Physics-based modeling of contaminant leaching in road construction

M.M. Masmoudi, X.L. Li, V.B. Boddeti, N.L. Lajnef
Michigan State University, East Lansing, Michigan, USA

ABSTRACT: Coal combustion products (CCPs) offer a sustainable solution for road construction but pose environmental risks due to trace element leaching. Predicting solute concentrations accurately is critical to mitigating these risks as it guides contamination assessment and effective remediation. However, the heterogeneity of soils and materials in pavement layers, along with environmental factors, hinders reliable predictive models. To address this, this study utilizes a novel method to predict seepage velocity as a field, a critical parameter for modeling contaminant transport, integrated into the advection–dispersion–reaction governing equation. Unlike traditional assumptions of constant velocity parameters, our approach provides accurate prediction of solute concentrations and valuable insights into when pollutants may reach groundwater. Numerical simulations demonstrate superior performance in estimating seepage velocity and pollutant concentrations. The findings underscore the potential of this methodology to enhance environmental risk assessment in pavement engineering and support the safe, sustainable use of CCPs in infrastructure development.

1 INTRODUCTION AND RELATED WORK

In civil engineering, the use of CCPs in road construction raises significant environmental concerns due to the potential leaching of harmful trace elements, such as cadmium, into the groundwater (Eighmy, 2001). Estimating the concentration of these solutes is crucial for assessing and mitigating the environmental impact of such materials. Several studies have modeled the leaching behavior of trace elements from CCPs in road construction, relying on known physical and chemical parameters, such as soil permeability, porosity, and hydraulic conductivity, to predict solute movement (Kim, 2002; Mudd, 2004; Praharaj, et al., 2002). However, these parameters are often difficult to measure accurately due to the inherent variability in soil composition, changes in material properties over short spatial scales, and the influence of environmental factors such as moisture content, temperature, and pressure gradients. Additionally, measuring these properties often requires controlled laboratory tests on extracted samples, which are not only time-consuming and labor-intensive but also impractical for in-field applications without causing damage to the pavement structure.

One of the most critical parameters for predicting the transport of contaminants is the seepage velocity, which influences the rate at which pollutants migrate through soil layers. This parameter is challenging to measure directly, as it is dependent on soil properties such as porosity**,** tortuosity**,** and hydraulic conductivity, all of which are spatially variable and difficult to obtain in practice. Without accurate knowledge of these parameters, reliable predictions of solute concentrations become extremely difficult, undermining risk assessments and decision-making in environmental management. Moreover, road materials and the underlying soil layers are inherently heterogeneous, meaning that the seepage velocity must be modeled as a field parameter to reflect this spatial variability.

In recent years, machine learning tools, particularly Physics-Informed Neural Networks (PINNs)**,** have shown promise in estimating the parameters needed to solve partial differential equations (PDEs) governing solute transport. PINNs have been used successfully to predict material properties and flow characteristics in civil engineering problems, leveraging available data and solving the PDEs directly without needing traditional numerical methods (Raissi, et al., 2019; Bolandi, et al., 2023). These models learn the underlying physical laws governing the system and can predict unknown parameters with fewer measurements, making them highly efficient. However, PINNs still require substantial data for training and may not be as efficient in cases with limited data. In our previous work, we demonstrated several methods that have significant potential in estimating parameters for partial differential equations (PDEs) using limited or sparse data. These methods have been

successfully applied across various domains, including beam vibration, heat conduction, and electrophysiology (Li, et al., 2022; Li, et al., 2024; Masmoudi, et al., 2024). Other traditional and alternative parameter estimation methods require significantly large amount of data or parameter measurements that are not feasible in most real world applications.

Our approach offers a more data-efficient solution by predicting the seepage velocity with significantly less data, using only response data (i.e., the observed concentration of pollutants over time) without the need for direct measurements of soil properties. This is particularly advantageous in real-world scenarios where obtaining extensive datasets for all relevant parameters can be costly and time-consuming. Furthermore, our method is capable of extrapolating future outcomes**,** allowing for predictions about when pollutants will reach the groundwater table. This predictive capability makes our approach a valuable tool for long-term environmental risk assessment, offering actionable insights into the environmental impact of using CCPs in road construction with far less data than traditional models.

2 MODELING SETUP

The proposed method discretizes PDEs spatially using Finite Difference discretization. The seepage velocity V_z is modeled by a feed-forward neural network with spatial coordinates inputs x and y, which correspond to the coordinates of the nodes. The network in this study consists of 6 layers, each of the middle four layers containing 50 neurons and featuring skip-connections. The network parameters are denoted as θ, and the seepage velocity is modeled as $V_z(x, y) = N(x, y, \theta)$. The estimated seepage velocity from the neural network is inserted into differential equations for forward inference (state variable prediction). The inference is compared with available observations in mini-batches to compute the loss. We employed the L1 loss function $L(\theta) = \sum \left| C_{pred} - C_{Ground\,Truth} \right|$ as it provided superior results compared to both L2 and normalized L2 loss. To optimize the parameters and train the network, the adjoint sensitivity method, combined with reverse-mode automatic differentiation techniques is utilized for solving differential equations and efficiently computing the gradient of the loss function through the PDE solver (Rackauckas, et al., 2019; Chen, et al., 2018).

3 PROBLEM SETUP

In previous studies (Li, et al., 2006; Zheng & Bennett, Gordon D, 2002), seepage velocity has been modeled as a scalar parameter, assuming that the CCP layer and subgrade layer exhibit homogeneity in key properties such as porosity, hydraulic conductivity, and hydraulic gradient. This simplification directly neglects the spatial variability that is often present in real-world systems. In this work, we address this limitation by modeling seepage velocity as a spatially dependent field parameter, capturing the inherent heterogeneity in these layers. By defining seepage velocity as a nonlinear function, we account for variations in material properties and hydraulic conditions, enabling a more realistic representation of the flow dynamics. This approach provides a significant improvement over traditional models, offering a framework better suited for analyzing complex subsurface systems.

The pavement configuration depicted in Fig. 1 is analyzed in this study, which focuses on modeling the concentration of cadmium (Cd) within the CCP and subgrade layers. The objective is to predict the timing and concentration at which this pollutant may reach the groundwater table. To achieve this, the advection–dispersion–reaction equation (ADRE) is solved within these layers (Bear, 2012).

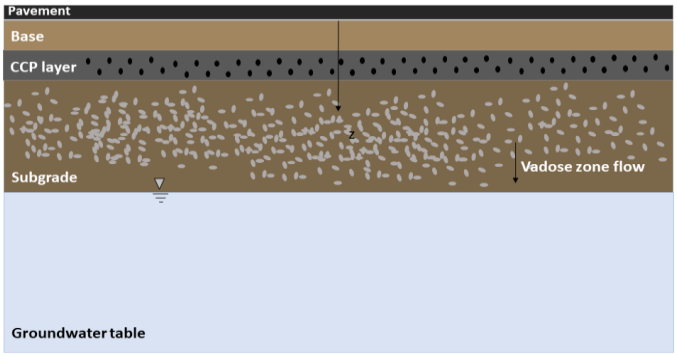

Figure 1. Layers of pavement: CCP layer and subgrade are considered heterogeneous materials, resulting in spatially dependent seepage velocity.

3.1 *Governing equations*

Leaching occurs as water moves downward through the CCP layer, following different patterns such as first-flush, lagged response, or empirically defined behaviors. First-flush leaching is modeled using the ADRE with linear, instantaneous, and reversible sorption, while lagged or other patterns are described empirically based on concentration and flow data. In the vadose zone below the CCP layer, transport is modeled using the ADRE for 1D steady-state vertical flow with 2D dispersion and linear, instantaneous sorption.

$$R\frac{\partial C}{\partial t} = D_x\frac{\partial^2 C}{\partial x^2} + D_z\frac{\partial^2 C}{\partial z^2} - v_z\frac{\partial C}{\partial z} \quad (1)$$

In this equation, C represents the solute concentration, t is time, x is the horizontal distance from the pavement centerline, z is the depth below ground surface, v_z is the vertical seepage velocity, D_x and D_z are the dispersion coefficients in the x and z directions, and R is the retardation factor. The vertical seepage velocity (v_z) is to be predicted. The hydrodynamic dispersion coefficients are calculated using

$D = \alpha\, v_z + \tau D_0$, where $\boldsymbol{\alpha}$ is the dispersivity (vertical or horizontal), and D_0 is the molecular diffusion coefficient (Leij, et al., 1991). It is assumed that chemical and biological reactions, which might alter or consume trace elements, are absent.

3.2 *Boundary and initial conditions*

To solve the ADRE, appropriate initial and boundary conditions are specified.

$$C(x, z, t = 0) = C_0 \tag{2}$$

$$\left(v_z C - D_z \frac{\partial C}{\partial z}\right)\Big|_{z=z_{top}} = v_z\, f(x, z_{top}, t) \tag{3}$$

$$\left(\frac{\partial C}{\partial x}\right)\Big|_{x=x_0, x_L} = 0 \tag{4}$$

Where, $x_0 = 0$, $x_L = 2m$, $z_{top,CCP} = 0m$, and $z_{top,Subgrade} = 0.3m$. $f(t)$ is the concentration from the previous layer and C_0 is the initial concentration. Equation (3) defines a flux boundary at the top, driven by seepage from the upper layer, while Equation (4) sets no-flux (Neumann) conditions at the sides, assuming horizontal symmetry or isolation.

3.3 *Problems parameters*

Table 1. Simulation Parameters.

Parameters	CCP Layer	Subgrade
Initial concentration $C_0(\mu g/l)$	4	0
Tortuosity τ	0.7	0.7
Longitudinal dispersivity α_L	0.1	0.1
Transverse dispersivity α_T	0.01	0.01
Retardation factor R	3.5	3.5
Molec`ular diffusion coefficienD_0 $(10^{-10} m^2/s)$	6.0	6.0

We conducted contaminant leaching analysis on the CCP and subgrade layers using a computational mesh with 400 nodes per layer. The system dimensions included a $2m$ width, a $0.3m$ deep CCP layer, and a $4.5m$ deep subgrade layer. The subgrade material was modeled as silt loam. The seepage velocity in the CCP layer was modeled with a nonlinear distribution ranging from 0.19 to 0.26, while for the subgrade layer, it ranged from 0.91 to 1.12.

4 NUMERICAL RESULTS

We demonstrate the efficacy of our method in predicting a nonlinear field seepage velocity distribution, achieving a mean absolute percentage error (MAPE) as low as 1.6%. We also analyze the system's response at $t = 40$ years, where the response mean absolute error (MAE) remains below $3 \times 10^{-4}\ \mu g/l$ when utilizing the predicted field seepage velocity.

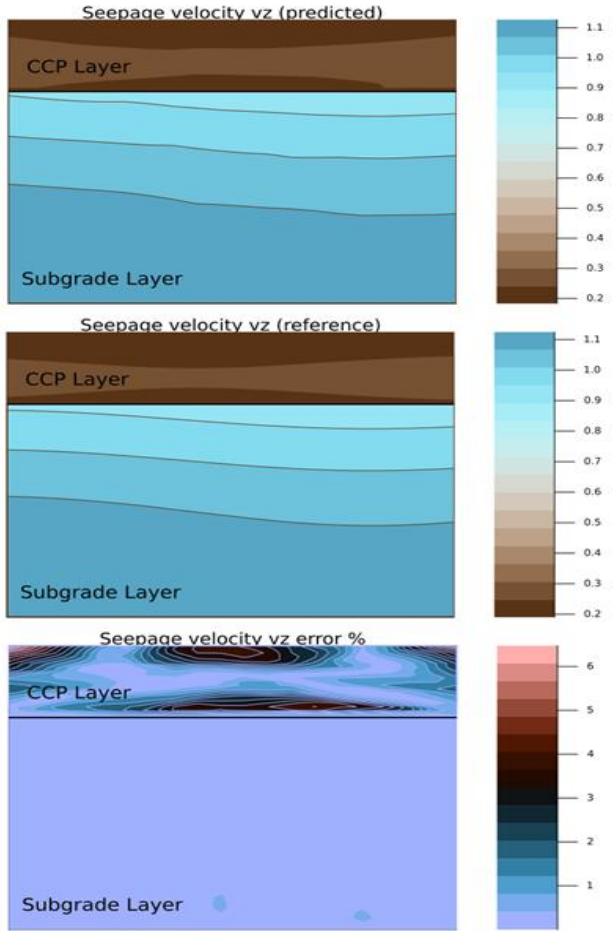

Figure 2. Accurate estimation of the field seepage velocity V_z with a MAE of 1.6%.

To provide a comprehensive comparison and emphasize the importance of modeling the parameter as a field rather than a scalar, we also present the system response based on scalar parameter prediction. In this case, a single neuron predicts the scalar parameter, which typically converges to the mean of the ground truth field parameter distribution. However, this approach results in substantial errors, with a forward inference MAE of $2.6 \times 10^{-2} \mu g/l$ compared to just $3 \times 10^{-4} \mu g/l$ for the field-based approach by our method. Furthermore, we evaluate our method's extrapolation performance at $t = 50$ years, where it achieves an MAE of $8 \times 10^{-3} \mu g/l$ relative to the ground truth. In contrast, the extrapolated response using the scalar parameter model shows a significantly higher error of $8 \times 10^{-5} \mu g/l$. We successfully trained our model using only 600 measurements per layer, distributed across both temporal and spatial dimensions.

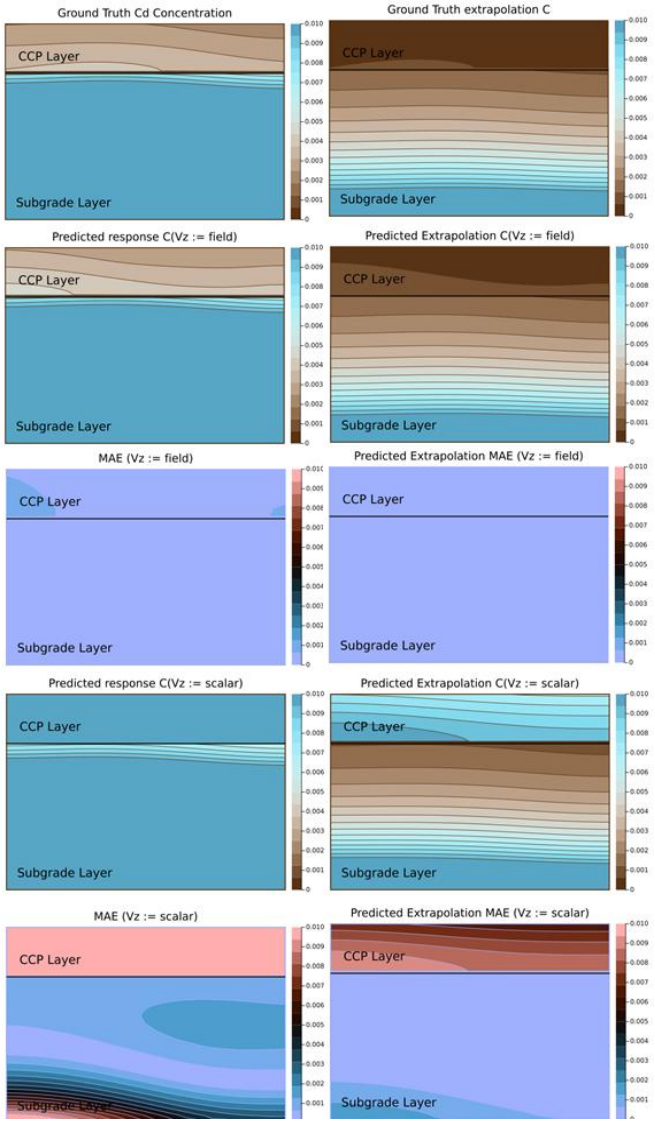

Figure 3. Forward inference using the predicted parameter shows our method's robust performance in both interpolation and extrapolation tasks. Field parameter modeling achieves a forward inference MAE of $3 \times 10^{-4} \mu g/l$ (first column), significantly outperforming scalar parameter modeling (MAE of $2.6 \times 10^{-2} \mu g/l$. Extrapolation at $t = 50$ years (second column) results in an MAE of $8 \times 10^{-3} \mu g/l$, compared to $8 \times 10^{-5} \mu g/l$ for the scalar model.

5 CONCLUSION

This paper introduces a novel approach for modeling seepage velocity as a spatially dependent field parameter, significantly improving prediction accuracy. The method reduces error to $3 \times 10^{-4} \mu g/l$, a substantial enhancement compared to the previous accuracy of $2.6 \times 10^{-2} \mu g/l$ in predicting contaminant transport through heterogeneous pavement layers. It outperforms scalar-based models and shows strong extrapolation capabilities, enabling reliable predictions of when and at what concentration pollutants reach groundwater. This makes it a valuable tool for timely interventions and mitigating groundwater contamination risks in CCP-based road construction.

REFERENCES

Bear, J., 2012. *Hydraulics of groundwater*. s.l.:Courier Corporation.

Bolandi, H. et al., 2023. Physics informed neural network for dynamic stress prediction. *Applied Intelligence*, Volume 53, pp. 26313-26328.

Chen, R. T., Rubanova, Yulia, Bettencourt, Jesse & Duvenaud, David K, 2018. Neural ordinary differential equations. *Advances in neural information processing systems*, Volume 31.

Eighmy, T. T. a. C. W. H. a. o., 2001. Framework for Evaluating Use of Recycled Materials in the Highway Environment. *United States. Department of Transportation. Federal Highway Administration.*

Kim, A. G., 2002. *CCB leaching summary: Survey of methods and results.* s.l., Proceedings of Technical Interactive Forum, Coal Combustion By-Products and Western Coal Mines.

Leij, F. J., Skaggs, Todd H & Van Genuchten, Martinus Th, 1991. Analytical solutions for solute transport in three-dimensional semi-infinite porous media. *Water resources research*, Volume 27, pp. 2719-2733.

Li, L., Benson, Craig H, Edil, Tuncer B & Hatipoglu, Bulent, 2006. *Groundwater impacts from coal ash in highways.* s.l., Proceedings of the institution of civil engineers-waste and resource management.

Li, X. et al., 2022. *NeuralSI: Structural Parameter Identification in Nonlinear Dynamical Systems.* s.l., European Conference on Computer Vision.

Li, X., Masmoudi, Mahdi, Lajnef, Nizar & Boddeti, Vishnu, 2024. *Estimating field parameters from multiphysics governing equations with scarce data*. s.l., ICLR 2024 Workshop on AI4DifferentialEquations In Science.

Masmoudi, M., Li, Xuyang, Lajnef, Nizar & Boddeti, Vishnu, 2024. *ParaFIND: Parameter Field Inference on Non-uniform Domains using Neural Network.* s.l., NeurIPS 2024 Workshop on Data-driven and Differentiable Simulations, Surrogates, and Solvers.

Mudd, G. M., 2004. Environmental geochemistry of leachate from leached brown coal ash. *Journal of environmental engineering*, Volume 130, pp. 1514-1526.

Praharaj, T., Powell, MA, Hart, BR & Tripathy, S, 2002. Leachability of elements from sub-bituminous coal fly ash from India. *Environment international*, Volume 27, pp. 609-615.

Rackauckas, C. et al., 2019. Diffeqflux. jl-A julia library for neural differential equations. *arXiv preprint arXiv:1902.02376.*

Raissi, M., Perdikaris, Paris & Karniadakis, George E, 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, Volume 378, pp. 686-707.

Zheng, C. & Bennett, Gordon D, 2002. Applied contaminant transport modeling. *Wiley-Interscience New York*, Volume 2.