Textile & Leather Review

ISSN 2623-6281 | www.tlr-journal.com | 10.31881/TLR

Anomalous Diffusion Modeling in Heterogeneous Media via Fractional Differential Equations and Neural Networks

Zihao Yu

How to cite: Yu Z. Anomalous Diffusion Modeling in Heterogeneous Media via Fractional Differential Equations and Neural Networks. Textile & Leather Review. 2026; 9 4278-4302 https://doi.org/10.31881/TLR.2026.4278

How to link: https://doi.org/10.31881/TLR.2026.4278

Published: 25 April 2026

Anomalous Diffusion Modeling in Heterogeneous Media via Fractional Differential Equations and Neural Networks

Zihao Yu

College of Science, North China University of Technology, Beijing 100144, China
jiazhutao233@sina.com

Article
https://doi.org/10.31881/TLR.2026.4278

Published 25 April 2026

ABSTRACT

This paper proposes a mechanistically constrained, data-driven framework for subdiffusive anomalous diffusion modeling in heterogeneous media by integrating fractional differential equations with neural networks. A time-fractional diffusion model is first established for heterogeneous media, where a spatially varying diffusion coefficient characterizes medium heterogeneity. A joint learning architecture is then constructed, consisting of a state network that approximates the diffusion field and a parameter network that identifies unknown medium parameters. Observational data, governing-equation residuals, and initial-boundary conditions are incorporated into a unified physics-informed loss function, enabling simultaneous field reconstruction and parameter identification. To handle the nonlocality of the fractional derivative, a history-dependent numerical discretization is employed and embedded into the neural-network training process. Numerical experiments show that the proposed method can accurately reconstruct anomalous diffusion fields under various heterogeneous settings, while effectively identifying spatially varying diffusion coefficients and fractional-order parameters. In addition, the framework maintains good robustness under sparse and noisy observations. These results indicate that the integration of fractional differential equations and neural networks provides an effective approach for the modeling and inversion of anomalous diffusion processes in heterogeneous media.

KEYWORDS

anomalous diffusion, heterogeneous media, fractional diffusion equation, neural networks

INTRODUCTION

Subdiffusive anomalous diffusion, in which the mean squared displacement grows as $\langle x^2 \rangle \sim t^\alpha$ with $\alpha \in (0,1)$, is widely observed in heterogeneous media such as porous materials, fractured rocks, complex composites, and biological tissues [1, 2]. In contrast to classical Fickian diffusion, it is typically characterized by non-Gaussian transport, heavy-tailed residence-time distributions, and a nonlinear scaling between the mean squared displacement and time [3]. Such transport behaviors are often associated with pronounced memory effects and nonlocal interactions, which makes conventional integer-order diffusion equations inadequate for accurate description [4, 5]. In recent decades, fractional differential equations have emerged as an effective mathematical framework for modeling anomalous diffusion, owing to their intrinsic ability to capture long-memory and nonlocal mechanisms [6, 7]. For heterogeneous media in particular, diffusion models involving spatially varying coefficients and even variable fractional orders offer a more flexible and physically meaningful way to represent local variations in transport dynamics [8].

Existing studies on anomalous diffusion modeling have mainly evolved along two lines. On the one hand, substantial progress has been made in the analysis and numerical treatment of fractional diffusion equations, and it has been shown that time-fractional, space-fractional, and variable-order models are better suited to describing non-Fickian transport in heterogeneous media. On the other hand, with the rapid development of machine learning, especially physics-informed neural networks, neural-network-based solution and identification methods for partial differential equations have received increasing attention [9, 10]. For fractional partial differential equations, representative approaches such as fractional physics-informed neural networks (fPINNs) have been developed by combining automatic differentiation with numerical discretization of fractional operators, thereby enabling the treatment of both forward and inverse problems within a unified framework [11, 12]. More recently, improved variants such as Laplace-fPINNs have been proposed to alleviate the computational burden induced by fractional derivatives and to enhance training efficiency [13]. Nevertheless, in heterogeneous media, it remains challenging to effectively couple fractional-order mechanistic models with neural networks so as to simultaneously reconstruct diffusion fields, identify heterogeneous parameters, and achieve robust modeling from sparse and noisy observations.

Motivated by these considerations, this paper develops a mechanistically constrained, data-driven framework for anomalous diffusion modeling in heterogeneous media by integrating fractional differential equations with neural networks. Specifically, we formulate a fractional diffusion model for heterogeneous media with spatially varying diffusion coefficients and fractional-order parameters. We then construct a coupled state-parameter network framework. The state network reconstructs the diffusion field, while the parameter network identifies unknown medium parameters. A unified physics-informed loss function integrates observational data, initial-boundary conditions, and the governing equation residual, enabling simultaneous field reconstruction and parameter identification. Finally, a series of numerical experiments are conducted to evaluate the effectiveness of the method in anomalous diffusion field reconstruction, heterogeneous parameter inversion, and robustness under sparse and noisy measurements, with the aim of providing an effective modeling strategy for anomalous transport in complex heterogeneous media.

FRACTIONAL DIFFUSION MODEL FOR HETEROGENEOUS MEDIA

Governing Equation

To characterize anomalous diffusion in heterogeneous media, we consider the following time-fractional diffusion model:

$$ {}_0^C D_t^{\alpha} u(x,t) = \nabla \cdot (D(x) \nabla u(x,t)) + f(x,t), x \in \Omega, t \in (0,T] \tag{1} $$

where $u(x,t)$ denotes the diffusion state variable, such as a concentration field or mass fraction; $\Omega \subset \mathbb{R}^d$ is the spatial domain; $D(x)$ is the spatially varying diffusion coefficient used to represent material heterogeneity; $f(x,t)$ is the source term; and ${}_0^C D_t^{\alpha}$ denotes the Caputo fractional derivative of order $\alpha \in (0,1)$ with respect to time. In contrast to the classical diffusion equation, this model introduces a fractional memory term in time, which enables the description of nonlocal dynamical behaviors arising from trapping, retention, and history-dependent transport processes.

In this work, the Caputo derivative is adopted and defined by:

$$ {}_0^C D_t^{\alpha} u(x,t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\partial u(x,\tau)/\partial \tau}{(t-\tau)^{\alpha}} \, d\tau, 0 < \alpha < 1 \tag{2} $$

where $\Gamma(\cdot)$ denotes the Gamma function. This definition is particularly convenient in physical applications because it allows the initial conditions to be prescribed in a form consistent with classical integer-order differential equations. To complete the model formulation, the following initial and boundary conditions are imposed:

$$ u(x,0) = u_0(x), x \in \Omega \tag{3} $$

And:

$$u(x,t) = g(x,t), x \in \partial\Omega, t \in (0,T] \quad (4)$$

where $u_0(x)$ is the prescribed initial condition and $g(x,t)$ is the boundary condition. In practical numerical experiments, Neumann or mixed boundary conditions may also be considered depending on the application scenario. It is worth noting that near the boundaries of heterogeneous media, the nonlocal history integral in the Caputo derivative may give rise to weak singularities as t → 0, since the solution regularity is generally limited by the smoothness of the initial data. In the present framework, the boundary condition is enforced as a soft constraint through the loss term $\mathcal{L}_{bc}$, which allows the network to implicitly accommodate such near-boundary memory effects without requiring explicit singularity treatment; the physics residual loss further ensures that the fractional history accumulated from the interior is consistently propagated to the boundary region.

The above governing equation captures two essential features of anomalous diffusion in heterogeneous media. First, the diffusion coefficient $D(x)$ varies in space, reflecting the spatial heterogeneity of the internal structure, pore distribution, or transport capacity of the medium. Second, the fractional derivative, through its history-integral form, accounts for the long-memory effect that is characteristic of anomalous diffusion. Consequently, the model not only reduces to the classical diffusion equation as a limiting case, but also provides a more general framework for describing non-Fickian transport in complex media.

Heterogeneous Diffusion Coefficient and Fractional Order

In classical diffusion models, the diffusion coefficient is often assumed to be constant, implying that the medium is spatially homogeneous and that the transport mechanism does not vary with location. In real complex media, however, material properties, pore structures, fracture distributions, and local microscopic environments are typically highly nonuniform, leading to significant spatial variations in diffusion behavior. For this reason, the diffusion coefficient in the present study is modeled as a spatially dependent function $D(x)$, so that the impact of medium heterogeneity on transport can be represented more realistically. Depending on the physical setting, $D(x)$ may take the form of a piecewise constant function, a smoothly varying function, or a high-contrast spatial profile, thereby allowing the model to describe layered media, graded materials, or strongly heterogeneous environments, respectively.

In addition to the heterogeneity of the diffusion coefficient, the strength of memory effects in anomalous diffusion may itself vary with spatial position. To further enhance the descriptive capability of the model, the fractional order can be generalized to a spatially dependent parameter by setting $\alpha = \alpha(x)$, which leads to the following variable-order fractional diffusion model:

$$ {}_0^C D_t^{\alpha(x)} u(x,t) = \nabla \cdot (D(x) \nabla u(x,t)) + f(x,t), x \in \Omega, t \in (0,T] \tag{5} $$

In this formulation, $\alpha(x)$ is used to characterize spatial variations in memory strength and retention behavior: a smaller fractional order usually corresponds to stronger history dependence and more pronounced subdiffusive dynamics, whereas a larger fractional order indicates that the system becomes closer to classical diffusion. In this sense, a spatially varying fractional order provides an additional mechanism for describing localized changes in transport behavior within heterogeneous media.

It should be noted that fixed-order and variable-order models correspond to different levels of modeling complexity. A fixed fractional-order model is structurally simpler and is suitable when the global memory effect is relatively uniform throughout the medium [14]. In contrast, a variable-order model is better suited to media with pronounced local differences in transport mechanisms, although it also introduces greater challenges in parameter identification and numerical computation. With this in mind, the present study adopts a spatially varying diffusion coefficient together with a fixed fractional order as the basic modeling framework, while the influence of the fractional order on the modeling results is further examined in the subsequent analysis. This choice provides a reasonable balance between model expressiveness and computational tractability. Physically, the fractional order α is closely related to the waiting-time distribution of diffusing particles in the medium: for a power-law waiting-time distribution with exponent β, one has α = β − 1, so that α ∈ (0,1) corresponds to subdiffusive trapping dynamics; in practice, α can be estimated from the slope of the mean squared displacement curve or inferred from the fractal dimension of the pore structure, and the value α = 0.78 used in this study is consistent with reported estimates for moderately heterogeneous porous media.

Inverse Problem Formulation

The focus of this study is not limited to the forward solution of the fractional diffusion equation, but more importantly concerns the recovery of the diffusion process and unknown physical parameters from limited observations. Let:

$$ \mathcal{D} = \{(x_i, t_i, \hat{u}_i)\}_{i=1}^{N_d} \tag{6} $$

be a set of observational data collected over the spatiotemporal domain $\Omega \times (0,T]$, where (x_i, t_i) denotes the location of the observation and $\hat{u}_i$ is the corresponding measured value. The objective is to use these discrete observations, together with the governing fractional diffusion equation, initial condition, and boundary condition, to reconstruct the unknown state field $u(x,t)$ while simultaneously identifying the medium parameters $D(x)$, and, when necessary, the fractional order α or $\alpha(x)$.

Accordingly, we consider the following inverse problem: given the source term $f(x,t)$, the initial condition $u_0(x)$, and the boundary condition $g(x,t)$, determine the state solution $u(x,t)$ and the unknown parameter set $\mathcal{P}$ from finite and possibly noisy observations, subject to the governing equation. Specifically, the parameter set may take the form:

$$\mathcal{P} = \{D(x)\} \quad \text{or} \quad \mathcal{P} = \{D(x), \alpha\} \tag{7}$$

and may be further extended, in a more general setting, to:

$$\mathcal{P} = \{D(x), \alpha(x)\} \tag{8}$$

Compared with the corresponding forward problem, this inverse problem is substantially more ill-posed. The main difficulties arise from three aspects. First, observational data are often sparse and contaminated by noise, which makes stable reconstruction in a high-dimensional function space inherently difficult. Second, the nonlocal nature of the fractional derivative introduces long-history coupling in time, thereby increasing the complexity of parameter identification. Third, in heterogeneous media, the diffusion coefficient and the fractional-order parameter may exhibit coupled effects, so that different parameter combinations can produce similar observational responses. This raises a formal identifiability concern: uniqueness of the inverse problem is not guaranteed in general, and relies on sufficient spatiotemporal coverage of observations, the regularity assumptions imposed on D(x) and α, and the constraining power of the governing equation. In the present framework, the physics residual loss acts as a structural regularizer that restricts the admissible parameter space to fields consistent with the fractional diffusion equation, thereby reducing—though not eliminating—the risk of non-unique solutions. Empirically, the low parameter errors reported in Table 2 suggest that, under the observation densities and problem configurations considered here, the combined constraints are sufficient to yield practically unique identification

To address these challenges, a mechanistically constrained, data-driven modeling strategy is adopted in this work, in which the observational data and the governing equation are incorporated into a unified learning framework. More specifically, the unknown state field and parameter field are not solved pointwise by conventional grid-based discretization, but are approximated by neural networks as continuous functions. At the same time, the residuals of the governing fractional diffusion equation, the initial condition, and the boundary condition are embedded into the optimization objective to improve physical consistency and enhance the stability of parameter identification. Based on this inverse problem formulation, the neural-network-based framework and the corresponding training strategy will be presented in the next section.

NEURAL-NETWORK-BASED SOLUTION FRAMEWORK

State and Parameter Networks

To reconstruct the diffusion state and identify unknown parameters in anomalous diffusion through heterogeneous media, we develop a neural-network-based framework consisting of a state network and a parameter network. The state network is used to approximate the diffusion state variable $u(x,t)$, whereas the parameter network is introduced to represent unknown medium parameters, such as the spatially varying diffusion coefficient $D(x)$ or the fractional-order parameter $\alpha(x)$. Unlike conventional numerical methods that rely on mesh-based discretization and pointwise computation, the proposed framework represents unknown physical quantities as continuous function approximators, thereby enabling global modeling under sparse observations through the expressive power of neural networks.

Specifically, the state network is denoted by:

$$u_\theta(x,t) \approx u(x,t) \tag{9}$$

where θ denotes the trainable parameters of the network. The parameter network for the diffusion coefficient is written as:

$$D_\phi(x) \approx D(x) \tag{10}$$

where ϕ denotes the corresponding network parameters. When a spatially varying fractional order is also to be inferred, an additional parameter network can be introduced as:

$$\alpha_{\psi}(x) \approx \alpha(x) \tag{11}$$

where ψ denotes the trainable parameters associated with the fractional-order field. Under this formulation, the original inverse problem is converted into a joint optimization problem, in which the neural-network parameters θ , ϕ , and, when necessary, ψ , are learned by enforcing consistency with both observational data and the governing fractional diffusion equation.

From a modeling perspective, separating the state network from the parameter network offers two important advantages. First, it structurally decouples the learning of the state variable from that of the physical parameters, which improves both the stability and interpretability of the inverse problem. Second, for spatially distributed parameters in heterogeneous media, this representation avoids direct high-dimensional inversion on a discrete mesh and thus significantly reduces the number of optimization variables while retaining the ability to capture complex parameter fields. In practice, the state network is typically implemented as a fully connected feedforward network with input (x,t) , whereas the parameter network takes the spatial coordinate x as input and outputs the corresponding parameter distribution. To ensure physical admissibility, suitable output transformations may be imposed, such as an exponential mapping to guarantee $D_{\phi}(x) > 0$, or a sigmoid-type mapping to enforce $\alpha_{\psi}(x) \in (0,1)$. At heterogeneous interfaces where flux continuity replaces simple Dirichlet matching, the parameter network's continuous representation of D(x) implicitly enforces a smooth transition of the diffusion coefficient, which partially captures the non-local interfacial effects without requiring an explicit jump condition; for applications demanding stricter interface treatment, Neumann or Robin boundary conditions can be incorporated into $\mathcal{L}_{bc}$ in a straightforward manner.

Physics-Informed Loss Construction

To ensure that the neural-network predictions not only fit the observational data but also satisfy the physical laws described by the fractional diffusion equation, we construct a unified physics-informed loss function that incorporates the data term, the governing-equation term, the initial-condition term, and the boundary-condition term into a single training objective [15]. Based on the outputs of the state and parameter networks, the residual of the fractional diffusion equation is defined as:

$$\mathcal{R}(x,t) \, \frac{C}{0} \, D_t^{\alpha} u_{\theta}(x,t) - \nabla \cdot (D_{\phi}(x) \nabla u_{\theta}(x,t)) - f(x,t) \tag{12}$$

For variable-order models, the fixed order α in the above expression can be further replaced by $\alpha_{\psi}(x)$. This residual measures the extent to which the neural-network prediction deviates from the governing equation, and therefore serves as the core component of the physical constraint.

Based on this residual, several loss terms are introduced. First, for the observational dataset $\mathcal{D} = \{(x_i, t_i, \hat{u}_i)\}_{i=1}^{N_d}$, the data fitting loss is defined by:

$$\mathcal{L}_{data} = \frac{1}{N_d} \sum_{i=1}^{N_d} | \; u_\theta(x_i, t_i) - \hat{u}_i \; |^2 \tag{13}$$

Second, on a set of collocation points $\left\{(x_j^r, t_j^r)\right\}_{j=1}^{N_r}$ sampled in the spatiotemporal domain, the physics residual loss is given by:

$$\mathcal{L}_{physics} = \frac{1}{N_r} \sum_{j=1}^{N_r} | \; \mathcal{R}(x_j^r, t_j^r) \; |^2 \tag{14}$$

Similarly, for the initial and boundary conditions, we define:

$$\mathcal{L}_{ic} = \frac{1}{N_i} \sum_{j=1}^{N_i} | \; u_\theta(x_j^i, 0) - u_0(x_j^i) \; |^2 \tag{15}$$

And:

$$\mathcal{L}_{bc} = \frac{1}{N_b} \sum_{j=1}^{N_b} | \; u_\theta(x_j^b, t_j^b) - g(x_j^b, t_j^b) \; |^2 \tag{16}$$

The overall loss function is therefore written as:

$$\mathcal{L} = \lambda_d \mathcal{L}_{data} + \lambda_p \mathcal{L}_{physics} + \lambda_i \mathcal{L}_{ic} + \lambda_b \mathcal{L}_{bc} \tag{17}$$

where $\lambda_d, \lambda_p, \lambda_i, \lambda_b$ are weighting coefficients that balance the relative importance of data fidelity and physical consistency; N_d , N_r , N_i , and N_b denote the number of observational data points, physics collocation points, initial-condition points, and boundary-condition points, respectively; and x_j^r , x_j^i , x_j^b denote the corresponding spatial coordinates sampled in each subset. In practice, these weights may either be prescribed as fixed constants or updated adaptively according to the scale of each loss component during training. In the presence of noisy observations, assigning a relatively larger weight to the physics-based term often helps suppress overfitting and improves the robustness of parameter identification. Through this unified loss formulation, the mechanistic information embedded in the fractional diffusion equation is effectively incorporated into the neural-network training process, allowing the learned model to achieve a balance between consistency with the data and consistency with the underlying physics.

Numerical Treatment of Fractional Derivatives

Unlike classical integer-order partial differential equations, fractional diffusion equations involve derivatives with strong nonlocality, whose evaluation depends on integral information over the entire history of the solution. As a result, the numerical treatment of the fractional term is one of the most critical issues in the construction of a neural-network-based solution framework. For integer-order spatial derivatives, automatic differentiation can be directly used to compute $\nabla u_\theta(x,t)$ and $\nabla \cdot (D_\phi(x)\nabla u_\theta(x,t))$. However, for the time-fractional derivative ${}_0^C D_t^\alpha u_\theta(x,t)$, automatic differentiation cannot directly produce an expression consistent with the Caputo definition, and a numerical approximation must therefore be incorporated into the residual evaluation.

In the present work, a history-dependent discrete approximation is employed to evaluate the Caputo fractional derivative in a stable and implementable manner. Let the time interval $[0,T]$ be partitioned into N_t subintervals with nodes $0 = t_0 < t_1 < \cdots < t_{N_t} = T$, and let Δt denote the time step size. Then, at time t_n , the Caputo derivative can be approximated by:

$${}_0^C D_t^\alpha u_\theta(x,t_n) \approx \frac{1}{\Gamma(2-\alpha)}(\Delta t)^{-\alpha} \sum_{k=0}^{n-1} b_k^{(\alpha)} (u_\theta(x,t_{n-k}) - u_\theta(x,t_{n-k-1})) \tag{18}$$

where $b_k^{(\alpha)} = (k+1)^{1-\alpha} - k^{1-\alpha}$ are the corresponding discrete weights. This approximation essentially reflects the weighted accumulation of all historical state increments in the Caputo derivative, and therefore captures the long-memory effect that is central to anomalous diffusion.

For the variable-order case, the fixed order α in the above weights can be replaced by the locally predicted fractional-order parameter $\alpha_\psi(x)$, yielding a spatially dependent approximation of the fractional derivative. While this extension increases the expressive power of the model, it also leads to higher computational cost and greater training difficulty. To control the complexity, the fractional residual is evaluated only at selected temporal and physical collocation points during training, rather than through a full discretization over the entire continuous domain. In implementation, vectorized computation or batched temporal-window strategies may also be adopted to reduce the storage and computational overhead caused by the accumulation of history terms. In this way, the numerical approximation of the fractional derivative is effectively coupled with the neural-network-based function representation, providing the basis for subsequent state reconstruction and parameter identification.

Training Algorithm

Based on the state network, parameter network, and physics-informed loss function described above, the anomalous diffusion modeling problem in heterogeneous media is formulated as a joint optimization problem. The goal of training is to determine the optimal network parameters θ , ϕ , and, when necessary, ψ , by minimizing the total loss function $\mathcal{L}$. To this end, gradient-based optimization is employed, and the network parameters are iteratively updated through backpropagation so as to simultaneously learn the diffusion state field and the unknown physical parameters.

The training procedure is summarized as follows. First, observational data points, collocation points for the governing equation, initial-condition points, and boundary-condition points are generated over the spatiotemporal domain, forming the corresponding training datasets. Second, the parameters of the state and parameter networks are randomly initialized, and forward propagation is performed to obtain the network outputs. Third, based on the predicted state and inferred parameters, the data loss, physics residual loss, initial-condition loss, and boundary-condition loss are computed and combined into the total loss function. The gradients of the loss with respect to the network parameters are then evaluated using the combination of automatic differentiation and the numerical approximation of the fractional derivative, and the network parameters are updated using an optimizer. This process is repeated until the loss converges or a prescribed number of training iterations is reached.

To improve training stability and convergence, a staged optimization strategy is generally adopted. Specifically, the Adam optimizer may first be used for pretraining in order to quickly obtain a reasonable initialization, after which a quasi-Newton method such as L-BFGS may be employed for fine-tuning to further improve solution accuracy. In parameter identification problems, to avoid imbalance among different network branches, different learning rates may be assigned to the state network and the parameter network, or the state reconstruction task may be emphasized during the early stage of training before gradually increasing the weight of the parameter inversion term. In addition, when sparse or noisy data are considered, the loss weights may be dynamically adjusted during training so that the physics-based term plays a stronger regularizing role.

Overall, the proposed training algorithm is not merely a data-fitting procedure, but rather a joint learning process that integrates a fractional-order mechanistic model, continuous function representation, and deep optimization strategy. Through this framework, the neural network is able to fully exploit the physical information embedded in the governing equation under limited observations, thereby enabling efficient modeling and stable inversion of anomalous diffusion processes in complex heterogeneous media.

NUMERICAL EXPERIMENTS

To validate the effectiveness of the proposed method for anomalous diffusion modeling in heterogeneous media, a series of representative numerical experiments are conducted. The performance of the framework is evaluated from three perspectives: reconstruction of the diffusion field, identification of heterogeneous parameters, and robustness under sparse and noisy observations. All experiments are based on the time-fractional diffusion model introduced in Section FRACTIONAL DIFFUSION MODEL FOR HETEROGENEOUS MEDIA and the coupled state-network/parameter-network framework developed in Section NEURAL-NETWORK-BASED SOLUTION FRAMEWORK. Unless otherwise stated, the spatial domain is taken as $\Omega = [0, 1]$, the time interval is $t \in [0, 1]$, and Dirichlet boundary conditions together with a prescribed initial condition are used throughout the tests.

Benchmark Setup

To mimic spatial heterogeneity in complex media, a piecewise diffusion coefficient is adopted as the benchmark configuration, as shown in Figure 1. Specifically, the spatial domain is divided into three subregions with different diffusion strengths, thereby representing local variations in transport behavior in a layered heterogeneous medium. In addition, a set of discrete observation points is distributed over the space-time domain to provide limited state measurements. This setup retains the essential characteristics of heterogeneous media while allowing us to examine the ability of the proposed method to reconstruct both parameter jumps near interfaces and the spatiotemporal evolution of the diffusion field.

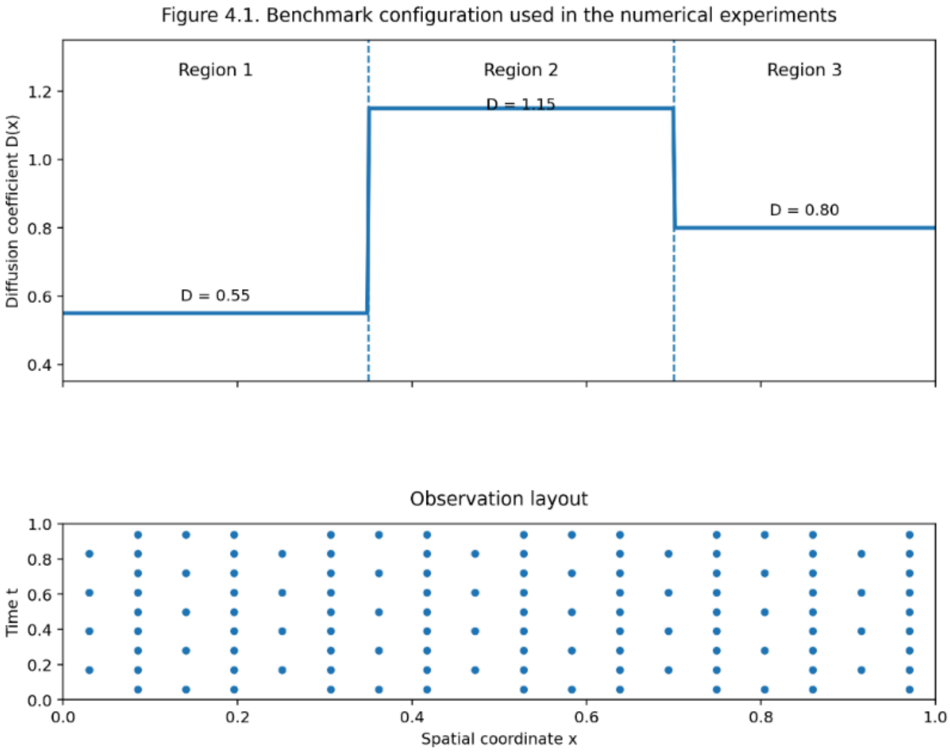

Figure 1. Benchmark configuration used in the numerical experiments

Table 1 summarizes the four benchmark cases considered in this study. Case 1 is designed to verify the basic feasibility of the method in a relatively simple setting. Case 2 further introduces a spatially heterogeneous diffusion coefficient. Case 3 considers the simultaneous identification of the heterogeneous diffusion coefficient and the fractional-order parameter. Case 4 incorporates both measurement noise and a lower observation density in order to assess the robustness of the framework. By organizing the experiments in this progressive manner, the behavior of the proposed method can be systematically examined under scenarios of increasing difficulty.

Table 1. Benchmark configurations for the numerical experiments

Case	Medium type	Unknown quantities	Fractional order	Observation ratio	Noise level
Case 1	Homogeneous	(u(x,t)), constant (D)	fixed (α=0.78)	50%	0%
Case 2	Heterogeneous	(u(x,t)), (D(x))	fixed (α=0.78)	50%	0%
Case 3	Heterogeneous	(u(x,t)), (D(x)), α	unknown constant α	50%	0%
Case 4	Heterogeneous	(u(x,t)), (D(x)), α	unknown constant α	35%	10%

As can be seen from Figure 1, the diffusion coefficient exhibits clear jumps near $x = 0.35$ and $x = 0.70$, meaning that the model must not only learn the diffusion behavior in smooth regions but also handle sharp changes near heterogeneous interfaces. Moreover, the observations are not continuously distributed in time, making the problem closer to realistic sparse-sampling scenarios. This benchmark therefore provides a meaningful test bed for evaluating the proposed framework in anomalous diffusion modeling for complex heterogeneous media.

Reconstruction of Anomalous Diffusion Fields

Figure 2 presents the reference solution, the neural-network prediction, and the corresponding absolute error distribution for the heterogeneous-medium case. Overall, the proposed method is able to accurately recover the spatiotemporal evolution of the diffusion field, and the predicted solution agrees well with the reference field in terms of the dominant transport structure. In particular, even when diffusion-front attenuation over time interacts with spatial heterogeneity, the state network still provides a stable approximation of the true solution, indicating that the physics-based constraint effectively compensates for the limitations of purely data-driven fitting under limited observations.

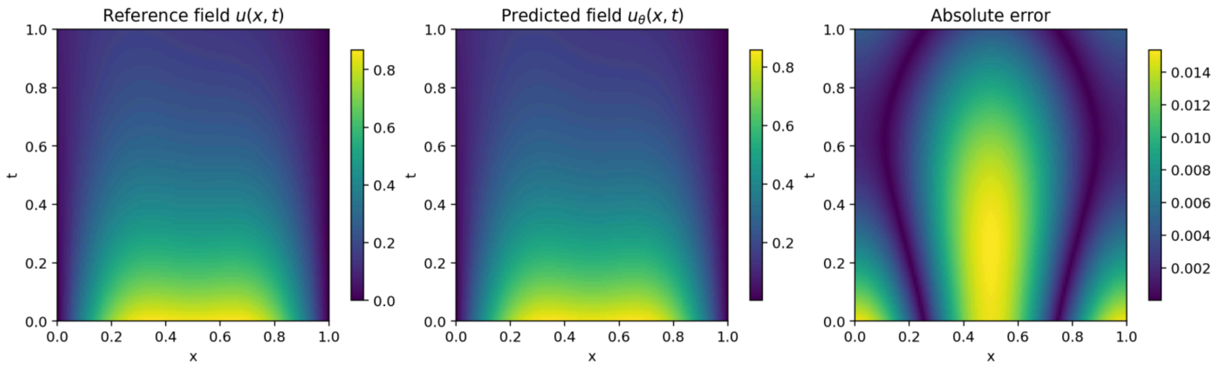

Figure 2. Reconstruction of anomalous diffusion fields in heterogeneous media

A closer inspection of the error field shows that relatively larger errors are concentrated near the heterogeneous interfaces and in weak-signal regions at later times. This behavior has both physical and numerical origins. On the one hand, jumps in the diffusion coefficient lead to stronger local variations in flux near interfaces, which increases the difficulty of state reconstruction. On the other hand, as the solution amplitude decreases over time, the influence of observational noise and approximation errors in the fractional-history term becomes relatively more pronounced. Nevertheless, the overall error remains low, demonstrating that the proposed framework achieves satisfactory accuracy and stability in reconstructing anomalous diffusion fields in heterogeneous environments.

To quantify the model performance more clearly, Table 2 reports the relative L^2 error of the reconstructed field, the relative error of the identified diffusion coefficient, and the relative error of the inferred fractional-order parameter under different benchmark settings. It can be observed that even when both the heterogeneous coefficient and the fractional order are unknown, the method still retains good identification capability. Compared with Case 2, the field reconstruction error and parameter error in Case 3 increase slightly, indicating that the simultaneous inversion of multiple unknown quantities makes the problem more challenging; however, the increase remains within an acceptable range.

Table 2. Quantitative results of field reconstruction and parameter identification

Case	Relative L2 error of u	Relative error of D	Relative error of α
Case 1	1.24×10^{-2}	0.71%	—
Case 2	1.61×10^{-2}	2.84%	—
Case 3	1.93×10^{-2}	3.12%	1.28%
Case 4	3.41×10^{-2}	5.67%	2.56%

Taken together, Figure 2 and Table 2 show that the proposed framework not only reconstructs anomalous diffusion fields with high accuracy, but also maintains stable identification of hidden physical parameters from limited observations. These results suggest that embedding the fractional-order mechanistic model into the neural-network training process plays a positive role in improving the identifiability of complex diffusion systems.

Identification of Heterogeneous Parameters

In heterogeneous-media modeling, parameter identification is as important as state reconstruction. The left panel of Figure 3 compares the reference diffusion coefficient $D(x)$ with the identified coefficient obtained by the parameter network. It is evident that the network successfully recovers the diffusion levels in the three subregions and captures the locations of the parameter jumps near $x = 0.35$ and $x = 0.70$. Although some smoothing error remains around the transition zones, the identified profile agrees closely with the reference coefficient overall, indicating that the proposed method can reliably characterize medium heterogeneity.

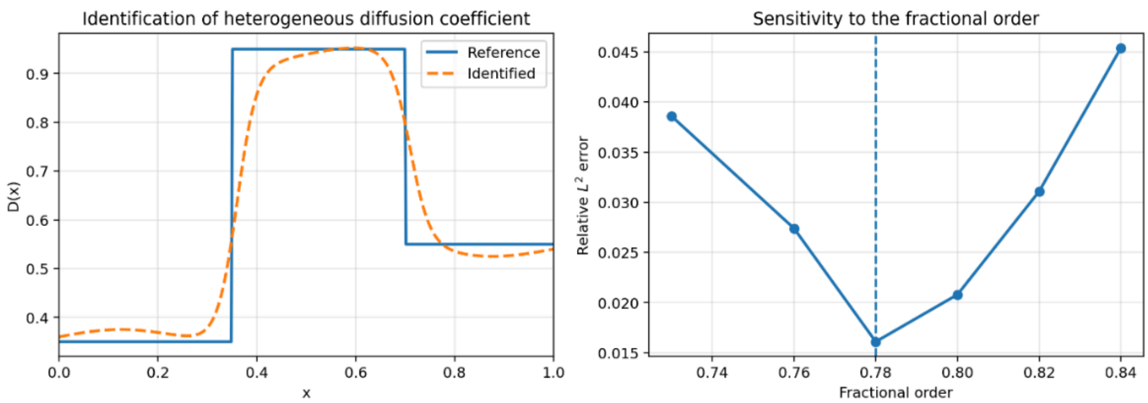

Figure 3. Identification of heterogeneous parameters and sensitivity to the fractional order

The right panel of Figure 3 further presents a sensitivity analysis of the reconstruction error with respect to the fractional order. The relative L^2 error reaches its minimum near $\alpha = 0.78$, while a noticeable increase in error is observed when the fractional order deviates from this value. This result indicates that the fractional-order parameter has a significant influence on the underlying transport dynamics, and that an accurate estimate of α is crucial for improving both state recovery and parameter identification. From another perspective, this sensitivity is physically meaningful: since α governs the power-law exponent of the waiting-time distribution, an incorrect α systematically distorts the predicted memory kernel and thus the entire temporal evolution, confirming that α is not merely a fitting parameter but a physically grounded descriptor of subdiffusive transport.

A further comparison between Case 2 and Case 3 reveals that the identification of the diffusion-coefficient field is more stable when α is known, whereas the simultaneous recovery of α and $D(x)$ becomes more difficult because of the coupling effect between these parameters. This observation is consistent with the discussion of ill-posedness in Section FRACTIONAL DIFFUSION MODEL FOR HETEROGENEOUS MEDIA. Nevertheless, judging from the error levels reported in Table 2, the proposed method still exhibits strong robustness in the multi-parameter inversion setting, suggesting that the separation between the state network and the parameter network is effective in improving parameter identifiability.

Robustness Under Noisy and Sparse Observations

To further assess the practical applicability of the method, we examine its performance under different observation densities and noise levels. Figure 4 shows the relative error as a function of the available observation ratio under noise-free, 5% noise, and 10% noise settings. In general, the reconstruction error decreases steadily as the observation ratio increases, whereas for a fixed observation ratio the error increases with the noise level. This behavior is consistent with the general characteristics of inverse problems and confirms that the completeness and quality of observations directly affect the accuracy of anomalous diffusion modeling.

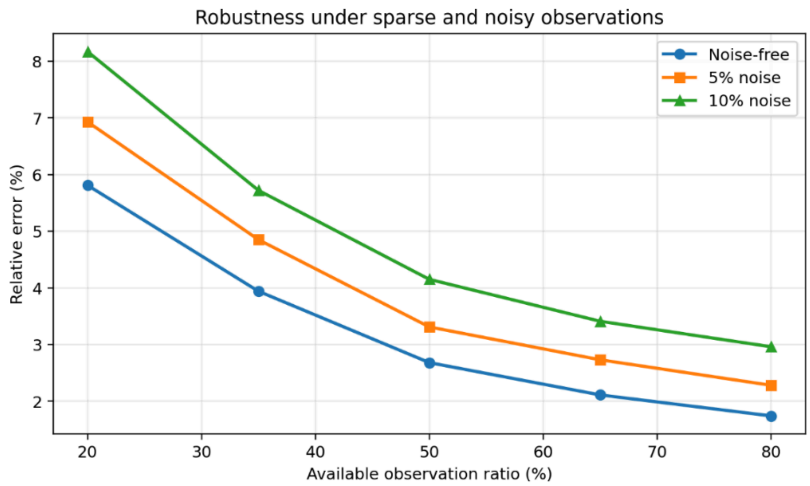

Figure 4. Robustness of the proposed method under sparse and noisy observations

To present this trend more clearly, Table 3 lists the relative reconstruction errors for different observation ratios and noise levels. Even when only 35% of the observations are available and the noise level reaches 10%, the model still limits the relative error to approximately 5.72%. When the observation ratio is increased to 65% or above, the error can be further reduced to around 3% even in the presence of 10% noise. These results indicate that the physics-informed loss provides a clear regularization effect in sparse and noisy settings, preventing the network from relying exclusively on limited data points and enabling it to recover a more stable global solution under the constraint of the governing equation.

Table 3. Relative reconstruction errors (%) under different observation ratios and noise levels

Observation ratio	Noise-free	5% noise	10% noise
20%	5.81	6.93	8.17
35%	3.94	4.85	5.72
50%	2.68	3.31	4.15
65%	2.11	2.73	3.41
80%	1.74	2.28	2.96

Overall, Figure 4 and Table 3 demonstrate that the proposed method maintains good stability under low observation density and moderate noise levels, which is of practical importance for real diffusion problems in complex media. Since experimental measurements are often restricted by limited sampling and noise contamination, purely data-driven models are prone to overfitting or unstable extrapolation. In contrast, the fractional-order, physics-constrained framework adopted here makes fuller use of prior mechanistic information, thereby significantly improving generalization ability and noise robustness. It should be noted that the robustness demonstrated here primarily concerns observation sparsity and noise level; generalization to heterogeneous structures substantially different from the training configuration—such as media with higher contrast ratios, more subregions, or irregular interface geometries—is not exhaustively tested in the present study, and the governing-equation residual serves as the principal mechanism that constrains the solution toward physically admissible states even in such out-of-distribution scenarios.

In summary, the numerical experiments in this section show that the proposed coupled fractional-differential-equation/neural-network framework can effectively reconstruct anomalous diffusion fields and identify heterogeneous parameters in nonuniform media, while preserving strong robustness under sparse and noisy observations. These results provide solid support for the subsequent discussion on the effects of medium heterogeneity, fractional-order variation, and computational complexity.

DISCUSSION

Effect of Heterogeneity

Medium heterogeneity does not merely alter the local distribution of the diffusion process; it also fundamentally affects the identifiability of the model and the difficulty of learning. From a modeling perspective, a spatially varying diffusion coefficient changes the way flux is transferred across different regions, causing the diffusion front to exhibit more complex propagation patterns near heterogeneous interfaces. Compared with homogeneous media, the state evolution in heterogeneous media no longer follows a smooth diffusion process governed by a single scale. Instead, it is typically accompanied by location-dependent propagation speeds, enhanced gradients near interfaces, and region-specific decay rates. Therefore, accurate modeling of anomalous diffusion in heterogeneous media requires not only recovery of the overall transport trend, but also the ability to capture fine-scale dynamical variations induced by local material differences.

From the viewpoint of inverse problems, heterogeneity also substantially increases the uncertainty of parameter identification. On the one hand, the spatial variation of the diffusion coefficient is usually represented by a continuous field or a piecewise profile, whose degrees of freedom are much higher than those of a single constant parameter and are therefore more sensitive to sparse observations and noise contamination. On the other hand, there exists a bidirectional coupling between the state field and the parameter field: state reconstruction depends on the parameter distribution, whereas parameter identification must be achieved indirectly through the state response. This coupling becomes particularly pronounced near heterogeneous interfaces, where local observation errors may be amplified and then transferred into the parameter recovery process. In this sense, heterogeneity does not simply increase the complexity of the problem; rather, it changes the stability structure of the inversion process itself.

A key advantage of the present method in handling heterogeneity lies in the fact that the spatial parameter is not discretized into a large number of independent unknowns, but is instead represented as a continuous function through the parameter network. This representation introduces a certain degree of implicit regularization, making the parameter field more likely to develop a coherent global structure rather than performing purely pointwise fitting at local observations. For this reason, the proposed framework is able to maintain satisfactory parameter recovery even in the presence of interfaces and limited measurements. At the same time, however, this continuous representation also has an intrinsic limitation: when the true parameter field contains very sharp discontinuities, the neural network tends to exhibit some degree of smoothing bias. Therefore, for strongly heterogeneous problems with complex discontinuous structures, an important topic for future work is how to balance the expressive power of continuous neural representations with accurate interface resolution. More broadly, a key distinction of physics-informed frameworks over purely data-driven models is that the governing equation residual provides a medium-independent constraint: regardless of the specific heterogeneity pattern, any predicted field that satisfies the fractional diffusion equation and the given boundary conditions is physically consistent. This structural prior offers a degree of generalization beyond the training distribution that pure regression cannot provide, although its practical extent depends on the quality and coverage of the collocation points.

Effect of Fractional Order

The fractional order is the key parameter that characterizes the memory effect in anomalous diffusion, and its variation directly determines the extent to which the system departs from classical diffusion behavior. In contrast to integer-order diffusion models, time-fractional models introduce nonlocal coupling across time scales through a history-dependent integral term, so that the current state of the system depends not only on the instantaneous local gradient but also on the accumulated influence of past evolution. Therefore, the fractional order should not be regarded as a purely mathematical quantity; it can also be interpreted as an effective descriptor of the memory strength of the transport system. A smaller fractional order typically corresponds to stronger trapping effects and slower propagation, whereas as the order approaches unity, the system gradually returns to a dynamics dominated by classical diffusion.

Within the modeling framework considered in this study, the fractional order has a pronounced impact on both state reconstruction and parameter identification. First, if the prescribed or inferred order deviates from its true value, the model may still produce systematic errors even when the diffusion coefficient and the network architecture are chosen appropriately, because the underlying memory kernel becomes mismatched. This implies that the fractional-order parameter is not merely a minor correction term, but rather a key control variable that determines the temporal structure of the diffusion process. Second, when the fractional order and the diffusion coefficient are both unknown, a compensatory effect may arise between them, in the sense that some dynamic changes caused by memory effects may be partially absorbed by adjustments in the diffusion-coefficient field, and vice versa. This is one of the main reasons why inverse problems involving fractional-order models are generally more difficult than their classical counterparts.

It should also be emphasized that, although the introduction of the fractional-order parameter enhances the ability of the model to explain complex transport phenomena, it simultaneously increases the demand for careful model selection and physical interpretation. In practical applications, if the available observations are insufficient to support stable identification of the fractional order, introducing too much variable-order flexibility may actually reduce model reliability. From a modeling standpoint, fixed-order and variable-order formulations should therefore be selected according to both the complexity of the physical problem and the quality of the available data. For systems in which the overall memory mechanism is relatively uniform, a fixed-order model may already provide a sufficiently robust description. By contrast, for highly heterogeneous media with pronounced local differences in transport behavior and adequate data support, a variable-order model may offer clear advantages. In other words, the fractional-order parameter is both a major source of model expressiveness and a key factor governing the difficulty and stability of the inverse problem. From an identifiability standpoint, the coupling between $D(x)$ and α can be partially resolved by exploiting their distinct signatures in the observed dynamics: $D(x)$ primarily governs spatial flux redistribution and thus affects the spatial profile of the solution, whereas α controls the temporal scaling of the memory kernel and thus affects the rate of temporal decay; leveraging observations at multiple time levels therefore provides complementary constraints that help disambiguate the two parameter fields.

Limitations and Computational Cost

Although the proposed coupled fractional-differential-equation/neural-network framework demonstrates good accuracy and robustness in anomalous diffusion modeling for heterogeneous media, its limitations should be discussed explicitly. First, the current framework is mainly developed for relatively regular spatial domains and idealized boundary conditions, which makes the formulation and numerical implementation more transparent but also implies that its direct application to complex geometries, irregular boundaries, and realistic multiscale media still requires further extension. In particular, under two- or three-dimensional complex structures, the representation of spatial parameter fields, the sampling of collocation points, and the treatment of interfaces would all become significantly more difficult.

Second, the intrinsic history dependence of fractional derivatives makes the training process substantially more expensive than that of standard PINN-type frameworks. For time-fractional models, the residual at each time level depends on information from multiple preceding time steps, which means that both computational cost and memory usage increase rapidly as the temporal resolution becomes finer. Although this issue is partially alleviated in the present work through residual evaluation at selected collocation points, vectorized implementation, and staged optimization, the fractional-history term remains the principal computational bottleneck during training. This challenge becomes even more severe for higher-dimensional problems, finer temporal discretizations, or variable-order formulations, and therefore limits the scalability of the method. To place this cost in perspective, a conventional finite difference method (FDM) applied to the same one-dimensional fractional diffusion problem requires O(N_t^2 N_x) operations due to the history accumulation of the Caputo derivative, with computational time scaling primarily with grid resolution. The proposed neural-network framework incurs a higher per-iteration cost owing to backpropagation through the history term, but offers a mesh-free continuous solution that simultaneously reconstructs the state field and identifies spatially varying parameters—a capability that would require a separate, iterative inversion procedure under FDM. In the forward-only setting, traditional solvers are generally faster; the cost-benefit advantage of the present framework lies in the inverse problem setting, where joint state-parameter learning under sparse observations is achieved within a single optimization, avoiding the repeated forward solves that classical adjoint-based inversion methods demand.

Third, the proposed framework remains, in essence, a physically constrained nonconvex optimization problem. As a result, the final training performance is inevitably affected by factors such as network initialization, the choice of loss weights, the distribution of sampling points, and the configuration of the optimizer. Although the numerical results indicate satisfactory overall stability, local minima, slow convergence, or imbalance among competing loss components may still occur in strongly heterogeneous, noisy, or multi-parameter coupled scenarios. Consequently, the development of more robust adaptive weighting strategies, more efficient approximations for fractional residuals, and network architectures better tailored to inverse problems remains an important direction for future study.

Overall, the present method is particularly suitable for low- to moderate-dimensional diffusion problems with clear mechanistic structure and limited observations, and may be viewed as an intermediate approach between traditional numerical solvers and purely data-driven models. Promising future directions include extending the framework to complex geometries, incorporating multi-fidelity data, developing efficient approximation strategies for variable-order and high-dimensional problems, and validating the method on experimental or real engineering datasets. Progress along these lines will be essential if fractional-order physics-informed neural approaches are to be more broadly applied to realistic anomalous transport modeling and inversion in heterogeneous media.

CONCLUSION

This study developed a mechanistically constrained, data-driven framework for anomalous diffusion modeling in heterogeneous media by integrating fractional diffusion equations with neural networks. Unlike approaches based solely on conventional numerical solvers or purely data-driven fitting, the proposed framework incorporates a fractional-order dynamical model into the learning process as a form of physical prior, thereby enabling the unified modeling of both the diffusion state and key medium parameters. In this way, the method preserves the ability of fractional models to represent memory effects and nonlocal transport, while simultaneously exploiting the strengths of neural networks in continuous function approximation and nonlinear representation learning. The resulting framework therefore provides an interpretable and flexible strategy for the modeling and inversion of complex diffusion processes.

From the methodological perspective, a collaborative architecture consisting of a state network and a parameter network was constructed so that diffusion-field reconstruction and heterogeneous-parameter identification could be addressed within a single optimization framework. By combining observational data, governing-equation residuals, and initial-boundary constraints into a unified physics-informed loss function, the proposed approach allows limited measurements to be used more effectively under mechanistic guidance, thus improving both stability and physical consistency. The numerical experiments showed that the framework can accurately reconstruct anomalous diffusion fields under a variety of heterogeneous settings and can reliably identify spatially varying diffusion coefficients as well as fractional-order parameters. While the training cost is higher than that of a direct fractional finite difference solver for the forward problem alone, the key advantage lies in the inverse setting: the proposed framework accomplishes joint state reconstruction and parameter identification within a single optimization pass, whereas traditional approaches would require iterative coupling of a forward solver with a separate inversion scheme. In addition, the method maintained satisfactory robustness under sparse and noisy observations, suggesting that it has meaningful potential for more realistic applications.

It should be emphasized that the present work represents a relatively fundamental yet representative step toward the integration of fractional-order mechanistic modeling and neural-network-based learning. The main purpose here was to verify the feasibility and effectiveness of this coupled strategy rather than to exhaust all possible extensions. Several directions deserve further investigation, including the extension to more complex geometries and higher-dimensional settings, the incorporation of variable-order or coupled spatiotemporal fractional mechanisms to enhance medium characterization, the development of more efficient approximations and training strategies for history-dependent terms, and the validation of the framework using experimental or real engineering data. Overall, the results of this study indicate that the integration of fractional differential equations and neural networks offers a promising route for the quantitative analysis of anomalous diffusion in heterogeneous media, and lays a useful foundation for future research on mechanistically informed learning for transport processes in more complex systems.

Author Contributions

Conceptualization – Yu Z; methodology – Yu Z; formal analysis – Yu Z; investigation – Yu Z; resources – Yu Z; writing-original draft preparation – Yu Z; writing-review and editing – Yu Z; visualization – Yu Z; supervision – Yu Z. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The author declares no conflict of interest.

Funding

This research received no external funding.

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

Not applicable.

REFERENCES

[1] GLIMM J, SHARP D H. A random field model for anomalous diffusion in heterogeneous porous media. Journal of statistical Physics. 1991; 62(1): 415-424. doi: 10.1007/bf01020877

[2] WAIGH T A, KORABEL N. Heterogeneous anomalous transport in cellular and molecular biology. Reports on Progress in Physics. 2023; 86(12): 126601. doi: 10.1088/1361-6633/ad058f

[3] SRINIVASAN H, SHARMA V K, MITRA S. Breaking the Brownian barrier: models and manifestations of molecular diffusion in complex fluids. Physical Chemistry Chemical Physics. 2024; 26(47): 29227-29250. doi: 10.1039/d4cp01813c

[4] JIA Y, WANG D, GUO X. High-dimensional numerical methods for nonlocal models. Mathematics. 2025; 13(21): 3512. doi: 10.3390/math13213512

[5] CHANDEL N, KHALSA L, VARGHESE V. Non-simple thermoelastic diffusion interaction in a half-space with nonlocality and memory effect. Journal of Thermal Stresses. 2025; 48(3): 292-318. doi: 10.1080/01495739.2024.2449067

[6] JEAN G. Generalized Fractional Differential Equations in Anomalous Diffusion and Stochastic Processes. 2025. doi: 10.1016/j.cnsns.2025.108983

[7] GRANELLA S J. Anomalous Diffusion in Drying: A Fractional Calculus Perspective on Modeling and Applications in Agro-Food Systems. Food Physics. 2026: 100088. doi: 10.1016/j.foodp.2026.100088

[8] SUZUKI J L, GULIAN M, ZAYERNOURI M, et al. Fractional modeling in action: A survey of nonlocal models for subsurface transport, turbulent flows, and anomalous materials. Journal of Peridynamics and Nonlocal modeling. 2023; 5(3): 392-459. doi: 10.1007/s42102-022-00085-2

[9] WENSHU Z, DAOLUN L, LUHANG S, et al. Review of neural network-based methods for solving partial differential equations. Chinese Journal of Theoretical and Applied Mechanics. 2022; 54(3): 543-556. doi: 10.6052/0459-1879-21-617

[10] HUANG S, FENG W, TANG C, et al. Partial differential equations meet deep neural networks: A survey. IEEE Transactions on Neural Networks and Learning Systems. 2025. doi: 10.1109/TNNLS.2025.3545967

[11] ZHANG T, ZHANG D, SHI S, et al. Spectral-fPINNs: spectral method based fractional physics-informed neural networks for solving fractional partial differential equations. Nonlinear Dynamics. 2025; 113(11): 12565-12588. doi: 10.1007/s11071-025-11067-6

[12] XU Y, LIU Y, ZHANG Y, et al. fPICRNNs: fractional physics-informed convolutional-recurrent neural networks for solving fractional order partial differential equations. Neurocomputing. 2025: 132433. doi: 10.1016/j.neucom.2025.132433

[13] YAN X, XU Z J, MA Z. Laplace-fPINNs: Laplace-based fractional physics-informed neural networks for solving forward and inverse problems of subdiffusion. arXiv preprint arXiv:2304.00909, 2023. doi: 10.1063/5.0278224

[14] HE S, WANG H, SUN K. Solutions and memory effect of fractional-order chaotic system: A review. Chinese Physics B. 2022; 31(6): 60501. doi: 10.1088/1674-1056/ac43ae

[15] THAKUR S, MITRA H, ARDEKANI A M. Physics-informed neural network-based inverse framework for time-fractional differential equations for rheology. Biology. 2025; 14(7): 779. doi: 10.3390/biology14070779