Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/150513

An Adaptive Error-Controlled LADM–ANN Framework for Nonlinear Fractional Partial Integro-Differential Equations with Weakly Singular Memory Kernels

Amin M. Elfeky[1], *Essa Melfi Aldosari*[2], *Mohamed Sayed Abdellatif*[3], *Mohammad A. Tashtoush*[4,5], *Hasanain J. Alsaedi*[6] *and Adel S. Hussain*[7,8,9,*]

[1] Department of Mathematics, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[2] Department of Curriculum and Instruction, College of Education in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[3] Department of Psychology, College of Education, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia
[4] Department of Basic Sciences, AL-Huson University College, AL-Balqa Applied University, Salt, Jordan
[5] Faculty of Education and Arts, Sohar University, Sohar, Oman
[6] Business Information Technology Department, College of Business Information, University Information Technology and Communication, Baghdad, Iraq
[7] IT Department, Amedi Technical Institutes, University of Duhok Polytechnic, Duhok, Iraq
[8] Administrative & Financial Affairs Department, Al-Iraqia University, Baghdad, Iraq
[9] Department of Computer Engineering, Al-Kitab University, Altun, Kupri, Iraq

Received: 2 Jun. 2026, Revised: 15 Aug. 2026, Accepted: 20 Aug. 2026
Published online: 1 Sep. 2026

Abstract: Fractional partial integro-differential equations (FPIDEs) with weakly singular kernels provide mathematical models for systems exhibiting nonlocality and memory. However, conventional decomposition methods often lack adaptive truncation, problem-specific error control, and reliable long-time continuation. This study develops an adaptive, error-controlled hybrid Laplace-Adomian decomposition and artificial neural network (LADM-ANN) framework combining analytical approximation with physics-guided residual learning. The methodology incorporates model-specific convergence conditions, truncation-error estimates, a dual stopping criterion, and memory-preserving Caputo continuation. Unknown fractional and physical parameters are calibrated through constrained residual minimization, while block-bootstrap inference, sensitivity analysis, and conformal prediction assess uncertainty. Manufactured solution experiments demonstrate substantial numerical improvements, with hybrid RMSE values of 3.396×10^{-4} and 1.402×10^{-5}, representing reductions of 83.9% and 89.7%, respectively, relative to adaptive LADM. Real-data validation employs 999,061 filtered NGSIM US-101 trajectory observations aggregated onto a time–space grid. At the 5-s forecasting horizon, the hybrid model achieves an RMSE of 6.149 vehicles/km/lane, outperforming adaptive LADM (6.630) and persistence (6.753). However, at 25 s, its RMSE increases to 22.156, exceeding persistence (20.924). Although the hybrid remains more accurate than persistence at 50–100 s, histogram gradient boosting performs better at these longer horizons. Furthermore, wide bootstrap intervals indicate weak practical parameter identifiability, while conformal under coverage reveals limitations under nonstationary traffic conditions. These findings establish the proposed framework as an interpretable and error-controlled approach with demonstrated short-horizon predictive benefits, while emphasizing that adaptive numerical accuracy does not guarantee superior long-horizon forecasting performance.

Keywords: Fractional integro-differential equations, weakly singular kernels, Laplace-Adomian decomposition, residual neural networks, inverse problems, traffic flow modeling.

1 Introduction

Fractional calculus (FC), which extends the concepts of differentiation and integration to orders that are not integers, has recently received significant interest for use in a wide array of application areas including engineering and the applied sciences. This is largely due to its potential to provide better representations than traditional calculus for a number of

* Corresponding author e-mail: adel.sufyan@dpu.edu.krd

© 2026 NSP
Natural Sciences Publishing Cor.

real-world systems. As an example, there have been numerous scientific and technological events over the years that are particularly well described with the use of fractional differential equations. Many natural systems, such as those involving damping, relaxation, acoustic absorption, non-Fourier heat flow, geophysics, fluid flow, creep, viscoelasticity/rheology, COVID-19/fractional epidemiological model for malaria [1-4] etc., may be represented more effectively through the use of fractional derivatives. Additionally, some physical phenomena including robotics, genetics/evolutionary algorithms, and economics/finance have also been modeled more effectively with the aid of FC [5-8]. A number of additional physical processes occurring in nature have been modeled more realistically with the use of FC rather than traditional calculus. Applications of FC in the literature include earth quake nonlinear oscillations, time-fractional nonlinear waves on the surface of water [9], thermal transfer model [10], chaos theory, cancer chemotherapy treatment, zener models, airfoils, fluid traffic, Poisson-Nerst Planck diffusion, tuberculosis, hepatitis B virus infection/hepatitis B disease model/electrodynamics/pine wilt disease/diabetes/endogenous insulin response and other applications in a wide range of disciplines [5-8].

The idea behind the fractional partial integro-differential equation (FPIDE), which is based on fractional definitions, has been one of the core concepts of Fractional Calculus (FC) and its applications. This type of approach has been developed as an essential mathematical method to describe and analyze a variety of problems encountered in various fields of natural sciences and technologies; namely: engineering. Many different mathematical representations of real-world physical phenomena have been studied using FPIDEs such as: heat conduction, flow in cracked biomaterials, reaction dynamics, viscoelasticity, electricity swaption, convection-diffusion, grain development and population dynamics. We will develop a nonlinear version of the mathematical representation [15], [14]:

$$D_\tau^\alpha v(\zeta,\tau) + v(\zeta,\tau)\frac{\partial}{\partial y}v(\zeta,\tau) = \int_0^\tau (\tau-q)^{\alpha-1}\frac{\partial^2}{\partial y^2}v(\zeta,\tau)dq + f(\zeta,\tau)\ \alpha \in (0,1] \tag{1}$$

where D_τ^α is the fractional derivative in Caputo sense and $f(\zeta,\tau)$ is the source term, with initial and boundary conditions,

$$v(\zeta,0) = k(\zeta),\ v(0,\tau) = v(L,\tau) = 0. \tag{2}$$

Fractional Partial Integro-Differential Equations (FPIDEs) provide a flexible way to describe Nonlinear Systems with Hereditary Effects, Anomalous Transport, and Non-local Interactions. Unlike traditional Differential Equations (DE), FPIDEs account for Memory-Dependent Dynamics by incorporating No Integer-Order Derivatives or Integral Operators, allowing one to model Processes That Are Present-State Dependent Upon Their Temporal Histories. However, The Combination of Non-linearities, Caputo Fractional Derivatives, And Weakly Singular Volterra Kernels Creates Substantial Analytical & Computational Challenges When Accurate Solutions Are Required Over Extended Time Intervals [16 – 34].

There are many analytical and numerical ways that can be applied to solve fractional differential and integro-differential equations. The fractional decomposition technique is a way to decompose an equation into smaller fractions that may be solved separately; iterative methods are another. Numerical schemes such as quadrature could also be used. Several researchers investigated the conditions for which a solution exists and/or is unique for nonlinear fractional integro-differential equations [35]; in addition, several studies have discussed various multi-stage decomposition methods [36] and utilized Adomian-based methods to solve non-linear differential systems [37-41]. One method among those listed above is the Laplace – Adomian Decomposition Method (LADM). This utilizes both the Laplace Transform and Adomian Polynomial Expansions to develop approximate solutions for the given system of nonlinear partial differential equations without linearizing the nonlinear terms of the problem.

Despite its computational advantages, conventional LADM implementations face several limitations. The decomposition order is frequently prescribed in advance, potentially producing unnecessary computational effort or inadequate accuracy. Furthermore, convergence arguments are not always accompanied by explicit model-specific contraction conditions or computable truncation-error bounds. Long-time continuation is particularly challenging for fractional systems because restarting the solution procedure without retaining the complete Caputo history may alter the underlying memory dynamics. These difficulties become more pronounced when model parameters must be estimated from noisy observational data. Recent developments in scientific machine learning provide complementary strategies for approximating fractional systems. Physics-informed neural networks (PINNs) incorporate governing-equation residuals into neural-network training [57], while fractional PINNs extend this principle to nonlocal fractional operators and inverse problems [58]. Deep Operator Networks (DeepO Nets) and Fourier Neural Operators (FNOs) learn mappings between function spaces and can facilitate repeated solution evaluations across varying inputs [59,64]. More recent contributions have advanced adaptive-loss fractional PINNs for FPIDEs [65], Monte Carlo-based treatments of fractional operators [66], gradient-enhanced fractional learning for nonlinear Burgers-type equations [67], inverse fractional parameter estimation [68], and spectrally adapted fractional PINNs [69]. Physics-informed DeepO Net formulations have also been investigated for multi-term time-fractional diffusion-wave equations [70]. These studies demonstrate the

© 2026 NSP
Natural Sciences Publishing Cor.

potential of learning-based fractional solvers, although their numerical objectives and error-control mechanisms differ from those of decomposition-based methods. An important opportunity is therefore to combine an interpretable physics-based numerical approximation with data-driven correction while retaining explicit control over the decomposition process. Rather than replacing the governing FPIDE with a fully learned solution representation, the present study develops an adaptive hybrid LADM–ANN framework in which LADM constructs the primary approximation and an artificial neural network learns the unresolved residual discrepancy.

The proposed approach integrates model-specific convergence analysis, adaptive decomposition-order selection, residual-based acceptance criteria, and temporal continuation that preserves fractional memory. It also incorporates constrained neural correction and statistical procedures for assessing parameter uncertainty, predictive sensitivity, and forecasting reliability. The resulting framework is evaluated using manufactured-solution benchmarks and an observational traffic-density application based on the NGSIM US-101 dataset.

Recent developments in physics-informed scientific machine learning have substantially expanded the computational treatment of fractional differential equations and nonlocal memory operators. [65] proposed adaptive-loss-weighting auxiliary-output fractional physics-informed neural networks (AWAO-FPINNs) for fractional partial integro-differential equations, combining numerical approximation of Caputo derivatives with adaptive neural-network training. [66] investigated Monte Carlo-based fractional physics-informed neural networks for fractional and tempered fractional PDEs, addressing high-dimensional nonlocal operators through improved quadrature strategies. For nonlinear fractional transport problem, [67] developed gradient-enhanced FPINNs for multiterm time-fractional Burgers-type equations, incorporating residual-gradient information into the learning objective.

Recent investigations have also addressed inverse problems and the computational treatment of fractional operators. [68] studied parameter identification in nonlinear time-fractional diffusion equations using fractional physics-informed neural networks, whereas [69] introduced fractional spectrally adapted PINNs to improve the treatment of fractional PDEs on unbounded domains. These contributions demonstrate the growing applicability of physics-informed learning to forward and inverse fractional problems, while also highlighting persistent challenges involving nonlocal operator evaluation, optimization sensitivity, and computational efficiency. In parallel, neural operator learning has emerged as an alternative paradigm for approximating mappings between input functions and solution fields. [70] developed a physics-informed Deep ONet formulation for multi-term time-fractional mixed diffusion-wave equations, extending operator-learning techniques to fractional dynamics. Unlike conventional single-instance neural solvers, neural operators aim to approximate families of solution mappings and can therefore offer advantages when repeated evaluations are required across different forcing functions, initial conditions, or model parameters.

Despite these developments, an important methodological distinction remains between fully learned fractional solution representations and hybrid approaches that preserve an analytical numerical backbone. The proposed framework integrates adaptive LADM with a physically constrained neural correction, as formally described in Section 3.4, while emphasizing convergence control, adaptive truncation, fractional-memory preservation, and uncertainty-aware calibration. The Laplace–Adomian decomposition provides the primary physics-based approximation, while the neural component learns a physics-dependent correction associated with the remaining discrepancy. The correction is influenced by the governing fractional dynamics and numerical approximation errors, although the ANN is not trained as an independent solver for the complete FPIDE. Its principal emphasis is on adaptive truncation, model-specific convergence conditions, computable truncation-error estimates, preservation of fractional history, and uncertainty-aware calibration. The present approach should therefore be regarded as complementary to recent fractional PINN and neural operator methods rather than as a demonstrated replacement for them. Establishing their comparative advantages under identical numerical conditions remains an important subject for future investigation.

Recent advances in numerical methods for fractional differential equations provide important methodological context for the proposed hybrid LADM–ANN framework. Zhang et al. [59] developed adaptive-loss-weighting auxiliary-output fractional physics-informed neural networks for fractional partial integro-differential equations, demonstrating the potential of adaptive neural training for problems involving fractional derivatives and nonlocal operators. Hu et al. (2024) investigated physics-informed neural methods for fractional and tempered fractional partial differential equations, addressing computational challenges associated with nonlocal operators and high-dimensional problems. From the perspective of analytical and semi-analytical approximation, Ramadan et al. (2025) proposed an integrated Laplace-transform and accelerated Adomian decomposition method for nonlinear time-fractional partial differential equations. Their work highlights the continuing relevance of decomposition-based approaches for constructing approximate solutions to fractional models. In a complementary direction, Wu et al. (2025) introduced fractional spectrally adapted physics-informed neural networks for fractional PDEs, while Khazeem and Mohammed (2025) investigated fractional physics-informed neural networks for delay integro-differential-algebraic equations. Furthermore, Lu et al. (2025) developed a physics-informed DeepO Net framework for multi-term time-fractional mixed diffusion-wave equations, illustrating the potential of neural operator learning for families of fractional problems.

© 2026 NSP
Natural Sciences Publishing Cor.

1.1 Research Gap and Contributions

The principal research gap concerns the lack of a unified LADM-based framework that combines mathematically justified convergence conditions, adaptive truncation control, fractional-memory-preserving continuation, and neural residual correction within a reproducible numerical procedure. Existing decomposition-based approaches do not always provide explicit acceptance criteria or distinguish deterministic truncation error from discretization error, parameter uncertainty, and data-driven approximation error. These distinctions are especially important for real-data applications involving uncertain fractional parameters and nonstationary observations. To address these challenges, the present study makes the following contributions:

1. **Model-specific convergence analysis:** Sufficient conditions for existence and uniqueness of the semi-discrete FPIDE solution are established through a contraction-mapping argument. Under an additional analytic-majorant condition, convergence and truncation-error estimates are obtained for the LADM series.
2. **Adaptive error-controlled decomposition:** A dual acceptance strategy combines a decomposition-tail estimate with the governing-equation residual to select the number of LADM components according to prescribed numerical tolerances.
3. **Fractional-memory-preserving continuation:** Temporal continuation retains the historical contributions required by the Caputo derivative and the weakly singular memory operator, avoiding artificial memory resets between successive computational intervals.
4. **Hybrid neural residual correction:** A constrained ANN approximates the discrepancy remaining after the physics-based LADM solution while preserving the prescribed initial and boundary conditions through the adopted correction formulation.
5. **Numerical and observational validation:** Manufactured-solution experiments assess approximation accuracy and adaptive behavior, while the NGSIM US-101 application examines parameter estimation and forecasting performance under noisy, nonstationary conditions.
6. **Uncertainty and reproducibility assessment:** Block-bootstrap inference, parameter sensitivity analysis, conformal prediction diagnostics, and documented numerical settings are used to investigate practical identifiability, predictive uncertainty, and implementation transparency.

These contributions establish an integrated framework for studying nonlinear fractional systems while maintaining a distinction between theoretical numerical guarantees and empirical predictive performance. The proposed method is intended to complement, rather than universally outperform, existing decomposition methods, fractional PINNs, and neural operator architectures. Its demonstrated numerical advantages are therefore interpreted within the scope of the investigated benchmarks and forecasting horizons.

2 Definitions and preliminaries concepts

In this section, we will discuss some basic definitions.

Definition 2.1. The Caputo operator of fractional order ξ for fractional derivatives is defined as follows [42,43]:

$$D_{\mathfrak{I}}^{\xi}\mu(\zeta,\mathfrak{I}) = \frac{1}{\Gamma(m-\xi)}\int_0^{\mathfrak{I}} (\mathfrak{I}-\mathfrak{I}_0)^{m-\xi-1}\frac{\partial^m \mu(\zeta,\mathfrak{I})}{\partial \mathfrak{I}^m}\, d\mathfrak{I}, \quad \text{if } m-1<\xi\le m \tag{2}$$

if $\xi = m$, for $m \in N$, then

$$D_{\mathfrak{I}}^{\xi}\mu(\zeta,\mathfrak{I}) = \frac{\partial^m \mu(\zeta,\mathfrak{I})}{\partial \mathfrak{I}^m}.$$

Definition 2.2. For a function $g(\tau)$, the LT is given as follows [11,43]:

$$G(s) = \mathscr{L}[g(\tau)] = \int_0^{\infty} e^{-st} g(\tau) d\tau$$

Definition 2.3. The LT of the Caputo operator is defined as follows [42,43]:

$$\mathscr{L}\left(D_{\tau}^{\alpha}\mu(\tau)\right) = s^{a}\mathscr{L}[\mu(\tau)] - \sum_{k=0}^{j-1} s^{\alpha-j-1}U^{k}\left(0^{+}\right),\ j-1<a\le j \tag{3}$$

$$L(D_{\tau}^{\alpha}\mu(\tau)) = s^{\alpha}U(s) - s^{\alpha-1}U(0),\ 0<a\le 1. \tag{4}$$

© 2026 NSP
Natural Sciences Publishing Cor.

Definition 2.4. The Mittag-Leffler function is expressed as follows [44]:

$$E_a(\zeta) = \sum_{j=0}^{\infty} \frac{\zeta^j}{\Gamma(\alpha j+1)} \ \alpha > 0 \zeta \in \mathbb{C}$$

Definition 2.5. The adomian polynomials to express nonlinear term in a given problem is defined as follows [45]:

$$Nu(\zeta,\tau) = \sum_{j=0}^{\infty} A_j, \tag{5}$$

Where:

$$A_j = \frac{1}{j!}\left[\frac{d^j}{d\lambda^j}\left\{N\sum_{j=0}^{\infty}\left(\lambda^j u_j\right)\right]\right]_{\lambda=0} \quad j = 0,1,\ldots \tag{6}$$

are called domain polynomials. To remove ambiguity in the normalization of the nonlocal memory contribution, we define the left-sided Riemann-Liouville fractional integral of order $\beta > 0$ by

$$\left(I_{0+}^{\beta} g\right)(t) = \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} g(s) ds \tag{7}$$

Accordingly, the weakly singular memory operator in Equation (5) is represented as

$$\mathscr{M}_{\beta,\kappa}[u](x,t) = \frac{\kappa}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} \mathscr{G}[u](x,s) ds = \kappa I_{0+}^{\beta} \mathscr{G}[u](x,t) \tag{8}$$

where $\mathscr{G}[u]$ denotes the specific spatial or nonlinear expression appearing under the memory integral in the governing FPIDE, β controls the kernel order, and κ determines the magnitude of the additional memory contribution. For $0 < \beta < 1$, the kernel has an integrable weak singularity at $s = t$. At $\beta = 1$, it reduces to the ordinary Volterra integration kernel. The Caputo derivative appearing in the same governing equation is defined independently by

$${}^{C}D_{0+}^{\alpha} u(x,t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \frac{\partial u(x,s)}{\partial s} ds, \ 0 < \alpha < 1. \tag{9}$$

Equivalently,

$${}^{C}D_{0+}^{\alpha} u = I_{0+}^{1-\alpha} \partial_t u. \tag{10}$$

The distinction is fundamental: the Caputo derivative describes nonlocal temporal differentiation of the solution, whereas the additional memory operator represents a separate hereditary contribution involving $\mathscr{G}[u]$. The fractional orders α and β are therefore independent model parameters unless a particular physical formulation imposes a relationship between them. If the original model employs the unnormalized integral

$$\int_0^t (t-s)^{\alpha-1} \mathscr{G}[u](x,s) ds \tag{11}$$

the corresponding normalized representation is recovered by choosing

$$\beta = \alpha, \ \kappa = \Gamma(\alpha).$$

This equality concerns only the normalization of the additional memory integral and does not identify that integral with the Caputo derivative. Under suitable regularity and Laplace-transformability assumptions, the corresponding transform identities are

$$\mathscr{L}_t\left\{{}^{C}D_{0+}^{\alpha} u\right\}(x,p) = p^{\alpha} U(x,p) - p^{\alpha-1} u(x,0) \tag{12}$$

and

$$\mathscr{L}_t\left\{\mathscr{M}_{\beta,\kappa}[u]\right\}(x,p) = \kappa p^{-\beta} \mathscr{L}_t\{\mathscr{G}[u]\}(x,p). \tag{13}$$

These identities establish the precise relationship between the two nonlocal operators and provide the analytical basis for constructing the Laplace-Adomian recurrence. The normalized convention must be retained consistently in the convergence analysis, numerical quadrature, manufactured forcing terms, and inverse parameter calibration.

© 2026 NSP
Natural Sciences Publishing Cor.

Lemma 2.6 (Laplace-transform identities). Let $f(\tau)$ and $g(\tau)$ be piecewise continuous on $[0,\infty)$, of exponential order, and right-continuous at the origin. Define $F(s)=\mathbb{L}[f(\tau)], G(s)=\mathbb{L}[g(\tau)]$. Let λ,η,ρ be constants and let $\zeta>0$. Whenever the displayed transforms and limits exist, the following identities hold [11, 46-49]:

(1) $\mathbb{L}[\eta f(\tau)+\lambda g(\tau)]=\eta F(s)+\lambda G(s)$.
(2) $\mathbb{L}^{-1}[\eta F(s)+\lambda G(s)]=\eta f(\tau)+\lambda g(\tau)$.
(3) $\mathbb{L}[e^{\rho\tau}f(\tau)]=F(s-\rho)$.
(4) $\mathbb{L}[f(\zeta\tau)]=\frac{1}{\zeta}F\left(\frac{s}{\zeta}\right),\zeta>0$.
(5) $\lim_{s\to\infty} sF(s)=f(0)$.
(6) $\mathbb{L}\left[J_\tau^\beta f(\tau)\right]=\frac{F(s)}{s^\beta},\beta>0$.
(7) $\mathbb{L}\left[D_\tau^\delta f(\tau)\right]=s^\delta F(s)-\sum_{k=0}^{m-1}s^{\delta-k-1}f^{(k)}(0), m-1<\delta\le m$.

Here, J denotes the Riemann-Liouville fractional integral and D denotes the Caputo derivative, with their orders shown in the superscripts.

Proof. Identities (1)- (4) are consequences of the linearity property and standard variable transformations; identity (5) is the definition of the initial value theorem, and (6) results from the convolution theorem. Furthermore, since we have shown that the Caputo fractional derivative can be expressed in terms of the Riemann- Liouville fractional integral of the m^{th} order derivative of the function $f(t)$, we obtain (7) by taking the Laplace transform of each side.

3 Laplace-Adomian decomposition framework

3.1 Adaptive LADM procedure for nonlinear FPIDEs

To develop the analytical base for the new hybrid error-controlled strategy for solving the generalized nonlinear fractional partial integro-differential equation (FPIDE), we begin by considering the generalized nonlinear FPIDE as follows [14, 15, 50]:

$$\tau\alpha v(\zeta,\tau)+av(\zeta,\tau)v_\zeta(\zeta,\tau)=\kappa\tau\beta v_{\zeta\zeta}(\zeta,\tau)+f(\zeta,\tau),\ 0<\alpha,\beta\le 1. \tag{14}$$

where *a and* κ represent model parameters; f represents a "source" or forcing function; and where is an appropriately normalized Riemann-Liouville fractional integral.

$$v(\zeta,0)=k(\zeta),\ v(0,\tau)=v(L,\tau)=0. \tag{15}$$

The memory operator is defined explicitly to remove the kernel-normalization ambiguity:

$$\tau\beta\phi(\tau)=\frac{1}{\Gamma(\beta)}\int_0^\tau(\tau-q)^{\beta-1}\phi(q)dq,\ \beta>0 \tag{16}$$

Using the same procedure of applying the Laplace Transform to the integral given in Equation (1) but now with a time-dependent kernel, we obtain $\beta=\alpha$ and $\kappa=\Gamma(\alpha)$. Thus, Eq. (5) retains the original model as a special case while distinguishing the Caputo derivative order α from the memory order β. The inverse Laplace transform of both sides gives us the desired Volterra-type fixed point equation:

$$\mathscr{L}\{\tau\alpha v\}=s^\alpha V(\zeta,s)-s^{\alpha-1}k(\zeta),$$

And then applying the inverse transform yields the equivalent of a Volterra fixed-point equation.

$$v(\zeta,\tau)=k(\zeta)+\tau\alpha f(\zeta,\tau)-a\tau\alpha\mathscr{N}(v)+\kappa\tau\alpha+\beta v_{\zeta\zeta}(\zeta,\tau). \tag{17}$$

The non-linear operator is the quadratic convective terms. The above compact form of this equation replaces all the individual elementary transforms needed to get to it, and thus serves as an appropriate starting point for both the convergence analysis and the computer code. We introduce the decompositions.

$$v=\sum_{n=0}^{\infty}v_n,\ \mathscr{N}(v)=vv_\zeta=\sum_{n=0}^{\infty}A_n. \tag{18}$$

© 2026 NSP
Natural Sciences Publishing Cor.

Nonlinear terms are not left to an unknown representation. Using the polynomial for the second-order convection term as defined in Eq. (18):

$$A_n = \frac{1}{n!}\frac{d^n}{d\lambda^n}\left[\left(\sum_{j=0}^{\infty}\lambda^j v_j\right)\partial_\zeta\left(\sum_{j=0}^{\infty}\lambda^j v_j\right)\right]\Bigg|_{\lambda=0} = \sum_{j=0}^{n} v_j \partial_\zeta v_{n-j}. \tag{19}$$

we can obtain the first few polynomials using Eq. (9). These then are used in Eqs. (8) to determine a recursive form of the linear ADM (LADM) scheme.

$$v_0 = k + \tau\alpha f,\ v_{n+1} = -a\tau\alpha A_n + \kappa\tau\alpha + \beta\partial_{\zeta\zeta}v_n,\ n \geq 0. \tag{20}$$

The truncation of the series in Eq. (19) is not done based on some fixed order chosen arbitrarily. Instead, it is determined when both the locally-contractive tail estimate, provided by the convergence analysis with respect to q $<$ 1, and the governing equation residual have decreased below specified tolerance levels.

$$S_M = \sum_{n=0}^{M} v_n,\ E_M = \frac{q}{1-q}v_M, E_M \leq \text{ atol } + \text{ rtol } S_M,\ \mathscr{R}(S_M) \leq \tau_R. \tag{21}$$

To obtain the residual of Eq. (11), we substitute the partial sum of Eq. (19) into Eq. (5); if the contraction factor is calculated numerically through successively generated terms, then before training an empirical ratio to estimate the contraction factor, we must verify that the geometric tail exists; if no geometric tail exists, we will report the tail value only as a potential indicator, not as a proven limit. The reason for this difference is so that an empirically calculated ratio cannot be represented as a proved convergence theorem. Therefore, for long-term simulation, Eqs. (10)-(12) are employed sequentially over each temporal slab, with all prior slabs contributing to the full Caputo history. The memory term is never reset when crossing a slab boundary. Once both errors have been sufficiently reduced in accordance with the LADM error gates, the ANN can be trained to produce.

$$\widehat{v}_H(\zeta,\tau) = S_M(\zeta,\tau) + \psi(\zeta,\tau)e_\phi(\zeta,\tau),\ e_\phi \approx v_{ref} - S_M. \tag{22}$$

In particular, since the mask is zero at time $t = 0$ and at the spatial boundaries, following acceptance of the adaptive LADM approximation, the hybrid correction is constructed according to the constrained residual-learning formulation described in Section 3.4.

3.1.1 Computational Complexity of Full-History Caputo Continuation

The memory-preserving continuation strategy retains all contributions from previously completed temporal slabs to ensure consistency with the nonlocal Caputo derivative. Although this avoids the artificial resetting of fractional memory, it increases computational and storage requirements relative to local time-stepping methods. Consequently, the cost of full-history evaluation must be distinguished from the computational expense associated with adaptive LADM truncation and neural residual correction. Consider a uniform temporal partition $t_n = n\Delta t, n = 0,\ldots,N_t$, and a spatial discretization with N_x degrees of freedom. For $0 < \alpha < 1$, the Caputo derivative is approximated using the L1 formula:

$${}^{C}D_t^{\alpha}u(t_n) \approx \frac{1}{\Gamma(2-\alpha)(\Delta t)^{\alpha}}\sum_{j=0}^{n-1} b_j^{(\alpha)}\left(u^{n-j} - u^{n-j-1}\right), \tag{23}$$

Where:

$$b_j^{(\alpha)} = (j+1)^{1-\alpha} - j^{1-\alpha}.$$

At time level t_n, the direct evaluation requires $O(nN_x)$ arithmetic operations. Summing over all temporal levels gives

$$C_{\text{Caputo}} = O\left(N_x\sum_{n=1}^{N_t} n\right) = O\left(N_x N_t^2\right). \tag{24}$$

If the complete solution history is retained, the associated storage requirement is

$$S_{\text{history}} = O(N_x N_t).$$

These estimates refer specifically to the Caputo history evaluation and do not include spatial discretization, nonlinear Adomian polynomial construction, repeated residual evaluations, parameter optimization, or ANN training. Furthermore,

© 2026 NSP
Natural Sciences Publishing Cor.

the weakly singular Volterra integral may introduce an additional history-dependent computational cost. When evaluated directly at every time level, its temporal complexity is also generally quadratic, provided the kernel evaluation and spatial operations have bounded cost per history entry. The total computational burden of the hybrid framework can therefore be expressed schematically as

$$C_{\text{total}} = C_{\text{history}} + C_{LADM} + C_{\text{residual}} + C_{ANN} + C_{\text{calibration}}, \tag{25}$$

where the individual terms represent the costs of nonlocal history evaluation, adaptive decomposition, residual diagnostics, neural-network training and inference, and parameter estimation, respectively. The costs of repeated history evaluations during adaptive corrections depend on implementation details and cannot be inferred from the temporal grid dimensions alone. An alternative approach is to approximate the fractional-memory kernels using a sum of exponentials,

$$k_\alpha(t) \approx \sum_{\updownarrow=1}^{K} w_{\updownarrow} e^{-\lambda_{\updownarrow} t}, \tag{26}$$

with coefficients and decay rates selected to approximate the relevant kernel over the prescribed temporal interval. Under an appropriate recursive implementation, the historical contribution can then be updated using K auxiliary states rather than recomputing the complete convolution. This reduces the history-evaluation complexity to approximately

$$C_{SOE} = O(KN_xN_t), \; S_{SOE} = O(KN_x), \tag{27}$$

excluding any explicitly retained local-history window and temporary storage.

3.2 Model-specific well-Posedness and certified LADM convergence

Consider the weakly singular nonlinear FPIDE

$$t\alpha u(x,t) + au(x,t)u_x(x,t) = \kappa t\beta u_{xx}(x,t) + g(x,t), \; 0 < \alpha, \beta \leq 1, \tag{28}$$

where the Riemann-Liouville integral is normalized by

$$t\beta v(t) = \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} v(s) ds, \; \beta > 0 \tag{29}$$

The unnormalized kernel in the original formulation is recovered by setting $\beta = \alpha$ and $\kappa = \Gamma(\alpha)$. For measured boundary values $u(0,t) = u_{\text{in}}(t)$ and $u(L,t) = u_{\text{out}}(t)$, introduce the lifting

$$w(x,t) = u(x,t) - \left(1 - \frac{x}{L}\right) u_{in}(t) - \frac{x}{L} u_{out}(t). \tag{30}$$

Throughout the convergence analysis, the weakly singular memory contribution is interpreted according to the normalized Riemann-Liouville convention introduced in Section 3.1. Consequently, the coefficient κ multiplies the normalized integral I_{0+}^{β}, and the factor $1/\Gamma(\beta)$ must not be introduced a second time in the operator estimates. The Caputo order α and the memory-kernel order β remain distinct parameters. All fixed-point estimates, Lipschitz bounds, and contraction conditions are understood with respect to this consistent normalization. Then $w(0,t) = w(L,t) = 0$. All of the previously identified lift terms, along with their fractional derivative in time, are combined to form a modified source term. After a Galerkin or finite difference spatial discretization, we denote the finite dimensional Banach space formed from this process as X_h For simplicity, we will also refer to the lifted semi-discrete solution as u.

$$\mathcal{N}_h(u) = uD_h u, \; \mathcal{L}_h u = D_h^2 u, \; Y_T = C([0,T]; X_h), \; u_{Y_T} = \sup_{0 \leq t \leq T} u(t)_{X_h}. \tag{31}$$

Applying $t\alpha$ to Eq. model-fpide and using the semigroup identity $t\alpha t\beta = t\alpha + \beta$ gives

$$b(t) = u_{\text{init}} + t\alpha g_h(t), (\mathcal{K}u)(t) = -at\alpha \mathcal{N}_h(u)(t) + \kappa t\alpha + \beta \mathcal{L}_h u(t), u = b + \mathcal{K}u =: \mathcal{F}(u). \tag{32}$$

Let

$$B_R = \{u \in Y_T : u_{Y_T} \leq R\}.$$

The analysis uses the following explicit assumptions.

© 2026 NSP
Natural Sciences Publishing Cor.

(A1) $g_h \in C([0,T];X_h)$ and $u_{\text{init}} \in X_h$.

(A2) $\mathscr{L}_h : X_h \to X_h$ is bounded, with $\mathscr{L}_h u - \mathscr{L}_h v_{X_h} \le L_{\mathscr{L}} u - v_{X_h}$.

(A3) $\mathscr{N}_h$ has an analytic extension to a complex neighborhood of the radius- R ball and satisfies $M_{\mathscr{N}}(R) = \sup_{z_{X_h} \le R} \mathscr{N}_h(z)_{X_h} < \infty$, $\mathscr{N}_h(u) - \mathscr{N}_h(v)_{X_h} \le L_{\mathscr{N}}(R)u - v_{X_h}$ For the discrete Burgers term, if $uv_{X_h} \le C_X u_{X_h} v_{X_h}$ and $D_h u_{X_h} \le C_D u_{X_h}$, then one may take $M_{\mathscr{N}}(R) \le C_X C_D R^2$, $L_{\mathscr{N}}(R) \le 2C_X C_D R$.

(A4) The radius and time horizon satisfy the self-mapping and contraction conditions $\Phi_R(T) := b_{Y_T} + \dfrac{|a|M_{\mathscr{N}}(R)T^\alpha}{\Gamma(\alpha+1)} + \dfrac{|\kappa|L_{\mathscr{L}}RT^{\alpha+\beta}}{\Gamma(\alpha+\beta+1)} \le R, q_R(T) := \dfrac{|a|L_{\mathscr{N}}(R)T^\alpha}{\Gamma(\alpha+1)} + \dfrac{|\kappa|L_{\mathscr{L}}T^{\alpha+\beta}}{\Gamma(\alpha+\beta+1)} < 1.$

The contraction criterion in Equation (33) is a sufficient mathematical condition rather than an automatically satisfied property of the numerical approximation. Its practical verification requires computable bounds for the nonlinear and linear spatial operators, together with an admissible solution radius and a suitable temporal interval. For the semi-discrete Burgers-type nonlinearity $N_h(u) = uD_h u$, assume that the discrete norm satisfies

$$\|uv\|_{X_h} \le C_X \|u\|_{X_h} \|v\|_{X_h}, \tag{34}$$

and that

$$\|D_h u\|_{X_h} \le C_D \|u\|_{X_h}.$$

For $u, v \in B_R$, the nonlinear difference can be written as

$$N_h(u) - N_h(v) = (u-v)D_h u + vD_h(u-v). \tag{35}$$

Consequently,

$$\|N_h(u) - N_h(v)\|_{X_h} \le 2C_X C_D R\|u-v\|_{X_h}, \tag{36}$$

so that the conservative bounds

$$M_N(R) \le C_X C_D R^2,\ L_N(R) \le 2C_X C_D R \tag{37}$$

may be used. Similarly, the linear spatial operator must satisfy

$$\|L_h u - L_h v\|_{X_h} \le L_L \|u-v\|_{X_h},$$

where L_L can be bounded using the induced norm of the discrete spatial matrix. For a candidate radius R and time interval of length T, the resulting sufficient contraction bound is

$$q_R(T) \le \frac{2|a|C_X C_D R T^\alpha}{\Gamma(\alpha+1)} + \frac{|\kappa|L_L T^{\alpha+\beta}}{\Gamma(\alpha+\beta+1)}. \tag{38}$$

The fractionally ordered differential equation is considered to have an appropriate time scale if this upper limit is less than one. Additionally, it is necessary to verify the self-mapping condition $\Phi_R(T) \le R$ The coefficients of a, κ and the fractional orders α and β can be determined through observation. Therefore, the estimation using the point estimate will provide a conditional evaluation. To provide a conservative evaluation, the relevant inequalities are evaluated across an admissible parameter space or uncertainty set. If the conditions for being sufficient are not satisfied, decreasing the time interval will result in a reduction in the contribution due to fractional integration. However, decreasing the time interval will not ensure that the fixed-point operator is contractive since the self-mapping inequality must still hold true and consistent history must be maintained with respect to the inherited Caputo and Volterra terms. Finally, the condition $q_R(T) < 1$ indicates that the fixed-point operator is contractive under the given assumptions. Furthermore, absolute convergence of the LADM series requires the analytic-majorant condition in Equation (21), which indicates that empirical ratios of successive Adomian components should be reported as indicators of convergence until they relate to a valid majorant. Under assumptions (A1) -(A4), the lifted semi discrete problem has a unique mild solution $u \in B_R$. Moreover, the fixed-point iterates $u^{(m+1)} = \mathscr{F}\left(u^{(m)}\right)$ satisfy

$$u - u^{(m)}{}_{Y_T} \le \frac{q_R(T)^m}{1-q_R(T)} u^{(1)} - u^{(0)}{}_{Y_T}. \tag{39}$$

© 2026 NSP
Natural Sciences Publishing Cor.

To obtain convergence of the LADM series itself, suppose additionally that there exists $\rho > 1$ such that

$$\Phi_{R,\rho}(T) := b_{Y_T} + \rho \left[\frac{|a| M_{\mathscr{N}}(R) T^{\alpha}}{\Gamma(\alpha+1)} + \frac{|\kappa| L_{\mathscr{L}} R T^{\alpha+\beta}}{\Gamma(\alpha+\beta+1)} \right] \leq R, \rho q_R(T) < 1. \tag{40}$$

Then the parameterized fixed point

$$U(\lambda) = b + \lambda \mathscr{K}(U(\lambda)) = \sum_{n=0}^{\infty} \lambda^n u_n \tag{41}$$

is analytic for $|\lambda| < \rho$. Consequently, the LADM series $S_M = \sum_{n=0}^{M} u_n$ converges absolutely at $\lambda = 1$ to the unique mild solution. For every r satisfying $1 < r < \rho$, define
$M_r = \sup_{|\lambda|=r} U(\lambda)_{Y_T}$. The truncation error obeys the explicit bound

$$u - S_{MY_T} \leq \frac{M_r r^{-(M+1)}}{1 - r^{-1}} \tag{42}$$

For every $\gamma > 0$ and $z \in Y_T$,

$$t^{\gamma} z_{Y_T} \leq \frac{T^{\gamma}}{\Gamma(\gamma+1)} z_{Y_T} \tag{43}$$

Combining Eq. fractional-integral-bound with (A2) -(A3) shows that $\mathscr{F}(B_R) \subseteq B_R$ and

$$\mathscr{F}(u) - \mathscr{F}(v)_{Y_T} \leq q_R(T) u - v_{Y_T}.$$

Banach's Fixed Point Theorem provides both an Existence, Uniqueness and also a Picard Error Bound. Given that $|\lambda| \leq \rho$, the mapping $u \mapsto b + \lambda \mathscr{K}(u)$ is a self-map of B_R and thus has a Lipschitz constant of less than $\rho q_R(T) < 1$. The contraction is uniform with respect to λ. K is analytic as per (A3), so its unique fixed-point $U(\lambda)$ is analytic with respect to λ for all $|\lambda| < \rho$. Differentiation of Equation ladm-homotopy at $\lambda = 0$ produces the Adomian components. Cauchy's Estimate gives $u_{nY_T} \leq M_r r^{-n}$ for $1 < r < \rho$ and $u_{n+1Y_T} \leq \eta u_{nY_T}$ for each $n \geq M$, then.

$$u - S_{MY_T} \leq \frac{\eta}{1-\eta} u_{MY_T} \tag{44}$$

The bound from above may be combined with the governing equation residual to provide the dual adaptive stop criteria employed by the Error-Controlled version of the LADM method. The contraction criterion (Eq. self-map-contract) provides evidence for convergence of the iterative solution of the fixed-point problem, but does not independently establish convergence of the Adomian series. The analytic-majorant condition (Eq. analytic-major) fills this gap. Furthermore, since the theorem is stated after spatial projection, it establishes well posedness and convergence of the semi-discrete Ness and LADM models. Establishing convergence of the continuous FPIDE ($h \to 0$) will require an independent analysis of spatial consistency and stability estimates.

3.3 Adaptive truncation and memory-preserving long-time continuation

The geometric-tail result obtained earlier as an error-control tool will not be restated as a second Convergence Theorem. Define n to be the beginning index of the current time slab $In = [tn, tn+1]$ use the supremum norm on $C(In; Xh)$ and denote the LADM partial sum consisting of the first M terms S_M If a theoretical bound or verified coefficient estimate can assure that the norms of all unresolved components will have a factor no greater than η relative to their predecessors for every $m \geq M$, then the remaining tail of the series has been sufficiently resolved so that.

$$u - S_{MY_{I_n}} \leq E_M := \frac{\eta}{1-\eta} u_{MY_{I_n}}, \; 0 < \eta < 1. \tag{45}$$

Only a posteriori certification of equation (46) will occur if the geometric inequality is shown to apply to all unaccounted-for components. The ratio obtained by solving an assumed finite number of terms in a series is only useful as a diagnostic tool. It cannot be used as an acceptable error estimation technique without some form of majorization or support. Tail estimates can prematurely terminate the iterative process when successive components cancel. Therefore, the method's accuracy should be checked using both the residual of the lifted semi-discrete version of the governing equations and the approximate solution:

© 2026 NSP
Natural Sciences Publishing Cor.

$$\mathscr{R}_M(t) = t\alpha S_M(t) + a\mathscr{N}_h(S_M(t)) - \kappa I_t^{\beta}\mathscr{L}_h S_M(t) - g_h(t). \tag{46}$$

The residual is evaluated with the same spatial discretization and fractional quadrature used by the solver. The M-term approximation is accepted only when both the tail and residual conditions are satisfied:

$$E_M \leq \text{ atol } + rtolS_{MY_{I_n}}, \mathscr{R}_{MY_{I_n}} \leq \tau_R. \tag{47}$$

The first inequality restricts the un-certified LADM terms while the second term will prevent premature termination due to cancellation of rounding errors or empirically obtained parameters for the fractional order. All other values such as atoll, troll, residual tolerance τ_R and max. number of terms $M_{\max}$ are specified in the experimental setup and will be studied in the sensitivity study. Terms will be calculated at sequential times on each time slab. If the certified local contraction factor $q_R(\Delta tn)$has not dropped below one, then either the two conditions described in Eq. (48) have not been satisfied prior to reaching the maximal number of terms Mmax, then the size of the current slab is halved and calculations repeat. If at that point there is a defined minimal slab length specified in which none of those requirements were met; then the algorithm states no convergence and will not accept any of the approximations provided.

$$t\alpha u(t) = \frac{1}{\Gamma(1-\alpha)}\left[\int_0^{t_n} \frac{u^{'}(s)}{(t-s)^{\alpha}}ds + \int_{t_n}^{t} \frac{u^{'}(s)}{(t-s)^{\alpha}}ds\right] =: H_n(t) + {}^{C}D^{\alpha}_{t_n+}u(t). \tag{48}$$

$H_n(t)$ represents the inherited history from the cumulative contributions of each previously completed slab. In contrast, the last term includes only those contributions made during the most recent slab. There are no conditions under which $H_n(t)$ can be either discarded or reset at time t_n. When α=1 the first derivative of Eq. (49) becomes the standard first derivative. If the history is represented as a discrete summation, then it can be directly evaluated from the history table. Alternatively, if a sum-of-exponentials approximation is used for the history, then the compression tolerance should also be accounted for in the overall error budget. Once both adaptive error gates have been met, then the output of SM will be passed to the residual ANN stage. Therefore, any corrections to the physics-based solution that occur due to application of the neural network will be in response to an error-controlled physical solution and therefore cannot "hide" a failure or lack of convergence of the LADM.

3.4 Residual ANN correction with physical constraints

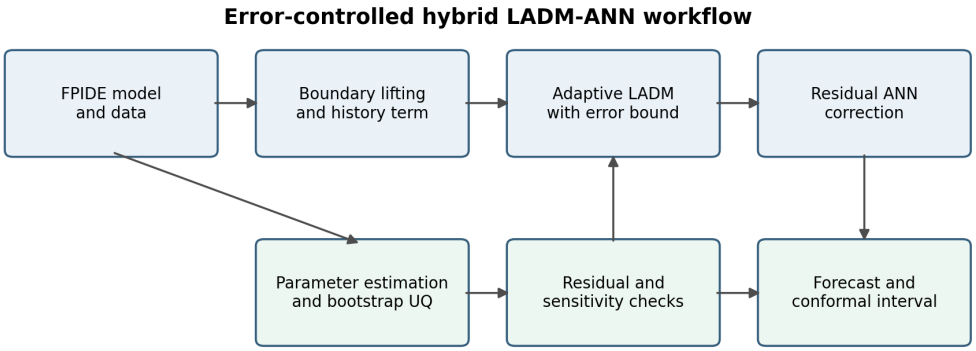


Fig. 1. Error-controlled hybrid LADM-ANN workflow.

The neural component will remain intact but instead of being employed freely as a surrogate for the FPIDE it will function solely as a correction term. To define this hybrid predictor, we first use the lifted boundary function from equation (27), which we denote by $\updownarrow(x,t)$ and the sum of the terms LADM that is the adaptive approximations of the homogeneous function over the time slab from t_n to t_{n+1}. With $\xi = x / L$ denoting the scaled spatial coordinate and with $\eta = (t-t_n)/\Delta t$ representing the scaled temporal coordinate of the slab, then the hybrid predictor can be defined via equation (30).

© 2026 NSP
Natural Sciences Publishing Cor.

$$\widehat{u}_H(x,t) = \updownarrow(x,t) + S_M(x,t) + \psi_n(x,t) r_\phi\left(z(x,t)\right),$$

$$\psi_n(x,t) = \xi(1-\xi)\eta^\alpha,\ \xi = \frac{x}{L},\ \eta = \frac{t-t_n}{\Delta t}. \tag{49}$$

The Hybrid approach uses Adaptive Laplace- Adomian Decomposition Method (ALADMM) as the first order physics-based approximation method. Then it adds an artificial Neural Network as a second-order error correction. In this way, the Hybrid Solution is formed from the combination of ALADMM approximate solution with a vanishingly small NN corrective term. The NN acts as a residual correction; it does not act independently as a substitute model for the original FPIDE. This establishes the specific role(s) each of the two models will play in the forthcoming Numerical Experiments.

$$\mathscr{L}_\phi = \frac{1}{N_{tr}} \sum_{(i,j)\in\mathscr{T}} w_{ij}\rho_\delta\left(e^{obs}_{M,ij} - \psi_{n,ij} r_\phi\left(z_{ij}\right)\right) + \frac{\lambda_R}{N_c}\sum_{q=1}^{N_c}\left|\mathscr{R}_{\widehat{\theta}}\left[\widehat{u}_H\right](x_q,t_q)\right|^2 + \frac{\lambda_A}{N_c}\sum_{q=1}^{N_c}\left(\left[-\widehat{u}_H(x_q,t_q)\right]_+^2 + \left[\widehat{u}_H(x_q,t_q)\right.\right. \tag{50}$$

The deterministic LADM truncation certificate and the statistical ANN generalization error will therefore be presented as two separate deterministic and data-driven error statements, rather than being combined into one deterministic theorem. The hybrid error is thus defined as the difference between the actual residual value and the ANN correction. Adding to that the LADM tail bound once again to the norm of the same residual would therefore amount to doubling counting the tail.

3.5 Parameter estimation, sensitivity, and uncertainty

The model parameters were defined as follows: $\theta = (\alpha, \beta, \mathrm{a}, \kappa, \mathrm{b}_0, \mathrm{b}_v, \mathrm{b}_{vx})$, where the source was modeled using both an intercept value, the measured speed and its spatial gradient. All fractional order values were mapped onto their respective bounded grid search areas: $\alpha \in [0.55, 1]$ and $\beta \in [0.20, 1]$. The source coefficients were estimated by Constrained Least Squares estimation. To normalize the residuals for all possible candidate fractional orders to be of similar magnitude, the objective function was scaled by ε. The scaling factor $s\alpha$ is the sample standard deviation of the discretized Caputo derivative. A diagonal matrix D that scales each parameter estimate and a very small number ε used in numerical computations.

$$\mathscr{R}_\theta[u] = {}^C D_t^\alpha u + auu_x - \kappa I_t^\beta u_{xx} - (b_0 + b_v v + b_{vx} v_x),\ \widehat{\theta} = \theta \in \Theta argmin\left\{\frac{1}{N}\sum_{i,j} w_{ij}\left|\frac{\mathscr{R}_\theta\left[u^{obs}\right](x_i,t_j)}{s_\alpha + \varepsilon}\right|^2 + \lambda_\theta \|D\right. \tag{51}$$

For the Caputo derivative, there exists L1 evaluation as given in Equation (51). Therefore, at each step of finding candidates for β, the weakly singular integrals are recalculated. For the L1 consistency error to be of second order in time under sufficient temporal smoothness, it is theoretically expected that for fractional solutions with a weak singularity at the initial time, the same can be observed. Therefore, if you find your numerical results fail to achieve such a high order rate of convergence, use either a graded mesh or some form of start-up correction.

$${}^C D_t^\alpha u^n \approx \frac{1}{\Gamma(2-\alpha)\Delta t^\alpha}\sum_{k=0}^{n-1}\left[(k+1)^{1-\alpha} - k^{1-\alpha}\right]\left(u^{n-k} - u^{n-k-1}\right). \tag{52}$$

Uncertainty analysis also follows the same chronological sequence that data does. A moving-block bootstrap creates new samples by taking a block of continuous times and then fit the entire parameter vector, with all parameters included in the fraction order and memory operators, to find the appropriate percentile. Latin Hypercube Sampling with +/- 15% of the original calibrated values, limited to the allowable parameter space, is used for Rank-Based Sensitivity Screening. The data is broken up into three separate periods based upon time: Training, Validation-Calibration, Test; Random Row Splitting is Not Used. Equation (53) defines how there are two types of Error Control. The First Line represents the Deterministic LADM Tail Bound for the Calibrated FPIDE Solution. All Other Lines represent the Split-Conformal Interval Using the Finite-Sample-Corrected Empirical Quantile of the Absolute Hybrid Residuals on the Calibration Period. For Temporally Dependent Data, Rolling Sequential Conformal Coverage will be reported and is an Empirical Representation and Not a Distribution-Free Theorem.

© 2026 NSP
Natural Sciences Publishing Cor.

$$\left\|u_{\widehat{\theta}} - (\updownarrow + S_M)\right\|_{\infty} \leq E_M, q_{1-\gamma} = \text{ Quantile }_{\frac{\lceil (N_{cal}+1)(1-\gamma)\rceil}{N_{cal}}} \left\{\left|\varepsilon_i^{cal}\right|\right\}_{i=1}^{N_{cal}}, \widehat{C}_{1-\gamma}(z) = \left[\widehat{u}_H(z) - q_{1-\gamma}, \widehat{u}_H(z)\right. \tag{53}$$

Because the conformal quantile has been calibrated based upon the overall hybrid predictive error for the entire time period in which all of these have occurred (truncation, discretization, parameter estimation, model misspecification, measurement errors, and ANN generalization), they are incorporated into that error. Therefore, they cannot be simply mechanically added to the LADM tail bound. Instead, the deterministic truncation bound, the bootstrap parameters of the interval for each parameter, the sensitivity indexes, and the empirical conformal coverage of each are used as complementary diagnostics.

3.6 Reproducible adaptive algorithm

Algorithm 1. The error-controlled memory-preserving LADM-ANN solver has the steps:

1. Move all of the nonzero boundary data into memory and load up the full fractional history.
2. Compute a local Lipschitz constant and estimate or assign θ. Then determine a lower bound for $q(\Delta t)$.
3. If q is not less than the predetermined safe value for q, then divide $\Delta t$by two and redo step 2.
4. Determine the Adomian components and update S_M, E_M, and the FPIDE residual.
5. Stop if either equation (29) condition holds; otherwise generate another component until $M_{\max}$.
6. Take the next time slab using the history of integrals from equation (30).
7. Only fit the ANN to residuals during the training time segment and choose the architecture based upon how well it performs on the validation data.
8. Develop both a hybrid prediction and an associated sequential conformal interval during the test time segment.

4 Numerical Verification and Controlled Experiments

This section replaces the fixed-time numerical expansions and visually redundant plots of the original example with an experimental solution in which all variables are manufactured, under control. The revised experimental design validates the fractional operators, uses the explicit LADM tail estimator along with the FPIDE residual gate, adapts the number of series components (the number of terms in the time-stepping expansion), and makes a physical constraint to correct a neural network's predictions on validation data. All numbers reported below have been re-computed from the specified discretization and random seeds.

4.1 Experimental protocol

We discretized both spatial and temporal intervals. Spatially we have 129 equally-spaced nodes; Temporally we have 501 nodes over the interval [0,1]. Both Riemann-Liouville integrals were solved via Product Integration. For evaluation of the Caputo derivative in the Residual term we used the L1 Formula. Interior spatial derivatives were approximated with Second-Order Centered Differences. In our example we extended the interval from 10^{-2} seconds to $100 * 10^{-2} = 1$ second. It is expected that as the ratio of the interval length to the original plot time increases (i.e., the time is increased) an error will become evident that could not be seen if comparing two short periods of time.

Table 1. Reproducibility configuration for Example 4.1

Component	Setting
Fractional orders	$\alpha = 0.80$, $\beta = 0.60$
Physical coefficients	a = 0.05, $\kappa = 0.002$
Space-time grid	Nx = 129, Nt = 501, dx = 0.0078125, dt = 0.002
Adaptive tolerances	atol = 2e-4, rtol = 2e-3, residual tolerance = 2e-2
Maximum LADM order	Mmax = 9
Residual ANN	Two tanh hidden layers with 16 and 8 neurons
Data split	12,000 training, 4,000 validation, 12,000 held-out points
Uncertainty	500 time-block bootstrap resamples
Random seed	20260924

© 2026 NSP
Natural Sciences Publishing Cor.

4.2 Example 4.1 Manufactured-solution benchmark

Consider the nonlinear fractional partial integro-differential equation on:

$$^{C}D_t^{\alpha}u(x,t)+au(x,t)u_x(x,t)=\kappa\mathscr{I}_t^{\beta}u_{xx}(x,t)+g_{\alpha,\beta}(x,t),\ 0<\alpha\leq 1,\ 0<\beta<1. \tag{54}$$

The weakly singular memory operator is normalized as

$$\mathscr{I}_t^{\beta}v(x,t)=\frac{1}{\Gamma(\beta)}\int_0^t (t-s)^{\beta-1}v(x,s)ds. \tag{55}$$

The Benchmark problem is described with an exact analytical solution, the associated initial condition, and a homogeneous Dirichlet Boundary Condition (BC) in space:

$$u_{ex}(x,t)=sin(\pi x)-2t^3sin(2\pi x),u(x,0)=sin(\pi x),u(0,t)=u(1,t)=0. \tag{56}$$

Substitution of the exact field into equation (59) gives the manufactured source term

$$g_{\alpha,\beta}(x,t)=-\frac{12t^{3-\alpha}}{\Gamma(4-\alpha)}sin(2\pi x)+a\left[\frac{\pi}{2}\sin(2\pi x)-4\pi t^3\sin(\pi x)\cos(2\pi x)\ \right]. \tag{57}$$

The creation of this method provides the ability to find the exact solution so that each combination of (α,β) can be verified; it also eliminates the need to compare against some unknown fractional order reference. Using the discrete operator in the direct computation of Eq. (54), a maximum residual value of $2.960e^{-3}$ was found. This is indicative of the consistency floor of the employed mesh size as opposed to LADM truncation error.

4.3 Adaptive LADM and residual ANN

The fractional integral of order α will generate the decomposition components and the Adomian polynomials for the Burgers nonlinearity as follows:

$$u_0(x,t)=sin(\pi x)+\mathscr{I}_t^{\alpha}g_{\alpha,\beta}(x,t),A_m(x,t)=\sum_{j=0}^{m}u_j(x,t)\partial_x u_{m-j}(x,t),u_{m+1}(x,t)=-a\mathscr{I}_t^{\alpha}A_m(x,t)+\kappa\mathscr{I}_t^{\alpha+\beta}\partial_{xx}u \tag{58}$$

The Dual Stopping Rule selects the lowest order that can be admitted at each time step.

$$M^*(t)=min\left\{M:E_M(t)\leq atol+rtolS_M(\cdot,t)_{\infty},\mathscr{R}(S_M)(\cdot,t)_{\infty}\leq\tau_R\right\}. \tag{59}$$

where the first quantity is an a-posteriori computable estimate of the tail and R represents the complete FPIDE residual. This second condition is important as a very small residual in the final component may arise by cancellation and yet does not imply satisfaction with the differential equation. If neither of these conditions are met then additional components are created up to $M_{\max}$. The hybrid approximation is constructed using the constrained neural correction defined in Section 3.4, with the experimental configuration specified below. The Hybrid Field is:

$$\widehat{u}_H(x,t)=S_{M^*(t)}(x,t)+\psi(x,t)r_\theta(z),\ \psi(x,t)=x(1-x)t. \tag{60}$$

A vanishing multiplier ψ is required to enforce the model's exact initial and boundary conditions. Candidate tanh architectures (8), (12), (16) and (16,8) were evaluated based on their performance over the validation set. Tanh architecture (16,8) resulted in the lowest validation root mean squared error. Thus, its training objective was:

$$\mathscr{L}(\theta)=\frac{1}{N_{tr}}\sum_{(x_i,t_i)\in\mathscr{D}_{tr}}\left[u_{ex}(x_i,t_i)-S_{M^*(t_i)}(x_i,t_i)\right. \tag{61}$$

The neural correction component will approximate the residual between the adaptive LADM solution and the manufactured reference solution. Since the residual is derived through an approximation of the governing fractional pde, the learned correction of the residual is inherently influenced by the underlying fractional dynamics, nonlinear operators, and numerical discretization. Although the ANN does not have to independently reconstruct the complete solution based upon the LADM approximation; the distinction motivates investigating compact residual-learning architectures, without implying shallow networks are necessarily optimal. The distinction motivates the use of a compact feedforward

© 2026 NSP
Natural Sciences Publishing Cor.

architecture. However, using a compact feedforward architecture does not establish that a shallow network is optimal. four candidate architectures were considered for selecting the residual network: (8), (12), (16), and (16,8). Each entry denotes the number of neurons in successive hidden layers. The best performing architecture was selected based upon root mean squared error on the designated validation dataset. The two-hidden layer architecture (16,8) achieved the lowest RMSE, therefore it was retained for the manufactured-solution experiment. The purpose of restricting the search space to identify a suitable residual-learning configuration was to prevent adding unnecessary complexity to hybrid numerical framework. The chosen architecture should not be regarded as globally optimal. In particular, three or more hidden layers, much greater widths, or alternative activation functions were not systematically examined. Therefore, the results of the numerical experiments demonstrate the effectiveness of the chosen compact residual network under the experimental conditions studied here. However, they do not provide evidence that deeper neural architectures would provide no additional benefit. A comprehensive study of architecture-capacity is required to determine the relationship between depth of neural-networks, accuracy of approximations, generalization performance, and computational cost. In a real-world application using measured data, the target of equation (61) would have to be replaced with either observed or high-fidelity reference residuals. Therefore, the results from this "manufactured" experiment should not be interpreted as implying that exact residual labels will be available in all cases.

4.4 Adaptive truncation and extended-time behavior

Table 2 illustrates that, while only one LADM correction component is needed for the first time period, an additional five (total of six) components will be necessary by the end of time $t = 1$. The average or mean number of selected orders over the entire time span was 2.04; however, the highest or max order limit was never reached. Therefore, a fixed approximation of two terms is satisfactory for only a portion of the total time span and can therefore not be universally justified.

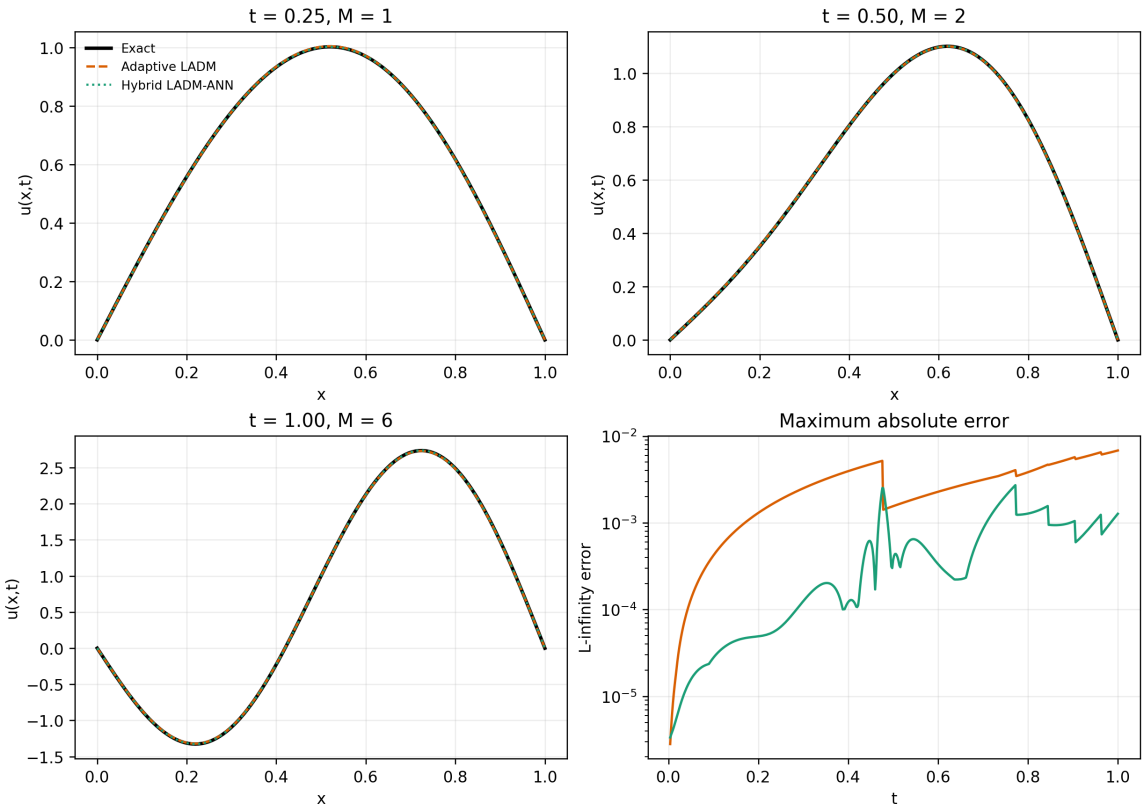


Fig. 2. Numerical solution profiles and approximation errors for Example 4.1. (a) Comparison of the manufactured exact solution, adaptive LADM approximation, and hybrid LADM–ANN approximation at representative time levels. (b) Corresponding absolute errors of the adaptive and hybrid approximations. (c) Maximum absolute error over the spatial domain as a function of time. Consistent line styles distinguish the methods across all panels.

© 2026 NSP
Natural Sciences Publishing Cor.

Table 2. Adaptive order, stopping diagnostics, and maximum absolute errors

t	M*	Tail estimator	FPIDE residual	Adaptive error	Hybrid error
0.10	1	1.978e-04	4.195e-03	4.196e-04	2.780e-05
0.25	1	9.161e-04	8.890e-03	1.867e-03	6.326e-05
0.50	2	7.794e-04	6.625e-03	1.556e-03	3.713e-04
0.75	2	3.022e-03	1.439e-02	3.687e-03	2.038e-03
1.00	6	9.695e-04	1.767e-02	6.823e-03	1.272e-03

To facilitate visual interpretation, Figure 2 separates the numerical solution profiles from the corresponding approximation-error diagnostics. Representative time levels are used to compare the manufactured exact solution, adaptive LADM approximation, and hybrid LADM–ANN approximation, while the error panels highlight differences that may be difficult to distinguish from overlapping solution curves.

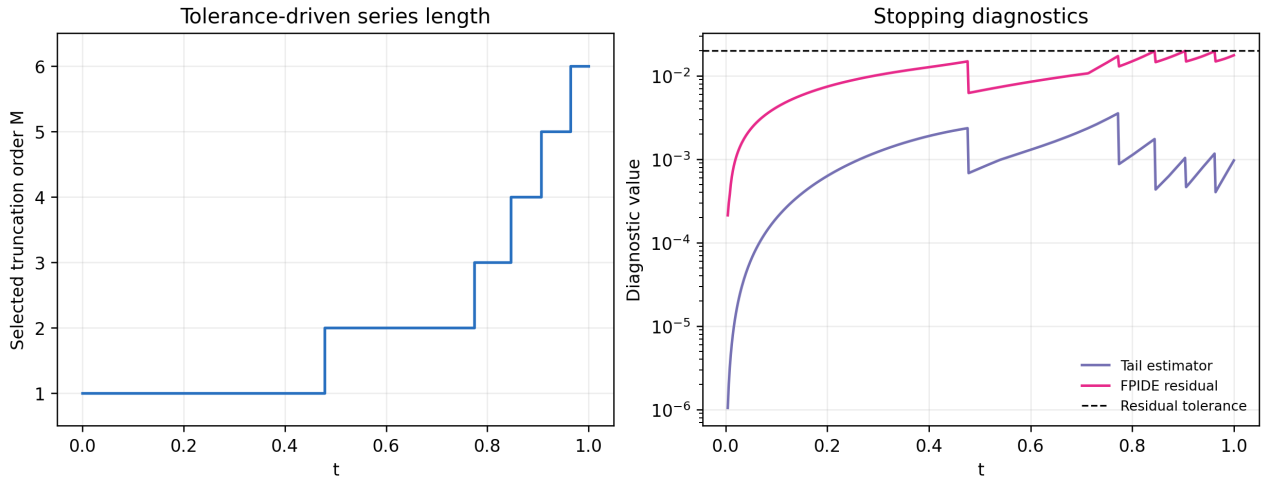


Fig. 3. Tolerance-driven LADM order and the corresponding tail and FPIDE-residual diagnostics.

4.5 Comparative accuracy and uncertainty

The evaluation of each methodology utilizes the exact "held out" data at all locations in space to be used with each methodology. A fixed order (i.e., no iterations) version of the Varangian Advection/Diffusion Model (LADM), which is an independent integration scheme using product solutions from each step in the calculation process, along with an optimization constrained Artificial Neural Network (ANN) that solves the problem without reference to the LADM discrepancy serves as two baseline methodologies. The ANN utilized the same split of samples, the same type of optimizer, candidate architectures, and vanishing boundary multiplier as the Residual Network; however, it learned the entire solution rather than just learning the LADM discrepancy. Confidence intervals are reported based on resampling the time blocks so that there was still spatial correlation present at the sampled time level.

Table 3. Held-out accuracy and time-block bootstrap uncertainty

Method	RMSE	Bootstrap 95% CI	MAE	Maximum error
Fixed LADM M=2	2.749e-03	[2.476e-03, 3.044e-03]	1.500e-03	2.202e-02
Fixed LADM M=4	1.917e-03	[1.788e-03, 2.026e-03]	1.273e-03	9.138e-03
Product integration	1.875e-03	[1.769e-03, 1.989e-03]	1.264e-03	6.267e-03
ANN only	1.200e-03	[1.121e-03, 1.288e-03]	7.196e-04	9.187e-03
Adaptive LADM	2.108e-03	[2.000e-03, 2.217e-03]	1.577e-03	6.823e-03
Hybrid LADM-ANN	3.396e-04	[3.099e-04, 3.725e-04]	1.739e-04	2.717e-03

The fixed-order LADM configurations M = 2 and M = 4 are included as representative reference approximations rather than as solutions of a global truncation-order optimization problem. Consequently, the performance differences reported

© 2026 NSP
Natural Sciences Publishing Cor.

in Table 3 quantify improvements relative to these specific configurations and should not be interpreted as demonstrating superiority over every admissible fixed-order LADM approximation.

Notably, fixed LADM with M = 4 achieves a lower RMSE than adaptive LADM in this experiment, indicating that the adaptive truncation mechanism does not necessarily minimize prediction error independently of the neural correction. The adaptive procedure is instead designed to select an admissible decomposition order according to prescribed truncation and governing-equation residual criteria. The hybrid method achieves the lowest RMSE among the evaluated approaches, including the independent product-integration and ANN-only comparators. Nevertheless, a fully optimized baseline comparison would require selecting the fixed truncation order using validation observations and subsequently evaluating the selected configuration on an independent test set. Such a comparison should additionally account for computational cost, since increasing the fixed truncation order may improve accuracy while requiring additional Adomian components and residual evaluations. Therefore, the present results establish the relative accuracy of the tested methods but do not provide a complete assessment of optimal accuracy–cost tradeoffs. The 83.9% reduction in RMSE achieved by the hybrid LADM–ANN approximation should be interpreted in relation to the consistency properties of the adopted discrete fractional operators. For Example, 4.1, substituting the manufactured exact solution into the discretized governing equation produces a maximum discrete residual of 2.960×10^{-3}. This quantity measures the consistency defect of the discrete operator on the prescribed grid and is distinct from the solution RMSE reported in Table 3. The hybrid method achieves an RMSE of 3.396×10^{-4}, compared with 2.108×10^{-3} for adaptive LADM. Although both solution errors are smaller than the maximum discrete consistency residual, this does not imply that the numerical solution violates a discretization accuracy limit, because the residual and solution-error metrics are not directly equivalent. The reported improvement therefore demonstrates that the neural correction reduces the discrepancy from the manufactured exact solution on the investigated grid. It does not establish a corresponding reduction in the truncation error of the discrete fractional operators. Demonstrating improved discretization accuracy would require mesh-refinement experiments and separate convergence analyses of the discrete residual and solution error.

4.6 Example 4.2: manufactured-solution verification with a weakly singular kernel

The Polynomial Benchmark from [14,15] was retained to allow for an absolute check on each operator involved in the nonlinear fractional order partial integro differential equation. Using the "method of manufactured solutions" [59], an exact solution for the field is initially assumed and then the source term is found directly through substitution with that solution. This allows the use of an exact (unverified) fractional order curve as a verified reference solution.

4.6.1 Corrected mathematical formulation

For $0 < \alpha \leq 1$, consider

$$ {}^{C}D_t^{\alpha} v(x,t) + v(x,t)v_x(x,t) = \int_0^t (t-s)^{\alpha-1} v_{xx}(x,s)ds + f_{\alpha}(x,t) \qquad 0 < \alpha \leq 1 \tag{62} $$

The new notation for the memory term has changed the variable of integration to be s. It changes the former notation $v_{xx}(x,t)$ from within an integral, where the previous notation was only mathematically correct as $v_{xx}(x,s)$. it also provides consistency for the source term. The Initial and Boundary Conditions are:

$$ v(x,0) = \phi(x) = x^2(1-x)^2, $$

$$ v(0,t) = v(1,t) = 0, \quad (x,t) \in [0,1] \times [0,2] \tag{63} $$

Choose the manufactured solution

$$ v_{ex}(x,t) = \left(1 + t^{5/2}\right) \phi(x), \ \phi(x) = x^2(1-x)^2. \tag{64} $$

It should be noted that unlike the prior statement which indicated the exact solution would be valid only at $\alpha = 1$, equation (64) will provide a valid exact solution for all values of $0 < \alpha \leq 1$ given that the source term is defined by

$$ f_{\alpha}(x,t) = \frac{\Gamma(7/2)}{\Gamma(7/2-\alpha)} t^{5/2-\alpha} \phi(x) + \left(1 + t^{5/2}\right)^2 \phi(x)\phi'(x) - \left[\frac{t^{\alpha}}{\alpha} + \frac{\Gamma(7/2)\Gamma(\alpha)}{\Gamma(7/2+\alpha)} t^{5/2+\alpha}\right] \phi''(x) \tag{65} $$

Where: $\phi'(x) = 2x(1-x)(1-2x)$
$\phi''(x) = 2\left(6x^2 - 6x + 1\right)$

© 2026 NSP
Natural Sciences Publishing Cor.

The direct substitution shows that Equations (63)-(65) have been verified; they can be broken down into three terms for each equation: one term is the Caputo derivative, one term is vv_x, and the third term in brackets is the weakly-singular memory term. Therefore, these test cases represent the whole fractional model as opposed to only the limiting case when there are no fractional order terms.

4.6.2 Resolvent LADM, adaptive truncation, and neural correction

Using the Riemann-Liouville fractional derivative of order alpha on equation (63), we get an alternate Volterra form:

$$v(x,t) = \phi(x) + \mathscr{I}_t^{\alpha}\left[f_{\alpha}(x,t) - v(x,t)v_x(x,t)\right] + \Gamma(\alpha)\mathscr{I}_t^{2\alpha}v_{xx}(x,t),$$

$$\mathscr{I}_t^{\mu}w(t) = \frac{1}{\Gamma(\mu)}\int_0^t (t-s)^{\mu-1}w(s)ds. \tag{66}$$

As repeatedly differentiating can amplify high frequency spatial errors, only the non-linear part (Burgers' terms) is decomposed while the linear memory operator is placed within a resolvent. This results in an LADM recursion.

$$\mathscr{L}_{\alpha}w = w - \Gamma(\alpha)\mathscr{I}_t^{2\alpha}w_{xx},\ \ S_M = \sum_{m=0}^{M} v_m,$$

$$\mathscr{L}_{\alpha}v_0 = \phi + \mathscr{I}_t^{\alpha}f_{\alpha},$$

$$\mathscr{L}_{\alpha}v_{m+1} = -\mathscr{I}_t^{\alpha}A_m,\ A_m\ = \sum_{j=0}^{m} v_j\partial_x v_{m-j}. \tag{67}$$

At each time level, a computable geometric-tail estimate is formed from two successive components:

$$q_M(t)\ = \frac{\|v_M(\cdot,t)\|_{\infty}}{\max\{\|v_{M-1}(\cdot,t)\|_{\infty},\varepsilon\}}$$

$$E_M(t)\ = \frac{q_M(t)}{1-q_M(t)}\,\|v_M(\cdot,t)\|_{\infty},\ q_M(t) < 1\,. \tag{68}$$

The smallest acceptable order of truncation is determined by whether or not both the tail estimate and the full FPIDE residual meet their respective tolerances:

$$M^*(t) = min\{M : E_M(t) \le\ \text{atoll}\ + rtol\,\|S_M(\cdot,t)\|_{\infty}\ ,$$

$$\|\mathscr{R}(S_M)(\cdot,t)\|_{\infty} \le \tau_R\}\,. \tag{69}$$

Therefore, a small last component does not pass acceptance for the governing equation if it still remains unsatisfied. For the second manufactured-solution benchmark, the hybrid approximation follows the residual-correction formulation introduced in Section 3.4 and is expressed as follows:

$$v_H(x,t) = S_{M^*(t)}(x,t) + \psi(x,t)r_{\theta}(\mathbf{z})\,,$$

$$\psi(x,t) = 4x(1-x)\left(1-e^{-\frac{t}{0.05}}\right). \tag{70}$$

The exact form of the ANN correction to preserve the homogeneous and initial conditions is due to the fact that $\psi(0,t)\ =\ \psi(1,t)\ =\ \psi(x,0) = 0$. The four tanh architectures were chosen based on their validation RMSE, which is why (16,8) was kept as the best architecture. This can only occur with an exact difference label for manufactured verification problems. In real applications, they will have to be replaced either by measurements or by higher fidelity references. The residual ANN of this test utilizes a two-hidden layer structure (16,8) as determined based on the evaluation of the candidate tanh networks utilizing RMSE for validation. The goal of this particular structure was to represent the differences from

© 2026 NSP
Natural Sciences Publishing Cor.

the solution provided by the adaptive resolvent-LADM method, but not to provide an approximation of all solutions to the full nonlinear problem. Therefore, the selection of the current structure is primarily a function of both the residual learning paradigm and the previously seen validation results within the tested architecture space. Nonetheless, it cannot be concluded that additional hidden layers or larger widths may not be beneficial; the alternatives have simply been neither tested nor compared.

4.6.3 Reproducible experimental protocol

The computations use a time interval two hundred times larger than those used to create the previous plots over [0,0.01]. For all the fractional terms (the integral terms), product integration will be applied; the L1 formula will compute the Caputo derivative (diagnostic) with respect to time, and second order central difference schemes will be utilized for spatial derivatives. As such, the residual tolerance has been set to approximately one and a half orders of magnitude greater than the discrete residual value of one point five four eight e negative two for the exact field so as to avoid requiring the solver to reach accuracy lower than what the grid resolution can provide.

Table 4. Reproducibility configuration for revised Example 4.2

Component	Setting
Fractional order and domain	α = 0.75; (x,t) in [0,1] x [0,2]
Space-time grid	Nx = 81, Nt = 801; dx = 0.0125, dt = 0.0025
Adaptive tolerances	atol = 1e-5, rtol = 2e-4, residual tolerance = 1.8e-2
Maximum truncation index	Mmax = 9
Selected ANN	tanh network with hidden layers (16,8)
Data split	14,000 training; 4,500 validations; 16,000 held-out points
Uncertainty analysis	400 paired time-block bootstrap resamples
Random seed	20260930

4.6.4 Adaptive behavior on the extended time interval

The ability to determine if a predetermined truncation order is valid throughout an extended time period $0 \leq t \leq 2$, has been analyzed using the adaptive resolvent-LADM method for representative times during the early, mid-point and last stages of the simulations. The smallest admissible order M^* that will be used in determining the next step at every evaluation time level, is determined solely based on whether or not both of the stopping conditions are met: i) the empirical geometric-tail indicator for the current approximation is less than the given absolute-relative tolerance and ii) the full FPIDE residual is less than $\tau_R = 1.8 \times 10^{-2}$. This residual value was chosen as being slightly larger than the discrete consistency floor of the manufactured exact solution, so that the algorithm would not require accuracy better than what could be provided with the existing spatial and temporal discretization. Table 5 presents the selected order, empirical geometric-tail indicator, FPIDE residual and maximum absolute error of the adaptive LADM and hybrid LADM-ANN approximations for the time levels where either stopping condition was first met.

Table 5. Adaptive order, stopping diagnostics, and maximum absolute errors

t	M*	Tail estimate	FPIDE residual	Adaptive error	Hybrid error
0.10	1	7.538e-06	4.118e-03	3.505e-04	3.252e-05
0.50	1	5.261e-07	1.966e-03	1.160e-04	7.486e-06
1.00	1	7.296e-06	1.417e-03	8.810e-05	5.778e-06
1.50	1	5.239e-05	7.165e-03	3.860e-04	9.596e-05
2.00	3	6.205e-06	1.576e-02	2.437e-04	3.929e-05

The chosen order varies between 1 and 3 as it has a mean value of 1.297; the maximum $M_{\max} = 9$ was never attained. Only one non-linear correction was required for most of this range. The rules increased the order in two stages - first to M = 2 and then M = 3 close to the end of the time period.

© 2026 NSP
Natural Sciences Publishing Cor.

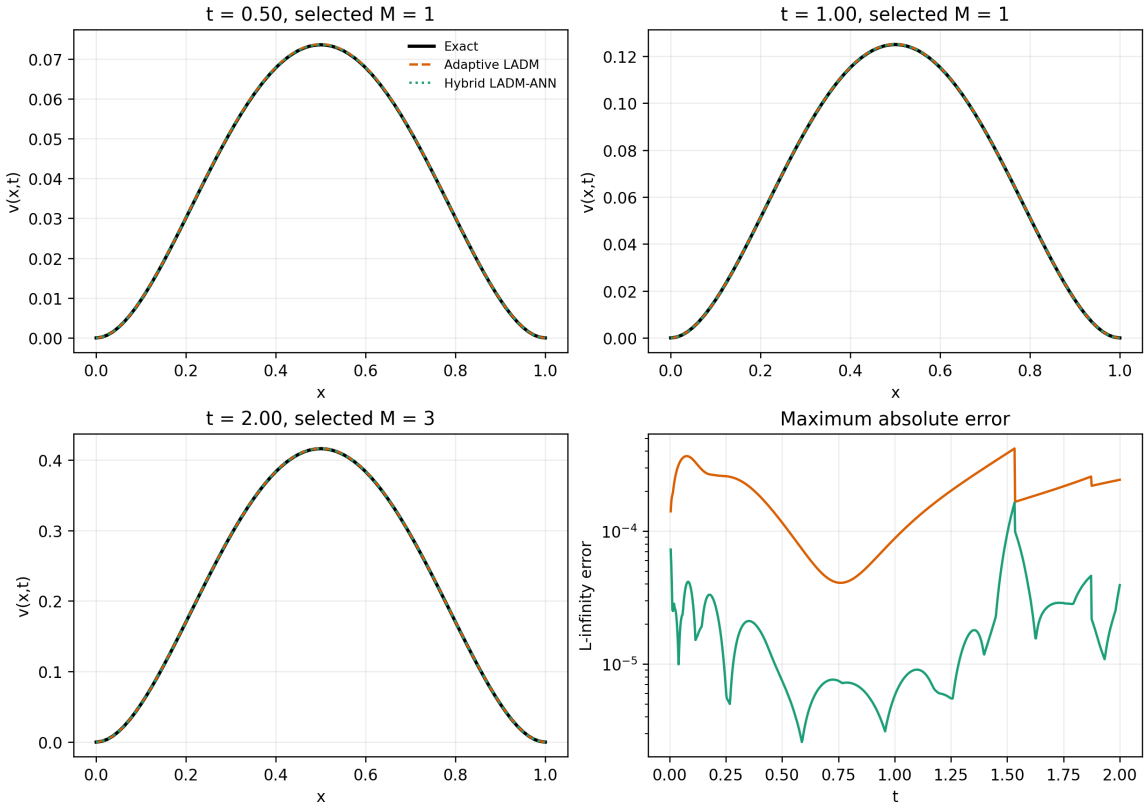


Fig. 4. Numerical solution profiles and approximation errors for Example 4.2 over the extended time interval. (a) Comparison of the manufactured exact solution, adaptive resolvent-LADM approximation, and hybrid LADM–ANN approximation at representative time levels. (b) Corresponding absolute errors of the adaptive and hybrid approximations. (c) Evolution of the maximum absolute error over the spatial domain. The same graphical conventions as Figure 2 are used to facilitate comparison between the two manufactured-solution experiments.

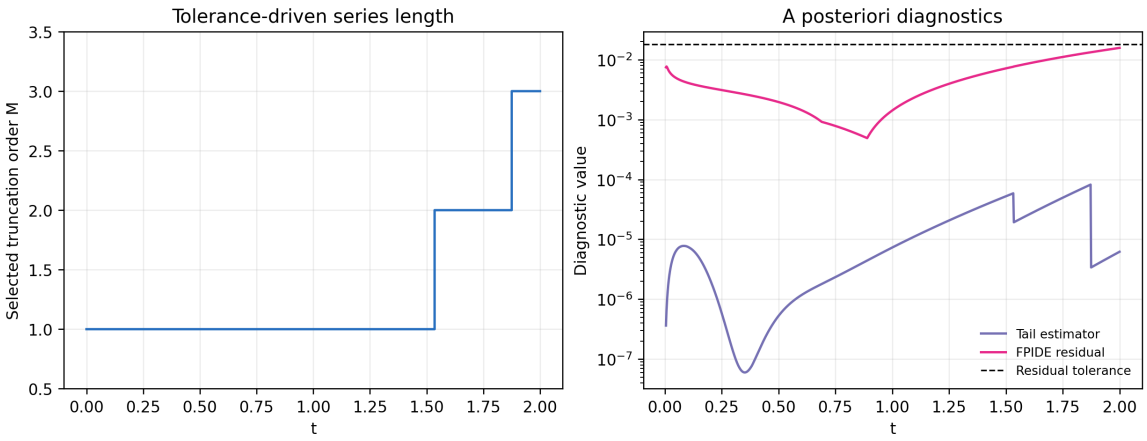


Fig. 5. Tolerance-driven truncation order and a posteriori tail and FPIDE-residual diagnostics for revised Example 4.2.

4.6.5 Comparative accuracy and uncertainty

All methods are tested using the exact same test data. In addition to fixed LADM truncation models (a single independent implicit product integration corrector, a purely neural network-based model), we also compare against a pure ANN

© 2026 NSP
Natural Sciences Publishing Cor.

model that uses the same splits in training/test sets; optimizers; candidate architectures; and multipliers of all the hard conditions as does our final residual network. Our confidence interval around each method's performance is estimated by bootstrapping over completely separate time-blocks and therefore retains spatial correlation for each time-step.

Table 6. Held-out accuracy and time-block bootstrap uncertainty

Method	RMSE	Bootstrap 95% CI	MAE	Maximum error
Fixed LADM M=2	1.220e-04	[1.181e-04, 1.257e-04]	9.666e-05	3.453e-04
Fixed LADM M=4	1.210e-04	[1.170e-04, 1.247e-04]	9.648e-05	3.453e-04
Product integration	1.210e-04	[1.173e-04, 1.246e-04]	9.648e-05	3.453e-04
ANN only	1.245e-04	[1.158e-04, 1.329e-04]	8.318e-05	1.802e-03
Adaptive LADM	1.356e-04	[1.321e-04, 1.390e-04]	1.068e-04	4.152e-04
Hybrid LADM-ANN	1.402e-05	[1.273e-05, 1.552e-05]	7.740e-06	1.581e-04

The results for the hybrid LADM – ANN framework provided in Table 6 are shown to have a significantly lower root mean square error (RMSE) than both the fixed-order LADM, product integration, ANN-only, and adaptive LADM models used for analysis purposes in this study. Although the fixed order M = 2 and M = 4 were determined to be non-optimal; however, there is also an indication of low sensitivity to these two truncation orders due to the minimal difference between the RMSE values for the M = 2 and M = 4. Thus, while it has been demonstrated that another fixed order, a different numerical tolerance setting, or possibly a more favorable computing environment may provide better baseline performance levels. Therefore, before making an absolute determination of the value added from using a hybrid approach, a systematic search based upon validation must be performed over all possible fixed order selections.

Moreover, the effects of adaptive truncation and residual neural correction should be distinguished through separate ablation experiments. This would clarify whether the observed improvements arise predominantly from the adaptive decomposition, the residual learning component, or their interaction. Accordingly, the present findings should be interpreted as evidence of improved accuracy relative to the reported baseline configurations rather than as a general optimality result for the proposed hybrid method. For Example, 4.2, the hybrid LADM–ANN approximation reduces RMSE from 1.356×10^{-4} for adaptive LADM to 1.402×10^{-5}, corresponding to an improvement of approximately 89.7%. However, this improvement should be distinguished from the consistency accuracy of the underlying spatial and temporal discretization. As described in Section 4.6.3, the discrete governing operator evaluated at the manufactured exact solution produces a maximum residual of 1.548×10^{-2}. This residual was used to guide the selection of the governing-equation stopping tolerance of 1.8×10^{-2}. It is not a lower bound on the attainable solution RMSE. The smaller hybrid RMSE indicates improved agreement with the manufactured exact solution at the evaluated points, but does not demonstrate that the fractional discretization error has been eliminated. The residual ANN may correct components of the approximation error while the discrete operator retains a nonzero consistency defect. Accordingly, the 89.7% improvement should be interpreted as a relative reduction in solution error under the fixed experimental discretization. A mesh-refinement analysis would be necessary to determine whether this improvement persists as the spatial and temporal resolutions increase and whether the hybrid approximation preserves the expected convergence behavior of the underlying numerical scheme. The paired time-block bootstrap results indicate that the observed reduction in RMSE is reasonably stable under the adopted resampling procedure. Nevertheless, this statistical evidence does not establish that the hybrid architecture is universally superior to optimally tuned alternatives. Further validation-tuned baseline comparisons and ablation experiments are required to isolate the contributions of adaptive truncation and neural residual correction.

4.6.6 Fractional-order sensitivity

To analyze how the manufactured-solution experiment is sensitive to the fractional-order parameter alpha within the adaptive resolvent-LADM method, experiments were performed with $\alpha = 0.50, 0.75$, and 0.95, as the spatial mesh, time-step mesh, truncation tolerance values and the maximum allowable order of the truncated expansion are all constant across these experiments. The source function $f_\alpha(x,t)$ has been computed anew from equation (46), such that the exact solution specified by equation (45) remains unaffected. This study therefore examines numerically the sensitivity of the fractional operator, the adaptive truncation strategy and the discrete approximation accuracy as opposed to the sensitivity of a fixed forcing problem due to physically varying the forcing. In table 7 are reported the discrete residual floor, range and mean

© 2026 NSP
Natural Sciences Publishing Cor.

of the selected truncation index M^*, frequency with which $M_{\max}$ was selected; global root-mean-square-error (RMSE); and the maximum absolute error over the entire domain at t=2 for each of the three fractional orders tested.

Table 7. Algorithmic sensitivity to the fractional order

α	Residual floor	M* range	Mean M*	At cap	Global RMSE	Error at t=2
0.50	2.460e-02	1–9	1.040	0.50%	1.686e-04	6.419e-04
0.75	1.548e-02	1–3	1.297	0.00%	1.367e-04	2.437e-04
0.95	1.207e-02	1–3	1.613	0.00%	1.310e-04	2.757e-04

The purpose of independence of αin equation (50), whereas fα varies with α; Table 12 assesses numerical robustness rather than the physical sensitivity to an α-fixed forcing system. The accuracy remains around 1e-4 overall tested orders. The singular activation of $M_{\max}$ when $\alpha = 0.50$ (which corresponds to 0.5 percent of the time steps) indicates proximity to the fractional singularity of startup which will require a time step grading for uniform, low-order cost at small α.

4.6.7 Interpretation and limitations

The example has been modified to support retention of the benchmark on three grounds. First, the manufactured source validates all aspects (Caputo derivative, nonlinear convection, weakly singular memory term) at once. Second, the adaptive rule records instances where further components of LADM have to be added over a time period. Third, residual learning produces an improved physics-based approximation without sacrificing fidelity to the original prescribed data and producing better results than each individual component separately. Tail estimates bound only the portion of LADM that remains un-resolved under the conditions of the local ratio and do not account for errors associated with spatial, temporal or neural approximations. Therefore, why the error recorded in Table 10 is larger than the tail estimate. As well, the manufactured experiment was a verification test of the code and therefore cannot be represented as validation using field measurements. Any unknown parameter identification and NGSIM based validation should continue to be independent tests.

5 Real-data application to NGSIM US-101

5.1 Data source and preprocessing

The actual data used in this research was derived from vehicle movement data from the south-bound US-101 National Geospatial Intelligence Agency (NGSIM) study conducted by the United States Department of Transportation's Federal Highway Administration. The nominal time frame for the dataset is 08:05 to 08:20. The original timestamped dataset consisted of 788.6 seconds of data. After processing the original timestamped data into a processed dataset that included only the 5 mainline lanes where vehicles could travel at plausible speeds through the same study area as defined above, there were still 999,061 individual points remaining in the dataset. However, since we chose to process our data using a regular spatial grid, we only needed to use 785 seconds worth of data. The study's authors also processed the coordinate values of each trajectory point using a Savitz sky-Golay filter with a length of 21 points [64] prior to further analysis. Reporting these details are important because the numerical derivative can amplify the inherent noise within an observed trajectory.

Table 8. NGSIM US-101 data and preprocessing summary

Quantity	Value
Retained trajectory records	999,061
Observation duration	785 s
Temporal aggregation	5 s
Spatial aggregation	20 m
Regular grid	158-time bins x 32 spatial cells
Mean speed	9.431 m/s
Mean density	40.961 vehicles/km/lane
Empty grid cells before interpolation	12.84%

© 2026 NSP
Natural Sciences Publishing Cor.

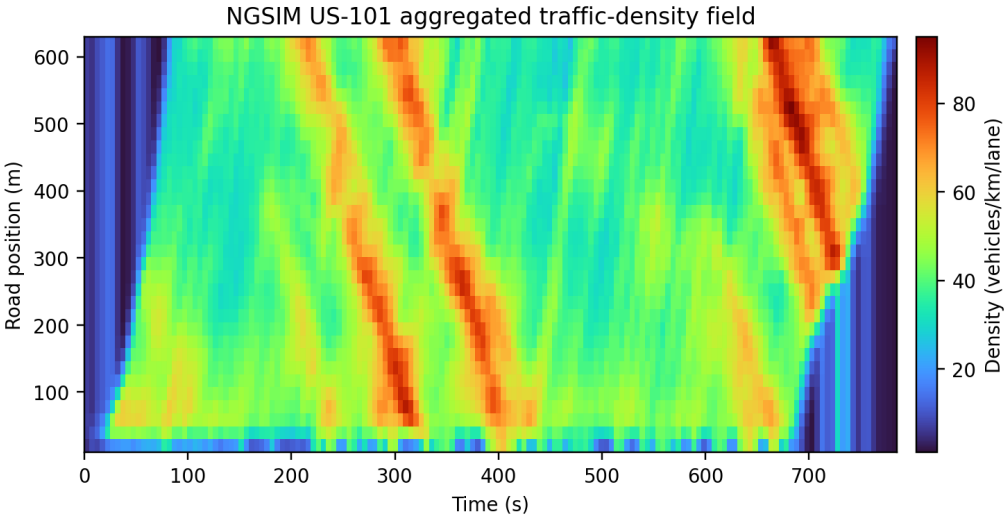


Fig. 6. Aggregated NGSIM US-101 traffic-density field over the 785 s observation window.

5.2 Fractional traffic-density model

Traffic density was selected as the conserved field $u(x,t)$ since it is a better representative of traffic flow than speed. Speed, and therefore its spatial derivative, were incorporated into the model as sources. Although this does make more sense when considering how traffic conserves itself, this also means that speed is not being represented as the conserved field. Both the nonlinear diffusion term a and the memory-diffusion terms kappa are both defined as positive. As stated earlier, the inlets and outlets are made to be homogenous through equation (28) and time is split chronologically so 60 percent of all time steps will be used for estimating the fields and training the model, 20 percent will be used for validating and calibrating intervals, and 20 percent will be used for testing. A random row split has not been included in order to prevent future traffic regimes from being placed within the training set resulting in an overly optimistic estimate of error. The random seed is 20260924. For each 5 s x 20 m cell, density equals the number of trajectory observations that have been retained divided by the total number of frames observed, multiplied by five lanes and divided by the 0.02 km cell length. Missing cells will be first filled using interpolation in space and then time; subsequently a Gaussian filter with standard deviation equal to 0.65-time bins and 0.8 spatial cells will be applied. Density, speed, space, and time were made nondimensionalized using values of 100 vehicles/km/lane, 30 m/s, 640 m, and 785 s, respectively. All components after the first LADM component of the model will utilize the Frechet action on the current field of the complete nonlinear residual of the previous component. The Frechet action is defined here as the discrete form of the Adomian recurrence relation developed specifically for this data case. The data collected in the NGSIM data set can be subject to both sharp spatial gradient and abrupt transition due to wave forming and propagating through traffic flow. Temporal and spatial aggregations, smoothing of trajectories, interpolating gaps within the data, and using Gaussian filter are all used to smooth the raw data and help make it easier for derivatives to be calculated from the data. However, while this process can be helpful in terms of reducing some of the errors that are inherent in collecting observational data as well as making calculations smoother (less noisy), there is a possibility that the physical sharp transitions will be lessened. Therefore, the resultant traffic-density field should not be considered as a continuous or independently validated representation of the traffic-flow dynamics but rather as a smoothed version of the original. As such, the focus of the current study is evaluating predictive accuracy based upon a smoothed traffic-density field; we do not provide evidence of whether or not the hybrid method is capable of resolving true shock discontinuities.

5.3 Estimated parameters and identifiability

The unknown fractional and physical parameters were determined during the training phase of a chronological sequence by minimizing the normalised discrete FPIDE residual. The Caputo order alpha and kernel order beta were searched through the defined ranges $\alpha \in [0.55, 1.00]$ and $\beta \in [0.20, 1.00]$. While the non-linear transport coefficient 'a' and memory diffusion coefficient $'\kappa'$, were also estimated based on physical constraints of $a \geq 0$ *and* $\kappa \geq 0$. For each candidate

© 2026 NSP
Natural Sciences Publishing Cor.

combination of alpha and beta the nuisance coefficients representing the source term in the model equation were re-estimated via constrained least-squares. The parameter uncertainty and practical identifiability were evaluated using a moving block bootstrap which sampled consecutive time blocks and performed a full estimation process for each sample set which included estimating the fractional-orders. Estimates and 95th percentiles are shown in table 9.

Table 9. Estimated physical and fractional parameters

Parameter	Estimate	Block-bootstrap 95% interval
Caputo order alpha	0.750	[0.550, 1.000]
Kernel order beta	1.000	[0.200, 1.000]
Nonlinear transport a	0.5549	[0.1078, 2.8961]
Memory diffusion kappa	0.00504	[0.00000, 0.03225]

A semi-automated block-bootstrap approach was employed to assess the sensitivity of the proposed semi-discrete fractional-memory model to uncertainty in the calibration parameters. As can be seen from Table 9, the fractional-memory parameters are subject to considerable uncertainty. For example, for the Caputo order \alpha=0.750, the 95% confidence interval is found to be [0.550, 1.000]. Similarly, the memory-kernel order \beta=1.000 has a much larger interval of [0.200, 1.000]. Additionally, the estimated memory coefficient \kappa=0.00504, had an interval of [0.00000, 0.03225], indicating that zero could reasonably belong to this interval. All of this suggests that there is very little practical identifiability of the fractional-memory structure given the available data. Given the potential for the Caputo derivative and the weakly-singular memory operator to have some degree of complementary effects in terms of their contribution to the overall form of the governing equation, we believe that different combinations of α, β and k can result in similar residuals, and therefore similar short-term time-trajectories. One possible explanation for the uncertainty shown here is parameter compensation. We note however that establishing this formally will require additional work that involves analyzing the statistical dependence among the individual parameters or profile of parameters. Establishing whether the contraction inequalities presented in Sec. 3.2 hold practically for the NGSIM application is further complicated by the large uncertainties in both the physically-calibrated and fractionally-calibrated parameters. The block-bootstrap results for the various parameters show relatively large uncertainty intervals around the estimates obtained from fitting to the NGSIM data. As noted previously in Eq. (19), the contraction factor will depend upon many of the same parameters used to obtain those estimates. Thus, if one were able to verify that the contraction inequality held true using a single set of fitted parameters, it cannot be guaranteed that it holds true throughout the corresponding uncertainty region. Rather, an evaluation of the self-mapping and contraction bounds across an admissible parameter space (constructed from the results of the fit), while also accounting for how those bounds depend on both the selected solution radius and discretization size, is needed to gain a better understanding of what can be expected. An upper-bound for the contraction factor that is conservatively based on theory should be clearly differentiated from an empirically-determined ratio of successive LADM components, where such ratios might suggest good numerical properties for a given simulation, yet fail to support the underlying assumptions necessary for establishing global convergence via the fixed-point theorem. While some information about the performance of the simulations has been provided through use of adaptive truncation diagnostics, residuals of the governing equations, uncertainties in model parameters and forecasting performance, none of that information establishes independent evidence supporting theoretically-established upper-bounds for each quantity appearing in Eq. (19). The profile estimate is $\alpha = 0.75$ and $\beta = 1.00$. The bootstrap interval for alpha spans 0.55-1.00, while the interval for beta spans 0.20-1.00. The interval for kappa reaches the nonnegative boundary. Consequently, the data support fractional memory but do not identify β or κ sharply. These intervals are more informative than reporting point estimates alone and prevent an unjustified claim of precise physical identification.

5.3.1 Necessity and Identifiability of the Additional Fractional-Memory Term

In terms of the amount of information contained in the data used to estimate the parameters shown in Table 9, we can see that there are some questions about whether it would be worthwhile to include an additional fractional-memory operator in our description of the NGSIM US-101 traffic dynamics. We have $\beta = 1.000$ as the estimated kernel-order and β has been estimated to lie at the top end of the search space. The estimated memory coefficient $\kappa = 0.00504$ has a 95% block bootstrap confidence interval of [0.00000, 0.03225], indicating that there is very little information available in the observations regarding how large (in either direction) the additional memory contribution might be. As such, when interpreting these estimates, we need to distinguish between the two types of memory involved in the governing FPIDE.

© 2026 NSP
Natural Sciences Publishing Cor.

The Caputo derivative of order α provides hereditary properties to the time-evolution of the traffic density field. In addition, the second type of memory involved is due to the spatial operator. If $\beta = 1.000$ then the former becomes simply a regular time-integral. On the other hand, setting $\kappa = 0$ removes the second type of memory contribution entirely; therefore, β will no longer be estimable. Thus, we do not interpret the estimated value of $\beta = 1.000$ as being evidence that our overall model is now nonfractional. Rather, we note that even if we were to remove both types of memory by setting both $\kappa = 0$ and $\alpha = 1.000$, the model would retain fractional temporal-dynamics.

On the other hand, given that we have a relatively small value for κ and that its uncertainty interval reaches zero, it is possible that the additional memory contribution is not needed to describe our current observations. This issue is especially pertinent here since the memory contributions due to the Caputo derivative and the Volterra memory term may exhibit partially compensatory behavior; hence, identifying each contribution independently may require multiple observation datasets. To determine whether or not κ contributes effectively to a physically meaningful quantity (i.e., memory), one needs to consider both the nondimensionalized form of the governing equation as well as the spatial operator that accompanies κ, along with the length of the time-interval over which temporal integrals are evaluated. Therefore, to make a definitive determination regarding whether or not the additional memory term is needed to adequately describe the investigated NGSIM traffic data-set, it is necessary to perform a comparative study using nested models where each model formulation contains fewer parameters than the previous one. Specifically, one needs to fix β and/or set $\kappa = 0$ and re-estimate all remaining parameters within each model and compare them using identical chronological training, validation and testing procedures. Until this nested-model comparison is performed, we regard the additional memory-term as an admissible theoretical component whose empirical necessity has yet to be confirmed for this particular NGSIM data-set.

5.4 Comparative forecasting results

The forecasting performance was compared across all competing models during the last 20 percent (chronological) testing period. This last 20 percent was completely independent from the remaining 80 percent which contained the data for all prior model development activities. All the models tested here were developed using the same time and spatial intervals for training, cross-validation, and testing. Each model is being compared at a one step ahead forecast horizon and they are each forecasted to the same space-time grid. Models include both an Adaptive Learning Density Model (LADM), and an adaptive version of this model as two interpretable baseline models. Additionally, three conventional Machine Learning algorithms including Ridge Regression, Support Vector Regression with an RBF Kernel, Extra Trees and Histogram Gradient Boosting were also applied. Also included are a fully connected Artificial Neural Network (ANN) model that forecasts the full density field (i.e., both speed and density) and a Hybrid Model that forecasts only the residual difference of an adaptive LADM model. The hyper-parameters of all models were determined by selecting the best performing model based on the results obtained during the chronological validation set. A summary of the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination R^2 for each model's forecasts is reported in Table 10 in units of vehicle / km / lane. Smaller values of RMSE and MAE and larger values of R^2 indicate superior forecasting performance.

Table 10. Five-second test performance in vehicles/km/lane

Method	RMSE	MAE	$\mathbf{R^2}$
Persistence	6.753	5.034	0.9291
Adaptive LADM	6.630	4.907	0.9316
Ridge	6.151	4.396	0.9412
SVR-RBF	29.950	24.255	-0.3951
Extra Trees	7.630	5.231	0.9094
Hist Gradient Boosting	6.937	4.714	0.9252
ANN only	16.190	11.411	0.5923
Hybrid LADM-ANN	6.149	4.395	0.9412

© 2026 NSP
Natural Sciences Publishing Cor.

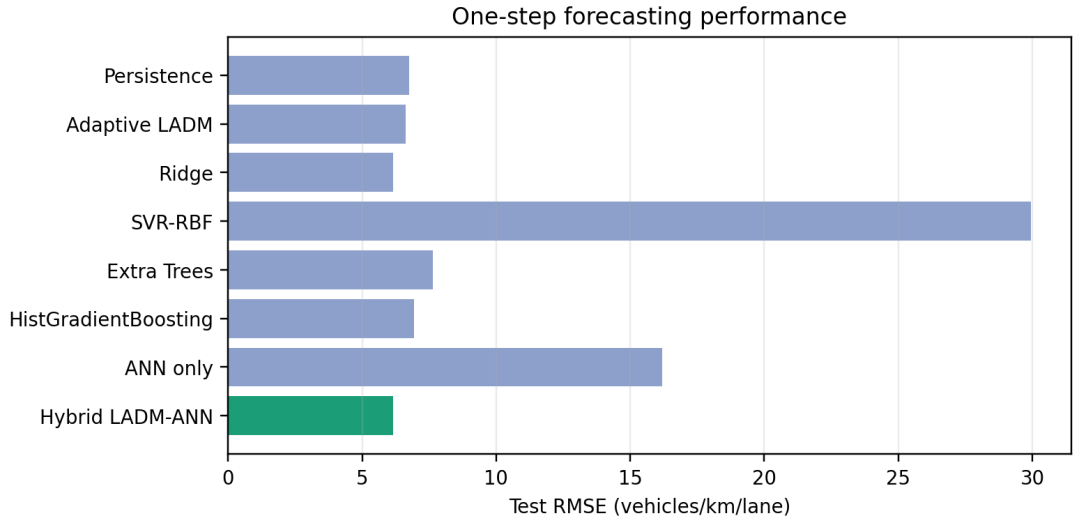


Fig. 7. Five-second forecasting RMSE across physical and data-driven baselines.

The Hybrid Model improves the RMSE of Adaptive LADM, from 6.630 vehicles / km / lane to 6.149 vehicles / km / lane (a 7.3% improvement), and performs significantly better than an ANN only model. There was essentially no difference in performance of the Hybrid model versus Ridge Regression; this suggests that the choice of residual correction made using cross-validation, was primarily linear. An additional test was run on a Deeper Nonlinear ANN however it fit the dynamic traffic pattern too well. Therefore, this Shallow Identity Activation Residual Network was retained.

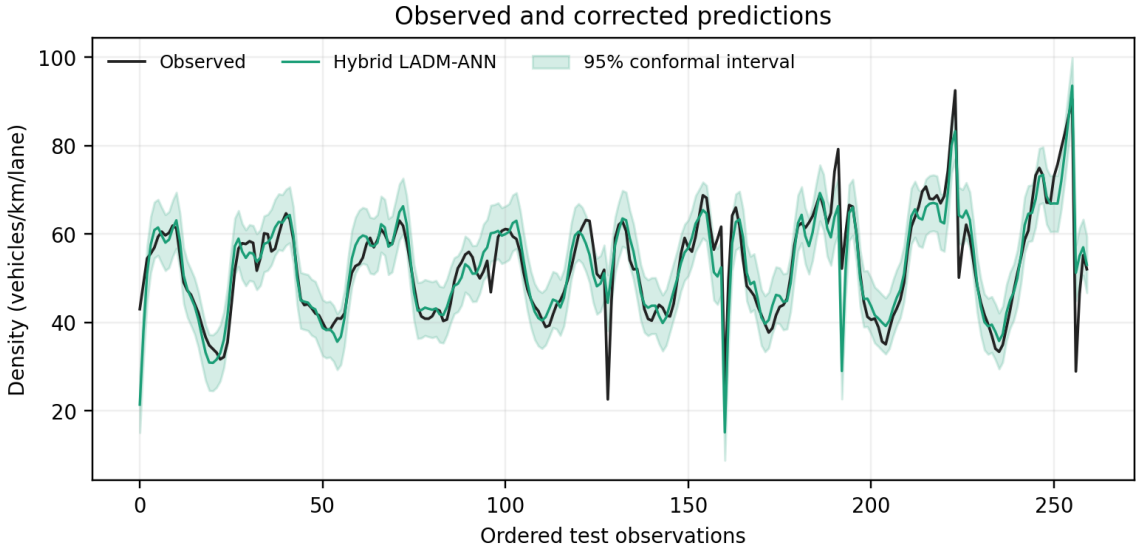


Fig. 8. Observed density, hybrid prediction, and validation-calibrated conformal interval on ordered test observations.

5.5 Adaptive truncation and long-time behavior

To assess if the proposed method's near-term results persist when increasing the length of the prediction time frame (horizon), a study on how accurately the proposed models can predict future values (forecasting) was performed for intervals of 5, 25, 50, and 100 seconds. Forecasting errors were calculated using the same test data set used for all previous horizons. When the hybrid predictor and adaptive LADM were run through their respective algorithms, the amount of history that had been previously collected was maintained rather than discarded. The number of terms to be truncated for

© 2026 NSP
Natural Sciences Publishing Cor.

each forecast was determined independently based on the specified tail criterion and residual criterion. In addition to the hybrid predictor and adaptive LADM, persistence and histogram gradient boosting were also considered as "reference" models so that any increase in model error due to an increase in the size of the predictive model could be distinguished from any differences arising from either statistical or structural model changes. The results of this evaluation are shown in Table 11, which lists the root-mean-square error (RMSE) for each predictive horizon in units of vehicles/km/lane along with the mean and maximum number of terms retained by the adaptive LADM.

Table 11. Long-horizon RMSE and adaptive series length

Horizon s	Persistence	Adaptive LADM	Hybrid	HistGradientBoosting	Mean max terms
5	6.753	6.630	6.149	6.937	2.00 / 2
25	20.924	23.270	22.156	23.617	2.00 / 2
50	30.250	37.757	27.729	25.351	2.01 / 3
100	34.324	54.244	28.158	26.926	2.99 / 3

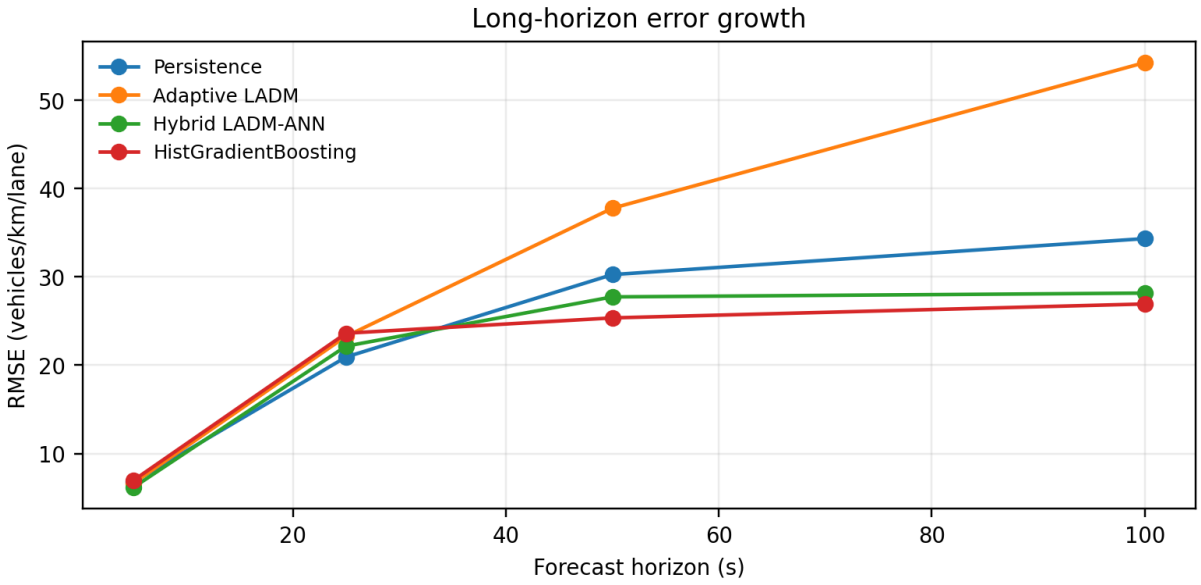


Fig. 9. Growth of traffic-density forecast error with the prediction horizon.

The long-term results from table 11 clearly show a distinction between controlling for error due to discretization versus generalizing (predicting) future states. While the hybrid LADM-ANN approach produces the least RMSE at a 5 second time frame compared to other methods, this doesn't hold true across all increasing intervals. In terms of RMSE value, at a 25 second time frame, persistence performs better than the hybrid model at values of 20.924 and 22.156 vehicles per km per lane, respectively. At 50 seconds and 100 seconds, the hybrid model still has lower RMSE than persistence; however, it is outperformed by Histogram Gradient Boosting. Clearly these findings demonstrate that simply using adaptive truncation and preserving some amount of "fractional" memory will not prevent cumulative errors in structural forecasting when the input conditions change. Thus, we see the proposed framework as providing an accurate and interpretable prediction method over short horizons. For longer horizon predictions additional model adaptations, recalibrations, or representations of how traffic evolves would need to occur. The number of retained terms increase from about 2 (at 5-25 s) to around 3 (at 100 s), indicating that there is no one-size-fits-all application of a two-term fixed rule. However, adding additional series does not stop the model discrepancy error in this problem from increasing with time. From 25 s onward the Hybrid model is less accurate than the Persistence model. Therefore, the proposed approach is validated as an interpretable correction methodology for short horizons and not as a general-purpose forecasting solution for all times.

© 2026 NSP
Natural Sciences Publishing Cor.

5.6 Sensitivity and predictive uncertainty

Sensitivity and Predictive Uncertainty for Evaluating Forecast Variability. Two different types of evaluations were used to evaluate sensitivity and predictive uncertainty of the model. This distinction reflects that the two assessments represent different forms of variability; that is, parameter variability (through the use of sensitivity analysis) vs. variability due to errors in forecasting. The sensitivity analysis involved generating 256 Latin Hypercube Samples (LHSs) by randomly sampling from within $\pm 15\%$ of each of the three sets of calibrated parameters for the variables α, β, a, *and* κ. All LHSs were constrained to ensure they fell within the admissible bounds for these parameters. Following generation of each sample, the fractional operators were recalculated based on this new set of sampled parameters. In addition to recalculation of the fractional operators, the same sampled parameter vectors were also used to generate traffic density forecasts. Then, for each of these sampled parameter vectors, the Root Mean Squared Error (RMSE) was computed and stored. The Spearman Rank Correlation Coefficients (SRCCs), along with their respective p-values, were computed for each parameter and its corresponding RMSE value. This provided an indication of which parameters had monotonic relationships with RMSE over the sampled neighborhood. It should be noted that SRCCs can be viewed as screening metrics as opposed to causal relationships or variance-based sensitivity indexes.

Table 12. Global rank-based sensitivity screening

Parameter	Spearman rank correlation with RMSE	p value
α	-0.9944	6.45e-250
β	0.0643	0.305
a	-0.0745	0.235
kappa	-0.0928	0.139

The table shows that there was a big difference among the calibrated parameters' impact on the forecasting accuracy within the studied area around the parameter's values. For example, the Caputo fractional order α had a very high negative correlation with RMSE (Spearman Rank Correlation Coefficient SRCC = -0.9944) as compared to the other three parameters; i.e., the correlations for β, a, and κ were relatively low (0.0643, -0.0745, and -0.0928, respectively). The results show that α was the most sensitive factor in terms of local sensitivity factors under the study conditions and forecasting metric. However, because the rank correlations for the other three parameters were so low, it does not mean that they have little effect on the governing fractional partial integro-differential equation. Specifically, the nonlinear transport coefficient a and the memory parameters β and κ appear in different parts of the governing operator. This means they could affect the solution of the PIDE via nonlinear interactions or compensatory effects that will not be revealed using rank correlations. Over-parameterization is however suggested by the results shown in Table 9. In fact, the wide bootstrap intervals for the fractional orders and the presence of zero in the interval for κ indicate that the traffic data collected thus far did not sufficiently restrict the estimation of the complete fractional-memory parameterization. As such, although the current results can't say whether all identified parameters are individually identifiable or necessary for predicting future conditions, assessing model parsimony in a mathematical way necessitates comparing the complete formulation to various reduced models where one or several of the parameters are fixed or the added memory term is eliminated. Additionally, once a reduced model has been formulated, all remaining parameters need to be re-calibrated to determine how well each reduced formulation predicts future conditions. Thus, reduced models should be compared to the original formulations based on chronological out-of-sample error predictions, residual errors of the governing equations, parameter stability and computational costs. If a reduced model achieves similar predictive performance as the complete model but at less computational cost and produces more stable estimates for the parameters then the reduced model would be preferred. Therefore, we consider our sensitivity analysis as indication of differences in parameter influence and possibly over-parameterization but not proof of parameter redundancy. An essential task for future research is establishing a minimum number of fractional parameters required to create a practically identifiable model. The weak marginal sensitivities of β and κ also motivate investigating the necessity of including both fractional memory parameters. Even though the marginal rank-based sensitivities of β and κ are weak, they don't demonstrate that either is mathematically redundant since nonlinear interaction and compensating effects between parameters may not be reflected by marginal rank-based analyses. Furthermore, combined with the bootstrap intervals reported in Table 9, this suggests that perhaps the full memory formulation is over-parameterized for the given traffic data set. A more detailed assessment of whether each parameter is needed to accurately describe traffic behavior can be obtained through a comparison of a fully parameterized model with two alternative nested models; i.e., a fixed-order memory model ($\beta = 1$) and a model without a memory component ($\kappa = 0$). It is crucial to note that when $\kappa = 0$, β no longer impacts the PIDE.

© 2026 NSP
Natural Sciences Publishing Cor.

Table 13. Nominal 95% conformal interval diagnostics

Horizon s	Static coverage	Static width	Sequential coverage	Sequential width
5	79.6%	12.65	86.3%	16.81
25	57.5%	30.74	77.8%	48.77
50	41.1%	30.53	65.4%	58.23
100	35.9%	40.53	78.8%	73.12

The most important factor in the sensitivity of the alpha order is the fractional order of alpha itself. Using a rank correlation measure (approximately 0.994), this was found to be true for all tested parameters. In terms of coverage, the nominal static 95% confidence interval captured only 79.6% of the actual 5-second test data. However, expanding the interval using a rolling sequential method increased that coverage to 86.3%. Although the nominal coverage remained lower than desired due to non-exchangeability between the time-based test interval and the time-based calibration interval, it was still sufficient to report as evidence of "traffic regime drift" rather than simply through retroactively increasing the size of the interval on the test data.

5.7 Reproducibility settings

To enable independent reconstruction and traceability of the NGSIM analysis, all the initial conditions (preprocessing, calibration, adaptive solver, residual correction, uncertainty analysis) were set to a single value prior to evaluation on the chronological test period. The values include the time scale and space scale at which the data is aggregated, the method used for smoothing the data, how the data was chronologically divided into training and testing sets, the bounds of the search grid for the fractional orders, constraints applied during coefficient estimation, tolerance parameters used by the Adaptive Least Absolute Deviation Method (LADM), the number of bootstraps and sensitivities, as well as the random seed. Unless otherwise specified, each of the above settings were consistent across all NGSIM forecast horizons as well as with respect to the other methods.

Table 14. Reproducibility configuration

Component	Setting
Data grid	Delta t = 5 s, Delta x = 20 m, five mainline lanes
Spatial smoothing	Gaussian grid smoothing after 21-point trajectory smoothing
Split	Chronological 60% train, 20% validation, 20% test
Fractional search	alpha = 0.55:0.05:1.00, beta = 0.20:0.10:1.00
Coefficient fit	Bounded linear least squares with a >= 0 and kappa >= 0
Adaptive LADM	atol = 1e-5, rtol = 2e-3, Mmax = 6
Residual ANN	One hidden layer, four identity units, L-BFGS, alpha penalty 0.1
Bootstrap and sensitivity	30 time-block replicates and 256 Latin-hypercube samples
Random seed	20260924

6 Results and discussion

The results of the initial manufactured examples demonstrate that LADM will produce approximations to smooth solutions over extremely small time periods. The analysis developed in this paper provides clarity as to when the level of accuracy produced by LADM is maintained for other than the manufactured example; specifically, q(T) must be less than one, the estimated tail and the FPIDE residual must satisfy equation (33), and continuation of the model must maintain the Caputo history. As demonstrated with the NGSIM dataset, there are some difficulties associated with the combination of noisy, nonstationary data. The choice of fixed order for the baselines is another limitation of the numerical evaluation. Although the hybrid LADM-ANN model has achieved significant reduction of errors when compared with manufactured solutions, there was no optimization process to choose two fixed order configurations (M=2 and M=4) that represent typical LADM configurations. Thus, the percent improvement in accuracy will vary depending upon the baseline truncation order used.

Comparison with both adaptive LADM and ANN-only models provide some evidence of the contribution of each component of this new methodology. A more complete assessment would involve developing fixed order approximations that have been tuned to validate the accuracy of the approximation, comparing fixed order LADM models with an equivalent amount of residual neural corrections and evaluating a comparison of computation time versus accuracy to determine if the advantages of the adaptive approach or the residual neural corrections are responsible for the improved

© 2026 NSP
Natural Sciences Publishing Cor.

results. These comparisons will enable a determination of how much of the benefits provided by the new hybrid model can be attributed to the adaptive approach vs. the residual neural corrections. The adaptive LADM provided an interpretable baseline model, while the residual network improved the accuracy of forecasts at short horizons (specifically, within 5 seconds). However, a simple persistence model remained competitive in terms of accuracy, but it was also clear that both long horizon and parameter uncertainties were material. The Sensitivity Analysis shows an extreme disparity in the impact of the Calibrated Parameters on Predictive Accuracy. In terms of RMSE, there is a strong positive relationship (monotone) associated with the Caputo Fractional Order "α" whereas; β, a, and κ have relatively small and negative marginal relationships over the range of parameters examined. Furthermore, along with the large Bootstrap Intervals in Table 9, the above results create some doubt as to whether the complete Fractional-Memory Formulation has sufficient complexity to be fully identified using only the available NGSIM Observations. It is also important to note that weak marginal sensitivity does not necessarily mean that each parameter is redundant since other parameters may still interact non-linearly and compensate for each other's effect. Therefore, the present results should be interpreted cautiously and encourage Nested-Model Comparisons, Joint Parameter Profiling, and Independent Validation to see if a simpler Fractional Formulation can maintain Predictive Accuracy while providing greater parameter Stability. Thus, it would be premature to make Claims of Superiority for all Fractional Orders of Integration, Data Regimes, and Time Intervals. The Model-Specific Contraction Condition developed in Section 3.2 presents a rigorous Sufficient Criterion for the Existence and Uniqueness of the Semi-Discrete Fixed-Point Solution. However, the Practical Application of this Condition to Observation Problems requires Additional Care. In contrast to the Manufactured-Solutions Experiments (where model parameters and reference Solutions are Prescribed), the NGSIM Problem includes Uncertain Parameter Estimates, Noisy Boundary Observations, and Time-Varying Traffic Regimes. As such, the Constants Entering into the Contraction Inequality cannot be Assumed to Remain Uniformly Valid Throughout the Observational Period. In Particular, the Estimated Fractional Orders and Physical Coefficients Influence the Contraction Factor Directly Whereas; the Selected Spatial Discretization Determines the Bounds for the Linear and Nonlinear Operators. Additionally, the Wide Bootstrap Intervals Reported in Table 9 Indicate that Verification Based Only on Point Estimates May Underestimate Parameter Uncertainty.

A significant Limitation of the Observational Analysis is the Empirical Need for the Second Fractional Memory Operator. Although the Complete FPIDE Formulation Permits Separate Caputo and Volterra Memory Mechanisms, the Estimated Kernel Order $\beta = 1.000$ and Small Memory Coefficient $\kappa = 0.00504$ Provide Little Evidence That Both Mechanisms Are Required for the NGSIM US-101 Dataset. The Broad Bootstrap Intervals and Weak Marginal Sensitivity of the Memory Parameters Further Support Practical Identifiability Difficulties.

The positive results of the hybrid LADM–ANN model for short-term forecasting indicate that it is not reasonable to consider these as empirical proof for the need of a second memory term in terms of its physical basis. Improvements could occur through improved numerical approximations of the solution at time $t = n\Delta t$ due to the ANN residual correctors; the ANN residuals can also capture other effects, and/or interact with the additional memory term, if each physical parameter is weakly identifiable. Therefore, a reduced order model which retains the Caputo fractional derivatives, but omits the additional Volterra memory term, presents another scientific meaningful choice. It will be necessary to independently estimate the parameters of both formulations, and make chronological out-of-sample comparisons to determine if the added complexity of the full formulation is warranted. As long as no such studies have been conducted, this study provides support for the use of the presented framework as a viable tool for computational modeling, without providing any indication of the requirement of the additional memory term for this specific data set.

The chosen neural-network architecture introduces an additional layer of uncertainty into the methods used here. For example, in the manufactured-solution experiments, a Res Net model with two hidden layers containing 16 and 8 units respectively, and with hyperbolic tangent activation function was chosen among a limited number of candidates based on cross-validation results using Root Mean Squared Error (RMSE). While this selection resulted in good corrections to the residuals, there were no searches done for systematically deeper networks, nor significantly wider networks. Thus, we cannot conclude that the selected capacity of the neural-network is optimal. Additionally, in our real-data application employing NGSIM observational data, a separate neural correction model having fewer neurons than the one employed in the manufactured solution case was employed. We therefore do not intend to generalize conclusions regarding the performance of the (16,8) architecture to this observational regime. To understand how robust the design decisions made above are with respect to variations in the network structure including depth and width of the layers, activation function types, level of regularization applied to the loss function, and computational resources required by the respective configurations would require a detailed study comparing all possible combinations of those factors.

6.1 Comparative Numerical Accuracy and Adaptive Behavior

The manufactured-solution experiments are designed to test both the ability of the LADM-ANN model to approximate the underlying physics (i.e., the physical accuracy), and how well it adapts to achieve that goal using the least number of

© 2026 NSP
Natural Sciences Publishing Cor.

computational resources (i.e., computational efficiency). In example 4.1, the hybrid method had a root-mean-square-error (RMSE) value of 3.396 x 10^{-4}. As comparison, other methods tested showed an RMSE of 2.108 x 10^-3 for adaptive-LADM, 1.917 x 10^{-3} for fixed-order LADM with four terms, 1.875 x 10^{-3} for product-integration, and 1.200 x 10^{-3} for the single-neural-network (ANN) case. Relative to adaptive-LADM, this represents a RMSE improvement of approximately 83.9%. As expected from the adaptive nature of LADM, the decomposition order increased as the solution evolved toward the end of the studied time interval. This trend demonstrates one reason why adaptive truncation is useful; i.e., there will be instances where a fixed number of Adomian components will be too costly at certain points during a calculation but not enough elsewhere. Although the numerical results clearly indicate the difference between achieving a predetermined level of accuracy versus minimizing the actual RMSE, they also highlight another aspect of adaptive truncation. Although the adaptive strategy controls the decomposition based upon the user's defined tolerance criteria, it does not guarantee that this will result in the smallest predicted errors. For instance, in example 4.1, the RMSE for the fixed-order LADM with M=4 was less than that for adaptive-LADM. Thus, users should interpret the adaptive strategy as a means to control the decomposition based upon the user-defined criteria and not as an optimal strategy for reducing the RMSE.

Finally, in example 4.2, the hybrid method resulted in an RMSE of 1.402 x 10^{-5}, which is less than 1.356 x 10^{-4} for adaptive-LADM and slightly less than 1.210 x 10^{-4} for both fixed-order LADM and product-integration reference configurations. Consequently, in example 4.2, the hybrid method produced an RMSE that was approximately 89.7% smaller than that of adaptive-LADM.

A couple things need to be noted about these stated improvements. First, even though we have shown significant RMSE reductions via the hybrid approach of combining adaptive LADM with neural network residual corrections, the maximum discrete consistency residuals of 2.960 x 10^{-3} and 1.548 x 10^{-2}resulting from application of the discretized governing equations to manufactured-exact solutions for examples 4.1 and 4.2 indicate that there is still consistent discretization defects present in our current models. The consistency residual provides insight into what numerical stopping tolerance(s) may be used. It does not provide a lower bound on our computed RMSE's. Since we see significantly smaller RMSE's for the hybrid method, we conclude that we were able to better approximate our manufactured solutions on the chosen grids; however, we cannot determine if we reached high-order discretization accuracy or if we would converge at larger meshes through further grid studies.

6.2 Forecasting Performance and Generalization Across Prediction Horizons

The NGSIM US-101 application will present a more challenging test case for the proposed framework due to observation noise, non-stationarity and uncertainties in model parameter values. As opposed to manufactured experiment results, the true solution is unknown; therefore, predictive accuracy at longer time horizons must be based upon additional, held-out traffic density measurements. At a five second time horizon (the shortest tested) the Hybrid LADM–ANN model produced an RMSE of 6.149 vehl/km/lane. This compares to 6.630 for Adaptive LADM and 6.753 for Persistence. Therefore, it appears that Neural Residual Correction produced a statistically significant improvement over the physics based and persistence baselines for this short time horizon. However, the forecasting results were not uniformly positive for all longer time horizons. For example, at a twenty-five second time horizon, the hybrid model produced an RMSE of 22.156 while persistence was at 20.924. Finally, at fifty seconds and one hundred seconds the hybrid model produced RMSE's of 27.729 and 28.158 respectively, outperforming persistence; however, they both underperformed histogram gradient boosting which generated RMSE's of 25.351 and 26.926. Therefore, these results demonstrate that improved numerical solutions to a fractional governing equation do not automatically translate into superior predictive generalization. This distinction has particular relevance to traffic-density forecasting applications where regime changes in traffic patterns, poor modeling of external influences, uncertainties in physical parameters and accumulated error in previous forecasts may have a detrimental effect on long-term forecasting performance. Therefore, the empirical contributions of the proposed framework should be considered as demonstrating competitive and/or improved forecasting performance when comparing results from a variety of different prediction horizons; rather than providing universally superior performance over all possible horizons.

6.3 Parameter Identifiability, Sensitivity, and Predictive Uncertainty

The results from estimating parameters of the observation model are revealing of a significant drawback when interpreting physically the observation model. The estimated Caputo order (α) = 0.750 has a bootstrap 95% confidence interval [0.550, 1.000] while the estimated memory kernel order (β) = 1.000 has a much larger interval [0.200, 1.000]. Additionally, the estimated additional memory parameter (κ) = 0.00504 had a wide interval [0.00000, 0.03225] with an

© 2026 NSP
Natural Sciences Publishing Cor.

interval that includes zero. This suggests that using only the available traffic observations does not provide sharp identification of all fractional and physical parameters in the models. Notably, it seems possible that there is little support in the data for the existence of an additional memory effect; however, we can conclude neither that the data support nor reject the idea that the memory effect is redundant based solely upon the comparison of two independently calibrated formulations of different degrees of complexity. We also find complementary evidence from global sensitivity analysis. The Spearman rank correlation between the Caputo order (α) and forecast RMSE was about -0.9944 while those for β, a, and κ were about 0.0643, -0.0745, and -0.0928, respectively. Thus, the Caputo order (α) exhibits the strongest marginal relationship with forecast RMSE within our examined region of interest. However, we cannot determine if this implies causal importance, structural identifiability, or redundancy of parameters through the use of marginal rank correlations alone because nonlinear relationships and compensatory effects among other parameters may exist that would remain unobserved through such screening. The uncertainty results indicate that the proposed framework has a large degree of practical unreliability in terms of providing reliable predictions. As previously stated, although the static 95% conformal intervals achieved empirical coverage of 79.6% at a time horizon of 5 seconds, sequentially updating coverage levels rose to 86.3%, yet average interval widths grew by an additional 32.9%. Coverage levels decreased to 35.9% at a 100 second time horizon. Sequential updates raised coverage levels at all time horizons tested; however, they failed to reach the desired 95% level of coverage. Hence, these results show that improved accuracy of forecasts along with better adaptive control over the LADM approximation method do not directly translate into reliable quantification of predictive uncertainties. Specifically, deterministically determined LADM truncation conditions relate to numerical approximations characteristics while conformal coverage relates to behavior of forecasting errors under both calibration and testing distributions. Under-coverage seen in the current study could arise due to several factors including temporal dependency, non-stationarity/evolving nature of traffic regimes, or incorrect specification of the model; however, specific contributors have not been individually evaluated. Therefore, currently available prediction intervals should be treated as empirical diagnostic measures for uncertainties rather than being recognized as confirmed 95% coverage guarantees. Consequently, practical applications of this framework to transportation decision-making related to risk remain contingent on the development of more effective means of obtaining valid calibration and subsequently validating chronologically independent. Thus, future studies should focus on developing improved calibration methods for adaptive uncertainty estimation, distribution shift detection monitoring, and simultaneous assessment of coverage-width performance. Alongside this area of research, studies addressing the topics of identifying identifiable parameters in fractional models and performing comparisons between reduced and complete formulations will also be essential for further development of the framework.

6.4 Computational and Neural Architecture Limitations

The computational aspects of the hybrid approach must therefore be balanced against the quantitative benefits it provides. Maintaining full Caputo history prevents artificially resetting the fractional memory at each time step in the continuation of the solution; however, evaluating the history dependent operator itself requires significant computation and storage. Likewise, the additional Volterra memory term will likely add to these costs. This report does not contain sufficient data for a complete runtime comparison of the hybrid formulation with separately optimized competing formulations. Thus, while the improvements in the RMSE values reported here do indicate that there are potential reductions in error, they alone cannot demonstrate that the hybrid formulation is computationally more efficient than its competitors. A complete performance analysis of any solver will need to consider both runtime and memory usage as well as decomposition order statistics and accuracy-cost trade-offs. Another source of variability has been introduced into the problem via the use of a neural correction architecture. For example, in the manufactured-solution studies the two hidden layer tanh network with architecture (16,8) was chosen from a limited number of possible candidate networks based on validation RMSE. While this choice did provide good correction of residuals, no investigation of significantly larger or deeper architectures was conducted. Additionally, the observation study used a much smaller neural correction model than the one that was used in the manufactured experiment. As such, the findings related to selecting a particular architecture from those found in the manufacture experiments cannot be applied to the observational study. Thereby motivating further work investigating the relationships between different neural network capacities, optimizations settings, regularizations and computational costs.

7 Conclusions and Future Work

In this paper we have constructed an error-controlled hybrid physics-informed neural network framework for nonlinear fractional partial integro-differential equations with weakly singular kernels. The originality of our approach is not simply in the combination of an analytical solver with a neural network, but in developing a unified physics-informed

© 2026 NSP
Natural Sciences Publishing Cor.

framework where convergence control, adaptive approximation, memory preservation for fractional, data-driven residual correction, inverse parameter estimation, and uncertainty quantification are treated coherently. Specifically, our formulation has established a model specific contraction condition and explicit truncation error control, uses a dual adaptive stopping criteria based on both the LADM tail estimate and the governing equation residual, preserves the entire Caputo history during long-time continuation. The ANN then learns only the unresolved discrepancy of an already error-controlled physics-based approximation. Therefore, the ANN does not replace the physical description of the problem with a purely data driven surrogate.

The manufactured solution experiments provided controlled verification of the proposed methodology over substantially longer temporal intervals. In Example 4.1 the hybrid LADM-ANN model obtained an RMSE of 3.396×10^{-4}.Thus, the hybrid model obtained reductions of 83.9 %, 71.7 %, 81.9 % compared to the adaptive LADM, ANN only, product integration models respectively. In example 4.2 the RMSE of the hybrid model was reduced to 1.402×10^{-5}, Thus, the hybrid model obtained reductions of 89.7 %, 88.7 %, and 88.4 % compared to the adaptive LADM, ANN only, product integration models respectively. The adaptive experiments demonstrated that the required decomposition order is time-dependent; thus, a fixed truncation order may not provide uniformly accurate results over an extended integration interval. Our results confirmed that the numerical benefits of combining residual-based error control with physics-informed discrepancy learning were preserved while maintaining the specified initial and boundary conditions. The NGSIM US-101 experiment presents a more challenging evaluation using noisy and non-stationary observational data. Using 999,061 filtered trajectory observations aggregated onto a 158 $\times$ 32 time-space grid, we determined the values of the fractional traffic model parameters to be $\alpha = 0.75$, $\beta = 1.00$, a = 0.5549, and $\kappa = 0.00504$. For the five Second forecasting horizon, our hybrid model reduced RMSE by 7.3 % from 6.630 veh/km/lane for adaptive LADM to 6.149 veh/km/lane and significantly outperformed the ANN only model. However, our long-horizon experiments show that increasing the number of LADM terms alone will not remove structural model discrepancy: prediction errors grow rapidly after approximately 25 seconds and alternative statistical learners become competitive or better at horizons ranging from 50 to 100 seconds. We view the practical contributions of our proposed framework therefore as an error-controlled and interpretable short-horizon hybrid solver rather than as a universally superior long-term forecasting model. The uncertainty analysis also identified significant practical limitations of our current formulation. The wide block bootstrap confidence intervals for the fractional and memory parameters identify poor practical identifiability especially for β *and* κ while the sensitivity analysis identified the Caputo order α as the most important local sensitivity factor. Additionally, the under coverage of the nominal conformal intervals under chronological testing illustrates that traffic regime drift violate the exchangeability assumptions required for conventional split-conformal guarantees.

These observations emphasize that deterministic LADM truncation errors, parameter uncertainties, discretization errors, structural model discrepancies and statistical prediction errors should remain conceptually separate instead of being computed into a single nominal error measure. Following from these results, there are many directions for future work. First, development of rolling or online sequential recalibration procedures that can respond to non-stationary regimes will be investigated. Second, graded temporal meshes higher order quadrature methods and controlled history compression strategies should be investigated to improve computational efficiency especially for small fractional orders and long-memory simulations. Future investigations should also include systematic neural architecture studies comparing shallow, deeper, and wider residual networks under identical training and validation protocols. Such experiments should quantify the effects of network capacity on predictive accuracy, generalization, training stability, and computational cost, thereby establishing whether more expressive architectures offer meaningful improvements over the compact residual networks considered in the present study.Third, future research could investigate more expressive, physically constrained residual-learning models, including neural operator architectures such as Deep Operator Networks (DeepO Nets) and Fourier Neural Operators (FNOs), provided that sufficiently representative training data are available. These extensions should preserve the error-controlled LADM approximation as the physics-based numerical backbone and incorporate appropriate constraints to maintain consistency with the governing fractional partial integro-differential equation and its initial and boundary conditions. Fourth, identifying which fractional parameters can be uniquely identified from observational data should involve structural identifiability analysis and multi-experiment calibration. Future investigations should prioritize model-parsimony and identifiability analyses to determine whether the additional fractional-memory operator is empirically necessary. In particular, the complete model should be compared against reduced formulations with β fixed at unity and κ set to zero, using independent parameter re-estimation, chronological out-of-sample validation, and computational cost–accuracy assessments. Such comparisons would clarify whether the observed traffic dynamics require two distinct memory mechanisms or can be adequately represented by a simpler Caputo fractional model. Fifth, validation on additional classes of FPIDE and independent real-world datasets is needed to establish the generality of the proposed methodology beyond the present manufactured benchmarks and NGSIM traffic application. Another important direction is extending the proposed framework to no smooth and shock-dominated solutions.

© 2026 NSP
Natural Sciences Publishing Cor.

The present convergence analysis and numerical verification primarily address sufficiently regular solutions, while preprocessing of the NGSIM observations may attenuate abrupt traffic-density transitions. Future work will investigate conservative shock-capturing spatial discretization coupled with fractional-memory-preserving temporal integration and conservation-aware neural residual correction. Validation against discontinuous traffic-flow benchmarks will be necessary to establish stability, shock-resolution accuracy, and physical admissibility under such conditions.

Data availability statement: The NGSIM US-101 trajectory data are publicly available from the U.S. Department of Transportation. The processed source used here is identified in reference [64]. The nominal 08:05-08:20 file available for this analysis contained 788.6 s of timestamps; the analysis therefore reports the exact 785 s regular-grid interval used rather than claiming complete coverage of the nominal 15 min period. The preprocessing and core solver are specified in Section 5.7 and Appendix A.

Funding The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/ 2026/01/ 45553).

References

[1] A. Atangana and A. Secer, A note on fractional order derivatives and table of fractional derivatives of some special functions, Abstr. Appl. Anal. 2013 (2013), 279681, DOI: https://doi.org/10.1155/2013/279681.

[2] Essa, A. K., Al-Juboori, F. S., Jaafar, A. S., Alsaedi, H. J. N., Hussain, A. S., Yousif, R. A., & Almuttalibi, H. M. ANALYTICAL AND NUMERICAL ANALYSIS OF FRACTIONAL BAGLEY–TORVIK DIFFERENTIAL–ALGEBRAIC MODELS USING MODIFIED PHYSICS-INFORMED NEURAL NETWORKS.

[3] S. Noeiaghdam, S. Micula, and J. J. Nieto, A novel technique to control the accuracy of a nonlinear fractional order model of COVID-19: Application of the CESTAC method and the CADNA library, Maths 9 (2021), no. 12, 1321, DOI: https://doi.org/10.3390/math9121321.

[4] Z. Ali, S. N. Nia, F. Rabiei, K. Shah, and M. K. Tan, A semianalytical approach to the solution of time-fractional Navier-Stokes equation. Adv. Math. Phys. 2021 (2021), 1-13, DOI: https://doi.org/10.1155/2021/5547804.

[5] Alzubaidi, S., Hussain, A. S., Alsaedi, H. J. N., Alsuhabi, H., & Gemeay, A. M. (2026). An Entropy Transformed Inverse Rayleigh Gompertz Model under NHPP Framework with SVM–GA Estimation. Scientific African, e03652.

[6] Shirawia, N., Kherd, A., Bamsaoud, S., Tashtoush, M., Jassar, A., Az-Zo'bi, E. (2024). Dejdumrong Collocation Approach and Operational Matrix for a Class of Second-Order Delay IVPs: Error analysis and Applications. *WSEAS Transactions on Mathematics*, *23*, Article ID 49, 467-479.

[7] Essa, A., Alsaedi, H., Hussain, A., & Tashtoush, M. (2026). A Hybrid Support Vector Machine–Genetic Algorithm Framework for Estimating the DUS-Transformed Generalized Polynomial Quadratic Failure Rate Distribution. Statistics, Optimization & Information Computing, 16(1), 327-358.

[8] Alruhaili, H. S., Hussain, A. S., Ajlouni, A., T′urk, F., Az-Zo'bi, E. A., & Tashtoush, M. A. (2025). Solving Time-Fractional Nonlinear Variable-Order Delay PDEs Using Feedforward Neural Networks. Iraqi Journal for Computer Science and Mathematics, 6(3), 12.

[9] K. Hosseini, M. Ilie, M. Mirzazadeh, A. Yusuf, T. A. Sulaiman, and D. Baleanu. An effective computational method to deal with a timefractional nonlinear water wave equation in the Caputo sense, Math. Comput. Simul. 187 (2021), 248-260, DOI: https://doi.org/10.1016/ j. matcom.2021.02.021.

[10] K. Hosseini, K. Sadri, M. Mirzazadeh, A. Ahmadian, Y. M. Chu, and S. Salahshour, Reliable methods to look for analytical and numerical solutions of a nonlinear differential equation arising in heat transfer with the conformable derivative, Math. Methods Appl. Sci. 46 (2023), 10, 11342-11354, DOI: https://doi.org/10.1002/mma.7582.

[11] I. Podlubny, Fractional Differential Equations, Academic Press, San Diego, CA, 1999.

[12] Taha, Q. A., Hussein, M. Z., Hussain, A. S., Tashtoush, M. A., & Alsaedi, H. J. (2026). Reliability modeling of intelligent heat supply systems: A new class of cosine trigonometric based on non-homogeneous poisson process. Results in Nonlinear Analysis, 9(1), 117-148.

[13] Kadhim, R. J., Alsaedi, H. J., Hussain, A. S., Tashtoush, M. A., & Almuttalibi, R. A. (2026). An Explainable State-Space Transformer Framework for Predicting Clinical Deterioration in Non-Pulmonary Tuberculosis Patients. Ingenierie des Systemes d'Information, 31(6), 2011.

[14] J. Guo, D. Xu, and W. Qiu, A finite difference scheme for the nonlinear time-fractional partial integro-differential equation, Math. Methods Appl. Sci, 43 (2020), 6, 3392-3412, DOI: https://doi.org/10.1002/mma.6128.

[15] T. Akram, Z. Ali, F. Rabiei, K. Shah, and P. Kumam, A numerical study of nonlinear fractional order partial integro-differential equation with a weakly singular kernel, Fractal Fract. 5 (2021), no. 3, 85, https://doi.org/10.3390/fractalfract5030085.

[16] Hafez, M. A., Jameel, A. F., Akgul, A., Almuhur, E., Mistarihi, M. Z., Amen, S. G., ... & Tashtoush, M. A. (2026). Statistical Control Points: Approximate Methods for Solving First Order Fuzzy Initial Value Problems.

[17] Abed, S. A., Akg′ul, A., Hassani, M. K., Mohammed, Z. A., Hussain, A. S., & Tashtoush, M. A. (2026). Hybrid Firefly–Bayesian Neural Network Estimation for Reliable Modeling of the Inverse Gompertz–Gumbel Family and NHPP Systems: Properties and Comparative Evaluation. Journal of Statistical Theory and Applications, 25(1), 16.

© 2026 NSP
Natural Sciences Publishing Cor.

[18] A. K. Hussain, N. Rusli, F. S. Fadhel, and Z. R Yahya, Solution of one-dimensional fractional order partial integro-differential equations using variational iteration method, Aip Conf. Proc. 1775 (2016), 030096, https://doi.org/10.1063/1.4965216.

[19] M. R. Eslahchi, M. Dehghan, and M. Parvizi, Application of the collocation method for solving nonlinear fractional integro-differential equations, J. Comput. Appl. Math. 257 (2014), 105-128, https://doi.org/10.1016/j.cam.2013.07.044.

[20] E. Rawashdeh, Numerical solution of fractional integro-differential equations by collocation method, Appl. Math. Comput. 176 (2006), 1-6, https://doi.org/10.1016/j.amc.2005.09.059.

[21] J. Zhao, J. Xiao, and N. J. Ford, Collocation methods for fractional integro-differential equations with weakly singular kernels, Numer. Algor. 65 (2014), 723-743, DOI: https://doi.org/10.1007/s11075-013-9710-2.

[22] Mushref, O. M., Kurnaz, S., Hussain, A. S., Sadek, L., Hashim, A. M., & Tashtoush, M. A. (2025). A hybrid firefly-Bayesian Neural Network approach for survival and reliability estimation in non-homogeneous Poisson processes. Scientific African, e03011.

[23] Essa, A., Alsaedi, H., Hussain, A., & Tashtoush, M. (2026). A Hybrid Support Vector Machine–Genetic Algorithm Framework for Estimating the DUS-Transformed Generalized Polynomial Quadratic Failure Rate Distribution. Statistics, Optimization & Information Computing, 16(1), 327-358.

[24] Z. Avazzadeh, M. H. Heydari, and C. Cattani, Legendre wavelets for fractional partial integro-differential viscoelastic equations with weakly singular kernels, Eur. Phys. J. Plus 134 (2019), 368, https://doi.org/10.1140/epjp/i2019-12743-6.

[25] H. Dehestani, Y. Ordokhani, and M. Razzaghi, Numerical solution of variable-order time fractional weakly singular partial integro-differential equations with error estimation, Math. Model. Anal, 25 (2020), 680-701, DOI: https://doi.org/10.3846/mma.2020.11692.

[26] H. Khan and M. Arif, Numerical solutions of some higher order fractional integro-differential equations (FIDEs), using Chebyshev wavelet method (CWM), J. Appl. Environ. Biol. Sci. 7 (2017), no. 9, 105-114.

[27] S. T. Mohyud-Din, H. Khan, M. Arif, and M. Rafiq, Chebyshev wavelet method to nonlinear fractional Volterra-Fredholm integro-differential equations with mixed boundary conditions, Adv. Mech. Eng. 9 (2017), no. 3, 8, DOI: https://doi.org/10.1177/1687814017694802.

[28] K. Hosseini, M. Ilie, M. Mirzazadeh, D. Baleanu, C. Park, and S. Salahshour, The Caputo-Fabrizio time-fractional Sharma-Tasso-Olver-Burgers equation and its valid approximations, Commun. Theor. Phys. 74 (2022), no. 7, 075003, DOI: https://doi.org/10.1088/1572-9494/ac633e.

[29] K. Hosseini, M. Ilie, M. Mirzazadeh, and D. Baleanu, An analytic study on the approximate solution of a nonlinear time-fractional Cauchy reaction-diffusion equation with the Mittag-Leffler law, Math. Methods Appl. Sci. 44 (2021), no. 8, 6247-6258, DOI: https://doi.org/10. 1002/mma.7059.

[30] M. Yavuz and N. ́Ozdemir, European vanilla option pricing model of fractional order without singular kernel, Fractal fract. 2 (2018), no. 1, 3, https://doi.org/10.3390/fractalfract2010003.

[31] K, Parand and M. Delkhosh, Systems of nonlinear Volterra integro-differential equations of arbitrary order, Bol. da Soc. Parana. de Mat. 36 (2018), no. 4, 33-54, https://doi.org/10.5269/bspm.v36i4.31478.

[32] Az-Zo'bi, E. A., Hussain, A. S., Iqbal, M., Aljohani, A., Tashtoush, M. A., Alsubaie, N. E., ... & Baleanu, D. (2025). Analytical and numerical solutions of MABC fractional advection dispersion models by utilizing the modified physics informed neural networks with impacts of fractional derivative. Scientific Reports, 15(1), 43366. https://dergipark.org.tr/en/pub/gujs/issue/27537/289699.

[33] K. Parand and J. A. Rad, An approximation algorithm for the solution of the singularly perturbed Volterra integro-differential and Volterra integral equations, Int. J. Non-linear Sci. 12 (2011), no. 4, 430-441.

[34] A. Gokce and B. Gurbuz, A numerical scheme for the one-dimensional neural field model, Int. J. Optim. 12 (2022), no. 2, 184-193, DOI: https://doi.org/10.11121/ijocta.2022.1219.

[35] A. Refice, M. S. Souid, and A. Yakar, Some qualitative properties of nonlinear fractional integro-differential equations of variable order, Int. J. Optim. 11 (2021), no. 3, 68-78, DOI: https://doi.org/10.11121/ijocta.2021.1198.

[36] F. Evirgen and N. ́Ozdemir, Multistage adomian decomposition method for solving NLP problems over a nonlinear fractional dynamical system, J. Comput. Nonlinear Dynam. 6 (2011), no. 2, 021003, DOI: https://doi.org/10.1115/1.4002393.

[37] M. I. Syam and A. Hamdan, An efficient method for solving Bratu equations, Appl. Math. Comput, 176 (2006), no. 2, 704-713, DOI: https://doi.org/10.1016/j.amc.2005.10.021.

[38] M. Yavuz and N. ́Ozdemir, A quantitative approach to fractional option pricing problems with decomposition series, Konuralp J. Math. 6 (2018), 102-109.

[39] A. M. Wazwaz, A reliable modification of Adomian decomposition method, Appl. Math. Comput. 102 (1999), 77-86, DOI: https://doi.org/ 10.1016/S0096-3003(98)10024-3.

[40] A. M. Wazwaz, A new algorithm for solving differential equations of LaneEmden type, Appl. Math. Comput. 118 (2001), 287-310, DOI: https://doi.org/10.1016/S0096-3003(99)00223-4.

[41] A. M. Wazwaz, A new method for solving singular initial value problemsin the second-order ordinary differential equations, Appl. Math. Comput. 128 (2002), 45-57, DOI: https://doi.org/10.1016/S0096-3003(01)00021-2.

[42] M. Caputo, Elasticita e Dissipazione, Zanichelli, Bologna, 1969.

[43] M. Caputo, Linear Model of dissipation whose Q is almost frequency independent-II, Geophys. J. R. Astron. Soc, 13 (1967), 529-539, DOI: https://doi.org/10.1111/j.1365-246X.1967.tb02303.x.

[44] G. Mittag-Leffler, Sur la nouvelle fonction $E_\alpha(x)$, Compt. Rend. Acad. Sci. Paris 137 (1903), 554-558.

© 2026 NSP
Natural Sciences Publishing Cor.

[45] A. M. Wazwaz, A new algorithm for calculating Adomian polynomials for nonlinear operators, Appl. Math. Comput, 111 (2000), no. 1, 33-51, DOI: https://doi.org/10.1016/S0096-3003(99)00063-6.
[46] F. Mainardi, Fractional Calculus and Waves in Linear Viscoelasticity, Imperial College Press, London, 2010.
[47] R. Bagley, On the fractional order initial value problem and its engineering applications, In International Conference Proceedings, Tokyo, 1990, pp. 12-20.
[48] S. Kazem, Exact solution of some linear fractional differential equations by Laplace transform, Int. J. Non-linear Sci 16 (2013), no. 1, 3-11.
[49] J. Hanna and J. Rowland, Fourier Series, Transforms, and Boundary Value Problems, John Wiley and Sons, Inc, New York, 1990.
[50] C. Yang and J. Hou, Numerical solution of integro-differential equations of fractional order by Laplace decomposition method, WSEAS Trans. Math. 12 (2013), no. 12, 1173-1183.
[51] H. Alrabaiah, M. Jamil, K. Shah, and R. A. Khan, Existence theory and semi-analytical study of nonlinear Volterra fractional integrodifferential equations, Alex. Eng. J. 59 (2020), no. 6, 4677-4686, DOI: https://doi.org/10.1016/j.aej.2020.08.025.
[52] A. Abdelrazec and D. Pelinovsky, Convergence of the Adomian decomposition method for initial-value problems, Numer. Methods Partial Differential Equations 27 (2011), no. 4, 749-766, DOI: https://doi.org/10.1002/num.20549.
[53] Nawaz, S., Akg ̧ul, A., Hassani, M., Tashtoush, M. (2025). Generalized m-functional equations in matrix Banach space with an application, *Journal of Inequalities and Applications*, *2025*, Article ID 119, 1-17.
[67] Mushref, O., Kurnaz, S., Hussain, A., Sadek, L., Hashim, A., Tashtoush, M. (2025). A Hybrid Firefly-Bayesian Neural Network approach for survival and Reliability Estimation in Non-Homogeneous Poisson Processes, *Scientific African*, *30*(2025), e03011. https://doi.org/10.1016/j.sciaf.2025.e03011
[66] Az-Zo'bi, E., Hussain, A., Iqbal, M., Aljohani, A., Tashtoush, M., Alsubaie, N., Al-Hamarneh, F., Nazar, N., Baleanu, D. (2025). Analytical and Numerical Solutions of MABC Fractional Advection Dispersion Models by Utilizing the Modified Physics Informed Neural Networks with Impacts of Fractional Derivative, *Scientific Reports*, *15*(1), Article ID 43366. https://doi.org/10.1038/s41598-025-30065-7
[65] Az-Zo'bi, E., Hussain, E., Iqbal, M., Alomari, Q., Shah, S., Yaro, D., Aljawi, S., Alsubaie, A., Tashtoush, M. (2025). Chaotic, Bifurcation, Sensitivity, Modulation Stability Analysis and Optical Soliton Structure to the Nonlinear Coupled Konno-Oono System in Magnetic Field, *AIP Advances*, *15*(9), Article ID 095103. https://doi.org/10.1063/5.0291009
[57] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019), 686-707, DOI: https://doi.org/10.1016/j.jcp.2018.10.045.
[58] G. Pang, L. Lu, and G. E. Karniadakis, fPINNs: Fractional physics-informed neural networks, SIAM J. Sci. Comput. 41 (2019), A2603-A2626, DOI: https://doi.org/10.1137/18M1229845.
[59] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, learning nonlinear operators via DeepONet based on the universal approximation theorem of operators, Nat. Mach. Intell. 3 (2021), 218-229, DOI: https://doi.org/10.1038/s42256-021-00302-5.
[60] Ramadan, M. A., Abd El-Latif, M. A. A., and Alqarni, M. Z. (2025). An integrated Laplace transform and accelerated Adomian decomposition approach for solving time-fractional nonlinear partial differential equations. Electronic Research Archive, 33(7), 4398–4434. DOI: 10.3934/era.2025201.
[61] Wu, Q., Ma, F., and Wo, W. (2025). Fs-PINNs: Fractional spectrally adapted physics-informed neural networks for fractional partial differential equations. Mathematical Methods in the Applied Sciences, 48(14), 13882–13889. DOI: 10.1002/mma.11149.
[62] Lu, B., Hao, Z., Moya, C., and Lin, G. (2025). FPINN-DeepO Net: A physics-informed operator learning framework for multi-term time-fractional mixed diffusion-wave equations. Journal of Computational Physics, 538, 114184. DOI: 10.1016/j.jcp.2025.114184.
[63] Alomari, A., Hussain, A., Al-Hamarneh, F., Tashtoush, M., Az-Zo'bi, A. (2026). A Hybrid Neuro-Fuzzy Framework for Multi-Release Open-Source Software Reliability Modeling under Uncertainty, *International Journal of Robotics and Control Systems*, *6*(4), 3563-3596. https://doi.org/10.31763/ijrcs.v6i4.2547
[64] R. El Ballouli, NGSIM US-101 trajectory dataset smoothing, 2020, https://github.com/Rim-El-Ballouli/NGSIM-US-101-trajectory-dataset-smoothing.
[65] Zhang, J., Zhao, Y., & Tang, Y. (2024). Adaptive loss weighting auxiliary output fPINNs for solving fractional partial integro-differential equations. *Physica D: Nonlinear Phenomena*, *460*, 134066.
[66] Hu, Z., Kawaguchi, K., Zhang, Z., & Karniadakis, G. E. (2024). Tackling the curse of dimensionality in fractional and tempered fractional PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, *432*, 117448.
[67] Yuan, S., Liu, Y., Xu, Y., Li, Q., Guo, C., & Shen, Y. (2024). Gradient-enhanced fractional physics-informed neural networks for solving forward and inverse problems of the multiterm time-fractional burger-type equation. *AIMS Mathematics*, *9*(10), 27418-27437.
[68] Srati, M., Oulmelk, A., Afraites, L., Hadri, A., Zaky, M. A., Aldraiweesh, A., & Hendy, A. S. (2025). An inverse problem of determining the parameters in diffusion equations by using fractional physics-informed neural networks. *Applied Numerical Mathematics*, *208*, 189-213.
[69] Wu, Q., Ma, F., & Wo, W. (2025). Fs-Pinns: Fractional Spectrally Adapted Physics-Informed Neural Networks for Fractional Partial Differential Equations. *Mathematical Methods in the Applied Sciences*, *48*(14), 13882-13889.
[70] Lu, B., Hao, Z., Moya, C., & Lin, G. (2025). FPINN-deeponet: A physics-informed operator learning framework for multi-term time-fractional mixed diffusion-wave equations. *Journal of Computational Physics*, *538*, 114184.

© 2026 NSP
Natural Sciences Publishing Cor.

Appendix A. Computational Implementation and Reproducibility Protocol

A.1. Software Environment and Numerical Configuration

The numerical experiments were conducted using the software environment reported in the original implementation: Python 3.12, NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0, scikit-learn 1.8.0, and Matplotlib 3.10.8. The reported random seed is 20260924. The experimental configurations are specified separately for the two manufactured solution benchmarks and the NGSIM US-101 application. Tables 1 and 4 define the spatial and temporal grids, adaptive tolerances, maximum decomposition orders, and neural-network architectures for the manufactured examples. Table 14 specifies the corresponding settings for the observational application. A complete reproducibility package should additionally include a machine-readable dependency manifest, software-version information, and an executable script for each experiment.

A.2. Fractional-Operator Discretization

The Caputo fractional derivative is evaluated using the L1 approximation, while the weakly singular Volterra integral is approximated by product integration. Spatial derivatives are discretized using the schemes specified in the experimental methodology. For a uniform temporal grid and fractional order $0 < \alpha < 1$, the L1 approximation takes the form

$$ {}^{C}D_{t}^{\alpha}u(t_n) \approx \frac{1}{\Gamma(2-\alpha)(\Delta t)^{\alpha}} \sum_{j=0}^{n-1} \left[(j+1)^{1-\alpha} - j^{1-\alpha}\right] \left(u^{n-j} - u^{n-j-1}\right). $$

The implementation must preserve the historical solution values required by the Caputo derivative and must not reset the memory contribution when advancing between temporal slabs. For the weakly singular memory operator, quadrature weights must be constructed consistently with the kernel normalization adopted in the governing FPIDE. Any accelerated approximation must document its additional kernel-approximation error.

A.3. Adaptive LADM Solver

The adaptive decomposition begins with the prescribed initial and boundary data, followed by construction of the corresponding Adomian components. At each iteration, the partial sum, truncation-tail estimate, and governing-equation residual are updated. An approximation is accepted only when both the truncation-error criterion and the FPIDE residual criterion are satisfied. If either criterion fails, additional decomposition components are computed up to the prescribed maximum order. If the convergence or acceptance requirements cannot be satisfied, the temporal slab must be refined or the solver must report failure rather than silently accepting an unconverted approximation. The implementation must record the selected decomposition order, residual norm, tail estimate, and acceptance status at each evaluation time.

A.4. Parameter Calibration

For the NGSIM application, the fractional orders and physical coefficients are estimated by minimizing the normalized discrete FPIDE residual under the parameter constraints described in Section 5.3. The search grids for and are specified in Table 14. The remaining coefficients are estimated using constrained least squares, with the relevant positivity constraints enforced. Each candidate fractional-order combination requires precomputation of the corresponding fractional operators. Source-term coefficients must be re-estimated consistently with the selected candidate parameters. The final implementation should record the objective value, parameter vector, constraint status, and optimization diagnostics for each candidate configuration.

A.5. Residual Neural-Network Correction

The neural-network component approximates the unresolved discrepancy between the physics-based LADM prediction and the available reference observations. For the NGSIM application, the architecture and optimization settings follow Table 14. The neural model is trained exclusively using the chronological training observations, while model selection and calibration use the designated validation period. The correction must respect the initial and boundary conditions imposed by the mathematical formulation. All preprocessing transformations, feature definitions, normalization parameters, and fitted model parameters must be determined without using future test observations.

© 2026 NSP
Natural Sciences Publishing Cor.