Physics-Informed Machine Learning for Poroelastic Consolidation Using an Analytical Digital Twin

Agnila Ghosh Surovi

December 10, 2025

Abstract

This work introduces a physics-informed machine learning (PIML) framework for studying one-dimensional poroelastic consolidation using a digital twin generated entirely from analytical solutions. The analytical model enables the construction of a clean and fully controlled synthetic dataset in which poroelastic layer thickness, porosity, elastic moduli, and hydraulic conductivity can be varied independently. Within this dataset, we explore consolidation behavior through dimensionality reduction, unsupervised clustering, regression models, and causal analysis in a single, coherent workflow. The analysis shows that key consolidation metrics, including t_{50} and t_{90}, can be predicted with high accuracy from material parameters, while inverse regression recovers permeability and compressibility related properties directly from early-time pore-pressure signatures. A directed acyclic graph (DAG) is then developed to represent physically meaningful causal relations among poromechanical variables. The overall framework provides a transparent and interpretable foundation for data-driven characterization and design of poroelastic and geomechanical materials.

Introduction

Poroelastic consolidation describes the coupled interaction between fluid flow and deformation in saturated porous media, and it plays a central role in fields ranging from geomechanics and hydrogeology to soft biological tissues such as cartilage, hydrogels, and the intervertebral disc [1, 2, 3, 4, 5]. The classical one-dimensional theories of Terzaghi and Biot have long provided the basis for estimating drainage behavior, effective stresses, and the time scale of settlement or creep [6, 1, 7, 8, 9]. Although the governing equations are mathematically well established, determining the permeability, compressibility, and other material parameters needed for predictive modeling often requires indirect measurements or experiments that are sensitive, expensive, and limited in scope [2, 10, 11, 12].

In parallel, machine learning has become increasingly common in applications involving porous materials, including pore-pressure forecasting, aquifer characterization, and settlement prediction [13, 14, 15, 11]. These approaches are capable of capturing nonlinear patterns in data, but many existing studies remain largely empirical and rely on numerically generated or experimentally measured datasets that may span only a narrow portion of the

relevant parameter space [16, 17, 18, 19]. Furthermore, most ML models do not incorporate the structure of the underlying poromechanical equations, making it difficult to interpret their predictions in a physical sense [13, 16].

Analytical solutions of the one-dimensional consolidation equations offer an opportunity to construct a digital twin in which the exact physics are preserved and the parameter space can be explored systematically [8, 7, 9]. Compared to finite-element-based digital twins which are often used in biomechanics and geomechanics, an analytical model avoids numerical artifacts and makes it straightforward to examine how changes in material properties influence pore-pressure evolution [4, 20, 19]. While some recent studies have used analytical or semi-analytical solutions within neural-network frameworks, these efforts have focused mainly on forward prediction or single-parameter identification, and have not connected analytical digital twins to broader questions such as regime classification, dimensionality reduction, or causal structure [21, 22, 23, 17].

Another limitation of the current literature is that pore-pressure time signatures themselves are seldom treated as structured features for inverse identification of intrinsic material properties. Many geotechnical and reservoir-related studies take porosity, density, or seismic attributes as inputs and attempt to estimate pore pressure or consolidation coefficients [14, 11, 12, 15]. Although this provides useful engineering predictions, it does not fully utilize the dynamic patterns present in the consolidation response. Operator-learning and physics-informed ML techniques have demonstrated the ability to match analytical solutions and estimate the coefficient of consolidation, but these methods typically operate within constrained parameter settings and do not form part of a larger workflow that integrates unsupervised learning, regression, and causal interpretation [21, 22, 23, 17, 24].

The aim of this study is to address these gaps by developing a physics-informed machine-learning approach built upon an analytical digital twin for one-dimensional consolidation [13, 25]. The digital twin provides a reliable and noise-free source of data; the resulting dataset allows for regime discovery, inverse prediction of material parameters, and the construction of a physics-informed causal graph. Together, these components offer a unified and interpretable framework for analyzing consolidation behavior and for linking pore-pressure evolution to underlying poromechanical properties.

Contributions

This study offers several methodological contributions to data-driven poromechanics. First, we construct a poroelastic digital twin using analytical consolidation solutions, allowing controlled exploration of material parameters without numerical artifacts. Second, we generate a large synthetic dataset that spans realistic ranges of geomechanical properties and captures a broad variety of consolidation responses. Third, we apply dimensionality reduction and K-means clustering to uncover low-dimensional structure and identify distinct consolidation regimes. Fourth, we train regression models that accurately predict consolidation metrics such as t_{50} and t_{90} and recover underlying material properties from pore-pressure time histories. Finally, we develop a physics-informed directed acyclic graph to formalize the causal relationships among material parameters, diffusion characteristics, and consolidation behavior. Taken together, these elements form a coherent and interpretable machine-learning

framework grounded in classical poromechanics.

Methods

Digital Twin: Analytical Terzaghi's one-dimensional consolidation

We consider a saturated poroelastic layer of thickness h subjected to a vertical load at its top surface and resting on an impermeable base. Vertical deformation is assumed to be one-dimensional, so that the relevant field variables are the excess pore pressure $p(z,t)$ and the vertical total stress $\sigma_{zz}(t)$, with z measured from the impermeable base ($z=0$) to the drained top surface ($z=h$). In the framework of linear poroelasticity, conservation of mass of fluid and solid together with Darcy's law leads, for one-dimensional flow, to the storage equation [26]

$$\alpha \frac{\partial \varepsilon}{\partial t} + S \frac{\partial p}{\partial t} = \frac{\partial}{\partial z}\left(\frac{k}{\gamma_f} \frac{\partial p}{\partial z} \right), \tag{1}$$

where ε is the vertical (volumetric) strain, α is Biot's coefficient, S is the storativity, $\bar{k}$ is the hydraulic conductivity (or mobility written as k/γ_f or κ/μ), and γ_f is the unit weight of the pore fluid and μ is the viscosity of the fluid.

The vertical strain rate is related to the vertical effective stress by the confined compressibility m_v,

$$\frac{\partial \varepsilon}{\partial t} = -\, m_v \frac{\partial \sigma'_{zz}}{\partial t} = -\, m_v \left(\frac{\partial \sigma_{zz}}{\partial t} - \alpha \frac{\partial p}{\partial t} \right), \tag{2}$$

with

$$m_v = \frac{1}{K + \frac{4}{3}G}, \tag{3}$$

where K and G are the drained bulk and shear moduli of the porous skeleton.

Biot's coefficient and the storativity are expressed in terms of the compressibilities of the solid particles (C_s), the pore fluid (C_f), and the porous medium as a whole ($C_m = 1/K$) as [26]

$$\alpha = 1 - \frac{C_s}{C_m}, \qquad S = nC_f + (\alpha - n)C_s, \tag{4}$$

where n is the porosity.

Eliminating $\partial\varepsilon/\partial t$ between (1) and (2) gives the general one-dimensional consolidation equation

$$\left(S + \alpha^2 m_v\right) \frac{\partial p}{\partial t} = \alpha m_v \frac{\partial \sigma_{zz}}{\partial t} + \frac{\partial}{\partial z}\left(\frac{k}{\gamma_f} \frac{\partial p}{\partial z} \right). \tag{5}$$

This is Verruijt's general differential equation for one-dimensional consolidation.

Terzaghi-type step loading

In the classical Terzaghi problem a vertical load of magnitude q is suddenly applied at $t=0$ and then kept constant, so that for $t>0$ the total stress σ_{zz} no longer varies with time. In

this case the first term on the right-hand side of (5) vanishes and the equation reduces to a diffusion equation for the excess pore pressure

$$\frac{\partial p}{\partial t} = c_v \frac{\partial^2 p}{\partial z^2}, \qquad c_v = \frac{\bar{k}}{S + \alpha^2 m_v}, \tag{6}$$

where c_v is the consolidation coefficient.

Immediately after application of the load, no drainage has yet occurred, so the initial excess pore pressure is spatially uniform and given by

$$p(z, 0) = p_0 = B\, q, \qquad B = \frac{\alpha m_v}{S + \alpha^2 m_v}, \tag{7}$$

where B is Skempton's coefficient for undrained loading. The boundary conditions for a layer with an impermeable base and a fully drained top surface read

$$\frac{\partial p}{\partial z}(z = 0, t) = 0, \qquad t > 0, \tag{8}$$

$$p(z = h, t) = 0, \qquad t > 0. \tag{9}$$

Solving (6) with initial condition (7) and boundary conditions (8)–(9) by separation of variables yields the familiar Fourier series solution for the transient pore pressure distribution

$$p(z, t) = p_0 \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{2k-1} \cos\left(\frac{(2k-1)\pi z}{2h}\right) \exp\left[-\frac{(2k-1)^2 \pi^2 c_v t}{4h^2}\right]. \tag{10}$$

This expression is identical to Verruijt's series representation of Terzaghi's one-dimensional consolidation solution for a layer with a drained top boundary and an impermeable base.

Extension to a time-dependent surface load

Because the governing equation (5) is linear, a general time-dependent surface load $\sigma_{zz}(t) = q(t)$ can be accommodated by superposition. At each instant the undrained response at the top surface is characterized by a time-dependent pore pressure scale

$$p_0(t) = \frac{\alpha m_v}{S + \alpha^2 m_v}\, q(t) = B\, q(t), \tag{11}$$

which reduces to (7) for a step load. In the present work we prescribe a sinusoidal ramp followed by a constant load,

$$q(t) = \begin{cases} \sin(5\pi t), & 0 \le t < t_0, \\ \sin(5\pi t_0), & t \ge t_0, \end{cases} \qquad t_0 = 0.1 \text{ s}, \tag{12}$$

so that

$$p_0(t) = \frac{\alpha m_v}{S + \alpha^2 m_v}\, q(t). \tag{13}$$

In our numerical implementation we evaluate $p_0(t)$ from (13) and insert it into the series solution (10) to generate a reference pore-pressure history at a given depth, which is then compared against the finite element results.

Parameter Ranges and Sensitivity Study Design

To define a representative parameter space for the digital twin, we first carried out an LLM-assisted literature mining exercise to collect reported values of porosity, permeability, skeleton modulus, grain modulus, and fluid modulus for saturated soils, weak rocks, and shallow aquifer systems. The resulting database spanned several orders of magnitude in stiffness and hydraulic properties. From this broad set, we then restricted attention to a subrange that (i) is consistent with typical geotechnical applications at depths of tens of metres, and (ii) produces non-trivial pore-pressure evolution over the simulation horizon $t \in [0, 30]$ s. Parameter combinations that led to unphysical Biot coefficients ($\alpha \notin [0, 1]$) or essentially drained or undrained behaviour within this time window were discarded and resampled [5, 8, 12, 2, 15, 9].

Layer thickness (h): The characteristic thickness of the consolidating layer was taken as

$$h \in [10,\ 40] \text{ m},$$

representing shallow soil or weak-rock columns of tens of metres. This range is typical of one-dimensional consolidation problems in embankments, shallow reservoirs, and aquifer compaction studies.

Porosity (n): Porosity was treated as a dimensionless volume fraction and sampled uniformly in

$$n \in [0.10,\ 0.90],$$

covering dense granular soils at the lower end and highly voided, weakly cemented materials at the upper end. These values are consistent with reported porosities for sands, silts, clays, and weak sedimentary rocks.

Poroelastic modulus (E): The elastic modulus of the porous material was sampled from

$$E \in [2 \times 10^4,\ 5 \times 10^4] \text{ kPa},$$

corresponding to stiffnesses in the range 20–50 MPa. This interval captures the transition from relatively soft, compressible soils to stiffer, cemented geomaterials, and is compatible with published ranges for young sedimentary formations and overconsolidated clays.

Solid grain bulk modulus (K_s): The intrinsic bulk modulus of the solid skeleton was taken as

$$K_s \in [1.5 \times 10^7,\ 9 \times 10^7] \text{ kPa},$$

which corresponds to approximately 15–90 GPa. These values fall within the range reported for common mineral constituents such as quartz- and feldspar-rich sandstones and carbonates.

Fluid bulk modulus (K_f): The bulk modulus of the pore fluid was sampled within

$$K_f \in [0.5 \times 10^6,\ 5 \times 10^6] \text{ kPa},$$

i.e. 0.5–5 GPa. This bracket includes water-like fluids at ambient conditions and allows for moderate variability associated with temperature, salinity, and mixtures with more compressible phases.

Hydraulic mobility ($\bar{k}$) and intrinsic permeability (κ): In the analytical consolidation model, hydraulic mobility is defined as

$$\bar{k} = \frac{\kappa}{\mu},$$

where κ is the intrinsic permeability and μ is the dynamic viscosity. Using $\mu = 10^{-6}$ kPa $\cdot$ s for a water-like fluid, we sampled the mobility over the logarithmic interval

$$\bar{k} \in [10^{-4},\, 10^{-1}]\ \mathrm{m}^4/(\mathrm{kN}\cdot\mathrm{s}).$$

This corresponds to an intrinsic permeability range of

$$\kappa = \bar{k}\,\mu \in [10^{-10},\, 10^{-7}]\ \mathrm{m}^2,$$

which spans materials from relatively tight, low-permeability soils to more permeable sands and weakly fractured formations. The mobility was drawn log-uniformly to ensure that each permeability decade is sampled with comparable frequency.

In the sensitivity study, parameters were sampled independently within these ranges to generate ensembles of pore-pressure histories and consolidation times. The chosen limits provide a balance between physical realism as informed by the literature mining and numerical practicality, ensuring that the resulting pore-pressure trajectories exhibit appreciable evolution over the simulation time window and are informative for subsequent dimensionality reduction, clustering, and regression.

To build a large training dataset, we randomly sampled 5000 sets of material, hydraulic, and geometric parameters across physically realistic ranges (Table 1). Each simulation produced a pore-pressure curve $p(t)$ sampled at 300 logarithmically spaced time points.

Dimensionality Reduction

To identify dominant modes of variability across the pressure-time responses, Principal Component Analysis (PCA) was applied. PCA represents each curve as a linear combination of orthogonal basis vectors (principal components) obtained from the eigen-decomposition of the covariance matrix. Formally, the standardized data matrix $X \in \mathbb{R}^{N\times T}$ is decomposed as

$$X = U\Sigma V^\top, \tag{14}$$

where the rows of V contain the principal directions and the columns of $U\Sigma$ represent the low-dimensional coordinates. PCA provides a non-parametric way to obtain an optimal (variance-maximizing) reduced representation of all simulated consolidation curves.

Clustering of Consolidation Behaviors

To classify families of consolidation responses, k-means clustering was performed in the PCA space. k-means partitions data into k disjoint groups by minimizing the total within-cluster variance,

Table 1: Parameter ranges used in the sensitivity study. All units use m, kPa, and kN.

Parameter	**Unit**	**Range Used**	**Physical Context**
Porosity (n)	–	0.10–0.90	Typical values for sands, silts, clays, and weak sedimentary rocks; excludes extreme hydrogel porosities.
Layer thickness (h)	m	10–40	Shallow soil or weak-rock columns representative of embankments and laboratory-scale consolidation layers.
Poroelastic modulus(E)	kPa	2×10^4–5×10^4	Elastic stiffness typical of soft to moderately stiff geomaterials (20–50 MPa).
Solid grain modulus (K_s)	kPa	1.5×10^7–9×10^7	Intrinsic mineral stiffness covering quartz, feldspar, and carbonate-rich grains (15–90 GPa).
Fluid bulk modulus (K_f)	kPa	0.5×10^6–5×10^6	Pore-fluid bulk modulus range from moderately compressible mixtures to water-like fluids (0.5–5 GPa).
Intrinsic permeability (κ)	m^2	10^{-10}–10^{-7}	Intrinsic permeability implied by the mobility range; corresponds to low-permeability soils to permeable sands.
Hydraulic mobility ($\bar{k}$)	$m^4/(kN{\cdot}s)$	10^{-4}–10^{-1}	Computed from $\kappa = \bar{k}\mu$ with viscosity $\mu = 10^{-6}$ kPa·s; sampled logarithmically.

$$\min_{\{\mathcal{C}_i\}} \sum_{i=1}^{k} \sum_{x\in\mathcal{C}_i} \|x - \mu_i\|^2, \tag{15}$$

where μ_i is the centroid of cluster $\mathcal{C}_i$. This approach groups responses with similar transient shapes, enabling unsupervised identification of behavior types (e.g., fast- vs. slow-draining materials). Cluster numbers were chosen using standard metrics such as inertia and silhouette scores.

Forward Regression for Consolidation Times

To develop data-driven surrogate models for consolidation metrics, we trained Random Forest (RF) regressors to predict the times to reach 50% and 90% dissipation (t_{50} and t_{90}). Random Forests are ensemble models consisting of multiple decision trees,

$$\hat{y} = \frac{1}{M} \sum_{m=1}^{M} f_m(x), \tag{16}$$

where each tree f_m is trained on a bootstrapped subset of data with randomized feature splits. RFs handle nonlinear interactions and heterogeneous feature importance naturally,

making them well suited for poromechanical parameter mappings. Model performance was evaluated using:

- R^2 (coefficient of determination),
- Mean Absolute Error (MAE).

These models form fast surrogates for physics-based consolidation predictions.

Inverse Modeling from Early-Time Pressure

To determine which physical parameters are inferable from partial observations, inverse Random Forest models were trained. The inputs consisted only of the early-time pressure samples (first 5–10 time points). For a given parameter θ, the inverse model takes the form:

$$\hat{\theta} = \mathcal{F}(p(t_1), p(t_2), \ldots, p(t_k)) \,, \tag{17}$$

where k is the number of early-time samples. This framework tests parameter identifiability: a parameter is considered identifiable if the inverse map exhibits low prediction error and a strong monotonic trend between true and predicted values.

Causal Analysis

To study structural relationships between parameters, storativity, diffusivity, and consolidation times, we combined correlation analysis with a physics-informed causal graph (directed acyclic graph, DAG).

Correlation. Pairwise Pearson correlations were computed for:

- physical parameters,
- derived quantities S and c,
- targets t_{50} and t_{90}.

Correlation provides linear associations but does not imply directionality or causation.

Causal Graph. The DAG was constructed using the known structure of the governing equations:

$$(C_m, C_s, C_f, n) \to S, \qquad (S, k_{bar}) \to c, \qquad (c, h) \to \{t_{50}, t_{90}\}. \tag{18}$$

Edges encode direct mechanistic influence based on Biot–Terzaghi theory. This hybrid physics–ML causal structure helps interpret feature importance, parameter identifiability, and clustering outcomes.

Results

Sensitivity of Consolidation Response to Material and Hydraulic Parameters

Figure 1 illustrates how the one-dimensional consolidation response changes as each physical parameter is varied independently. The results reflect the theoretical dependence of pore-pressure dissipation on geometry, hydraulic mobility, and storativity.

Effect of layer thickness h. Increasing the drainage path length significantly delays dissipation. A thicker layer requires fluid to travel a longer distance to reach the drained boundary, reducing the effective hydraulic gradient and decreasing the diffusion rate. As seen in the top-left panel, larger h shifts the entire curve to the right, increasing both the time to peak pressure and the duration of the decay phase. This behavior follows directly from the classical scaling

$$t \sim \frac{h^2}{c_v},$$

which shows that consolidation time increases quadratically with thickness.

Effect of porosity n. Porosity influences the fluid storage capacity through the storativity

$$S = n\,C_f + (\alpha - n)\,C_s.$$

Varying n slightly modifies the peak pressure but has almost no effect on the dissipation rate. This weak sensitivity arises because porosity affects storage but not the dominant transport parameter $c_v = k/[\gamma_f(S + \alpha^2 m_v)]$. Consequently, all curves nearly overlap.

Effect of solid bulk modulus K_s. Changes in K_s alter the solid compressibility C_s and Biot coefficient α, but the impact on the consolidation curve is minimal. In one-dimensional drainage, deformation is severely constrained, so the pore pressure evolution is governed primarily by fluid flow rather than solid elasticity. The nearly identical responses confirm that K_s plays a secondary role.

Effect of fluid bulk modulus K_f. Increasing K_f (reducing fluid compressibility) modifies storativity and slightly affects the magnitude of the early-time pressure peak. However, the overall shape and dissipation rate remain similar. Since fluid compressibility appears only through S, its influence is weaker than that of mobility or geometry.

Effect of mobility $\bar{k}$. Hydraulic mobility exhibits the strongest influence on the consolidation response. Because the consolidation coefficient is

$$c_v = \frac{\bar{k}}{S + \alpha^2 m_v},$$

increasing mobility accelerates fluid diffusion directly. The sensitivity curves clearly show:

- low mobility → slow dissipation and a long pressure tail,
- intermediate mobility → moderate decay rate,
- very high mobility → rapid pressure drop.

This confirms mobility as the dominant rate-controlling parameter in consolidation.

Effect of Young's modulus E. Variations in E affect the confined compressibility

$$m_v = \frac{1}{K + \frac{4}{3}G},$$

but the resulting changes in pressure response are small. Because deformation is constrained in 1D consolidation and drainage is governed primarily by hydraulic processes, elasticity plays a minor role. Only slight shifts in the pressure peak are observed.

The sensitivity study demonstrates that the consolidation response is controlled mainly by hydraulic mobility and layer thickness, with storativity parameters contributing moderately and elastic parameters having weak influence. These trends are fully consistent with Biot–Terzaghi theory, confirming that the digital twin captures the essential physics of one-dimensional consolidation.

These trends validate that the digital twin spans both fast-draining and slow-draining behaviors and provides a physically representative dataset.

PCA-Based Dimensionality Reduction and Clustering of pore-pressure responses

Figure 2 shows the two-dimensional embedding of all simulated pore-pressure time histories using the first two principal components (PC1 and PC2). Remarkably, despite being generated from a six-dimensional parameter space, all responses collapse onto a smooth one-dimensional manifold in the $(\mathrm{PC1}, \mathrm{PC2})$ plane. This indicates that the high-dimensional variability of the consolidation response is effectively governed by a single dominant mode of variation. Physically, this structure is consistent with classical diffusion-type solutions, where pore-pressure responses differ primarily by a horizontal shift (effective drainage timescale) and a vertical scaling (initial pressurization level).

Applying k-means clustering to the PCA scores separates the manifold into distinct groups associated with slow, intermediate, and fast dissipation regimes. The cluster at the right-most end of the manifold corresponds to responses with delayed peak pressure and long-duration tail decay, while the cluster on the left corresponds to rapidly dissipating profiles. This suggests that the clustering captures meaningful differences in consolidation rate, rather than noise or sampling artifacts.

To interpret the physical meaning of the PCA manifold, each parameter was visually mapped onto the PCA coordinates (Figure 3). Layer thickness h exhibits a strong monotonic trend along the manifold: thicker layers shift the response toward larger PC1 values, reflecting slower drainage timescales. This observation is fully consistent with Terzaghi's consolidation theory, where the characteristic time is proportional to h^2/c.

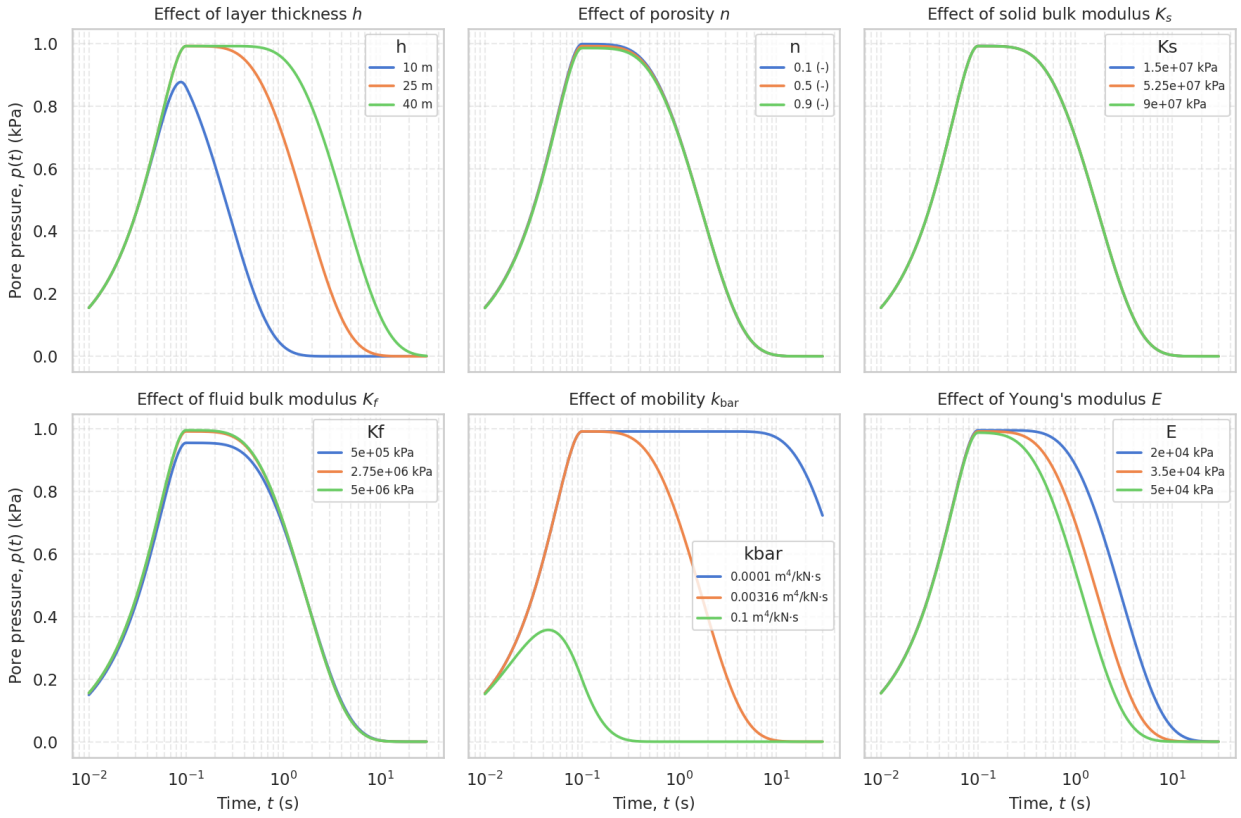

Figure 1: Parametric sensitivity of the pore-pressure response predicted by the analytical digital twin. Each subplot illustrates the effect of varying a single parameter—layer thickness h, porosity n, solid bulk modulus K_s, fluid bulk modulus K_f, hydraulic mobility k_{bar}, and Young's modulus E—while keeping all other parameters fixed. The results show that mobility and drainage path length dominate the diffusion time scale, storativity parameters produce moderate changes in peak pressure, and elastic parameters have only a minor influence on the consolidation curve.

In contrast, porosity n, solid bulk modulus K_s, fluid bulk modulus K_f, and Young's modulus E show little systematic variation along the manifold, indicating that these parameters influence pore-pressure magnitude far less than they influence timescale. Mobility k_{bar} shows a clearer trend: higher mobility shifts curves toward the left (faster dissipation), while low mobility produces long tails and moves the curves toward the right. This reinforces the dominant role of the consolidation coefficient c and hydraulic conductivity in shaping the temporal response.

Overall, the PCA and clustering analyses reveal that (i) pore-pressure profiles exist on a low-dimensional physical manifold, (ii) the manifold is organized primarily by the governing drainage timescale h^2/c, and (iii) material parameters such as porosity or solid stiffness influence only secondary, higher-order variations in the shape of the response.

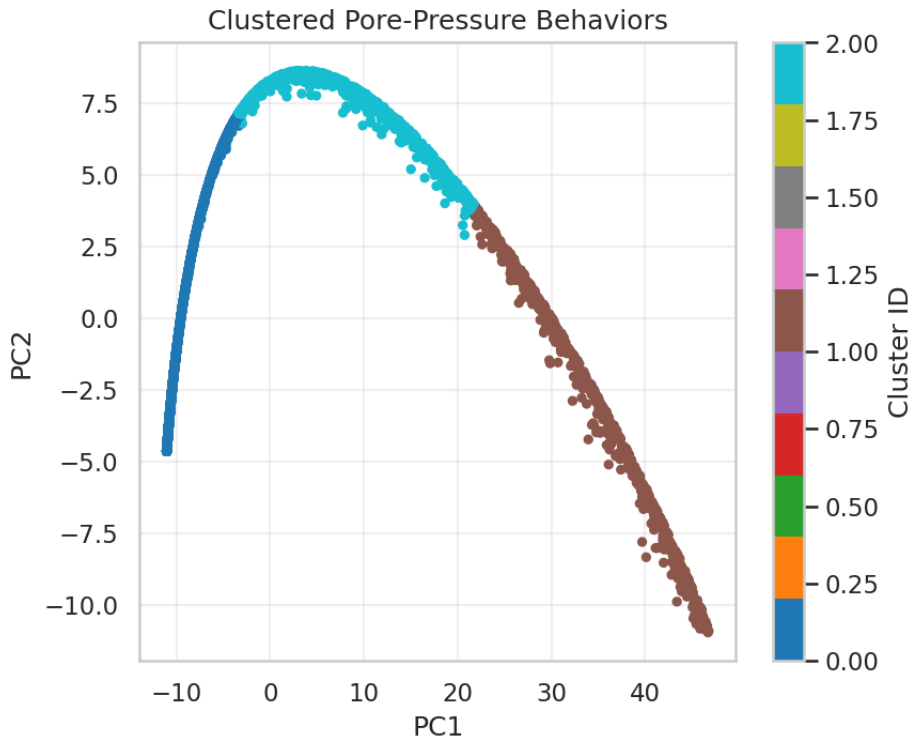

Figure 2: Two-dimensional PCA embedding of all pore-pressure time histories. Despite the seven-dimensional input parameter space, all curves collapse to a one-dimensional nonlinear manifold. Colors indicate k-means clusters, representing distinct consolidation regimes (fast, intermediate, slow dissipation).

Forward regression for predicting consolidation timescales

Figures 4 and 5 show the forward regression performance of the Random Forest model trained to predict the characteristic consolidation times t_{50} and t_{90} from the full set of material and geometric parameters. To stabilize the highly skewed distribution of the target, the regression was performed on $\log(t)$ and subsequently exponentiated to obtain predictions. This greatly improved accuracy and reduced heteroscedastic bias at large values.

Both scatter plots demonstrate that the model closely follows the 1:1 line, indicating excellent predictive capability over the entire range. The coefficient of determination R^2 exceeded 0.98 for both metrics, with mean absolute errors on the order of 0.5 s for t_{50} and 0.7 s for t_{90}. Slight underprediction is observed at very long timescales ($t > 25$ s), which corresponds to low-conductivity and large-thickness regimes where the dataset is more sparse. Nevertheless, the overall predictive fidelity remains high, confirming that the pore-pressure responses contain sufficient information to recover global consolidation times.

The associated feature importance rankings, shown in Figures 6 and 7, reveal clear physical trends. For the prediction of t_{50}, the consolidation coefficient c dominates the importance distribution, contributing nearly 80% of the model's explanatory power. This result aligns directly with Terzaghi's classical relation $t \propto h^2/c$, emphasizing that diffusion speed is the primary control on intermediate consolidation times. The next two contributors are the peak

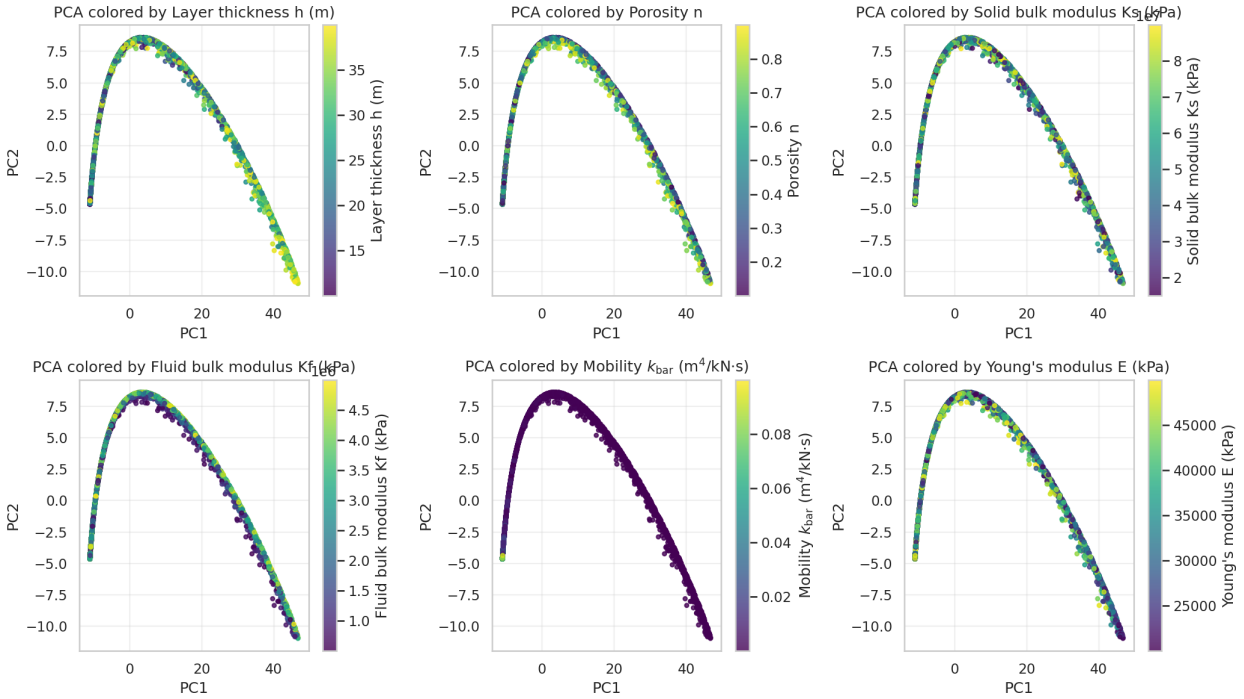

Figure 3: Physical interpretation of the PCA manifold obtained from pore-pressure responses. Each subplot colors the PCA scores by a different material or geometric parameter. Layer thickness h and mobility k_{bar} show strong ordering along the manifold, indicating their dominant control of the consolidation timescale. Other parameters exhibit weak correlation, indicating secondary influence on the response shape.

pore pressure $p_{\max}$, which reflects the initial pressurization level, and the layer thickness h, which governs the effective drainage path.

For the prediction of t_{90}, the relative importance shifts: $p_{\max}$ becomes the dominant feature, followed by c, and then h. The greater influence of $p_{\max}$ on t_{90} suggests that long-term dissipation is more sensitive to initial storage capacity and compressibility-related effects, particularly in low-permeability or highly pressurized regimes. Parameters such as intrinsic permeability, mobility, porosity, and solid/fluid bulk moduli exhibit negligible contribution in both cases, reinforcing the finding that the timescale of consolidation is governed primarily by the hydraulic diffusivity c and geometrical length scale h, with only minor secondary influences from other material parameters.

Overall, the forward regression analysis confirms that (i) consolidation timescales are highly predictable from material inputs, (ii) the learned model recovers well-known physical dependencies, and (iii) machine learning identifies the same dominant variables that govern classical consolidation theory.

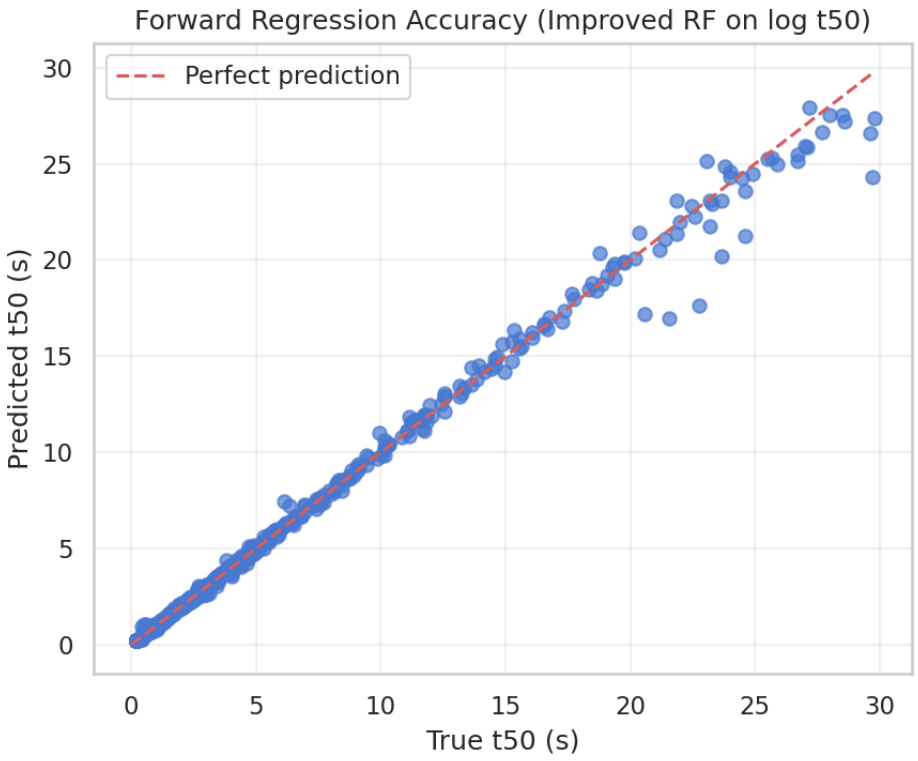

Figure 4: Forward regression accuracy for the prediction of t_{50}. The Random Forest model trained on $\log(t_{50})$ achieves excellent agreement with the true values and closely follows the perfect-prediction line.

Inverse identification of material parameters using early-time pore-pressure data

Figure 8 presents the inverse regression performance for recovering the underlying material and geometric parameters from the early-time pore-pressure response. The Random Forest inverse model was trained to map the first few temporal measurements of $p(t)$ to the set of input parameters $\{k_{\text{bar}}, n, h, K_s, K_f, E\}$, in order to assess which quantities are practically identifiable from the transient pressure signal.

The inverse prediction of the mobility parameter k_{bar} exhibits the strongest performance, with predictions closely following the 1:1 reference line and relatively low scatter. This is expected since k_{bar} directly controls the hydraulic diffusivity c and therefore strongly affects the rate at which pore pressure begins to dissipate. The early-time pressure decay contains sufficient information to infer k_{bar} with high accuracy.

Porosity n is also reasonably identifiable: although some scatter is present, a clear monotonic trend is observed. This reflects the role of n in governing storativity S, which affects the amplitude and curvature of the initial pore-pressure rise. However, because n interacts with C_f and C_s through the expression $S = nC_f + (\alpha - n)C_s$, its influence is partially masked, leading to moderate estimation uncertainty.

Parameter identification becomes increasingly challenging for the geometric thickness h. Although the model captures the overall trend, substantial scatter remains across the

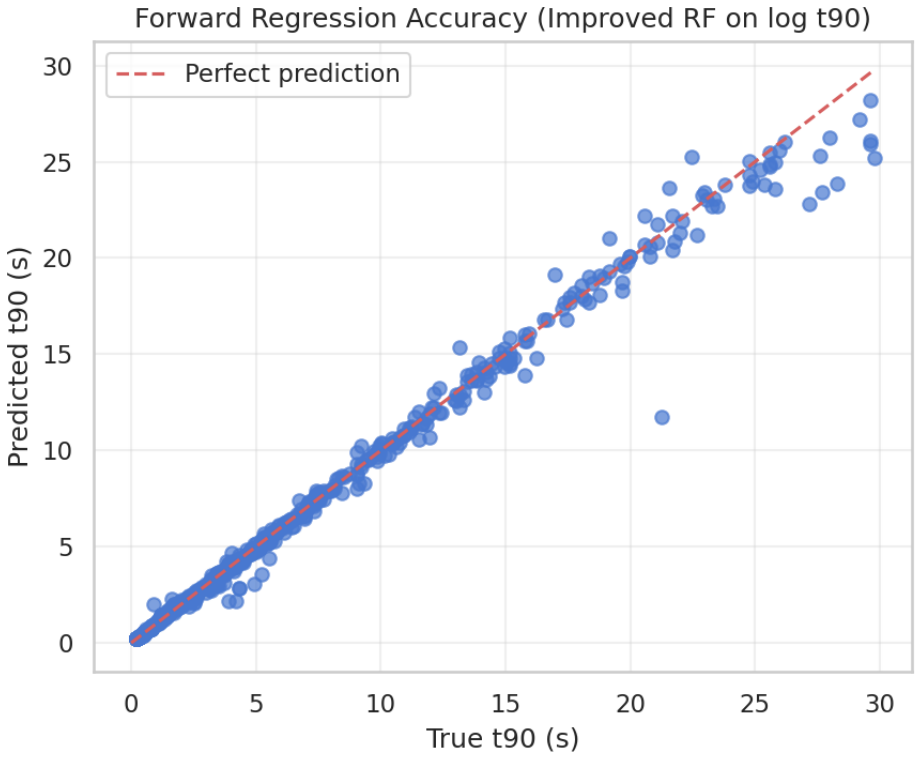

Figure 5: Forward regression accuracy for the prediction of t_{90}. The model exhibits similarly strong performance as for t_{50}, with slightly larger scatter at long times due to sparsity in the dataset.

full range. Early-time pore-pressure data contain limited information about h, since the effect of drainage length becomes more pronounced only at intermediate and late times. Consequently, h is only weakly identifiable from the first portion of the transient response.

For the solid bulk modulus K_s and fluid bulk modulus K_f, the inverse prediction accuracy is noticeably lower. Both parameters contribute to the overall compressibility of the mixture but influence the pressure evolution indirectly and nonlinearly. Their effects can be partially compensated by changes in n or α, making them difficult to disentangle from pressure data alone. The inverse model therefore shows broad scatter and weak alignment with the 1:1 line. Similarly, the Young's modulus E exhibits poor identifiability, reflecting its minimal influence on the pore-pressure field under one-dimensional consolidation loading.

Overall, the inverse modeling results reveal a strong hierarchy of identifiability: parameters that directly control the hydraulic diffusivity (e.g., k_{bar}) or storage properties (e.g., n) can be recovered with reasonable accuracy, whereas parameters contributing to mechanical stiffness (K_s, K_f, E) or long-timescale effects (h) cannot be robustly identified from early-time measurements alone. These findings highlight both the potential and inherent limitations of using transient pore-pressure responses for material characterization.

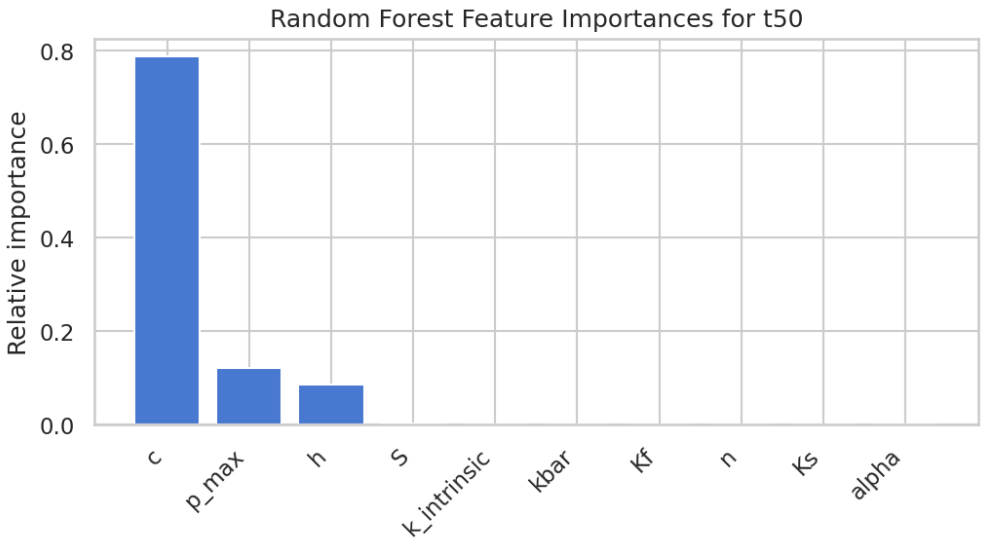

Figure 6: Feature importance for the Random Forest model predicting t_{50}. The consolidation coefficient c overwhelmingly dominates, followed by peak pore pressure $p_{\max}$ and layer thickness h, consistent with the governing diffusion physics.

Correlation analysis and physics-informed causal structure

Figure 9 presents the correlation matrix for all material parameters and consolidation metrics (t_{50} and t_{90}). Several important patterns emerge. First, the mobility parameter k_{bar} exhibits an extremely strong positive correlation with the consolidation coefficient c ($r \approx 0.96$), which is expected given the theoretical dependence $c = k_{\text{bar}}/(\alpha^2 m + S)$. This strong relationship confirms that variations in hydraulic conductivity directly translate into changes in the effective diffusivity of the medium.

Storativity S shows moderate correlation with porosity ($r = 0.48$) and strong negative correlation with the fluid bulk modulus ($r = -0.67$), consistent with its analytical definition $S = nC_f + (\alpha - n)C_s$. These relationships indicate that variations in fluid compressibility play a substantial role in governing how much fluid is stored or released during consolidation, even if they do not strongly impact the overall drainage timescale.

The consolidation times t_{50} and t_{90} show moderate negative correlation with the consolidation coefficient c ($r = -0.38$ and $r = -0.45$, respectively), reflecting the physical scaling $t \propto h^2/c$. A weakly positive correlation with layer thickness h is also observed, again consistent with classical consolidation theory. Notably, most other material parameters display correlations close to zero with t_{50} and t_{90}, suggesting that their effects on the timescale are either highly nonlinear or secondary to the dominant roles of c and h.

To complement the statistical correlation view, a physics-informed causal graph was constructed and is shown in Figure 10. In this representation, causal relationships are derived directly from Biot's poroelasticity theory and the constitutive definitions of the material parameters. Parameters such as porosity, solid bulk modulus, and fluid bulk modulus influence storativity S, while intrinsic permeability and mobility directly affect the consolidation coefficient c. Both c and h act as causal parents of the observable quantities t_{50} and t_{90},

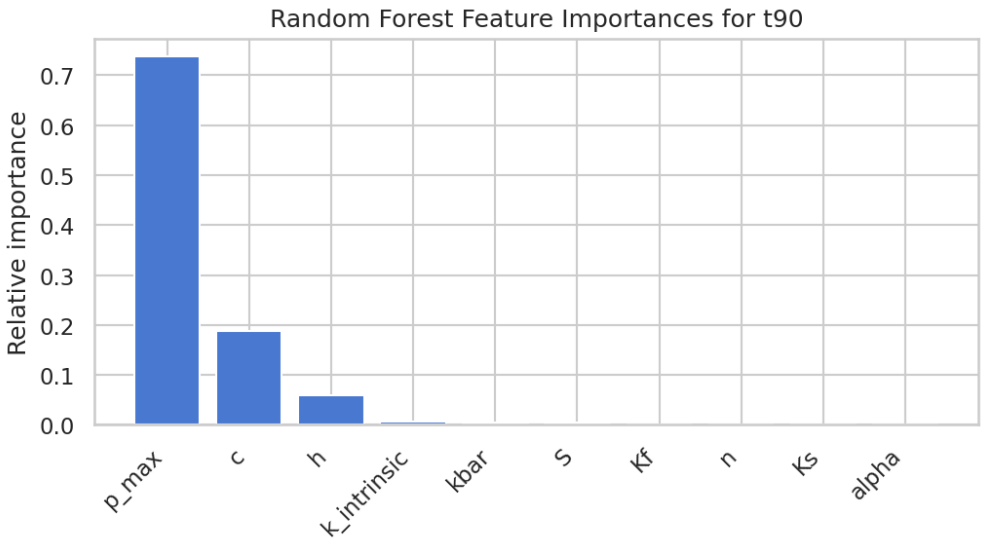

Figure 7: Feature importance for the prediction of t_{90}. At long times, $p_{\max}$ becomes the dominant predictor, followed by c and h, reflecting the increased sensitivity of late-stage dissipation to initial pressurization.

reflecting their theoretically grounded roles in controlling consolidation rate.

Comparing the correlation matrix with the causal graph reveals a consistent picture: (i) c is the dominant governing parameter for consolidation timing; (ii) S plays a secondary but important role through storage effects; and (iii) many material parameters have weak direct correlations with t_{50} and t_{90} because their influence is mediated through c or S, rather than acting directly. The causal graph also clarifies situations where correlation alone may be misleading—for example, porosity appears weakly correlated with consolidation times, yet it is causally upstream of S and therefore affects the response indirectly.

Overall, the combined statistical and physics-based analysis provides a coherent understanding of how material and geometric parameters shape consolidation behavior. The results confirm that the dominant factors controlling the rate of dissipation are hydraulic diffusivity c and drainage length h, while other parameters influence the response primarily through their effects on storativity or initial pressurization.

Conclusion

This study presented an integrated physics-based learning framework for analyzing one-dimensional consolidation behavior, combining analytical poroelastic theory with large-scale digital-twin data generation and multi-modal data analysis. Using the closed-form solution to Biot's consolidation equation, more than 5000 high-fidelity pore-pressure histories were generated across a wide range of material and geometric parameters. This dataset enabled a systematic investigation into the physical drivers of consolidation, the dimensionality of the response space, and the predictive capacity of modern regression and inverse-identification tools.

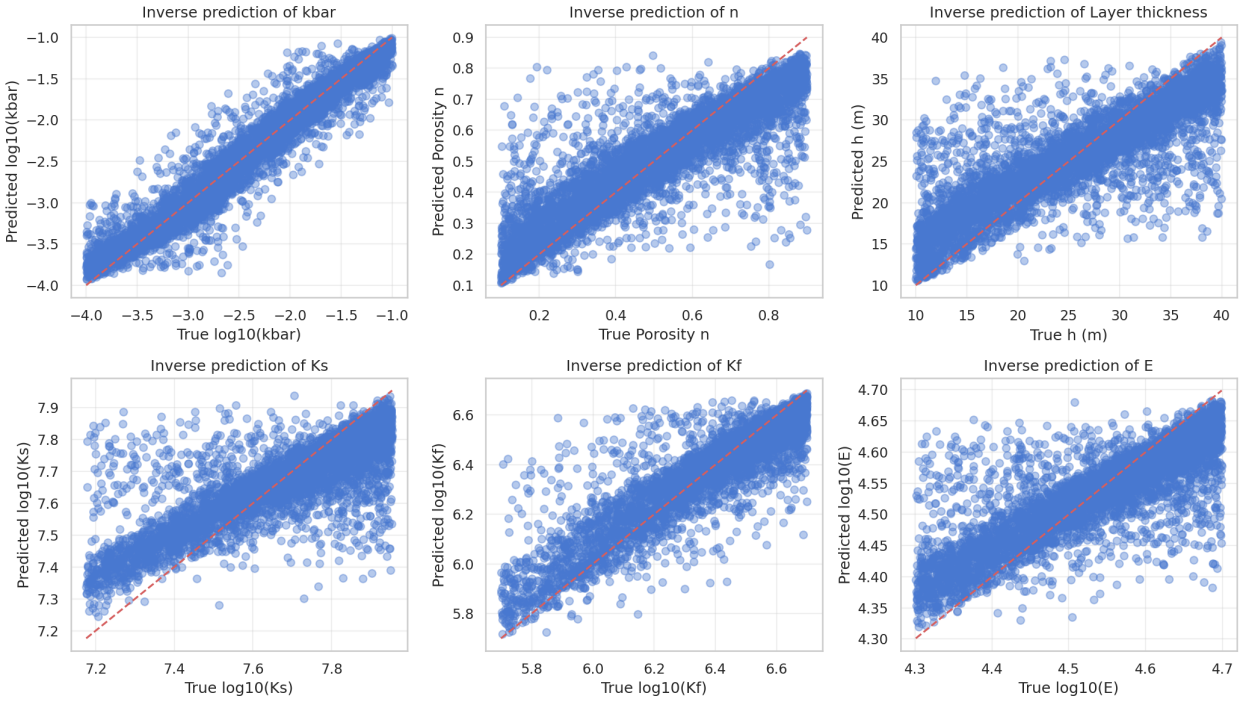

Figure 8: Inverse prediction accuracy for recovering material and geometric parameters from early-time pore-pressure data. Mobility k_{bar} and porosity n are reasonably identifiable, whereas layer thickness h, solid bulk modulus K_s, fluid bulk modulus K_f, and Young's modulus E show large scatter, indicating weak identifiability from the early-time response.

Sensitivity analysis demonstrated that consolidation responses are dominated by the drainage length h and the consolidation coefficient c, while other parameters influence only secondary features of the solution. Principal component analysis revealed that all pore-pressure histories collapse onto a smooth one-dimensional manifold, highlighting the fundamentally low-dimensional nature of consolidation dynamics. The manifold was shown to be organized primarily by h^2/c, consistent with classical diffusion theory, while clustering identified distinct behavioral regimes corresponding to fast, intermediate, and slow drainage processes.

Forward regression using Random Forest models achieved excellent predictive accuracy for the characteristic times t_{50} and t_{90}, recovering both the correct scaling behavior and the dominant physical dependencies. Feature importance ranking confirmed the theoretical expectations, with c and h emerging as the leading predictors across all models. Inverse modeling of material parameters from early-time pore pressures revealed a clear hierarchy of identifiability: hydraulic parameters such as k_{bar} and porosity n can be recovered with reasonable accuracy, whereas solid and fluid moduli (K_s, K_f, E) are weakly identifiable due to their indirect or nonlinear contributions. These findings provide practical guidance for parameter estimation in poroelastic systems and highlight the inherent limitations of relying solely on early-time data.

Correlation analysis and a physics-informed causal graph further clarified the relationships among parameters, reinforcing that many observable trends arise from indirect depen-

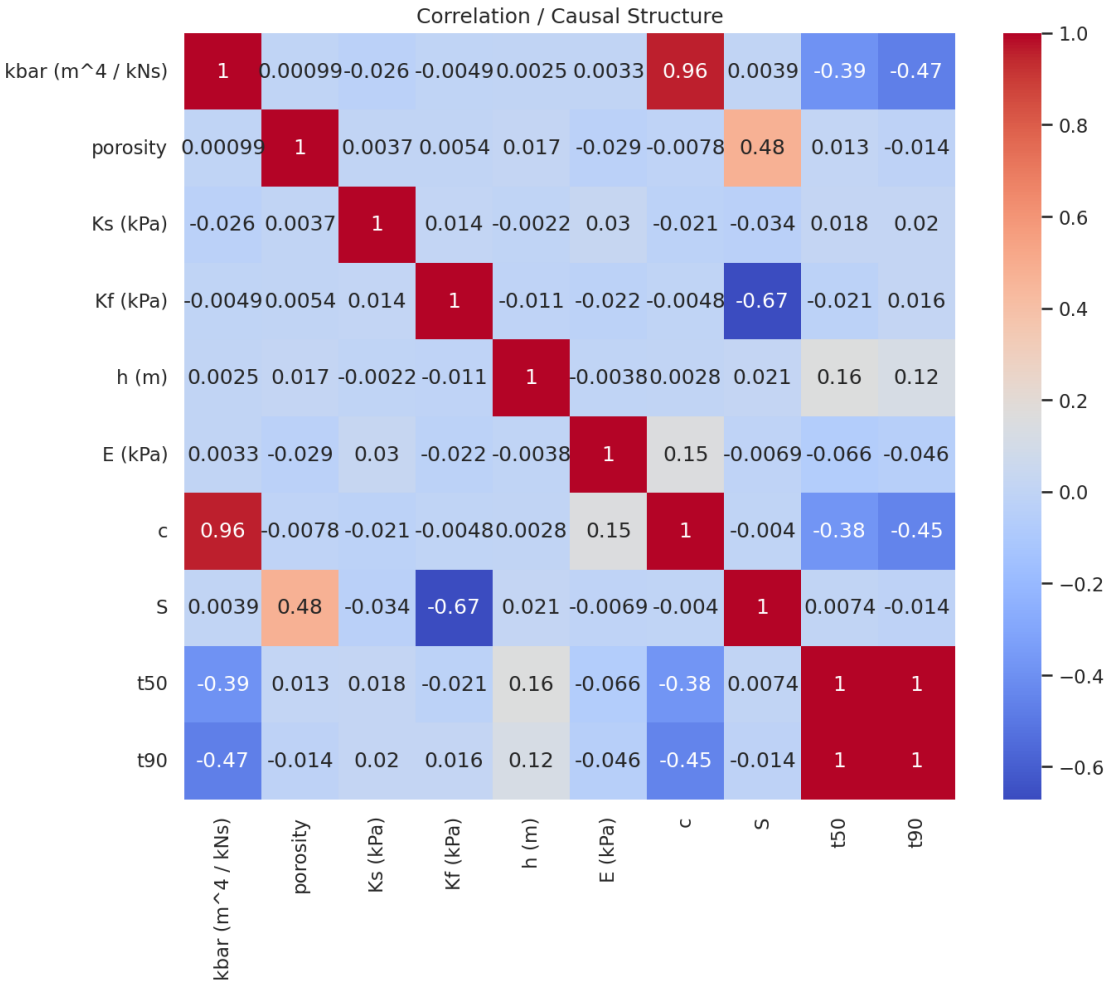

Figure 9: Correlation matrix of all material parameters and consolidation timescales. Mobility k_{bar} shows strong correlation with the consolidation coefficient c, while porosity and fluid compressibility largely control storativity S. Consolidation times t_{50} and t_{90} correlate primarily with c and h, consistent with diffusion-based theory.

dencies mediated through storativity S and diffusivity c. Together, these results demonstrate that combining analytical poroelastic theory with machine learning offers a powerful pathway for understanding, predicting, and inverting consolidation behavior.

Overall, the framework presented here advances both the physical interpretation and data-driven modeling of consolidation phenomena. The insights gained from this work provide a foundation for future extensions to nonlinear, anisotropic, and multiphase poromechanical systems, and open new opportunities for data-efficient parameter identification, uncertainty quantification, and model-based decision making in geomechanics and soft-matter engineering.

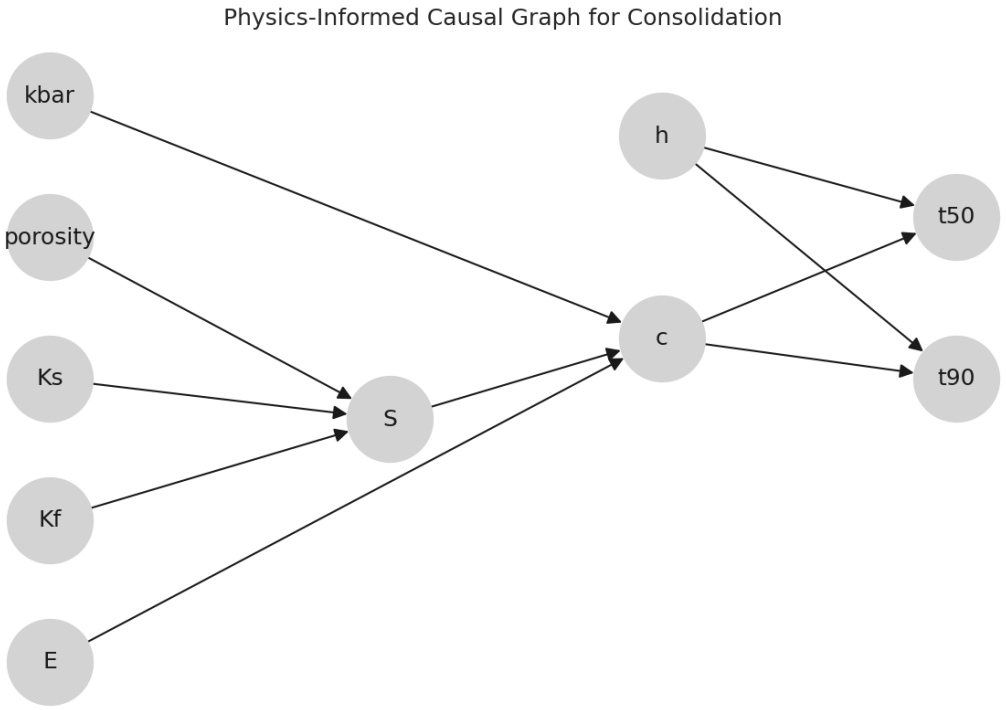

Figure 10: Physics-informed causal structure for consolidation. Material parameters influence storativity S and diffusivity c, which in turn govern the observable consolidation metrics t_{50} and t_{90}. The structure clarifies indirect effects that are not easily identified from correlations alone.

References

[1] M. A. Biot, "General theory of three-dimensional consolidation," *Journal of applied physics*, vol. 12, no. 2, pp. 155–164, 1941.

[2] J. A. Kluge, N. C. Rosiello, G. G. Leisk, D. L. Kaplan, and A. L. Dorfmann, "The consolidation behavior of silk hydrogels," *Journal of the mechanical behavior of biomedical materials*, vol. 3, no. 3, pp. 278–289, 2010.

[3] T. Nguyen and A. Oloyede, "Predictive rheological models for the consolidation behaviour of articular cartilage under static loading," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 215, no. 6, pp. 565–577, 2001.

[4] D. Chou, Y.-D. Li, Z. Mustansar, and C.-Y. Chung, "Using a poroelastodynamic model to investigate the dynamic behaviour of articular cartilage," *Computer Methods and Programs in Biomedicine*, vol. 233, p. 107481, 2023.

[5] S. Yin, M. B. Dusseault, and L. Rothenburg, "Coupled multiphase poroelastic analysis of reservoir depletion including surrounding strata," *International Journal of Rock Mechanics and Mining Sciences*, vol. 44, no. 5, pp. 758–766, 2007.

[6] K. Terzaghi and R. B. Peck, "Soil mechanics," *Engineering Practice. John Wiley and Sons, Inc., New York*, 1948.

[7] W. Lo, R. I. Borja, J.-H. Deng, and J.-W. Lee, "Poroelastic theory of consolidation for a two-layer system with an upper unsaturated soil and a lower saturated soil under fully permeable boundary conditions," *Journal of Hydrology*, vol. 596, p. 125700, 2021.

[8] W. Pariseau, "Poroelastic-plastic consolidation—analytical solution," *International journal for numerical and analytical methods in geomechanics*, vol. 23, no. 7, pp. 577–594, 1999.

[9] M. Kazemian, R. Penta, H. Dehghani, A. Hassani, and A. M. Goudarzi, "An analytical-based computational framework to study poroelastic column consolidation for solid compressibility at large deformations," *International Journal of Solids and Structures*, p. 113436, 2025.

[10] M. H. Esteki, A. A. Alemrajabi, C. M. Hall, G. K. Sheridan, M. Azadi, and E. Moeendarbary, "A new framework for characterization of poroelastic materials using indentation," *Acta biomaterialia*, vol. 102, pp. 138–148, 2020.

[11] M. Mittal, S. C. Satapathy, V. Pal, B. Agarwal, L. M. Goyal, and P. Parwekar, "Prediction of coefficient of consolidation in soil using machine learning techniques," *Microprocessors and Microsystems*, vol. 82, p. 103830, 2021.

[12] M. D. Nguyen, B. T. Pham, T. T. Tuyen, H. P. Hai Yen, I. Prakash, T. T. Vu, K. Chapi, A. Shirzadi, H. Shahabi, J. Dou, *et al.*, "Development of an artificial intelligence approach for prediction of consolidation coefficient of soft soil: a sensitivity analysis," *The Open Construction & Building Technology Journal*, vol. 13, no. 1, 2019.

[13] M. Delpisheh, B. Ebrahimpour, A. Fattahi, M. Siavashi, H. Mir, H. Mashhadimoslem, M. A. Abdol, M. Ghorbani, J. Shokri, D. Niblett, *et al.*, "Leveraging machine learning in porous media," *Journal of Materials Chemistry A*, 2024.

[14] H. Huang, J. Li, H. Yang, B. Wang, R. Gao, M. Luo, W. Li, G. Zhang, and L. Liu, "Research on prediction methods of formation pore pressure based on machine learning," *Energy Science & Engineering*, vol. 10, no. 6, pp. 1886–1901, 2022.

[15] A. Alghamdi, M. A. Hesse, J. Chen, U. Villa, and O. Ghattas, "Bayesian poroelastic aquifer characterization from insar surface deformation data. 2. quantifying the uncertainty," *Water Resources Research*, vol. 57, no. 11, p. e2021WR029775, 2021.

[16] M. Vasilyeva and A. Tyrylgin, "Machine learning for accelerating macroscopic parameters prediction for poroelasticity problem in stochastic media," *Computers & Mathematics with Applications*, vol. 84, pp. 185–202, 2021.

[17] S. Park, Y. Shin, and J. Choo, "Deep operator network for surrogate modeling of poroelasticity with random permeability fields," *arXiv preprint arXiv:2509.11966*, 2025.

[18] A. Gamra, J. Ninić, and B. Ghiassi, "Optimising synthetic datasets for machine learning-based prediction of building damage due to tunnelling," *Tunnelling and Underground Space Technology*, vol. 152, p. 105961, 2024.

[19] K. Mirniazy, M. Karamzadeh, A. Molaei, I. Shiri, and J. Dargahi, "Supervised deep learning with finite element generated data for force estimation in robotic-assisted liver surgery," *Available at SSRN 4102847*, 2022.

[20] R. Li, C.-S. Zhang, and Z. Chen, "A stokes–dual–porosity–poroelasticity model and discontinuous galerkin method for the coupled free flow and dual porosity poroelastic medium problem," *Journal of Scientific Computing*, vol. 102, no. 2, p. 41, 2025.

[21] Y. W. Bekele, "Physics-informed deep learning for one-dimensional consolidation," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 13, no. 2, pp. 420–430, 2021.

[22] H. Guo and Z.-Y. Yin, "A novel physics-informed deep learning strategy with local time-updating discrete scheme for multi-dimensional forward and inverse consolidation problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 421, p. 116819, 2024.

[23] B. Yuan, A. Heitor, H. Wang, and X. Chen, "Physics-informed deep learning to solve three-dimensional terzaghi consolidation equation: forward and inverse problems," *arXiv preprint arXiv:2401.05439*, 2024.

[24] S. Kubo, M. Nishio, J. Inoue, T. Miyamoto, and P.-j. Chun, "Effective prediction of ground consolidation settlement by physics-informed neural networks," *Intelligence, Informatics and Infrastructure*, vol. 6, no. 1, pp. 137–149, 2025.

[25] S. Urcun, P.-Y. Rohan, W. Skalli, P. Nassoy, S. P. Bordas, and G. Sciumè, "Digital twinning of cellular capsule technology: Emerging outcomes from the perspective of porous media mechanics," *PLoS One*, vol. 16, no. 7, p. e0254512, 2021.

[26] A. Verruijt, "Theory and problems of poroelasticity," *Delft University of Technology*, vol. 71, p. 465, 2013.