Feature engineering using homogenization theory with multiscale perturbation analysis for supervised model-based learning of physical clogging condition in seepage filters

Chew, Alvin Wei Ze; Law, Adrian Wing-Keung

2019

Chew, A. W. Z., & Law, A. W.-K. (2019). Feature engineering using homogenization theory with multiscale perturbation analysis for supervised model-based learning of physical clogging condition in seepage filters. *Journal of Computational Science*, 32, 21–35.
doi:10.1016/j.jocs.2019.02.003

<https://hdl.handle.net/10356/144554>

<https://doi.org/10.1016/j.jocs.2019.02.003>

© 2019 Elsevier B.V. All rights reserved. This paper was published in *Journal of Computational Science* and is made available with permission of Elsevier B.V.

Downloaded on 26 Apr 2023 14:28:12 SGT

Feature engineering using homogenization theory with multiscale perturbation analysis for supervised model-based learning of clogging behavior in seepage filters

Alvin Wei Ze CHEW¹, *Adrian Wing-Keung LAW^{1, 2}

*Corresponding author's email: cwklaw@ntu.edu.sg

¹ School of Civil and Environmental Engineering, Nanyang Technological University, N1-01c-98, 50 Nanyang Avenue, Singapore 639798

² Environmental Process Modelling Centre (EPMC), Nanyang Environment and Water Research Institute (NEWRI), 1 Cleantech Loop, CleanTech One, #06-08, Singapore 637141

Abstract

Aquifer recharge and recovery systems (ARRS) in natural or engineered aquifers are gaining attention worldwide due to climate changes and associated rising demands for larger system storage capacity for potable water needs. ARRS can broadly be analysed as seepage depth filters whereby engineering predictions of their complex physical clogging behavior continue to be challenging, which has since hindered the predictive maintenance of these large-scale systems. Upon detections of their stagnated low-flux condition in ARRS, i.e. breakthrough stage, immediate maintenance must be carried out which can be both expensive and time-consuming to restore their normal operations. To address this current problem statement, we first leverage the homogenization theory with the multiscale perturbation analysis as the feature engineering step to reduce the complexity of the physical clogging behavior in ARRS. The analytical approach systematically derives a unique homogenized representation, by volume-averaging the fluid and transport properties inside a pre-defined periodic cell domain, which then quantifies the clogging dynamics at the macroscale. Following which, a series of physical parameters is identified from the derived homogenized representation to build a pre-processed input layer into our own multi-layered neural network (NN) architecture for predictive analysis. Measured data extracted from the literature is then used to train and verify the NN model. By far, the model yields an average error deviation of 20% between the model's predictions and the respective measurements for an optimized set of hyperparameters tested. We then quantitatively discuss how a trained NN model can be adhered to predict the timing for a concerned ARRS to reach its breakthrough stage for a range of operational conditions. Finally, for the first time in the literature, we also demonstrate how the homogenized representation can be useful to determine an arbitrary filter's critical

reaction rate and diffusion coefficient which are responsible for the filter's physical clogging behavior over time.

Keywords: homogenization theory; multi-scale perturbation; supervised machine learning; aquifer recharge and recovery systems; mass transport

***This paper is an extended version of our conference paper 327 (Chew and Law., 2018) which was presented at ICCS 2018 in Wuxi, China. ***

1. Introduction

1.1 Background

With an increasing global population and uncertainties due to climatic changes [1-2], i.e. droughts, excessive rainfalls etc., there is now a greater demand for aquifer recharge and recovery systems (ARRS) in natural or engineered aquifers worldwide [3] to contribute stable supplies of potable water to local communities. ARRS can store massive volumes of stormwater for future uses and provide the following environmental benefits [2-4]: (a) pre-treatment of infiltrations with minimum energy requirements; (b) reduction in seawater intrusions; and (c) minimization in land subsistence. In comparison to open-surface storages, ARRS also minimizes the evaporation rate and avoids additional site-constructions. Countries which are currently deploying ARRS include the United States, Australia, Netherlands and Israel [3]. In addition, ARRS are most suitable for deployment in arid locations, near coastal regions, and regions having significant wet and dry weather conditions [3,5].

ARRS can generally be analysed as seepage depth filters whereby creeping flow condition occurs inside the global pore spaces inherent to any random porous media system. Notwithstanding the benefits provided by ARRS, there are still two compelling problems faced by engineers today [3,5], namely: (a) the complex clogging of these systems during operations due to the infiltration and physical trapping of turbidity particles, biofouling due to microbial growth, and precipitation of particles due to pH changes and redox conditions; and (b) the pollution of the stored groundwater due to the injection of contaminated waters or the unexpected leakage of the initially removed contaminants within the filter. For the latter problem, a source optimal control model was developed most recently to simulate the flow and transport processes in porous media of groundwater systems which link the predictions to pollutant concentration deviation, abatement cost and pollution tax [6]. Issue (a) has been an ongoing research which, however, still remains difficult to address effectively [3, 7] and thus establishes the problem statement in this study.

1.2. Problem statement

At present, engineers are still unable to predict the timing, with sufficient accuracy, for a concerned ARRS system to reach its maximum clogging stage, i.e. stagnated global permeability within the filter, based on the known parameters during the early stages of operations as illustrated in Figure 1. This behavior is commonly known as the filter's breakthrough stage. Upon detections of breakthrough, the concerned ARRS undergoes maintenance which can be operationally expensive. Examples of the maintenance acts include scrubbing, well jetting, vacuum pumping [3,8], mechanical surging and heating [3,9]. Hence, maintenance of ARRS can be both time- and cost-intensive in terms of the labour and machinery required, and logistical problems can occur in unexpected breakthrough events. Lengthy downtime for ARRS is particularly undesirable as it affects the supplies of potable water, especially for arid regions. These concerns underline the need for an engineering model which can perform predictive maintenance of ARRS with sufficiently good accuracy.

Overall, the model needs to be able to estimate the operational clogging rate inside ARRS which, however, remain challenging due to the complex random packing of the porous media inside the depth filters. The random packing complicates the inherent fluid dynamics and transport of the influent's contaminants within the filter's open spaces. Prediction of the physical clogging rate during the early stages of operations is desirable as it enables the model to extrapolate the effects of the continuous clogging on the filter's global permeability coefficient for determining the best possible timing to perform maintenance. Experiments have since been carried out to analyse the clogging behavior in lab-scale soil filter setups which emulated ARRS [10-14]. In those works, the authors examined varying types of contaminants treated by soil filters and discussed the complexity of the specific type of clogging behavior involved. However, an engineering model having the above-described predictive capability is still missing in today's context which thus constitutes the focus of this study.

1.3. Scope and objective of study

Machine learning (ML) has garnered increasing attention due to significant improvements to their inherent algorithms which can improve the predictive controls of various water-related industrial applications. Several examples include drinking water networks [15] and in-situ bioremediation for petroleum contamination [16]. To

further harness ML's advantages, we recognize an important opportunity to couple ML with traditional analytical methods which may better address complex engineering problems. An excellent example is the most recent work by [17] whereby the authors developed a novel proof-of-concept neural ordinary differential equation (ODE) to model complex time-series problems with deep learning. Their work has since gained attention due to the revolution made to traditional neural networks modelling. In our analysis, we leverage the homogenization theory with the multiscale perturbation analysis as the feature engineering step to first reduce the complexity of the physical clogging behavior in ARRS which then enable us to pre-identify a series of physical parameters required to build the engineering predictive model for ARRS. We hypothesize that the proposed approach is most suitable to address engineering problems having limited dataset.

Generally, the homogenization theory involves the derivations of average homogenized representations, having rapidly varying coefficients, which quantify the macroscopic behavior with a sufficiently small microstructure, i.e. a periodic cell domain. To systematically derive the homogenized representations, we adhere to the multiscale perturbation analysis for averaging the microscale variations inside the periodic cell by using a sufficiently small parameter to perform the asymptotic expansions of the concerned fluid and transport parameters. Consequently, the macroscale behavior is then encapsulated within the cell domain which thus reduces the inherent complexity of the original problem.

For porous media flow systems, the same analytical method systematically captures the macroscale seepage filtration behavior to within a pre-defined microscale cell problem, which contains singular or multiple arbitrary media grains, while imposing the periodic boundary conditions for the fluid and transport properties within the same cell to derive the homogenized representations. We then pre-identify the important physical parameters from the homogenized representation which are associated with the seepage filtration problem in ARRS. Dimensional analysis is carried out with those identified parameters to build the pre-processed input layer into our own multi-layered neural network (NN) architecture for supervised model-based learning to predict the breakthrough time of a concerned ARRS and thus achieve predictive maintenance. To the best of our knowledge, the proposed methodology has not been performed

before for ARRS systems, or for environmental and chemical engineering applications in general.

1.4 Literature review

In the following, we describe previous research studies which used the homogenization theory with the multiscale perturbation analysis to model the flow and transport behaviours in porous media systems, and how their research generally differs from ours in this study.

Notably, [18-19] derived the homogenized Darcy's Law equation for saturated porous media by considering the flow past a periodic array of rigid media, followed by the numerical computation of the hydraulic conductivity, i.e. permeability tensor, inside the periodic cell domain. [20-22] derived the convective-dispersive equations exactly and analysed the dispersion of a passive solute in seepage flow inside a spatially periodic domain. [23] adopted similar methodologies to analyse the transport of volatile organic compounds during soil vapor extraction, while [24] extensively computed the permeability and dispersion of heat or solute in the periodic cell. [25] derived the characteristic models for four varying flow phenomena inside the microscale domain to analyse the formation damage in the macroscopic porous media due to erosion and deposition of solid particles. [26] investigated the transport of contaminants in fractured porous media under varying local Peclet (Pe) numbers, based on the assumption that both convection and molecular diffusion were of equal importance inside the periodic domain. More recently, the effects of non-uniform flow and dispersion were also analysed in heterogeneous porous media by [27] which involved the Fourier series analysis. Likewise, [28] numerically resolved a system of homogenized equations, via the finite-element method, to compute the fluid permeability through a periodic array of cylinders having varying densities and angles. The above-mentioned studies have since provided an extensive quantitative understanding of the flow and transport mechanisms in porous media systems which can be useful for the engineering designs of depth filters in general. However, accounting for the dynamic clogging condition is still missing in those models.

[29-30] adopted an extensive sequential homogenization technique to build macroscopic engineering models which quantify the reactive transport in porous media systems including polydisperse media. Notably, these models can derive suitable effective homogenized equations for varying Peclet and Damköhler numbers in relations to contaminants hydrology. Likewise, however, their analyses have not yet been extended to the actual operational conditions of large-scale filters such as ARRS in terms of model calibration or estimations of field parameters. Hence, using their models to perform the predictive maintenance of real-world depth filter systems (lab-scale, pilot-scale, or industrial-scale) may still be a major limitation.

Only in recent years, the analytical technique is being adhered to model depth filter's clogging condition by estimating the transient variations to the microscale fluid properties within the periodic cell domain. For example, [31] analysed the transport of colloids and investigated the variation to the microstructure during the attachment and detachment of colloidal particles in a two-dimensional (2D) saturated porous media structure by coupling the surface reaction rate and Nernst-Planck equations. [32] demonstrated how a decreasing porosity gradient along the filter's depth can maximise its trapping capability. They later modelled the changes to the microscale media properties to quantify the filter's blockage [33]. The theoretical novelty of these works is well-appreciated as they can analytically estimate the porosity distribution which can achieve homogeneous clogging inside the filter. Cross-validation or calibration of their analytical predictions with available measured data has, however, not been carried out by far which is required for extending the model's usefulness for practicality. [34] investigated the effects of varying microstructure on the filter's removal performance and effective homogenized parameters for a range of filtration scenarios such as constant-flux and constant-pressure flow cases. The authors, however, acknowledged that it still remains difficult to use their analytical model to predict experimental data which includes determining the filter's realistic parameters such as the adsorption coefficient etc.

In our study, we model the physical clogging behavior of ARRS systems as an idealized deposition of turbidity particles at the boundaries of collector spheres under periodic flow conditions. At any specific time, the assumed deposition stage results in larger collector spheres which reduce the periodic cell's porosity, and hence the

macroscale filter's permeability due to the imposed homogenization theory. The major difference of our research with previous notable works [29-30, 32-33] is that we avoid solving for the homogenized parameters (permeability, diffusion coefficient etc.) exactly, or via traditional numerical means such as using computational fluid dynamics (CFD), which often require the user to adopt an idealized cell domain that may not be emulative of real-world depth filters generally. Instead, we maintain an arbitrary arrangement of the media grains inside the periodic cell domain and adopt measured data to perform supervised model-based learning with the derived homogenized representation for quantifying the physical clogging behavior of ARRS. Subsequently, we demonstrate how the built engineering model can be useful to decision-makers to predict the timing for a concerned ARRS to reach its breakthrough stage for predictive maintenance, and how the derived homogenized representation can be useful to estimate the filter's realistic parameters responsible for its physical clogging stage over time.

This paper is structured as follows. Section 2 describes the developmental details of the theoretical model to emulate the physical clogging condition in seepage filters (ARRS). The key homogenized equations are then derived in Section 3. Details relating to the pre-processing step and the multi-layered neural network model are provided in Section 4. In Section 5, supervised model-based learning with the homogenized representation is carried out with measured data from the literature to build the engineering model which can perform the predictive maintenance of ARRS systems. The model then uses available data to iterate the critical reaction rate and diffusion coefficient responsible for the physical clogging behavior inside the concerned filter. Quantitative details are also provided to demonstrate how engineers can use the model to estimate the timing for a concerned ARRS to reach its breakthrough stage for a range of operational conditions. Finally, Section 6 provides an overview of our study.

2. Development of theoretical model

2.1. Seepage filtration model

In actuality, porous media grains, i.e. sand, soil materials [35], are randomly packed inside the macroscale aquifer which complicate the fluid dynamics and contaminants' transport within the available void spaces of the system. To reduce the overall complexity of the seepage filtration problem as part of the proposed feature engineering extraction step, we model the aquifer as a network of non-overlapping three-dimensional (3D) rigid collector spheres which emulate the deployed media. For the homogenization theory to be valid, we assume a periodic microstructure (Figure 2) which is arbitrarily represented by a 3D box having $O(l)$, whereby l is the characteristic size of the media grains and is very much smaller than the characteristic aquifer's depth of L . We note that this configuration differs substantially from idealized packing geometries adopted in previous notable works whereby the homogenized parameters such as the effective diffusion coefficients, permeability etc. can be determined exactly or numerically [29-30, 32-33] for modelling the clogging condition. Instead, our current packing geometry aligns more towards the random configuration adhered in the recent work of [34] for examining the fluid and mass transport problems. In our analysis, a very small length scale $\left(\left(\varepsilon = \frac{l}{L}\right) \ll 1\right)$ is defined to perform the multiscale perturbation analysis which systematically discretizes the macroscale problem into a series of smaller scale problems for simplifications. Generally, the characteristic size (l) of each periodic cell is computed as follows:

$$l = \sqrt[3]{\frac{V_s}{1 - \theta_0}} \quad (2.1)$$

where V_s is the total defined volume of collector spheres present in each cell domain which depends directly on the sphere's initial diameter size (d_0^*), and θ_0 is the filter's initial clean bed porosity value.

The solid spheres within each periodic cell domain are defined as $\Omega_s \subset \mathbb{R}^3$. The interface between the solid and fluid phases is defined as $\Omega_{fs} \subset \mathbb{R}^3$. Coupling both phases then results in the whole periodic cell which is defined as $\Omega \subset \mathbb{R}^3$. To maintain a disordered arrangement of the collector spheres in each cell domain, the average pore-scale separation between the grains/collector spheres (μ) must be much smaller than l [37] as shown in Figure 2. This assumption is valid for real-world seepage filters as the media grains are usually randomly compacted under fully-saturated conditions.

We underline that the current packing configuration in Figure 2 is only arbitrary, i.e. random, as there can be many different packing variations inside the periodic cell domain. However, regardless of the packing choice, the pore-scale fluid properties (pressure, velocity and concentration) are assumed to be spatially periodic from cell to cell while varying slowly over L . Finally, the physical trapping of fine turbidity particles over time is modelled by idealized transient increases in the sizes of the collector spheres along the macroscale media depth.

The fluid motion at very low microscale Reynolds numbers ($\ll 1$) in the available pore space of each periodic cell domain is governed by the incompressible unsteady-state Stokes' equation in (2.2) and mass continuity equation in (2.3) respectively.

$$\frac{\partial u_i^*}{\partial t^*} = -\frac{\partial p^*}{\partial x_i^*} + \frac{\partial}{\partial x_j^*} \left(\nu \frac{\partial u_i^*}{\partial x_j^*} \right), \quad x^* \in \Omega_f \quad (2.2)$$

$$\frac{\partial u_i^*}{\partial x_i^*} = 0, \quad x^* \in \Omega_f \quad (2.3)$$

where x^* is the position vector, u^* the velocity vector, ν the fluid kinematic viscosity, and p^* the fluid pressure.

As discussed, the clogging behavior inside the macroscale filter is due to the complex trapping, i.e. deposition, of fine turbidity particles present in the recharge. These particles are initially transported inside the pore spaces of the filter via advection and diffusion:

$$\frac{\partial c^*}{\partial t^*} + \frac{\partial (c^* u_i^*)}{\partial x_i^*} = \frac{\partial}{\partial x_i^*} \cdot \left(D_p^* \frac{\partial c^*}{\partial x_i^*} \right), \quad x^* \in \Omega_f \quad (2.4)$$

where c^* is the concentration of fine turbidity particles and t^* is time. D_p^* in (2.4) represents an unknown coefficient responsible for the complex diffusion of the turbidity particles inside the seepage filter, due to both their molecular diffusion and the zig-zag trajectories inherent to the randomly packed filter. For random media packing, the zig-zag effect has been confirmed to be dominant [38-39]. By far, there have been selected studies in the literature which provided quantitative representations to estimate D_p^* for depth filters settings [40-41]. It is worth noting that [40-41] only used a singular media grain (collector sphere) to model the diffusion transport mechanism.

To the best of our knowledge, a generalised equation to predict the complex D_p^* value inside real-world seepage filters having random packing arrangement is not available at present.

In terms of interfacial mass transport, we assume that the turbidity particles undergo a first-order deposition process at the sphere's boundary (S^*), as defined in (2.5), due to the complex diffusion mechanism occurring inside the available pore spaces of the periodic cell domain.

$$-\frac{\frac{\partial S^*}{\partial x_i^*}}{|\nabla S^*|} \cdot \left(D_p^* \frac{\partial c^*}{\partial x_i^*} \right) = k_{fs} c^*, \quad x^* \in \Omega_{fs} \quad (2.5)$$

where $\frac{\partial S^*}{\partial x_i^*}$ is the outward normal vector acting on the sphere's boundary, and k_{fs} the unknown reaction rate [LT^{-1}] responsible for the interfacial transport.

As the size of collector spheres (d^*) grows along the macroscale media depth due to the deposition process, i.e. clogging, there will be a resulting outward normal velocity (an unknown quantity) acting at the boundary of each growing sphere as defined in the following:

$$-\frac{\frac{\partial S^*}{\partial x_i^*}}{|\nabla S^*|} \left(\frac{\partial d^*}{\partial t^*} \right) = u_i^*, \quad x^* \in \Omega_{fs} \quad (2.6)$$

To couple the growth of the sphere's initial diameter along the media depth to the accumulation of turbidity particles around the sphere's boundary, we adhere to the following volume conservation principle:

$$\frac{\partial d^*}{\partial t^*} = -\frac{1}{\rho_p} \frac{\frac{\partial S^*}{\partial x_i^*}}{|\nabla S^*|} \cdot \left(D_p^* \frac{\partial c^*}{\partial x_i^*} \right), \quad x^* \in \Omega_{fs} \quad (2.7)$$

where ρ_p is the bulk density of the deposits of turbidity particles.

(2.2 – 2.4) are well-known representations which are commonly known as the Navier-Stokes fluid flow and mass transport equations respectively. We still retain these fundamental equations in the above derivations for the following reasons:

- i. (2.2 – 2.3) quantify the creeping flow condition inherent to the filter's global void spaces
- ii. (2.4) quantifies the general mass transport of turbidity particles inside the void spaces of the porous media system before undergoing an assumed first-order deposition process at the surface boundaries of the collector spheres in (2.5)
- iii. Upon particles' deposition, the overall diameters of the collector spheres grow over time as represented in (2.6), which growth rate is estimated by the volume conservation principle in (2.7)

We do note that (2.5 – 2.7) are not commonly used in standard model derivations for modelling the filter's physical clogging condition. The closest forms are found in [29-30, 32-33] which, however, have not been cross-validated with measured data or being used for predictions with real-world operational conditions. Finally, the following additional assumptions are adhered for our conceptual model: (i) macroscale pressure gradient acting upon the seepage filter is balanced by the viscous resistance around each growing microscale collector sphere during the transient clogging build-up; and (ii) detachment of the deposited turbidity particles inside the seepage filter is not considered over time.

2.2. Normalization procedures

In this section, we provide the important scaling variables to normalize the fundamental equations of (2.2 – 2.7) and the required length- and time-scales to quantify the macroscale physical clogging behavior in ARRS. Major differences of our analysis from previous notable studies will also be indicated along the subsequent discussions.

To first perform the normalization procedures, we adopt the following scaling variables: (i) $c^* = c_o c$; (ii) $t^* = T t$; (iii) $u_i^* = U u_i$; (iv) $x_i^* = l x_i$; (v) $p^* = P p$; and (vi) $d^* = l d$; whereby T , U and P are the respective scales for the time, velocity and pressure parameters. c_o is the influent concentration of fine turbidity particles. Three basic macroscopic time scales are also defined as follows: (a) convection time scale (T_c) in (2.8); (b) reaction time scale (T_R) in (2.9); and (c) diffusion time scale (T_D) in (2.10);

$$T_c = \frac{L}{U} \quad (2.8)$$

$$T_R = \frac{L}{k_{fs}} \quad (2.9)$$

$$T_D = \frac{(L)^2}{D_p^*} \quad (2.10)$$

As stated above, we expect a balance between the macroscale pressure gradient acting upon the seepage filter and the viscous flow resistance around each growing microscale collector sphere in (2.11). Since $d^* = O(l)$, we can represent the characteristic length of each collector sphere as l to facilitate the subsequent analysis, i.e.

$$\frac{P}{L} \sim O\left(\frac{vU}{l^2}\right) \quad (2.11)$$

From (2.11), the scaling variable P has the ordering importance of $\frac{vUL}{l^2}$. The dimensionless microscale Reynolds number (Re_l) and Peclet number (Pe_l) are also defined in (2.12) and (2.13) respectively as follows:

$$Re_l = \frac{Ul}{v} \quad (2.12)$$

$$Pe_l = \frac{Ul}{D_p^*} \quad (2.13)$$

We also propose a relationship between the dimensionless macroscale ($D_{a,L}$) and microscale Damköhler ($D_{a,l}$) number in the following:

$$D_{a,L} = \frac{T_D}{T_R} = \frac{k_{fs}l}{\varepsilon D_p^*} = \frac{D_{a,l}}{\varepsilon} \quad (2.14)$$

$D_{a,l}$ in (2.14) is defined as the ratio between the rate of particles deposition at the boundary of each sphere multiplying the characteristic length of the periodic cell domain, and the complex diffusion coefficient responsible for the deposition process. The macroscale convection time scale of T_c of (2.8) is used to normalize the mass transport problem in (2.4) and (2.6 – 2.7) as we are affirmative that the physical

clogging behavior occurs over the convection time scale as evident from earlier experimental studies [10-14]. In the previous works of [32-33], the authors instead used the contaminant's reaction time inside the fluid phase which is expected to be much longer than the fluid residence time. Hence, our homogenized representation derived in this study will be different from theirs.

As the collector spheres grow slowly over time, the cell's porosity will reduce concurrently which affect the seepage velocity. However, we avoid using any time scaling for the flow problem as we are considering quasi-steady-state creeping flow conditions in this analysis. As discussed previously, we adopt the spatial scaling of $x_i^* = lx_i$ to normalize the spatial terms of (2.2 – 2.7). By performing the above normalizations, the final respective dimensionless forms for (2.2) to (2.7) are defined as follows:

$$\varepsilon^2 \frac{\partial u_i}{\partial t} = -\frac{\partial p}{\partial x_i} + \varepsilon \frac{\partial}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} \right), \quad x \subset \Omega_f \quad (2.15a)$$

$$\frac{\partial u_i}{\partial x_i} = 0, \quad x \subset \Omega_f \quad (2.15b)$$

$$\varepsilon \frac{\partial c}{\partial t} + \frac{\partial(cu_i)}{\partial x_i} = Pe_l^{-1} \frac{\partial}{\partial x_i} \left(\frac{\partial c}{\partial x_i} \right), \quad x \subset \Omega_f \quad (2.15c)$$

$$-\frac{\frac{\partial S}{\partial x_i}}{|\nabla S|} \cdot \left(\frac{\partial c}{\partial x_i} \right) = \varepsilon D_{a,L} c, \quad x \subset \Omega_{fs} \quad (2.15d)$$

$$-\varepsilon \frac{\frac{\partial S}{\partial x_i}}{|\nabla S|} \frac{\partial d}{\partial t} = u_i, \quad x \subset \Omega_{fs} \quad (2.15e)$$

$$\varepsilon \frac{1}{\beta} \frac{\partial d}{\partial t} = -\frac{\frac{\partial S}{\partial x_i}}{|\nabla S|} \cdot \left(\frac{\partial c}{\partial x_i} \right), \quad x \subset \Omega_{fs} \quad (2.15f)$$

where $\beta = \frac{c_o D_p^*}{\rho_p U l}$. Note that again this unique parameter differs from that of [31-33] due to the different timescale used for normalization in our study.

The derived ordering importance in (2.15a) is based on (2.11), and its inherent Re_l number is already taken to be sufficiently small from the use of (2.2) previously due to creeping flow conditions ($Re_l \ll 1$) assumptions for seepage filters. In the later section,

we shall demonstrate the following perturbation links to quantify the physical clogging stage: (a) between the flow and transport problems as represented by (2.15a – b) and (2.15c) respectively; (b) between the growth rate of the collector sphere in (2.15f) and the deposition rate of turbidity particles at the sphere’s boundary in (2.15d); and (c) between the outward velocity acting at the sphere’s boundary in (2.15e) and the flow problem in (2.15a – b). Finally, we compact the original complexity of the physical clogging problem by homogenizing (2.15a – f) via the multiscale perturbation analysis. Prior to performing the above-mentioned details, we first extensively discuss the key assumptions adhered in our analysis, namely: (a) the use of the periodic cell domain; (b) the localization principle; (c) discretization of the macroscale problem into a sequence of smaller-scale problems; and (d) the types of scaling techniques adopted.

2.3. Key assumptions

2.3.1. Periodic cell domain

The complex macroscale system is hierarchical and a sufficiently small portion of the macroscale problem is defined with a periodic unit-cell in our study (Figure 2) as demonstrated earlier. To fulfil the periodic boundary condition for the cell, rapid variations to the pore-scale fluid properties are averaged out within the cell domain having a small length scale (l) while the same fluid properties vary slowly over the macroscale length scale. By assuming periodic conditions, the effective fluid parameters associated with the macroscale problem can then be derived systematically from the cell problem via the volume averaging technique. The unit-cell is a convenient choice for volume averaging as it reduces the complexity of mathematical derivations and splits the macroscale problem into a hierarchy of scales which then allows the decomposition of the concerned parameters into high and low spatial frequency components.

We do highlight that the choice of the unit cell often invites debate as balancing the need to capture the important features of the macroscale problem while avoiding the need to solve highly complex formulations is difficult. One major disadvantage of using the periodic boundary conditions is the inability to capture anisotropic features. Overall, the average quantities defined inside a periodic cell domain correspond to the

volume averaging procedure over the same cell which is independent of the averaging support and its geometry. Alternatively, a non-periodic cell can be used whereby the values or/and fluxes of the fluid parameters jump at the cell's boundaries. However, this may result in errors for problems close to the percolation threshold.

In summary, imposing the periodic boundary conditions for the unit cell must have two key pre-requisites: (a) the size of the unit cell must be sufficiently small, i.e. $\varepsilon \rightarrow 0$, for very weak or negligible variations of the fluid properties and their respective fluxes within the cell; and (b) any rapid fluctuations to the values and fluxes of the fluid properties must be averaged out within the unit-cell to maintain continuity/conservation at the boundaries of the same cell.

2.3.2. Localization principle

The localization principle lies with constructing a homogenized model which is local in both space and time, whereby the fluid properties vary slowly over the macroscale space and time by using small dimensionless perturbation parameters respectively. This localization assumption is most important for the upscaling procedure with the homogenization theory but also imposes the largest constraints on the perturbation parameters whereby they approach the value of zero. Applying the localization principle to the fluid properties results in their average quantities defined within the periodic cell after volume integration over a suitable local variable which is usually the length scale of the concerned unit cell.

Typically, the addition between the average quantities and appropriate perturbation decompositions is used to separate the low- and high-spatial frequencies of the fluid properties. Depending on the user and the nature of the problem, this constraint can be relaxed to an extent to derive more complex and difficult formulations which can have a broader domain of validity. However, additional computational cost must be incurred to solve the more difficult formulations. In our analysis, we avoid making assumptions about the unknown homogenized parameters, and instead adopt supervised model-based training to calibrate the unknown parameters with available measured data.

2.3.3. Discretization of macroscale problem into sequence of smaller-scale problems

By considering the limit of $\varepsilon \rightarrow 0$, the macroscale problem is systematically discretized into a sequence of smaller scale problems. For example, from macroscale to microscale to mesoscale etc. A sufficiently small value of ε enables the user to clearly separate the complex macroscale problem into distinct scales for systematic analysis before appropriate upscaling, via the volume averaging technique, for engineering predictions. The selection of the ε value often requires a good mix of domain knowledge, intuition and a clear understanding of the engineering problem for appropriate downscaling.

2.3.4. Scaling techniques

The multiscale perturbation analysis requires careful selections of the scaling variables to formulate the dimensionless terms, prior to the homogenization step, as each combination results in a different set of effective homogenized equations which are representative of the macroscale behavior. Furthermore, each scaling decision needs an appropriate timescale relationship, i.e. diffusion, advection or reaction related etc., with the macroscale length scale. Generally, for mass transport problems in porous media which are governed by advection and diffusion mechanisms, the following example cases can be proposed with respect to the microscale Peclet number (Pe_l) and macroscale Peclet number (Pe_L) whereby both scales are uniquely separated by ε .

- If diffusion dominates within the local cell problem, i.e. $Pe_l = O(\varepsilon^2)$, then it will also dominate at the macroscale $Pe_L = O(\varepsilon)$ whereby only tortuosity effects are important.
- If diffusion dominates within the local cell problem, i.e. $Pe_l = O(\varepsilon)$, then both advection and diffusion are equally important at the macroscale $Pe_L = O(1)$ whereby an advection-dispersion equation exists having only tortuosity effects.
- If diffusion and advection are in balance within the local cell problem, i.e. $Pe_l = O(1)$, then advection dominates at the macroscale $Pe_L = O(\varepsilon^{-1})$ whereby an advection-dispersion equation exists having both hydrodynamic dispersion and tortuosity effects.

3. Homogenization procedures

We first introduce the multiple-scale coordinates of x and $x' = \varepsilon x$, whereby x is the fast variable defined within the periodic cell and x' is the slow variable spanning across the macroscale seepage filter. The perturbation expansions for the fluid parameters (which are all cell-periodic) are expressed as follows

$$G = G^{(0)} + \varepsilon G^{(1)} + \varepsilon^2 G^{(2)} + \dots \quad (3.1)$$

where G can be p , c and u_i .

The following spatial derivative is adopted to perform the multiscale perturbation expansion of (2.15a – f).

$$\frac{\partial}{\partial x_i} \rightarrow \frac{\partial}{\partial x_i} + \varepsilon \frac{\partial}{\partial x_i'} \quad (3.2)$$

To homogenize the fluid flow and transport problems with the multiscale perturbation analysis, readers are referred to Supplementary Section A and our previous paper [41] for the detailed mathematical details involved to derive the respective homogenized representations which are shown in the next two sub-sections. These equations effectively quantify the core physics of the fluid and mass transport problems responsible for the physical clogging behavior at the macroscale.

3.1. Fluid flow problem

For the fluid flow problem, the homogenized effective equation in (3.3), derived at the ordering importance of ε , relates the spatial variation of the dimensionless seepage velocity (RHS of (3.3)) to the direct multiplication between the growth rate of the collector sphere, i.e. the dimensionless outward velocity acting at the sphere's boundary, and the slowly changing cell's porosity (LHS of (3.3)).

$$\frac{|S(x, t)|}{|\Omega|_V} \left(\frac{\partial d}{\partial t} \right) = \frac{\partial u_i^{(0)}}{\partial x_i'} = - \frac{\partial}{\partial x_i'} \left(\tilde{k}_{ij}(x, t) \frac{\partial p^{(0)}}{\partial x_j'} \right) \quad (3.3)$$

where $S(x, t)$ is the slowing changing boundary of the sphere over space and time which affects the volume of fluid inside the cell domain, $|\Omega|_V$ the fixed volume of the cell domain, and $\tilde{k}_{ij}(x, t)$ the slowly changing permeability of the cell problem over

space and time. Overall, (3.3) demonstrates the direct perturbation link between the outward velocity acting at the sphere's boundary in (2.15e) and the flow problem in (2.15a – b) at the ordering importance of ε for the filter's physical clogging stage.

3.2. Mass transport problem

From the direct coupling of (2.15d) and (2.15f), it is obvious that the growth rate of the collector sphere is only directly linked to the deposition rate of turbidity particles (quantified by $D_{a,L}$) at the ordering importance of ε . Again, by using (3.1) and (3.2) to perform the multiscale perturbation expansion of (2.15c) while relating to (2.15d) and (2.15f), we derive (3.4) and (3.5) (see Supplementary Section A) for the ordering importance of ε and ε^2 respectively.

$$\theta(x, t) \frac{\partial c^{(0)}}{\partial t} + \tilde{u}_i^{(0)}(x, t) \frac{\partial c^{(0)}}{\partial x_i'} = -Pe_l^{-1} \frac{|S(x, t)|}{|\Omega|_V} \frac{1}{\beta} \frac{\partial d}{\partial t}, \quad (3.4)$$

$$\begin{aligned} \theta(x, t) \frac{\partial c^{(1)}}{\partial t} + \tilde{u}_i^{(0)}(x, t) \frac{\partial c^{(1)}}{\partial x_i'} + \tilde{u}_i^{(1)}(x, t) \frac{\partial c^{(0)}}{\partial x_i'} \\ = Pe_l^{-1} \theta(x, t) \frac{\partial}{\partial x_i'} \left(\frac{\partial c^{(0)}}{\partial x_i'} \right) - Pe_l^{-1} \frac{|S(x, t)|}{|\Omega|_V} D_{a,L} c^{(1)}, \end{aligned} \quad (3.5)$$

From these equations, it can be observed that the perturbation link between the respective homogenized transport and flow problems is achieved via the slowly changing seepage flow velocity parameter, i.e. $\tilde{u}_i^{(0)}(x, t)$. We then consider the linear time-derivative of the averaged concentration parameter, i.e. $\tilde{c} = \theta(x, t)(c^{(0)} + \varepsilon c^{(1)} + \dots)$, as:

$$\frac{\partial \tilde{c}}{\partial t} = \theta(x, t) \frac{\partial c^{(0)}}{\partial t} + \varepsilon \theta(x, t) \frac{\partial c^{(1)}}{\partial t} + O(\varepsilon^2) \quad (3.6)$$

By using the above (3.4) and (3.5), we arrive at the final form of (3.7), accurate up to $O(\varepsilon^2)$, for (3.6) which quantifies the averaged behavior of turbidity particles (transport and trapping) inside seepage filters.

$$\frac{\partial \tilde{c}}{\partial t} = -\frac{\partial(\tilde{u}_i c)}{\partial x_i'} - D_{a,L} Pe_l^{-1} \frac{|S(x, t)|}{|\Omega|_V} c + \varepsilon Pe_l^{-1} \theta(x, t) \frac{\partial}{\partial x_i'} \left(\frac{\partial c}{\partial x_i'} \right) + O(\varepsilon^2), \quad (3.7)$$

At this stage, $D_{a,L}$ and Pe_l^{-1} remain important unknowns in (3.7) for real-world seepage filters. During the depth filtration process, we expect the trapping of turbidity particles to be most significant near to the injection point. Hence, spatial variation to the seepage velocity is also expected which further modifies (3.7) into:

$$\frac{\partial \tilde{c}}{\partial t} = -\tilde{u}_i(x, t) \frac{\partial c}{\partial x_i'} - c \frac{|S(x, t)|}{|\Omega|_V} \frac{\partial d}{\partial t} - D_{a,L} Pe_l^{-1} \frac{|S(x, t)|}{|\Omega|_V} c + \varepsilon Pe_l^{-1} \theta(x, t) \frac{\partial}{\partial x_i'} \left(\frac{\partial c}{\partial x_i'} \right) + O(\varepsilon^2), \quad (3.8)$$

For one-dimensional seepage filter which results have been reported in the literature, we can assume quasi-steady-state clogging condition and compare the terms in (3.8) at the leading order to obtain:

$$0 = -\tilde{u}_K(x_K, t) \frac{\partial c}{\partial x_K'} - c \frac{|S(x_K, t)|}{|\Omega|_V} \frac{\partial d}{\partial t} - D_{a,L} Pe_l^{-1} \frac{|S(x_K, t)|}{|\Omega|_V} c, \quad (3.9)$$

where x_K is along the seepage direction towards the discharge point of the filter. Using (A14) in its averaged form for unidirectional flow, (3.9) becomes:

$$0 = \tilde{k}(x_K, t) \frac{\partial p}{\partial x_K'} \frac{\partial c}{\partial x_K'} - \frac{|S(x_K, t)|}{|\Omega|_V} \left(\frac{\partial d}{\partial t} + D_{a,L} Pe_l^{-1} \right) c, \quad (3.10)$$

where $\tilde{k}$ is the average permeability coefficient of the one-dimensional seepage filter which changes slowly over space and time.

(3.10) generally represents an inhomogeneous first-order differential equation for c in the following form:

$$0 = \tilde{k}(x_K, t) \frac{\partial p}{\partial x_K'} \frac{\partial c}{\partial x_K'} - A(x_K, t) c, \quad (3.11)$$

where $A(x_K, t) = \frac{|S(x_K, t)|}{|\Omega|_V} \left(\frac{\partial d}{\partial t} + D_{a,L} Pe_l^{-1} \right)$ which is directly associated with the physical parameters responsible for operating seepage filters.

Integrating (3.11) and adhering to the influent's boundary condition of $c(0) = 1$ yields:

$$\tilde{k}(x_K, t) = \frac{A(x_K, t)}{\ln(c) \frac{\partial p}{\partial x_K'}} x_K' + C_1 = - \frac{m(x_K, t)}{\frac{\partial p}{\partial x_K'}} x_K' + C_1, \quad (3.12)$$

where $m(x_K, t)$ is the lumped calibration function of $-\frac{A(x_K, t)}{\ln(c)}$ which in turn comprises of $D_{a,L}$, Pe_l^{-1} and $|S(x_K, t)|$, and C_1 is an unknown integration constant which varies between 0 and 1 since $\tilde{k}(x_K, t)$ is a dimensionless parameter and too varies between 0 and 1. Note that $\ln(c)$ is always negative, i.e. $c < 1$, at the filter's discharge point during its effective physical clogging stage, and $\frac{\partial p}{\partial x_K'}$ is also a negative quantity along the media depth as gravitational force dictates the driving force for the depth filtration process.

By leveraging the homogenization theory with the multiscale perturbation analysis systematically, the calibration function of $m(x_K, t)$ is derived in (3.12) which compacts the complexity of the physical clogging problem inside the one-dimensional (1D) seepage filters. In the subsequent sections, we focus on the maximum dimensionless media depth, i.e. $x_K = \varepsilon^{-1}$, and demonstrate how supervised model-based learning is adhered to train $m(\varepsilon^{-1}, t)$ with available dataset for 1D seepage filters from the literature.

4. Results and discussions

4.1. Supervised model-based learning

We further pre-process (3.12) via dimensional analysis to build a pre-processed input layer into our own multi-layered neural network (NN) architecture for supervised model-based learning. The entire pre-processing procedure involved to reduce the initial unknown input layer, relating to the physical clogging problem of ARRS, to five unique input nodes for our selected NN model is illustrated in Figure 4.

4.1.1. Pre-processing step with dimensional analysis

The unknown field parameters of $D_{a,L}$, Pe_l^{-1} , $|S(x_K, t)|$ and $\frac{\partial d}{\partial t}$ in (3.12) are associated with the known physical parameters for seepage filters, namely: (i) initial grain size (d_0^*); (ii) media depth (L_m); (iii) hydraulic loading rate (V_s); (iv) influent's concentration (c_0); (v) size of turbidity particles (d_p); (vi) density of turbidity particles (ρ_p); (vii) gravitational acceleration (g); (viii) kinematic viscosity (ν); and (ix) filtration time (t^*).

Readers are referred to Table 1 which explains the inclusion of each parameter in affecting $\tilde{k}(x_K, t)$. Adopting the repeating variables of ρ_p , g and L_m for the Buckingham pi-group theorem, the following pi-terms are obtained as:

$$\begin{aligned} \pi_1 &= \tilde{k}(x_K, t), & \pi_2 &= \frac{d_0^*}{L_m}, & \pi_3 &= \frac{V_s}{(L_m)^{0.5} g^{0.5}}, & \pi_4 &= \frac{c_0}{\rho_p}, & \pi_5 &= \frac{t^* g^{0.5}}{(L_m)^{0.5}}, \\ \pi_6 &= \frac{d_p}{L_m}, & \pi_7 &= \frac{\nu}{(L_m)^{1.5} g^{0.5}} \end{aligned} \quad (4.2)$$

At this stage, access to large volumes of real-time operational data for field applications of ARRS is difficult as they are usually not available to the public. Hence, we extract available dataset from the literature for the experimental analysis of ARRS [11-12] as summarized in Figure 3a and Table 2. Closer examination of the raw dataset in the Figure 3a indicates a characteristic exponential decay for $\tilde{k}(\varepsilon^{-1}, t)$, particularly for the dataset of [11] (Figures 3a to 3c in their paper). Such characteristic profile hinders data extrapolation for the later stages of operations if the predictive model is only trained by dataset associated with the early stages of operations. To address this, we linearize the general problem (Figure 3b), to the best possible extent, by taking the natural logarithm of both sides in (3.12) to obtain:

$$\ln(\tilde{k}(x_K, t)) = \ln\left(-\frac{m(x_K, t)}{\frac{\partial p}{\partial x_K'}} x_K' + C_1\right) = B(x_K, t), \quad (4.3)$$

For convenience, we consider the lumped term of $\ln\left(-\frac{m(x_K, t)}{\frac{\partial p}{\partial x_K'}} x_K' + C_1\right)$ as $B(x_K, t)$ which is also related to the same physical parameters of the above-mentioned (i) to (ix) for $\tilde{k}(x_K, t)$. Hence, we retain the dimensionless pi-terms from (4.2) to model $B(x_K, t)$. The following assumptions are also adhered to establish the respective quantitative relationship between each independent pi-term and $B(x_K, t)$ whereby each condition generally leads to higher $\ln(\tilde{k})$ values over time, i.e. gradual reduction of $\tilde{k}$, in (4.3).

- i. Invert $\frac{d_0^*}{L_m}$ as smaller grains tend to result in higher trapping efficacy of the turbidity particles inside the seepage filter

- ii. Higher V_s values increase the mass flux of turbidity particles into the filter which result in greater mass of particles to be trapped per unit time
- iii. Higher c_0 values have the tendency for higher mass/volume of turbidity particles to be trapped inside the filter
- iv. Smaller ρ_p values usually result in greater number of turbidity particles to be trapped inside the filter for each mass/volume of influent filtered
- v. Invert $\frac{d_p}{L_m}$ as smaller-sized turbidity particles tend to result in greater number of particles to be trapped inside the filter for each mass/volume of influent filtered
- vi. Invert $\frac{\nu}{(L_m)^{1.5}g^{0.5}}$ as lower fluid viscosity tends to result in higher removal efficacy of turbidity particles inside the filter for each mass/volume of influent filtered

In summary, $B(x_K, t)$ can be modelled as follows:

$$B(x_K, t) = f\left(\frac{L_m}{d_0^*}, \frac{V_s}{(L_m)^{0.5}g^{0.5}}, \frac{t^*g^{0.5}}{(L_m)^{0.5}}, \frac{c_0}{\rho_p}, \frac{L_m}{d_p}, \frac{(L_m)^{1.5}g^{0.5}}{\nu}\right) \quad (4.4)$$

In (4.4), there are two possible ways to build the pre-processed input layer into our own multi-layered NN architecture for supervised model-based learning of the filter's physical clogging problem, namely: (a) establish six independent nodes whereby each node represents a unique dimensionless pi-term; or (b) formulate the simplest monomial form for (4.4) in (4.5) whereby it represents a unique input node into the NN model.

$$B(x_K, t) = n_1 \left(\frac{L_m}{d_0^*} \cdot \frac{V_s}{(L_m)^{0.5}g^{0.5}} \cdot \frac{t^*g^{0.5}}{(L_m)^{0.5}} \cdot \frac{c_0}{\rho_p} \cdot \frac{L_m}{d_p} \cdot \frac{(L_m)^{1.5}g^{0.5}}{\nu} \right)^{n_2} \quad (4.5)$$

whereby n_1 and n_2 are unknown calibration constants with specific emphasis of n_1 being the weighted average of the compacted representation as the only input feature. We note that (4.5) can have a biased constant term after calibration with the NN module. In the next sub-section, we explain the selection of the NN architecture in our study.

4.1.2. Selection of multi-layered neural network (NN) architecture

At this stage, the limited training dataset extracted prevents us from using convolutional or recurrent neural network architectures in a satisfactory manner. Both

architectures are more suitable for deep learning purpose which is out-of-scope in our present study. Hence, we restrict our supervised model-based learning to a relatively small neural network structure (see Table 3 for descriptions) having three hidden layers whereby each has three, three and two nodes respectively, while the pre-processed input layer is based on (4.5) as discussed earlier. We note that the current selection of the NN model is only based on trial and error and varying the size of the hidden layer extensively is not the focus. Rather, we vary the respective values of the hyperparameters (training period, learning rate, number of epochs and number of mini batches) for the selected NN structure to determine the most suitable combination to analyse the filter's physical clogging problem. Finally, the current NN algorithm uses the Xavier Initialization as it maintains enough activation values and backpropagation gradients during the training step. Hence, the solution converges more quickly as compared to that of the standard random initialization which tends to converge more slowly and can easily fall into an ineffective local minimum during the early stages of the training step.

At the same time, we avoid directly using simple regression analysis or least squares method to fit the available measured data for the $\ln(\tilde{k})$ parameter in (4.3) as both techniques result in an empirical equation which, however, does not link the calibrated coefficients to the known physical parameters associated with the filter's physical clogging problem. Hence, predictive maintenance is not possible as the resulting equation is unable to relate the hybrid variations of the physical parameters to the breakthrough timings for a given ARRS. On the contrary, the NN algorithm, having an optimized set of hyperparameters, enables us to vary the pre-processed input layer flexibly and predict their respective effects on the filter's breakthrough timing. This advantage of combining the analytical feature engineering step with the selected NN architecture will be demonstrated in the later sections.

4.1.3. Training dataset

Prior to the training step, the number of measured data points is limited as evident in Figure 3 which hinders the initial model-based learning. To increase the size of the training dataset, we adhere to the following procedures (commonly known as "handling the missing data" step in machine learning) to impute logical values into the pool of training data.

- i. Extract the available data points for $\ln(\tilde{k}(\varepsilon^{-1}, t))$ at the maximum media depth associated with a defined training period from each experiment
- ii. Perform linear regression analysis of $\ln(\tilde{k}(\varepsilon^{-1}, t))$ over time of the defined training period as extracted from step (i)
- iii. Check the respective high coefficient of determination (R^2) values; ensure that R^2 should be of minimum value of 0.90. If so, proceed to step (v)
- iv. Attempt a higher-order regression analysis or other types of machine learning models (decision trees, support vector machines etc.) of the measured data and perform step (ii – iii) again using the alternate machine learning model
- v. Adopt the derived fitting model from step (ii) to impute more data values for $\ln(\tilde{k}(\varepsilon^{-1}, t))$, via interpolation, associated with the same training period defined from step (i)

In our analysis, we adhere to the above steps (i – v) to generate a minimum of 1000 training data points for each of the referenced experiments of [11-12] as shown in Figures 5 and 6 respectively. We note that the current training period constitutes from the first 30 to 35% of the total operational period for each referenced experiment. The generated data is then adhered to train the NN model (Figure 4) for performing supervised model-based learning of the physical clogging problem in the referenced experiments. We emphasize that the above procedure to impute more training data, via regression analysis, does not directly contribute to building up the predictive model as the empirical coefficients derived in Figures 5 and 6 cannot be linked with the known physical parameters in (4.4) as discussed.

4.2. Training and testing steps

4.2.1. Variations of hyperparameters

To first determine the optimized set of hyperparameters to train the proposed multi-layered NN predictive model in Figure 4, we vary the following parameters (see Table 3), namely: (a) the learning rate (α_1) of values 0.01, 0.001 and 0.0001; (b) the sizes of mini-batches (α_2) of values of 4, 8 and 16; and (c) the number of epochs (α_3) of values 100, 300 and 500. For this purpose, we only examined the extensive variations of the

hyperparameters using the measured data from [11] (Figure 4c in their paper) and [12] (Figure 3b in their paper). For the latter component, we only use the raw dataset from their actual Figure 3b to examine the hyperparameters' variations as the other dataset from Figures 3a and 3c in their paper [12] were also derived from the same filter configuration previously. Hence, the optimized set of hyperparameters determined for [12] (Figure 3b in their paper) from our analysis should also work for the other dataset derived from the same filter setup (Figures 3a and 3c in their paper).

For each combination of the varied hyperparameters, the model is being trained for several times with continuous updating of its model parameters from the previous step till the RMSE value from the validation step is sufficiently small, i.e. ideally less than 10% error deviation. We note that 80% of the training dataset, after imputation, is fixed to train the NN model while the remaining 20% is used for validation during the same training period. In Tables 4a and 4b, we list the respective RMSE values computed for the comparison between the model predictions and the respective measured data from the entire training period (training plus validation steps) for the respective experimental works of [11] (Figure 4c in their paper) and [12] (Figure 3b in their paper). At present, the combination of $\alpha_1 = 0.001$, $\alpha_2 = 4$ and $\alpha_3 = 300$ results in the lowest possible RMSE value from Tables 4a and 4b respectively. Hence, the combination set is taken to be the optimized set of parameters to train the selected NN model in Figure 4.

4.2.2. Variations of training periods

To justify the initial selection of the 30% to 35% of the total operational period for model-training in each referenced experiment, we further shorten the training period to 10% and 20% respectively and determine the associated RMSE values computed for the entire training period (10% and 20% of total operational period) and the subsequent testing period (remaining 90% and 80% respectively). Tables 5a and 5b show that the RMSE values from the training and testing steps, from the respective referenced experiments, gradually reduce with an increasing training period as expected. However, training the NN model with dataset from a large percentage of the entire operational period invalidates the overall objective of building an engineering model which can achieve predictive maintenance for ARRS. On the other hand, too short of a training period (10% and 20%) results in high inaccuracies in the subsequent testing stages as shown in Tables 5a and 5b. Hence, at this stage, the selected 30%

to 35% initial training period is most appropriate to maintain a balance between the computed RMSE from the later testing period and the need to predict the filter's breakthrough timing at a reasonably early stage of its operations.

4.2.3. Predictive capabilities of NN model

Subsequently, we adopt the optimized combination of hyperparameters ($\alpha_1 = 0.001$, $\alpha_2 = 4$ and $\alpha_3 = 300$) to train the model's predictive accuracy for the remaining 65% to 70% of the total operational periods in the respective referenced experiments. Figure 7a shows reasonable agreement ($\sim 17.0\%$ error deviation) with the bulk of the raw dataset [11] (Figure 4c in their paper) as the best-fit linear trend for $\ln(\tilde{k}(\varepsilon^{-1}, t))$ remained relatively constant over the entire experimental run. The sudden jump in the values for $\ln(\tilde{k}(\varepsilon^{-1}, t))$ in the time range of $10 \leq t \leq 15$ in Figure 7a could be caused by the unexpected rapid detachment of the initially trapped contaminants inside the experimental filter which resulted in a rise in its global porosity, and hence its average permeability. On the contrary, the average error deviations for Figures 7b to 7d (24.2%, 24.6% and 16.3% respectively) are generally higher than that of Figure 7a due to the stagnating tail-end which may be indicative of a dominant balance between the trapping and detachment rates of contaminants, hence leading to minimum transient changes to the filter's average permeability. At this stage, however, the present model does not account for the above-mentioned scenarios due to the adhered assumptions to build the theoretical model in Section 2 previously.

4.2.4. Limitations of NN model

The compelling limitation to the current engineering model is its need to have at least 30% to 35% of the training dataset to perform the supervised model-based training as illustrated in Figure 7. The relative extensive training period may not be viewed as sufficiently early for predictive maintenance to occur. To address this limitation, the following are viable improvements to be further explored: (a) further reduce the complexity of the multiple unknowns in $B(x_K, t)$ in (4.3). For example, quantifying the growth rate of the collector spheres with greater precision by using a closer representation of the actual clogging behavior, and its subsequent effects on the periodic cell's porosity and hence the global average permeability over time; (b)

increase data collection associated with the early periods of operations for ARRS and perform additional rounds of supervised training which may reduce the time length required for effective training; and (c) introducing a unique parameter to the model formulation which accounts for the possible detachment of the initially trapped contaminants over time.

4.3. Predictive maintenance of ARRS

4.3.1. Effects of variations in c_0 and d_p

To illustrate the concept of predictive maintenance, we adhere to the trained NN model built from the dataset of [12] (Figure 3d in their paper) as the model demonstrates the best accuracy as shown in Figure 6d relative to that of Figures 6b and 6c. We first examine the varying effects of c_0 and d_p on the timing (t_c) for the seepage filter to reach its breakthrough stage at the maximum media depth. As discussed previously, stagnation in the filter's average permeability coefficient may imply that the removal rate of turbidity particles inside the ARRS is balanced off by its detachment rate as there is likely no more transient changes to the filter's global porosity. For the simulation runs, all other operational conditions of V_s , $\frac{\partial p}{\partial x_3}$, v and ρ_p are maintained from that in Table 2. Using the trained NN model derived for Figure 7d earlier, Table 6 lists the respective values of t_c for each unique combination of c_0 and d_p values as shown. Generally, the simulation results indicate close weightages to the effects contributed by $\frac{c_0}{\rho_p}$ and $\frac{L_m}{d_p}$ whereby higher c_0 and lower d_p values result in the seepage filter reaching its stagnation zone at a faster rate, i.e. smaller t_c values, as expected from the explanations in Table 1.

4.3.2. Effects of variation in V_s

Engineered aquifers enable operators to vary the system's hydraulic loading rates to meet the required productivity rate of potable water downstream. Intuitively, higher V_s generally results in faster rate of physical clogging inside ARRS which can then

expedite the timing of which maintenance must be carried out. Likewise, we adopt the same trained NN model from Figure 7d to examine the varying effects of V_s on t_c for treating both low- and high-concentrated influents of $0.001kg/m^3$ and $0.100kg/m^3$ having the average d_p of $50\mu m$. All other operational conditions are again unchanged from Table 2. The range of V_s examined is between $0.200m/h$ to $0.300m/h$ with a step size of $0.01m/h$, which then enables us to maintain the creeping flow conditions ($Re_l \ll 1$) inside the same seepage filter of. Table 7 shows that the correlation between V_s and t_c is less obvious when all other conditions are constant. The results indicate that a common high c_0 results in close t_c values for all V_s variations, hence confirming that c_0 is relatively more influential than V_s in affecting t_c . On the other hand, for the treatment of low concentrated influents, the effects of V_s are more prominent whereby higher V_s generally results in lower t_c as listed in Table 7. Overall, the same procedures can be applied to adopt the trained NN model for examining the varying effects of v and ρ_p on t_c while keeping the other physical parameters constant.

In summary, after satisfactory model verification from the testing period, the trained NN model can be useful for environmental engineers to predict the timing for a concerned seepage filter (ARRS) to reach its breakthrough stage for any other combinations of $\frac{L_m}{d_0^*}$, $\frac{V_s t^*}{L_m}$, $\frac{c_0}{\rho_p}$, $\frac{L_m}{d_p}$ and $\frac{(L_m)^{1.5} g^{0.5}}{v}$ in (4.5). By doing so, ample time is then provided to engineers to better manage the logistics (manpower, allocation of machinery, back-up plans etc.) involved in predictive maintenance. We do, however, note that some additional analysis should be done to account for the risk of the current error deviation on the respective predicted t_c values.

4.4. Estimations of filter's complex parameters

As a recap, $\tilde{k}(x_K, t)$ can be expressed as:

$$\tilde{k}(x_K, t) = - \frac{\frac{1}{\ln(c)} \frac{|S(x_K, t)|}{|\Omega|_V} \left(\frac{\partial d}{\partial t} + D_{a,L} Pe_l^{-1} \right)}{\frac{\partial p}{\partial x_K'}} x_K' + C_1, \quad (5.1)$$

In (2.5), we assume a first-order deposition process (reaction rate) of the turbidity particles at the boundary of each collector sphere along the macroscale media depth. Hence, the growth rate of the collector spheres can be approximated to follow first-order trend too. At this stage, $|S(x_K, t)|$ in (5.2) depends directly on the user's choice for the number of collector spheres (n) in each periodic cell domain which then affects the total volume of spheres with respect to (2.1), while $\ln(c)$ is known temporal profile for the following reasons: (a) effective clogging stage during operations whereby the removal efficacy of turbidity particles can be estimated beforehand, or (b) based on previous or current data. At the same time, $\frac{\partial p}{\partial x_K'}$ is also known beforehand as ARRS is usually operated under constant-pressure condition as shown in Table 2. Overall, (5.2) shows that the complexity of the physical clogging problem in ARRS can further be reduced to an unknown $k_{fs}(x_K)$ parameter which is expected to vary along the media depth inside the depth filter.

To simplify (5.2) as our first approximations, $\frac{\partial d}{\partial t}$ is taken to be sufficiently small by assuming that effective clogging occurs over a long operational period whereby the collector spheres grow slowly over time. By focusing on the maximum media depth of $x_K = \varepsilon^{-1}$, it can easily be shown that $k_{fs}(\varepsilon^{-1})$ must become a time variable in order for $\tilde{k}(\varepsilon^{-1}, t)$ to reduce gradually over time during the effective physical clogging stage whereby $\ln(c)$ is a known value under constant-pressure flow conditions. At the same time, $\frac{|S(x_K, t)|}{|\Omega|_V}$ can be approximated as $(1 - \theta_{t^*})$ whereby θ_{t^*} is the reducing porosity of the periodic cell domain over time. By doing so, (5.2) is modified into (5.3). In the following, we are concerned with critical $(k_{fs,c})$ reaction rate responsible for the filter's effective clogging behavior at the maximum media depth in the same referenced experiments of [11] (Figure 4c in their paper) and [12] (Figures 3a to 3c in their paper).

$$\tilde{k}(\varepsilon^{-1}, t) = - \frac{\frac{1 - \theta_{t^*} k_{fs}(\varepsilon^{-1}, t)}{\ln(c)} \frac{U\varepsilon}{\frac{\partial p}{\partial x_K'}}}{\frac{\partial p}{\partial x_K'}} \varepsilon^{-1} + C_1 \quad (5.3)$$

where C_1 represents the initial dimensionless average permeability coefficient having the value of unity.

For sand and soil filters, it is typical to assume that the critical porosity (θ_c) is close to 0.1 [32-33] which directly leads to stagnation in the filter's average permeability coefficient. In this analysis, the critical permeability coefficient (dimensionless value) is taken to be 0.2 for the referenced experiments with respect to Figure 3a. Table 8 lists the respective $k_{fs,c}$ values approximated from (5.3) for the referenced experiments. Since diffusion is assumed to be the dominant transport mechanism leading to the physical clogging (deposition) of the turbidity particles at both the microscale and macroscale, the microscale Damköhler ($D_{a,l}$) and macroscale Damköhler ($D_{a,L}$) numbers from (2.14) are taken to be ordering importance of ε^2 and ε respectively. By doing so, we can approximate the unknown diffusion coefficient values, as also indicated in Table 8, in the original forms of (2.4 – 2.5, 2.7) by using the relationship from (2.14). At the same time, to justify that diffusion is the focus mechanism at the microscale, very small microscale Peclet numbers (Pe_l), of orders of magnitude close to 10^{-8} , are estimated from (2.13) by using the approximated diffusion values as indicated in Table 8. The predictions are intuitive by far as we would expect Pe_l to be much less than the value of unity for creeping flow conditions, i.e. $Re_l \ll 1$, inside the periodic cell domain.

5. Closing remarks

In light of climate change and the rising populations globally, aquifer recharge and recovery systems (ARRS), i.e. seepage filters, in natural or engineered aquifers are garnering increasing attention worldwide to expand the water storage capacity in local communities. Besides their abilities to store stormwater for future uses, they can provide other benefits such as pre-treatment of low energy requirements and minimizing land subsistence. However, the complex physical clogging behavior resulting from the infiltration step remains intractable to handle till today which thus stymies the predictive maintenance of ARRS. Consequently, this may affect the supplies of potable water in the events of unexpected maintenance. The concern generally derives from the current inability to forecast the timing for ARRS to reach its breakthrough stage based on the known physical parameters associated with seepage filtration a priori.

To address the overall problem statement, we combine the homogenization theory with the multiscale perturbation analysis as the feature engineering step to first compact the complexity of the physical clogging problem in ARRS. Systematically, the homogenized representation, which quantifies the macroscale clogging behavior in ARRS, enables us to extract the relevant features to predict the timing, via supervised model-based learning, for the concerned ARRS to reach its breakthrough stage. Prior to the training step, the pool of training data is limited. Hence, we propose a series of steps to impute more training data, via interpolation with regression analysis, for training a multi-layered neural network (NN) architecture which has a pre-processed input layer by performing dimensional analysis with the extracted input features from the earlier analytical step.

For each referenced experiment from the literature, the NN model is best trained with dataset associated with the first 30 to 35% of the total operational period. After training, the model's predictive capability is then verified for the remaining 65 to 70% of the same total period. By far, the overall agreement is relatively satisfactory with an average deviation of 20% for the referenced experiments. Following which, we propose how the homogenized representation is useful to estimate an arbitrary filter's critical reaction rate and diffusion coefficient values which are responsible for its physical clogging occurring over time.

Acknowledgements

The first author is grateful to NTU, Singapore for the 4-year Nanyang President Graduate Scholarship (NPGS) for his Ph.D. studies.

References

- [1] Vorosmarty, C.J., Green, P., Salisbury, J., Lammers, R.B., Global water resources: vulnerability from climate change and population growth, *Science*, v289, 2000. DOI: <https://doi.org/10.1126/science.289.5477.284>.
- [2] Bouwer, H., Artificial recharge of groundwater: hydrogeology and engineering, *Hydrogeology Journal*, v10, 2002. DOI: <https://doi.org/10.1007/s10040-001-0182-4>.

- 852 [3] Jeong, H. Y., Jun, S.-C., Cheon, J.-Y., Park, M., A review on clogging
853 mechanisms and managements in aquifer storage and recovery (ASR) applications,
854 *Geosciences Journal*, v22 (4), 2018. DOI: <https://doi.org/10.1007/s12303-017-0073-x>.
- 855 [4] Masciopinto, C., Vurro, M., Palmisano, V.N., Liso, I.S., A suitable tool for
856 sustainable groundwater management, *Water Resources Management*, v31, 2017.
857 DOI: <https://doi.org/10.1007/s11269-017-1736-0>.
- 858 [5] Bloetscher, F., Sham, C.H., Danko III, J.J., Ratick, S., Lessons learned from
859 aquifer storage and recovery (ASR) systems in the United States, *Journal of Water*
860 *Resource and Protection*, v6, 2014. DOI: <https://doi.org/10.4236/jwarp.2014.617146>.
- 861 [6] Yuan, Y., Liang, D., Zhu, H. (2017), Optimal control of groundwater pollution
862 combined with source abatement costs and taxes, *Journal of Computational Science*,
863 v20, 2017. DOI: <https://doi.org/10.1016/j.jocs.2017.03.014>.
- 864 [7] Rinck-Pfeiffer, S., Ragusa, S., Sztajn bok, P., Vandeveld, T., Interrelationships
865 between biological, chemical, and physical processes as an analog to clogging in
866 aquifer storage and recovery (ASR) wells, *Water Research*, v34, 2000. DOI:
867 [https://doi.org/10.1016/S0043-1354\(99\)00356-5](https://doi.org/10.1016/S0043-1354(99)00356-5).
- 868 [8] Martin, R., Clogging Issues Associated with Managed Aquifer Recharge
869 Methods, *IAH Commission on Managing Aquifer Recharge*, Australia, 2013.
- 870 [9] Driscoll, F.G., Groundwater and Wells, second ed., Johnson Screens, St. Paul,
871 1986.
- 872 [10] Hajra, M. G., Reddi, L. N., Glasgow, L. A., Xiao, M., Lee, I. M., Effects of Ionic
873 Strength on Fine Particle Clogging of Soil Filters, *Journal of Geotechnical and*
874 *Geoenvironmental Engineering*, v128 (8), 2002. DOI:
875 [https://doi.org/10.1061/\(ASCE\)1090-0241\(2002\)128:8\(631\)](https://doi.org/10.1061/(ASCE)1090-0241(2002)128:8(631)).
- 876 [11] Zhou, J., Zheng, X., Flury, M., Lin, G., Permeability changes during remediation
877 of an aquifer affected by sea-water intrusion: A laboratory column study, *Journal of*
878 *Hydrology*, v376 (3), 2009. DOI: <https://doi.org/10.1016/j.jhydrol.2009.07.067>.
- 879 [12] Wang, Z., Du, X., Yang, Y., Ye, X., Surface clogging process modeling of
880 suspended solids during urban stormwater aquifer recharge, *Journal of Environmental*
881 *Sciences*, v24(8), 2012. DOI: [https://doi.org/10.1016/S1001-0742\(11\)60961-3](https://doi.org/10.1016/S1001-0742(11)60961-3).
- 882 [13] Seifert, D., Engesgaard, P., Sand box experiments with bioclogging of porous
883 media: Hydraulic conductivity reduction, *Journal of Contaminant Hydrology*, v136-137,
884 2012. DOI: <https://doi.org/10.1016/j.jconhyd.2012.04.007>.
- 885 [14] Page, D., Vanderzalm, J., Miotliński, K., Barry, K., Dillon, P., Lawrie, K., Brodie,
886 R. S., Determining treatment requirements for turbid river water to avoid clogging of
887 aquifer storage and recovery wells in siliceous alluvium, *Water Research*, v66, 2014.
888 DOI: <https://doi.org/10.1016/j.watres.2014.08.018>.

- 889 [15] Grosso, J. M., Ocampo-Martínez, C., Puig, V., Learning-based tuning of
 890 supervisory model predictive control for drinking water networks, *Engineering*
 891 *Applications of Artificial Intelligence*, v26(7), 2013. DOI:
 892 <https://doi.org/10.1016/j.engappai.2013.03.003>.
- 893 [16] Hu, Z., Chan, C. W., In-situ bioremediation for petroleum contamination: A fuzzy
 894 rule-based model predictive control system, *Engineering Applications of Artificial*
 895 *Intelligence*, vol. 38, 2015. DOI: <https://doi.org/10.1016/j.engappai.2014.10.019>.
- 896 [17] Chen, T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D.. Neural Ordinary
 897 Differential Equations. *arXiv preprint arXiv:1806.07366*, 2018.
- 898 [18] Mei, C.C., Auriault, J. L., Ng, C.O., Some Applications of the Homogenization
 899 Theory, In J. W. Hutchinson and T. Y. Wu (Eds.), *Advances in Applied Mechanics*,
 900 v32, 1996. DOI: [https://doi.org/10.1016/S0065-2156\(08\)70078-4](https://doi.org/10.1016/S0065-2156(08)70078-4).
- 901 [19] Mei, C.C., and Vernescu, B., Seepage in Rigid Porous Media, *Homogenization*
 902 *Methods for Multiscale Mechanics*, 2012a. DOI:
 903 https://doi.org/10.1142/9789814282451_0003.
- 904 [20] Mei, C.C., Method of homogenization applied to dispersion in porous media,
 905 *Transport in Porous Media*, v9 (3), 1992. DOI: <https://doi.org/10.1007/BF00611970>.
- 906 [21] Mei, C.C., Auriault, J. L., Ng, C.O., Some Applications of the Homogenization
 907 Theory, In J. W. Hutchinson and T. Y. Wu (Eds.), *Advances in Applied Mechanics*,
 908 v32, 1996. DOI: [https://doi.org/10.1016/S0065-2156\(08\)70078-4](https://doi.org/10.1016/S0065-2156(08)70078-4).
- 909 [22] Mei, C.C., and Vernescu, B., Dispersion in Periodic Media or Flows,
 910 *Homogenization Methods for Multiscale Mechanics*, 2012b. DOI:
 911 https://doi.org/10.1142/9789814282451_0004.
- 912 [23] Ng, C.O., Mei, C.C., Homogenization theory applied to soil vapor extraction in
 913 aggregated soils, *Physics of Fluids*, vol. 8(9), 1996. DOI:
 914 <https://doi.org/10.1063/1.869017>.
- 915 [24] Lee, C. K., Sun, C.-C., Mei, C. C., Computation of permeability and
 916 dispersivities of solute or heat in periodic porous media, *International Journal of Heat*
 917 *and Mass Transfer*, v39 (4), 1996. DOI: [https://doi.org/10.1016/0017-9310\(95\)00174-](https://doi.org/10.1016/0017-9310(95)00174-3)
 918 [3](https://doi.org/10.1016/0017-9310(95)00174-3).
- 919 [25] Bouddour, A., Auriault, J. L., Mhamdi-Alaoui, M., Erosion and deposition of solid
 920 particles in porous media: Homogenization analysis of a formation damage, *Transport*
 921 *in Porous Media*, v25 (2), 1996. DOI: <https://doi.org/10.1007/BF00135852>.
- 922 [26] Royer, P., Auriault, J.-L., Lewandowska, J., Serres, C., Continuum Modelling of
 923 Contaminant Transport in Fractured Porous Media, *Transport in Porous Media*, vol.
 924 49(3), 2002. DOI: <https://doi.org/10.1023/A:1016272700063>.

925 [27] Xu, Z., Meakin, P., Upscaling of solute transport in heterogeneous media with
 926 non-uniform flow and dispersion fields, *Applied Mathematical Modelling*, v37 (18),
 927 2013. DOI: <https://doi.org/10.1016/j.apm.2013.03.070>.

928 [28] Chamsri, K., Bennethum, L. S., Permeability of fluid flow through a periodic
 929 array of cylinders, *Applied Mathematical Modelling*, v39 (1), 2015. DOI:
 930 <https://doi.org/10.1016/j.apm.2014.05.024>.

931 [29] Battiato, I., Tartakovsky, D. M., Applicability regimes for macroscopic models
 932 of reactive transport in porous media, *Journal of Contaminant Hydrology*, v120-121,
 933 2011. DOI: <https://doi.org/10.1016/j.jconhyd.2010.05.005>.

934 [30] Korneev, S., Battiato, I., Sequential Homogenization of Reactive Transport in
 935 Polydisperse Porous Media, *Multiscale Modeling & Simulation*, vol. 14(4), 2016. DOI:
 936 <https://doi.org/10.1137/16M1074278>.

937 [31] Ray, N., van Noorden, T., Frank, F., and Knabner, P., Multiscale Modeling of
 938 Colloid and Fluid Dynamics in Porous Media Including an Evolving Microstructure,
 939 *Transport in Porous Media*, v95(3), 2012. DOI: [https://doi.org/10.1007/s11242-012-](https://doi.org/10.1007/s11242-012-0068-z)
 940 [0068-z](https://doi.org/10.1007/s11242-012-0068-z).

941 [32] Dalwadi, M. P., Bruna, M., Griffiths, I. M., A multiscale method to calculate filter
 942 blockage, *Journal of Fluid Mechanics*, v809, 2016. DOI:
 943 <https://doi.org/10.1017/jfm.2016.656>.

944 [33] Dalwadi, M. P., Griffiths, I. M., Bruna, M., Understanding how porosity gradients
 945 can make a better filter using homogenization theory, *Proceedings of the Royal*
 946 *Society A: Mathematical, Physical and Engineering Science*, v471(2182), 2015. DOI:
 947 <https://doi.org/10.1098/rspa.2015.0464>.

948 [34] Printsypar, G., Bruna, M., Griffiths, I. M.. The influence of porous-medium
 949 microstructure on filtration. *Journal of Fluid Mechanics*, v861, 2019. DOI:
 950 <https://doi.org/10.1017/jfm.2018.875>.

951 [35] Rieu, M., G. Sposito., Fractal Fragmentation, Soil Porosity, and Soil Water
 952 Properties: I. Theory, *Soil Sci. Soc. Am. J.* v55, 1991. DOI:
 953 <https://doi.org/10.2136/sssaj1991.03615995005500050007x>.

954 [36] Bruna, M., Chapman, S., Diffusion in Spatially Varying Porous Media, *SIAM*
 955 *Journal on Applied Mathematics*, v75 (4), 2015. DOI:
 956 <https://doi.org/10.1137/141001834>.

957 [37] Y.Davit, C.G.Bell, H.M.Byrne, L.A.C.Chapman, L.S.Kimpton, G.E.Lang, ...
 958 Quintard, M. (2013). Homogenization via formal multiscale asymptotics and volume
 959 averaging: how do the two techniques compare? *Advances in Water Resources*, v62,
 960 2013. DOI: <https://doi.org/10.1016/j.advwatres.2013.09.006>.

- [38] Long, W., Hilpert, M., A Correlation for the Collector Efficiency of Brownian Particles in Clean-Bed Filtration in Sphere Packings by a Lattice-Boltzmann Method, *Environ Sci Technol*, v43 (12), 2009. DOI: <https://doi.org/10.1021/es8024275>.
- [39] Montessori A., Prestininzi P., La Rocca M., Falcucci G., Succi S., Kaxiras E. (2016), Effects of Knudsen diffusivity on the effective reactivity of nanoporous catalyst media, *Journal of Computational Science*, vol. 17 (part 2), 2016. DOI: <https://doi.org/10.1016/j.jocs.2016.04.006>.
- [40] McDowell-Boyer, L. M., Hunt, J. R., Sitar N., Particle Transport Through Porous Media, *Water Resources Research*, v22 (13), 1986. DOI: <https://doi.org/10.1029/WR022i013p01901>.
- [41] Yao, K. M., Habibian, M. T., O'Melia, C. R., Water and waste water filtration. Concepts and applications. *Environmental science & technology*, v5(11), 1971. DOI: <https://doi.org/10.1021/es60058a005>
- [42] Chew, A. W. Z., Law, A. W.-K., Multiscale Homogenization of Pre-treatment Rapid and Slow Filtration Processes with Experimental and Computational Validations, *LNCS*, v10861, 2018. DOI: https://doi.org/10.1007/978-3-319-93701-4_66.

983 List of Figures.

- FIGURE 1. Generalized illustration of the current problem statement described for ARRS operations
- FIGURE 2. Simplified representation of disordered packing geometry of media grains (represented by collector spheres) for ARRS system
- FIGURE 3. Raw dataset for $\tilde{k}(\varepsilon^{-1}, t)$ derived from the experimental works of [11] (Figure 4c in their paper) and [12] (Figures 3a to 3c in their paper): (a) before linearization; and (b) after linearization by taking natural logarithm of raw values
- FIGURE 4. Illustration of feature engineering method to build pre-processed input layer for multi-layered neural network (NN) architecture comprising of 3 hidden layers for supervised model-based training
- FIGURE 5. Regression analysis of raw dataset for $\ln(\tilde{k}(\varepsilon^{-1}, t))$ parameter derived from first 30% operational phase of the experimental run from [11] (Figure 4c in their paper)
- FIGURE 6. Regression analysis of raw dataset for $\ln(\tilde{k}(\varepsilon^{-1}, t))$ parameter derived from first 35% operational phase of the experimental runs from [12] (Figures 3a to 3c in their paper)
- FIGURE 7. Comparison between measured and predicted values from NN model for $\ln(\tilde{k}(\varepsilon^{-1}, t))$ parameter during training and testing stages for the experimental conditions of: (a) [11] (Figure 4c in their paper); and (b-d) [12] (Figures 3a to 3c in their paper).

985 List of Tables.

TABLE 1.	Description of the quantitative relationship between each extracted physical parameter and $m(x_K, t)$ derived in (3.12)
TABLE 2.	Summary of operational conditions relating to the experimental runs of [11] (Figure 4c in their paper) and [12] (Figures 3a to 3c in their paper) to emulate ARRS seepage filters
TABLE 3.	Values of hyper-parameters involved to train multi-layered neural network (NN) model for supervised model-based learning
TABLE 4.	Values of RMSE computed for the entire training step (training plus validations) for varying combinations of the hyper-parameters used: (a) using measured data from [11] (Figure 4c in their paper); and (b) using measured data from [12] (Figure 3b in their paper).
TABLE 5.	Effects of varying training periods on RMSE computed for the entire training and testing steps: (a) using measured data from [11] (Figure 4c in their paper); and (b) using measured data from [12] (Figure 3b in their paper).
TABLE 6.	Summary of predicted timings from NN predictive model for ARRS configuration of [12] (Figure 3c in their paper) to reach its stagnated low-flux condition (breakthrough stage) for varying c_0 and d_p values
TABLE 7.	Summary of predicted timings from NN predictive model for ARRS configuration of [12] (Figure 3c in their paper) to reach its stagnated low-flux condition (breakthrough stage) for varying c_0 and d_p values
TABLE 8.	Summary of estimated critical reaction rate and diffusion coefficients values responsible for the physical clogging stages in the referenced experiments

986

987

988 List of Figures

989

990 FIGURE 1.

FIGURE 2.

FIGURE 3.

FIGURE 4.

FIGURE 5.

1002 FIGURE 6.

1003

1004
1005

FIGURE 7.

1006 List of Tables

1007

TABLE 1.

Parameter	Association with $m(x_K, t)$	Explanations
d_0^*	Embedded within Pe_l^{-1} and $ S(x_K, t) $	Affects flow pattern inside porous media and growth rate of collector spheres whereby smaller grains result in lower microscale Reynolds number and improves the filter's trapping efficacy. Vice versa for bigger grains.
L_m	Embedded within $D_{a,L}$	Affects the volume of deposited turbidity particles inside porous media whereby larger media depth generally leads to higher volume of deposited turbidity particles for an unknown deposition rate. Vice versa for smaller media depths.
V_s	Embedded within Pe_l^{-1}	Affects flow pattern inside porous media whereby lower hydraulic loading rate result in lower microscale Reynolds number and improves the filter's trapping efficacy. Vice versa for higher hydraulic loading rate.
c_0	Embedded within $ S(x_K, t) $	Affects growth rate of collector spheres whereby higher influent concentration of turbidity particles results in higher growth rate of spheres for a given trapping efficacy. Vice versa for smaller influent concentrations.
ρ_p	Embedded within $D_{a,L}$	Affects the complex diffusion coefficient of the turbidity particles whereby higher bulk particle density generally leads to higher probability of contact of the particles with the media grains. Vice versa for lower bulk particle density.
d_p	Embedded within $D_{a,L}$	Affects the complex diffusion coefficient of the turbidity particles whereby larger turbidity particles generally leads to higher probability of contact of the particles with the media grains. Vice versa for smaller turbidity particles.
t	Embedded within $ S(x_K, t) $	Affects the growth rate of collector spheres
ν	Embedded within $D_{a,L}$	Affects the complex diffusion coefficient of the turbidity particles whereby lower fluid viscosity generally leads to higher probability of contact of the particles with the media grains. Vice versa for higher fluid viscosity.
g	Embedded within $D_{a,L}$	Dictates the gravitational force acting on the turbidity particles

1008

1009

TABLE 2.

Run no.	d_0^* (mm)	L_m (m)	V_S (m/h)	$d_p(\mu m)$	$\frac{\partial p}{\partial x_3'}$ (dimensionless)	c_0 (kg/m ³)	ν (m ² /s)	ρ_p (kg/m ³)	Re_l (dimensionless)	References
1	0.130	0.250	0.286	0.450	1.25×10^{-4}	0.100	9.33×10^{-7}	1030	$O(10^{-2}) \ll 1$	Zhou et al., 2009
2	0.100	0.500	0.271	44.3, 62.5, 38.5	2.00×10^{-4}	0.100	8.92×10^{-7}	2650	$O(10^{-3}) \ll 1$	Wang et al., 2012

1010

1011

TABLE 3.

Hyper-parameter	Value
Number of hidden layers	3
Size of hidden layer (nodes)	3, 3, 2 (sequentially)
Learning rate (α)	0.01, 0.001, 0.0001
Size of mini batches	4, 8, 16
Number of epochs	100, 300, 500
Activation function	Rectified Linear Unit

1012

1013

TABLE 4a.

Combination	α_1 (learning rate)	α_2 (sizes of mini-batches)	α_3 (number of epochs)	RMSE (training)
1	0.01	4	100	0.249
2	0.001	4	100	0.249
3	0.0001	4	100	0.0302
4	0.01	8	100	0.249
5	0.001	8	100	0.249
6	0.0001	8	100	0.0300
7	0.01	16	100	0.249
8	0.001	16	100	0.0302
9	0.0001	16	100	0.0303
10	0.01	4	300	0.249
*11	0.001	4	300	0.0302
12	0.0001	4	300	0.0302
13	0.01	8	300	0.249
14	0.001	8	300	0.0302
15	0.0001	8	300	0.0303
16	0.01	16	300	0.249
17	0.001	16	300	0.0302
18	0.0001	16	300	0.0302
19	0.01	4	500	0.249
20	0.001	4	500	0.0302
21	0.0001	4	500	0.0302
22	0.01	8	500	0.249
23	0.001	8	500	0.0302
24	0.0001	8	500	0.0302
25	0.01	16	500	0.249
26	0.001	16	500	0.0302
27	0.0001	16	500	0.0302

1014

* lowest possible RMSE

1015

1016

TABLE 4b.

Combination	α_1 (learning rate)	α_2 (sizes of mini-batches)	α_3 (number of epochs)	RMSE (training)
1	0.01	4	100	0.134
2	0.001	4	100	1.14
3	0.0001	4	100	1.14
4	0.01	8	100	1.14
5	0.001	8	100	0.136
6	0.0001	8	100	1.14
7	0.01	16	100	0.0793
8	0.001	16	100	0.0805
9	0.0001	16	100	1.14
10	0.01	4	300	1.14
*11	0.001	4	300	0.0744
12	0.0001	4	300	1.14
13	0.01	8	300	1.14
14	0.001	8	300	0.0857
15	0.0001	8	300	1.14
16	0.01	16	300	1.14
17	0.001	16	300	1.14
18	0.0001	16	300	0.0998
19	0.01	4	500	1.14
20	0.001	4	500	1.14
21	0.0001	4	500	1.14
22	0.01	8	500	1.14
23	0.001	8	500	1.14
24	0.0001	8	500	1.14
25	0.01	16	500	1.14
26	0.001	16	500	1.14
27	0.0001	16	500	1.14

* lowest possible RMSE

1017

1018

1019

1020

TABLE 5a.

Training period (%)	RMSE (training step)	Testing period (%)	RMSE (testing step)
10	0.0351	90	0.742
20	0.0316	80	0.219
35	0.0302	65	0.170

1021

1022

TABLE 5b.

Training period (%)	RMSE (training step)	Testing period (%)	RMSE (testing step)
10	0.102	90	0.604
20	0.0812	80	0.499
35	0.0744	65	0.246

1023

1024

TABLE 6.

Combination	$c_0(kg/m^3)$	$d_p (\mu m)$	$t_c (h)$
1	0.100	10	~ 9
2	0.100	20	~ 18
3	0.100	30	~ 27
4	0.100	40	~ 36
5	0.100	50	~ 44
6	0.100	60	~ 54
7	0.100	70	~ 62
8	0.100	80	~ 72
9	0.100	90	~ 81
10	0.100	100	~ 90
11	0.001	10	~ 900
12	0.001	20	~ 1800
13	0.001	30	~ 2700
14	0.001	40	~ 3500
15	0.001	50	~ 4500
16	0.001	60	~ 5400
17	0.001	70	~ 6300
18	0.001	80	~ 7200
19	0.001	90	~ 8000
20	0.001	100	~ 9000

1025

1026

1027

TABLE 7.

Combination	$c_0(kg/m^3)$	$d_p(\mu m)$	$V_s(m/h)$	$t_c(h)$
1	0.100	50	0.200	~44
2	0.100	50	0.210	~42
3	0.100	50	0.220	~40
4	0.100	50	0.230	~40
5	0.100	50	0.240	~38
6	0.100	50	0.250	~36
7	0.100	50	0.260	~34
8	0.100	50	0.270	~33
9	0.100	50	0.280	~32
10	0.100	50	0.290	~31
11	0.100	50	0.300	~30
12	0.001	50	0.200	~4400
13	0.001	50	0.210	~4300
14	0.001	50	0.220	~4200
15	0.001	50	0.230	~4100
16	0.001	50	0.240	~4000
17	0.001	50	0.250	~3900
18	0.001	50	0.260	~3900
19	0.001	50	0.270	~3800
20	0.001	50	0.280	~3700
21	0.001	50	0.290	~3700
22	0.001	50	0.300	~3600

1028

1029

1030

TABLE 8.

Referenced experiment	$k_{fs,c} \text{ (ms}^{-1}\text{)}$	$D_{fs,c} \text{ (m}^2\text{s}^{-1}\text{)}$	Pe_l
Zhou et al. (2009, Figure 4c)	5.54×10^{-5}	2.66×10^{-2}	3.88×10^{-7}
Wang et al. (2012, Figures 3a to 3c)	1.06×10^{-5}	2.64×10^{-2}	2.85×10^{-7}

1031