure 3.1 shows the overall research design.

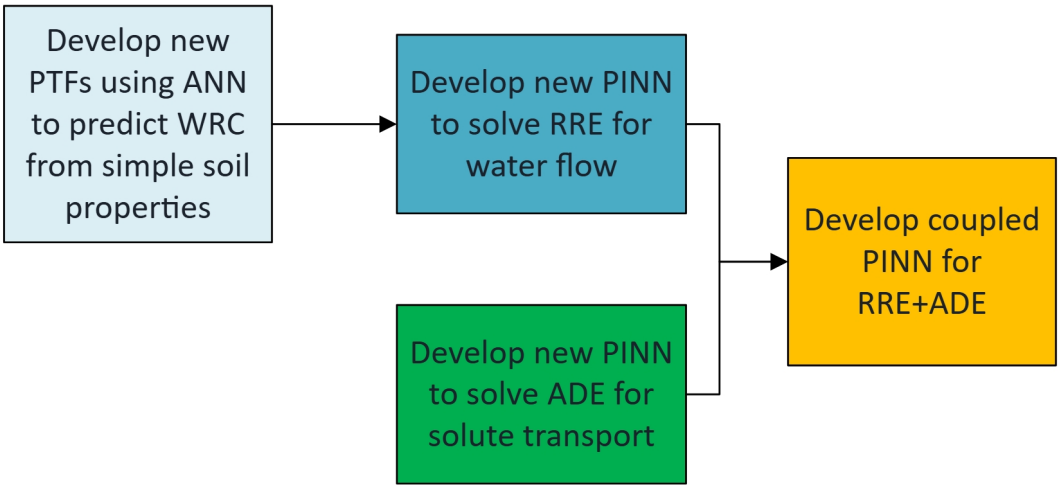


Figure 3.1: Overall Research Design (PTF: Pedo-Transfer Functions; NN: Neural Network; PINN: Physics Informed NN; RRE: Richardson-Richards Equation; ADE: Advection-Diffusion Equation)

Guided by the statistical learning theory, the model development in this thesis is framed as a process of systematically reducing predictive error. The error can be decomposed into three components, which are approximation error, estimation error, and irreducible error. The magnitude of the approximation error is related to how well the developed model can approximate the underlying functions, while the estimation error is related to how efficiently the model development process utilises the available dataset. The irreducible error is caused by the noise that

exists in the data and cannot be eliminated during model development. The models built in this thesis focus on reducing approximation and estimation errors across all subprojects.

Two new machine learning models are first developed to predict WRC from readily available soil properties, using available databases. The first model is general in scope while the second is aimed specifically at clay and clay-rich soils. These can be used when direct measurements of WRC is not possible. Then, a newly introduced machine learning framework, PINN, is employed to model solute transport by solving ADE. This framework is also applied to model unsaturated flow governed by RRE. Finally, the ADE PINN solver and RRE PINN solver are combined to simulate the coupled water and solute transport in unsaturated soils.

This chapter presents various aspects of this methodology and provides a rationale for key methodological choices made in the thesis.

3.2 Artificial Neural Network

ANN is a machine learning algorithm for classification or regression tasks. The development of an ANN is inspired by the animal brain's biological neural network (BNN) (Figure 3.2). The most basic element of an ANN model is the artificial neuron. As shown in Figure 3.3, the artificial neuron comprises several inputs, weights for each input, a summation function, a bias, an activation function, and outputs. Weighted inputs are summed and augmented by a bias value before passing through the activation function. The activation function is an important step that adds nonlinearity to the process, allowing the network to better capture complex relationships in the training data. The selection of an activation function for the hidden neuron can significantly influence the outcome of the whole model [246]. The suitability of each activation function is dependent on the specific dataset and input, and output parameters [247]. In this study, several widely used activation functions are assessed for their suitability, including Sigmoid,

hyperbolic tangent (Tahn), rectified linear unit (ReLU), and leaky rectified linear unit (Leaky ReLU) (Figure 3.4). Mathematically, the data transformation process of an artificial neuron shown in Figure 3.3 can be expressed by:

$$y_j = f[(\sum_{i=1}^{n} W_{ij} x_i) + b_j] \tag{3.1}$$

where x_i is the i^{th} input, W_{ij} is the weight connecting the i^{th} input to the j^{th} neuron, and b_j is the bias associated with the j^{th} neuron, n is the number of inputs for this neuron. The weight (W_{ij}) and bias (b_j) are adjusted iteratively during the training process until the predicted output from the network is close enough to the data. Then, the combination of weights and bias is considered representative of the complex relationship between input and output data. The output of each artificial neuron serves as the input of the next available artificial neuron and so on, as illustrated in Figure 3.3.

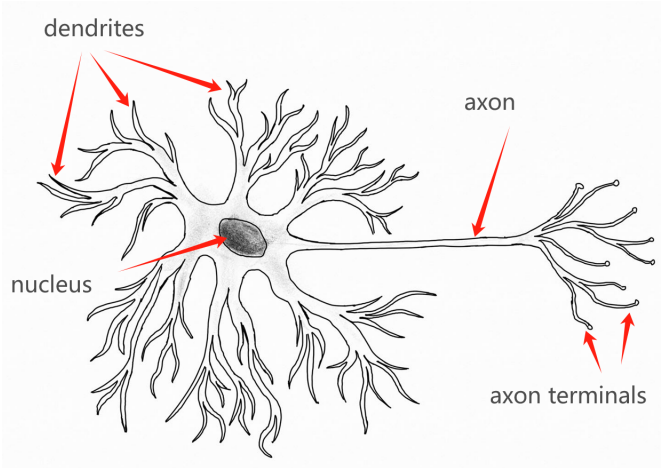


Figure 3.2: Schematic of a biological neural network

ANN models are always composed of three layers: input layer, hidden layer, and output layer. Each layer consists of several neurons. The schematics of the ANN structures are shown in Fig. 3.5 for both a single hidden layer (SNN; Figure 3.5a) and multiple hidden layers (DNN; Figure 3.5b). The network takes inputs $x_1, x_2, \ldots, x_n$, and approximates the solution $u(x_1, x_2, \ldots, x_n)$. This is achieved by passing the inputs through multiple layers of interconnected neu-

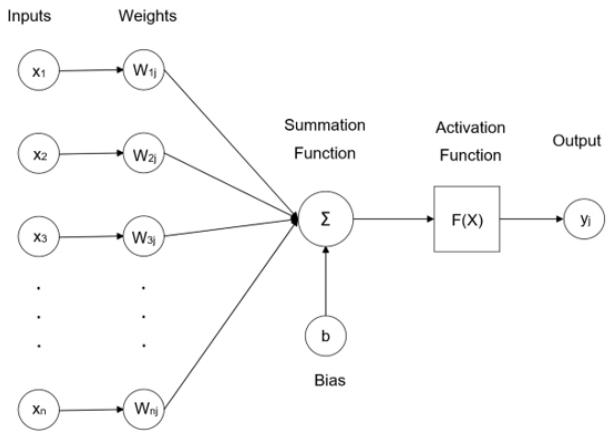
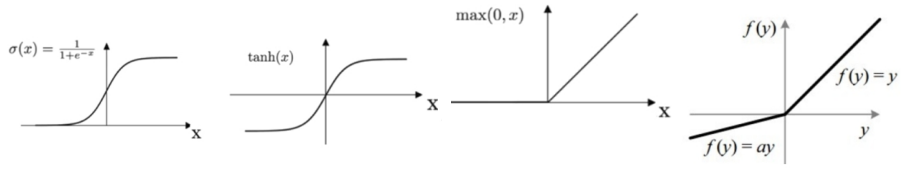


Figure 3.3: Schematic of an artificial neuron's working principle



(a) (b) (c) (d)

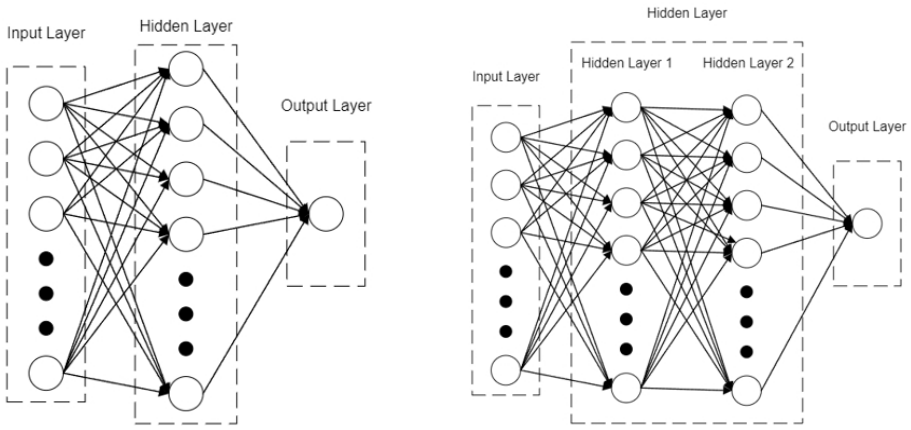
Figure 3.4: Activation functions compared in this study. (a) Sigmoid (b)Tahn (c) ReLU (d) Leaky ReLU.

rons, where weighted sums and activation functions are applied at each layer to transform the data and produce the final output. Mathematically, the approximate solution $u(x,t)$ can be represented by the formula:

$$u(x_1,\ldots,x_n) = \mathbf{W}_{nl}\,\sigma\Big(\mathbf{W}_{nl-1}\,\sigma\big(\cdots\sigma\big(\mathbf{W}_1\begin{bmatrix} x_1 \\ \vdots \\ x_{nl}\end{bmatrix} + b_1\big)\cdots + b_{nl-1}\big) + b_{nl}\Big) \tag{3.2}$$

where nl is the total number of hidden layers, $\mathbf{W_i}$ is the weights matrix of the i^{th} hidden layer, b_i is the biases matrix of the i^{th} hidden layer, and σ is the activation function applied to the neuron in order to inject nonlinearity into the network.

Both SNN and DNN are composed of several artificial neurons organised in



(a) (b)

Figure 3.5: Structure of neural network models built in this study. (a) SNN (b) DNN.

a three-layer structure. The neurons in adjacent layers are fully connected. The weights and biases in each neuron are adjustable to let the approximation of the network be closer to the expected outcome, as measured by a loss function [206]. In the present study, the mean squared error (MSE) is used as a loss function:

$$MSE = \frac{1}{n_a}\sum_{i=1}^{n_a}(p_i - a_i)^2 \tag{3.3}$$

where p_i is the predicted data of the i^{th} data point, a_i is the actual value of the i^{th} data point, and n_a is the total number of data points to be analysed.

3.3 Water Retention Curve

Water retention behaviour is a fundamental soil hydraulic property that describes how soil retains water at different suction pressures under unsaturated conditions. It is mathematically represented by the water retention curve, a non-linear function that describes the relationship between soil suction pressure and water content. Traditionally, determining the WRC involves measuring soil water content across a wide range of matric potentials, followed by curve fitting using empirical models.

Over decades of development, several empirical models have been developed to

describe the WRC. Three widely used empirical models are the Brooks-Corey, Van Genuchten and Fredlund-Xing models. The Brooks-Corey model was introduced in 1964 by Professors Royal Harvard Brooks and Arthur Thomas Corey [248]. This four-parameter water retention model is defined as:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{(\alpha_{bc}|h|)^{\lambda_{bc}}} \tag{3.4}$$

where h is the suction pressure, $\theta(h)$ is the volumetric water content with respect to a specific h, θ_r is the residual water content (m^3/m^3), θ_s is the saturated water content (m^3/m^3), α_{bc} is the empirical fitting parameter of the Brooks-Corey model related to air entry pressure, and λ_{bc} is another fitting parameter.

Another well-established WRC model is the Van-Genuchten (VG) model, which was developed by Professor Martinus Th. Van Genuchten in 1980 [249]. The general form of this formulation is expressed as follows:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha_{vg}|h|)^{n_{vg}}]^{m_{vg}}} \tag{3.5}$$

where α_{vg} is the Van Genuchten model fitting parameter related to the air-entry pressure, n_{vg} and m_{vg} are the other two fitting parameters which are often assumed to be related by: $m_{vg} = 1 - 1/n_{vg}$. The extended VG model can be used to derive the hydraulic conductivity function (HCF), another important hydraulic parameter, based on the effective saturation, and is expressed as:

$$K(h) = K_s \cdot S \cdot \left[1 - \left(1 - S^{1/m_{vg}}\right)^{m_{vg}}\right]^2, \tag{3.6}$$

where K_s [L T^{-1}] is the saturated hydraulic conductivity, and S is the effective saturation, defined as:

$$S = \frac{\theta - \theta_r}{\theta_s - \theta_r}, \tag{3.7}$$

Another frequently used WRC formulation, known as the Fredlund and Xing

model, was proposed in 1992 [250]. The relationship between soil suction and volumetric water content is formulated as:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha_{fx}|h|)^{n_{fx}}]}(1 + (\beta_{fx}|h|)^{m_{fx}})^{-\lambda_{fx}} \tag{3.8}$$

where α_{fx} is the Fredlund and Xing model fitting parameter related to the air-entry pressure, β_{fx}, and m_{fx} are two fitting parameters related to pore structure, n_{fx} and λ_{fx} are another two fitting parameters.

3.4 General PTF models for WRC

3.4.1 Data Source

The dataset used for building PTFs for WRC is extracted from UNSODA 2.0 (unsaturated soil database 2.0) [251]. This dataset contains measured soil hydraulic properties and pedological information about 790 samples collected from around the world. UNSODA 2.0 is widely used in the modelling and investigation of unsaturated zones [153, 168, 252, 253]. According to Feike [254], UNSODA database is assembled from contributions by individual scientists and published peer-reviewed publications. The data is carefully judged by the database managers before adding any entry to the database, based on criteria previously agreed at a soil-hydraulic workshop in Riverside in 1989 [254].

3.4.2 Models

To investigate the effects of different parameters on SWRC prediction, thirteen continuous models (M1 to M13) with different sets of input parameters, shown in Table 3.1, are built. In addition to suction, a maximum of eight input parameters are included, namely soil texture (%clay, %sand, %silt), dry density (DD), porosity (PO), specific gravity (SG), depth of sample (DS) and organic content of sample (OM). In all models, the output is the volumetric water content. The robustness of ANN to multicollinearity enables the simultaneous inclusion of some

correlated variables, such as the 3 soil fractions [255, 256]. Additionally, the robustness of ANN to multicollinearity allows the inclusion of correlated features during prediction [257]. This is important because in this study, sand, silt, and clay contents were all included as input variables, which clearly leads to multicollinearity. Although these three fractions should theoretically add up to 100%, measurement methods and inherent errors often lead to small deviations, so the total sum does not always equal exactly 100%. Therefore, excluding one component may affect the true relationships among soil texture components and target variables. Including all three components has been the conventional practice in previous WRC PTF studies, as it allows for a more accurate representation of soil texture [148, 164, 165, 258]. This has also been confirmed in this project: models built with two out of the three texture variables have been found to lose predictive accuracy. Hence, all 3 variables are kept in the models.

Table 3.1 summarises the input parameters considered for each model (from M1 to M13). This selection of models is made so as to allow an assessment of the effects on the accuracy of each variable (other than the 3 soil texture variables) by comparing the performances of models with and without that variable. When one of the three parameters of dry density, specific gravity and porosity is missing while the other two are available, it is possible to infer the former from the latter as follows:

$$DD = SG \times (1 - PO) \tag{3.9}$$

The total dataset is divided into two dedicated subsets: a training dataset and a testing dataset. The training dataset is used during the training phase, followed by testing using the testing dataset. In addition, 20% of data in the training dataset is used for validation during the training process. In other words, the training dataset is divided into a "training" subset and a "validation" subset. The model is trained on the "training" subset and then evaluated on the "validation" subset to fine-tune the model parameters and select the best model. This

approach has two advantages. First, it ensures that the model is evaluated during the training process, which is essential to prevent overfitting and ensure the model has good generalisation performance. Second, it allows the use of available data efficiently, which is particularly important when the dataset is not very large. The same testing dataset is used for all thirteen models to ensure consistency in performance evaluation and comparison of models.

Table 3.1: Input parameters combination of models M1 to M13

Input	Clay%	Silt%	Sand%	DD	PO	SG	Depth	OM	Suction
M1	✓	✓	✓						✓
M2	✓	✓	✓	✓					✓
M3	✓	✓	✓	✓	✓				✓
M4	✓	✓	✓	✓	✓	✓			✓
M5	✓	✓	✓	✓	✓	✓	✓		✓
M6	✓	✓	✓	✓	✓	✓	✓	✓	✓
M7	✓	✓	✓	✓		✓			✓
M8	✓	✓	✓	✓		✓	✓		✓
M9	✓	✓	✓	✓		✓	✓	✓	✓
M10	✓	✓	✓	✓			✓		✓
M11	✓	✓	✓	✓			✓	✓	✓
M12	✓	✓	✓		✓				✓
M13	✓	✓	✓		✓	✓	✓		✓

If either dry density, specific gravity, or porosity is not available, and the other two values are known, the absent parameter can be calculated using the relationship outlined in Equation 3.9, resulting in 3217 data points for all models. Points with missing input are also removed from training datasets, however as different models have different input parameters, the training datasets for different models will not be identical to each other (after points with missing data are removed). Therefore, the higher the number of input variables used in the model, the higher the likelihood that some data points are incomplete, and therefore the lower the number of data points available for training. Distribution of the soil samples in the dataset over the range of soil texture, based on USDA classification, is shown in Figure 3.6, revealing a paucity of clay samples in the data. The number of available training data points for each model is summarised in Figure 3.7. On the other hand, a reasonably uniform distribution of data points over the suction head

axis for both training (3217 data points) and testing (122 data points) datasets can be seen in Figure 3.8. These figures also show that the range of volumetric water contents in the training dataset for model M1 is wider than that of the more restricted testing dataset used for all models including M1, for the reasons described above.

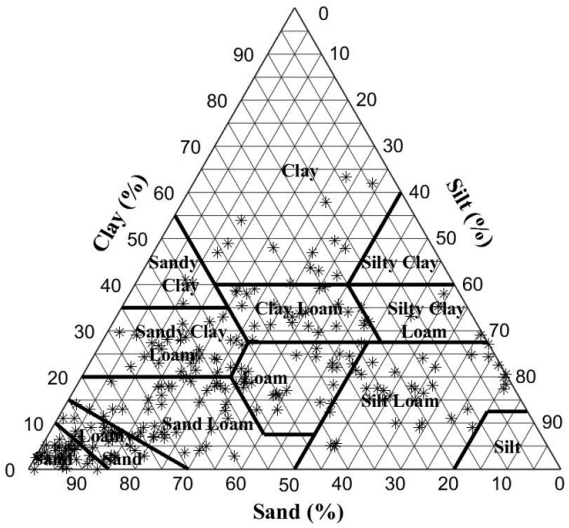


Figure 3.6: Soil texture ternary diagrams for soil samples

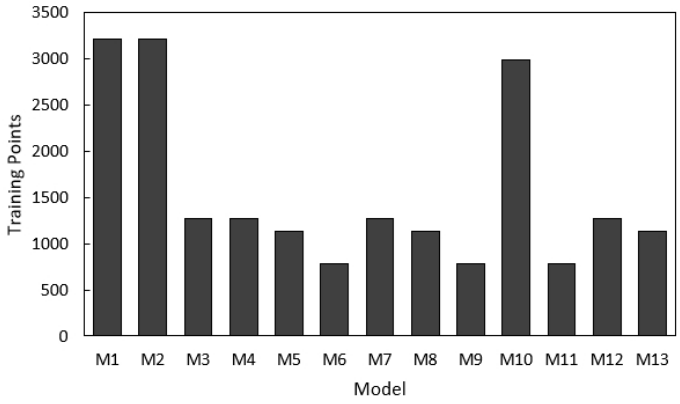
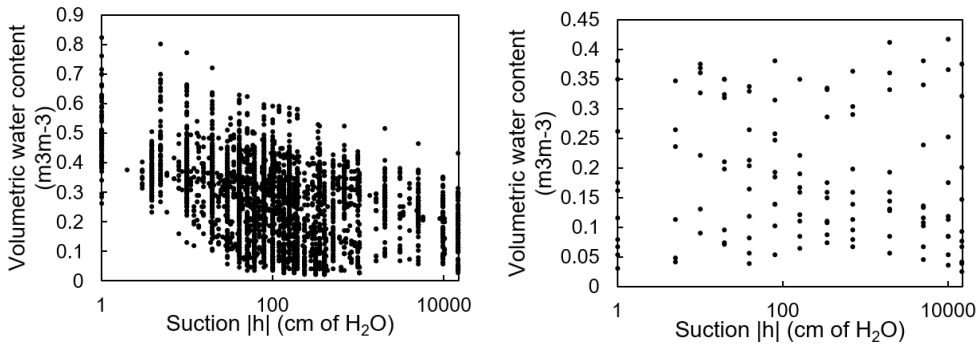


Figure 3.7: Number of available training data points for all models



(a) (b)

Figure 3.8: Distribution of datapoints over wide ranges of volumetric water content (m^3m^{-3}) and pressure heads |h| (cm of H_2O). (a) Training Dataset (for model M1), (b) Testing Dataset (for all models).

3.4.3 Development of PTF models for General WRC estimation

For each of the 13 models, one SNN-PTF and one DNN-PTF are built and evaluated, with a total of 26 ANN models hence constructed (SNN-PTF-M1 to SNN-PTF-M13 and DNN-PTF-M1 to DNN-PTF-M13). Four activation functions, Sigmoid (Figure 3.4a), Tanh (Figure 3.4b), ReLU (Figure 3.4c), and Leaky ReLU (Figure 3.4d), are implemented, their performances evaluated, and the best-performing function selected.

As a first step in data preprocessing, data points with missing variables or some variables falling outside an acceptable range (e.g., volumetric water content values less than 0 or exceeding 1) are removed. After examining the data for outliers, the potential influence of outliers is mitigated by applying a logarithmic transformation to the matric potential data, and all input variables are normalised to the range of 0 to 1.

Hyperparameter tuning is applied whereby the number of neurons used in each hidden layer, the gradient coefficient of Leaky ReLU activation function, and training-related parameters (e.g., learning rate, batch size) are carefully selected to optimise the performance of the machine learning model.

Model development has been conducted using Python 3.8, machine learning

platform TensorFlow [259], application programming interface (API) Keras, and Python library Scikit-Learn [260]. The correlation between different input variables in the training dataset is evaluated using the correlation coefficient (r):

$$r = \frac{\sum(V1_i - \bar{V1})(V2_i - \bar{V2})}{\sqrt{\sum(V1_i - \bar{V1})^2 \sum(V2_i - \bar{V2})^2}} \tag{3.10}$$

where $V1_i$ is a given input variable, $\bar{V1}$ is the mean of $V1_i$ in the dataset, $V2_i$ is another input variable, $\bar{V2}$ is the mean of $V2_i$. r is also used to assess the correlation between two input variables. r in Eq. 3.10 is also used to assess model performance by replacing $V1$ and $V2$ with p and t, respectively. The performance of ANN models is also evaluated by the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n_a}\sum_{i=1}^{n_a}(p_i - a_i)^2} \tag{3.11}$$

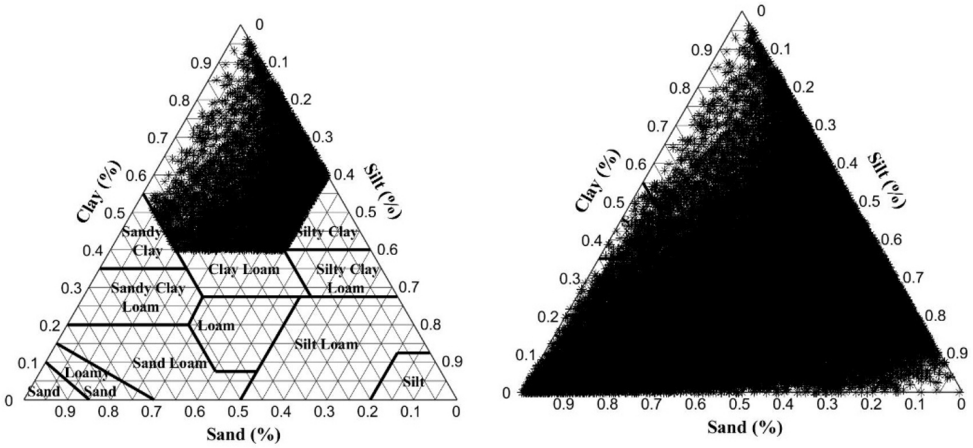
where p_i is the i_{th} predicted value, a_i is the i_{th} actual value.

3.5 PTF models for WRC of clayey soils

3.5.1 Data Source

The datasets used in the study of WRC prediction for clay soils are extracted from the National Cooperative Soil Survey (NCSS) Soil Characterisation Database. Compared to UNSODA, the NCSS Database provides a larger and more comprehensive dataset for clay soils, which is beneficial for developing and validating predictive models. The NCSS database consists of soil properties collected from laboratories across the United States. For our analyses, only data featuring soil texture, soil density, suction and water content were used. Two datasets are extracted from the NCSS database for the training of three PTF applications. Dataset C1 (Figure 3.9a), which is a subset of NCSS database with only clayey soil data retained. Dataset C2 consists of data from all soil texture classes (Figure 3.9b). Both datasets contain soil water retention data in 7 predefined matric potentials (6, 10, 33.3, 100, 200, 500, and 1500 kPa). Datasets 1 and 2 have

12,799 and 119,112 data points, respectively.



(a) (b)

Figure 3.9: Soil texture ternary diagrams of datasets used for (a) Dataset C1, and (b) Dataset C2. Based on USDA taxonomy, soil particles are classified as clay (< 0.002 mm), silt ($0.002 - 0.05$ mm), and sand (> 0.05 mm), depending on their particle diameters.

3.5.2 Development of PTF models for Clay WRC estimation

As shown in Table 3.2, data in the two datasets is unevenly distributed across the 7 predefined potentials, with the majority of data points concentrated at 200 and 1500 kPa. Specifically, in the clay-focussed Dataset C1, measurements at 1500 kPa account for 70.5% of the data, while in Dataset C2, 70.2% of the data is recorded at the same potential. Such an imbalance may significantly influence the performance of ANN-based PTFs, as dominant data points tend to bias model training and limit generalisation [261].

Table 3.2: Number of available data points for each matric potential of both datasets

Matric Potential (kPa)	6	10	33.3	100	200	500	1500
Dataset C1	72	126	469	198	2822	87	9025
Dataset C2	1384	2254	3860	3188	23457	1337	83632

To address this issue, point PTFs are developed in this study, where separate ANN models are trained for each matric potential. Two PTF systems are con-

structed: PTF-C1, which is trained using clayey soils (Dataset C1), and PTF-C2, which is trained using samples across all soil texture classes (Dataset C2). The architecture diagram of PTF-C1 is shown in Figure 3.10, and PTF-C2 has the same architecture. Each PTF consists of seven independent ANN sub-models corresponding to the seven matric potentials, denoted as PTF-C1a to PTF-C1g for PTF-C1, and PTF-C2a to PTF-C2g for PTF-C2. For example, PTF-C1a predicts the soil water content at 6 kPa matric potential.

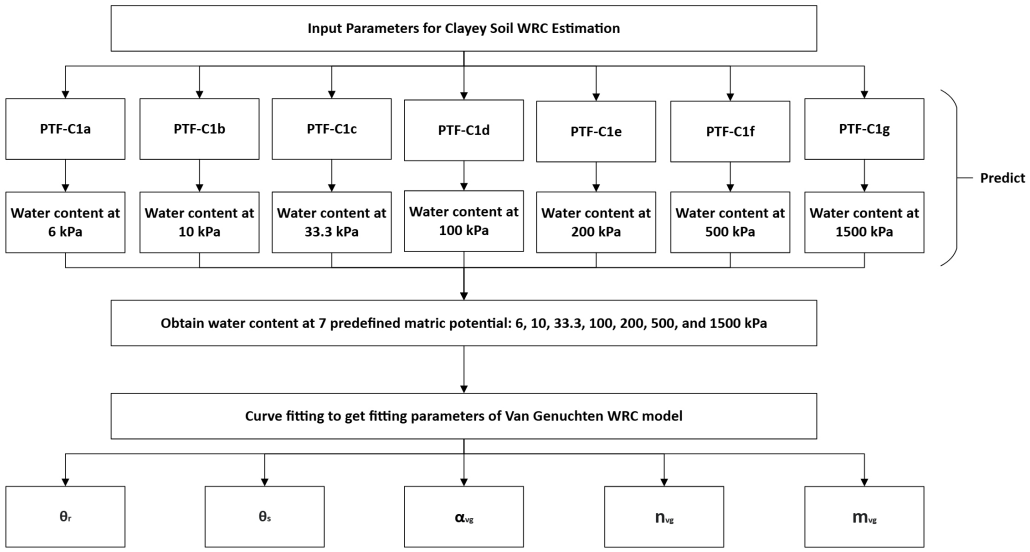


Figure 3.10: The architecture diagram of PTF-C1

All sub-models adopt a feedforward neural network architecture with two hidden layers to capture the nonlinear relationships between soil properties and water retention. The selection of two hidden layers enhances the model's capability to approximate the complex nature of soil WRC. Due to the diversity of the datasets and in order to maximize the number of available training samples, a simple input parameter combination is used. The inputs are bulk density and soil texture (percentage of clay, silt, and sand). Leaky ReLU is used as the activation function for all hidden layers based on the activation analysis in previous general PTF development.

After the model is trained, each PTF is capable of predicting soil water content at seven predefined matric potentials. The predicted data points are subsequently

fitted to the Van Genuchten equation (Eq. 3.5) using the Levenberg–Marquardt curve-fitting algorithm, allowing estimation of the VG WRC model parameters. This parametric representation is more convenient because it makes it easier to integrate the WRC information into soil hydraulic modelling applications, without affecting the prediction outcomes of the WRC model.

Nine independent clayey soil samples are used to evaluate the performance of the two PTF applications developed in this section. Table 3.3 describes the soil properties of these nine samples. None of the samples used for evaluation are included in the training dataset, so all PTF applications are unfamiliar with these new samples.

Table 3.3: Properties of clay samples used for performance evaluation of PTFs

Sample	Clay (%)	Silt (%)	Sand (%)	Bulk Density (g/cm^3)
1	56.7	27.0	16.3	1.36
2	56.1	39.9	4.0	1.32
3	58.6	32.0	9.4	1.34
4	43.9	34.8	21.3	1.08
5	64.5	31.9	3.6	1.30
6	55.6	37.0	7.4	1.43
7	55.6	39.9	4.5	1.52
8	70.2	28.0	1.8	0.96
9	67.2	23.1	9.7	1.19

The predicted results for both PTFs are directly expressed as water contents at seven predefined matric potentials (6, 10, 33.3, 100, 200, 500, and 1500 kPa). Therefore, the model performance can be evaluated by directly comparing the predicted water contents with the experimentally measured values at the corresponding matric potentials. The following error metrics are used: mean absolute error (MAE), RMSE (Eq. 3.11), and % Error. MAE has the following formula:

$$\text{MAE} = \frac{\sum_{i=1}^{n_a} |p_i - a_i|}{n_a} \tag{3.12}$$

The % Error is expressed as:

$$\%\ \text{Error} = \frac{\sum_{i=1}^{n_a} |p_i - a_i|}{\sum_{i=1}^{n_a} |a_i|} \times 100 \tag{3.13}$$

In addition to predicting point-wise water contents, the PTFs can also estimate the fitting parameters of the VG WRC model. These predicted parameters can be used to build the entire WRC curve. Therefore, model performance can also be evaluated by comparing the predicted WRC curves with those derived from experimental data, providing a parametric perspective on accuracy. The predicted WRC curves are evaluated using the same overall error metrics: MAE (Eq. 3.12), RMSE (Eq. 3.11), and %Error (Eq.3.13).

3.6 Physics-informed neural network

3.6.1 PINN Algorithm

A typical PINN network consists of three building blocks: a neural network, a physics-informed network, and a loss function (Figure 3.11). The neural network on the left-hand side of Figure 3.11 is used to predict the value of the state variable $u(x,t)$ in the PDE. Then, the predicted $u(x,t)$ is differentiated by an algorithmic differentiation technique as shown on the right-hand side of Figure 3.11 to estimate $u_t = \frac{\partial u}{\partial t}$, $u_x = \frac{\partial u}{\partial x}$, and $u_{xx} = \frac{\partial^2 u}{\partial x^2}$ [262]. In a conventional soft-constraint approach, a loss function L_{total} is calculated next to assess the extent to which the current solution satisfies the differential equation and initial and boundary conditions:

$$L_{\text{total}} = w_I L_I + w_B L_B + w_R L_R, \tag{3.14}$$

where L_I is the loss term for the initial condition, L_B is the loss term for the boundary conditions, L_R is the loss term for the PDE constraint, w_I, w_B and w_R are weights in the loss function. The mean squared error (MSE) is commonly selected as a loss function. Then, the total loss L_{Total} is used as a penalty term to optimise the neural network. Designating the set of weights and biases in each neuron of the neural network as η, the formula for the η optimisation can be written as:

$$\hat{\eta} = \arg\min \left[L_{\text{total}}(\eta) \right], \tag{3.15}$$

where $\hat{\eta}$ is the optimised parameter collection of the neural network. After η is optimised, the trained network can simulate the complex physics that the PDE represents.

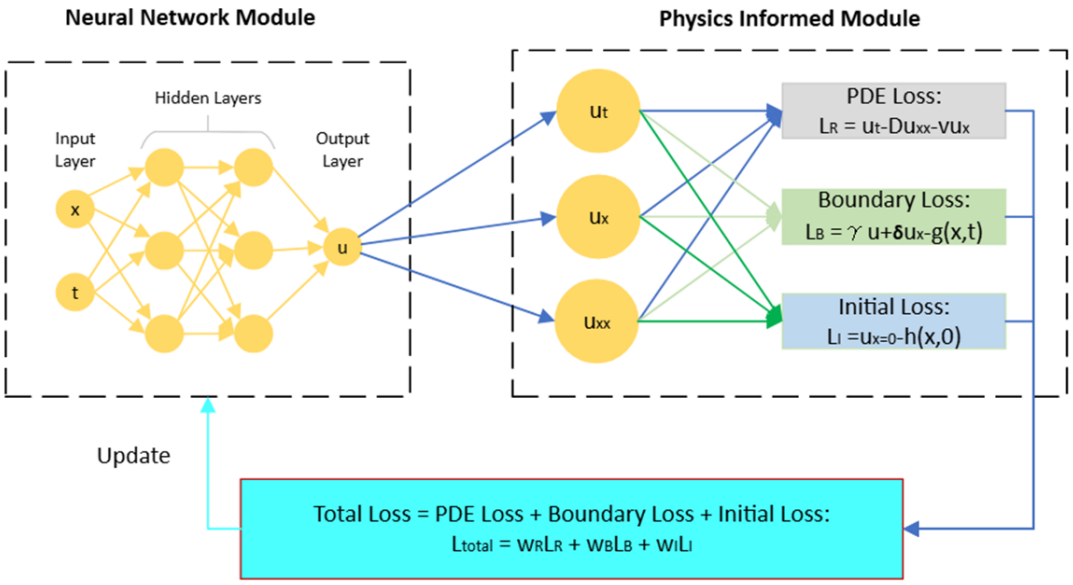


Figure 3.11: Structure of a typical, soft-constraint PINN network

3.6.2 Soft and Hard Constraints and Distance Functions

As mentioned earlier, during the training process of a traditional PINN, the BC and IC of the problem are not strictly enforced but rather included in the loss function as terms to be minimised, hence treated as soft constraints. However, previous studies have shown that the contribution and significance of each residual term can become skewed during the optimisation process, with the loss associated with some components remaining high [207]. Adjusting the weights of the three components to balance their influence can be difficult, as there is no universal rule for selecting optimal weights to achieve desired accuracy and convergence [207, 263, 264]. Instead, an empirical trial-and-error process is followed, which is highly specific to each case and problem setup.

To address this problem, an alternative approach has been proposed in the literature which consists of strictly enforcing the BC and/or IC, and the method is referred to as hard-constraint (HC) PINN [208, 265]. The implementation of an HC PINN is illustrated in Figure 3.12. An additional HC module is inserted between the neural network and physics-informed modules that ensures the constraints are enforced.

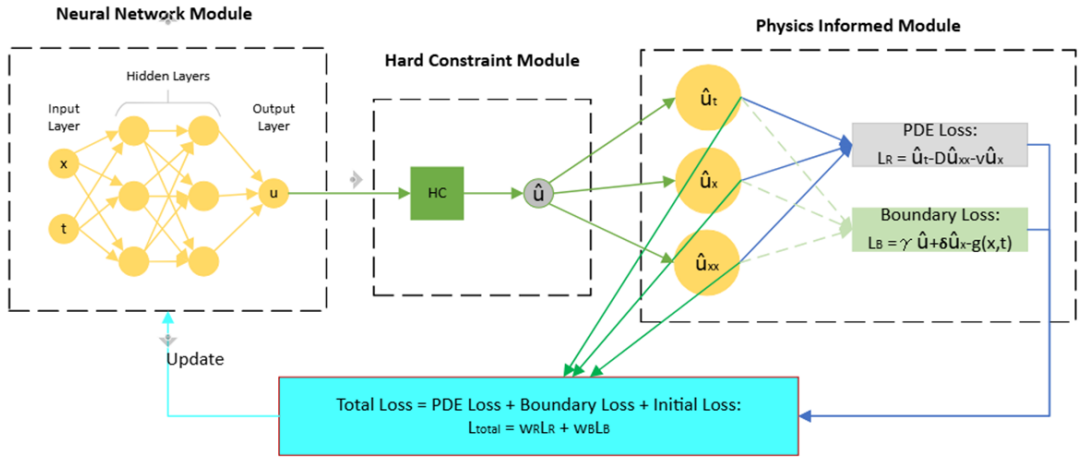


Figure 3.12: The structure of PINN with hard constraint

Mathematically, the transformation in HC layer can be expressed as follows. For an IC:

$$\hat{u}(x,t) = u^{\partial t}(x,0) + l(t)NN(x,t;\eta) \tag{3.16}$$

And, for a Dirichlet-type BC:

$$\hat{u}(x,t) = u^{\partial\Omega}(x,t) + l^{\partial\Omega}(x)NN(x,t;\eta) \tag{3.17}$$

where $u^{\partial t}(x,0)$ is the user-specified value of the IC at any coordinate $(x,0)$, $u^{\partial\Omega}(x,t)$ is the user-specified BC, $l(t)$ and $l^{\partial\Omega}(x)$ are the distance functions for the IC and BC, respectively. The distance function is an important component of an HC framework. Its primary role is to ensure that the solution at initial and boundary points remains unaffected by the neural network's output, enforcing exact solutions at these locations. This function also allows the neural network

to represent the results at all other points while maintaining the specified values at the initial and boundary conditions. The most stringent distance function that meets these two requirements [195] is:

$$l(t) = \begin{cases} 0 & \text{if } t = 0 \\ 1 & \text{otherwise} \end{cases} \tag{3.18}$$

$$l^{\partial\alpha}(x) = \begin{cases} 0 & \text{if } x \in \partial\Omega \\ 1 & \text{otherwise} \end{cases} \tag{3.19}$$

However, Eq. 3.18 and Eq. 3.19 cannot be used as distance functions in practice because they are not differentiable everywhere, a key requirement across the entire neural network to allow optimisation. So far, to the best of the authors' knowledge, only linear distance functions have been used in HC-PINN implementation reported in the literature. The linear distance function for initial conditions can be defined as:

$$l(t) = \frac{t - t_{\min}}{t_{\max} - t_{\min}} \tag{3.20}$$

where $t_{\min}$ [T] and $t_{\max}$ [T] are the minimum and maximum time, respectively. For Dirichlet-type BC:

$$l^{\partial\Omega}(x) = \left(\frac{x-a}{b-a}\right)^{i_a} \left(\frac{b-x}{b-a}\right)^{i_b} \tag{3.21}$$

where $\Omega = [a, b]$ with a [L] and b [L] as the coordinates of the points at which the two BCs are applied, i_a (i_b) is a switch set to 1 if a Dirichlet BC applies at boundary $x = a$ ($x = b$) and zero otherwise. In the case where $i_a = i_b = 0$, Eq. 14 yields 1 but this is irrelevant since, in this case, no Dirichlet BC applies at either end, and the hard constraint equation no longer applies.

A linear distance function seems the simplest form complying with differentiability requirements, but other functions are possible. While this has been

recognised by some researchers (e.g., [195]), no attempt to date has been made to implement non-linear functions [211, 212, 265]. In this study, two other forms are tested, logarithmic and exponential, in order to assess their effects on the accuracy of HC-PINN predictions. The general forms of the equations considered here are given in Table 3.4. The exponential distance function is especially pertinent for ADE problems, given that analytical solutions of simple forms of the equation are exponential in nature.

Table 3.4: Distance functions investigated in this study ($\alpha_t > 0$ and $\alpha_x > 0$ are shape parameters for IC and BC distance functions, respectively; i_a/i_b are 1 or 0 depending on whether the BC at points a/b is Dirichlet-type or not).

Name	**Equation $l(t)$ or $l^{\partial\alpha}(x)$**
Initial Conditions Distance Function $l(t)$	Linear: $\dfrac{t - t_{\min}}{t_{\max} - t_{\min}}$
	Logarithmic: $\log_{a_t}\left(1 + \dfrac{t - t_{\min}}{t_{\max} - t_{\min}}\right)$
	Exponential: $1 - e^{-a_t\left(\frac{t - t_{\min}}{t_{\max} - t_{\min}}\right)}$
Boundary Conditions Distance Function $l^{\partial\alpha}(x)$	Linear: $\left(\dfrac{x-a}{b-a}\right)^{i_a}\left(\dfrac{b-x}{b-a}\right)^{i_b}$
	Logarithmic: $\left[\log_{a_x}\left(\dfrac{x-a}{b-a}+1\right)\right]^{i_a}\left[\log_{a_x}\left(\dfrac{b-x}{b-a}+1\right)\right]^{i_b}$
	Exponential: $\left(1 - e^{-a_x\frac{x-a}{b-a}}\right)^{i_a}\left(1 - e^{-a_x\frac{b-x}{b-a}}\right)^{i_b}$

Neumann-type boundary conditions are more difficult to impose as HC because they are a function of a gradient term at the boundary, a variable not directly generated by the neural network. Hence, in this study, Dirichlet-type BCs are imposed as hard constraints, while Neumann-type BCs are included in the loss function as soft constraints.

3.6.3 Sampling Mechanisms

In a data-driven machine learning method, the quality of data plays an important role in the training performance as discussed in section 3.4.1 and 3.5.1. Although PINNs are not purely data-driven methods, their training quality also depends on how effectively the underlying physical laws are enforced across the domain. To ensure accurate enforcement of these laws, sampling of collocation points becomes a critical factor contributing to the success of PINN-based approaches [205, 266]. Since the introduction of PINNs, sampling strategies have received significant attention from researchers due to their strong influence on model performance. Broadly, these methods can be classified into two main categories: fixed sampling approaches, where collocation points remain unchanged during training, and re-sampling (including adaptive resampling) approaches, where points are updated after certain training intervals. These two categories are discussed next.

Fixed Sampling

The fixed sampling strategy involves generating a set of collocation points prior to training, which remain unchanged throughout the entire training process. This approach is widely adopted in the PINN research community due to its simplicity and ease of implementation. The two most commonly used methods for fixed sampling are equi-spaced uniform grids and uniformly random sampling [2, 239, 267, 268, 269]. Their definitions are as follows:

- **Equi-spaced Uniform Grid**: The domain is divided into evenly spaced intervals along both the spatial and temporal dimensions, forming a structured grid. The sampled collocation points are located at the intersection points of this grid. If the domain is one-dimensional, the collocation points are equally spaced points between the boundaries.

- **Uniformly Random Sampling**: Collocation points are generated by drawing samples from a continuous uniform distribution over the domain,

meaning that every location within the domain has an equal probability of being selected. For example, in `PyTorch` (a machine learning library in python) [270], to generate N_c collocation points over a 1D domain $[a, b]$, the following expression can be used:

$$x_c = a + (b - a) \times \texttt{torch.rand}(N_c, 1) \tag{3.22}$$

x_c is the collocation points collection, $\texttt{torch.rand}(N_c, 1)$ generates a tensor of shape $(N_c, 1)$ containing N_c random values sampled from the range $[0, 1]$. This function produces N_c points that are independently distributed over the interval $[a, b]$.

In this thesis, uniformly random sampling is selected as the primary strategy for generating collocation points. This method is widely adopted as the standard choice in the PINN literature [271, 272, 146], and was originally introduced by Raissi et al. [2, 11] in their papers that proposed PINN framework. Since then, it has been adopted in the majority of subsequent studies due to its simplicity, flexibility, and effectiveness in capturing general solution features across a variety of problems. Equi-spaced uniform grids have also been explored in some later works and have shown good results in specific cases [269, 273, 274]. However, many studies have pointed out that this sampling method is only suitable for those cases with smooth solutions. It performs poorly when the target solution exhibits sharp gradients in their solution [239, 274]. Other sampling techniques, such as Latin Hypercube Sampling (LHS) [264], have also been introduced to improve coverage of the domain. However, these methods are less commonly used in the PINN literature, and current empirical results have not demonstrated significant performance advantages over more conventional strategies such as uniformly random sampling [239]. For these reasons, this thesis adopts uniformly random sampling as the baseline sampling strategy.

Adaptive Resampling

Adaptive resampling uses the residual information during training to identify regions with higher errors. At certain training intervals, a fraction of the collocation points is replaced, allowing the sampling distribution to gradually shift towards areas with larger errors. This enables the model to allocate more computational resources to better approximate these challenging regions.

Two adaptive resampling methods are tested in this study. The first method explored is the residual-based adaptive refinement (RAR) method [237, 145, 275]. RAR aims to increase the collocation points coverage in regions with large PDE residuals. While many RAR implementations achieve this by continuously adding new points to high-residual regions, the method proposed by Zapf et al. [237] maintains a fixed number of collocation points throughout training in order to control computational cost while achieving comparable accuracy. The approach adopted in this study follows the same principle, but with some modifications. After each training stage, the PDE residuals are evaluated at all current collocation points. A proportion of collocation points with the highest residuals is retained, while the remaining points are replaced by new samples drawn uniformly from the entire spatial-temporal domain. The algorithm is shown as follows:

Algorithm 1 Residual-Based Adaptive Refinement Algorithm. Adapted from the method in [237], with modifications applied in this study.

1: Uniformly sample N_c collocation points from the domain to form the initial training dataset $\mathcal{T}_{\text{initial}}$.
2: Train the PINN for a predefined number of iterations.
3: **repeat**
4: Evaluate the PDE residuals $|\mathcal{R}_i|$ at each collocation point.
5: Retain the top $N_{\text{keep}} = k_1 N_c$ points with the largest residuals to form $\mathcal{T}_{\text{keep}}$.
6: Uniformly resample $N_{\text{resample}} = (1 - k_1)N_c$ new points from the domain to form $\mathcal{T}_{\text{rand}}$.
7: Update the collocation set: $\mathcal{T}_{\text{new}} = \mathcal{T}_{\text{keep}} \cup \mathcal{T}_{\text{rand}}$.
8: Train the PINN for a predefined number of iterations.
9: **until** The maximum number of iterations is reached.

In Algorithm 1 described above, two special cases can be identified. When $k_1 = 0$, full replacement is implemented, where all collocation points are resam-

pled without considering the residual information. In this case, the residual hot spots do not influence the sampling strategy. Resampling is conducted only to update the training dataset and prevent excessive dependence on a fixed set of collocation points. The other special case is when $k_1 = 1$, in which no resampling is performed and the entire training process relies on a single, fixed set of collocation points.

The second method examined in this study is a probability-based resampling strategy, which is referred to as the residual-based adaptive distribution (RAD) method in some publications [210, 236, 238]. In this approach, the probability of sampling each point is determined based on its residual value and can be expressed as:

$$p(x,t) \propto \varepsilon^{k_2}(x,t), \tag{3.23}$$

where $\varepsilon(x,t)$ denotes the point-wise PDE residual, and k_2 is a resampling scaling factor that controls the sharpness of the probability distribution. Larger values of k_2 amplify the contribution of high-residual points, thereby increasing the density of resampled points in regions with larger residual errors. Because the resampling is based on probability, this adaptive strategy enables the model to focus more on difficult regions with higher residuals, while still maintaining a certain level of global coverage across the entire domain. In this study, the resampling procedure is performed every 5000 epochs and continues until either the maximum number of training epochs is reached or the model converges. To investigate the influence of the scaling factor on model performance, five different k_2 values (0.5, 1, 1.5, 2, 2.5) are tested. The detailed algorithm of the RAD method is presented as follows:

Algorithm 2 Residual-based Adaptive Distribution (RAD), adapted from [210, 236, 238]

1: Uniformly sample N_c collocation points from the domain to form the initial training dataset $\mathcal{T}_{\text{initial}}$.
2: Train the PINN for a predefined number of iterations.
3: **repeat**
4: Evaluate the PDE residuals $\varepsilon(x_i, t_i)$ at each collocation point.
5: Compute the sampling probability for each point according to:

$$p(x_i, t_i) = \frac{\varepsilon^{k_2}(x_i, t_i)}{\sum_{j=1}^{N_c} \varepsilon^{k_2}(x_j, t_j)}$$

where $k_2 \geq 0$ is the resampling scaling factor.
6: Resample N_c new collocation points according to the probability distribution $p(x_i, t_i)$ to form $\mathcal{T}_{\text{new}}$.
7: Replace the collocation set with $\mathcal{T}_{\text{new}}$.
8: Train the PINN for a predefined number of iterations.
9: **until** The maximum number of iterations is reached.

3.7 PINN for Advection-Diffusion Equation

3.7.1 Governing Equation

A general form of the one-dimensional version of the ADE and its initial and boundary conditions can be written as [276]:

$$u_t + N_x[u] = 0, \quad x \in \Omega,\ t \in [0, t_{\text{lim}}], \tag{3.24}$$

$$N_x[u] = -\nabla \cdot (D\nabla u) + \nabla \cdot (vu), \tag{3.25}$$

$$f_x[u] = -D\nabla u + vu, \tag{3.26}$$

$$u(x, 0) = h(x, 0), \quad x \in \Omega, \tag{3.27}$$

$$\gamma \cdot u(x, t) + \delta \cdot \nabla u(x, t) = g(x, t), \quad x \in \partial\Omega \ \&\ t \in [0, t_{\text{lim}}]. \tag{3.28}$$

where x [L] is a space coordinate in the interval [a,b] with $L_d = b - a$, t [T] is a temporal coordinate, Ω is a spatial domain, $\partial\Omega$ is the boundary of Ω, t_{lim} is the upper bound of time, $N_x[u]$ is the spatial differential operator, u [M.L^{-3}] is the state variable that denotes contaminant concentration in transport problems, u_t [M.L^{-3}.T^{-1}] is the first-order time derivative of u, D [L^2.T^{-1}] is the hydrodynamic dispersion coefficient, v [L.T^{-1}] is the advection coefficient, $h(x, 0)$

and $g(x,t)$ represent the initial and boundary conditions, respectively. γ and δ are coefficients that control the boundary behaviour. When $\gamma = 1$ and $\delta = 0$, the condition reduces to a Dirichlet-type boundary condition. When $\gamma = 0$ and $\delta = 1$, it becomes a Neumann-type boundary condition. The steady-state version of the ADE can be obtained by setting the first term on the left-hand side of Eq. 3.24 to zero. Note that, although the ADE carries the term "diffusion" but not "dispersion", it is understood throughout this thesis that D in the equation is the hydrodynamic dispersion coefficient which incorporates the effects of both molecular diffusion and mechanical dispersion.

3.7.2 ADE under Unsaturated Conditions

In contrast to saturated conditions that the soil pore space is filled with water, the pore space in unsaturated soil is only partially filled. ADE under unsaturated soil conditions is different from its saturated version due to the varied dynamics across the entire domain. The governing equation of solute transport by considering the varied water content in unsaturated soil is expressed as [277]:

$$\frac{\partial(\theta u)}{\partial t} = \frac{\partial}{\partial x}\left(\theta D_{\text{eff}} \frac{\partial u}{\partial x}\right) - \frac{\partial(qu)}{\partial x} \tag{3.29}$$

where D_{eff} is the effective dispersion coefficient, q is the water flux in the x direction, and the water flux information can be determined from the simulation of the water movement in unsaturated soil with RRE. The effective dispersion coefficient comprises two components: mechanical dispersion and molecular diffusion. D_{eff} can be expressed as [278, 279]:

$$D_{\text{eff}} = D_L \cdot \frac{|q|}{\theta} + D_m \cdot \tau(\theta) \tag{3.30}$$

where D_L is the longitudinal dispersivity coefficient, D_m is the molecular diffusion coefficient, and $\tau(\theta)$ is the tortuosity factor which is given as $\tau(\theta) = \frac{\theta^{10/3}}{\theta_s^2}$ [279]. In this thesis, the molecular diffusion coefficient D_m is assumed to be

neglected because the focus of this study is on solute transport processes primarily contributed by advection and mechanical dispersion, both of which are mainly driven by water flow.

3.7.3 Implementation and validation of PINN for ADE

The algorithms described above are implemented in Python. Early stopping is employed in training with a patience value of 20 adopted (i.e., training was stopped when no decline in loss was recorded over 20 consecutive epochs). The choice of patience value is a compromise between adequate fitting, the need to reduce the risk of overfitting (the model becoming too tailored to the training data and hence losing some generalisability) and computational economy.

For validation, both steady-state and transient ADE problems, with both Dirichlet and Neumann boundary conditions, are considered, as shown in the three case problems described in Table 3.5. In each case, predictions of the proposed PINN algorithm are compared to those of the reference solutions, either analytical closed-form solutions or solutions derived from numerical solvers, with sources shown in Table 3.5. The performance of HC-PINNs under various distance functions described above, are compared to each other and to that of traditional soft-constraint PINN, under low and high-Peclet number conditions. Note that the four problems shown in Table 3.5 all assume saturated conditions for simplicity. However, the unsaturated version of the ADE, with D_{eff} and v varying in time and space, will be used in solving the coupled problems, as described later in this chapter.

To evaluate the global performance of the proposed model, three commonly used error metrics are employed: MAE (Eq. 3.12), RMSE (Eq. 3.11), and the Coefficient of Determination (R^2). These metrics provide quantitative insights into the accuracy and reliability of the model's predictions. MAE offers a direct measure of the average absolute difference between predicted and actual values. RMSE incorporates squared differences, which penalise larger errors more

Table 3.5: Problems simulated with PINN

Case Number	**1**	**2a**	**2b**	**3**
Case Name	Steady-State Dirichlet-Dirichlet	Transient Dirichlet-Dirichlet	Transient Dirichlet-Dirichlet	Transient Dirichlet-Neumann
L_d [L]	1-100	1	1	1-20
BC at x=0:	Dirichlet: u(0)=0	Dirichlet: u(0)=1	Dirichlet: u(0)=1	Dirichlet: u(0)=1
BC at x=L_d:	Dirichlet: u(L_d)=1	Dirichlet: u(L_d)=0	Dirichlet: u(L_d)=0	Neumann: f_x=0
IC (t=0):	Not Relevant	u(x)=x	u(x)=0	u(x)=0
t_{max} [T]	Not Relevant	2	1	10
Hydrodynamic Dispersion Coefficient D (L^2T^{-1})	0.1	0.01	0.1, 0.01	0.01-0.05
Velocity v (LT^{-1})	0.05-1000	1	10, 5	0.01-1
P_e	0.5-10^6	100	100, 500	1-1000
Reference Solution	Analytical solution of ordinary differential equation (Appendix A)	Numerical by FEM-based SPAS	Numerical by FEM-based SPAS	Analytical solution from van Genuchten and Alves

severely. Lastly, R^2 quantifies how well the predictions approximate the actual data by comparing model variance to the total variance in the data. R^2 is shown as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n_a}(p_i - a_i)^2}{\sum_{i=1}^{n_a}(p_i - \bar{a})^2} \tag{3.31}$$

To quantify the point-wise error at random locations within the domain, two additional local metrics are used: Point-Wise Absolute Error (PWAE) and Point-Wise Absolute Percentage Error (PWAPE). These metrics provide a more detailed, location-specific evaluation of model performance. PWAE measures the absolute deviation at each point, while PWAPE expresses this deviation as a percentage of the actual value, allowing for better interpretability and comparison across scales. The equations for these metrics are:

$$\text{PWAE}_i = |p_i - a_i| \tag{3.32}$$

$$\text{PWAPE}_i = \frac{|p_i - a_i|}{|a_i|} \tag{3.33}$$

3.8 PINN for Richard's Equation

3.8.1 Governing Equation

The Richards Equation is derived from the continuity equation and Darcy's law. The mass balance of water in a controlled volume of soil can be written as [225]:

$$\frac{\partial \theta}{\partial t} = -\nabla \cdot \mathbf{q} \tag{3.34}$$

where θ $[L^3L^{-3}]$ is the volumetric water content of the soil, t [T] is the time variable, and $\mathbf{q}$ $[LT^{-1}]$ is the water flux. The water flux term in Eq. 3.34 can be derived from Darcy's law under unsaturated conditions, and in this case, soil hydraulic conductivity is dependent on the water content in the soil. The water flux is expressed as:

$$\mathbf{q} = -K(\theta) \cdot \nabla H \tag{3.35}$$

where $K(\theta)$ $[LT^{-1}]$ is the unsaturated hydraulic conductivity as a function of the volumetric water content of the soil, and H [L] is the total pressure head. Using the relationship in Eq. (3.35) to replace the water flux term in Eq. 3.34, the Richards Equation can be written as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x}\left[K(\theta)\left(\frac{\partial h(\theta)}{\partial x} + 1\right)\right] \tag{3.36}$$

where x [L] is the soil depth, and $h(\theta)$ [L] is the soil water pressure head, which is dependent on the volumetric water content of the soil. As shown in Eq. 3.36, the solution of the Richards Equation requires the specification of the WRC ($h(\theta)$) and HCF ($K(\theta)$). In this study, the widely used Van Genuchten WRC model and the Van Genuchten-Mualem HCF model are implemented (see Eq. 3.5 and Eq. 3.6).

3.8.2 Implementation of PINN solution for RRE

The overall structure of the PINN algorithm is illustrated in Figure 3.13. A hard constraint module is incorporated between the neural network and the PDE loss function, transforming the conventional soft-constraint framework into a hard-constraint formulation. To investigate potential improvements over the baseline model, three adaptive resampling strategies discussed in section 3.6.3 are implemented and evaluated. The performance of each PINN variant developed in this study is assessed in comparison with the conventional soft-constraint PINN.

In the hard-constraint formulation, the feedforward neural network outputs the pressure head $h(\theta)$. The boundary and initial condition residuals are exactly satisfied, so the total loss consists only of the PDE residual. In RRE, the distance function used is the linear one for the spatial domain, and an exponential one with a shape factor $\alpha_t = 100$ is used for the temporal domain. Since Eq. 3.36 depends on θ and $K(\theta)$, the network output $h(\theta)$ is used to compute these variables via the Van Genuchten WRC and Van Genuchten-Mualem HCF models, defined in Eq. 3.5 and Eq. 3.6, respectively. The corresponding derivatives are then calculated to evaluate the PDE residual, given by:

$$L_{\text{PDE}} = \frac{\partial \theta}{\partial t} - \frac{\partial}{\partial x}\left[K(\theta)\left(\frac{\partial h(\theta)}{\partial x} + 1\right)\right] \tag{3.37}$$

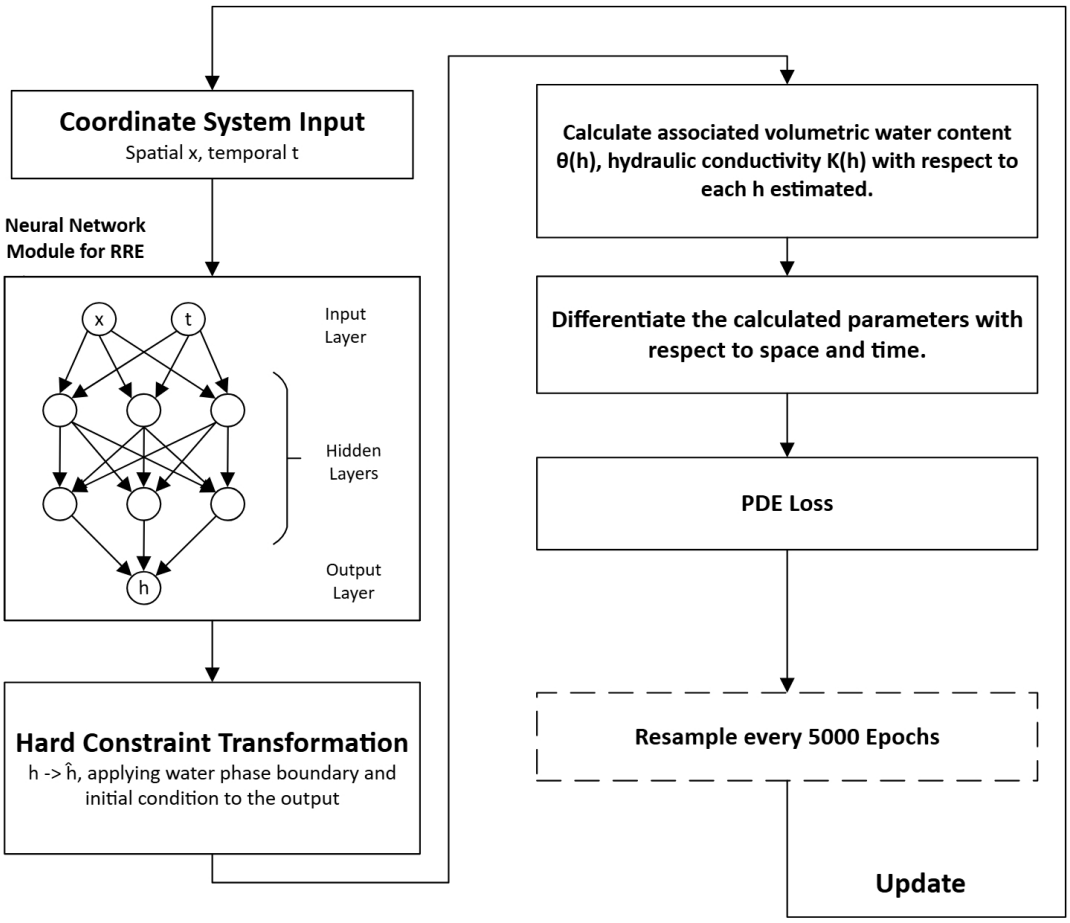


Figure 3.13: Model implementation flowchart, including hard constraint and adaptive resampling.

3.8.3 Model Evaluation

A test case study, shown in Figure 3.14 is used to assess the performance of the various PINN algorithms explored in this paper. PINN predictions are compared to a generalised analytical solution for this problem developed by Warrick et al. [280] and detailed in Appendix B.

The PINN solution and sensitivity analysis results are compared with the generalized solution provided by Warrick et al. [280], as stated in the previous section. We selected twelve time steps ranging from 0.5 to 6 days ($\Delta t = 0.5$). At each time step, predictions of the analytical solution and PINN solver are compared at three reference volumetric water content levels (0.24, 0.31, and 0.38), hence a total of 36 observation points (3 points per time step).

The absolute error (AE) at each observation point is defined as:

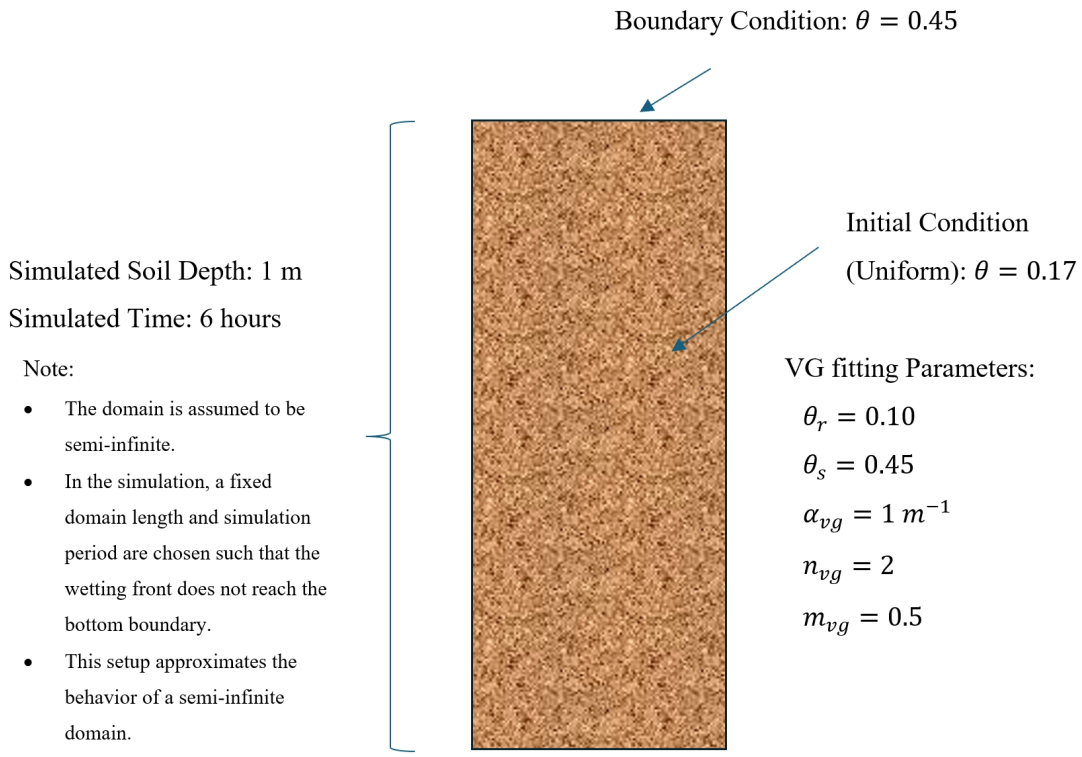


Figure 3.14: RRE test case

$$AE(\theta_{i,j}) = |x_{\text{pred}}(\theta_{i,j}) - x_{\text{true}}(\theta_{i,j})| \tag{3.38}$$

where i is the time step number ($1 \sim 12$), j is the observation index within each time step ($1 \sim 3$), $\theta_{i,j}$ is the reference water content at observation point (i, j), x_{pred} is the depth predicted by the PINN model, and x_{true} is the target value from the analytical solution. To evaluate the overall model performance, a global %Error (Eq. 3.13) is used.

However, a limitation of the analytical solution incorporated in this study is that it cannot capture the sharp gradient near the wetting front. To address this limitation, the PINN predictions are also compared against a FEM solution obtained using HYDRUS software [281]. The FEM solution is first evaluated at the same observation points used in the analytical comparison to verify its consistency with the analytical results under a conservatively refined FEM mesh. Once a strong agreement is achieved between the FEM and analytical values at these

reference points, the FEM solution is used as an additional benchmark to assess the performance of the PINN model near the wetting front. This comparison helps to determine whether the PINN model can replicate the sharp transitions at the wetting front.

3.9 Coupled Water Flow and Solute Transport by PINN

The final part of the methodology in this thesis brings together three key developments described earlier to illustrate how machine learning can be used to solve the coupled problem of water flow and solute transport in unsaturated soils.

Parameters of the WRC and HCF are first predicted by the neural network PTF from soil data input. Next, the WRC parameters are provided as part of the input to the RRE PINN solver, alongside geometry, saturated hydraulic conductivity, saturated and residual water content, and initial and boundary conditions to the water flow problem. Finally, the distribution in time and space of the seepage velocities and water content is used as part of the input to the unsaturated version of the ADE (Eq. 3.29) to solve the solute transport problem. Additional inputs include the diffusion coefficient, dispersivity (both dependent on water content), and the saturated volumetric water content. Figure 3.15 describes the combination of machine learning components solving the problem. All of the components shown in Figure 3.15 have already been described earlier in this chapter. However, the PINN solver of the ADE has been presented in a saturated soil context. Modification of the ADE hence needs to be made to cater for unsaturated problems as described in Section 3.7.2.

The individual components of the coupled system (Figure 3.15) have been systematically evaluated from Chapters 4 to 7. For the coupled analysis, the focus is to assessing the performance of the integrated system as a whole. To this end, results predicted by the coupled system are compared against those generated by using the numerical solver Hydrus-1D. This comparison serves to evaluate the level of agreement between the machine learning approach and traditional FEM

numerical modelling techniques.

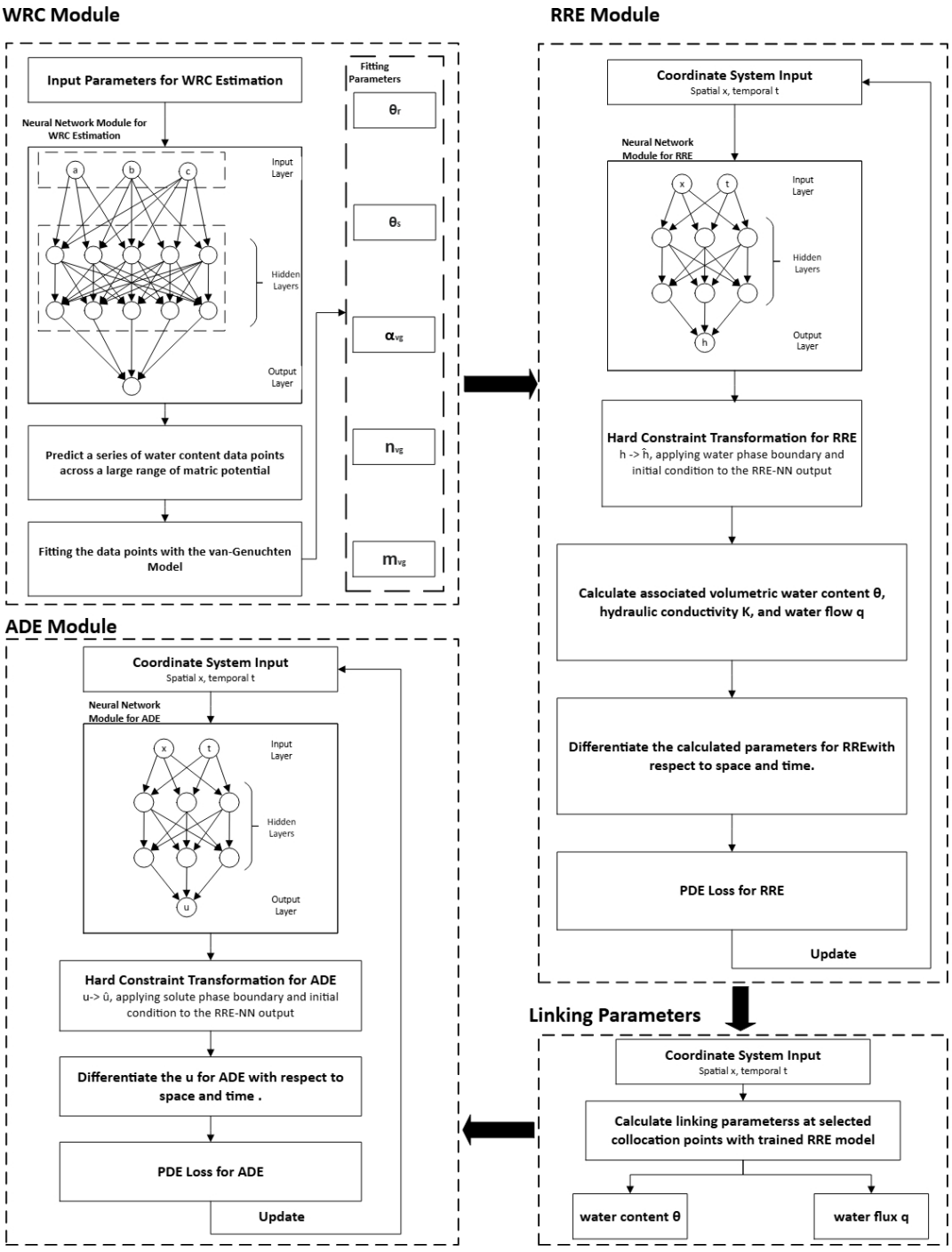


Figure 3.15: The Flowchart that illustrates the one-way coupling mechanism between RRE and the unsaturated ADE. The chart also incorporates a WRC-PTF model to allow users to predict WRC when needed.

Chapter 4

Water Retention Curves Prediction for all soil types

4.1 Overview

This chapter presents the evaluation of the newly developed PTF-based neural network models for predicting WRC using both SNN and DNN. The models are trained and tested using datasets from UNSODA 2.0 and include various combinations of easy-to-measure soil properties as input variables. The performance of the models is systematically assessed across a wide range of model architectures and input combinations, addressing the limitations of previous PTF studies.

The results are organised as follows: first, the correlation between various input parameters and WRC outputs is examined. Next, the performances of different activation functions are compared. The study then evaluates and compares the predictive capabilities of both SNN and DNN structures across 13 different models (M1 to M13). Subsequently, the influence of individual input variables such as dry density, porosity, specific gravity, depth, and organic content on model accuracy is systematically analysed. Finally, the performance of the best models is benchmarked against existing PTF-ANN models reported in the literature.

4.2 Binary Correlations between Input and Output Variables

Table 4.1 shows the correlation matrix r between all considered input variables and output data (volumetric water content) in the dataset used in this study. It is widely accepted that an r value exceeding 0.7 signifies a very strong correlation, while a value between 0.4 and 0.7 reflects a moderate to strong correlation ranging

from strong to moderate. The presence of a plus (+) or minus (–) sign indicates whether the correlation is positive or negative, respectively. As expected, the table shows that soil texture data (% sand, % silt, % clay) and dry density correlate more strongly with the output than other input variables. The soil sample depth correlates least with volumetric water content ($r = -0.041$).

Note that the presence of multicollinearity in the data (especially caused by the presence of three texture variables) has already been discussed in section 3.4 and the rationale for keeping all three texture variables presented.

Table 4.1: Correlation matrix for input variables and output data; the output data is volumetric water content.

	Clay%	Silt%	Sand%	DD	PO	SG	Depth	OM	Suction	Output θ
Clay%	1									
Silt%	0.367	1								
Sand%	-0.756	-0.886	1							
DD	-0.354	-0.424	0.479	1						
PO	-0.032	-0.056	0.060	0.445	1					
SG	0.237	0.335	-0.379	-0.868	-0.254	1				
Depth	-0.053	-0.090	0.092	0.261	0.384	-0.362	1			
OM	0.496	0.215	-0.437	-0.656	-0.574	0.532	-0.381	1		
Suction	0.095	0.206	-0.192	-0.097	-0.045	0.056	0.025	0.017	1	
Output θ	0.562	0.449	-0.596	-0.447	-0.123	0.358	-0.041	0.363	-0.210	1

4.3 Performance of Activation Functions

Table 4.2 summarises the performance of four candidate activation functions considered in this study. ReLU and Leaky ReLU demonstrated a reduced RMSE compared to Sigmoid and Tahn in predicting WRC. The model built with Leaky ReLU, as a modification of ReLU, has a lower RMSE value of 0.042 m^3m^{-3} compared with the model built by ReLU (0.051 m^3m^{-3}). This finding agrees with results reported by Xu et al. [282] and Dubey and Jain [283]. In the rest of this chapter, Leaky ReLU activation function with a carefully tuned slope value is

selected to develop the SNN-PTF and DNN-PTF models.

Table 4.2: Evaluation of the suitability of activation functions for WRC estimation by using SNN-PTF model. The input parameters considered here are those for M2, i.e., soil texture data (%sand, %silt, %clay) + dry density.

Activation Function	**Testing**	
	MSE	**RMSE**
Sigmoid	0.0041	0.064
Tahn	0.0042	0.065
ReLU	0.0026	0.051
Leaky ReLU	0.0018	0.042

4.4 Overall Models Performance

The relationships between predicted volumetric water content (m^3m^{-3}) and target volumetric water content (m^3m^{-3}) of the training dataset are shown in the form of scatter plots (Figure 4.1) and correlation coefficients (Table 4.3). Most models have an r value (Eq. 3.10) greater than 0.9 except SNN-PTF-M1, M3 and DNN-PTF-M1, and these three exceptions still have r values greater than 0.86. Scatter plots in Figure 4.1 illustrate the good agreement between prediction and measurement.

Table 4.3: Binary correlation coefficient (r) between predicted volumetric water content (m^3m^{-3}) and target volumetric water content (m^3m^{-3}) evaluated by the testing dataset for both SNN-PTF and DNN-PTF at all 13 models.

Inputs	**SNN-PTF**	**DNN-PTF**	**Inputs**	**SNN-PTF**	**DNN-PTF**
M1	0.869	0.884	M8	0.926	0.954
M2	0.954	0.925	M9	0.952	0.944
M3	0.869	0.941	M10	0.940	0.936
M4	0.945	0.918	M11	0.942	0.941
M5	0.952	0.955	M12	0.951	0.961
M6	0.952	0.936	M13	0.910	0.933
M7	0.945	0.916			

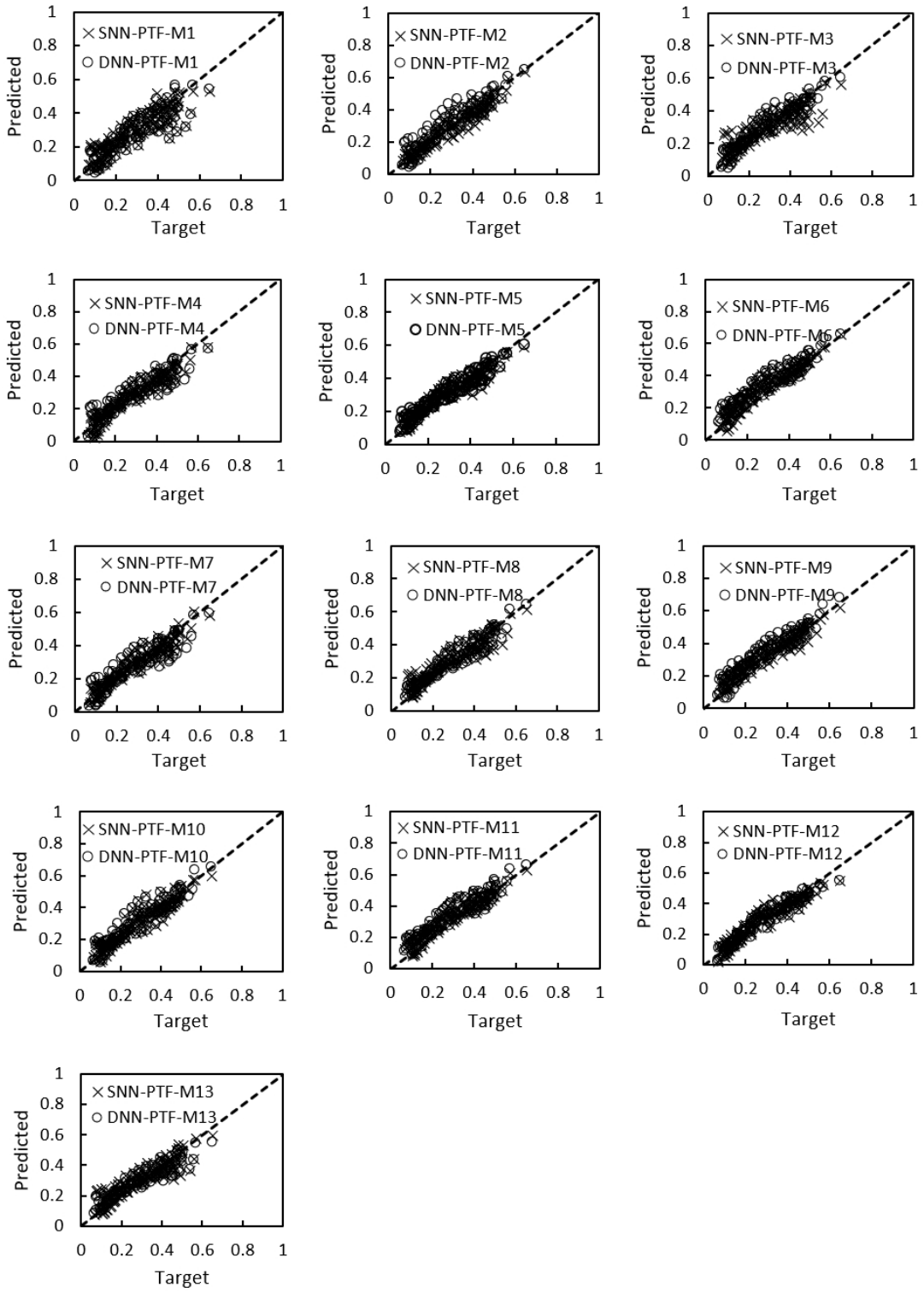


Figure 4.1: Scatter Plots of the testing dataset's predicted volumetric water content (m^3m^{-3}) against target volumetric water content (m^3m^{-3}). Legends in each subfigure indicate the scatter plot of SNN and DNN of models M1 to M13.

4.5 Deep versus Shallow Neural Networks

Table 4.4 shows the performance evaluation of SNN-PTF and DNN-PTF for 13 models from M1 to M13. Taking the RMSE of testing dataset as a criterion of accuracy, Table 4.3 shows that DNN-PTF has lower RMSE (Eq. 3.11) values than SNN-PTF by more than 24% in four models (M3, M8, M12, M13) and by less than 4% in another two (M1 and M5). The largest improvement is obtained for M12, when texture and porosity are adopted as input variables, with a 32% reduction in RMSE of the testing dataset when shifting from SNN to DNN. On the other hand, SNN-PTF performs better in seven out of thirteen models. However, only in three of these seven models, does SNN-PTF achieve an RMSE that is more than 20% lower than its corresponding value for DNN-PTF. Hence, whether improvement in accuracy can be achieved by shifting from SNN to DNN depends on the specific set of input variables chosen, even when using the same dataset.

4.6 Influence of Input Variables

Table 4.4 and Table 4.5 provide insights into the effect on prediction accuracy of introducing different input variables by comparing models M1 to M13. The base case is model M1 which includes only three texture parameters as input, with different input variables added in different models from M2 to M13. Table 4.5 lists 17 pairs of models for which the only difference is the addition of a single variable. For example, the only difference between M1 and M2 is the addition of dry density to the latter; hence, a comparison of the performances of these two models provides some evidence for the importance of this particular input variable in predicting WRC.

Comparing models M1 and M2, the RMSE value of SNN-PTF drops from 0.069 to 0.042 m^3m^{-3} (38.5% Improvement), and the value for DNN-PTF drops from 0.066 to 0.054 m^3m^{-3} (19.2% Improvement) when evaluated for the testing dataset. In fact, model M2 performs significantly better than M1, in training

Table 4.4: Root mean squared error (RMSE: m^3m^{-3}) for SNN-PTF and DNN-PTF for training and testing datasets.

Model	RMSE of SNN-PTF (E_{SNN})		RMSE of DNN-PTF (E_{DNN})		$(E_{DNN}-E_{SNN})/E_{SNN}\times100\%$	
	Training	Testing	Training	Testing	Training	Testing
M1	0.064	0.069	0.064	0.066	0.8%	-3.5%
M2	0.054	0.042	0.052	0.054	-3.1%	26.9%
M3	0.073	0.070	0.056	0.048	-23.9%	-31.3%
M4	0.060	0.047	0.069	0.057	15.2%	21.3%
M5	0.056	0.046	0.062	0.045	11.2%	-1.7%
M6	0.039	0.057	0.044	0.063	12.8%	10.5%
M7	0.060	0.046	0.068	0.059	14.9%	28.1%
M8	0.061	0.056	0.059	0.042	-2.8%	-24.3%
M9	0.035	0.051	0.041	0.057	14.6%	11.4%
M10	0.049	0.049	0.049	0.050	-0.4%	2.0%
M11	0.041	0.057	0.040	0.059	-3.0%	2.6%
M12	0.055	0.057	0.043	0.039	-20.9%	-31.6%
M13	0.068	0.072	0.059	0.054	-13.0%	-24.7%

and testing, in both SNN and DNN, hence highlighting the importance of dry density as a predictor of water retention behaviour. This trend is also evidenced by comparing models M13 and M5, whereby the inclusion of dry density leads to a 36% and 16.5% performance improvement for testing datasets of SNN and DNN, respectively. This finding agrees with those reported by Schaap and Bouten [148], and Rastgou et al. [155]. Haghverdi et al. [164], on the other hand, did not find a significant advantage in using dry density. However, in a later study, also led by Haghverdi, but using a different dataset, the introduction of dry density was found to lead to a significant improvement in accuracy [258].

Table 4.4 and Table 4.5 show that the picture is more mixed when dry density is added to models that include porosity as part of the input variables (M12 to M3 and M13 to M5). Improvements are found when moving from M13 to

M5, for both testing and training in SNN, but only for the testing dataset for DNN. On the other hand, moving from M12 to M3 increases the RMSE of SNN and DNN, for both training and testing, i.e., lowering prediction accuracy. The strong predictive power of dry density and porosity in relation to WRC is consistent with fundamental mechanisms. Changes in dry density or porosity are indicators of soil structural change [284, 285, 286, 287, 288]. From a microscopic perspective, matric potential arises from the tensile capillary forces in the liquid bridges between soil particles. Soil densification process shortens the liquid bridges and leads to an increase of the capillary force [289]. Therefore, if the soil is compacted, the same level of matric potential can be obtained with fewer liquid bridges. At the macroscopic scale, soil densification reduces the total pore volume, and the relative fraction of small pores becomes larger in the pore-size distributions [290]. This macro-scale change typically leads to a decrease in soil water content at any given matric potential [291, 292].

Conversely, Table 4.5 also shows that the addition of porosity to the list of input variables yields a significant improvement in prediction accuracy in both SNN and DNN, for training and testing datasets, provided dry density is not included in the input variables (M1, M12 and M13). However, when dry density is included, adding porosity does not generally yield improvements (M2 to M3, M8 to M5, M7 to M4 and M9 to M6). In fact, Table 4.4 shows that the best-performing shallow and deep PTFs are models SNN-M2 and DNN-M12 with RMSEs of 0.042 m^3m^{-3} (SNN-M2) and 0.039 m^3m^{-3} (DNN-M12). Models M2 and M12 differ from the base model M1 (3 texture variables as the only input, in addition to suction) by adding one input variable, namely dry density for M2 and porosity for M12.

Concerning specific gravity, Table 4.5 shows no clear pattern concerning the benefits, or lack thereof, of adding it to the list of input variables. Depth, on the other hand, seems to add to the accuracy of prediction in DNN-PTF but not SNN-PTF. This may be due to the ability of DNN to capture more complex

data relationships, compared to SNN. Soil sample's depth can affect the soil's overburden stress, leading to different porosities which may affect the shape of the WRC when expressed in terms of volumetric water content [293]. However, in the datasets under consideration, depth ranges between 0 and 325cm and a stronger relationship between depth and WRC may emerge if data with a wider range of depth is considered, reflecting deeper soils with higher overburden stress.

Finally, organic content can affect soil particle size distribution, soil porosity, hydraulic conductivity, and soil strength [6]. Including organic content as an input variable in this study, seems to improve the prediction of the training dataset but has the opposite effect on the testing dataset. This may be due by data overfitting caused by the large number of samples with missing values of organic content in the overall dataset, with models containing OM (M6, M8 and M11) as an input variable, reduced to around 700 points, from over 1000 for all other models.

4.7 Comparison with PTF in the Literature

The best performing PTF in this study, based on RMSE of testing data, is DNN-M12 (RMSE=0.03914 m^3m^{-3}) which takes soil texture and porosity as input variables. SNN-M2, using dry density instead of porosity as input, is the second most accurate model, with an RMSE of 0.042 m^3m^{-3}. Table 4.6 compares the performances of these two models to those of PTF-WRC published in the literature, as evaluated by RMSE of testing data. Different studies use different databases and different sizes of training and testing datasets which limits the value of the comparison. In addition, it is not always possible to compare the sizes of the datasets as some studies only report the number of soil samples, rather than the number of data points, used in building the models. Nevertheless, Table 4.6 shows that the RMSE reported for the two best-performing models in this study compares favourably with those published in the literature.

Table 4.5: Percent difference in RMSE between pairs of models; a negative difference indicates an improvement in performance by moving from "model without" to "model with"; each row shows a pair of models in which the only difference is the addition of the variable in question (e.g., the only difference between models M4 and M5 is the addition of depth as input parameter in the latter, as is the case for M7 and M8 and for M2 and M10).

Additional	Models		SNN-PTF		DNN-PTF	
Variable	Without	With	Train	Test	Train	Test
Dry Density	M1	M2	-15.6%	-38.5%	-18.9%	-19.2%
	M12	M3	34.6%	22.5%	29.6%	23.0%
	M13	M5	-17.7%	-36.0%	5.2%	-16.5%
Depth	M4	M5	-6.5%	-1.7%	-9.8%	-20.4%
	M7	M8	3.3%	21.1%	-12.6%	-28.5%
	M2	M10	-9.0%	16.0%	-6.5%	-6.7%
Porosity	M2	M3	36.6%	66.2%	7.3%	-10.0%
	M8	M5	-8.8%	-17.9%	4.4%	6.6%
	M7	M4	0.8%	1.2%	1.1%	-4.2%
	M1	M12	-14.4%	-16.6%	-32.9%	-40.9%
	M9	M6	8.9%	12.0%	7.2%	11.1%
Specific Gravity	M3	M4	-18.8%	-33.1%	23.0%	18.0%
	M11	M9	-13.9%	-11.6%	1.7%	-4.0%
	M2	M7	10.1%	9.8%	30.5%	10.8%
Organic Content	M5	M6	-30.8%	23.5%	-29.8%	39.0%
	M8	M9	-42.1%	-9.4%	-31.7%	33.4%
	M10	M11	-15.8%	17.4%	-18.0%	18.0%

Table 4.6: Comparison of performance of best ANN model developed in this study to models published in the literature evaluated by RMSE m^3m^{-3} of testing dataset (HC = heat capacity, FSI = Free swelling index).

Model	**Algorithm**	**PTF Type**	**Input Variables**	**RMSE**	**Dataset Size**
SNN-M2	ANN with Leaky ReLU	Continuous	%Clay, %Silt, %Sand, DD	0.042	3217 training data points and 122 testing data points, the dataset is extracted from UNSODA
DNN-M12	ANN with Leaky ReLU	Continuous	%Clay, %Silt, %Sand, PO	0.039	1277 training data points and 122 testing data points, the dataset is extracted from UNSODA
Schaap et al. [149]	Model 1	ANN	%Clay, %Silt, %Sand	0.076	UNSODA dataset is used for model development
	Model 2	ANN Parametric	%Clay, %Silt, %Sand, DD	0.068	UNSODA dataset is used for model development
Bayat and Zadeh [6]	Model 1	ANN	%Clay, %Silt, %Sand, DD	0.036	98 training samples and 50 testing samples, samples are collected from six provinces of Iran
	Model 2	ANN Parametric	%Clay, %Silt, %Sand, HC	0.038	98 training samples and 50 testing samples, samples are collected from six provinces of Iran
Javanshir et al. [294]	Model 1	ANN	%Clay, %Silt, %Sand, DD	0.047	90 training samples and 40 testing samples, samples are collected from five provinces of Iran
	Model 2	Regression Parametric	%Clay, %Silt, %Sand, DD	0.049	90 training samples and 40 testing samples, samples are collected from five provinces of Iran
	Model 3	ANN	%Clay, %Silt, %Sand, DD, FSI	0.043	90 training samples and 40 testing samples, samples are collected from five provinces of Iran
	Model 4	Regression Parametric	%Clay, %Silt, %Sand, DD, FSI	0.048	90 training samples and 40 testing samples, samples are collected from five provinces of Iran

4.8 Summary

The performances of single (SNN-PTF) and double (DNN-PTF) hidden layer ANN models using thirteen combinations of up to eight input parameters (in addition to suction) have been compared. Differences in performance between SNN-PTF and DNN-PTF are found to be significant in seven models, but which of the two (SNN vs DNN) is more favourable is highly case-dependent. It is hence recommended to give the freedom to the ANN model to exceed one hidden layer when building WRC PTF, and to consider the number of hidden layers as an additional hyperparameter to be tuned during model development. The study has also found that, of the five input parameters added to texture, only porosity and dry density yield a consistent and significant improvement in accuracy. In fact, based on the dataset used in this study, models with texture variables and only one of the two variables of porosity or dry density appear sufficient for generating accurate predictions of WRC.

The above findings are based on a database of around 3200 points and are hence inevitably limited by the breadth of data used. Indeed, other potentially useful predictors, such as initial water content, soil mineralogy, cation exchange capacity, wetting and drying paths, and method of soil compaction, could not be explored here because of data limitations. For future studies, there is a need for creating more comprehensive datasets from the literature. Specifically, further enriching datasets with more data points, a wider range of values for variables and additional potentially useful input variables are arguably the most effective measures for further improving the prediction accuracy of PTFs based on neural networks. An extensive experimental literature on the water retention curve exists in peer-reviewed journals. While some of the water retention data in that literature can be found in databases such as UNSODA, much of it remains undigitized and not yet fully exploited in PTF model building. This presents a valuable opportunity in this space, provided digitised data undergoes strict quality assurance evaluation. Such evaluation is necessary because of the wide range

of methods used to measure WRC in the laboratory and the different forms in which this data is presented. In addition to the data-related limitations, an important architecture limitation must be acknowledged. The ANN models used in this study are a type of implicit model that does not contain explicit interaction terms between input variables. Therefore, any interaction can only be captured indirectly, making it difficult to rank feature importance. In future research, incorporating feature importance ranking could add to our understanding of how different soil properties contribute to WRC predictions with AI. Developments in the interpretability tools of ANN, such as Shapley Additive explanations (SHAP) method, can offer new insights and their application to WRC PTF estimation is worth in future study.

Chapter 5

Water Retention Curves Prediction for Clay Soil

5.1 Overview

This chapter presents the evaluation of a newly developed PTF-based neural network model specifically trained on clayey soil datasets for predicting the WRC of clayey soils. The performance of the ANN model trained exclusively on clayey soil data (PTF-C1) is compared with that of a model trained on a broader dataset (PTF-C2) containing various soil textures. Each PTF model consists of seven sub-models, and each sub-model is used to estimate the water content at a predefined matric potential. The evaluation is performed based on nine clay soil samples as described in the methodology chapter.

The evaluation of the developed PTF models was conducted in three stages. Firstly, the predictive performance of the sub-models (C1a to C1g for PTF-C1, and C2a to C2g for PTF-C2), which estimate water content at seven predefined matric potentials (6, 10, 33.3, 100, 200, 500, and 1500 kPa), was assessed. Secondly, for each PTF model, the predicted water contents at the seven predefined matric potentials were used to fit the VG WRC model to obtain the corresponding fitting parameters and generate the corresponding WRC. The fitted curves were compared with the curve generated by directly fitting the experimental data points to assess the overall performance of PTF models.

5.2 Evaluation of Sub-model Predictions at Predefined Matric Potentials

The performance evaluation for each sub-model is summarised in Table 5.1. The table presents the MAE (Eq. 3.12), RMSE (Eq. 3.11), and % Error(Eq. 3.13) for the prediction of each sub-model at its predefined matric potential. Overall, the clay-focused PTF-C1 has lower errors across all metrics in 5 out of the 7 sub-models. In sub-models corresponding to 6 kPa and 500 kPa, PTF-C2 has a slightly lower error than PTF-C1.

The mean %Error for PTF-C1 is 10.5%, while the error is 13.2% for PTF-C2. PTF-C1 has approximately 20% lower errors than PTF-C2 among all assessment metrics, MAE, RMSE, and %Error. It is also noticeable that the %Error of PTF-C1 remain relatively stable across most matric potentials with a %Error of 10%. The only exception is at 100kPa, and the %Error is 13.2% for that case. In contrast, PTF-C2 shows a much wider range of variation, with %Error ranging from as low as 7% (at 6 and 500 kPa) to as high as 19.5% (at 33.3 kPa).

Table 5.1: Prediction errors of PTF-C1 and PTF-C2 models for water content estimation at different matric potentials. MAE, RMSE, and %Error are reported for each matric potential, along with the overall mean values.

Matric Potential (kPa)	MAE		RMSE		% Error	
	PTF-C1	PTF-C2	PTF-C1	PTF-C2	PTF-C1	PTF-C2
6	0.057	0.041	0.067	0.051	9.8%	7.0%
10	0.052	0.084	0.083	0.103	10.4%	16.8%
33.3	0.044	0.079	0.050	0.122	10.8%	19.5%
100	0.045	0.049	0.051	0.064	13.2%	14.4%
200	0.031	0.042	0.035	0.054	9.7%	13.3%
500	0.032	0.025	0.032	0.025	9.0%	7.0%
1500	0.024	0.025	0.031	0.032	9.4%	9.9%
Mean	0.042	0.052	0.062	0.086	10.5%	13.2%

5.3 Estimation of van Genuchten Model Parameters Using PTF-Predicted Data

After assessing the prediction accuracy of individual sub-models, the predicted water contents were further used to estimate the van Genuchten model parameters through curve fitting as the model architecture shown in Figure 3.10. The VG fitting parameters derived from PTF-C1 and PTF-C2 predictions are summarised in Table 5.3, which also includes the reference parameters obtained by directly fitting the experimental data. Using these parameters, the WRC can be built based on the VG WRC model. Figure 5.1 presents the WRC predictions for the first test sample. The figure includes three lines, two WRC curves generated from PTF-C1 and PTF-C2, and one curve generated by direct fitting to experimental data. The figure also shows the experimental data points and the predicted water contents at the seven predefined matric potentials. As shown in the figure, the fitting quality for each set of data is great, and the fitted curve can reflect the trend that their associated data points present. Then, the two fitted curves are compared with the reference curve. Both curves overestimate the saturation water content, as indicated by the horizontal lines on the left. However, PTF-C1 provides a better approximation of the air entry value, capturing the knee of the curve more accurately. While both PTF-C1 and PTF-C2 perform well in capturing the steep decline after air entry, they begin to deviate near the dry end. In this region, both deviate from the residual water content, with PTF-C1 showing slightly greater deviation than PTF-C2.

Furthermore, the error of the fitted WRCs generated by the PTF-C1 and PTF-C2 models from the reference curves obtained through direct fitting was evaluated using MAE, RMSE, and % Error. These metrics were calculated based on the fitted curves for each of the nine soil samples, and the results are summarised in Table 5.3. The table also includes the average performance across all samples for both models.

While PTF-C1 outperforms PTF-C2 in most cases with lower MAE and %Er-

Table 5.2: Fitting parameters estimated for the nine test samples using PTF-C1 and PTF-C2 predictions, together with the reference parameters obtained by direct fitting of the experimental data. The percentage difference of each predicted fitting parameters are also presented

Parameters		1	2	3	4	5	6	7	8	9
Direct Fit	θ_r	0.237	0.237	0.203	0.274	0.283	0.201	0.167	0.282	0.287
	θ_s	0.482	0.566	0.493	0.710	0.621	0.472	0.428	0.879	0.675
	α	6.992	6.339	12.905	12.747	8.579	9.799	4.393	10.543	9.899
	n	1.780	1.569	1.584	1.537	1.551	1.708	1.679	1.473	1.752
PTF-C1	θ_r	0.175	0.226	0.106	0.144	0.212	0.000	0.137	0.334	0.291
		-26.2%	-4.6%	-47.8%	-47.4%	-25.1%	-100%	-18.0%	18.4%	1.4%
	θ_s	0.512	0.549	0.537	0.616	0.534	0.487	0.368	0.749	0.690
		6.2%	-3.0%	8.9%	-13.2%	-14.0%	3.2%	-14.0%	-14.8%	2.2%
	α	10.138	8.693	30.087	8.563	10.958	23.075	1.902	18.651	9.442
		45.0%	37.1%	133.1%	-32.8%	27.7%	135.5%	-56.7%	76.9%	-4.6%
	n	1.503	1.592	1.195	1.417	1.366	1.151	1.466	1.541	1.788
		-15.6%	1.5%	-24.6%	-7.8%	-11.9%	-32.6%	-12.7%	4.6%	2.1%
PTF-C2	θ_r	0.202	0.172	0.208	0.229	0.202	0.174	0.165	0.306	0.266
		-14.8%	-27.4%	2.5%	-16.4%	-28.6%	-13.4%	-1.2%	8.5%	-7.3%
	θ_s	0.397	0.423	0.412	0.477	0.441	0.380	0.349	0.578	0.462
		-17.6%	-25.3%	-16.4%	-32.8%	-29.0%	-19.5%	-18.5%	-34.2%	-31.6%
	α	2.009	2.293	2.059	1.680	2.274	2.012	1.804	1.543	1.925
		-71.3%	-63.8%	-84.0%	-86.8%	-73.5%	-79.5%	-58.9%	-85.4%	-80.6%
	n	1.661	1.426	1.596	1.892	1.456	1.502	1.549	1.717	1.742
		-6.7%	-9.1%	0.8%	23.1%	-6.1%	-12.1%	-7.7%	16.6%	-0.6%

ror, the RMSE values between the two PTF models are often quite similar. For example, in Sample 5, PTF-C1 and PTF-C2 have MAE values of 0.055 and 0.086, respectively, reflecting a 35% relative improvement. However, the RMSE values (0.580 vs. 0.560) are nearly the same, indicating similar deviation magnitudes. From an overall perspective, the mean RMSE values of the two models are nearly identical (0.546 vs. 0.534), but PTF-C1 achieves a lower mean MAE (0.038 vs.0.052) and a lower mean %Error (10.0% vs. 13.6%). This indicates that PTF-C1 shows an improved predictive performance than PTF-C2.

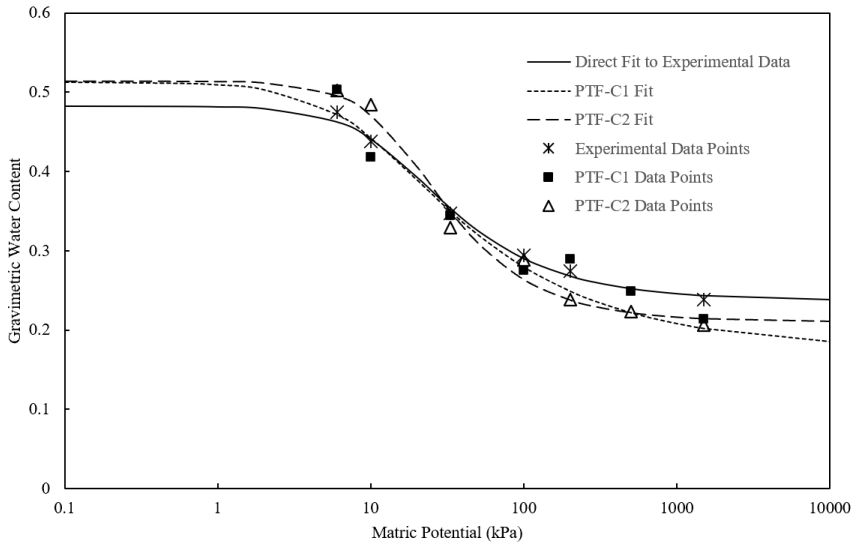


Figure 5.1: Comparison of the fitted WRC from PTF-C1 and PTF-C2 with the curve derived from direct fitting of experimental data points. The experimental data points and the predicted water contents at seven predefined matric potentials (6, 10, 33.3, 100, 200, 500, and 1500 kPa) for PTF-C1 and PTF-C2 are also presented. *Note: The experimental data point at 500 kPa is unavailable for this soil sample, so there are only six experimental data points.

Table 5.3: The accuracy of the fitted WRC for the PTF-C1 and PTF-C2 models. MAE, RMSE, and %Error are presented for each of the nine soil samples, along with the overall mean values.

Sample	MAE		RMSE		% Error	
	PTF-C1	PTF-C2	PTF-C1	PTF-C2	PTF-C1	PTF-C2
1	0.015	0.024	0.474	0.479	4.6%	7.3%
2	0.032	0.053	0.546	0.534	8.2%	13.7%
3	0.043	0.024	0.475	0.464	14.0%	7.9%
4	0.060	0.087	0.603	0.584	13.7%	19.9%
5	0.055	0.086	0.580	0.560	13.0%	20.2%
6	0.026	0.016	0.449	0.447	8.9%	5.3%
7	0.025	0.019	0.427	0.431	8.5%	6.4%
8	0.079	0.121	0.739	0.724	14.6%	22.3%
9	0.007	0.036	0.617	0.587	1.6%	8.6%
Mean	0.038	0.052	0.546	0.534	10.0%	13.6%

5.4 Summary

This chapter developed and evaluated two PTF models to investigate whether excluding non-clayey soils from the training dataset improves the accuracy of ANN-based WRC predictions for clayey soils. Between the two PTF models examined in this chapter, the PTF trained exclusively on clayey soil data (PTF-C1) achieves over 20% improvement on average compared to the model trained on a mixed-texture dataset (PTF-C2), in which only 10.7% of the samples were clayey soils. The results highlight the advantage of using soil-type-specific training data. Although both models shared the same structure, the model trained exclusively on clayey soils demonstrated better WRC prediction accuracy across most evaluation scenarios, including both point-wise assessment at predefined matric potentials and the overall fitting of the WRC curve. It must be acknowledged that each empirical WRC model has its own domain of suitability. In this study, the van Genuchten model is adopted to represent the hydraulic behaviour of all types of soil. However, this empirical model is not suitable for all soil types, preventing it from fully capturing the true hydraulic behaviour. Furthermore, even in soils for which the van Genuchten model is considered most appropriate, its fixed mathematical form inherently restricts the flexibility of the model and introduces structural uncertainties to the fitting parameters. When these parameters are subsequently used in hydraulic modelling, the uncertainties in these parameters will be propagated through the modelling process and may affect the modelling outcomes. In future work, instead of selecting only one form of empirical WRC model for all soil samples, automatically selecting the most suitable empirical WRC model that best fit the specific soil characteristics may lead to more reliable models.

Chapter 6

Advection-Diffusion Modelling with Machine Learning

6.1 Overview

This chapter presents the evaluation of the newly proposed hard-constraint PINN incorporating a Peclet-adjusted exponential distance function for solving the ADE under both steady-state and transient conditions. The model's performance is systematically assessed across a wide range of Peclet numbers, including high P_e scenarios that are particularly challenging for conventional PINN approaches and traditional numerical methods. The evaluation is conducted under Dirichlet-Dirichlet and Dirichlet-Neumann boundary conditions.

The results are organised as follows: first, the model is tested for steady-state ADE problems with P_e ranging from 1 to 1,000,000, and the effects of different distance function formulations are examined. An analysis of the optimised shape factor and its relationship with P_e is also presented. Next, the model's performance is evaluated for transient cases under Dirichlet-Dirichlet boundary conditions, including an investigation of shape factors in both spatial and temporal domains. Finally, the model is assessed for transient Dirichlet-Neumann problems to assess its predictive capability under mixed boundary conditions.

6.2 Case 1: Steady-State with Dirichlet-Dirichlet Boundary Conditions

The ability of the three different types of distance functions associated with HC to simulate case 1 was assessed in two stages. In the first stage, optimal values of parameter α_x for the exponential and logarithmic functions were determined.

In the second stage, the performances of four versions of the PINN algorithm, namely soft constraints and HC with linear, optimal exponential, and optimal logarithmic, were compared. Figure 6.1 shows convergence histories (loss versus epoch) of exponential and logarithmic functions for different values of α_x, as well as loss versus α_x at different epochs, for $P_e = 50$ and $P_e = 100$. In addition, two indicators are introduced to further quantify the speed of convergence of PINN under different distance functions: E_1 and E_5, defined as the number of epochs at which the loss function first drops below 1 and 5, respectively; the smaller E_1 and E_5, the faster the convergence. Figure 6.2 shows E_1 and E_5 versus α_x for both exponential and logarithmic distance functions. Four observations can be drawn from Figure 6.1 and Figure 6.2.

First, it is clear that α_x has a strong effect on both PINN accuracy (Figure 6.1) and speed of convergence (Figure 6.2). Second, it can be seen from the right-hand side graphs of Figure 6.1 (loss versus epoch) that, even after the loss has dropped to low values, some fluctuations still occur at later epochs. It is notable, however, that these fluctuations are far more pronounced in the logarithmic distance functions, and rather mild in the exponential one.

Third, the right-hand side graph of Figure 6.1c shows that for the logarithmic distance function at $P_e = 50$, the range of α_x corresponding to a basin of low loss values widens with the number of epochs completed. The same observation can be made for $P_e = 100$ except that some fluctuations occur before the low loss values stabilise by the time 1000 epochs are reached. Based on these two cases, an optimal value of $\alpha_x = 25$ is chosen for the logarithmic function.

Finally, a much stronger pattern emerges in Figure 6.1a, b and Figure 6.2 for the exponential function with a minima of loss, E_1 and E_5 as a function of α_x clearly visible. This optimal minimum seems to correspond $\alpha_x = P_e$. To further test this finding and explore whether it applies to higher Peclet numbers, Figure 6.3 shows loss versus α_x for the exponential function at $P_e = 200$ and $P_e = 500$. Although the loss function is generally larger under $P_e = 500$, compared to

lower Peclet numbers, and experiences more fluctuations, the pattern of optimal $\alpha_x = P_e$ holds well for both cases considered. Based on the above discussion, the following form is adopted as optimal:

$$l^{\partial\Omega}(x) = \left[1 - e^{-P_e \frac{x-a}{b-a}}\right]^{i_a} \left[1 - e^{-P_e \frac{b-x}{b-a}}\right]^{i_b} \tag{6.1}$$

For the second stage of evaluation, Figure 6.4 compares the performances of four different versions of PINN: soft constraint, and HC based on linear, logarithmic, and exponential functions. A soft-constraint approach struggles to minimise the loss function, and the loss-versus-epoch curve, after an initial drop, remains flat at a rather high value of loss. By comparison, a hard-constraint approach based on a linear distance function achieves a much better reduction in loss function after 400 and 4700 epochs, for $P_e = 50$ and 100, respectively. The logarithmic distance function achieves even lower loss values at earlier epochs but suffers from significant fluctuations at later epochs. Hence, the exponential distance function performs best by far in terms of faster convergence and a much lower level of fluctuations. While the value of minimised loss for the exponential function is higher than the average value for the logarithmic function, the former remains much more reliable, especially given the fluctuations of the logarithmic function.

Finally, the performance of the hard-constraint PINN with exponential distance function is tested for a range of Peclet numbers. Figure 6.5 shows the comparison between PINN predictions and analytical solutions for P_e values ranging from 0.5 to 5000, all demonstrating good agreement. Table 6.1 summarises the model performance across a broader P_e range from 1 to 1 million. Each case was repeated five times, and the reported accuracy is the average of each case. Table 6.1 also presents the convergence speed with respect to each P_e, average time (AVG) and standard deviation (STD). For $P_e < 500$, no clear pattern is observed. The slowest convergence occurs at $P_e = 1$, where early stopping was not triggered. This is because the loss continued to decrease throughout the training process. Since the model allows continuous refinement, it kept improving for

this relatively smooth case. Introducing a target accuracy threshold could help terminate training earlier in practical use. Nevertheless, this case also achieves the lowest error. For $P_e > 500$, convergence speed generally declines as Peclet number increases.

Table 6.1: Case 1, Accuracy and Convergence Speed of Hard-Constraint PINN with Exponential Distance Function at Different Peclet Numbers.

Parameters					**Accuracy**			**CPU Time (s)**	
D	**V**	**L**	P_e	α_x	**MAE**	**RMSE**	R^2	**AVG**	**STD**
0.1	0.1	1	1	1	1.2×10^{-5}	1.4×10^{-5}	1.000	77.9	1.9
0.1	0.5	1	5	5	2.2×10^{-4}	2.7×10^{-4}	1.000	11.8	4.4
0.1	1	1	10	10	6.9×10^{-4}	8.8×10^{-4}	1.000	9.5	1.9
0.1	5	1	50	50	5.3×10^{-3}	6.5×10^{-3}	0.996	20.9	3.4
0.1	10	1	100	100	3.8×10^{-3}	4.7×10^{-3}	0.996	19.7	2.9
0.1	50	1	500	500	1.4×10^{-3}	1.7×10^{-3}	0.997	3.9	0.8
0.1	100	1	1000	1000	7.5×10^{-4}	8.9×10^{-4}	0.998	4.0	0.7
0.1	1000	1	10000	10000	3.1×10^{-4}	3.6×10^{-4}	0.998	6.4	3.0
0.1	1000	2	20000	20000	6.1×10^{-4}	6.9×10^{-4}	0.990	7.7	2.4
0.1	1000	5	50000	50000	2.9×10^{-4}	3.3×10^{-4}	0.998	14.3	2.4
0.1	1000	10	100000	100000	2.4×10^{-4}	2.8×10^{-4}	0.998	27.0	7.8
0.1	1000	100	1000000	1000000	1.3×10^{-3}	1.4×10^{-3}	0.960	37.5	2.1

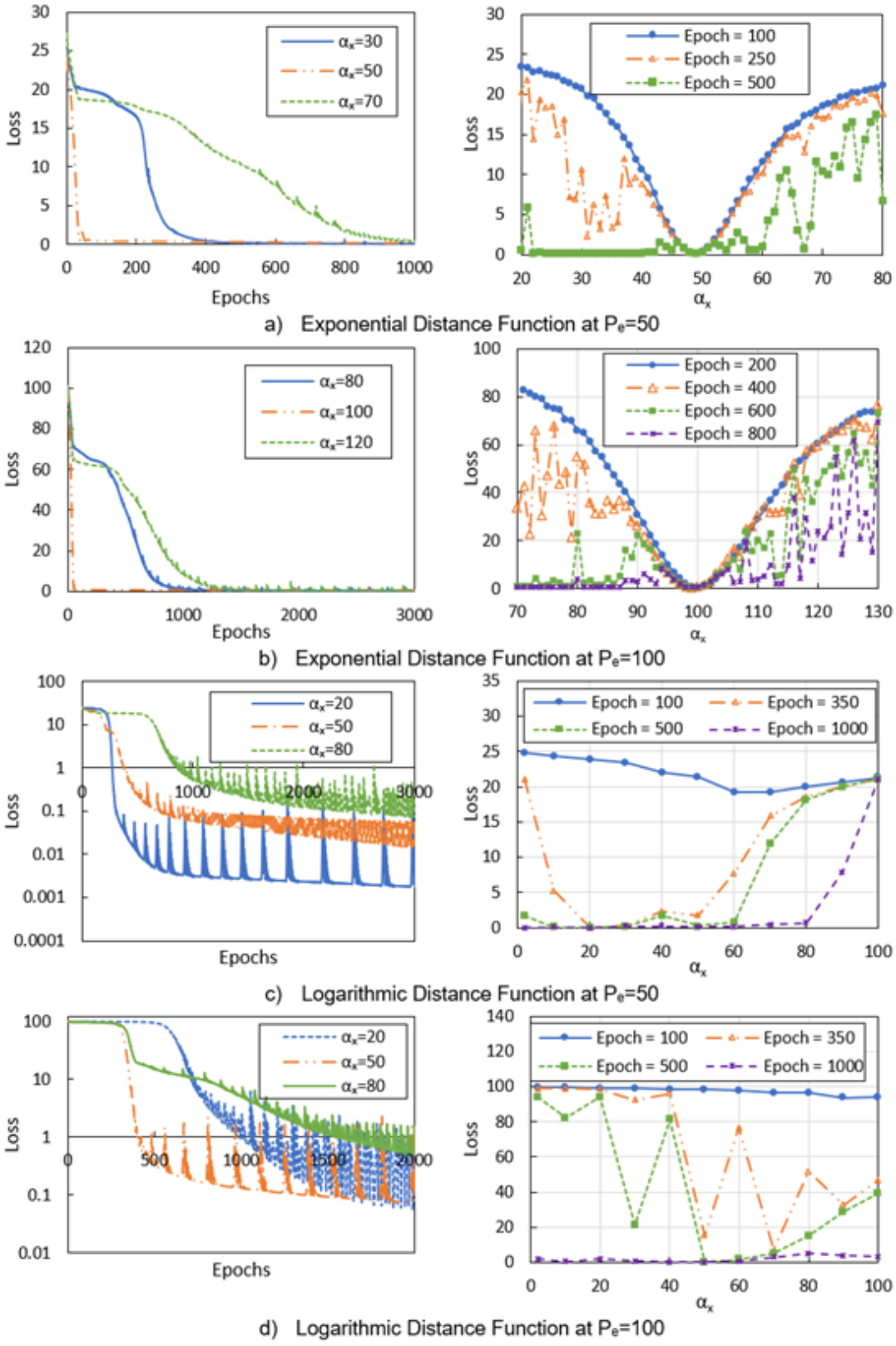


Figure 6.1: Case 1, Results showing the performance of the PINN solver under different distance function formulations and Peclet number cases using various values of α_x. The left-hand side figures show the loss versus epochs for three values of α_x; the right-hand side figures show the loss versus α_x at four different epochs.

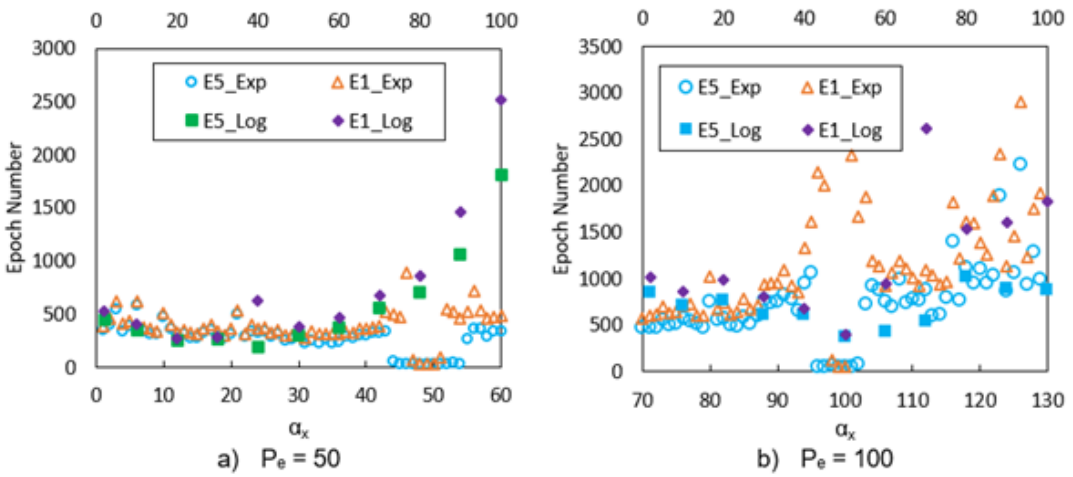


Figure 6.2: Case 1, E_1 and E_5 versus α_x for a) $P_e = 50$ and b) $P_e = 100$.

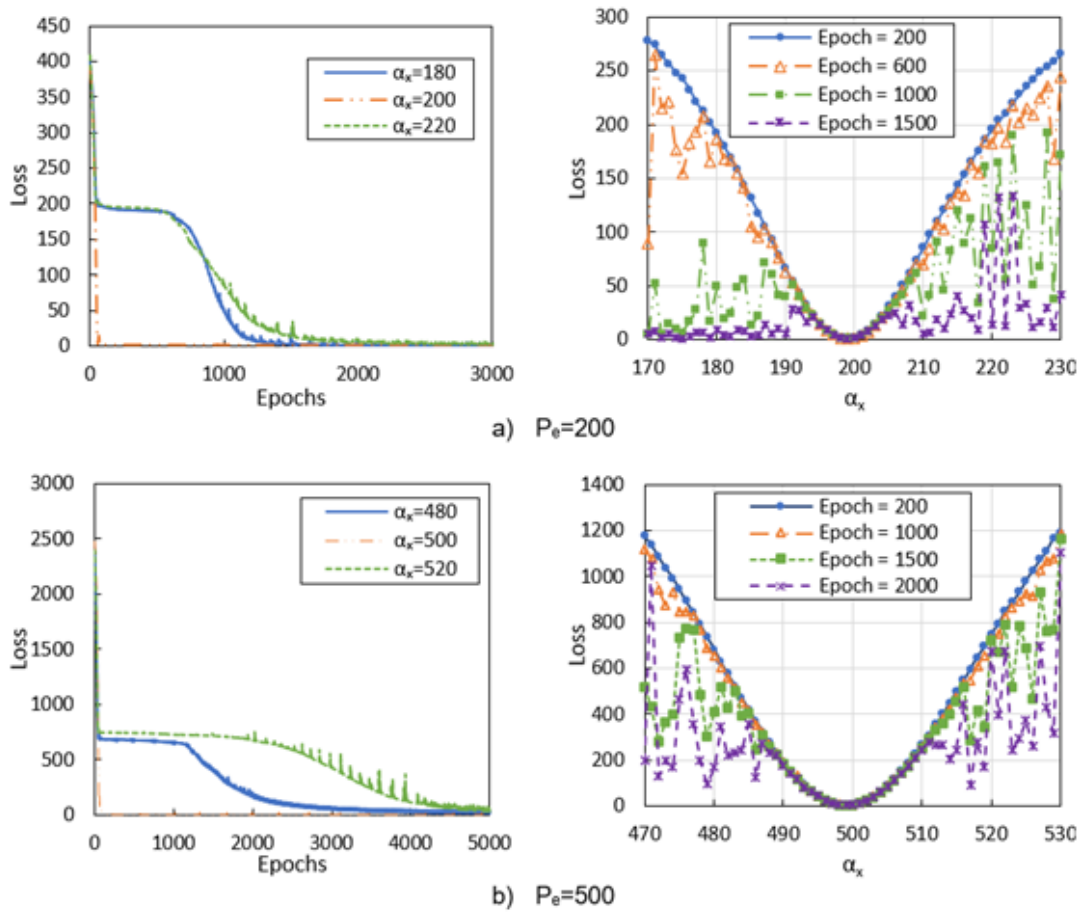


Figure 6.3: Case 1, a) $P_e = 200$ and b) $P_e = 500$, showing the performance of PINN solver under exponential BC distance function using different values of α_x. Left-hand side figures show loss versus epochs for three α_x; right-hand side figures show loss versus α_x at four different epochs.

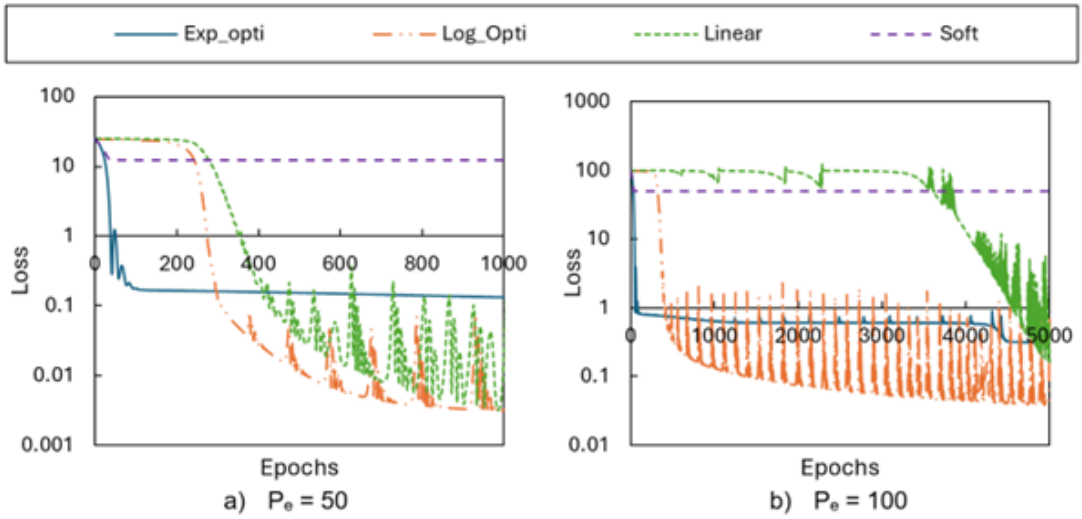


Figure 6.4: Case 1, loss history comparison among different PINN versions (with soft and hard constraints) for a) $P_e = 50$ and b) $P_e = 100$.

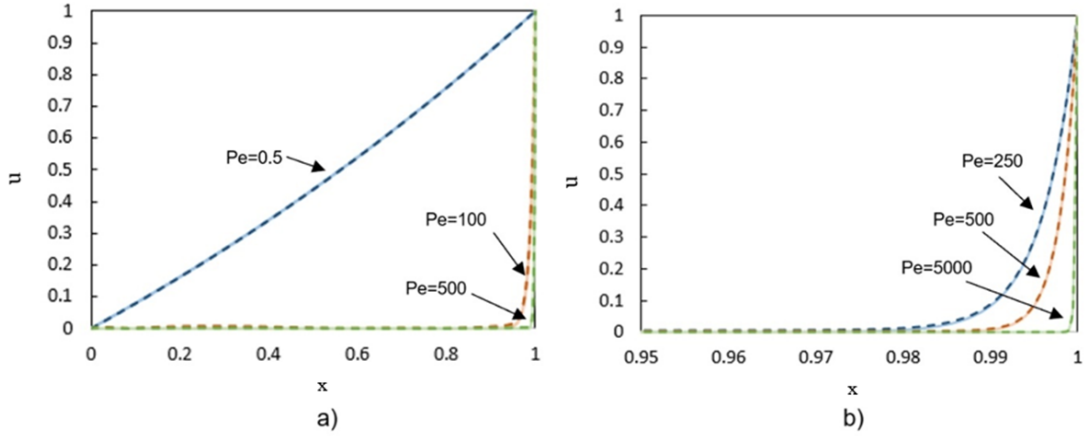


Figure 6.5: Case 1, comparison between PINN predictions and analytical solutions. a) including cases of $P_e = 0.5$, 100, and 500 and b) zooming to the range of 0.95∼1 in the spatial domain for high P_e cases comparison at $P_e = 250$, 500, and 5000. **Note: The solid line is the analytical solution, and the dashed line is the PINN solution.*

6.3 Case 2: Transient Case Study with Dirichlet-Dirichlet Boundary Conditions

Cases 2a and 2b, shown in Table 6.2, are time-dependent variations of case 1 analysed in the previous section. The only difference between cases 2a and 2b lies in their initial conditions. Case 2a has a linearly distributed initial concentration between two boundaries at x=0 and x=1, which ensures compatibility between initial and boundary conditions. Case 2b has a uniform initial concentration of zero across the spatial domain, creating a discontinuity at x=0 between the boundary (u=1) and initial condition (u=0). These two cases are used to assess the effects of different distance functions for initial and boundary conditions on the performance of PINN. Ten tests are conducted for this purpose, as summarised in Table 6.2. Tests 2.1-2.6 and 2.7-2.10 are based on cases 2a and 2b, respectively. All tests have $P_e = 100$, except for Test 2.10 conducted under $P_e = 500$. HC-PINN predictions in all tests are compared to those of a FEM model discretised with a highly refined mesh established through a mesh-convergence analysis.

Table 6.2: Summary of tests conducted on Cases 2a and 2b to determine optimal BC and IC distance functions ($L_d = 1$ in all problems; for Case 2a: $t_{max} = 2$, $D = 0.01$ and $v = 1$; for Case 2b: $t_{max} = 1$, $D = 0.1$ and $v = 10$, except for 2.10 where $t_{max} = 1$, $D = 0.01$ and $v = 5$)

Case	Test	Aim	BC Distance Function	IC Distance Function	P_e	Agreement with FEM
2a	2.1	Assess the performance of PINN with different combinations of BC and IC distance functions	Linear	Linear	100	Poor
	2.2		Linear	Exponential $\alpha_t = 50$	100	Poor
	2.3		Exponential $\alpha_x = P_e$	Linear	100	Good
	2.4		Exponential $\alpha_x = P_e$	Exponential $\alpha_t = 50$	100	Good
	2.5	Determine optimal α_x for BC distance function	Exponential $\alpha_x \epsilon [86,115]$	Linear	100	Good
	2.6	Determine optimal α_t for IC distance function	Exponential $\alpha_x = P_e$	Exponential $\alpha_t \epsilon [85, 115]$	100	Good
2b	2.7	Determine the best IC distance function	Exponential $\alpha_x = P_e$	Linear	100	Poor
	2.8		Exponential $\alpha_x = P_e$	Exponential $\alpha_t \epsilon [1,150]$	100	Good
	2.9	Determine the best BC distance function	Exponential $\alpha_x \epsilon [86,115]$	Exponential $\alpha_t = 50$	100	Good
	2.10		Exponential $\alpha_x \epsilon [486,515]$	Exponential $\alpha_t = 50$	500	Good

6.3.1 Case 2a: Linear initial condition distribution: compatible IC and BC

Figure 6.6 shows results from Tests 2.1 to 2.4 which consider different combinations of linear and exponential BC and IC distance functions. In Tests 2.2 and 2.4, an exponential IC distance function is used with $\alpha_t = 100$, while in Tests 2.2 and 2.3, a linear IC distance function is used. Figure 6.6 clearly shows that, as under steady-state, exponential BC distance functions yield highly accurate solutions, regardless of whether linear or exponential IC distance function is used, at least for the case under study. PINN with linear BC distance function, on the other hand, fails to produce accurate predictions.

With the exponential BC distance function confirmed as superior, the aim of Test 2.5 is to investigate whether $\alpha_x = P_e$ remains optimal, by examining a range of values on either side, namely $[P_e - 15, P_e + 15]$. A linear IC distance function is used in this test. Figure 6.7 confirms the optimality of $\alpha_x = P_e$ which produces the lowest loss at the fastest rate.

In Test 2.6, exponential BC and IC distance functions are adopted with $\alpha_x = 100$ and the effect of changing the value of α_t between $[\alpha_t - 15, \alpha_t + 15]$ investigated. Results, shown in Figure 6.8a, do not reveal any distinct relationship between loss and α_t, except for lower and relatively more constant loss versus α_t at higher epochs. Figure 6.8a also confirms that a linear IC distance function performs at least as well as exponential ones, while exhibiting fewer fluctuations with epochs.

These results hence confirm the conclusion drawn from Tests 2.1 to 2.4 that a combination of exponential BC distance function with $\alpha_x = P_e$ and a linear IC distance function is optimal, at least for Dirichlet-Dirichlet BCs with no discontinuity between BC and IC.

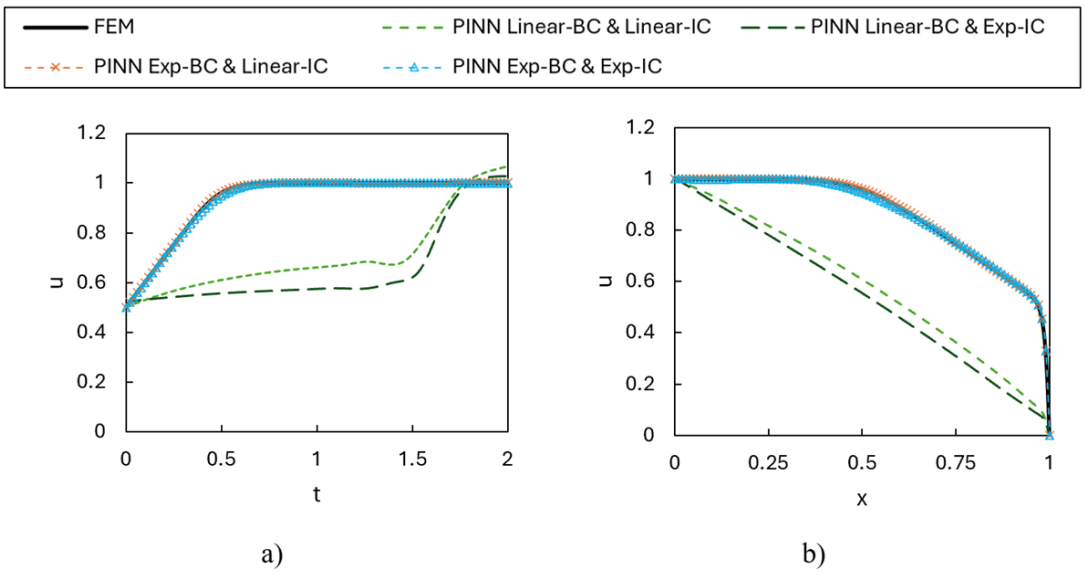


Figure 6.6: Case 2a, P_e=100, results of Tests 2.1 to 2.4 showing performance of PINN solver under different combinations of linear and exponential distance functions: a) u versus t at x=0.5, and b) u versus x at t=0.5. **Note: Linear-BC = Linear BC distance function, Linear-IC = Linear IC distance function, Exp-BC = Exponential BC distance function, and Exp-IC = Exponential IC distance function.*

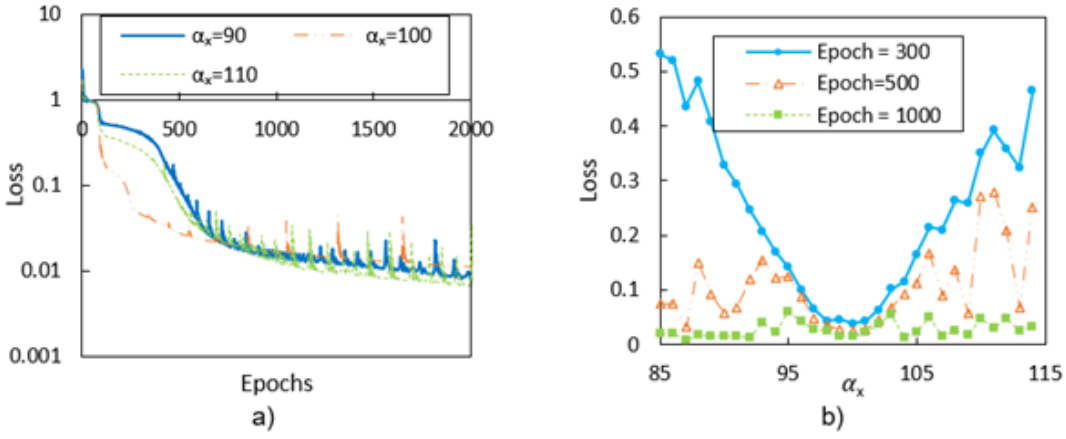


Figure 6.7: Case 2a, P_e=100, results of tests 2.5 showing performance of PINN solver under BC exponential and IC linear distance functions with different values of α_x: a) loss versus epochs for three α_x and b) loss versus α_x at three different epochs.

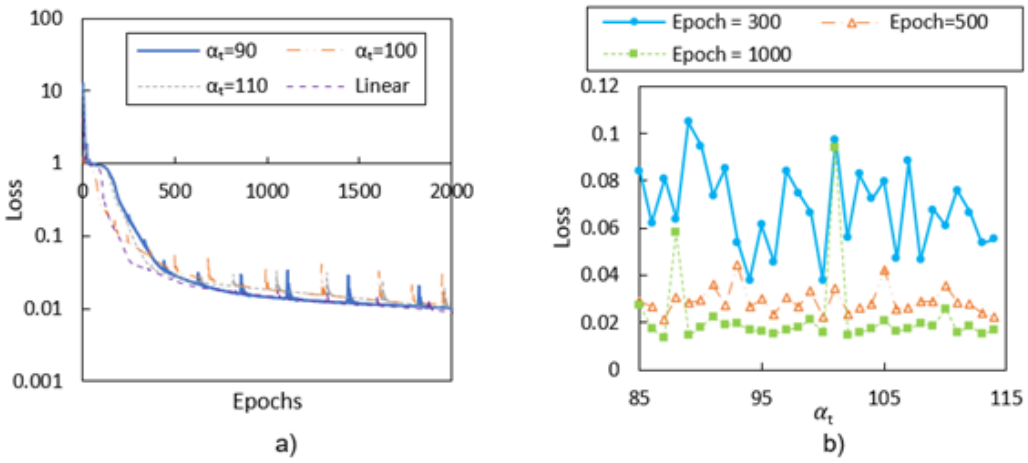


Figure 6.8: Case 2a, $P_e = 100$, results of tests 2.6 showing performance of PINN solver under BC exponential and IC exponential distance functions with different values of α_t: a) loss versus epochs for three α_t and b) loss versus α_t at three different epochs.

6.3.2 Case 2b: Uniform initial condition distribution: non-compatible BC and IC

Tests 2.7 to 2.10 analyse the performance of linear and exponential distance functions (BC and IC) under the conditions of Case 2b, where there is a discontinuity at one boundary between BC and IC, which leads to incompatibility between BC and IC hard constraints.

Using an exponential BC distance function with $\alpha_x = P_e$, Tests 2.7 and 2.8 assessed the performance of linear and exponential IC distance functions, respectively. Starting with the linear IC distance function Figure 6.9 shows the contours of u versus x and t. It is clear from the figure that IC conditions dominated the solution in earlier times, with the PINN solver unable to enforce the BC of u=1 at x=1 until much later.

Moving to exponential IC distance function, Figure 6.10 shows the results of Test 2.8, with a range of α_t values tested. The figure indicates that an IC exponential distance function with low $\alpha_t = 10$ achieves the lowest loss and fastest convergence. However, it is important to note that as α_t declines and approaches 0, the IC distance function reduces to 0 (see Table 3.4) which is unsuitable for enforcing initial conditions as a hard constraint.

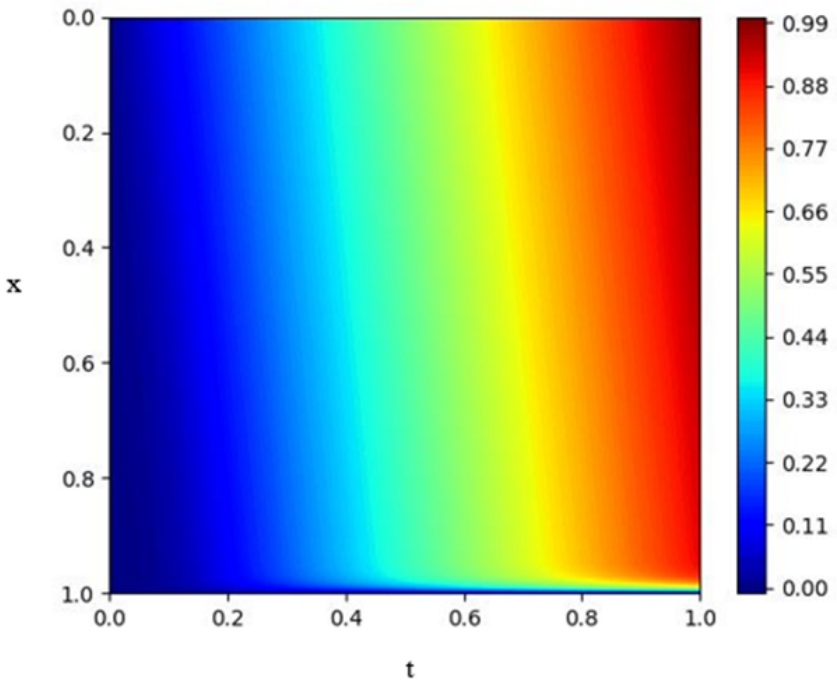


Figure 6.9: Case 2b, $P_e = 100$, contour result of test 2.7 of PINN solver under exponential BC distance function and linear IC distance function.

The effect of decreasing α_t can be seen by examining the different contours in Figure 6.11. The incompatibility between IC and BC in case 2b leads to a transition zone in the PINN result field. The width of the transition zone is significantly influenced by the choice of α_t. For instance, the length of the transition zone is around 0.45 when $\alpha_t = 10$, 0.08 when $\alpha_t = 50$, and 0.04 when $\alpha_t = 100$. To achieve a reasonably small transition zone, α_t greater than 50 appears necessary.

Next, adopting an IC exponential distance function with a fixed $\alpha_t = 50$, Tests 2.9 and 2.10 explore whether $\alpha_x = P_e$ continues to be optimal for the BC exponential distance function, under $P_e = 100$ and $P_e = 500$. Figure 6.12 confirms that this is the case in both tests, with the lowest loss and fastest convergence observed under $\alpha_x = P_e$. It is noticeable however that the minimal loss achieved is significantly higher under Pe=500 compared to $P_e = 100$. Finally, adopting exponential distance functions for both BC and IC, with $\alpha_x = P_e$ and $\alpha_t > 50$, Figure 6.13 shows excellent agreement between the proposed PINN algorithm and FEM results.

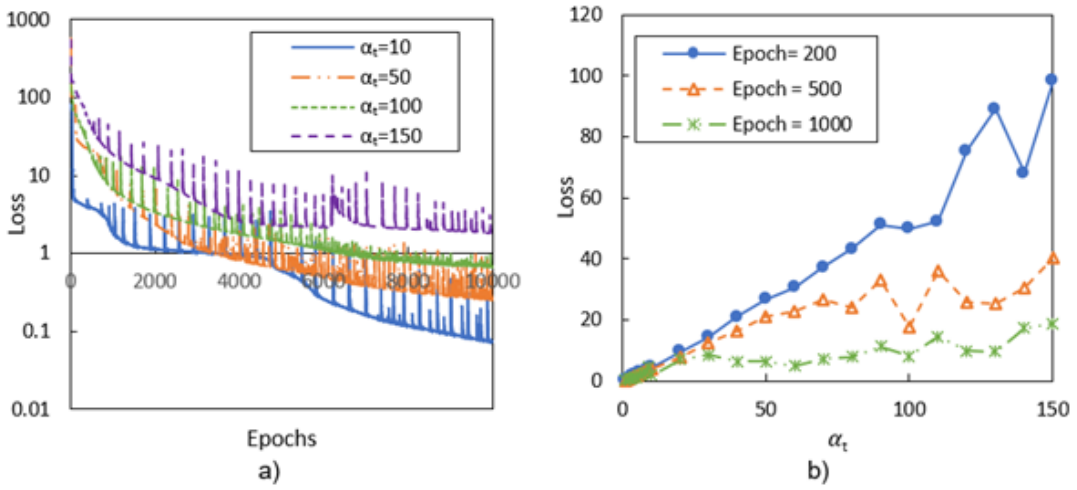


Figure 6.10: Case 2b, $P_e = 100$, results of test 2.8 showing performance of PINN solver under exponential BC distance function and exponential IC distance function using different values of α_t: a) loss versus epochs for four α_t and b) loss versus α_t at three different epochs.

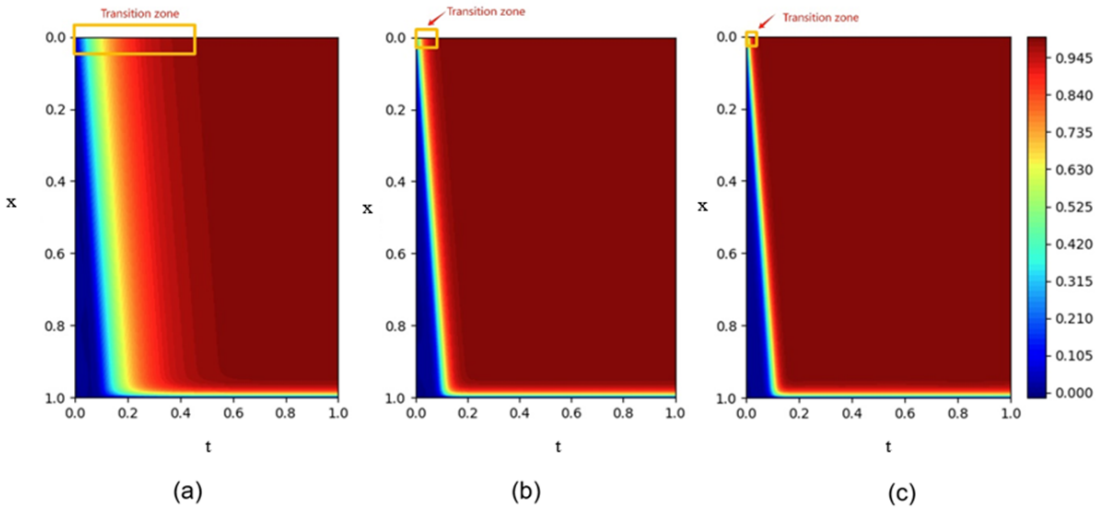


Figure 6.11: Case 2b, $P_e = 100$, results of test 2.8 showing performance of PINN solver under exponential BC distance function and exponential IC distance function, illustrating the width of the transition zone affected by different selection of α_t values (a)10, (b) 50, and (c) 100.

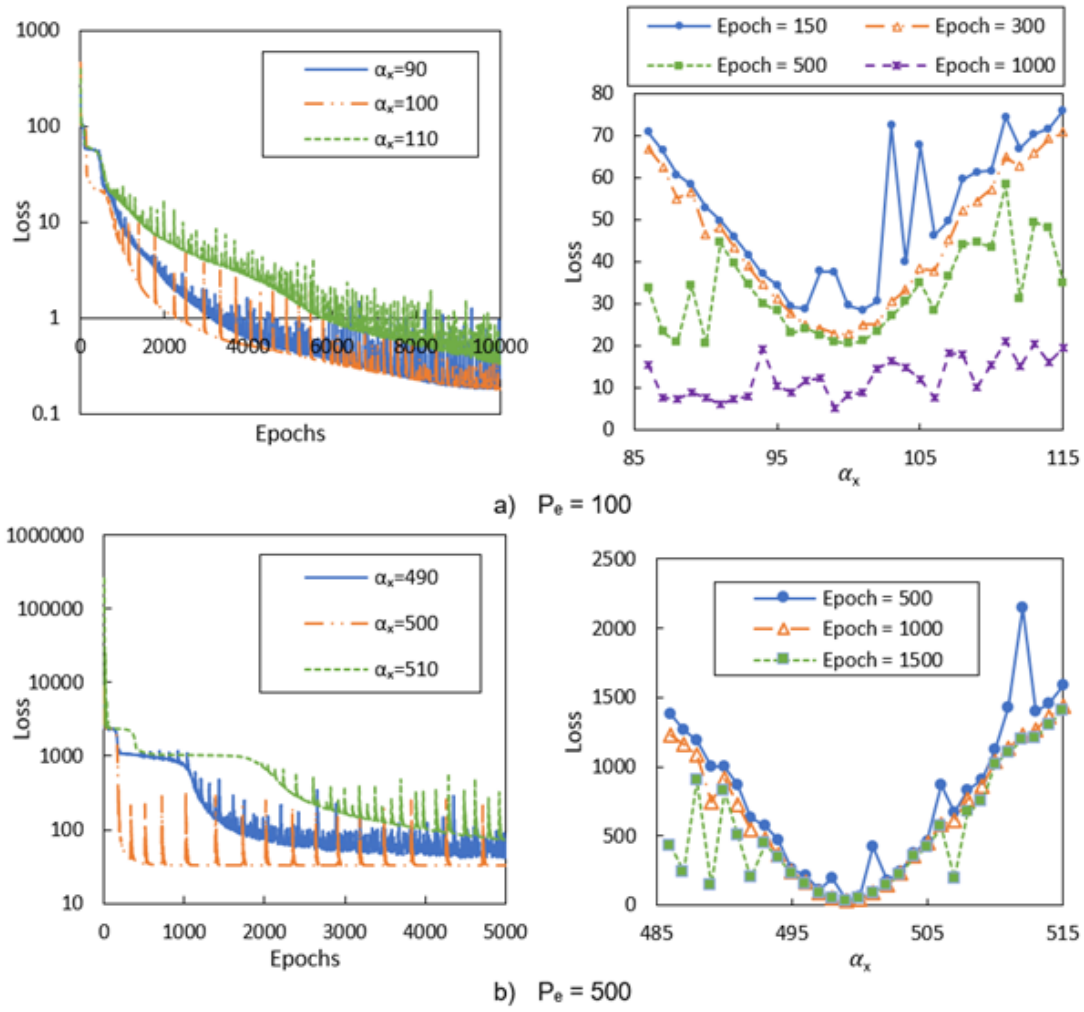


Figure 6.12: Case 2b, results of test 2.9 a) $P_e = 100$ and 2.10 b) P_e =500, showing performance of PINN solver under exponential BC distance function and exponential IC distance function using different values of α_x. Left-hand side figures show loss versus epochs for three values of α_x; right-hand side figures show loss versus α_x at three different epochs.

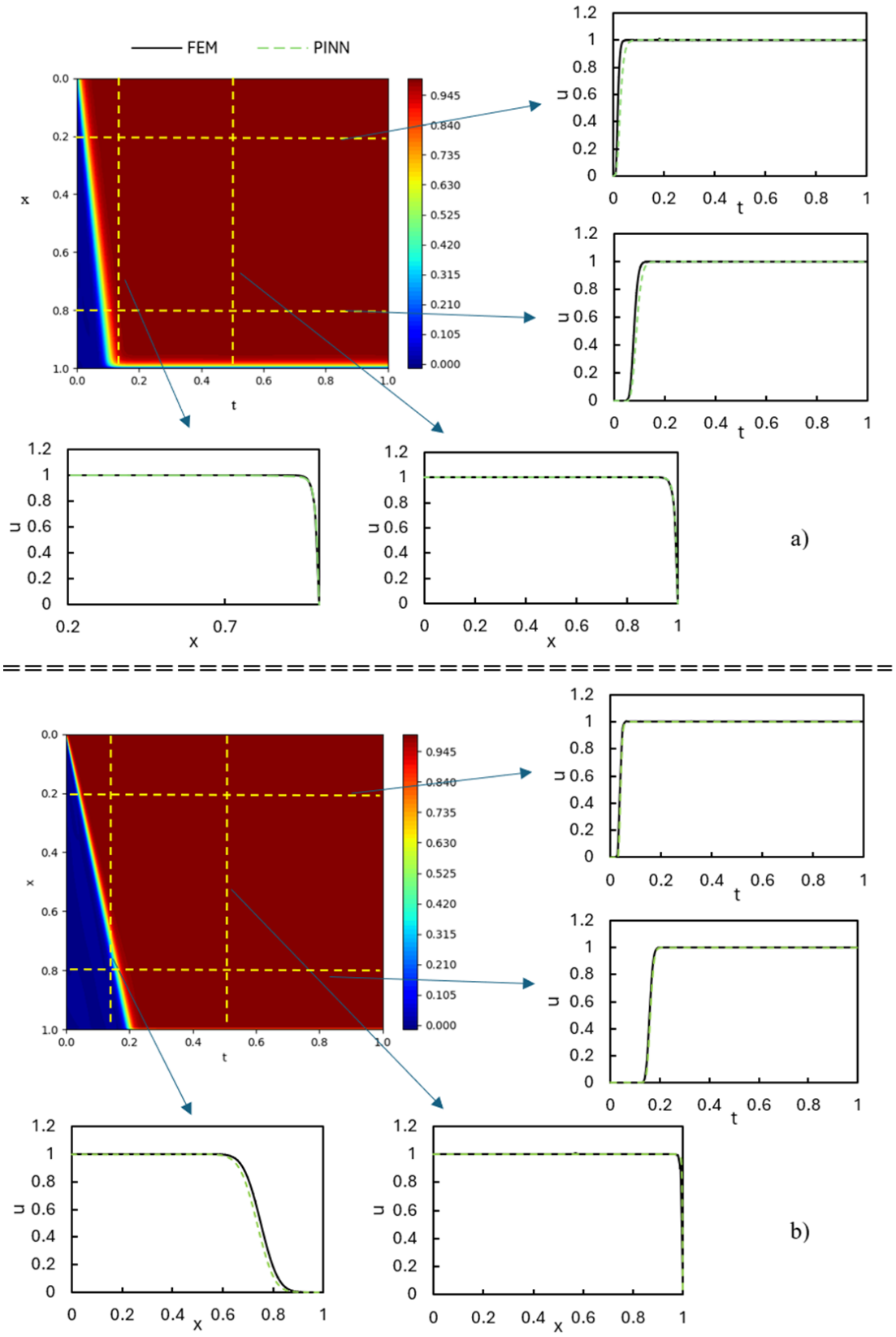


Figure 6.13: Case 2b, comparison of optimal PINN solver results with FEM solutions at selected observation points in both temporal and spatial domains for: a) $P_e = 100$ and b) $P_e = 500$

6.4 Case 3: Transient Case Study with Dirichlet-Neumann Boundary Conditions

Case 3 investigates a one-dimensional transient ADE with a Dirichlet boundary condition at $x = 0 (u = 1)$ and a Neumann boundary condition at $x = 1$ (zero flux). As summarised in Table 6.3, nine tests with Pe ranging from 1 to 1000 were conducted to evaluate the ability of the proposed PINN to solve transient ADE under Dirichlet-Neumann Boundary conditions.

Table 6.3 shows that MAE ranges from 6.89×10^{-3} to 4.94×10^{-2}, while the RMSE spans from 1.82×10^{-2} to 4.94×10^{-2}. The R^2 values for all cases are consistently high, close to 1.0, showing that even as P_e increases and advection becomes more dominant, PINN maintains its predictive accuracy.

Figure 6.14 shows contour plots of $u(x, t)$ from PINN and analytical solution, as well as plots of pointwise absolute error, for Tests 3.1, 3.5 and 3.9, with $P_e = 1, 100$ and 1000, respectively. In all cases, there is a good agreement between PINN analytical solutions across the domain, and the absolute error is relatively low. It is noticeable that near $x = 0, t = 0$, higher absolute errors occur that dissipate over time. This discrepancy arises from the mismatch between the boundary condition at $x = 0$ and the initial condition at time 0. This is similar to behaviour observed in Case 2a.

Table 6.3: Case 3, Discrepancies between PINN and Analytical Solution for Problems with increasing Peclet Numbers under Transient Conditions with Dirichlet-Neumann Boundary Condition.

	Parameters					**Accuracy**		
Test	**D**	**v**	**L**	**T**	P_e	**MAE**	**RMSE**	R^2
3.1	1	1	1	3	1	3.72×10^{-2}	4.94×10^{-2}	0.969
3.2	0.2	1	1	3	5	7.05×10^{-3}	1.91×10^{-2}	0.995
3.3	0.1	1	1	3	10	6.89×10^{-3}	1.82×10^{-2}	0.996
3.4	0.02	1	1	3	50	7.08×10^{-3}	2.02×10^{-2}	0.996
3.5	0.01	1	1	3	100	7.47×10^{-3}	2.36×10^{-2}	0.995
3.6	0.005	1	1	3	200	7.30×10^{-3}	2.54×10^{-2}	0.995
3.7	0.004	1	1	3	250	8.16×10^{-3}	2.89×10^{-2}	0.993
3.8	0.002	1	1	3	500	7.59×10^{-3}	2.90×10^{-2}	0.994
3.9	0.001	1	1	3	1000	9.25×10^{-3}	3.55×10^{-2}	0.991

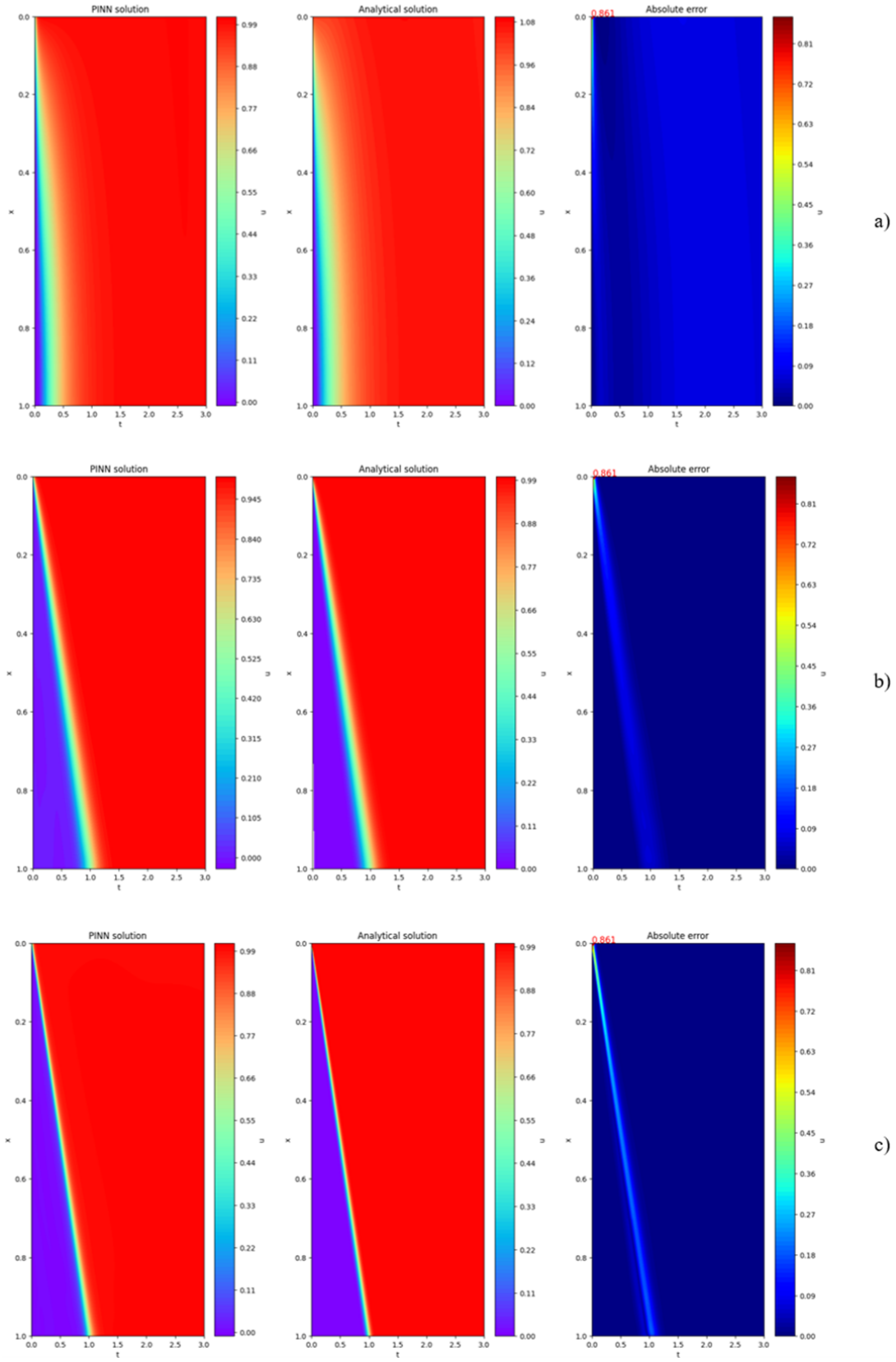


Figure 6.14: Case 3, comparison between PINN solver predictions and analytical solutions, along with pointwise absolute error for a) $P_e = 1$, b) $P_e = 100$, and c) $P_e = 500$.

6.5 Summary

Distance functions in hard-constraints PINN are under-researched in the literature and several distance functions have been systematically explored in this study. Under steady-state, the HC-PINN model with an exponential distance function, using a shape factor equal to the Peclet number is found to produce by far the best performance.

As a result, a new mixed-constraints-PINN model, based on exponential distance functions, is proposed and shown to perform remarkably well in solving the advection-diffusion equation under both steady-state and transient conditions across a wide range of P_e values, from 1 to 1 million. This is a significant improvement relative to current PINN solvers and shows that a mixed-constraint PINN with exponential distance functions can tackle both diffusion/dispersion-dominated and advection-dominated cases. Furthermore, the method is able to capture transient behaviour in cases where there is incompatibility between boundary and initial conditions, as well as cases of mixed, Dirichlet-Neumann, boundary conditions.

However, there are still some limitations of this project. First, there are noticeable fluctuations observed in the loss history of HC-EXP-PINN models. Similar fluctuations during PINN models training have been widely reported in the literature although the reason for these fluctuations is still underexplored. In future research, understanding the mechanisms behind these fluctuations is key to enhancing the training reliability of PINN [295, 296, 297]. Second, Given the versatility of the method, future research should also explore its effectiveness in 2D and 3D space, as well as its ability to capture real-life behaviour characterised by heterogeneous and/or time-dependent seepage velocity fields. In addition, an extension of the method to unsaturated conditions in which the diffusion and/or mechanical dispersion coefficients may depend on water content and hence become time-dependent would be an important step towards a practical application of PINN to practical problems.

This page was intentionally left blank.

Chapter 7

Richard's Equation Modelling with Machine Learning

7.1 Overview

This chapter presents results of the development of PINN solvers for the RRE. Specifically, it evaluates the effectiveness of two candidate approaches for improving the prediction accuracy of RRE PINN solvers, namely hard constraint and adaptive resampling. The first part of the chapter focuses on the hard constraint approach, by comparing its predictions to both analytical and FEM solutions to assess its accuracy. In addition, a comparison is made between the accuracies of hard constraint and soft constraint models to evaluate any gains achieved through a hard constraint approach.

The second part of the chapter presents the results of two adaptive resampling algorithms applied to the hard-constraint PINN solver: RAR and RAD. The evolution of the collocation point distributions over the training intervals is presented and the accuracy of the two models, as well as their effectiveness in improving prediction accuracy are discussed. Finally, a summary of findings is presented, and recommendations for optimal modelling strategies for the PINN-RRE model are made.

7.2 Hard vs. Soft Constraint PINN Performance

The contour plots of model predictions are presented in Figure 7.1, where subfigure (a) shows the results of the soft constraint PINN, and subfigure (b) shows the results of the HC-PINN. Both subfigures illustrate the spatial-temporal distribution of pressure head and volumetric water content. Clear differences can

be observed between the two models. In particular, the HC-PINN predicts a much sharper wetting front compared to the soft constraint PINN. This indicates that the HC-PINN produces a much narrower transition zone between the saturated and unsaturated regions, whereas the soft constraint PINN generates a considerably smoother and wider transition zone.

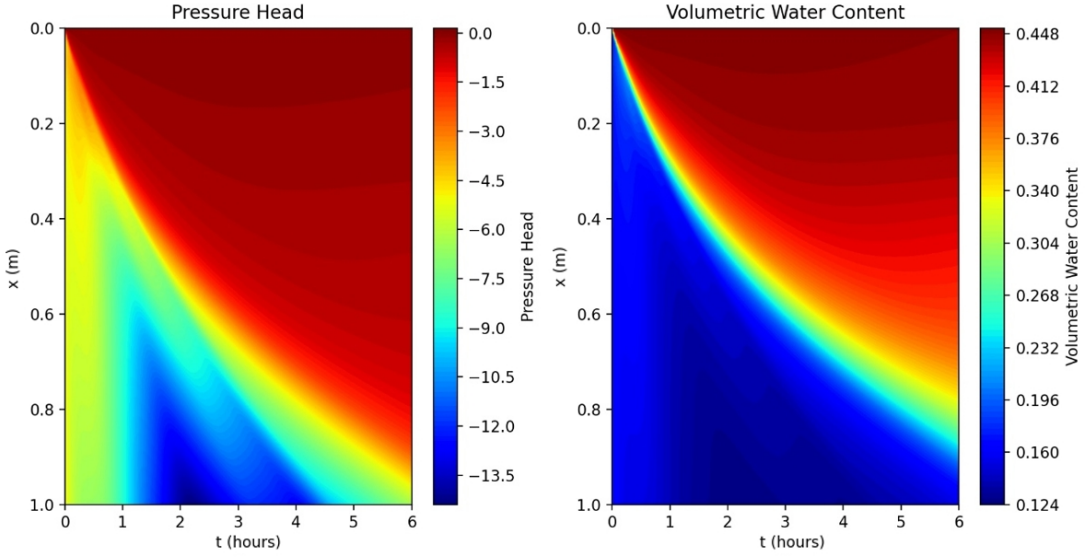


Figure 7.1: Contours of pressure head and volumetric water content predicted by PINN solutions of the Richard's equation: a) baseline PINN with soft constraints b) HC-PINN.

To assess the accuracy of the two models, their predictions are compared

against the reference solution. As explained in the methodology section, the analytical solution is limited to 36 observation points in this thesis and cannot fully evaluate whether the models can accurately capture the sharp wetting front and reproduce the water content behaviour near both the saturated and dry ends of the infiltration profile. Therefore, a combination of analytical observation points and FEM predictions is used as the reference solution for comparison.

First, a comparison between the soft constraint PINN model and the reference solution is shown in Figure 7.2. At early time steps ($t = 1$ to $t = 3$), both the FEM model and the soft-PINN model align well with the analytical data points. As time progresses ($t = 4$ to $t = 6$), the soft-PINN starts deviating from the analytical solution, while the FEM model continues to align well with the analytical data. In addition, across all time steps, near both the saturated and dry ends of the infiltration profile, the soft-PINN fails to reproduce the transition patterns captured by the FEM model, which exhibits much sharper transitions and steeper gradients.

Next, the HC-PINN model is also compared with the reference solutions. The comparison with the analytical solution is shown in Figure 7.3. Except for the initial time step at $t = 0.5$, the HC-PINN predictions closely match the analytical observations at all other time steps. At the available observation points, the HC-PINN solution aligns more closely with the analytical solution than the FEM model. Furthermore, when comparing the full profiles with the FEM solution, the HC-PINN model successfully captures the key features observed in the FEM results across the entire infiltration process. Specifically, the HC-PINN accurately reproduces the steep wetting front transition and properly reflects the water content behaviour near both the saturated and dry ends. This demonstrates that the hard constraint approach enables the HC-PINN to effectively represent both the sharp gradients and the overall infiltration dynamics.

The performance of both models is also captured in Table 7.1, which reports the %Error (Eq. 3.13) at each time step as well as the overall average error.

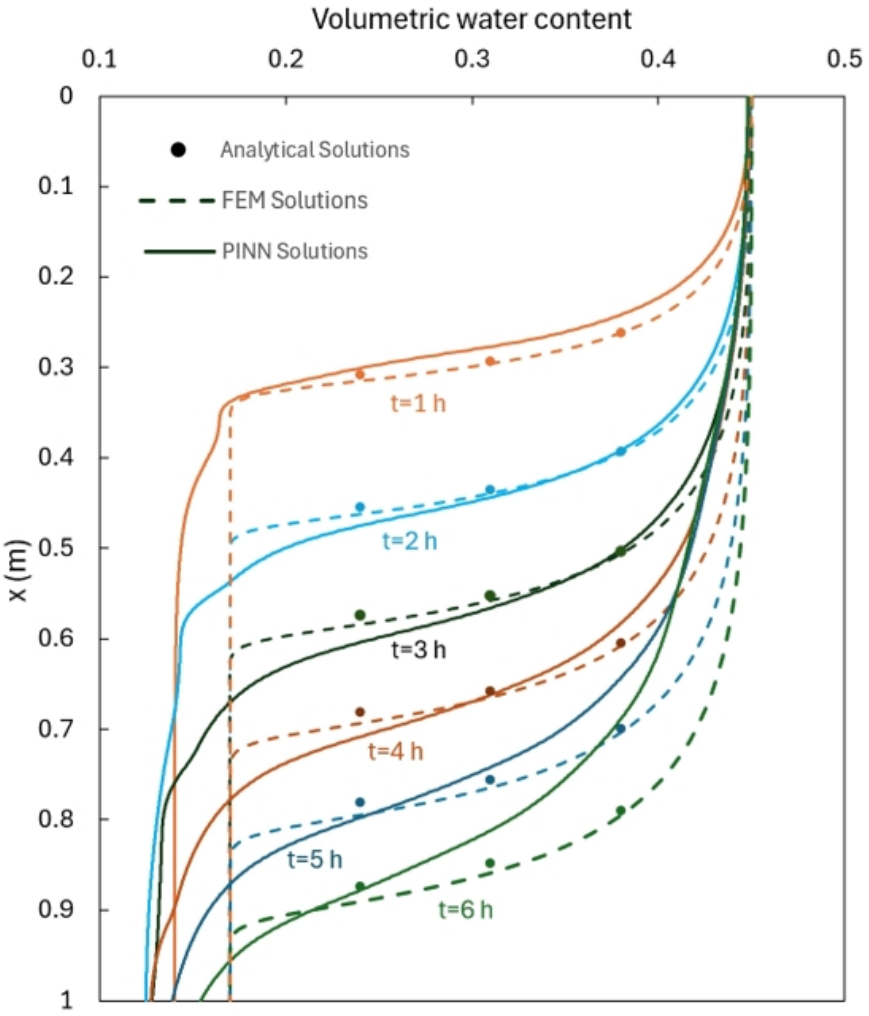


Figure 7.2: Comparison of predictions of analytical solutions, FEM model, and baseline soft constraint PINN model at different time steps ranging from t = 1 to 6.

As shown, the overall percentage error decreases from 4.02% for the soft-PINN model to 1.28% with the introduction of hard constraints. In addition, the error at nearly all time steps is significantly reduced under the hard constraint approach, except at $t = 1.5$, where a slight increase is observed.

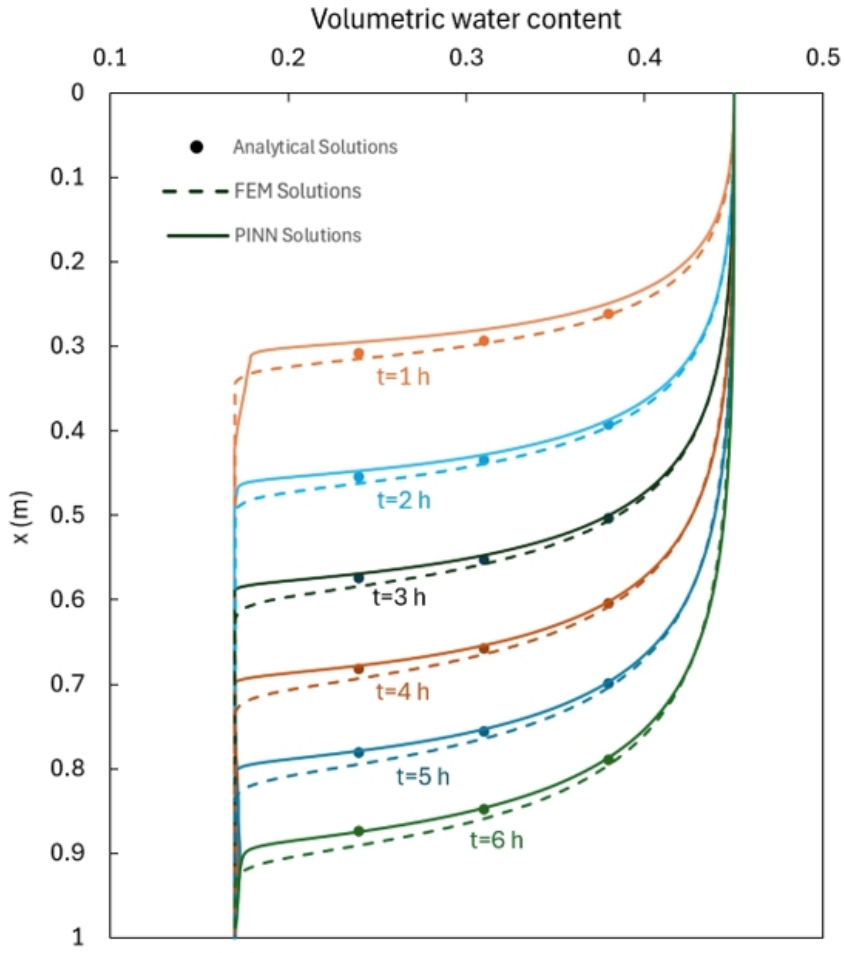


Figure 7.3: Comparison of predictions by analytical solution, FEM model, and HC-PINN model at different time steps ranging from t = 1 to 6.

Table 7.1: Comparison of percentage error between soft constraint PINN and HC-PINN at different time steps.

Time (h)	**Soft Constraint PINN Error (%)**	**HC-PINN Error (%)**
0.5	18.99	13.90
1.0	5.05	4.59
1.5	1.74	2.41
2.0	2.37	1.61
2.5	2.82	1.21
3.0	2.81	0.94
3.5	2.71	0.69
4.0	2.68	0.55
4.5	3.03	0.46
5.0	3.91	0.32
5.5	4.78	0.29
6.0	5.64	0.27
Average	**4.02**	**1.28**

7.3 Effect of Adaptive Resampling on PINN

In the previous section, it has been demonstrated that the application of hard constraints can significantly improve the prediction accuracy of the PINN solver for the Richards equation. Based on this improved baseline, the HC-PINN model is used as the reference framework to further evaluate the effectiveness of two adaptive resampling strategies, RAR and RAD. THe two algorithms have been presented in section 3.6.3. The results of these two methods are presented in the following subsections.

7.3.1 RAR Method

The RAR resampling method is first tested to examine whether it can further improve the accuracy of the PINN solver in solving the Richards equation. Figure 7.4 illustrates the evolution of collocation points under different k_1 values at various training stages. In this figure, the first column shows the initial distribution of training collocation points, the second column presents the distribution after the first resampling stage, and the third column shows the distribution after the second resampling stage. Each row corresponds to a different k_1 value, with k_1 increasing from 0 to 1 from top to bottom.

The topmost row represents a special case where $k_1 = 0$, meaning full replacement is performed at each resampling stage, resulting in purely uniform sampling across the spatial-temporal domain. In this case, the distribution remains random and is not influenced by the residual distribution. As k_1 increases, a larger proportion of collocation points with higher PDE residuals are retained after each resampling. This leads to a visible accumulation of collocation points along regions where residuals are higher. For cases with $k_1 = 0.6$ and 0.8, a clear accumulation pattern is observed, where points increasingly cluster along high-residual zones, while certain regions, particularly in the lower-left corner, exhibit lower sampling density. Finally, at $k_1 = 1.0$, another special case is observed where no resampling is performed, and the collocation point distribution remains

identical throughout all resampling stages.

The overall prediction accuracy of the RAR method for different k_1 values is presented in Figure 7.5. The red dashed line represents the baseline HC-PINN error, which is 1.28%. It can be observed that five cases achieve lower prediction errors than the baseline, all of which correspond to k_1 values smaller than 0.5. The special case of $k_1 = 0$, full replacement, yields the best performance with the lowest prediction error. When comparing the cases with k_1 values between 0.1 and 0.4, it is noticeable that partial adaptive resampling does not further improve the performance beyond the full replacement case. Although these cases still achieve slightly lower errors than the HC-PINN baseline, the improvement is likely attributed more to the benefit of periodically refreshing the collocation points rather than to the adaptive concentration guided by residual information. In other words, simply updating the training samples across the entire domain at certain intervals contributes to better generalisation, while adaptive resampling based on residuals does not provide additional gains.

Furthermore, as k_1 increases beyond 0.5, the prediction errors begin to rise and consistently exceed the baseline HC-PINN error. This suggests that increased attention on high-residual regions may lead to insufficient global coverage, causing the model to lose its ability to represent the full solution domain. Therefore, based on the results of the RAR method, maintaining sufficient global coverage through unbiased resampling appears to be more important for improving the prediction performance. The use of the RAR method (except the special case with $k_1 = 0$) does not contribute positively to the overall model accuracy.

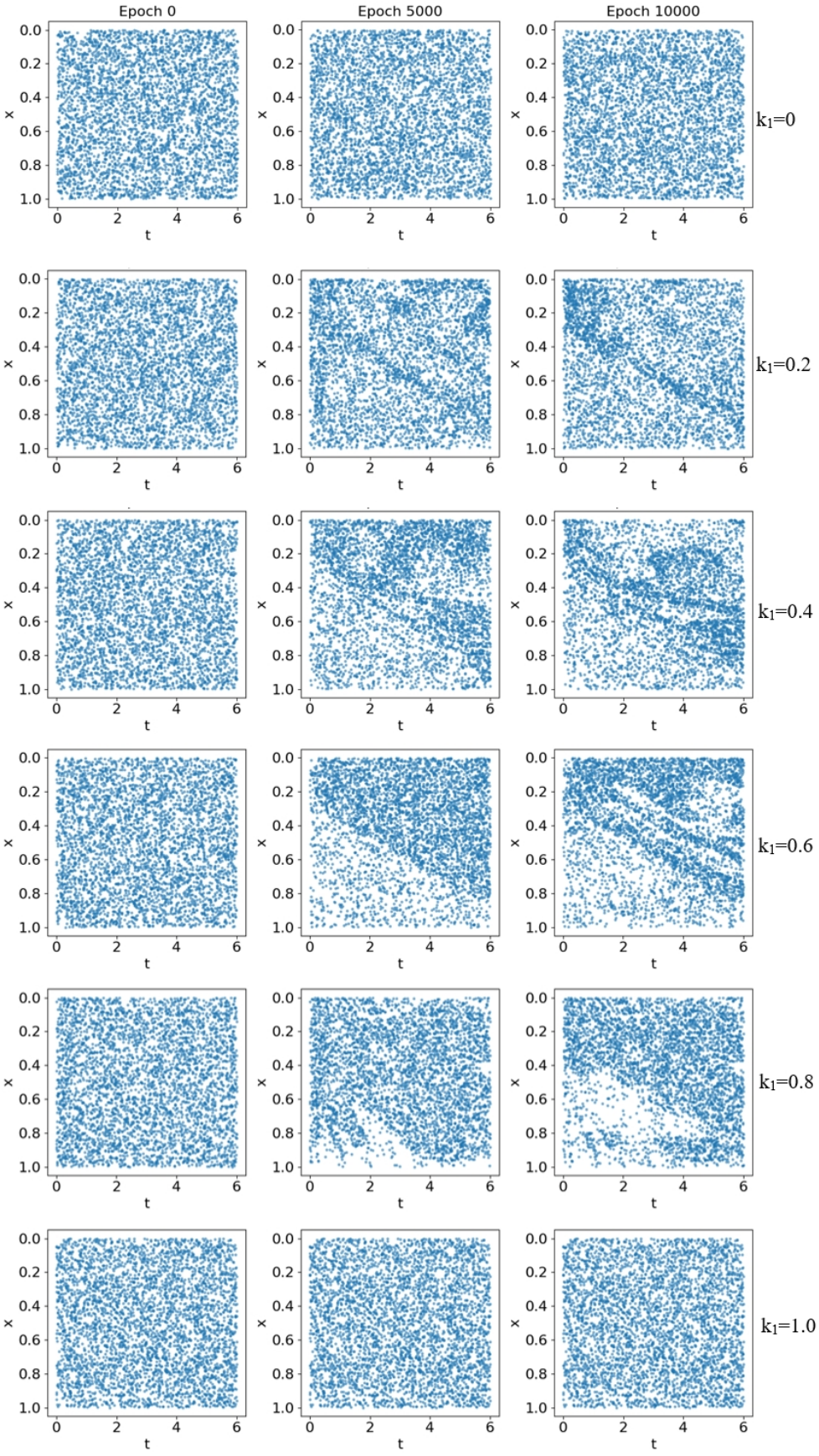


Figure 7.4: Evolution of collocation points during training using the RAR resampling method. The first column shows the initial distribution; the second and third columns show the distributions after the first and second resampling, respectively. From top to bottom, five cases are presented with increasing k_1 values from 0 to 1 in increments of 0.2.

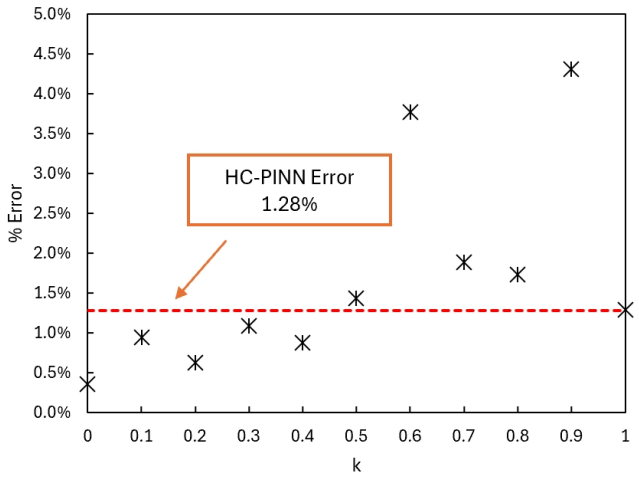


Figure 7.5: Comparison of percentage errors for eleven RAR resampling cases with k_1 values ranging from 0 to 1, shown against baseline HC-PINN model.

7.3.2 RAD Method

The RAD method is also investigated to assess its impact on prediction accuracy. Figure 7.6 illustrates the evolution of collocation point distributions under different k_2 values at various training stages. Similar to the presentation in the RAR result section, the first column shows the initial uniform distribution of collocation points, and the second and third columns display the distribution after the first and second resampling stages, respectively. Each row corresponds to a different k_2 value, ranging from 0.5 to 2.5.

As shown in the figure, when k_2 increases from 0.5 to 1.5, a clear accumulation pattern begins to emerge along the wetting front region, which is consistent with the wetting front location observed in the contour map presented in Figure 7.1b. When k_2 is further increased to 2.0 and 2.5, the collocation points become highly concentrated within localised high-residual regions, while other regions have reduced sampling density.

The prediction errors for all RAD cases with different k_2 values are presented in Figure 7.7. As shown in the figure, all five cases result in much higher %Error compared to the baseline HC-PINN model. In other words, the application of the RAD method does not lead to improved prediction performance. Consistent with the discussion in the previous subsection, these results further strengthen the importance of maintaining sufficient generalisation and diversity in the collocation point distribution. In all RAD cases, the selection of collocation points biased towards high-residual regions leads to an increase in the overall %Error. The failure of adaptive resampling observed aligns with the findings of Daw et al. [274], which show that too many collocation points in high residual regions can result in other areas being under sampled. In the case of the RRE solver, the sharp gradients tend to draw sampling points into narrow regions. Due to the differences in resampling strategies, this problem is more significant when using RAD than RAR. The research from Gao et al. [298] also supported this point.

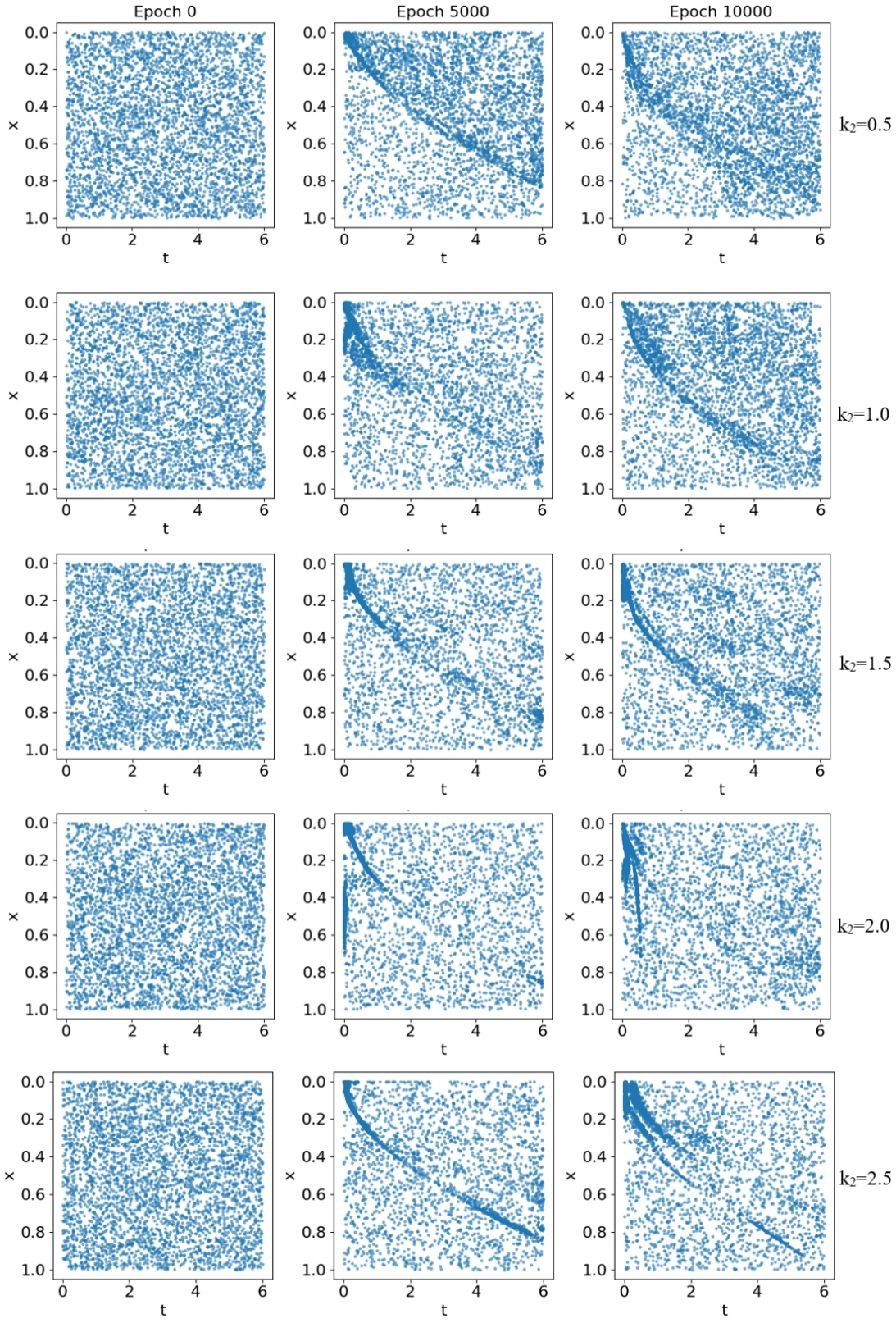


Figure 7.6: Evolution of collocation points during training using the RAD resampling method. The first column shows the initial distribution; the second and third columns show the distributions after the first and second resampling, respectively. From top to bottom, five cases are presented with increasing k_2 values from 0.5 to 2.5 in increments of 0.5.

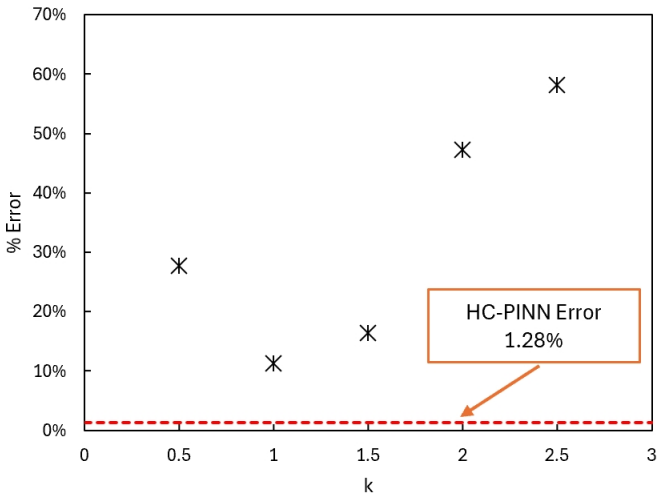


Figure 7.7: Comparison of percentage errors for five resampling cases with k values ranging from 0.5 to 2.5, along with a comparison to the performance of the HC-PINN model.

7.4 Comparison of Different Resampling Strategies

As shown in Table 7.1, although the HC-PINN model achieves strong overall performance with a relatively low total %Error of 1.28%, it performs not so well at early time steps, with errors of 13.90% at $0.5h$ and 4.59% at $1.0h$. As discussed in Section 7.3.1, incorporating the RAR strategy into HC-PINN with $k_1 < 0.5$ can result in lower overall errors compared to the baseline HC-PINN. In contrast, in Section 7.3.2, applying the RAD strategy with HC-PINN across all tested values of k_2 from 0.5 to 2.5 yields no improvement and shows a much higher %Error. To better understand these behaviour when applying adaptive resampling method, we show the predicted water content profile across the entire soil column at different time steps with the different resampling configurations with HC-PINN in Figure 7.8. In Figure 7.8, five sets of solution profiles are presented for each of the six selected time steps. These include the baseline HC-PINN model, three cases of HC-PINN with RAR using $k_1 = 0$, 0.5, and 0.9, as well as the HC-PINN model with RAD using $k_2 = 1$. Subfigures (a) to (f) correspond to time steps from $t = 0.5h$ to $t = 5.0h$.

From Figure 7.8, the first noticeable observation is that the RAR $k_1 = 0.9$ model consistently underestimates the solution across all time steps, and the prediction near the dry-end boundary is also not accurate. With the exception of the RAR $k_1 = 0.9$ model, the predictions of all remaining resampling models at $t = 0.5h$ are closer to the analytical solution than the baseline HC-PINN. At $t = 1.0h$, the RAD $k_2 = 1$ model shows the best agreement with the analytical solutions, while the predictions of RAR $k_1 = 0$ and 0.5 models are very similar to the HC-PINN baseline. Although the RAD $k_2 = 1$ model performs well at $t = 0.5h$ and $1.0h$, clear differences from the baseline start to appear at $t = 2.0\,h$, and the gap becomes larger at later time steps. The RAR $k_1 = 0$ and 0.5 models remain consistently close to the baseline across all time steps. As shown in Table 7.1, the HC-PINN already achieves very low %Error at these later stages, so these models are also accurate.

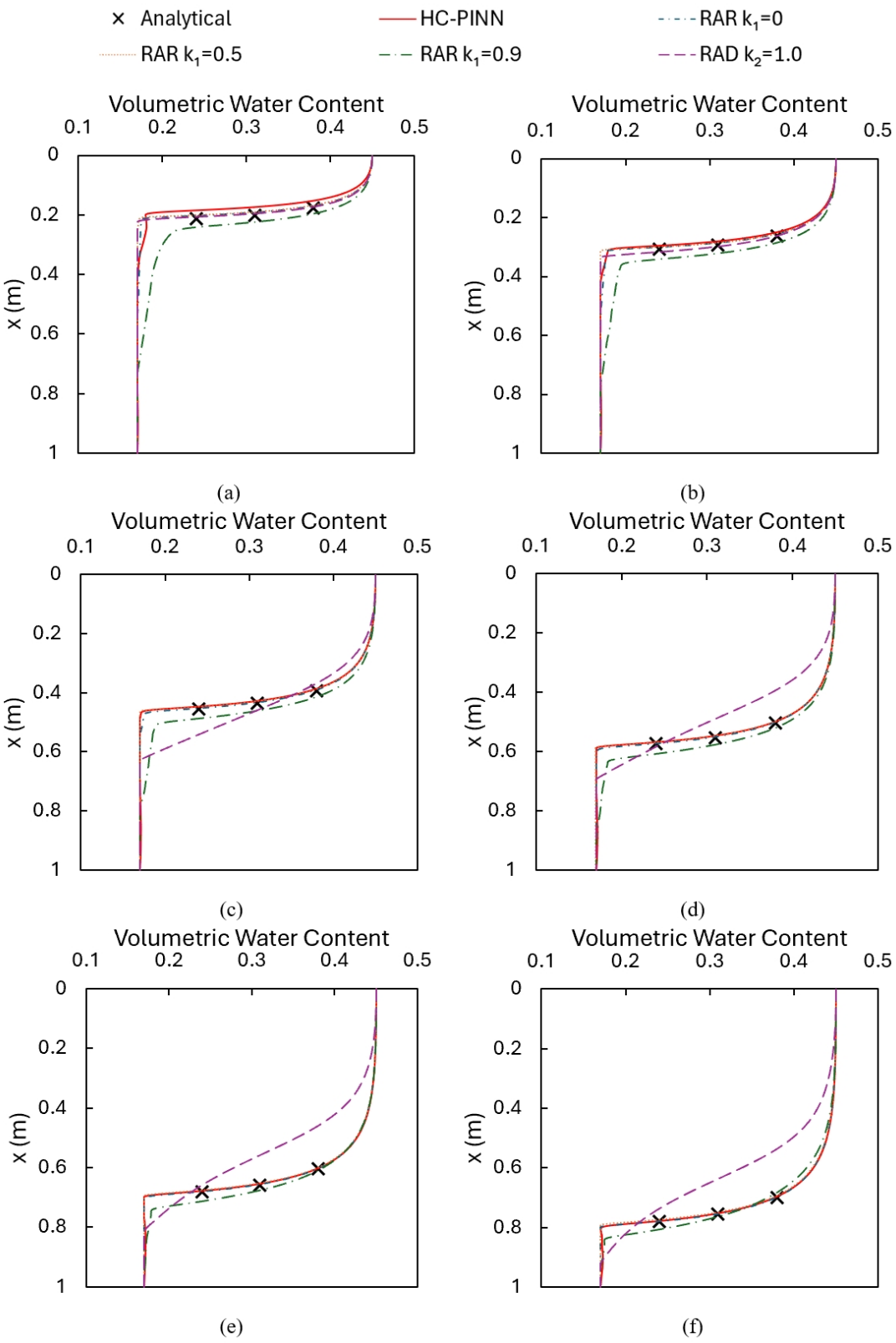


Figure 7.8: Comparison of solutions obtained by HC-PINN, HC-PINN with RAR ($k_1 = 0, 0.5, 0.9$), and HC-PINN with RAD ($k_2 = 1$) at (a) $t = 0.5h$, (b) $t = 1.0\,h$, (c) $t = 2.0\,h$, (d) $t = 3.0\,h$, (e) $t = 4.0\,h$, and (f) $t = 5.0\,h$.

7.5 Summary

In this chapter, the hard constraint and adaptive resampling method for solving the RRE have been systematically evaluated. First, the application of hard constraints has shown a significant improvement in the prediction accuracy compared to the soft constraint PINN. As the comparison with the combined analytical and FEM reference solution, the HC-PINN model reproduces the sharp wetting front and aligns well with the available analytical observation points, outperforming the FEM solver.

Following the application of the hard constraint, two adaptive resampling strategies, RAR and RAD, have been tested to assess whether they can enhance model performance. However, only a special case of the RAR method, where $k_1 = 0$, shows clear improvement in prediction accuracy. In this case, all collocation points are periodically replaced by uniformly sampled points from the entire domain without utilising any residual information. For all other cases where sampling is biased toward high-residual regions, no performance improvement is observed. These results indicate that maintaining sufficient global coverage in the collocation sampling is essential for ensuring the prediction accuracy of RRE.

In summary, the combination of hard constraints and periodic full replacement across the entire domain is recommended to improve the prediction accuracy of the RRE-PINN solver. However, it must be acknowledged that the resampling approach in this thesis is applied in combination with the hard constraint framework. Our primary objective is to evaluate whether adaptive resampling can bring additional benefits when hard constraint is already been implemented. If the HC framework is not used, whether adaptive resampling can improve the performance of a soft-constrained model for RRE solver remains unclear. The behaviour of the resampling approaches observed in this infiltration problem should not be taken as evidence against their potential applicability to a broader range of problems. Its effectiveness may vary depending on the nature of the problem. In future work, it is necessary to investigate more problems with sharp gradients in the

solution field to assess the effectiveness of this approach to this type of problems. Another limitation is that the hydraulic characteristics of the soil are described by using VG WRC model and VG HCF model when solving Richard's Equation In future research, the selection of different empirical models and how their associated uncertainties propagate the performance of the RRE solver is worth investigating.

Chapter 8

Coupling of ADE and RRE model

8.1 Overview

This chapter presents the application of the PINN framework for simulating coupled water (RRE) and solute transport (ADE) in unsaturated soil. The aim is to make a preliminary assessment as to whether PINN is capable of modelling this coupled set of equations by testing a relatively simple problem.

A problem in which a partially dry soil is infiltrated by polluted water is simulated. The simulation is conducted in a one-dimensional vertical soil column with a total depth of 1 m and a total simulation time of 6 hours. The soil hydraulic properties are given in the form of VG WRC model (Eq 3.5) and HCF model (Eq 3.6) with the following parameters: residual water content $\theta_r = 0.10$, saturated water content $\theta_s = 0.45$, shape parameters $\alpha_{vg} = 1m^{-1}$, $n_{vg} = 2$, $m_{vg} = 1–1/n_{vg} = 0.5$, saturated hydraulic conductivity $K_s = 0.0216m/h$. For the water movement component, the initial condition for volumetric water content of the soil column profile is set as $\theta = 0.17$. The boundary condition at the soil surface ($x = 0\,m$) is maintained at saturated water content ($\theta = 0.45$), while the bottom boundary ($x = 1\,m$) is assumed to be free drainage. For the solute transport component, the initial concentration is $u = 0kg/m^3$ across the soil column, and the boundary condition of the top surface ($x = 0\,m$) is set as $u = 1\,kg/m^3$. The longitudinal dispersivity coefficient is $D_L = 0.1$ m, while molecular diffusion is neglected because mechanical dispersion and advection are assumed to dominate the solute transport process under the simulated flow conditions.

The evolutions of the key state variables, including pressure head, water con-

tent, water flux, and solute concentration, are presented in a series of contour plots. Then, PINN predictions are compared with a reference solutions obtained from the Hydrus FEM model for all state variables. The comparisons are conducted at five selected depths of the soil column (x = 0 m, 0.25 m, 0.5 m, 0.75 m, and 1.0 m) to evaluate the model's performance.

8.2 Evolution of State variables

With the model parameters and boundary conditions defined, the coupled water and solute transport processes were simulated by using the PINN framework. The hard constraint approach was applied to both water flow (governed by RRE) and solute transport component (governed by ADE). For the solute transport component, the newly developed Peclet-adjusted exponential distance function was also used. However, in this problem setup, the water flux varies over space and time and resulting in a non-uniform water flux profile in the soil column, and hence a variable Peclet number. Hence, a value of $\alpha_x = \alpha_t = 100$ was used for solving the unsaturated ADE with the hard constraint method. For the water flow component, a linear distance function was applied to the spatial domain and an exponential distance function with $\alpha_t = 100$ was used, consistently with the convergence observed in the RRE-only PINN model.

To understand the overall dynamics of this coupled process, Figure 8.1 presents the contour plots showing the evolution of four key state variables, including pressure head, volumetric water content, water flux, and solute concentration. Figure 8.1a shows the pressure head distribution in the spatial-temporal domain. The pressure head distribution can be further converted into the corresponding evolution of volumetric water content, as shown in Figure 8.1b. The wetting front is clearly observed, with a sharp transition zone between the saturated and unsaturated regions. As water infiltrates deeper soil, the transition zone gradually broadens but remains relatively confined. Figure 8.1c shows the water flux contours. In the initial time steps, due to the water content gap between the top

surface ($x = 0m$) boundary condition and the relatively dry initial setup, a high water flux is observed near the soil surface shortly after infiltration begins. As water infiltrates the soil column, the water flux starts to stabilise. As the linking variables in the coupled system, the predicted volumetric water content (Figure 8.1b) and water flux (Figure 8.1c) provide the necessary input information for solving the ADE in the coupled analysis. The distribution of solute concentration in the domain is presented in Figure 8.1d. At the initial time steps, mild numerical oscillations in the solute distribution are observed near the soil surface. This seems to coincide with areas with potentially high Peclet numbers and may be due to non-optimal choice of shape parameters for the distance functions. As time progresses, oscillations are diminished, and the solute gradually moves from the soil top surface ($x = 0m$) to the deeper soil over time.

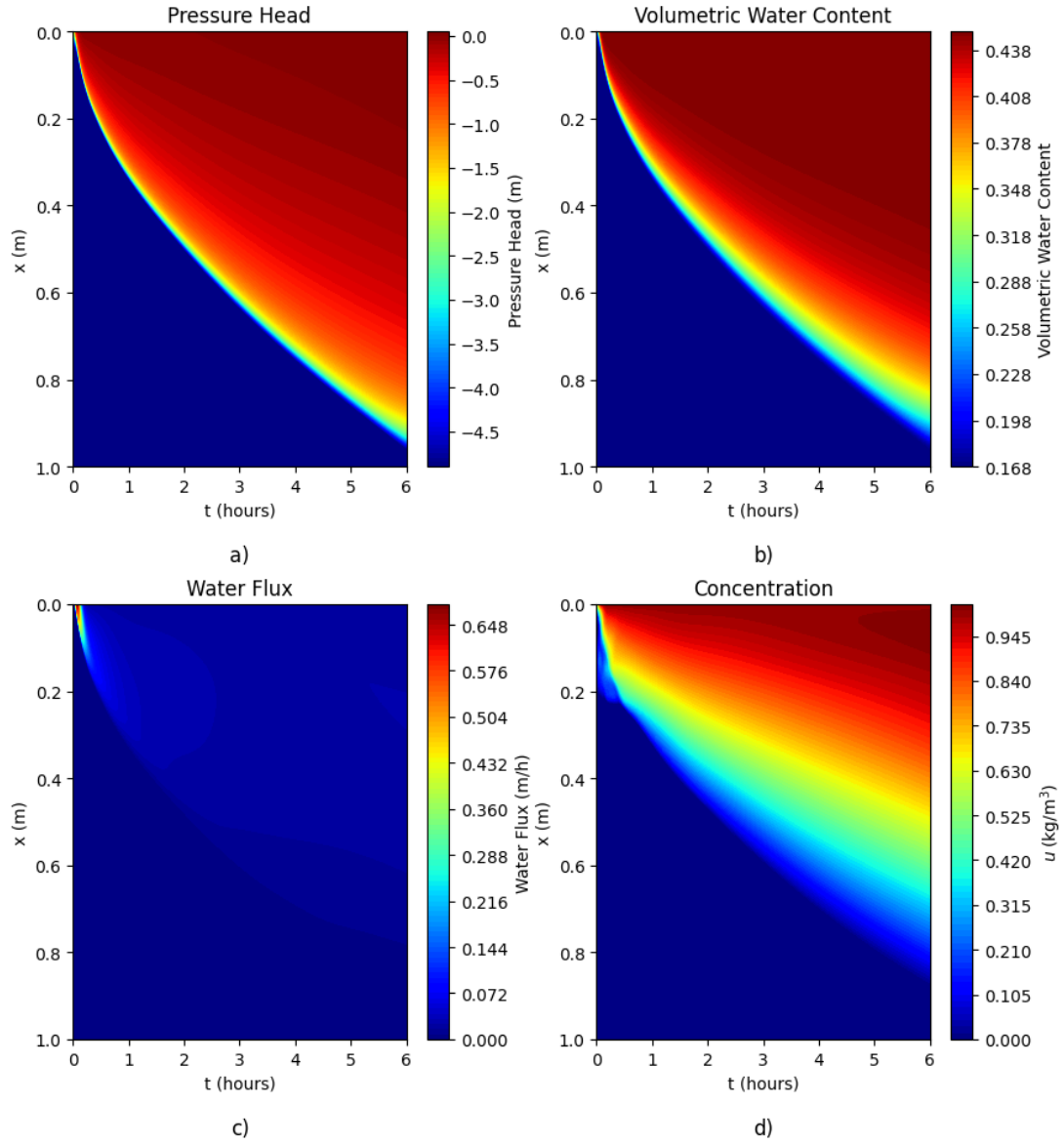


Figure 8.1: Spatial-temporal domain evolution of four state variables reproduced by the coupled PINN model. (a) Pressure head (m); (b) Volumetric water content (m^3/m^3); (c) Water flux (m/h); and (d) Solute concentration (kg/m^3).

8.3 Model Prediction Comparison with FEM Solutions

Figure 8.2 presents the PINN predictions of the temporal evolution of the four state variables at five elevations, along with their comparisons against the FEM (Hydrus-1D) reference solutions. The evolution of pressure head is shown in Figure 8.2a. The PINN solutions closely match the FEM results at the soil surface ($x = 0\,m$) and at shallow depth ($x = 0.25\,m$). However, as the wetting front propagates to deeper layers, discrepancies between the PINN and FEM solutions become slightly larger but remain close. A similar trend is observed in the evolution of volumetric water content, as shown in Figure 8.2b. The agreement between the two models is very good at the surface and declines somewhat at further depth.

For water flux (Figure 8.2c), a logarithmic scale is used on the vertical axis to better capture the wide range of values. One of the most noticeable differences between PINN and FEM solutions occurs near the soil surface at early time steps. Due to the initial water content gap between the imposed boundary condition and the initial water content, the FEM solution shows a very sharp drop from a very high initial flux to a relatively stable lower value. In contrast, the PINN prediction exhibits a rapid initial increase, followed by a sharp decrease before reaching the stable value. Among the five depths, the best agreement between PINN and FEM is observed at $x = 0.25m$, followed by $x = 0.5m$ and $x = 0.75m$. For the $x = 1.0m$ case, the wetting front has not reached this depth within the simulation period, and both the PINN and FEM models correctly predict zero values, which is physically consistent with the infiltration process.

Finally, the solute concentration comparisons are shown in Figure 8.2d. The overall trend of solute transport is consistent with the patterns observed in the water flow variables. However, the discrepancies between PINN and FEM are more noticeable for solute concentration than water content and water flux. This may be caused by the accumulation of linking variables in the one-way coupling mechanism. The two linking variables, water content and water flux, may provide

imperfect information to the ADE due to inaccuracies in their predictions. As the coupled simulation progresses, the accumulation of errors from both variables propagates into the solute concentration predictions, making the discrepancies more obvious in the solute transport results.

However, Figure 8.2d does not show the mild numerical oscillations observed in the solute distribution near the soil surface at the initial time steps, as seen in Figure 8.1d. The oscillations occurred between the boundary ($x = 0\,m$) and the observation elevation ($x = 0.25\,m$), and there are no additional observation points placed between them, so that the oscillations are not captured in Figure 8.2d. To better illustrate the spatial and temporal characteristics of these oscillations, Figure 8.3a presents the concentration distribution at various depths near the soil surface ($x = 0$, $x = 0.05$, 0.1, 0.15, and $0.2\,m$), while Figure 8.3b shows the concentration profiles at early time steps ($t = 0.1$, 0.2, 0.3, 0.4, and $0.5\,h$). As shown in the figure, the mild oscillations can be observed in both space and time.

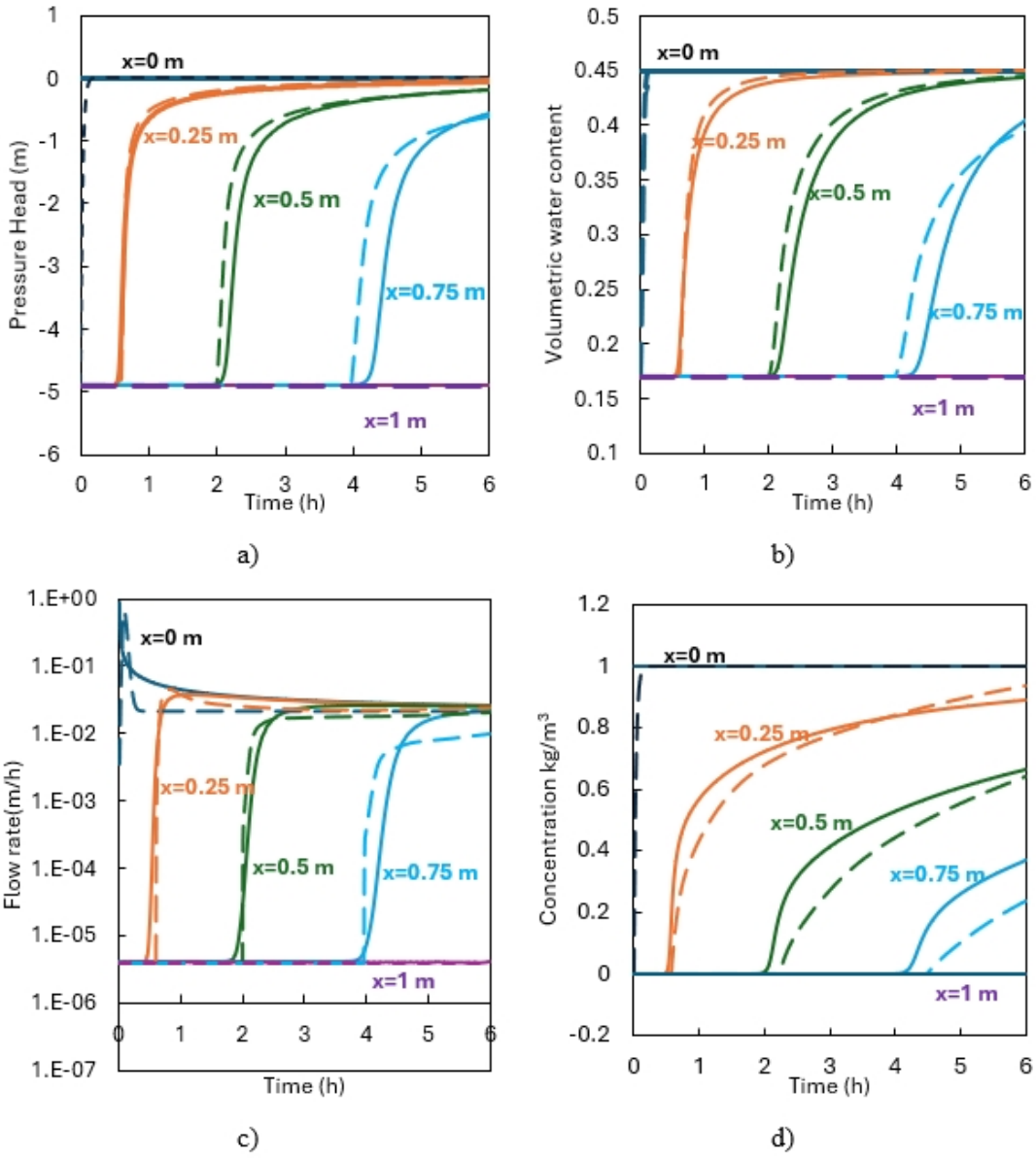


Figure 8.2: Comparison of model predictions between the PINN predictions and the FEM results generated by HYDRUS) at five elevations (x = 0 m, 0.25 m, 0.5 m, 0.75 m, and 1.0 m) for four state variables: (a) pressure head h, (b) volumetric water content θ, (c) water flux q, and (d) solute concentration u. *Notes:* dashed lines represent PINN predictions; Solid lines represent FEM solutions.

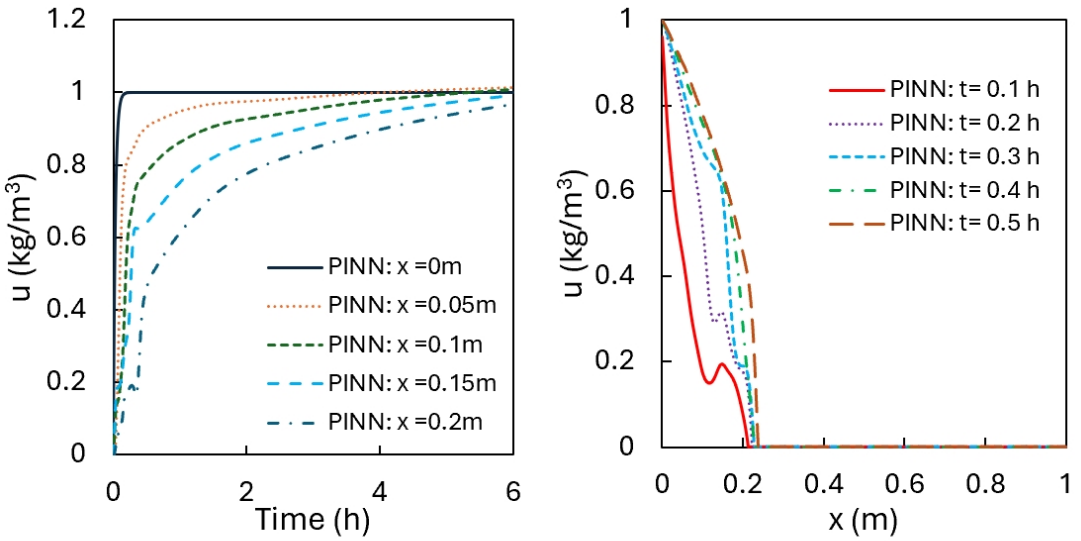


Figure 8.3: The distribution of concentration profile along the additional a) Spatial observation points($x = 0$, $x = 0.05$, 0.1, 0.15, and $0.2\,m$); b) Temporal observation points ($t = 0.1$, 0.2, 0.3, 0.4, and $0.5\,h$)

8.4 Summary

In this section, a coupled PINN framework was applied to simulate water flow and solute transport in unsaturated soil. The evolution of four key state variables, pressure head, water content, water flux, and solute concentration, was first presented through a series of contour plots. Then, detailed comparisons between PINN predictions and FEM solutions were conducted at five selected depths in the soil column (0, 0.25, 0.5, 0.75, and 1.0 m).

Overall, the simulation has demonstrated the potential for using PINN to solve the coupled RRE-ADE problem. Coupling adopted here is in one direction, with flow rate and water content calculated by RRE transferred to the ADE. Hence, uncertainties in both hydraulic outcomes will propagate into the ADE, making the errors in solute transport prediction larger.

This promising start clearly requires further investigation. Specifically, the following two research pathways can be followed:

a) further improve PINN prediction accuracy for the coupled RRA-ADE by investigating non-linear distance functions when water flux varies in time and space. b) evaluate the ability of the proposed coupled PINN to solve more complex problems (e.g., more complex boundary conditions, multi-layered soils) c) extend the proposed coupled PINN to 2D and 3D space.

This page was intentionally left blank.

Chapter 9

Conclusions and Future Research

9.1 Conclusions

9.1.1 Review of Research

This thesis has investigated the use of AI techniques for simulating water flow (RRE) and solute transport (ADE) in unsaturated soils. Water and solute movements in porous media are important in a wide range of environmental and engineering applications, including water resources management, groundwater protection, and prevention and remediation of soil contamination. Traditional numerical methods, such as FEM and FDM, have been widely applied in this field. However, these methods often have difficulties when dealing with nonlinear governing equations, highly dynamic behaviours, and situations where observational data are limited, and those three factors are all present in the problems considered here.

In this context, AI-based modelling approaches provide an alternative pathway that integrates data-driven learning with the underlying physical mechanisms. In this study, AI-driven frameworks have been developed, optimised and evaluated to address the challenges of simulating water and solute transport in unsaturated soils.

In addition, WRC, which characterises the hydraulic behaviour of unsaturated soils by describing the relationship between soil water content and matric potential, plays an important role in the analysis and modelling of unsaturated flow. However, obtaining accurate WRC parameters is often challenging, as it typically requires time and resource consuming laboratory experiments. To overcome this

limitation and improve the availability of soil hydraulic data, an AI-based predictive model, referred to as WRC-PTF, has been developed to estimate WRC parameters efficiently. The developed model offers an efficient approach to parameter estimation and delivers input data for subsequent water flow simulations.

9.1.2 Research Work and Contributions

This thesis made several original contributions. First, A machine learning-based predictive model was developed for estimating WRC parameters, achieving competitive accuracy compared to existing models. The first part of this work is presented in Ding et al, [299], where a general-purpose AI-based PTF model was developed using a large international soil dataset, demonstrating strong predictive capability across a wide range of soil textures. In addition to evaluating model accuracy, the study identified key input features that contribute to accurate WRC predictions. It was found that models incorporating texture variables and either porosity or dry density are sufficient to generate reliable predictions, thereby simplifying input requirements while maintaining model performance. This study also investigated the effect of model complexity by allowing deeper neural network structures, and recommended giving the ANN model the flexibility to exceed one hidden layer, treating the number of hidden layers as an additional hyperparameter to be tuned during model development. In addition, a clay-specific model was constructed using a dedicated clayey soil dataset to further enhance WRC predictions for fine-textured soils, as presented in Ding et al. [300]. This targeted approach improved prediction accuracy for clayey soils, achieving an average 20% improvement compared to the model trained on the full dataset, in which clayey soils accounted for only 10.7% of the total samples.

Secondly, a PINN model was constructed for solving the ADE to simulate solute transport processes. In particular, this study addressed a key challenge associated with high Peclet number conditions, where solute transport becomes strongly advection-dominated. This scenario leads to steep concentration fronts

near boundaries. Such features often result in numerical oscillations and convergence difficulties in conventional numerical methods. A new PINN algorithm was proposed that can accurately solve, for the first time, both steady-state and transient ADE over a large Pe range. A novel method combining hard-constrained PINN with a newly introduced Peclet-adjusted exponential distance function was proposed. The shape factor of the exponential distance function was dynamically adjusted based on the P_e value of the problem to achieve best performance. Extensive analysis was conducted to assess the extent to which the method enhances model stability and accuracy, and to identify optimal relationships between shape factor and P_e values. The proposed solver performed consistently well for a wide range of P_e values from 1 to 1 million, successfully tackling both Dirichlet-Dirichlet and Dirichlet-Neumann boundary conditions in steady-state and transient scenarios.

In addition, a PINN model was developed to solve RRE for simulating unsaturated soil water flow processes. The nonlinear nature of the RRE and its sensitivity to both boundary and initial conditions often create challenges for PINN training, particularly in balancing the contributions of different loss terms. To address the issue of loss competition during training, a hard-constraint formulation was also introduced into the model architecture, consistent with the ADE solver developed earlier. In this approach, the initial and boundary conditions were first reformulated into pressure head form and then directly embedded into the network solution structure as hard constraints. The incorporation of these hard constraints allowed the developed PINN-RRE model to accurately solve the Richards equation while simultaneously enhancing training stability. By reducing the number of loss terms to a single residual loss, this method eliminated competition among loss components and removed the problem of the model's sensitivity to weight selection when a soft-constraint approach is followed. The developed model successfully captured the nonlinear dynamics of unsaturated flow and produced accurate predictions for the test case examined in this study.

Finally, building on the RRE and ADE developments, a PINN solver for the coupled water flow and solute transport problem was presented for the first time. The one-way coupling was established through shared variables, including water flux and water content distributions, in space and time. The two variables were computed by the RRE PINN solver then input into the ADE PINN solver. A relatively simple test case was used to evaluate the performance of the coupled PINN solver and reasonable agreement was found between PINN predictions and an FEM solution. While further refinement and validation are no doubt required before PINN can be deployed to solve this kind of coupled problem, results from this thesis are a promising start.

9.1.3 Limitations

Research in this thesis was constrained by several limitations in each of the four core developments (WRC prediction, PINN ADE solver, PINN RRE solver and PINN coupled RRE-ADE solver).

In the estimation of WRC parameters, the developed model was data-driven, meaning that the projection between input features and predictions depends heavily on the quality and completeness of the available data. Although the dataset used in this study is reputable and provides wide global coverage, certain limitations remain. First, there is an imbalance in the geographic distribution of samples, as some regions of the world are underrepresented due to differences in research activity and economic development, resulting in fewer available data entries. Second, inconsistencies in data formats, unit systems, and measurement standards across different national databases and laboratory protocols introduce additional complexity in data integration. These variations complicate the process of unifying and standardising the dataset, potentially introducing noise or errors during the transformation and unification processes. Consequently, the breadth and consistency of the dataset limit the generalisation of the developed PTF models, particularly for soil types or conditions not well represented in the

current database. Furthermore, several potentially important predictor variables, such as initial water content, soil mineralogy, cation exchange capacity, wetting and drying hysteresis, and compaction method, could not be incorporated into the model because they were not included consistently in the database.

In the modelling of water flow and solute transport, this study was restricted to one-dimensional simulations under idealised conditions. The primary focus of this work was to evaluate the feasibility, stability, and accuracy of applying PINN for modelling water flow and solute transport in unsaturated soils. Working with one dimension allowed for a detailed investigation of several key challenges that PINN faces, including balancing of loss, hard constraint implementation, training stability, and convergence issues. However, it must be acknowledged that one-dimensional simulations represent a simplification of real-world conditions. In natural soil systems, water flow and solute transport are often multidimensional and have more complex boundary conditions. These factors were not considered in this study. Therefore, extending the developed framework to high-dimensional domains remains an important direction for future research.

Although the AI models developed in this thesis achieve improved performance by reducing both approximation error and estimation error, several additional sources of uncertainty remain unaddressed. First, in terms of approximation error, the ANN models used in this thesis learn feature interactions implicitly. As a result, potential interaction effects among input variables cannot be explicitly captured, leading to some uncertainties related to the model structure. Second, in terms of estimation error, uncertainties can also arise from variability introduced by random weight initialisation, stochastic gradient descent, and other sources of randomness in the training procedure, even when the training dataset is fixed. Although these factors are usually not the major source of uncertainties, they can still have an effect on predictions.

While the accuracy of the developed PINN models was evaluated against analytical solutions and selected FEM simulations, no validations were performed

using real experimental or field data. As a result, the applicability of the proposed models to real-world problems remains to be further verified. In addition, due to resource limitations and the time constraints associated with a PhD research project, the scope of this study was necessarily limited. For example, comprehensive testing of model scalability, computational efficiency, and large-scale applicability of the PINN framework could not be fully explored within the current work.

9.2 Future Research

In future research, I will first address each of the limitations acknowledged and discussed in Section 9.1.3. In addition, three potential directions for extending the current work are discussed in the following paragraphs. First, for WRC estimation, future work should focus on expanding and enriching existing datasets. A considerable amount of potentially valuable data remains underutilised due to difficulties in data transformation, inconsistencies in formats, and the absence of standardised processing protocols [301, 302, 303, 304]. Developing unified procedures to standardise, convert, and integrate these datasets could substantially increase the volume of available data and improve the overall predictive capability of WRC models. There is a need to develop a unified framework that can facilitate the standardisation and transformation of soil data entries across different data formats, unit systems, measurement protocols, and testing standards. Such a framework would enable researchers to fully utilise a broader range of available data for data-rich analyses and model development. This approach would not only benefit WRC estimation but also support a wide range of soil studies where large empirical datasets are essential for developing predictive models and conducting comprehensive analyses. Some very recent studies (published in 2024 and 2025) have already started to tackle this field of research [305, 306, 307]. Another possible direction is to develop region-specific models by fine-tuning the global WRC model. The global model benefits from a large number of data rows for

a diverse set of soils whereas regional datasets are often limited in size, particularly in underdeveloped regions. However, soil properties are strongly regionally specific. Fine-tuning the pre-trained global model to adapt to local conditions represents a potentially useful future research direction.

Second, for the modelling with PINN, this thesis focuses exclusively on one-dimensional problems for both water flow and solute transport. This dimensional simplification allows the research can focus on solving key challenges, advection-dominated ADE and the prediction difficulty and numerical instability of RRE with PINN. After achieving progress with the one-dimensional model, future research should aim to extend the current developments to two- and three-dimensional domains. Although a few studies have demonstrated ADE modelling in higher dimensions, the advection-dominated scenarios remain unexplored [189, 190, 308]. For water transport, modelling of RRE by PINN remains limited and the success of the RRE solver developed in this thesis for the one-dimensional case provides a basis for future extension to higher dimensions. In this thesis, the Peclet-adjusted distance function demonstrated significant improvements in solving ADE problems under both steady-state and transient conditions across a wide range of Peclet numbers. However, its applicability to more complex scenarios, such as unsaturated conditions where diffusion and mechanical dispersion coefficients are functions of water content and may vary over time, remains to be investigated. Some existing studies have attempted to model unsaturated conditions ADE using conventional methods like FEM and FDM approaches [277, 309, 310, 311]. Whether machine learning–based methods, such as PINNs, can offer improved modelling capabilities or new insights in these contexts remains to be determined. Addressing these aspects would represent an important step toward advancing the practical applicability of PINNs for real-world water flow and solute transport problems.

Third, computational cost remains an important consideration for PINN-based modelling, as well as for most AI-based models. Although the current

framework demonstrates competitive accuracy, there is still considerable room for improving computational efficiency, particularly when scaling up to larger domains, higher dimensions, or more complex coupled systems. The topic has recently attracted increasing attention, and several studies have begun to explore ways to improve computational efficiency. Some recent studies explore the advanced optimisation algorithms to achieve faster convergence [312, 313, 314]. Some researchers are attempting to increase computational efficiency by utilising the parallel computing capabilities supported by recent developments in AI hardware [315, 316, 317]. These studies highlight the ongoing progress in improving the computational efficiency of PINN, which continues to advance as new methods and technologies are developed and incorporated.

Overall, the growing interest in AI within the geoenvironmental and geotechnical research communities highlights its great potential for advancing knowledge and best practice in the field. There is increasing realisation in both communities that access to comprehensive, high-quality, and well-organised datasets is an important objective in itself that can significantly enhance the prediction accuracy and generalisation of AI models. Then, the incorporation of physical laws with data-driven models can open up major opportunities for tackling complex and challenging problems. This thesis has highlighted the power and promise of combining high-quality data-driven models with physics-informed neural networks. It is a small step towards maximising the benefits of artificial intelligence in the theory and practice of geoenvironmental and geotechnical engineering.

REFERENCES

[1] A. Tafteh, N. Davatgar, and A. Sedaghat, "Estimation of important points on soil water retention curve (SWRC): comparison experimental-physical models and data mining technique," *Arabian Journal of Geosciences*, vol. 15, no. 10, p. 968, May 2022. [Online]. Available: https://doi.org/10.1007/s12517-022-10232-0

[2] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations," Nov. 2017. [Online]. Available: http://arxiv.org/abs/1711.10561

[3] C. V. Henri and E. Diamantopoulos, "Unsaturated Transport Modeling: Random-Walk Particle-Tracking as a Numerical-Dispersion Free and Efficient Alternative to Eulerian Methods," *Journal of Advances in Modeling Earth Systems*, vol. 14, no. 9, p. e2021MS002812, 2022, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021MS002812.

[4] M. W. Farthing and F. L. Ogden, "Numerical Solution of Richards' Equation: A Review of Advances and Challenges," *Soil Science Society of America Journal*, vol. 81, no. 6, pp. 1257–1269, 2017.

[5] M. Van Genuchten and W. Alves, *Analytical Solutions of One Dimensional Convective Dispersive Solute Transport Equations*, Jun. 1982, vol. 1661, journal Abbreviation: United States Department of AgricultureTechnical Bulletin Publication Title: United States Department of AgricultureTechnical Bulletin.

[6] H. Bayat and G. Ebrahim Zadeh, "Estimation of the soil water retention curve using penetration resistance curve models," *Computers and Electron-*

ics in Agriculture, vol. 144, pp. 329–343, Jan. 2018.

[7] W. Su, Y.-J. Cui, P.-J. Qin, F. Zhang, W.-M. Ye, and N. Conil, "Application of instantaneous profile method to determine the hydraulic conductivity of unsaturated natural stiff clay," *Engineering Geology*, vol. 243, pp. 111–117, Sep. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0013795218304952

[8] A. El-Zein, "Exponential finite elements for diffusion–advection problems," *International Journal for Numerical Methods in Engineering*, vol. 62, no. 15, pp. 2086–2103, 2005, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.1249. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.1249

[9] Z. H. Qiu, L. C. Wrobel, and H. Power, "Numerical solution of convection–diffusion problems at high Péclet number using boundary elements," *International Journal for Numerical Methods in Engineering*, vol. 41, no. 5, pp. 899–914, 1998.

[10] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, "Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next," *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, Jul. 2022. [Online]. Available: https://doi.org/10.1007/s10915-022-01939-z

[11] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999118307125

[12] T. Bandai and T. A. Ghezzehei, "Physics-Informed Neural Networks With Monotonicity Constraints for Richardson-Richards Equation: Estimation of Constitutive Relationships and Soil Water Flux Density From Volumetric Water Content Measurements," *Water Resources Research*, vol. 57, no. 2, p. e2020WR027642, 2021, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020WR027642. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020WR027642

[13] ——, "Forward and inverse modeling of water flow in unsaturated soils with discontinuous hydraulic conductivities using physics-informed neural networks with domain decomposition," *Hydrology and Earth System Sciences*, vol. 26, no. 16, pp. 4469–4495, Aug. 2022, publisher: Copernicus GmbH. [Online]. Available: https://hess.copernicus.org/articles/26/4469/2022/

[14] Y. Chen, Y. Xu, L. Wang, and T. Li, "Modeling water flow in unsaturated soils through physics-informed neural network with principled loss function," *Computers and Geotechnics*, vol. 161, p. 105546, Sep. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0266352X23003038

[15] I. Depina, S. Jain, S. Mar Valsson, and H. Gotovac, "Application of physics-informed neural networks to inverse problems in unsaturated groundwater flow," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 16, no. 1, pp. 21–36, Jan. 2022, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17499518.2021.1971251. [Online]. Available: https://doi.org/10.1080/17499518.2021.1971251

[16] P. Haruzi and Z. Moreno, "Modeling Water Flow and Solute Transport in Unsaturated Soils Using Physics-Informed Neural Networks Trained With Geoelectrical Data," *Water Resources Research*, vol. 59, no. 6, p. e2023WR034538, 2023,

_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023WR034538. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/2023WR034538

[17] W. A. Günther, M. H. Rezazade Mehrizi, M. Huysman, and F. Feldberg, "Debating big data: A literature review on realizing value from big data," *Journal of Strategic Information Systems*, vol. 26, no. 3, pp. 191–209, Sep. 2017. [Online]. Available: https://www.scopus.com/pages/publications/85030871734

[18] D. Ruppert, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction," *Journal of the American Statistical Association*, vol. 99, no. 466, pp. 567–567, Jun. 2004. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1198/jasa.2004.s339

[19] V. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, Sep. 1999. [Online]. Available: http://ieeexplore.ieee.org/document/788640/

[20] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967. [Online]. Available: https://ieeexplore.ieee.org/document/1053964

[21] E. Fix and J. Hodges, "Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties," USAF School of Aviation Medicine, Randolph Field, Texas, USA, Tech. Rep. Technical Report No. 4, 1951.

[22] N. Bhatia and Vandana, "Survey of Nearest Neighbor Techniques," Jul. 2010, arXiv:1007.0085 [cs]. [Online]. Available: http://arxiv.org/abs/1007.0085

[23] S. Zhang, "Challenges in KNN Classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 10, pp. 4663–4675, Oct. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9314060

[24] N. Kouiroukidis and G. Evangelidis, "The Effects of Dimensionality Curse in High Dimensional kNN Search," in *2011 15th Panhellenic Conference on Informatics*, Sep. 2011, pp. 41–45. [Online]. Available: https://ieeexplore.ieee.org/document/6065061

[25] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: https://doi.org/10.1007/BF00994018

[26] G. A. Senthil, S. Geerthik, R. Karthikeyan, and G. Keerthana, "Face Recognition based Automated Smart Attendance using Hybrid Machine Learning Algorithms and Computer Vision," *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pp. 606–611, Jun. 2024, conference Name: 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) ISBN: 9798350375190 Place: Salem, India Publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/document/10574896/

[27] S. B. Chaabane, M. Hijji, R. Harrabi, and H. Seddik, "Face recognition based on statistical features and SVM classifier," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 8767–8784, Mar. 2022. [Online]. Available: https://link.springer.com/10.1007/s11042-021-11816-w

[28] A. Wibowo Haryanto, E. Kholid Mawardi, and Muljono, "Influence of Word Normalization and Chi-Squared Feature Selection on Support Vector Machine (SVM) Text Classification," *2018 International Seminar on Application for Technology of Information and Communication*, pp. 229–233, Sep. 2018, conference Name: 2018 International Seminar on Application for Technology of Information and Communication (iSemantic) ISBN: 9781538674864 Place: Semarang Publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/document/8549748/

[29] M. Haddoud, A. Mokhtari, T. Lecroq, and S. Abdeddaïm, "Combining

supervised term-weighting metrics for SVM text classification with extended term representation," *Knowledge and Information Systems*, vol. 49, no. 3, pp. 909–931, Dec. 2016. [Online]. Available: http://link.springer.com/10.1007/s10115-016-0924-1

[30] M. A. El-Rashidy, R. G. Mohamed, N. A. El-Fishawy, and M. A. Shouman, "An effective text plagiarism detection system based on feature selection and SVM techniques," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 2609–2646, Jan. 2024. [Online]. Available: https://link.springer.com/10.1007/s11042-023-15703-4

[31] I. Chivers and J. Sleightholme, "An Introduction to Algorithms and the Big O Notation," in *Introduction to Programming with Fortran: With Coverage of Fortran 90, 95, 2003, 2008 and 77*, I. Chivers and J. Sleightholme, Eds. Cham: Springer International Publishing, 2015, pp. 359–364. [Online]. Available: https://doi.org/10.1007/978-3-319-17701-4_23

[32] N. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods." Cambridge University Press, Mar. 2000, edition: 1. [Online]. Available: https://www.cambridge.org/core/product/identifier/9780511801389/type/book

[33] A. Althnian, D. AlSaeed, H. Al-Baity, A. Samha, A. B. Dris, N. Alzakari, A. Abou Elwafa, and H. Kurdi, "Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain," *Applied Sciences*, vol. 11, no. 2, p. 796, Jan. 2021, publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2076-3417/11/2/796

[34] C. A. Ramezan, T. A. Warner, A. E. Maxwell, and B. S. Price, "Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data," *Remote Sensing*, vol. 13, no. 3, p. 368, Jan. 2021,

publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2072-4292/13/3/368

[35] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, Mar. 1986. [Online]. Available: https://doi.org/10.1007/BF00116251

[36] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. New York: Chapman and Hall/CRC, 1984.

[37] L. L. Custode and G. Iacca, "Evolutionary Learning of Interpretable Decision Trees," *IEEE Access*, vol. 11, pp. 6169–6184, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10015004

[38] V. G. Costa and C. E. Pedreira, "Recent advances in decision trees: an updated survey," *Artificial Intelligence Review*, vol. 56, no. 5, pp. 4765–4800, May 2023. [Online]. Available: https://doi.org/10.1007/s10462-022-10275-5

[39] A. Amro, M. Al-Akhras, K. E. Hindi, M. Habib, and B. A. Shawar, "Instance Reduction for Avoiding Overfitting in Decision Trees," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 438–459, Jan. 2021, publisher: De Gruyter. [Online]. Available: https://www.degruyterbrill.com/document/doi/10.1515/jisys-2020-0061/html

[40] M. Bramer, "Avoiding Overfitting of Decision Trees," in *Principles of Data Mining*, M. Bramer, Ed. London: Springer, 2013, pp. 121–136. [Online]. Available: https://doi.org/10.1007/978-1-4471-4884-5_9

[41] R. Rivera-Lopez, J. Canul-Reich, E. Mezura-Montes, and M. A. Cruz-Chávez, "Induction of decision trees as classification models through metaheuristics," *Swarm and Evolutionary Computation*, vol. 69, p. 101006, Mar. 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2210650221001681

[42] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324

[43] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, Apr. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0924271616000265

[44] Q. Cheng, Z. Chunhong, and L. Qianglin, "Development and application of random forest regression soft sensor model for treating domestic wastewater in a sequencing batch reactor," *Scientific Reports*, vol. 13, no. 1, p. 9149, Jun. 2023. [Online]. Available: https://www.nature.com/articles/s41598-023-36333-8

[45] W. Zhang, C. Wu, H. Zhong, Y. Li, and L. Wang, "Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization," *Geoscience Frontiers*, vol. 12, no. 1, pp. 469–477, Jan. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1674987120300669

[46] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, "An improved random forest based on the classification accuracy and correlation measurement of decision trees," *Expert Syst. Appl.*, 2024. [Online]. Available: https://www.semanticscholar.org/paper/An-improved-random-forest-based-on-the-accuracy-and-Sun-Wang/dcb8dc81ae3c162249383cddd9321bdd97940332

[47] S. Gündoğdu, "Efficient prediction of early-stage diabetes using XGBoost classifier with random forest feature selection technique," *Multimedia Tools and Applications*, vol. 82, no. 22, pp. 34 163–34 181, Sep. 2023. [Online]. Available: https://link.springer.com/10.1007/s11042-023-15165-8

[48] R. Zhou, J. Liang, Q. Chen, H. Tian, C. Yang, and C. Liu, "A 3-Gene Random Forest Model to Diagnose Non-obstructive Azoospermia

Based on Transcription Factor-Related Henes," *Reproductive Sciences*, vol. 30, no. 1, pp. 233–246, Jan. 2023. [Online]. Available: https://link.springer.com/10.1007/s43032-022-01008-8

[49] A. R. Linero and Y. Yang, "Bayesian Regression Tree Ensembles that Adapt to Smoothness and Sparsity," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 80, no. 5, pp. 1087–1110, Nov. 2018. [Online]. Available: https://doi.org/10.1111/rssb.12293

[50] L. Mentch and G. Hooker, "Quantifying Uncertainty in Random Forests via Confidence Intervals and Hypothesis Tests," Sep. 2015, arXiv:1404.6473 [stat]. [Online]. Available: http://arxiv.org/abs/1404.6473

[51] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, no. 2, pp. 197–227, Jun. 2016. [Online]. Available: http://link.springer.com/10.1007/s11749-016-0481-7

[52] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional Neural Networks for Speech Recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1533–1545, Oct. 2014, conference Name: IEEE/ACM Transactions on Audio, Speech, and Language Processing.

[53] S. Javanmardi, S.-H. Miraei Ashtiani, F. J. Verbeek, and A. Martynenko, "Computer-vision classification of corn seed varieties using deep convolutional neural network," *Journal of Stored Products Research*, vol. 92, p. 101800, May 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022474X21000394

[54] J. Momoh and R. Button, "Design and analysis of aerospace DC arcing faults using fast fourier transformation and artificial neural network," in *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491)*, vol. 2, Jul. 2003, pp. 788–793 Vol. 2.

[55] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech Recognition Using Deep Neural Networks: A Systematic Review," *IEEE Access*, vol. 7, pp. 19 143–19 165, 2019, conference Name: IEEE Access.

[56] K. Sabanci, A. Kayabasi, and A. Toktas, "Computer vision-based method for classification of wheat grains using artificial neural network," *Journal of the Science of Food and Agriculture*, vol. 97, no. 8, pp. 2588–2593, 2017, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jsfa.8080. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/jsfa.8080

[57] V. Yadav, R. Padhi, and S. N. Balakrishnan, "Robust/Optimal Temperature Profile Control of a High-Speed Aerospace Vehicle Using Neural Networks," *IEEE Transactions on Neural Networks*, vol. 18, no. 4, pp. 1115–1128, Jul. 2007, conference Name: IEEE Transactions on Neural Networks.

[58] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical Image Analysis using Convolutional Neural Networks: A Review," *Journal of Medical Systems*, vol. 42, no. 11, p. 226, Oct. 2018. [Online]. Available: https://doi.org/10.1007/s10916-018-1088-1

[59] J. C. Jan, S.-L. Hung, S. Y. Chi, and J. C. Chern, "Neural Network Forecast Model in Deep Excavation," *Journal of Computing in Civil Engineering*, vol. 16, no. 1, pp. 59–65, Jan. 2002, publisher: American Society of Civil Engineers. [Online]. Available: https://ascelibrary.org/doi/10.1061/%28ASCE%290887-3801%282002%2916%3A1%2859%29

[60] L. Sou-Sen and L. Hsien-Chuang, "Neural-network-based regression model of ground surface settlement induced by deep excavation," *Automation in Construction*, vol. 13, no. 3, pp. 279–289, May 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580503000189

[61] B. Ghanbarian, V. Taslimitehrani, G. Dong, and Y. A. Pachepsky, "Sample dimensions effect on prediction of soil water retention curve and saturated

hydraulic conductivity," *Journal of Hydrology*, vol. 528, pp. 127–137, Sep. 2015.

[62] V. Novák and H. Hlaváčiková, "Soil-Water Retention Curve," in *Applied Soil Hydrology*, V. Novák and H. Hlaváčiková, Eds. Cham: Springer International Publishing, 2019, pp. 77–96.

[63] H. Vereecken, M. Weynants, M. Javaux, Y. Pachepsky, M. G. Schaap, and M. van Genuchten, "Using Pedotransfer Functions to Estimate the van Genuchten–Mualem Soil Hydraulic Properties: A Review," *Vadose Zone Journal*, vol. 9, no. 4, pp. 795–820, 2010, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.2136/vzj2010.0045. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.2136/vzj2010.0045

[64] K. R. Aljanabi and O. M. AL-Azzawi, "Neural network application in forecasting maximum wall deflection in homogenous clay," *International Journal of Geo-Engineering*, vol. 12, no. 1, p. 29, Oct. 2021. [Online]. Available: https://doi.org/10.1186/s40703-021-00158-z

[65] P. Mishra, P. Samui, and E. Mahmoudi, "Probabilistic Design of Retaining Wall Using Machine Learning Methods," *Applied Sciences*, vol. 11, no. 12, p. 5411, Jan. 2021, number: 12 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2076-3417/11/12/5411

[66] I. A. Basheer, L. N. Reddi, and Y. M. Najjar, "Site Characterization by Neuronets: An Application to the Landfill Siting Problem," *Groundwater*, vol. 34, no. 4, pp. 610–617, 1996, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1745-6584.1996.tb02048.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-6584.1996.tb02048.x

[67] W. G. Baxt, "The Role of the Artificial Neural Network in the Characterisation of Complex Systems and the Prediction of Disease,"

in *Artificial Neural Networks in Biomedicine*, ser. Perspectives in Neural Computing, P. J. G. Lisboa, E. C. Ifeachor, and P. S. Szczepaniak, Eds. London: Springer, 2000, pp. 25–37. [Online]. Available: https://doi.org/10.1007/978-1-4471-0487-2_3

[68] C. H. Juang, T. Jiang, and R. A. Christopher, "Three-dimensional site characterisation: neural network approach," *Géotechnique*, vol. 51, no. 9, pp. 799–809, Nov. 2001, publisher: ICE Publishing. [Online]. Available: https://www.icevirtuallibrary.com/doi/10.1680/geot.2001.51.9.799

[69] P. Samui and T. G. Sitharam, "Site Characterization Model Using Artificial Neural Network and Kriging," *International Journal of Geomechanics*, vol. 10, no. 5, pp. 171–180, Oct. 2010, publisher: American Society of Civil Engineers. [Online]. Available: https://ascelibrary.org/doi/10.1061/%28ASCE%291532-3641%282010%2910%3A5%28171%29

[70] A. Varley, A. Tyler, L. Smith, and P. Dale, "Development of a neural network approach to characterise 226Ra contamination at legacy sites using gamma-ray spectra taken from boreholes," *Journal of Environmental Radioactivity*, vol. 140, pp. 130–140, Feb. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0265931X14003452

[71] N. Y.M. and I. Basheer, *A Neural Network Approach for Site Characterization and Uncertainty Prediction*, Jul. 1996.

[72] H. Lee and J. Song, "Introduction to convolutional neural network using Keras; an understanding from a statistician," *Communications for Statistical Applications and Methods*, vol. 26, no. 6, pp. 591–610, Nov. 2019, publisher: Korean Statistical Society. [Online]. Available: http://www.csam.or.kr/journal/view.html?doi=10.29220/CSAM.2019.26.6.591

[73] T. A. Pham, H.-B. Ly, V. Q. Tran, L. V. Giap, H.-L. T. Vu, and H.-A. T. Duong, "Prediction of Pile Axial Bearing Capacity Using Artificial Neural Network and Random Forest," *Applied Sciences*, vol. 10, no. 5, p. 1871, Jan.

2020, number: 5 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2076-3417/10/5/1871

[74] E. Díaz, V. Brotons, and R. Tomás, "Use of artificial neural networks to predict 3-D elastic settlement of foundations on soils with inclined bedrock," *Soils and Foundations*, vol. 58, no. 6, pp. 1414–1422, Dec. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038080618301215

[75] Z. H. Alzahamie and H. A. Abdul-Husain, "Artificial neural network for prediction of liquefaction triggering based on CPT data," *Journal of Physics: Conference Series*, vol. 1973, no. 1, p. 012197, Aug. 2021, publisher: IOP Publishing. [Online]. Available: https://dx.doi.org/10.1088/1742-6596/1973/1/012197

[76] T. A. Pham, "Application of Feedforward Neural Network and SPT Results in the Estimation of Seismic Soil Liquefaction Triggering," *Computational Intelligence and Neuroscience*, vol. 2021, p. e1058825, Oct. 2021, publisher: Hindawi. [Online]. Available: https://www.hindawi.com/journals/cin/2021/1058825/

[77] A. T. Y. Tung, Ya Yung Wang, and F. S. Wong, "Assessment of liquefaction potential using neural networks," *Soil Dynamics and Earthquake Engineering*, vol. 12, no. 6, pp. 325–335, Jan. 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/026772619390035P

[78] Z. Liao and Z. Liao, "Slope stability evaluation using backpropagation neural networks and multivariate adaptive regression splines," *Open Geosciences*, vol. 12, no. 1, pp. 1263–1273, Jan. 2020, publisher: De Gruyter Open Access. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/geo-2020-0198/html?lang=en

[79] J. Meng, H. Mattsson, and J. Laue, "Three-dimensional slope stability predictions using artificial neural networks," *Interna-*

tional Journal for Numerical and Analytical Methods in Geomechanics, vol. 45, no. 13, pp. 1988–2000, 2021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nag.3252. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/nag.3252

[80] M. G. Sakellariou and M. D. Ferentinou, "A study of slope stability prediction using neural networks," *Geotechnical & Geological Engineering*, vol. 23, no. 4, pp. 419–445, Aug. 2005. [Online]. Available: https://doi.org/10.1007/s10706-004-8680-5

[81] M. Azarafza, M. Azarafza, H. Akgün, P. M. Atkinson, and R. Derakhshani, "Deep learning-based landslide susceptibility mapping," *Scientific Reports*, vol. 11, no. 1, p. 24112, Dec. 2021, number: 1 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41598-021-03585-1

[82] M. Mehrabi and H. Moayedi, "Landslide susceptibility mapping using artificial neural network tuned by metaheuristic algorithms," *Environmental Earth Sciences*, vol. 80, no. 24, p. 804, Nov. 2021. [Online]. Available: https://doi.org/10.1007/s12665-021-10098-7

[83] A. Beucher, C. B. Rasmussen, T. B. Moeslund, and M. H. Greve, "Interpretation of Convolutional Neural Networks for Acid Sulfate Soil Classification," *Frontiers in Environmental Science*, vol. 9, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fenvs.2021.809995

[84] Y. Cal, "Soil classification by neural network," *Advances in Engineering Software*, vol. 22, no. 2, pp. 95–97, Jan. 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/096599789400035H

[85] P. Srivastava, A. Shukla, and A. Bansal, "A comprehensive review on soil classification using deep learning and computer vision techniques," *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 14 887–14 914, Apr. 2021. [Online]. Available: https://doi.org/10.1007/s11042-021-10544-5

[86] M. Adil, R. Ullah, S. Noor, and N. Gohar, "Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design," *Neural Computing and Applications*, Sep. 2020.

[87] S. Hayou, A. Doucet, and J. Rousseau, "On the Selection of Initialization and Activation Function for Deep Neural Networks," Oct. 2018. [Online]. Available: http://arxiv.org/abs/1805.08266

[88] H. Yu, "Network Complexity Analysis of Multilayer Feedforward Artificial Neural Networks," in *Applications of Neural Networks in High Assurance Systems*, J. Schumann and Y. Liu, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 41–55.

[89] A. J. Thomas, M. Petridis, S. D. Walters, S. M. Gheytassi, and R. E. Morgan, "Two Hidden Layers are Usually Better than One," *Engineering Applications of Neural Networks*, pp. 279–290, 2017.

[90] Y. Hayashi, M. Sakata, and S. I. Gallant, "Multi-Layer Versus Single-Layer Neural Networks and an Application to Reading Hand-Stamped Characters," in *International Neural Network Conference: July 9–13, 1990 Palais Des Congres — Paris — France.* Dordrecht: Springer Netherlands, 1990, pp. 781–784.

[91] E. Tadmor, "A review of numerical methods for nonlinear partial differential equations," *Bulletin of the American Mathematical Society*, vol. 49, no. 4, pp. 507–554, Jan. 2012. [Online]. Available: http://www.ams.org/jourcgi/jour-getitem?pii=S0273-0979-2012-01379-4

[92] D. P. Young, R. G. Melvin, M. B. Bieterman, F. T. Johnson, S. S. Samant, and J. E. Bussoletti, "A locally refined rectangular grid finite element method: Application to computational fluid dynamics and computational physics," *Journal of Computational Physics*, vol. 92, no. 1, pp. 1–66, Jan. 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/002199919190291R

[93] T. Strouboulis, K. Copps, and I. Babuška, "The generalized finite element method," *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 32, pp. 4081–4193, May 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782501001888

[94] Y. Zang, G. Bao, X. Ye, and H. Zhou, "Weak adversarial networks for high-dimensional partial differential equations," *Journal of Computational Physics*, vol. 411, p. 109409, Jun. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999120301832

[95] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*. Courier Corporation, Jun. 1994, google-Books-ID: y77n2ySMJHUC.

[96] N. R. Morgan and B. J. Archer, "On the Origins of Lagrangian Hydrodynamic Methods," *Nuclear Technology*, vol. 207, no. sup1, pp. S147–S175, Dec. 2021, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00295450.2021.1913034. [Online]. Available: https://doi.org/10.1080/00295450.2021.1913034

[97] T. E. Tezduyar and K. Takizawa, "Space–time computations in practical engineering applications: a summary of the 25-year history," *Computational Mechanics*, vol. 63, no. 4, pp. 747–753, Apr. 2019. [Online]. Available: https://doi.org/10.1007/s00466-018-1620-7

[98] T. Tezduyar, S. Aliabadi, M. Behr, A. Johnson, and S. Mittal, "Parallel finite-element computation of 3D flows," *Computer*, vol. 26, no. 10, pp. 27–36, Oct. 1993. [Online]. Available: https://ieeexplore.ieee.org/document/237441

[99] S. R. Morab, A. Sharma, and J. S. Murallidharan, "Fully finite volume method on a curvilinear grid-based arbitrary Lagrangian Eulerian approach for computational fluid flexible-structure interaction," *Computer Physics Communications*, vol. 296, p. 109054, Mar. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010465523003995

[100] E. Dick, "Introduction to Finite Element Methods in Computational Fluid Dynamics," in *Computational Fluid Dynamics*, J. F. Wendt, Ed. Berlin, Heidelberg: Springer, 2009, pp. 235–274. [Online]. Available: https://doi.org/10.1007/978-3-540-85056-4_10

[101] J. A. Vrugt, "Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation," *Environmental Modelling & Software*, vol. 75, pp. 273–316, Jan. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364815215300396

[102] E. Ermakova and M. Rynkovskaya, "Modern software features for shape optimization of shells," *Structural Mechanics of Engineering Constructions and Buildings*, vol. 19, pp. 220–232, Sep. 2023.

[103] M. Assad, R. A. Hawileh, J. A. Abdalla, and F. Abed, "Heat Transfer Analysis of Reinforced Concrete Walls in ANSYS and ABAQUS: A Comparative Study," in *2022 Advances in Science and Engineering Technology International Conferences (ASET)*, Feb. 2022, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/9735001

[104] T. G. Ritto and F. A. Rochinha, "Digital twin, physics-based model, and machine learning applied to damage detection in structures," *Mechanical Systems and Signal Processing*, vol. 155, p. 107614, Jun. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327021000091

[105] I. Poutiainen, P. Tanskanen, and G. Marquis, "Finite element methods for structural hot spot stress determination—a comparison of procedures," *International Journal of Fatigue*, vol. 26, no. 11, pp. 1147–1157, Nov. 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142112304000908

[106] J. R. Wu and Q. S. Li, "Structural parameter identification and damage detection for a steel structure using a two-stage finite element model updating method," *Journal of Constructional Steel Research*, vol. 62, no. 3, pp. 231–239, Mar. 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0143974X05001306

[107] M. Ashraf, L. Gardner, and D. A. Nethercot, "Finite element modelling of structural stainless steel cross-sections," *Thin-Walled Structures*, vol. 44, no. 10, pp. 1048–1062, Oct. 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263823106001741

[108] P. Dong, "A structural stress definition and numerical implementation for fatigue analysis of welded joints," *International Journal of Fatigue*, vol. 23, no. 10, pp. 865–876, Nov. 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S014211230100055X

[109] T. Nakata, H. Liu, and R. J. Bomphrey, "A CFD-informed quasi-steady model of flapping-wing aerodynamics," *Journal of Fluid Mechanics*, vol. 783, pp. 323–343, Nov. 2015.

[110] K. Zhang, S. Hayostek, M. Amitay, W. He, V. Theofilis, and K. Taira, "On the formation of three-dimensional separated flows over wings under tip effects," *Journal of Fluid Mechanics*, vol. 895, p. A9, Jul. 2020.

[111] A. Rizzi, "Separated and vortical flow in aircraft aerodynamics: a CFD perspective," *The Aeronautical Journal*, vol. 127, no. 1313, pp. 1065–1103, Jul. 2023. [Online]. Available: https://www.cambridge.org/core/journals/aeronautical-journal/article/separated-and-vortical-flow-in-aircraft-aerodynamics-a-cfd-perspective/BE5838AD5348C2BE0FD026371C0B6B75

[112] O. Baysal and M. E. Eleshaky, "Aerodynamic design optimization using sensitivity analysis and computational fluid dynamics," *AIAA Journal*, vol. 30, no. 3, pp. 718–725, Mar. 1992, publisher: American

Institute of Aeronautics and Astronautics. [Online]. Available: https://arc.aiaa.org/doi/10.2514/3.10977

[113] Z. Lyu and J. R. R. A. Martins, "Aerodynamic Design Optimization Studies of a Blended-Wing-Body Aircraft," *Journal of Aircraft*, vol. 51, no. 5, pp. 1604–1617, Sep. 2014, publisher: American Institute of Aeronautics and Astronautics. [Online]. Available: https://arc.aiaa.org/doi/10.2514/1.C032491

[114] A. Benaouali and S. Kachel, "Multidisciplinary design optimization of aircraft wing using commercial software integration," *Aerospace Science and Technology*, vol. 92, pp. 766–776, Sep. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1270963818324131

[115] S. A. Maas, B. J. Ellis, G. A. Ateshian, and J. A. Weiss, "FEBio: Finite Elements for Biomechanics," *Journal of Biomechanical Engineering*, vol. 134, no. 011005, Feb. 2012. [Online]. Available: https://doi.org/10.1115/1.4005694

[116] E. J. Rayfield, "Finite Element Analysis and Understanding the Biomechanics and Evolution of Living and Fossil Organisms," *Annual Review of Earth and Planetary Sciences*, vol. 35, no. Volume 35, 2007, pp. 541–576, May 2007, publisher: Annual Reviews. [Online]. Available: https://www.annualreviews.org/content/journals/10.1146/annurev.earth.35.031306.140104

[117] A. Samani, J. Bishop, M. Yaffe, and D. Plewes, "Biomechanical 3-D finite element modeling of the human breast using MRI data," *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 271–279, Apr. 2001. [Online]. Available: https://ieeexplore.ieee.org/document/921476

[118] M. Bendjaballah, A. Shirazi-Adl, and D. Zukor, "Biomechanics of the human knee joint in compression: reconstruction, mesh generation and finite element analysis," *The Knee*, vol. 2, no. 2, pp. 69–79, Jun.

1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/096801609500018K

[119] H. G. Delorenzi, "Energy release rate calculations by the finite element method," *Engineering Fracture Mechanics*, vol. 21, no. 1, pp. 129–143, Jan. 1985. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0013794485900608

[120] V. Cotoni, P. Shorter, and R. Langley, "Numerical and experimental validation of a hybrid finite element-statistical energy analysis method," *The Journal of the Acoustical Society of America*, vol. 122, no. 1, pp. 259–270, Jul. 2007. [Online]. Available: https://doi.org/10.1121/1.2739420

[121] M. Ali, A. Qamhiyah, D. Flugrad, and M. Shakoor, "Theoretical and finite element study of a compact energy absorber," *Advances in Engineering Software*, vol. 39, no. 2, pp. 95–106, Feb. 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0965997807000130

[122] D. V. Griffiths and P. A. Lane, "Slope stability analysis by finite elements," *Géotechnique*, vol. 49, no. 3, pp. 387–403, Jun. 1999. [Online]. Available: https://doi.org/10.1680/geot.1999.49.3.387

[123] S. Marras, J. F. Kelly, M. Moragues, A. Müller, M. A. Kopera, M. Vázquez, F. X. Giraldo, G. Houzeaux, and O. Jorba, "A Review of Element-Based Galerkin Methods for Numerical Weather Prediction: Finite Elements, Spectral Elements, and Discontinuous Galerkin," *Archives of Computational Methods in Engineering*, vol. 23, no. 4, pp. 673–722, Dec. 2016. [Online]. Available: https://doi.org/10.1007/s11831-015-9152-1

[124] D. Ngo and A. C. Scordelis, "Finite Element Analysis of Reinforced Concrete Beams." [Online]. Available: https://www.concrete.org/publications/internationalconcreteabstractsportal.aspx?m=details&id=7551

[125] O. Isbilir and E. Ghassemieh, "Finite Element Analysis of Drilling of Carbon Fibre Reinforced Composites," *Applied Composite Materials*,

vol. 19, no. 3, pp. 637–656, Jun. 2012. [Online]. Available: https://doi.org/10.1007/s10443-011-9224-9

[126] A. Abdulle and Y. Bai, "Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems," *Journal of Computational Physics*, vol. 231, no. 21, pp. 7014–7036, Aug. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999112001209

[127] Z. Li and F. Han, "The peridynamics-based finite element method (PeriFEM) with adaptive continuous/discrete element implementation for fracture simulation," *Engineering Analysis with Boundary Elements*, vol. 146, pp. 56–65, Jan. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0955799722003423

[128] C. A. Duarte and D. J. Kim, "Analysis and applications of a generalized finite element method with global–local enrichment functions," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 6, pp. 487–504, Jan. 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782507003404

[129] C. Farhat, P. Avery, T. Chapman, and J. Cortial, "Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency," *International Journal for Numerical Methods in Engineering*, vol. 98, no. 9, pp. 625–662, 2014, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.4668. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.4668

[130] J. Kudela and R. Matousek, "Recent advances and applications of surrogate models for finite element method computations: a review,"

Soft Computing, vol. 26, no. 24, pp. 13 709–13 733, Dec. 2022. [Online]. Available: https://doi.org/10.1007/s00500-022-07362-8

[131] J. Aarnes and B.-O. Heimsund, "Multiscale Discontinuous Galerkin Methods for Elliptic Problems with Multiple Scales," in *Multiscale Methods in Science and Engineering*, B. Engquist, O. Runborg, and P. Lötstedt, Eds. Berlin, Heidelberg: Springer, 2005, pp. 1–20.

[132] N. Wakabayashi, M. Ona, T. Suzuki, and Y. Igarashi, "Nonlinear finite element analyses: Advances and challenges in dental applications," *Journal of Dentistry*, vol. 36, no. 7, pp. 463–471, Jul. 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S030057120800105X

[133] A. Klawonn, M. Lanser, and O. Rheinbach, "Toward Extremely Scalable Nonlinear Domain Decomposition Methods for Elliptic Partial Differential Equations," *SIAM Journal on Scientific Computing*, vol. 37, no. 6, pp. C667–C696, Jan. 2015, publisher: Society for Industrial and Applied Mathematics. [Online]. Available: https://epubs.siam.org/doi/10.1137/140997907

[134] O. Lakkis and T. Pryer, "A Finite Element Method for Nonlinear Elliptic Problems," *SIAM Journal on Scientific Computing*, vol. 35, no. 4, pp. A2025–A2045, Jan. 2013, publisher: Society for Industrial and Applied Mathematics. [Online]. Available: https://epubs.siam.org/doi/10.1137/120887655

[135] H. Eivazi, M. Tahani, P. Schlatter, and R. Vinuesa, "Physics-informed neural networks for solving Reynolds-averaged Navier$\unicode{x2013}$Stokes equations," *Physics of Fluids*, vol. 34, no. 7, p. 075117, Jul. 2022, arXiv:2107.10711 [physics]. [Online]. Available: http://arxiv.org/abs/2107.10711

[136] X. Jin, S. Cai, H. Li, and G. E. Karniadakis, "NSFnets (Navier-Stokes Flow nets): Physics-informed neural networks for the incompressible

Navier-Stokes equations," *Journal of Computational Physics*, vol. 426, p. 109951, Feb. 2021, arXiv:2003.06496 [physics]. [Online]. Available: http://arxiv.org/abs/2003.06496

[137] S. Li and X. Feng, "Dynamic Weight Strategy of Physics-Informed Neural Networks for the 2D Navier–Stokes Equations," *Entropy*, vol. 24, no. 9, p. 1254, Sep. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9497516/

[138] C. Chang, Z. Xin, and T. Zeng, "A conservative hybrid deep learning method for Maxwell–Ampère–Nernst–Planck equations," *Journal of Computational Physics*, vol. 501, p. 112791, Mar. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999124000408

[139] S. Piao, H. Gu, A. Wang, and P. Qin, "A Domain-adaptive Physics-informed Neural Network for Inverse Problems of Maxwell's Equations in Heterogeneous Media," Aug. 2023. [Online]. Available: http://arxiv.org/abs/2308.06436

[140] X. Jiang, D. Wang, Q. Fan, M. Zhang, C. Lu, and A. P. Tao Lau, "Solving the Nonlinear Schrödinger Equation in Optical Fibers Using Physics-informed Neural Network," in *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, Jun. 2021, pp. 1–3. [Online]. Available: https://ieeexplore.ieee.org/document/9489953

[141] K. Shah, P. Stiller, N. Hoffmann, and A. Cangi, "Physics-Informed Neural Networks as Solvers for the Time-Dependent Schr\textbackslash"odinger Equation," Oct. 2022. [Online]. Available: http://arxiv.org/abs/2210.12522

[142] W.-X. Yuan, R. Guo, and Y.-N. Gao, "Physics-informed Neural Network method for the Modified Nonlinear Schrödinger equation," *Optik*, vol. 279, p. 170739, May 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030402623002358

[143] B. Huang and J. Wang, "Applications of Physics-Informed Neural Networks in Power Systems - A Review," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 572–588, Jan. 2023, conference Name: IEEE Transactions on Power Systems.

[144] Y. Sun, U. Sengupta, and M. Juniper, "Physics-informed deep learning for simultaneous surrogate modeling and PDE-constrained optimization of an airfoil geometry," *Computer Methods in Applied Mechanics and Engineering*, vol. 411, p. 116042, Jun. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782523001664

[145] J. Yu, L. Lu, X. Meng, and G. E. Karniadakis, "Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 393, p. 114823, Apr. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782522001438

[146] S. Cai, Z. Wang, S. Wang, P. Perdikaris, and G. E. Karniadakis, "Physics-Informed Neural Networks for Heat Transfer Problems," *Journal of Heat Transfer*, vol. 143, no. 060801, Apr. 2021. [Online]. Available: https://doi.org/10.1115/1.4050542

[147] W. Yan, Y. Yan, P. Shen, and W.-H. Zhou, "A hybrid physical data informed DNN in axial displacement prediction of immersed tunnel joint," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 17, no. 1, pp. 169–180, Jan. 2023, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17499518.2023.2169941. [Online]. Available: https://doi.org/10.1080/17499518.2023.2169941

[148] M. G. Schaap and F. J. Leij, "Using neural networks to predict soil water retention and soil hydraulic conductivity," *Soil and Tillage Research*, vol. 47, no. 1, pp. 37–42, Jun. 1998.

[149] M. G. Schaap, F. J. Leij, and M. T. van Genuchten, "rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions," *Journal of Hydrology*, vol. 251, no. 3, pp. 163–176, Oct. 2001.

[150] M. Maček, J. Smolar, and P. Ana, *Extension of measurement range of dewpoint potentiometer and evaporation method*, 2013.

[151] X. Zhang, M. Mavroulidou, M. Gunn, and Z. Cabarkapa, "Experiences gained using various axis translation technique apparatus to determine soil water retention curve," *Unsaturated Soils: Experimental Studies in Unsaturated Soils and Expansive Soils*, pp. 269–274, 2010.

[152] H. Bayat, M. R. Neyshaburi, K. Mohammadi, N. Nariman-Zadeh, M. Irannejad, and A. S. Gregory, "Combination of artificial neural networks and fractal theory to predict soil water retention curve," *Computers and Electronics in Agriculture*, vol. 92, pp. 92–103, Mar. 2013.

[153] I. Campos-Guereta, A. Dawson, and N. Thom, "An alternative continuous form of Arya and Paris model to predict the soil water retention curve of a soil," *Advances in Water Resources*, vol. 154, p. 103968, Aug. 2021.

[154] A. C. da Silva, R. A. Armindo, B. Minasny, and C. L. Prevedello, "Evaluating the Splintex model for estimating the soil water retention curve for a wide range of soils," *Soil and Tillage Research*, vol. 209, p. 104974, May 2021.

[155] M. Rastgou, H. Bayat, M. Mansoorizadeh, and A. S. Gregory, "Estimating the soil water retention curve: Comparison of multiple nonlinear regression approach and random forest data mining technique," *Computers and Electronics in Agriculture*, vol. 174, p. 105502, Jul. 2020.

[156] M. Acutis and M. Donatelli, "SOILPAR 2.00: software to estimate soil hydrological parameters and functions," *European Journal of Agronomy*, vol. 18, no. 3, pp. 373–377, Jan. 2003.

[157] J. Bouma, "Using Soil Survey Data for Quantitative Land Evaluation," *In: Advance in soil science, B.A. Stewart (ed.). Springer-Verlag New York. Vol. 9 (1989) 225-239.*, vol. 9, 1989.

[158] P. M. Nguyen, A. Haghverdi, J. de Pue, Y.-D. Botula, K. V. Le, W. Waegeman, and W. M. Cornelis, "Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils," *Biosystems Engineering*, vol. 153, pp. 12–27, Jan. 2017.

[159] C. D. Børgesen and M. G. Schaap, "Point and parameter pedotransfer functions for water retention predictions for Danish soils," *Geoderma*, vol. 127, no. 1, pp. 154–167, Jul. 2005.

[160] Y. A. Pachepsky, D. Timlin, and G. Varallyay, "Artificial Neural Networks to Estimate Soil Water Retention from Easily Measurable Data," *Soil Science Society of America journal*, vol. 60, no. 3, pp. 727–733, 1996.

[161] J. Tomasella, Y. Pachepsky, S. Crestana, and W. J. Rawls, "Comparison of Two Techniques to Develop Pedotransfer Functions for Water Retention," *Soil Science Society of America journal*, vol. 67, 2003.

[162] H. Bayat, M. R. Neyshabouri, K. Mohammadi, and N. Nariman-Zadeh, "Estimating Water Retention with Pedotransfer Functions Using Multi-Objective Group Method of Data Handling and ANNs," *Pedosphere*, vol. 21, no. 1, pp. 107–114, Feb. 2011.

[163] B. Ghanbarian-Alavijeh, A. Liaghat, G.-H. Huang, and M. T. Van Genuchten, "Estimation of the van Genuchten Soil Water Retention Properties from Soil Textural Data," *Pedosphere*, vol. 20, no. 4, pp. 456–465, Aug. 2010.

[164] A. Haghverdi, W. M. Cornelis, and B. Ghahraman, "A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data," *Journal of Hydrology*, vol. 442-443, pp. 46–54, Jun. 2012.

[165] B. Minasny and A. McBratney, "The Neuro-m Method for Fitting Neural Network Parametric Pedotransfer Functions," *Published in Soil Sci. Soc. Am. J*, vol. 66, pp. 352–361, 2002.

[166] H.-G. Han and J.-F. Qiao, "A structure optimisation algorithm for feedforward neural network construction," *Neurocomputing*, vol. 99, pp. 347–357, Jan. 2013.

[167] K. O. Achieng, "Modelling of soil moisture retention curve using machine learning techniques: Artificial and deep neural networks vs support vector regression models," *Computers & Geosciences*, vol. 133, p. 104320, Dec. 2019.

[168] Rudiyanto, B. Minasny, N. W. Chaney, F. Maggi, S. Goh Eng Giap, R. M. Shah, D. Fiantis, and B. I. Setiawan, "Pedotransfer functions for estimating soil hydraulic properties from saturation to dryness," *Geoderma*, vol. 403, p. 115194, Dec. 2021.

[169] M. Weynants, H. Vereecken, and M. Javaux, "Revisiting Vereecken Pedotransfer Functions: Introducing a Closed-Form Hydraulic Model," *Vadose Zone Journal*, vol. 8, no. 1, pp. 86–95, 2009, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.2136/vzj2008.0062. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.2136/vzj2008.0062

[170] A. Singh, A. Haghverdi, H. S. Öztürk, and W. Durner, "Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System: I. The Soil Water Retention Curve," *Water*, vol. 12, no. 12, p. 3425, Dec. 2020, number: 12 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2073-4441/12/12/3425

[171] S. Chakraverty, B. K. Sahoo, T. D. Rao, P. Karunakar, and B. K. Sapra, "Modelling uncertainties in the diffusion-advection equation for radon transport in soil using interval arithmetic," *Journal of Environmental*

Radioactivity, vol. 182, pp. 165–171, Feb. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0265931X16307548

[172] J. Eliasson, I. M. Watson, and K. Weber, "Chapter 5 - In Situ Observations of Airborne Ash From Manned Aircraft," in *Volcanic Ash*, S. Mackie, K. Cashman, H. Ricketts, A. Rust, and M. Watson, Eds. Elsevier, Jan. 2016, pp. 89–98. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780081004050000094

[173] V. S. Yakovleva and R. I. Parovik, "Solution of diffusion-advection equation of radon transport in many-layered geological media," *Nukleonika*, vol. Vol. 55, No. 4, pp. 601–606, 2010. [Online]. Available: http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BUJ7-0014-0097

[174] L. M. M. L. Bakker, N. Xiao, A. A. F. v. d. Ven, M. Schaap, F. N. v. d. Vosse, and C. A. Taylor, "Image-based blood flow estimation using a semi-analytical solution to the advection–diffusion equation in cylindrical domains," *Journal of Fluid Mechanics*, vol. 924, p. A18, Oct. 2021.

[175] Y. Bazilevs, V. M. Calo, T. E. Tezduyar, and T. J. R. Hughes, "YZ discontinuity capturing for advection-dominated processes with application to arterial drug delivery," *International Journal for Numerical Methods in Fluids*, vol. 54, no. 6-8, pp. 593–608, 2007, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/fld.1484. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/fld.1484

[176] Z. Horvat, M. Isic, and M. Spasojevic, "Two dimensional river flow and sediment transport model," *Environmental Fluid Mechanics*, vol. 15, no. 3, pp. 595–625, Jun. 2015. [Online]. Available: https://doi.org/10.1007/s10652-014-9375-y

[177] A. T. N. Papanicolaou, M. Elhakeem, G. Krallis, S. Prakash, and J. Edinger, "Sediment Transport Modeling Review—Current and Future Developments," *Journal of Hydraulic Engineering*, vol. 134, no. 1, pp.

1–14, Jan. 2008, publisher: American Society of Civil Engineers. [Online]. Available: https://ascelibrary.org/doi/10.1061/%28ASCE%290733-9429%282008%29134%3A1%281%29

[178] M. Dehghan, "Numerical solution of the three-dimensional advection–diffusion equation," *Applied Mathematics and Computation*, vol. 150, no. 1, pp. 5–19, Feb. 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0096300303001930

[179] J. Dilip Kumar, K. Atul, and Y. Raja Ram, "Analytical Solution to the One-Dimensional Advection-Diffusion Equation with Temporally Dependent Coefficients," *Journal of Water Resource and Protection*, vol. 2011, Jan. 2011, publisher: Scientific Research Publishing. [Online]. Available: http://www.scirp.org/journal/PaperInformation.aspx?PaperID=3781

[180] J. Gao, Z. Li, Z. Chen, Y. Zhou, W. Liu, L. Wang, and J. Zhou, "Deterioration of groundwater quality along an increasing intensive land use pattern in a small catchment," *Agricultural Water Management*, vol. 253, p. 106953, Jul. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378377421002183

[181] A. Corey and B. Auvermann, "Transport by advection and diffusion revisited," *Vadose Zone Journal*, vol. 2, no. 4, pp. 655–663, 2003.

[182] J. Ren, Y. Chen, X. Zhang, Y. Zhao, S. Zhu, G. Zhou, B. Li, Z. Cha, and K. Zhang, "An advection-diffusion equation-based approach to discern the meteorological factor effects on particle concentrations," *Atmospheric Research*, vol. 299, p. 107213, Apr. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169809523006105

[183] S. Savović and A. Djordjevich, "Finite difference solution of the one-dimensional advection–diffusion equation with variable coefficients in semi-infinite media," *International Journal of Heat and Mass*

Transfer, vol. 55, no. 15, pp. 4291–4294, Jul. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0017931012002323

[184] S. Venkatesan, J. Droniou, I. Roy, D. Robert, and A. Zhou, "Application of diffusion-advection equations to in-field monitoring of soil suction profiles," *Computers and Geotechnics*, vol. 139, p. 104329, Nov. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0266352X2100327X

[185] D. Dapelo, S. Simonis, M. J. Krause, and J. Bridgeman, "Lattice-Boltzmann coupled models for advection–diffusion flow on a wide range of Péclet numbers," *Journal of Computational Science*, vol. 51, p. 101363, Apr. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877750321000557

[186] R. Ewing and H. Wang, "A summary of numerical methods for time-dependent advection-dominated partial differential equations," *Journal of Computational and Applied Mathematics*, vol. 128, no. 1, pp. 423–445, Mar. 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377042700005227

[187] Y. Gorb, D. Nam, and A. Novikov, "Numerical simulations of diffusion in cellular flows at high Péclet numbers," *Discrete and Continuous Dynamical Systems - B*, vol. 15, no. 1, pp. 75–92, Oct. 2010, publisher: Discrete and Continuous Dynamical Systems - B. [Online]. Available: https://www.aimsciences.org/en/article/doi/10.3934/dcdsb.2011.15.75

[188] H. Riou and P. Ladevèze, "A new numerical strategy for the resolution of high-Péclet advection–diffusion problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 241-244, pp. 302–310, Oct. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782512002022

[189] Q. He and A. M. Tartakovsky, "Physics-Informed Neural Network Method for Forward and Backward Advection-Dispersion Equations," *WATER RESOURCES RESEARCH*, vol. 57, no. 7, p. e2020WR029479, Jul. 2021, num Pages: 20 Place: Washington Publisher: Amer Geophysical Union Web of Science ID: WOS:000680092200007. [Online]. Available: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DynamicDOIArticle&SrcApp=WOS&KeyAID=10.1029%2F2020WR029479&DestApp=DOI&SrcAppSID=EUW1ED0C68h9lucVVBPhj91NMVCvE&SrcJTitle=WATER+RESOURCES+RESEARCH&DestDOIRegistrantName=American+Geophysical+Union

[190] Q. Hou, Z. Sun, L. He, and A. Karemat, "Orthogonal grid physics-informed neural networks: A neural network-based simulation tool for advection–diffusion–reaction problems," *Physics of Fluids*, vol. 34, no. 7, p. 077108, Jul. 2022. [Online]. Available: https://doi.org/10.1063/5.0095536

[191] N. Sukumar and A. Srivastava, "Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks," *Computer Methods in Applied Mechanics and Engineering*, vol. 389, p. 114333, Feb. 2022, arXiv:2104.08426 [cs, math]. [Online]. Available: http://arxiv.org/abs/2104.08426

[192] T. De Wolff, H. C. Lincopi, L. Martí, and N. Sanchez-Pi, "MOPINNs: an evolutionary multi-objective approach to physics-informed neural networks," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. Boston Massachusetts: ACM, Jul. 2022, pp. 228–231. [Online]. Available: https://dl.acm.org/doi/10.1145/3520304.3529071

[193] J. M. Hanna, J. V. Aguado, S. Comas-Cardona, R. Askri, and D. Borzacchiello, "Sensitivity analysis using Physics-informed neural networks," *Engineering Applications of Artificial Intelligence*, vol. 135, p.

108764, Sep. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197624009229

[194] W.-F. Hu, Y.-J. Shih, T.-S. Lin, and M.-C. Lai, "A shallow physics-informed neural network for solving partial differential equations on surfaces," Jan. 2023. [Online]. Available: http://arxiv.org/abs/2203.01581

[195] X. Li, J. Deng, J. Wu, S. Zhang, W. Li, and Y.-G. Wang, "Physical informed neural networks with soft and hard boundary constraints for solving advection-diffusion equations using Fourier expansions," *Computers & Mathematics with Applications*, vol. 159, pp. 60–75, Apr. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0898122124000348

[196] F. Regazzoni, S. Pagani, A. Cosenza, A. Lombardi, and A. Quarteroni, "Numerical Analysis–A physics-informed multi-fidelity approach for the estimation of differential equations parameters in low-data or large-noise regimes." *Rendiconti Lincei-Matematica E Applicazioni*, vol. 32, no. 3, pp. 437–471, Sep. 2021, publisher: European Mathematical Society Publishing House.

[197] J. Xie, H. Li, S. Su, J. Cheng, Q. Cai, H. Tan, L. Zu, X. Qu, and H. Han, "Quantitative analysis of molecular transport in the extracellular space using physics-informed neural network," *COMPUTERS IN BIOLOGY AND MEDICINE*, vol. 171, p. 108133, Mar. 2024, num Pages: 12 Place: Oxford Publisher: Pergamon-Elsevier Science Ltd Web of Science ID: WOS:001185982900001.

[198] X. Zhu, J. Liu, X. Ao, S. He, L. Tao, and F. Gao, "A Best-Fitting B-Spline Neural Network Approach to the Prediction of Advection–Diffusion Physical Fields with Absorption and Source Terms," *Entropy*, vol. 26, p. 577, Jul. 2024.

[199] A. T. A. Gomes, L. M. da Silva, and F. Valentin, "Physics-Aware Neural Networks for Boundary Layer Linear Problems," Jul. 2022. [Online].

Available: http://arxiv.org/abs/2208.12559

[200] M. Sikora, P. Krukowski, A. Paszyńska, and M. Paszyński, "Comparison of Physics Informed Neural Networks and Finite Element Method Solvers for advection-dominated diffusion problems," *Journal of Computational Science*, vol. 81, p. 102340, Sep. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877750324001339

[201] T.-Y. Hsieh and T.-H. Huang, "A multiscale stabilized physics informed neural networks with weakly imposed boundary conditions transfer learning method for modeling advection dominated flow," *ENGINEERING WITH COMPUTERS*, May 2024, num Pages: 35 Place: New York Publisher: Springer Web of Science ID: WOS:001214804600001.

[202] J. Niu, W. Xu, H. Qiu, S. Li, and F. Dong, "1-D coupled surface flow and transport equations revisited via the physics-informed neural network approach," *Journal of Hydrology*, vol. 625, p. 130048, Oct. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022169423009903

[203] H. Bo, H. Hua, H. Han, S. He, Y. Zhou, S. Liu, and Z. Zuo, "Physics-informed neural networks for advection–diffusion–Langmuir adsorption processes," *Physics of Fluids*, vol. 36, p. 81906, Aug. 2024.

[204] S. Basir and I. Senocak, "Critical Investigation of Failure Modes in Physics-informed Neural Networks," Jun. 2022. [Online]. Available: http://arxiv.org/abs/2206.09961

[205] A. S. Krishnapriyan, A. Gholami, S. Zhe, R. M. Kirby, and M. W. Mahoney, "Characterizing possible failure modes in physics-informed neural networks," Nov. 2021. [Online]. Available: http://arxiv.org/abs/2109.01050

[206] Q. Wang, Y. Ma, K. Zhao, and Y. Tian, "A Comprehensive Survey of Loss Functions in Machine Learning," *Annals of Data Science*, vol. 9, 2022.

[207] J. Farmer, C. A. Oian, B. A. Bowman, and T. Khan, "Empirical loss weight optimization for PINN modeling laser bio-effects on human skin for the 1D heat equation," *Machine Learning with Applications*, vol. 16, p. 100563, Jun. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827024000392

[208] S. Alkhadhr and M. Almekkawy, "Physics-Informed Neural Networks with Resampling Technique to Model Ultrasound Wave Propagation of a Multi-Element Transducer," in *2022 IEEE INTERNATIONAL ULTRASONICS SYMPOSIUM (IEEE IUS)*. New York: IEEE, 2022, iSSN: 1948-5719 Num Pages: 4 Series Title: IEEE International Ultrasonics Symposium Web of Science ID: WOS:000896080400021. [Online]. Available: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DynamicDOIConfProc&SrcApp=WOS&KeyAID=10.1109%2FIUS54386.2022.9957203&DestApp=DOI&SrcAppSID=EUW1ED0C73bsmp87vmxzp2RXMP0fy&SrcJTitle=2022+IEEE+INTERNATIONAL+ULTRASONICS+SYMPOSIUM+%28IEEE+IUS%29&DestDOIRegistrantName=Institute+of+Electrical+and+Electronics+Engineers

[209] R. Leiteritz and D. Pflüger, "How to Avoid Trivial Solutions in Physics-Informed Neural Networks," Dec. 2021. [Online]. Available: http://arxiv.org/abs/2112.05620

[210] L. Lu, X. Meng, Z. Mao, and G. Karniadakis, "DeepXDE: A Deep Learning Library for Solving Differential Equations," *SIAM Review*, vol. 63, pp. 208–228, Feb. 2021.

[211] T. Alkhalifah and X. Huang, "Direct Imaging Using Physics Informed Neural Networks," in *2022 IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, ICIP*. New York: IEEE, 2022, pp. 2781–2785, iSSN: 1522-4880 Num Pages: 5 Series Title: IEEE International Conference on Image Processing ICIP Web of Science ID: WOS:001058109502174.

[212] X. Zhang, Y. Zhu, J. Wang, L. Ju, Y. Qian, M. Ye, and J. Yang, "GW-PINN: A deep learning algorithm for solving groundwater flow equations," *ADVANCES IN WATER RESOURCES*, vol. 165, p. 104243, Jul. 2022, num Pages: 29 Place: Oxford Publisher: Elsevier Sci Ltd Web of Science ID: WOS:000812933500003. [Online]. Available: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=WOS&KeyAID=10.1016%2Fj.advwatres.2022.104243&DestApp=DOI&SrcAppSID=EUW1ED0C73bsmp87vmxzp2RXMP0fy&SrcJTitle=ADVANCES+IN+WATER+RESOURCES&DestDOIRegistrantName=Elsevier

[213] R. Mojgani, M. Balajewicz, and P. Hassanzadeh, "Kolmogorov n–width and Lagrangian physics-informed neural networks: A causality-conforming manifold for convection-dominated PDEs," *Computer Methods in Applied Mechanics and Engineering*, vol. 404, p. 115810, Feb. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782522007666

[214] Z. Wu, L. Jun Jiang, S. Sun, and P. Li, "A Hard Constraint and Domain-Decomposition- Based Physics-Informed Neural Network Framework for Nonhomogeneous Transient Thermal Analysis," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 14, no. 7, pp. 1215–1226, Jul. 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10562356

[215] R. Chen, H. Li, J. Wang, X. Guo, and Y. Xiang, "Evaluating soil water movement and soil water content uniformity under sprinkler irrigation with different soil texture and irrigation uniformity using numerical simulation," *Journal of Hydrology*, vol. 626, p. 130356, Nov. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022169423012982

[216] X. Ma, K. A. Sanguinet, and P. W. Jacoby, "Direct root-zone irrigation outperforms surface drip irrigation for grape yield and crop

water use efficiency while restricting root growth," *Agricultural Water Management*, vol. 231, p. 105993, Mar. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378377419312892

[217] D. K. Vishwakarma, R. Kumar, S. A. Abed, N. Al-Ansari, A. Kumar, N. L. Kushwaha, D. Yadav, A. Kumawat, A. Kuriqi, A. Alataway, A. Z. Dewidar, and M. A. Mattar, "Modeling of soil moisture movement and wetting behavior under point-source trickle irrigation," *Scientific Reports*, vol. 13, no. 1, p. 14981, Sep. 2023, publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41598-023-41435-4

[218] J. Elam and C. Björdal, "A review and case studies of factors affecting the stability of wooden foundation piles in urban environments exposed to construction work," *International Biodeterioration & Biodegradation*, vol. 148, p. 104913, Mar. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S096483051931248X

[219] C. W. W. Ng and Y. W. Pang, "Influence of Stress State on Soil-Water Characteristics and Slope Stability," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 126, no. 2, pp. 157–166, Feb. 2000, publisher: American Society of Civil Engineers. [Online]. Available: https://ascelibrary.org/doi/10.1061/%28ASCE%291090-0241%282000%29126%3A2%28157%29

[220] R. J. Sterrett and T. B. Edil, "Ground-Water Flow Systems and Stability of a Slope," *Groundwater*, vol. 20, no. 1, pp. 5–11, 1982, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1745-6584.1982.tb01324.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-6584.1982.tb01324.x

[221] D. G. Toll, Z. Abedin, J. Buma, Y. Cui, A. S. Osman, and K. K. Phoon, "The impact of changes in the water table and soil moisture on structural stability of buildings and foundation systems:

Systematic review CEE10-005 (SR90)," Jan. 2012. [Online]. Available: https://durham-repository.worktribe.com/output/1636508

[222] S. Bachmair, M. Weiler, and G. Nützmann, "Controls of land use and soil structure on water movement: Lessons for pollutant transfer through the unsaturated zone," *Journal of Hydrology*, vol. 369, no. 3, pp. 241–252, May 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022169409001115

[223] S. M. Seyedpour, M. Janmaleki, C. Henning, A. Sanati-Nezhad, and T. Ricken, "Contaminant transport in soil: A comparison of the Theory of Porous Media approach with the microfluidic visualisation," *Science of The Total Environment*, vol. 686, pp. 1272–1281, Oct. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969719321138

[224] R. X. Zhou, J. Wang, C. J. Tang, Y. P. Zhang, X. A. Chen, X. Li, Y. Y. Shi, L. Wang, H. B. Xiao, and Z. H. Shi, "Identifying soil water movement and water sources of subsurface flow at a hillslope using stable isotope technique," *Agriculture, Ecosystems & Environment*, vol. 343, p. 108286, Feb. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167880922004352

[225] L. A. Richards, "Capillary Cconduction of Liquids through porous mediums," *Physics*, vol. 1, no. 5, pp. 318–333, Nov. 1931. [Online]. Available: https://pubs.aip.org/jap/article/1/5/318/1025629/CAPILLARY-CONDUCTION-OF-LIQUIDS-THROUGH-POROUS

[226] J. R. Philip, "Theory of Infiltration," in *Advances in Hydroscience*, V. T. Chow, Ed. Elsevier, Jan. 1969, vol. 5, pp. 215–296. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9781483199368500106

[227] A. J. Pullan, "The quasilinear approximation for unsaturated porous media flow," *Water Resources Re-*

search, vol. 26, no. 6, pp. 1219–1234, 1990, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/WR026i006p01219. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/WR026i006p01219

[228] R. Srivastava and T.-C. J. Yeh, "Analytical solutions for one-dimensional, transient infiltration toward the water table in homogeneous and layered soils," *Water Resources Research*, vol. 27, no. 5, pp. 753–762, 1991, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/90WR02772. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/90WR02772

[229] A. W. Warrick and A. Amoozegar-Fard, "Infiltration and drainage calculations using spatially scaled hydraulic properties," *Water Resources Research*, vol. 15, no. 5, pp. 1116–1120, 1979, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/WR015i005p01116. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/WR015i005p01116

[230] C. Miller, C. Dawson, M. Farthing, T. Hou, J. Huang, C. Kees, C. Kelley, and H. P. Langtangen, "Numerical Simulation of Water Resources Problems: Models, Methods, and Trends," *Advances in Water Resources*, Jan. 2013.

[231] Y. Zha, J. Yang, L. Yin, Y. Zhang, W. Zeng, and L. Shi, "A modified Picard iteration scheme for overcoming numerical difficulties of simulating infiltration into dry soil," *Journal of Hydrology*, vol. 551, pp. 56–69, Aug. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022169417303712

[232] S. F. Ahmed, M. S. B. Alam, M. Hassan, M. R. Rozbu, T. Ishtiak, N. Rafa, M. Mofijur, A. B. M. Shawkat Ali, and A. H. Gandomi, "Deep learning modelling techniques: current progress, applications,

advantages, and challenges," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 13 521–13 617, Nov. 2023. [Online]. Available: https://doi.org/10.1007/s10462-023-10466-8

[233] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6472238

[234] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi, and M. Dehmer, "An Introductory Review of Deep Learning for Prediction Models With Big Data," *Frontiers in Artificial Intelligence*, vol. 3, Feb. 2020, publisher: Frontiers. [Online]. Available: https://www.frontiersin.orghttps://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2020.00004/full

[235] J. M. Hanna, J. V. Aguado, S. Comas-Cardona, R. Askri, and D. Borzacchiello, "Residual-based adaptivity for two-phase flow simulation in porous media using Physics-informed Neural Networks," *Computer Methods in Applied Mechanics and Engineering*, vol. 396, p. 115100, Jun. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S004578252200295X

[236] M. A. Nabian, R. J. Gladstone, and H. Meidani, "Efficient training of physics-informed neural networks via importance sampling," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 8, pp. 962–977, 2021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12685. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.12685

[237] B. Zapf, J. Haubner, M. Kuchta, G. Ringstad, P. K. Eide, and K.-A. Mardal, "Investigating molecular transport in the human brain from MRI with physics-informed neural networks," *Scientific Reports*, vol. 12, no. 1, p. 15475, Sep. 2022.

[238] S. Zeng, Z. Zhang, and Q. Zou, "Adaptive deep neural networks methods for high-dimensional partial differential equations," *Journal of Computational Physics*, vol. 463, p. 111232, Aug. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999122002947

[239] C. Wu, M. Zhu, Q. Tan, Y. Kartha, and L. Lu, "A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks," *Computer Methods in Applied Mechanics and Engineering*, vol. 403, p. 115671, Jan. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782522006260

[240] K. R. Sheets and J. M. H. Hendrickx, "Noninvasive Soil Water Content Measurement Using Electromagnetic Induction," *Water Resources Research*, vol. 31, no. 10, pp. 2401–2409, 1995, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/95WR01949. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/95WR01949

[241] D. Michot, Y. Benderitter, A. Dorigny, B. Nicoullaud, D. King, and A. Tabbagh, "Spatial and temporal monitoring of soil water content with an irrigated corn crop cover using surface electrical resistivity tomography," *Water Resources Research*, vol. 39, no. 5, 2003, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2002WR001581. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/2002WR001581

[242] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, Jan. 2005. [Online]. Available: http://epubs.siam.org/doi/book/10.1137/1.9780898717921

[243] A. Binley, S. S. Hubbard, J. A. Huisman, A. Revil, D. A. Robinson, K. Singha, and L. D. Slater, "The emergence of hydrogeophysics for

improved understanding of subsurface processes over multiple scales," *Water Resources Research*, vol. 51, no. 6, pp. 3837–3866, 2015, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015WR017016. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017016

[244] S. P. Friedman, "Soil properties influencing apparent electrical conductivity: a review," *Computers and Electronics in Agriculture*, vol. 46, no. 1, pp. 45–70, Mar. 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169904001255

[245] Z. Moreno, A. Arnon-Zur, and A. Furman, "Hydro-geophysical monitoring of orchard root zone dynamics in semi-arid region," *Irrigation Science*, vol. 33, no. 4, pp. 303–318, Jul. 2015. [Online]. Available: https://doi.org/10.1007/s00271-015-0467-3

[246] Y. Wang, Y. Li, Y. Song, and X. Rong, "The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition," *Applied Sciences*, vol. 10, no. 5, p. 1897, 2020.

[247] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*, 2020.

[248] R. H. Brooks, A. T. A. T. Corey, and P. Colorado State University, "Hydraulic properties of porous media," 1964, publisher: Colorado State University. Libraries. [Online]. Available: http://hdl.handle.net/10217/61288

[249] M. Van Genuchten, "A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils1," *Soil Science Society of America Journal*, vol. 44, Sep. 1980.

[250] D. Fredlund and A. Xing, "Equations for the Soil–Water Characteristic Curve," *Canadian Geotechnical Journal - CAN GEOTECH J*, vol. 31, pp.

521–532, Aug. 1994.

[251] A. Nemes, M. G. Schaap, F. J. Leij, and J. H. M. Wösten, "Description of the unsaturated soil hydraulic database UNSODA version 2.0," *Journal of Hydrology*, vol. 251, no. 3, pp. 151–162, Oct. 2001.

[252] J. De Pue, Y.-D. Botula, P. M. Nguyen, M. Van Meirvenne, and W. M. Cornelis, "Introducing a Kriging-based Gaussian Process approach in pedotransfer functions: Evaluation for the prediction of soil water retention with temperate and tropical datasets," *Journal of Hydrology*, vol. 597, p. 125770, Jun. 2021.

[253] W. Liu, X. Luo, F. Huang, and M. Fu, "Prediction of soil water retention curve using Bayesian updating from limited measurement data," *Applied Mathematical Modelling*, vol. 76, pp. 380–395, Dec. 2019.

[254] J. L. Feike, *The UNSODA unsaturated soil hydraulic database : user's manual*. Version 1.0. Cincinnati, Ohio : National Risk Management Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, [1996], 1996.

[255] R. D. De Veaux and L. H. Ungar, "Multicollinearity: A tale of two nonparametric regressions," in *Selecting Models from Data*, ser. Lecture Notes in Statistics, P. Cheeseman and R. W. Oldford, Eds. New York, NY: Springer, 1994, pp. 393–402.

[256] M. Paliwal and U. A. Kumar, "Assessing the contribution of variables in feed forward neural network," *Applied Soft Computing*, vol. 11, no. 4, pp. 3690–3696, Jun. 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494611000585

[257] J. Y.-L. Chan, S. M. H. Leow, K. T. Bea, W. K. Cheng, S. W. Phoong, Z.-W. Hong, and Y.-L. Chen, "Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review," *Mathematics*, vol. 10, no. 8, p.

1283, Jan. 2022, publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2227-7390/10/8/1283

[258] A. Haghverdi, H. S. Öztürk, and W. Durner, "Measurement and estimation of the soil water retention curve using the evaporation method and the pseudo continuous pedotransfer function," *Journal of Hydrology*, vol. 563, pp. 251–259, Aug. 2018.

[259] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, and X. Zhang, "TensorFlow: A system for large-scale machine learning," 2016.

[260] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, and G. Louppe, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, 2012.

[261] M. Pietrzyk, Madej, Rauch, and D. Szeliga, *Computational Materials Engineering: Achieving High Accuracy and Efficiency in Metals Processing Simulations*, Jan. 2015, pages: 381.

[262] C. C. Margossian, "A review of automatic differentiation and its efficient implementation," *WIREs Data Mining and Knowledge Discovery*, vol. 9, no. 4, p. e1305, 2019, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1305. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1305

[263] Z. K. Lawal, H. Yassin, D. T. C. Lai, and A. Che Idris, "Physics-Informed Neural Network (PINN) Evolution and Beyond: A Systematic Literature Review and Bibliometric Analysis," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 140, Dec. 2022, number:

4 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2504-2289/6/4/140

[264] J. Sun, Y. Chen, and X. Tang, "Physics-Informed Neural Networks with Two Weighted Loss Function Methods for Interactions of Two-Dimensional Oceanic Internal Solitary Waves," *Journal of Systems Science and Complexity*, vol. 37, no. 2, pp. 545–566, Apr. 2024. [Online]. Available: https://doi.org/10.1007/s11424-024-3500-x

[265] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, "PHYSICS-INFORMED NEURAL NETWORKS WITH HARD CONSTRAINTS FOR INVERSE DESIGN\ast," *Siam Journal on Scientific Computing*, vol. 43, no. 6, pp. B1105–B1132, 2021, place: Philadelphia Publisher: Siam Publications WOS:000736742800001. [Online]. Available: https://epubs.siam.org/doi/10.1137/21M1397908

[266] A. D. Jagtap, K. Kawaguchi, and G. E. Karniadakis, "Adaptive activation functions accelerate convergence in deep and physics-informed neural networks," *Journal of Computational Physics*, vol. 404, p. 109136, Mar. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021999119308411

[267] S. Yang, H. Kim, Y. Hong, K. Yee, R. Maulik, and N. Kang, "Data-driven physics-informed neural networks: A digital twin perspective," *Computer Methods in Applied Mechanics and Engineering*, vol. 428, p. 117075, Aug. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782524003311

[268] A. Celaya, D. Fuentes, and B. Riviere, "An Adaptive Collocation Point Strategy For Physics Informed Neural Networks via the QR Discrete Empirical Interpolation Method," Jan. 2025. [Online]. Available: https://arxiv.org/abs/2501.07700v3

[269] A. Forootani, H. Kapadia, S. Chellappa, P. Goyal, and P. Benner, "GS-PINN: Greedy Sampling for Parameter Estimation in Partial Differential Equations," May 2024. [Online]. Available: https://arxiv.org/abs/2405.08537v1

[270] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," Dec. 2019. [Online]. Available: https://arxiv.org/abs/1912.01703v1

[271] X. Zhao, Z. Gong, Y. Zhang, W. Yao, and X. Chen, "Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105516, Jan. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197622005061

[272] X. Chen, X. Chen, W. Zhou, J. Zhang, and W. Yao, "The heat source layout optimization using deep learning surrogate modeling," *Struct. Multidiscip. Optim.*, vol. 62, no. 6, pp. 3127–3148, Dec. 2020. [Online]. Available: https://doi.org/10.1007/s00158-020-02659-4

[273] L. Russo, "Parsimonious Physics-Informed Random Projection Neural Networks for Initial-Value Problems of ODEs and index-1 DAEs," *Cornell University - arXiv*, Mar. 2022. [Online]. Available: https://www.academia.edu/92420511/Parsimonious_Physics_Informed_Random_Projection_Neural_Networks_for_Initial_Value_Problems_of_ODEs_and_index_1_DAEs

[274] A. Daw, J. Bu, S. Wang, P. Perdikaris, and A. Karpatne, "Rethinking the Importance of Sampling in Physics-informed Neural Networks," Jul. 2022. [Online]. Available: https://arxiv.org/abs/2207.02338v1

[275] S.-M. Qin, M. Li, T. Xu, and S. Dong, *RAR-PINN algorithm for the data-driven vector-soliton solutions and parameter discovery of coupled nonlinear equations*, Apr. 2022.

[276] G. I. Taylor, "Dispersion of soluble matter in solvent flowing slowly through a tube," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 219, no. 1137, pp. 186–203, Aug. 1953, publisher: Royal Society. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rspa.1953.0139

[277] J. a. M. T. v. Genuchten, "Contaminant Transport in the Unsaturated Zone: Theory and Modeling," in *The Handbook of Groundwater Engineering*, 3rd ed. CRC Press, 2016, num Pages: 34.

[278] N. Guven, "Dynamics of fluids in porous media," Jan. 1972. [Online]. Available: https://www.academia.edu/3367245/Dynamics_of_fluids_in_porous_media

[279] R. J. Millington and J. P. Quirk, "Permeability of porous solids," *Transactions of the Faraday Society*, vol. 57, no. 0, pp. 1200–1207, Jan. 1961, publisher: The Royal Society of Chemistry. [Online]. Available: https://pubs.rsc.org/en/content/articlelanding/1961/tf/tf9615701200

[280] A. W. Warrick, D. O. Lomen, and S. R. Yates, "A Generalized Solution to Infiltration," *Soil Science Society of America Journal*, vol. 49, no. 1, pp. 34–38, 1985.

[281] J. Šimůnek, M. T. van Genuchten, and M. Šejna, "Recent Developments and Applications of the HYDRUS Computer Software Packages," *Vadose Zone Journal*, vol. 15, no. 7, p. vzj2016.04.0033, 2016, _eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2136/vzj2016.04.0033. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.2136/vzj2016.04.0033

[282] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical Evaluation of Rectified Activations in Convolutional Network," 2015.

[283] A. K. Dubey and V. Jain, "Comparative Study of Convolution Neural Network's Relu and Leaky-Relu Activation Functions," *Applications of Computing, Automation and Wireless Systems in Electrical Engineering*, pp. 873–880, 2019.

[284] L. R. Ahuja, F. Fiedler, G. H. Dunn, J. G. Benjamin, and A. Garrison, "Changes in Soil Water Retention Curves Due to Tillage and Natural Reconsolidation," *Soil Science Society of America Journal*, vol. 62, no. 5, pp. 1228–1233, 1998.

[285] C. F. Stange and R. Horn, "Modeling the Soil Water Retention Curve for Conditions of Variable Porosity," *Vadose Zone Journal*, vol. 4, no. 3, pp. 602–613, 2005, _eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2136/vzj2004.0150. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.2136/vzj2004.0150

[286] E. Rabot, M. Wiesmeier, S. Schlüter, and H. J. Vogel, "Soil structure as an indicator of soil functions: A review," *Geoderma*, vol. 314, pp. 122–137, Mar. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016706117310273

[287] W. Hu, R. Cichota, M. Beare, K. Müller, J. Drewry, and A. Eger, "Soil structural vulnerability: Critical review and conceptual development," *Geoderma*, vol. 430, p. 116346, Feb. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S001670612300023X

[288] M. B. Almendro-Candel, I. G. Lucas, J. Navarro-Pedreño, A. A. Zorpas, M. B. Almendro-Candel, I. G. Lucas, J. Navarro-Pedreño, and A. A. Zorpas, "Physical Properties of Soils Affected by the Use of Agricultural

Waste," in *Agricultural Waste and Residues*. IntechOpen, Aug. 2018. [Online]. Available: https://www.intechopen.com/chapters/61756

[289] S. Yang and T.-H. Lu, "Study of Soil-Water Characteristic Curve Using Microscopic Spherical Particle Model," *Pedosphere*, vol. 22, no. 1, pp. 103–111, Feb. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1002016011601963

[290] S. Assouline, "Modeling the Relationship between Soil Bulk Density and the Water Retention Curve," *Vadose Zone Journal*, vol. 5, no. 2, pp. 554–563, May 2006. [Online]. Available: https://acsess.onlinelibrary.wiley.com/doi/10.2136/vzj2005.0083

[291] J. R. Nimmo, "Modeling Structural Influences on Soil Water Retention," *Soil Science Society of America Journal*, vol. 61, no. 3, pp. 712–719, 1997.

[292] A. R. Dexter, "Soil physical quality: Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth," *Geoderma*, vol. 120, no. 3, pp. 201–214, Jun. 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016706103002891

[293] Y. Zhao, Y. Cui, H. Zhou, X. Feng, and Z. Huang, "Effects of void ratio and grain size distribution on water retention properties of compacted infilled joint soils," *Soils and Foundations*, vol. 57, no. 1, pp. 50–59, Feb. 2017.

[294] S. Javanshir, H. Bayat, and A. S. Gregory, "Effect of free swelling index on improving estimation of the soil moisture retention curve by different methods," *CATENA*, vol. 189, p. 104479, Jun. 2020.

[295] Z. Hu, Z. Yang, Y. Wang, G. E. Karniadakis, and K. Kawaguchi, "Bias-Variance Trade-Off in Physics-Informed Neural Networks with Randomized Smoothing for High-Dimensional PDEs," *SIAM Journal on Scientific Computing*, vol. 47, no. 4, pp. C846–C872, Aug. 2025, publisher: Society for Industrial and Applied Mathematics. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/23M1621356

[296] X. Liu, X. Zhang, W. Peng, W. Zhou, and W. Yao, "A novel meta-learning initialization method for physics-informed neural networks," *Neural Computing and Applications*, vol. 34, no. 17, pp. 14 511–14 534, Sep. 2022. [Online]. Available: https://doi.org/10.1007/s00521-022-07294-2

[297] J. Barry-Straume, A. Sarshar, A. A. Popov, and A. Sandu, "Physics-Informed Neural Networks for PDE-Constrained Optimization and Control," *Communications on Applied Mathematics and Computation*, Aug. 2025. [Online]. Available: https://doi.org/10.1007/s42967-025-00499-x

[298] Z. Gao, T. Tang, L. Yan, and T. Zhou, "Failure-Informed Adaptive Sampling for PINNs, Part II: Combining with Re-sampling and Subset Simulation," *Communications on Applied Mathematics and Computation*, vol. 6, no. 3, pp. 1720–1741, Sep. 2024. [Online]. Available: https://doi.org/10.1007/s42967-023-00312-7

[299] X. Ding and A. El-Zein, "Predicting soil water retention curves using machine learning: A study of model architecture and input variables," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108122, Jul. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095219762400280X

[300] ——, "Water Retention Curves of Clayey Soils by Artificial Neural Networks with Uneven Datasets," in *Challenges and Innovations in Geomechanics*. Cham: Springer International Publishing, 2023, pp. 45–52.

[301] K.-K. Phoon, J. Ching, and Z. Cao, "Unpacking data-centric geotechnics," *Underground Space*, vol. 7, no. 6, pp. 967–989, Dec. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2467967422000514

[302] K.-K. Phoon, , and W. Zhang, "Future of machine learning in geotechnics," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 17, no. 1, pp. 7–22, Jan. 2023, publisher: Taylor &

Francis _eprint: https://doi.org/10.1080/17499518.2022.2087884. [Online]. Available: https://doi.org/10.1080/17499518.2022.2087884

[303] Y. Otake and Y. Honjo, "Challenges in geotechnical design revealed by reliability assessment: Review and future perspectives," *Soils and Foundations*, vol. 62, no. 3, p. 101129, Jun. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038080622000373

[304] S. M. Harle and R. L. Wankhade, "Machine learning techniques for predictive modelling in geotechnical engineering: a succinct review," *Discover Civil Engineering*, vol. 2, no. 1, p. 86, May 2025. [Online]. Available: https://doi.org/10.1007/s44290-025-00224-w

[305] S. Wu, Y. Otake, Y. Higo, and I. Yoshida, "Pathway to a fully data-driven geotechnics: Lessons from materials informatics," *Soils and Foundations*, vol. 64, no. 3, p. 101471, Jun. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038080624000490

[306] N. Monte, F. Bucci, F. A. Mevoli, M. Santangelo, P. Reichenbach, L. Di Matteo, and I. Marchesini, "A dataset of geotechnical parameters based on international literature to characterise lithotypes in Italy," *Scientific Data*, vol. 11, no. 1, p. 1371, Dec. 2024, publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41597-024-04095-1

[307] A. Tombari, L. Stefanini, G. L. D. Nicosia, L. M. J. Holland, and M. Dobbs, "Geotechnical data-driven possibility reliability assessment," *Computers and Geotechnics*, vol. 185, p. 107311, Sep. 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0266352X25002605

[308] Y. Ranjith, "2D and 3D Physics Informed Neural Networks to Model Pollution Spread with Obstructions," in *Proceedings of the IEEE/ACM 10th International Conference on Big Data Computing, Applications and Technologies*, ser. BDCAT '23. New York, NY, USA: Association

for Computing Machinery, Apr. 2024, pp. 1–2. [Online]. Available: https://dl.acm.org/doi/10.1145/3632366.3632380

[309] E. B. Diaw, F. Lehmann, and P. Ackerer, "One-dimensional simulation of solute transfer in saturated–unsaturated porous media using the discontinuous finite elements method," *Journal of Contaminant Hydrology*, vol. 51, no. 3, pp. 197–213, Oct. 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169772201001292

[310] J. Rekha, S. P. Suma, B. Shilpa, U. Khan, S. M. Hussain, A. Zaib, and A. M. Galal, "Solute transport exponentially varies with time in an unsaturated zone using finite element and finite difference method," *International Journal of Modern Physics B*, vol. 37, no. 09, p. 2350089, Apr. 2023, publisher: World Scientific Publishing Co. [Online]. Available: https://www.worldscientific.com/doi/10.1142/S0217979223500893

[311] M. K. Singh, A. Chatterjee, and P. Kumari, "Mathematical Modeling of One-Dimensional Advection Dispersion Equation in Groundwater Contamination Using Different Velocity and Dispersion for Different Zones," in *Applications of Fluid Dynamics*, M. Singh, B. Kushvah, G. Seth, and J. Prakash, Eds. Singapore: Springer, 2018, pp. 585–592.

[312] T. Fu, B. Yuan, W. Zhang, J. Wang, and S. Liu, "Particle swarm optimization-based PINN for direct-simulation of fluids flow in subsurface reservoir," *Geoenergy Science and Engineering*, vol. 254, p. 214010, Nov. 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2949891025003689

[313] S. S M, P. Kumar, and V. Govindaraj, "A novel optimization-based physics-informed neural network scheme for solving fractional differential equations," *Engineering with Computers*, vol. 40, no. 2, pp. 855–865, Apr. 2024. [Online]. Available: https://doi.org/10.1007/s00366-023-01830-x

[314] H. Wu, H. Luo, Y. Ma, J. Wang, and M. Long, "RoPINN: Region Optimized Physics-Informed Neural Networks," Oct. 2024, arXiv:2405.14369 [cs]. [Online]. Available: http://arxiv.org/abs/2405.14369

[315] Z. Hao, J. Yao, C. Su, H. Su, Z. Wang, F. Lu, Z. Xia, Y. Zhang, S. Liu, L. Lu, and J. Zhu, "PINNacle: A Comprehensive Benchmark of Physics-Informed Neural Networks for Solving PDEs," *Advances in Neural Information Processing Systems*, vol. 37, pp. 76 721–76 774, Dec. 2024. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2024/hash/8c63299fb2820ef41cb05e2ff11836f5-Abstract-Datasets_and_Benchmarks_Track.html

[316] P. Escapil-Inchauspé and G. A. Ruz, "h-Analysis and data-parallel physics-informed neural networks," *Scientific Reports*, vol. 13, no. 1, p. 17562, Oct. 2023, publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41598-023-44541-5

[317] X. Wang, X. Zhu, X. Meng, Z. Zhu, S. Zhang, and T. Song, "Inference and training acceleration of deep learning partial differential equation solver," *The Journal of Supercomputing*, vol. 81, no. 5, p. 733, Apr. 2025. [Online]. Available: https://doi.org/10.1007/s11227-025-07204-y

APPENDICES

A Steady-state Analytical solution of ADE

The PDE for the 1D ADE sets the time derivative term u_t in Eq. 1 to be 0, transforming the transient formula into its steady-state form as follows:

$$N_x[u] = -\left[\nabla \cdot (D\nabla u) + \nabla \cdot (vu)\right] = 0, \quad x \in \Omega,\ t \in [0, t_{\text{lim}}] \tag{A.1}$$

To solve Eq. A.1 analytically, we assume:

$$u = Ae^{\phi x} \tag{A.2}$$

Substituting Eq. A.2 into Eq. A.1:

$$DA\phi^2 e^{\phi x} - vA\phi e^{\phi x} = 0 \tag{A.3}$$

$$Ae^{\phi x}(D\phi^2 - v\phi) = 0 \tag{A.4}$$

$$D\phi^2 - v\phi = 0 \tag{A.5}$$

The equation above is a quadratic equation, the solution of ϕ can be easily calculated as follows:

$$\phi_{1,2} = \frac{v \pm v}{2D} \tag{A.6}$$

The closed form solution of concentration at any given location x can be

expressed as:

$$u(x) = A_1 e^{\phi_1 x} + A_2 e^{\phi_2 x} \tag{A.7}$$

Let we define two new variables:

$$p_{(1,x)} = e^{\phi_1 x}; \quad p_{(2,x)} = A_2 e^{\phi_2 x} \tag{A.8}$$

Substituting the solution from Eq. A.6 to Eq. A.8, we got:

$$p_{(1,x)} = e^{\frac{v}{D}x}; \quad p_{(2,x)} = e^{0 \cdot x} = 1 \tag{A.9}$$

Thus, the closed-form solution can be rewritten as:

$$u(x) = A_1 p_{(1,x)} + A_2 p_{(2,x)} \tag{A.10}$$

The values of A_1 and A_2 can be calculated based on the boundary conditions set up. The Dirichlet boundary conditions are applied at both ends of the domain and expressed as: $u(x_1 = a) = u_a$ and $u(x_2 = b) = u_b$. Then, applying the boundary conditions to Eq. A.10, we can get a pair of equations with two unknowns:

$$\begin{cases} u_a = A_1 + A_2 \\ u_b = A_1 e^{\frac{vL_d}{D}} + A_2 \cdot 1 \end{cases} \tag{A.11}$$

Solving the equation:

$$\begin{cases} A_2 = \frac{u_a e^{\frac{vL}{D}} - u_b}{e^{\frac{vL_d}{D}} - 1} \\ A_1 = u_a - A_2 \end{cases} \tag{A.12}$$

After substituting A_1 and A_2 into Eq. A.10, we can get the closed form solution as:

$$u(x) = \left(u_a - \frac{u_a e^{\frac{vL_d}{D}} - u_b}{e^{\frac{vL_d}{D}} - 1} \right) e^{\frac{v}{D}x} + \frac{u_a e^{\frac{vL_d}{D}} - u_b}{e^{\frac{vL_d}{D}} - 1} \tag{A.13}$$

B Generalised Solution for RRE

The performance of the PINN model for solving RRE is evaluated by comparing its predictions with a generalised solution proposed by Warrick et al. [280]. The generalised solution uses a reduced form framework, which simplifies the RRE (Eq.3.36) by replacing its dimensional parameters by dimensionless ones. The reduced form parameters are defined as follows:

$$W = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{B.1}$$

$$T = \frac{\mu K_s t}{\theta_s - \theta_r} \tag{B.2}$$

$$X = \mu x \tag{B.3}$$

$$K^* = \frac{K}{K_s} \tag{B.4}$$

$$h^* = \mu h \tag{B.5}$$

where W, T, X, K^* and h^* are all dimensionless parameters, representing respectively water content, time, soil depth, hydraulic conductivity and suction pressure. $\mu, [L^{-1}]$ is a positive scaling factor for length. Note that the dimensionless water content W is the degree of saturation of the soil. Substituting these dimensionless variables into Eq.3.36, the reduced form of RRE is derived as:

$$\frac{\partial W}{\partial T} = \frac{\partial}{\partial X}(K^* \frac{\partial h^*}{\partial X}) - \frac{\partial K^*}{\partial X} \tag{B.6}$$

The reduced form RRE is subject to the following initial and boundary conditions, given by:

$$W(X,0) = W_i \qquad \text{(initial condition)} \tag{B.7}$$

$$W(0,T) = W_0 \qquad \text{(boundary condition at } X = 0) \tag{B.8}$$

$$W(X \to \infty, T) = W_\infty \qquad \text{(boundary condition at } X \to \infty) \tag{B.9}$$

where W_i is the initial degree of saturation, and W_0 and W_∞ are the degrees of saturation are at $X = 0$ and $X = \infty$, respectively. The spatial domain is semi-infinite, extending to $X \to \infty$, implying that the comparison between the PINN model and the generalised solution is valid only prior to the system reaching hydraulic equilibrium.

Similar to the dimensional form of RRE, the reduced form variables can also be interrelated using empirical WRC models, such as Van Genuchten or Brooks-Corey model. In this analysis, the reduced form Van Genuchten model is adopted, expressed as:

$$W = \frac{1}{(1 + |h^*|^n)^m} \tag{B.10}$$

The solution to Eq.B.6 provides the infiltration profile in reduced form, expressed as the soil depth $X(W)$ corresponding to a given water content W. The solution is given by:

$$X(W) = \lambda(W)T^{0.5} + \chi(W)T + \psi(W)T^{1.5} + ... \tag{B.11}$$

$X(W)$ can usually be approximated with sufficient accuracy using the first three terms of the series. The coefficients $\lambda(W)$, $\chi(W)$, $\psi(W)$ in the first three terms can be determined numerically by using the method outlined by Philip [226], or obtained from the lookup table as presented in Table B.1. Once the solutions in reduced form have been determined, the results field can be transformed back to their dimensional forms by using Eq. B.1 to Eq. B.5.

Table B.1: Values of $\lambda(W)$, $\chi(W)$ and $\psi(W)$ defining moisture profiles for $W^* = 0.25$, 0.5, and 0.75. [280]

		λ			χ			ψ		
n	W_i	$W^* = 0.25$	0.50	0.75	$W^* = 0.25$	0.50	0.75	$W^* = 0.25$	0.50	0.75
1.1	0.0	0.314	0.314	0.312	0.266	0.266	0.269	0.200	0.200	0.200
	0.1	0.316	0.316	0.313	0.282	0.282	0.288	0.258	0.258	0.255
	0.2	0.321	0.320	0.316	0.299	0.300	0.305	0.336	0.336	0.334
	0.3	0.329	0.328	0.320	0.317	0.319	0.334	0.443	0.443	0.445
1.25	0.0	0.475	0.472	0.446	0.185	0.187	0.208	0.160	0.161	0.171
	0.1	0.503	0.498	0.463	0.198	0.202	0.229	0.183	0.186	0.202
	0.2	0.537	0.527	0.483	0.213	0.220	0.256	0.213	0.218	0.244
	0.3	0.579	0.562	0.504	0.232	0.245	0.292	0.250	0.261	0.301
1.5	0.0	0.725	0.710	0.642	0.229	0.239	0.280	0.119	0.121	0.130
	0.1	0.772	0.748	0.665	0.250	0.266	0.317	0.137	0.140	0.154
	0.2	0.828	0.791	0.691	0.278	0.302	0.364	0.159	0.165	0.186
	0.3	0.895	0.840	0.720	0.316	0.352	0.426	0.188	0.198	0.232
1.75	0.0	0.884	0.857	0.763	0.284	0.301	0.355	0.113	0.112	0.118
	0.1	0.941	0.901	0.790	0.315	0.340	0.404	0.129	0.129	0.140
	0.2	1.008	0.951	0.819	0.355	0.392	0.465	0.150	0.152	0.171
	0.3	1.088	1.000	0.853	0.409	0.461	0.528	0.183	0.217	0.217
2.0	0.0	0.990	0.955	0.846	0.331	0.354	0.413	0.110	0.107	0.111
	0.1	1.054	1.003	0.876	0.369	0.402	0.469	0.127	0.124	0.127
	0.2	1.128	1.058	0.909	0.419	0.464	0.539	0.146	0.145	0.165
	0.3	1.215	1.120	0.947	0.487	0.546	0.628	0.171	0.174	0.216
2.5	0.0	1.120	1.077	0.957	0.401	0.429	0.489	0.107	0.100	0.105
	0.1	1.191	1.130	0.992	0.449	0.488	0.552	0.122	0.122	0.127
	0.2	1.273	1.191	1.031	0.513	0.564	0.628	0.150	0.152	0.171
	0.3	1.370	1.260	1.077	0.598	0.661	0.728	0.160	0.161	0.202