fractal and fractional

MDPI

Article

Equation-Free Screening of Mittag–Leffler Compatibility from Multi-Horizon Relaxation Profiles

N'Gbo N'Gbo [1,2], Andrei Velichko [3] and Bruno Carpentieri [4,*]

[1] School of Science and Engineering, International University of Grand-Bassam, Route de Bonoua, Grand-Bassam BP 564, Côte d'Ivoire; npaulrene2@outlook.com or ngbo.n@iugb.edu.ci
[2] Laboratory for Intelligence and Mathematics (LIMAs), Route de Bonoua, Grand-Bassam BP 564, Côte d'Ivoire
[3] Institute of Physics and Technology, Petrozavodsk State University, 33 Lenina Prospekt, Petrozavodsk 185910, Republic of Karelia, Russia; velichkogf@gmail.com
[4] Faculty of Engineering, Free University of Bozen-Bolzano, Via Bruno Buozzi 1, 39100 Bozen–Bolzano, Italy
* Correspondence: bruno.carpentieri@unibz.it

Abstract

We study an equation-free screen for finite-window Mittag–Leffler compatibility, in which multi-horizon relaxation profiles are constructed out of sample by k-nearest-neighbor prediction from ensembles of scalar trajectories and compared with candidate temporal laws. The screen assesses compatibility, not the generating mechanism. In a matched, noise-free calibration, it reproduces the generating Mittag–Leffler behavior to numerical precision, while relaxation-rate heterogeneity shifts the fitted Mittag–Leffler order toward lower values. For noise-free classical distributed-rate ensembles, a Mittag–Leffler preference over a single exponential disappears when a biexponential alternative is included. Under observational noise, this recovery is incomplete: with three candidate families, the Mittag–Leffler law is still selected in 60% of runs for a classical ensemble, against 72% for a fractional reference. In the tested noise-free settings, simpler profile constructions reproduce the screening decisions, whereas kNN predicts individual trajectories far more accurately than a constant predictor. The screening interpretation of a single-record example, initially investigated as a chaotic test case, is withdrawn as evidence about the continuous system after numerical verification revealed a nonrobust simulated regime. Mittag–Leffler compatibility is therefore a comparative, condition-dependent screening outcome, not identification of fractional origin.

Keywords: fractional dynamics; Mittag–Leffler relaxation; k-nearest neighbors; time-series analysis; noise robustness; fractional screening; Caputo derivative; model selection

Check for updates

Academic Editor: Ricardo Almeida

Received: 18 September 2026
Revised: 21 October 2026
Accepted: 21 October 2026
Published: 26 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Fractional calculus provides a natural mathematical framework for systems in which the present state depends on a distributed history rather than only on an instantaneous local state. Caputo and related fractional operators have therefore been used to represent hereditary effects, anomalous relaxation, viscoelasticity, transport, electrochemical memory, and nonlinear dynamical behavior [1–5]. A canonical case is the linear Caputo relaxation problem ${}_C D^{\alpha}_{0,t} x = -\lambda x$, whose solution is the Mittag–Leffler law $x(t) = x_0 E_{\alpha}(-\lambda t^{\alpha})$, with the ordinary exponential recovered in the integer-order limit $\alpha = 1$ [4–6]. In this canonical setting, the governing fractional dynamics are linked analytically to a scalar relaxation response.

This explicit link suggests a data-driven question. When the governing equations are unknown and only scalar observations are available, can the data reveal a temporal pattern compatible with this canonical fractional response? Such compatibility could motivate further investigation of fractional models.

The difficulty is that such a pattern is not unique to fractional dynamics. Relaxation curves that resemble a Mittag–Leffler law can also be produced by classical mechanisms: stretched exponentials, mixtures or distributions of ordinary exponential rates, aggregation, structural changes, and nonstationarity can all generate similarly broad or curved responses over a finite observation window [7–10], and observational noise can further obscure their differences. Agreement with a Mittag–Leffler-type pattern can therefore indicate compatibility with a fractional description, but cannot by itself identify the mechanism that generated the data. It is useful to separate three questions that are often merged. The first is whether the observed temporal behavior is compatible with a fractional-type law. The second is whether a fractional generating mechanism is better supported than plausible classical alternatives. The third is what the parameters of that mechanism are, including the fractional order. Each question demands stronger evidence than the one before it.

Existing fractional-system identification methods work within a specified fractional model class: they estimate its order, coefficients, hidden states, or transfer function from input–output or state data [11–15]. Related model-based work also addresses observability and initial-state reconstruction in time-fractional systems [16]. These methods are powerful when a credible governing model is available, but they operate after a fractional model class has already been posited. In many applications only a single scalar quantity is measured and no governing ODE, PDE, or FDE has been selected. In such settings, a preliminary data-driven assessment of the first question can be useful before committing to a particular governing model [17–19]. The purpose of this work is to provide such an assessment: an earlier, equation-free screening layer for settings in which the appropriate governing model class is still uncertain.

The screening principle builds on the use of multi-horizon prediction as an equation-free probe of temporal dynamics, an idea we previously exploited to estimate largest Lyapunov exponents from one-dimensional chaotic time series [20]. Rather than fitting a governing differential equation to the observations, we use direct-horizon kNN prediction to construct, out of sample, an empirical relaxation profile from an ensemble of scalar trajectories: each test trajectory is represented only by its observed history, and its future is predicted from training trajectories with similar histories. Candidate temporal laws are then fitted to this profile as reference families to quantify finite-window compatibility, not to identify the governing equations directly. If the Mittag–Leffler law describes the profile better than the classical alternatives considered, a fractional description may warrant further model-based investigation. Because screening is performed on the aggregate profile, it is not evident how much the history-conditioned kNN construction contributes to the screening decision compared with simpler alternatives, including a constant training-mean predictor and the directly observed ensemble average; we examine this question explicitly.

The practical value of such a screen depends on understanding both its reliability and its limitations. The aim of this study is therefore twofold: to propose the screen and to examine its reliability and limitations under increasingly demanding conditions. We address four research questions:

RQ1 *Calibration.* Can a multi-horizon kNN profile recover the expected Mittag–Leffler compatibility in a controlled matched-family relaxation setting?

RQ2 *Specificity and candidate-set dependence.* How do relaxation-rate heterogeneity, non-fractional curved mechanisms, and observational noise affect an apparent Mittag–

Leffler preference, and how does that preference change when the candidate set is broadened?

RQ3 *Interpretation and robustness.* How robust are the screening decision and the fitted Mittag–Leffler shape parameter to observational noise and analysis choices? When the data do not arise from a single matched law, how should the fitted parameter be interpreted?

RQ4 *Profile construction.* To what extent does the kNN construction change the screening decision relative to simpler profile constructions, and what does it add at the level of individual-trajectory prediction?

As a controlled foundation, we deliberately begin with the canonical linear Caputo relaxation problem introduced above. Its simplicity is precisely what makes it useful: the generating law is known analytically, so the screening principle can first be calibrated under matched conditions and then challenged progressively by rate heterogeneity, nonfractional alternatives, observational noise, and sensitivity to analysis choices, while the kNN construction is compared with simpler profile constructions. We also report a single-record example generated numerically from a nonlinear fractional system initially investigated as a chaotic test case. Verification with an independent predictor–corrector implementation showed that the sustained oscillatory regime produced by the original explicit discretization is not numerically robust; we therefore retain this example only as a cautionary illustration of how numerical artifacts in simulated observations can affect data-driven screening.

Together, these experiments probe the conditions under which a finite-window compatibility screen is informative and those under which its interpretation becomes ambiguous. The resulting framework is positioned between descriptive time-series analysis and full fractional-system identification: it is intended to motivate subsequent mechanistic modeling, not to replace it.

Section 2 describes the screening principle, the benchmarks, and the protocols. Section 3 reports the results, and Section 4 discusses their implications and limitations. Section 5 concludes.

2. Materials and Methods

Section 2.1 states the screening procedure and the separation between model fitting and evaluation (Figure 1). Sections 2.2 and 2.3 define the stable benchmarks and their observable, together with the simpler profile constructions used for comparison (RQ1, RQ2, RQ4). Section 2.4 specifies the candidate families and their optimization, and Sections 2.5 and 2.6 the noise, specificity, and sensitivity analyses (RQ2, RQ3). Section 2.7 describes the numerical-reliability check of a single-record example.

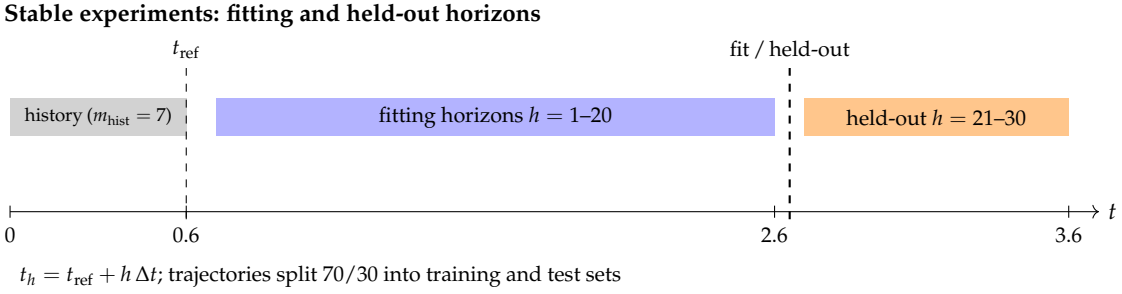


Figure 1. Separation of training and test data in the stable experiments: trajectories are split 70/30 into training and test sets; the candidate relaxation laws are fitted on horizons $h = 1$–20 and evaluated on the held-out horizons $h = 21$–30, on the absolute target time $t_h = t_{\mathrm{ref}} + h\,\Delta t$.

2.1. Screening Principle and Data-Separation Protocols

The screening procedure has three steps. First, an empirical multi-horizon relaxation profile is constructed from an ensemble of scalar trajectories by direct-horizon kNN prediction. Second, candidate temporal laws are fitted to this profile over a prescribed fitting interval. Third, the fitted laws are compared on that interval and on later horizons not used for fitting.

Because the screen is intended to characterize out-of-sample behavior, the profile must be constructed without information from the evaluation data entering the predictor. The kNN predictor and all standardization quantities therefore use training data only. Forecasting is also *direct-horizon*: each horizon h is predicted separately from the history vector, rather than recursively from predictions at shorter horizons. The profile at horizon h thus reflects the empirical predictive relationship at that horizon, not errors accumulated by iterating a one-step predictor.

In the stable branch, the data are separated at the trajectory level: 70% of the trajectories train the kNN predictor, and the remaining 30% are used to compute the profile for horizons $h = 1, \ldots, 30$. Candidate laws are fitted on horizons 1–20 and then, with their parameters held fixed, evaluated on horizons 21–30. We call this second stage a *horizon-extrapolation check*: it asks whether a law fitted to the earlier part of the profile continues to describe its later part. Because all horizons are computed from the same test trajectories, they are statistically dependent, and the later horizons are not treated as an independent validation sample.

Within each experiment, all candidate laws are fitted to the same profile with the same least-squares criterion (Section 2.4). Because the screen asks which reference family continues to describe the profile beyond the horizons used to fit it, the primary comparison criterion for the baseline stable experiments is the horizon-extrapolation RMSE on horizons 21–30. The corresponding intervals are adjusted in the sensitivity analyses described in Section 2.6. In-sample criteria need not agree with the extrapolation criterion: a more flexible family can fit the fitting horizons more closely and yet extrapolate less well. AIC and BIC computed on the fitting horizons are therefore reported only as descriptive complexity-penalized summaries, not as formal likelihood-based tests, because neighboring horizons are statistically dependent. Neither held-out RMSE nor the in-sample criteria can identify the generating mechanism: they assess compatibility with the constructed profile.

The screening decision is therefore comparative rather than a formal test of fractional origin. A lower fitting or extrapolation error indicates compatibility with a candidate family relative to the alternatives considered; it does not identify a unique generating mechanism or establish a physical fractional order.

2.2. Stable Benchmark Design

To interpret what the screen detects, it must first be tested on data whose generating law is known. The stable benchmarks therefore start from the most favorable situation and progressively relax or replace its idealizations. Three principal cases separate three questions: exact matched-family calibration (Case A), robustness to moderate rate heterogeneity within a fractional relaxation family (Case B), and specificity against a purely classical distributed-rate mechanism (Case C). Two additional shape controls probe a different issue: whether strong curvature alone can generate an apparent Mittag–Leffler-compatible profile, and whether a correctly specified curved classical family is recovered when it is included among the candidates.

2.2.1. Principal Stable Benchmark Sequence

For Cases A and B, trajectories are generated from the analytical fractional relaxation law

$$x_j(t) = x_{0,j} E_\alpha(-\mu_j t^\alpha), \tag{1}$$

where E_α denotes the one-parameter Mittag–Leffler function and $x_{0,j} \sim \mathcal{U}[0.5, 3]$. We use $N_{\text{traj}} = 500$ trajectories for each parameter setting. The Mittag–Leffler function is evaluated with the robust algorithm of Garrappa [21], so no time-integration discretization is introduced in the stable data generation. Only the interval $t \in [0, 3.6]$, spanned by the history window and the 30 prediction horizons (Section 2.3), enters the analysis; samples beyond $t = 3.6$ that the scripts may generate are not used.

Case A: exact matched-family calibration.

Case A is the best-case reference. The data are generated from the same Mittag–Leffler family that is later fitted, with $\alpha \in \{0.6, 0.7, 0.8, 0.9, 1\}$ and a common relaxation coefficient

$$\mu_j = \mu = 0.15 \tag{2}$$

for all trajectories. Because μ is common to all trajectories,

$$x_j(t) = x_{0,j} f(t), \qquad f(t) = E_\alpha(-\mu t^\alpha),$$

so every history is a rescaled copy of the same temporal template. Under this separable structure, the trajectory-dependent amplitude factor introduced by the kNN averaging is absorbed by the fitted amplitude C. Near-exact recovery is therefore the expected behavior here: Case A provides a calibration check of the numerical screening pipeline under exact controlled conditions, and its success is not evidence of general identifiability.

Case B: fractional relaxation with moderate rate heterogeneity.

Case A is deliberately ideal. Case B removes its multiplicative separability while remaining close to the reference relaxation scale: each trajectory is assigned its own rate,

$$\mu_j \sim \mathcal{U}[0.05, 0.25], \tag{3}$$

with a common fractional order α. The interval is centered at 0.15 and introduces moderate inter-trajectory heterogeneity without changing the underlying fractional relaxation family. Case B therefore tests whether the Mittag–Leffler-compatible profile remains detectable when the ensemble can no longer be written as $x_j(t) = x_{0,j} f(t)$ with one common temporal factor f.

Case C: classical distributed-rate challenge.

Case B still lies within a fractional generating family. Case C asks the stronger specificity question: can rate heterogeneity alone produce the same apparent signature when every individual trajectory is classical? Each trajectory is purely exponential,

$$x_j(t) = x_{0,j} \exp(-\mu_j t), \tag{4}$$

but the rates are log-uniformly distributed according to

$$\mu_j = 0.15\, s^{u_j}, \qquad u_j \sim \mathcal{U}[-1, 1], \qquad s \in \{3, 5, 10\}. \tag{5}$$

The distribution is centered geometrically at the same reference value 0.15, while s controls the spread. The spread is deliberately much wider than in Case B: Case C is not a perturbation of the reference configuration but an increasingly demanding classical control as s grows. Case C is first used in the same binary exponential-versus-Mittag–Leffler comparison as Cases A and B and is then revisited with a broader candidate set that includes a biexponential classical alternative; the fitting families themselves are defined in Section 2.4.

2.2.2. Separable Nonfractional Shape Controls

Cases A–C vary the rate structure of the ensemble. The two shape controls ask a different question: can a nonfractional but strongly curved temporal law mimic Mittag–Leffler compatibility, and is a curved classical law recognized when it is itself among the candidates? Both controls retain the separable form $x_j(t) = x_{0,j} f(t)$, preserving the ensemble structure of Case A while changing the temporal law.

The first control uses the generalized exponential law

$$x(t) = x_0 \exp\left(-\mu t^{\beta}\right), \tag{6}$$

with $\mu = 0.15$, $\beta = 0.4$ for stretched relaxation, and $\beta = 2$ for compressed relaxation; since μ multiplies t^{β}, it has dimension $[t]^{-\beta}$, so the same value has a different interpretation for the two values of β. These controls are false-positive shape tests: in particular, the stretched case asks whether a strongly curved but nonfractional decay can appear more Mittag–Leffler-compatible than a single exponential when the candidate set is restricted.

The second control uses the rational/Hill-type law

$$x(t) = \frac{x_0}{1 + (t/\mu)^{\beta}}, \tag{7}$$

with $\beta = 2$ and $\mu = 0.5$. In this parameterization, μ is a time-scale parameter rather than the coefficient appearing in the exponent of Equation (6). This experiment has a different role from the generalized-exponential false-positive test: the generating rational/Hill family is itself included among the fitted candidates. It is therefore a matched-family control designed to test whether broadening the candidate set can recover the appropriate curved classical family, starting from initial guesses that exclude the generating pair (Section 2.4).

2.3. Stable kNN Relaxation Profile and Horizon Extrapolation

All stable benchmarks are evaluated through the same observable, the kNN relaxation profile. This subsection defines it and fixes the time coordinate on which the candidate laws are evaluated. For each stable trajectory, the first $m_{\mathrm{hist}} = 7$ samples form the history vector, and future states at horizons $h = 1, \ldots, 30$ are the prediction targets. Using the trajectory-level split of Section 2.1, a Euclidean K-nearest-neighbor predictor with $K = 3$ maps each test history directly to the target at horizon h by averaging the corresponding targets of its three nearest training histories.

The stable observable is a kNN-derived relaxation profile rather than a forecast-error profile. Let $\widehat{x}_i^{(h)}$ denote the direct-horizon kNN prediction for test trajectory i at horizon h, and let $x_{i,\mathrm{ref}}$ denote the final state included in its history. The profile is defined as

$$D_{\mathrm{pred}}(h) = \frac{\frac{1}{N_{\mathrm{test}}} \sum_i \left| \widehat{x}_i^{(h)} \right|}{\frac{1}{N_{\mathrm{test}}} \sum_i \left| x_{i,\mathrm{ref}} \right|}. \tag{8}$$

Thus, $D_{\mathrm{pred}}(h)$ measures the average amplitude remaining in the kNN prediction relative to the reference state. It is a normalized relaxation-to-equilibrium profile; it is neither a Lyapunov separation between neighboring trajectories nor a forecast-error measure.

The time coordinate requires care. Unlike the exponential, Mittag–Leffler relaxation does not satisfy a semigroup property, so the clock cannot simply be reset at the end of the history window while retaining the same scalar relaxation law: the candidate laws must be evaluated on the absolute target time. Because the final point of the history window occurs at

$$t_{\mathrm{ref}} = (m_{\mathrm{hist}} - 1)\Delta t = 0.6,$$

the target time is

$$t_h = t_{\mathrm{ref}} + \tau_h, \qquad \tau_h = h\Delta t. \tag{9}$$

With $\Delta t = 0.1$, fitting horizons $h = 1, \ldots, 20$ therefore correspond to $t_h \in [0.7, 2.6]$, whereas horizons $h = 21, \ldots, 30$ correspond to $t_h \in [2.7, 3.6]$.

These two intervals are the fitting and horizon-extrapolation intervals of Section 2.1. The data-separation logic of Section 2.1 is summarized in Figure 1.

The normalization in Equation (8) affects the profile amplitude but not the physical time coordinate. Since every candidate relaxation family contains a free amplitude parameter C, the normalization factor is absorbed into C, whereas the time argument remains the absolute target time t_h.

Because the screening decision is taken on the aggregate profile, we compare the kNN construction with two simpler constructions, each passed through the same fitting and extrapolation pipeline. Let $x_i^{(h)}$ denote the observed state of test trajectory i at horizon h, and let $\bar{x}_{\mathrm{ref}} = \frac{1}{N_{\mathrm{test}}} \sum_i |x_{i,\mathrm{ref}}|$ be the denominator of Equation (8). The constant training-mean predictor assigns to every test trajectory the mean training target $\bar{x}_{\mathrm{train}}^{(h)}$, giving

$$D_{\mathrm{const}}(h) = \frac{\left|\bar{x}_{\mathrm{train}}^{(h)}\right|}{\bar{x}_{\mathrm{ref}}}; \tag{10}$$

it is an out-of-sample predictor that ignores the individual histories. The direct ensemble average,

$$D_{\mathrm{direct}}(h) = \frac{1}{\bar{x}_{\mathrm{ref}}} \frac{1}{N_{\mathrm{test}}} \sum_i \left|x_i^{(h)}\right|, \tag{11}$$

uses the observed futures of the test trajectories and is therefore a descriptive reference rather than an out-of-sample competitor. To separate the screening decision from the accuracy of individual predictions, we also record, for each of the kNN and constant predictors, the normalized individual-trajectory error

$$E(h) = \frac{1}{\bar{x}_{\mathrm{ref}}} \left[\frac{1}{N_{\mathrm{test}}} \sum_i \left(x_i^{(h)} - \hat{x}_i^{(h)}\right)^2 \right]^{1/2}, \tag{12}$$

where $\hat{x}_i^{(h)}$ denotes the prediction of the predictor being evaluated. This error measures individual-trajectory prediction accuracy, not the accuracy of the aggregate screening decision. The comparison uses the noise-free Cases A–C, the same seeds, and the exponential and Mittag–Leffler families.

For the observational-noise experiments, the same profile definition is used with the perturbed observations in both the numerator and the reference-state denominator; the noise protocol is specified in Section 2.5. The candidate families and their optimization procedures are specified in Section 2.4.

2.4. Candidate Model Families and Optimization

Because a compatibility result is meaningful only relative to the alternatives considered, each profile is compared with several reference families fitted under a common criterion. This subsection specifies the objectives, the candidate families, and the optimization. The candidate laws are fitted only to the profile points in the fitting interval. The residual objective is defined on the profile itself,

$$Q_{\mathrm{lin}} = \frac{1}{n}\sum_{i=1}^{n}\left[G_i - \widehat{G}(t_i)\right]^2, \tag{13}$$

where $G_i = D_{\mathrm{pred}}(h_i)$ and $t_i = t_{h_i}$ is the absolute target time defined in Equation (9). The stable candidate families are

$$G_{\mathrm{exp}}(t) = C\exp(-\lambda t), \tag{14}$$

$$G_{\mathrm{ML}}(t) = CE_{\alpha}(-\lambda t^{\alpha}), \tag{15}$$

$$G_{\mathrm{str}}(t) = C\exp(-\lambda t^{\beta}), \tag{16}$$

$$G_{\mathrm{RH}}(t) = \frac{C}{1+(t/\mu)^{\beta}}, \tag{17}$$

$$G_{2\mathrm{exp}}(t) = C_1\exp(-\lambda_1 t) + C_2\exp(-\lambda_2 t), \qquad 0 < \lambda_1 < \lambda_2. \tag{18}$$

The exponential and Mittag–Leffler models are the common baseline comparison in Cases A–C. The biexponential model is included in the focused noise-free specificity comparisons of Cases A and B at $\alpha = 0.8$ and Case C at $s = 3$, as well as in the noisy three-family comparisons of Section 2.5. The generalized-exponential and rational/Hill families are used in the separable nonfractional controls described in Section 2.2.2.

All fits were performed using MATLAB R2016a (version 9.0.0.341360) and its `fminsearch` implementation of the Nelder–Mead algorithm, which is unconstrained. Positivity is therefore enforced by exponential parameterizations and finite parameter intervals by an objective penalty; Table 1 lists the transformations, admissible regions, and starting values actually used. In particular, the parameterization $\alpha = \exp(p_3)$ enforces only $\alpha > 0$, and the restriction $0.05 \le \alpha \le 1$ is imposed separately by assigning the objective the value 10^{100} outside that interval.

Table 1. Candidate-model optimization settings. All fits use MATLAB `fminsearch` (Nelder–Mead) with `MaxFunEvals` = 5000, `MaxIter` = 2000, and `TolFun` = `TolX` = 10^{-10}. The solution with the smallest finite objective over the listed starts is retained. All fits minimize Q_{lin}.

Model	Parameterization and Admissible Region	Initialization Actually Used
Exponential	$C = e^{p_1}$, $\lambda = e^{p_2}$; $C, \lambda > 0$ by construction.	One start: $C_0 = \max(G_1, 10^{-8})$, $\lambda_0 = 0.1$.
Mittag–Leffler	$C = e^{p_1}$, $\lambda = e^{p_2}$, $\alpha = e^{p_3}$. The log transform enforces only positivity; $0.05 \le \alpha \le 1$ is imposed separately by assigning the objective the value 10^{100} outside the interval.	Six *paired* starts, all with $C_0 = \max(G_1, 10^{-8})$: $(\lambda_0, \alpha_0) = (0.05, 0.2), (0.10, 0.5), (0.15, 0.7), (0.20, 0.85), (0.25, 0.95), (0.30, 1.0)$.
Generalized exponential	$C = e^{p_1}$, $\lambda = e^{p_2}$, $\beta = e^{p_3}$; $0.05 < \beta < 3$ imposed by objective penalty 10^{100}.	Three paired starts: $(\lambda_0, \beta_0) = (0.05, 0.30), (0.10, 0.50), (0.20, 0.80)$, with $C_0 = \max(G_1, 10^{-8})$.
Rational/Hill	$C = e^{p_1}$, $\beta = e^{p_2}$, $\mu = e^{p_3}$; $0.05 < \beta < 10$ and $\mu > 0.01$ imposed by objective penalty 10^{100}.	6×6 Cartesian grid (36 starts): $\beta_0, \mu_0 \in \{0.25, 0.4, 1.0, 1.5, 3.0, 4.0\}$, with $C_0 = \max(G_1, 10^{-8})$.
Biexponential	$C_1 = e^{p_1}$, $C_2 = e^{p_2}$, $\lambda_1 = e^{p_3}$, $\lambda_2 = \lambda_1 + e^{p_4}$; thus $C_1, C_2 > 0$ and $0 < \lambda_1 < \lambda_2$ hold by construction.	$2 \times 3 \times 3$ Cartesian grid (18 starts): $w \in \{0.3, 0.7\}$, $\lambda_{1,0} \in \{0.02, 0.05, 0.1\}$, $(\lambda_2 - \lambda_1)_0 \in \{0.1, 0.3, 1.0\}$; $C_{1,0} = wG_1$, $C_{2,0} = (1-w)G_1$.

The exponential baseline, a two-parameter monotone family, uses a single start; the curved families use multistart fitting, because different initializations can terminate at

different solutions and objective values. For the stable Mittag–Leffler fits in Cases A and B, the final objectives across the six starts spread by approximately 4.0×10^{-4} and 4.4×10^{-4}, while the best solution is many orders of magnitude smaller; this spread confirms the need for multistart optimization, and the reported fit is the best finite-objective solution across starts. For the biexponential fits, 12/18, 9/18, and 3/18 starts terminated with `exitflag` = 1 in Cases A, B, and C, the others at the iteration/function-evaluation limit; nevertheless, the final objectives differed by at most 7.12×10^{-8}, 7.21×10^{-8}, and 7.88×10^{-10}. Numerically invalid model evaluations receive a large penalty, the start with the lowest objective is always retained, and complete per-start diagnostics are reported in Supplementary Table S1 and in the archived optimizer logs. The Mittag–Leffler function is evaluated with the algorithm of Garrappa [21].

For model comparison, RMSE is computed on the profile scale. AIC and BIC use the nominal number of fitted parameters and serve only as descriptive summaries (Section 2.1); at parameter boundaries or degenerate limits the effective model dimension may be smaller than the nominal count. In the matched rational/Hill control, where the residuals are at numerical-precision scale, information-criterion values are not assigned statistical meaning.

2.5. Observational-Noise and Uncertainty Protocol

The benchmarks of Sections 2.2 and 2.3 are noise-free. To assess whether the screen remains informative under observational noise, the analysis is repeated across independent noise realizations rather than relying on a single noisy run. This subsection specifies the noise model and the uncertainty analysis. Noise is added to the analytical Case A benchmark (common rate $\mu = 0.15$) as observation noise only, without process noise in the underlying dynamics: for trajectory j, the clean samples are perturbed according to

$$x_j^{\mathrm{obs}}(t_n) = x_j(t_n) + \varepsilon_{j,n}, \qquad \varepsilon_{j,n} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_j^2), \qquad \sigma_j = 0.03\,|x_{0,j}|. \tag{19}$$

Thus, the noise standard deviation is fixed in time along each trajectory and equals 3% of that trajectory's initial amplitude; it is not 3% of the instantaneous state magnitude. Noise samples are independent across time and across trajectories.

For each random seed, the trajectory-level training/test partition is generated before the noise realization is added, after which the complete ensemble is perturbed. Consequently, both the kNN training data and the kNN test data contain noisy observations. The relaxation profile retains the definition in Equation (8), but in the noisy experiment both the predicted future amplitudes in the numerator and the reference-state amplitudes in the denominator are formed from the perturbed observations. The same training-only standardization and direct-horizon kNN construction are then applied as in the noise-free stable analysis.

The uncertainty study uses 50 independent random seeds for each generating order $\alpha \in \{0.6, 0.7, 0.8, 0.9, 1\}$. The runs are paired across α: for a given seed, the initial conditions, trajectory partition, and additive noise realization are kept the same, while only the clean generating trajectory changes with α. Consequently, differences across generating orders cannot be attributed to different ensembles, train/test partitions, or noise realizations; the pairing isolates the change associated with the clean relaxation law as α varies.

For each seed we record the fitted order α_{fit} and the horizon-extrapolation difference

$$\Delta\mathrm{RMSE} = \mathrm{RMSE}^{*}_{\mathrm{exp}} - \mathrm{RMSE}^{*}_{\mathrm{ML}}, \tag{20}$$

where the superscript $*$ denotes evaluation on horizons 21–30 without refitting. Positive values indicate a smaller horizon-extrapolation RMSE for the Mittag–Leffler model. Across the 50 paired runs, we report the mean together with the 95% confidence interval for the

mean, based on the Student t distribution with 49 degrees of freedom, of both ΔRMSE and α_{fit}, and separately count wins, losses, and exact ties. Operationally, a tie is recorded when $|\Delta\mathrm{RMSE}| < 10^{-8}$ or when α_{fit} lies within 10^{-6} of 1, where the Mittag–Leffler model reduces numerically to the exponential model.

To complement the across-realization analysis with an estimate of test-ensemble variability conditional on a fixed fitted predictor, a trajectory bootstrap is performed for one representative noisy run. The fitted kNN training set is held fixed, while the test trajectories are resampled with replacement. For each of $B = 500$ bootstrap resamples, the noisy relaxation profile is recomputed, the exponential and Mittag–Leffler models are refitted on horizons 1–20, and ΔRMSE is reevaluated on horizons 21–30. The reported bootstrap interval is the percentile 95% interval of these 500 values. It therefore quantifies uncertainty conditional on the estimated training predictor and the selected representative run. It is not interpreted as a bootstrap confidence interval for the full pipeline, because the training trajectories, trajectory split, and model-building stage are not resampled.

The noise-free specificity analysis showed that a classical distributed-rate ensemble can favor the Mittag–Leffler law over a single exponential, and that this preference disappears once a biexponential alternative is admitted (Section 3.3). To test whether this conclusion survives observational noise, the same noise model, the same 50 seeds, and the same fitting and extrapolation horizons are applied to two classical generators: Case B with $\alpha = 1$, whose trajectories are exponentials with rates $\mu_j \sim \mathcal{U}[0.05, 0.25]$, and Case C with $s = 3$. Each noisy profile is compared first with the exponential and Mittag–Leffler families, as above, and then with the exponential, Mittag–Leffler, and biexponential families. A broader candidate set that rejects the Mittag–Leffler law for classical data could also do so for fractional data. To check that the three-family comparison retains sensitivity to fractional relaxation, it is also applied to the fractional Case B with $\alpha = 0.8$; this single reference is a targeted check, not a general assessment of sensitivity.

The selection rule was fixed before the results were examined. For each seed, the family with the lowest horizon-extrapolation RMSE is selected, and families whose RMSE differs from the lowest by less than 10^{-8} are counted as tied. The convention that a Mittag–Leffler fit with $|\alpha_{\mathrm{fit}} - 1| < 10^{-6}$ coincides with the exponential applies only to that pair and does not create a tie with the biexponential. Selection frequencies are reported with 95% Wilson score intervals, and the paired held-out RMSE differences between families with 95% Student t confidence intervals for their means. Seeds with $|\alpha_{\mathrm{fit}} - 1| < 10^{-6}$ are also counted as boundary hits, separately from ties defined by RMSE. For the binary comparisons, exact sign tests and Wilcoxon signed-rank tests on the paired differences, excluding ties under the stated binary rule, are reported as supporting, unadjusted analyses. Because the biexponential is the most flexible family, its fits are monitored for convergence and degeneracy (a normalized component weight or relative rate gap below 10^{-3}); these diagnostics are reported in the Supplementary Material.

2.6. Stable Sensitivity Protocol

The practical value of the screen also depends on its robustness to reasonable analysis choices. To assess this, the stable pipeline is rerun on Case B at the representative generating order $\alpha = 0.8$, changing one analysis setting at a time relative to the baseline configuration

$$K = 3, \qquad m_{\mathrm{hist}} = 7, \qquad \Delta t = 0.1, \qquad \text{fitting horizons 1–20.}$$

The four one-at-a-time perturbations are

$$K \in \{1, 3, 5, 10\}, \tag{21}$$

$$m_{\mathrm{hist}} \in \{4, 7, 10\}, \tag{22}$$

$$\text{fitting horizons} \in \{1\text{–}15,\ 1\text{–}20,\ 1\text{–}25\}, \tag{23}$$

$$\Delta t \in \{0.05, 0.1, 0.2\}. \tag{24}$$

This design provides a local sensitivity assessment around the baseline configuration rather than an exhaustive exploration of interactions among analysis choices. Case B is used because, unlike Case A, it is not reduced to a single temporal template by separability. The same underlying data-generation setting and the same ten random seeds are used across configurations, so each row is compared against the same baseline experiment.

For the sampling-interval perturbation, the physical history, forecast, and fitting windows are kept fixed rather than keeping the number of samples fixed. Consequently, changing Δt also changes the discrete history length and the number of forecast/fitting horizons. Specifically,

$$\Delta t = 0.05: \quad m_{\text{hist}} = 13, \quad H_{\max} = 60, \quad \text{fitting horizons 1–40},$$

whereas

$$\Delta t = 0.2: \quad m_{\text{hist}} = 4, \quad H_{\max} = 15, \quad \text{fitting horizons 1–10}.$$

These choices preserve the baseline physical windows exactly: the history span is $(m_{\text{hist}} - 1)\Delta t = 0.6$, the maximum forecast span is $H_{\max}\Delta t = 3.0$, and the fitting span is 2.0 time units in all three sampling configurations. Thus, the sampling-interval rows are not independent perturbations of m_{hist}: they probe a change in temporal resolution while preserving the physical windows.

For each configuration, the same stable fitting procedure described in Sections 2.3 and 2.4 is applied, and the fitted order α_{fit}, the horizon-extrapolation ΔRMSE of Equation (20), and the descriptive information-criterion difference are summarized over the repeated seeds. Since the number of fitted profile points changes when Δt or the fitting range changes, AIC differences are interpreted only within configurations with the same number of fitted points and are not compared numerically across all sensitivity rows.

2.7. Numerical Reliability of a Single-Record Example

A screen applied to simulated observations can be no more reliable than the simulation that produced them. This matters for the single-record example, a scalar series generated from a nonlinear fractional system initially investigated as a chaotic test case. The system is the fractional system analyzed in [22],

$$\begin{cases} {}_C D^{\alpha}_{0,t} x = ax - by - yz, \\ {}_C D^{\alpha}_{0,t} y = cx, \\ {}_C D^{\alpha}_{0,t} z = -dz + y^2, \end{cases} \tag{25}$$

with $a = -2$, $b = -6.4$, $c = 1$, and $d = 1$; related Caputo fractional-order chaotic systems have also been investigated in [23]. It is integrated to $T = 200$ from $\mathbf{x}_0 = (0.2, 0.2, 0.2)^{\mathsf{T}}$ with $\alpha = 0.916$, and the x-component is observed after the first 40 time units are discarded as transient.

The record was originally generated with the explicit product-integration (fractional Euler) scheme [24] at step $\Delta t = 0.01$. To verify whether the regime it produces is robust, we use Garrappa's MATLAB implementation `fde12` of the Adams–Bashforth–Moulton method [24], both in predictor-only mode, which reproduces the original explicit discretization, and in predictor–corrector (PECE) mode, with $\Delta t \in \{0.01, 0.005, 0.0025, 0.00125\}$. To assess whether the observed regime persists under numerical refinement, we compare both summary diagnostics and differences between numerical solutions: the standard

deviation of x over $t \geq 40$ and over the final window $180 \leq t \leq 200$, the distance of the final state from the nearest nonzero equilibrium, and the root-mean-square differences between PECE solutions at successive step sizes on a common grid. A run with an iterated corrector at $\Delta t = 0.005$ and a control run at $\alpha = 0.93$ complete the check.

These diagnostics are complemented by the local stability of the nonzero equilibria $E_{\pm} = (0, \pm\sqrt{6.4}, 6.4)$. For a commensurate Caputo system, an equilibrium is locally asymptotically stable if every eigenvalue λ_j of the Jacobian satisfies $|\arg \lambda_j| > \alpha\pi/2$ [2]. This criterion characterizes local behavior only and does not determine the global dynamics from a given initial condition. The kNN screening protocol originally applied to this record is documented in the Supplementary Material.

3. Results

The results progress from matched-family calibration through increasingly demanding tests of heterogeneity and specificity, followed by comparisons of profile constructions, observational noise, and analysis sensitivity. The section concludes with a numerical-reliability check of the single-record example.

3.1. Matched-Family Calibration

Case A provides a matched, noise-free calibration of the complete stable pipeline. Because the trajectories are generated analytically in an exactly matched setting, the fitted-order discrepancy is observed at the 10^{-11} scale. Across $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$, the fitted Mittag–Leffler order therefore reproduces the generating value to numerical precision, and the fitted rate is $\lambda_{\mathrm{ML}} = 0.150000$ in every case. The corresponding Mittag–Leffler horizon-extrapolation residuals are likewise at numerical-precision scale (Table 2); they are not interpreted as evidence of statistical discrimination. At $\alpha = 1$, the fitted order reaches the exponential limit and the two families coincide to numerical precision. Representative profiles for Cases A–C are shown in Figure 2.

Table 2. Exact matched calibration (Case A). Values of α_{fit}, RMSE, ΔAIC, and ΔBIC are shown for the first seed; the maximum absolute order deviation is taken over ten independent seeds. The seed-to-seed dispersion of α_{fit} is at the 10^{-11} scale. RMSE* is evaluated on horizons 21–30 without refitting. Negative ΔAIC and ΔBIC favor the Mittag–Leffler model.

α	α_{fit}	$\max \lvert\alpha_{\mathrm{fit}} - \alpha\rvert$	$\mathrm{RMSE}^*_{\mathrm{exp}}$	$\mathrm{RMSE}^*_{\mathrm{ML}}$	ΔAIC	ΔBIC
0.6	0.600000	1.5×10^{-11}	1.485×10^{-2}	5.393×10^{-14}	−964.07	−963.07
0.7	0.700000	1.0×10^{-11}	1.388×10^{-2}	2.925×10^{-13}	−919.13	−918.14
0.8	0.800000	2.2×10^{-11}	1.124×10^{-2}	2.040×10^{-13}	−922.00	−921.01
0.9	0.900000	2.7×10^{-11}	6.683×10^{-3}	1.316×10^{-12}	−889.10	−888.10
1.0	1.000000	1.2×10^{-10}	2.694×10^{-12}	5.295×10^{-13}	–	–

At $\alpha = 1$, the Mittag–Leffler family reduces to the exponential special case. Both residuals are at numerical-precision scale, so information-criterion differences are not interpreted.

Because the generator and fitted family coincide and the Case A ensemble is multiplicatively separable, this near-exact recovery is treated as a matched-family calibration result rather than as evidence of general fractional-order identifiability.

3.2. Heterogeneous Relaxation Rates

Case B breaks the multiplicative separability of Case A by assigning each trajectory its own relaxation rate $\mu_j \sim \mathcal{U}[0.05, 0.25]$ while keeping a common generating order. The fitted Mittag–Leffler order remains close to the generating value across the sweep, but with a small systematic negative offset of approximately 0.02 (Table 3). The fitted rate remains close to the mean generating rate, with $\lambda_{\mathrm{ML}} \approx 0.149$. Once separability is removed, the fitted

order in this heterogeneous ensemble thus behaves as an effective profile parameter rather than reproducing the generating order exactly.

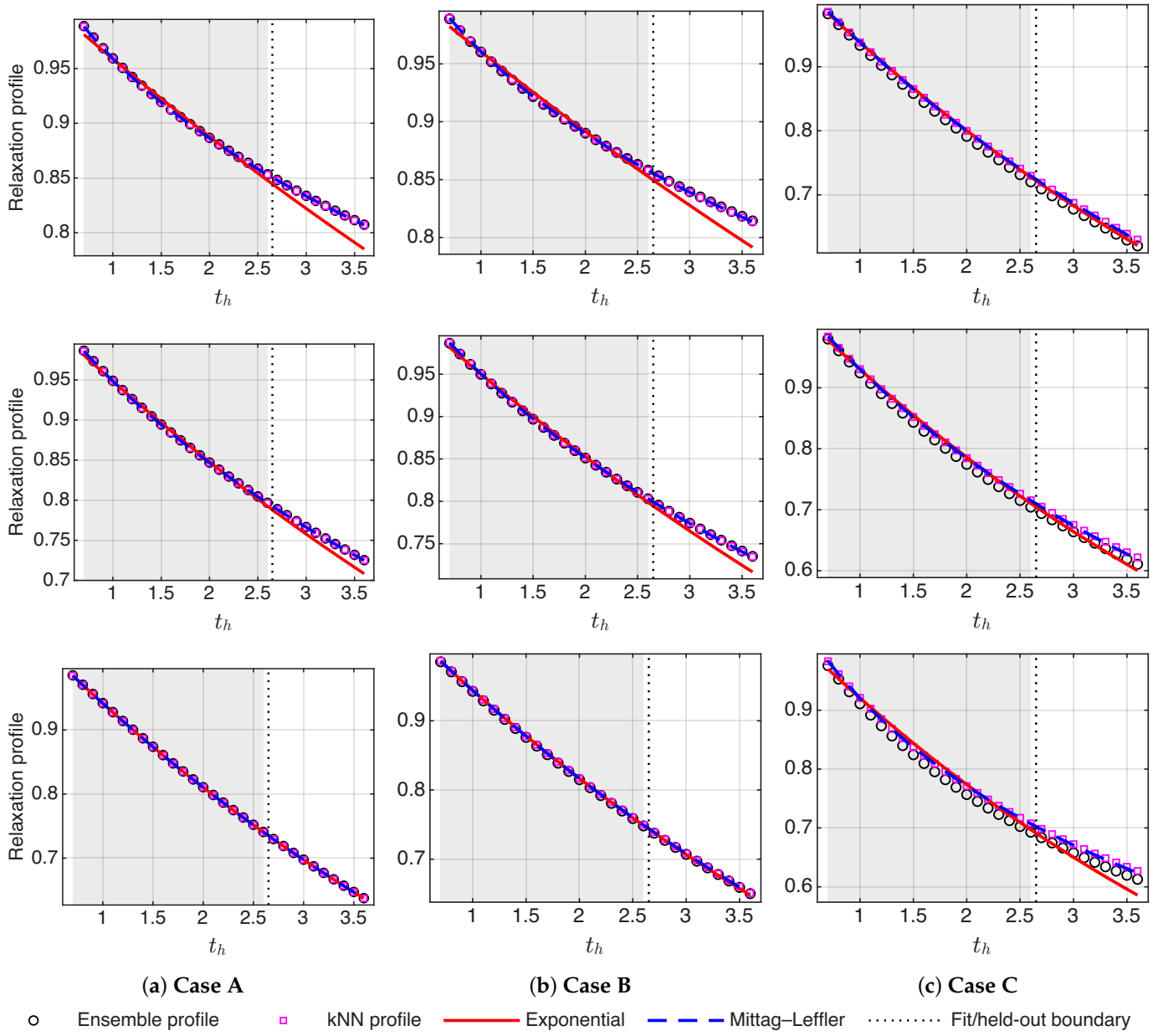


Figure 2. Stable relaxation profiles for the three benchmark cases. Each panel shows the directly computed ensemble profile and the kNN-predicted profile, together with the exponential and Mittag–Leffler laws fitted on horizons $h = 1$–20 (shaded) and continued over the held-out horizons $h = 21$–30, on the absolute target time $t_h = t_{\mathrm{ref}} + h\,\Delta t$. (**a**) Case A, homogeneous rate $\mu = 0.15$; (**b**) Case B, heterogeneous rates $\mu_j \sim \mathcal{U}[0.05, 0.25]$; in both columns, from top to bottom, $\alpha = 0.6$, 0.8 and 1. (**c**) Case C, purely exponential trajectories with log-uniform rates (Equation (5)); from top to bottom, $s = 3$, 5 and 10.

The Mittag–Leffler model has lower horizon-extrapolation error than the single exponential for every generating order in this heterogeneous ensemble. Notably, this remains true at $\alpha = 1$, where every individual trajectory is exponential but the ensemble contains a distribution of rates: the fitted order is 0.9768 ± 0.0009 and $\Delta\mathrm{AIC} = -94.6 \pm 0.4$. This integer-order case anticipates the more stringent classical distributed-rate specificity test in Section 3.3.

Table 3. Case B with heterogeneous relaxation rates $\mu_j \sim \mathcal{U}[0.05, 0.25]$. RMSE values are for the first random seed; α_{fit} and ΔAIC are mean $\pm$ standard deviation over ten seeds. RMSE* is evaluated on horizons 21–30 without refitting. Negative ΔAIC favors the Mittag–Leffler model.

α	α_{fit}	$\alpha_{\text{fit}} - \alpha$	$\text{RMSE}^*_{\text{exp}}$	$\text{RMSE}^*_{\text{ML}}$	ΔAIC
0.6	0.5828 ± 0.0010	-0.017	1.496×10^{-2}	5.508×10^{-5}	-245.2 ± 2.8
0.7	0.6805 ± 0.0010	-0.020	1.423×10^{-2}	9.654×10^{-5}	-222.4 ± 2.5
0.8	0.7789 ± 0.0010	-0.021	1.202×10^{-2}	1.606×10^{-4}	-198.3 ± 2.0
0.9	0.8777 ± 0.0010	-0.022	7.933×10^{-3}	2.390×10^{-4}	-167.0 ± 1.7
1.0	0.9768 ± 0.0009	-0.023	1.882×10^{-3}	3.449×10^{-4}	-94.6 ± 0.4

3.3. Specificity Tests and Candidate-Set Comparisons

The classical distributed-rate control in Case C first exposes a limitation of the binary exponential-versus–Mittag–Leffler comparison. All trajectories are strictly exponential, but their rates are drawn from the log-uniform family defined in Section 2.2.1. Nevertheless, the Mittag–Leffler model is preferred over the single exponential in all ten runs at each tested spread factor (Table 4). As the rate distribution widens from $s = 3$ to $s = 10$, the fitted order decreases monotonically from 0.933 to 0.668. The separation between the exponential and Mittag–Leffler continuations over the held-out horizons grows with s (Figure 2c).

Table 4. Case C: binary exponential-versus–Mittag–Leffler comparison for classical distributed-rate ensembles. The trajectories are purely exponential, with $\mu_j = 0.15 s^{u_j}$, $u_j \sim \mathcal{U}[-1, 1]$. $\Delta\text{RMSE} = \text{RMSE}^*_{\text{exp}} - \text{RMSE}^*_{\text{ML}}$, in units of 10^{-3}; positive values indicate lower horizon-extrapolation RMSE for the Mittag–Leffler fit. Values are mean $\pm$ standard deviation over ten seeds.

s	Rate Range	α_{fit}	ΔRMSE	ML Wins
3	$[0.05, 0.45]$	0.933 ± 0.003	5.36 ± 0.35	10/10
5	$[0.03, 0.75]$	0.847 ± 0.008	12.53 ± 1.27	10/10
10	$[0.015, 1.5]$	0.668 ± 0.022	26.43 ± 3.20	10/10

The same comparison changes when a more flexible classical relaxation family is included. Table 5 compares the exponential, Mittag–Leffler, and biexponential fits for Case A and Case B at $\alpha = 0.8$, and for Case C at $s = 3$. In Case B, the Mittag–Leffler model remains the lowest-error model on the horizon-extrapolation interval in all ten runs. Case A shows the same ranking, but it is the matched-family, noise-free calibration of Section 3.1, so its machine-precision Mittag–Leffler residuals are not interpreted as a statistically meaningful advantage over the biexponential model. In Case C, by contrast, the biexponential model is best in all ten runs: its mean held-out RMSE is 0.0129×10^{-3}, compared with 0.992×10^{-3} for the Mittag–Leffler model and 6.36×10^{-3} for the single exponential. Thus, in the noise-free setting, the preference observed in the binary Case C comparison does not persist after expansion of the candidate set. For Case C, the apparent Mittag–Leffler preference is therefore a property of this restricted candidate comparison, not evidence that the classical ensemble is fractionally generated. Section 3.5 examines whether this conclusion survives observational noise.

The separable nonfractional shape controls give two complementary checks. For the stretched-exponential generator ($\beta = 0.4$), the Mittag–Leffler fit produces an interior fitted order below one and markedly improves on the single exponential. For the compressed control ($\beta = 2$), the fitted order reaches $\alpha_{\text{fit}} = 1$, so the Mittag–Leffler model reduces to the exponential limit; however, neither family describes this control well, with training RMSE 2.925×10^{-2} and horizon-extrapolation RMSE 9.913×10^{-2} for both. The boundary solution therefore reflects failure of the tested exponential/Mittag–Leffler families to capture this compressed shape rather than successful discrimination.

Table 5. Expanded candidate comparison for Case A and Case B at $\alpha = 0.8$, and Case C at $s = 3$. Held-out RMSE values are in units of 10^{-3} and are mean $\pm$ standard deviation over ten seeds. "Wins" gives the number of runs in which the model has the lowest held-out RMSE. Case A is included only as a matched-family calibration reference; its machine-precision Mittag–Leffler residuals are interpreted as in Section 3.1, not as evidence of statistical discrimination.

Case	Model	Held-Out RMSE	Wins
A	Exponential	11.242 ± 0.002	0/10
	Mittag–Leffler	$<10^{-9}$	10/10
	Biexponential	0.7285 ± 0.0002	0/10
B	Exponential	12.07 ± 0.32	0/10
	Mittag–Leffler	0.181 ± 0.011	10/10
	Biexponential	0.764 ± 0.022	0/10
C	Exponential	6.36 ± 0.43	0/10
	Mittag–Leffler	0.992 ± 0.079	0/10
	Biexponential	0.0129 ± 0.0013	10/10

The rational/Hill control serves a different purpose because the generating family itself is included among the candidates. The fit recovers $\beta_{\text{fit}} = 2.000000$ and $\mu_{\text{fit}} = 0.500000$, with training and horizon-extrapolation RMSEs of 4.764×10^{-13} and 1.279×10^{-13}, respectively. These machine-precision residuals are treated only as a same-family numerical recovery check; information-criterion differences at this scale are not assigned statistical meaning.

3.4. Contribution of the kNN Construction

Because the screening decision is taken on the aggregate profile, the kNN profile was compared with the constant training-mean predictor and the direct ensemble average (Section 2.3) in all thirteen noise-free configurations of Cases A–C, with ten seeds each. The three constructions lead to the same exponential-versus–Mittag–Leffler outcome in every seed of every configuration (Table 6). For example, in Case A at $\alpha = 1$, all ten seeds yield ties under each construction. Their fitted orders are also close: the largest difference in the mean fitted order, 0.015, occurs in Case C at $s = 5$, and in no configuration does it change the screening outcome. The constructions differ sharply, however, at the level of individual trajectories. The mean held-out individual error of the kNN predictor ranges from 0.002 in Case A to 0.077 in Case C at $s = 10$, whereas that of the constant predictor ranges from 0.29 to 0.50, between about 5.6 and 135 times larger. Part of this advantage reflects the amplitude information contained in each observed history, which the constant predictor ignores. The complete summary for all thirteen configurations is given in the Supplementary Material.

Table 6. Comparison of profile constructions in representative noise-free configurations (ten seeds each). α_{fit}: mean fitted order for the kNN, constant training-mean (Const), and direct ensemble-average (Direct) profiles. W/L/T: Mittag–Leffler wins, losses, and ties on the held-out horizons, identical for the three constructions. E: mean held-out individual-trajectory error, Equation (12), for the kNN and constant predictors; it is not defined for the direct average, which uses the observed futures.

Case	Parameter	α_{fit}			W/L/T	E
		kNN	Const	Direct		kNN/Const
A	$\alpha = 0.8$	0.800	0.800	0.800	10/0/0	0.0024/0.318
B	$\alpha = 0.8$	0.779	0.775	0.775	10/0/0	0.040/0.339
B	$\alpha = 1$	0.977	0.972	0.973	10/0/0	0.053/0.320
C	$s = 3$	0.933	0.924	0.926	10/0/0	0.063/0.354
C	$s = 5$	0.847	0.832	0.836	10/0/0	0.070/0.416
C	$s = 10$	0.668	0.659	0.667	10/0/0	0.077/0.499

3.5. Observational Noise and Uncertainty

The next question is whether the matched-family screening behavior of Case A persists once observational noise is introduced. With trajectory-scaled noise, the fitted order remains close to the generating value over the fractional cases, while the horizon-extrapolation advantage of the Mittag–Leffler fit decreases as α approaches the integer-order limit (Table 7). In contrast to Case B, no comparable systematic negative offset is observed here across $\alpha = 0.6$–0.9: the corresponding estimates are 0.607, 0.706, 0.805, and 0.904. For $\alpha = 0.6, 0.7, 0.8$, and 0.9, the 95% confidence interval for $\Delta\mathrm{RMSE} = \mathrm{RMSE}^{*}_{\mathrm{exp}} - \mathrm{RMSE}^{*}_{\mathrm{ML}}$ remains above zero across the fifty-seed experiment. At $\alpha = 1$, the mean difference is slightly negative, -0.39×10^{-3}, with a 95% confidence interval $[-0.74, -0.03] \times 10^{-3}$; 30 of the 50 runs are exact numerical ties, 15 favor the exponential model, and 5 favor the Mittag–Leffler model. All 30 ties at $\alpha = 1$ are boundary hits, with α_{fit} within 10^{-6} of 1; at $\alpha = 0.9$ a single seed reaches the boundary, and none does at lower orders.

Table 7. Noisy stable experiment with $\sigma_j = 0.03|x_{0,j}|$. Values are based on fifty independent seeds. $\Delta\mathrm{RMSE} = \mathrm{RMSE}^{*}_{\mathrm{exp}} - \mathrm{RMSE}^{*}_{\mathrm{ML}}$ is evaluated on horizons 21–30 without refitting and is reported in units of 10^{-3}. Positive values indicate lower horizon-extrapolation RMSE for the Mittag–Leffler model. Bracketed values are 95% confidence intervals for the mean across the fifty realizations.

α	α_{fit} [95% CI]	ΔRMSE [95% CI]
0.6	0.607 [0.585, 0.629]	11.48 [10.82, 12.15]
0.7	0.706 [0.689, 0.723]	10.49 [9.81, 11.16]
0.8	0.805 [0.791, 0.820]	7.87 [7.22, 8.53]
0.9	0.904 [0.892, 0.916]	3.48 [2.82, 4.14]
1.0	0.987 [0.981, 0.993] †	−0.39 [−0.74, −0.03] ‡

† The fitted order is constrained by $\alpha_{\mathrm{fit}} \leq 1$, and 30 of the 50 fits at $\alpha = 1$ lie on this boundary. The reported Student t interval summarizes the across-seed mean of the constrained fitted orders; it should not be interpreted as a confidence interval for the generating order or for an unconstrained estimator. ‡ For $\alpha = 1$, the 50 runs comprise 5 Mittag–Leffler wins, 15 exponential wins, and 30 exact numerical ties ($|\Delta\mathrm{RMSE}| < 10^{-8}$ or $|\alpha_{\mathrm{fit}} - 1| < 10^{-6}$).

Figure 3 displays the same held-out comparison together with the trajectory-bootstrap interval for one representative run. The two intervals quantify different sources of uncertainty: the first summarizes variation across fifty independently generated realizations, whereas the bootstrap interval is conditional on the estimated kNN predictor and fixed training set of a single representative realization. In the latter calculation, test trajectories are resampled with replacement and the profile and model fits are recomputed for each bootstrap sample.

The noise-free specificity result of Section 3.3 was then tested under the same noise model. In the binary comparison, the classical controls frequently favor the Mittag–Leffler law. In Case C with $s = 3$, it is selected in 46 of the 50 runs (92%, 95% Wilson interval [81, 97]%), and the mean ΔRMSE, 3.45×10^{-3} with 95% confidence interval $[2.86, 4.03] \times 10^{-3}$, is clearly positive. In Case B with $\alpha = 1$, it is selected in 30 runs (60%, [46, 72]%), with 10 losses and 10 ties; the wins outnumber the losses, but the confidence interval of the mean difference, $[-0.01, 0.79] \times 10^{-3}$, includes zero, so a mean advantage is not established. In this setting, 10 of the 50 fits reach the boundary $\alpha_{\mathrm{fit}} = 1$. The supporting sign and Wilcoxon tests for all noisy settings are reported in the Supplementary Material.

Table 8 reports the three-family comparison. Under noise, the biexponential no longer dominates the classical Case C as it did without noise: the Mittag–Leffler law is selected in 30 of the 50 runs (60%, [46, 72]%), the biexponential in 17, and the exponential in 3, and the Mittag–Leffler model has, on average, a lower held-out RMSE than the biexponential. In Case B with $\alpha = 1$, the selections are spread across the three families, with 7 three-way

ties in which the fitted families coincide numerically. In the fractional reference, Case B with $\alpha = 0.8$, the Mittag–Leffler law is selected in 36 of the 50 runs (72%, [58, 83]%). The Mittag–Leffler selection frequencies of the classical Case C and of the fractional reference therefore have overlapping intervals. In all three settings, the biexponential usually attains the lowest unpenalized fitting-horizon RMSE, whereas the horizon-extrapolation rankings often differ.

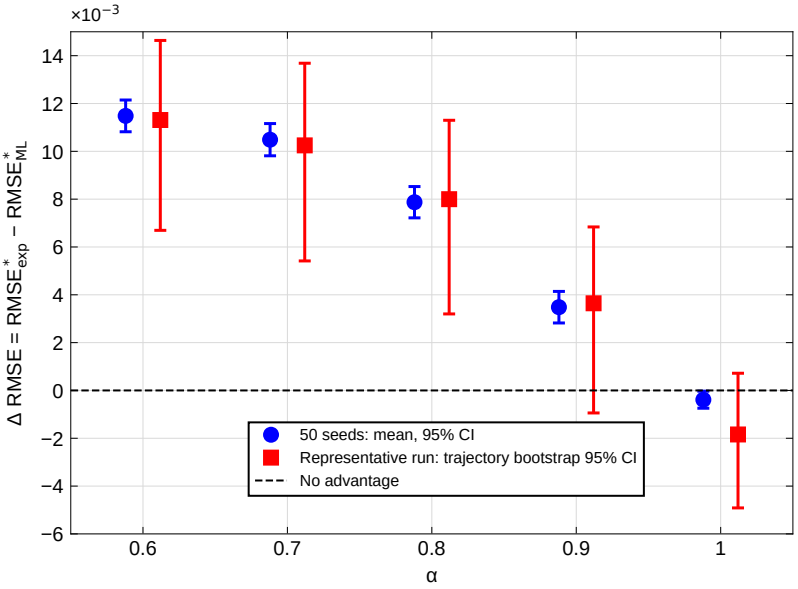


Figure 3. Held-out difference $\Delta\text{RMSE} = \text{RMSE}^*_{\text{exp}} - \text{RMSE}^*_{\text{ML}}$ versus the generating order for the noisy stable experiment. The first series shows the mean and the 95% confidence interval for the mean across fifty independent realizations. The second shows the percentile 95% trajectory-bootstrap interval for one representative realization, conditional on its estimated kNN predictor and fixed training set. The two intervals therefore quantify different sources of uncertainty and should not be read as competing estimates of the same sampling distribution.

Table 8. Three-family comparison under observational noise ($\sigma_j = 0.03|x_{0,j}|$, fifty seeds). Counts of runs in which each family has the lowest held-out RMSE, with 95% Wilson intervals for the selection frequency; ties: lowest RMSE shared within 10^{-8}. $\Delta_{\text{ML,Bi}} = \text{RMSE}^*_{\text{ML}} - \text{RMSE}^*_{\text{Bi}}$, mean with 95% confidence interval, in units of 10^{-3}; negative values favor the Mittag–Leffler model.

Setting	Exp	ML	Bi	Ties	$\Delta_{\text{ML,Bi}}$
Classical, Case B, $\alpha = 1$	8 [0.08, 0.29]	15 [0.19, 0.44]	20 [0.28, 0.54]	7	−0.14 [−0.36, 0.08]
Classical, Case C, $s = 3$	3 [0.02, 0.16]	30 [0.46, 0.72]	17 [0.22, 0.48]	0	−0.69 [−1.17, −0.22]
Fractional, Case B, $\alpha = 0.8$	1 [0.00, 0.10]	36 [0.58, 0.83]	13 [0.16, 0.40]	0	−1.56 [−2.16, −0.97]

3.6. Stable Sensitivity

To test whether the preference observed for heterogeneous Case B depends on one particular analysis configuration, we examine the Case B sensitivity study at $\alpha = 0.8$. It shows no change in the qualitative conclusion across the tested variations of K, history length, fitting interval, and sampling interval. Across all configurations, α_{fit} remains between 0.7756 and 0.7823, and the Mittag–Leffler fit has lower horizon-extrapolation RMSE than the single exponential in all ten seeds. The complete numerical results are reported in Supplementary Table S7.

As specified in Section 2.6, the sampling-interval rows preserve the physical windows rather than m_{hist}, and ΔAIC is compared only between configurations with the same number of fitted points.

3.7. Numerical Reliability of the Single-Record Example

In its original analysis, the single-record example was screened through its kNN forecast-error growth profile. Over the pre-transition interval, a saturating logistic curve, a classical sigmoidal reference whose relation to Mittag–Leffler-based sigmoids has been studied in [25,26], described that profile better than the exponential and Mittag–Leffler laws, but the ranking changed with the solver step (Supplementary Material). The verification of Section 2.7 shows why.

The predictor-only mode of `fde12` reproduces the original explicit runs: at $\Delta t = 0.01$, 0.005, and 0.0025, the post-transient standard deviations agree with those of the original computations to within 10^{-4}. As the step is refined, however, this regime changes qualitatively (Table 9): its late-window standard deviation decreases from 2.37 at $\Delta t = 0.01$ to 0.44 at $\Delta t = 0.00125$, and its final state approaches the nonzero equilibrium E_-. The PECE solutions show substantially greater stability under refinement: their summary diagnostics vary little over the same step-size range, and the root-mean-square differences between successive PECE solutions, 0.034, 0.0084, and 0.0022, decrease by factors close to four under step halving. The smaller late-window variability is consistent with a damped regime over the simulated time interval. An iterated corrector at $\Delta t = 0.005$ gives closely agreeing summary statistics, whereas the control at $\alpha = 0.93$ retains a sustained large-amplitude oscillatory regime over the simulated interval, with a post-transient standard deviation of about 2.44; the predictor–corrector scheme therefore does not generically suppress oscillations.

Table 9. Predictor-only and PECE integration of System (25) at $\alpha = 0.916$. SD: standard deviation of x over $t \geq 40$ (post-transient) and over $180 \leq t \leq 200$ (late window); distance: final distance from the nearest nonzero equilibrium.

Δt	Predictor-Only			PECE		
	SD, Post	SD, Late	Distance	SD, Post	SD, Late	Distance
0.01	2.345	2.367	5.04	0.929	0.276	0.492
0.005	2.353	2.322	3.80	0.926	0.274	0.489
0.0025	1.417	0.744	1.28	0.925	0.274	0.488
0.00125	1.117	0.444	0.577	0.925	0.273	0.488

The local stability analysis is consistent with these observations. At $E_\pm$, the Jacobian eigenvalues are approximately -3.482 and $0.241 \pm 1.902\mathrm{i}$, giving the critical order $\alpha_c = (2/\pi)\,|\arg(0.241 + 1.902\mathrm{i})| \approx 0.91985$. The order $\alpha = 0.916$ lies just below this threshold, so the nonzero equilibria are locally asymptotically stable, whereas $\alpha = 0.93$ lies above it. This local result does not determine the global dynamics, but together with the refinement evidence it shows that the sustained oscillatory regime of the original record is not numerically robust. The forecast-error profile and model ranking computed from that record are therefore withdrawn as evidence about the continuous fractional system; they are retained in the Supplementary Material only to document the sensitivity of the original discretization.

4. Discussion

4.1. Controlled Stable Experiments

The comparison of profile constructions (Section 3.4) clarifies the role of kNN in the stable screening results. In all tested noise-free configurations, the kNN, constant training-mean, and direct-average profiles led to the same exponential-versus–Mittag–Leffler decision, although their fitted orders differed slightly. This agreement suggests that features of the aggregate relaxation profiles sufficient for the binary screening decision are preserved across the three constructions under these conditions; the present experiments do

not identify which features are responsible. The kNN construction is therefore not necessary to reproduce these screening decisions, but it remains a useful practical construction: among the three alternatives tested, it alone is both out of sample and conditioned on the observed history of each test trajectory. It can therefore be applied to new trajectories when their histories, but not their futures, are available, and it predicts individual trajectories far more accurately than the constant predictor. Part of that advantage reflects the amplitude information contained in each history, and the comparison does not separate it from information about relaxation shape.

The near-exact recovery obtained in the matched Case A calibration (Section 3.1) is expected under the conditions of that experiment: the generating and fitted families coincide, the trajectories are noise-free, and the ensemble is multiplicatively separable. The result therefore provides a calibration check of the numerical implementation under matched conditions, but provides little evidence for general fractional-order identifiability.

Case B (Section 3.2) changes the interpretation because the ensemble profile is no longer generated by a single Mittag–Leffler law: individual trajectories share a common fractional order but carry different relaxation rates. The fitted ensemble-level order is systematically shifted downward relative to the trajectory-level generating order: the heterogeneous mixture yields an ensemble-level fitted shape parameter below the common generating order of its components. This offset should not be read as an estimation error relative to a uniquely defined ensemble-level order, because no such single-order generating law exists for the mixed profile.

The same direction of departure becomes more pronounced in the classical distributed-rate control of Case C (Section 3.3), where rate heterogeneity alone drives the fitted shape parameter well below the integer-order limit even though every trajectory is purely exponential. This progression motivates the specificity analysis in Section 4.2, where the role of the candidate model set is considered explicitly.

4.2. Specificity and Candidate-Set Breadth

Case C provides the direct test of candidate-set dependence. In this classical distributed-rate control (Section 3.3), the initial exponential-versus-Mittag–Leffler comparison favors the Mittag–Leffler family even though every trajectory is generated by a classical exponential mechanism. Without noise, that preference disappears once the biexponential family is included. The binary comparison therefore produced a candidate-set-dependent positive screen: the narrower candidate set made a classical mixture appear Mittag–Leffler-compatible. For $0 < \alpha < 1$, $E_\alpha(-t^\alpha)$ is completely monotone and admits a nonnegative spectral representation as a continuous superposition of exponential relaxations [4,5]. Relaxation shape alone is therefore, in general, insufficient to identify fractional origin uniquely; this provides a mathematical counterpart to the candidate-set dependence observed in Case C.

Case C thus directly demonstrates that model preference can change with candidate-set breadth. The monotone decrease of α_{fit} with increasing rate spread in Case C is consistent with the fitted shape parameter absorbing unresolved rate heterogeneity, but the present experiments do not establish that interpretation as a general mechanism.

Under observational noise, this recovery of specificity is incomplete (Section 3.5). With the biexponential admitted, the Mittag–Leffler law is still selected in 30 of 50 runs for the classical Case C ensemble, and it has, on average, the lower held-out error, much as in the fractional reference, where it is selected in 36 of 50 runs. Although these frequencies do not establish equivalent selection probabilities, their proximity in the tested configurations indicates that Mittag–Leffler selection alone provides limited discrimination between classical rate heterogeneity and fractional relaxation. The biexponential attains the

lowest fitting-horizon error in most runs yet is selected less often on the held-out horizons; this contrast is consistent with part of its additional flexibility being spent on the noise, although the experiments do not isolate that mechanism. Broadening the candidate set therefore reduces false-positive screens but does not, by itself, guarantee discrimination once the observations are noisy; the results indicate limited specificity under the tested conditions, not a general indistinguishability of the two mechanisms.

Operationally, a positive Mittag–Leffler screen should therefore be challenged, where possible, against flexible nonfractional alternatives before being interpreted as evidence favoring a fractional generating mechanism, although under observational noise such a challenge may not suffice. No finite candidate set can be exhaustive, so even after such comparisons a preference supports compatibility within the tested model set rather than unique mechanism identification.

4.3. Numerical Provenance of Simulated Data

Data-driven screening methods are often developed and tested on simulated observations, and their conclusions can then be no more reliable than the simulations themselves. The single-record example shows that numerical artifacts in simulated observations can propagate into screening conclusions: a coarse explicit discretization produced a sustained oscillatory regime on which the screening analysis returned a definite, though resolution-dependent, model ranking, whereas step refinement and an independent predictor–corrector scheme did not confirm that regime (Section 3.7). The example is particularly instructive because the chosen order, $\alpha = 0.916$, lies close to the local stability threshold $\alpha_c \approx 0.91985$ of the nonzero equilibria, where small numerical errors can plausibly have qualitative consequences. Simulated records intended for screening should therefore be checked under step refinement and, where possible, with an independent numerical scheme, especially when the parameters lie close to a stability boundary. Agreement across such checks does not by itself establish the scientific interpretation of a screening outcome, whereas disagreement prevents confident attribution to the continuous model without further numerical verification.

4.4. Scope and Limitations

The proposed workflow is a screening procedure for finite-window profile compatibility, not a mechanism-identification method: a positive screen indicates compatibility relative to the tested alternatives, and the fitted shape parameter should not be interpreted as a physical fractional order unless that interpretation is supported by independent modeling evidence.

The nonfractional controls include heterogeneous classical rate mixtures, biexponential alternatives, and separable curved shape controls. They are nevertheless not exhaustive. In particular, nonlinear classical generators and explicitly nonstationary alternatives remain outside the present study and constitute natural extensions of the specificity analysis. More generally, no practical analysis can include every competing family, so a positive screen cannot exclude all nonfractional explanations.

Observational noise introduces a further practical limitation. The Mittag–Leffler family converges to the exponential family as $\alpha \to 1$, so discrimination necessarily vanishes at the integer-order limit even for exact data. Observational noise makes the residual difference between the two families harder to resolve already before that limit is reached, despite the fitted order remaining close to the generating value over the fractional cases. A weak or vanishing preference near the integer-order regime should therefore not be interpreted as evidence against either family.

Finally, the uncertainty analysis in the stable noisy experiment has two different levels. Variation across repeated realizations is summarized over independent seeds, whereas the trajectory bootstrap for the representative run is conditional on the estimated training predictor and fixed training set. The latter is therefore not a full-pipeline uncertainty estimate.

The analyses added in this study have their own scope. The comparison of profile constructions covers thirteen noise-free configurations and includes no amplitude-aware predictor, so it neither extends automatically to noisy profiles nor separates amplitude from shape information. The noisy specificity analysis uses a single observational-noise level and three targeted generating settings, and its fractional reference, Case B with $\alpha = 0.8$, does not establish general sensitivity. Several noise-free analyses, including the comparison of profile constructions and the stable sensitivity study, rest on ten seeds per configuration. Formal paired hypothesis tests for fitted-order differences have not been conducted; the corresponding estimates and intervals are interpreted descriptively, particularly near the constrained integer-order boundary. Finally, the numerical-reliability check concerns one nonlinear system, one initial condition, and a limited selection of fractional orders. These limits bound the strength of any inferential or mechanistic claim drawn from the fitted profiles and should be taken into account when the method is used as a comparative screening tool.

5. Conclusions

This study develops and evaluates an equation-free, finite-window compatibility screen based on multi-horizon relaxation profiles constructed by kNN prediction and on the comparative fitting of reference temporal laws. The results support a deliberately limited interpretation of that screen. In the matched, noise-free calibration, the fitted Mittag–Leffler family reproduces the generating behavior to numerical precision; this result is a calibration check rather than evidence of general fractional-order identifiability.

When the individual trajectories retain a common fractional order but differ only in their relaxation rates, the ensemble-level fitted shape parameter moves systematically below the common generating order. The same directional effect becomes stronger for purely classical distributed-rate ensembles. The present results therefore show that relaxation-rate heterogeneity alone can produce an apparently Mittag–Leffler-compatible ensemble profile even when the individual components are purely exponential.

The central result is that model preference is conditional on the candidate set and on the observation conditions. Without noise, the apparent Mittag–Leffler preference obtained for a classical distributed-rate ensemble against a single exponential disappears when a flexible biexponential alternative is added. Under observational noise, broadening the candidate set reduces such false-positive screens but does not remove them: the Mittag–Leffler law is still selected in a majority of runs for the classical ensemble, at a frequency close to that obtained for a fractional reference. A positive Mittag–Leffler screen should therefore be interpreted only relative to the candidate families actually tested and the conditions of observation, and not as unique identification of the mechanism generating the data.

In the tested noise-free configurations, the binary screening decision is reproduced by simpler profile constructions, including a constant training-mean predictor and the directly observed ensemble average. The kNN construction is therefore not necessary for these decisions; its value lies in providing an out-of-sample, history-conditioned profile and substantially more accurate individual-trajectory predictions than the constant training-mean predictor.

The single-record example, initially investigated as a chaotic test case, is retained only as a cautionary illustration. Step refinement and an independent predictor–corrector scheme showed that the sustained regime produced by the original explicit discretization is not numerically robust, and its screening results are withdrawn as evidence about the continuous system. Screening conclusions drawn from simulated data therefore require careful verification of the numerical procedure used to generate those data.

Taken together, the results support the use of the method as a comparative finite-window compatibility screen. Its outcomes must be interpreted relative to the candidate set, with attention to relaxation-rate heterogeneity, and kept distinct from physical fractional-order identification; observational noise weakens discrimination both near the integer-order limit and between classical mixtures and fractional relaxation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/fractalfract10110745/s1. Table S1: Per-start biexponential optimizer diagnostics for Cases A and B at $\alpha = 0.8$ and Case C at $s = 3$ (main run); Table S2: Fitted parameters for the expanded candidate comparison of Table 5 (mean $\pm$ standard deviation over ten seeds); Table S3: Comparison of profile constructions in all thirteen noise-free configurations (ten seeds each); Table S4: Binary exponential-versus–Mittag–Leffler comparison under observational noise; Table S5: Paired held-out RMSE differences in the noisy three-family comparison; Table S6: Biexponential optimizer diagnostics in the noisy three-family comparison; Table S7: One-at-a-time sensitivity analysis on Case B at $\alpha = 0.8$; Table S8: Reproduction of the original explicit runs by the predictor-only mode of `fde12`; Table S9: Predictor–corrector (PECE) refinement at $\alpha = 0.916$; Table S10: Optimization settings of the original single-record fits; Table S11: Physical-time split of the original record; Table S12: Pre-transition model comparison for the original kNN forecast-error growth profile using the physical-time split with embargo; Table S13: Optimizer diagnostics for the original pre-transition fits at the three solver steps; Table S14: Sensitivity of the original pre-transition comparison to the training allocation of the physical-time split; Table S15: Solver-step comparison for the original fractional system; Table S16: Original record, $\Delta t = 0.01$: sensitivity of the pre-transition fits to the fitting window and to the error metric; Figure S1: kNN forecast-error growth profile of the original record (baseline split) with the exponential, Mittag–Leffler, and logistic laws fitted over the pre-transition interval; Figure S2: Identifiability of the Mittag–Leffler order for the baseline original fit; Figure S3: Solver-step comparison for the original fractional system. The Supplementary Material also cites the work of Apelblat [27].

Author Contributions: Conceptualization, N.N., A.V. and B.C.; methodology, N.N., A.V. and B.C.; software, N.N. and A.V.; validation, N.N., A.V. and B.C.; formal analysis, N.N., A.V. and B.C.; investigation, N.N. and A.V.; data curation, N.N., A.V. and B.C.; visualization, N.N.; writing—original draft preparation, N.N. and A.V.; writing—review and editing, N.N., A.V. and B.C.; supervision, B.C.; funding acquisition, B.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the European Regional Development Fund (ERDF) 2021–2027 under Project AI4AM—EFRE1052. This work was supported by the Open Access Publishing Fund of the Free University of Bozen-Bolzano.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The numerical experiments in this study are fully synthetic. The MATLAB scripts, generated data, and optimizer logs required to reproduce the reported results are openly available on Zenodo at https://doi.org/10.5281/zenodo.23494805. The repository README files map each script and output file to the corresponding tables and figures of the manuscript and the Supplementary Material. The Mittag–Leffler function is evaluated with the MATLAB implementation of Garrappa [21], which is not redistributed; the README indicates where to obtain it.

Acknowledgments: Bruno Carpentieri is a member of the Gruppo Nazionale per il Calcolo Scientifico (GNCS) of the Istituto Nazionale di Alta Matematica (INdAM).

Conflicts of Interest: The authors declare no financial conflicts of interest. B.C. serves as a Guest Editor of a Special Issue in *Fractal and Fractional*; he will not participate in the editorial handling, peer review, or decision-making process for this manuscript. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

AIC	Akaike information criterion
BIC	Bayesian information criterion
FDE	fractional differential equation
kNN	k-nearest neighbors
ML	Mittag–Leffler
RMSE	root mean square error

References

1. Podlubny, I. *Fractional Differential Equations*; Mathematics in Science and Engineering; Academic Press: San Diego, CA, USA, 1999; Volume 198.
2. Diethelm, K. *The Analysis of Fractional Differential Equations*; Lecture Notes in Mathematics; Springer: Berlin, Germany, 2010; Volume 2004. [CrossRef]
3. Herrmann, R. *Fractional Calculus: An Introduction for Physicists*, 2nd ed.; World Scientific: Singapore, 2014. [CrossRef]
4. Mainardi, F.; Gorenflo, R. On Mittag-Leffler-type functions in fractional evolution processes. *J. Comput. Appl. Math.* **2000**, *118*, 283–299. [CrossRef]
5. Mainardi, F. Why the Mittag-Leffler function can be considered the queen function of the fractional calculus? *Entropy* **2020**, *22*, 1359. [CrossRef] [PubMed]
6. Gorenflo, R.; Kilbas, A.A.; Mainardi, F.; Rogosin, S. *Mittag-Leffler Functions, Related Topics and Applications*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2020. [CrossRef]
7. Metzler, R.; Klafter, J. From stretched exponential to inverse power-law: Fractional dynamics, Cole-Cole relaxation processes, and beyond. *J. Non-Cryst. Solids* **2002**, *305*, 81–87. [CrossRef]
8. Granger, C.W.J.; Joyeux, R. An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1980**, *1*, 15–29. [CrossRef]
9. Delle Monache, D.; Grassi, S.; Santucci de Magistris, P. *Testing for Level Shifts in Fractionally Integrated Processes: A State Space Approach*; CREATES Research Paper 2015-30; Department of Economics and Business Economics, Aarhus University: Aarhus, Denmark, 2015.
10. Mainardi, F. A note on the equivalence of fractional relaxation equations to differential equations with varying coefficients. *Mathematics* **2018**, *6*, 8. [CrossRef]
11. Poinot, T.; Trigeassou, J.C. Identification of fractional systems using an output-error technique. *Nonlinear Dyn.* **2004**, *38*, 133–154. [CrossRef]
12. Li, X.; Rosenfeld, J.A. Fractional order system identification with occupation kernel regression. *IEEE Control Syst. Lett.* **2022**, *6*, 19–24. [CrossRef]
13. Zhang, T.; Lu, Z.R.; Liu, J.K.; Liu, G. Sparse identification of fractional chaotic systems based on the time-domain data. *Chin. J. Phys.* **2024**, *89*, 160–173. [CrossRef]
14. Duhé, J.F.; Victor, S.; Melchior, P.; Abdelmoumen, Y.; Roubertie, F. Recursive system identification of continuous-time fractional systems for all parameter estimation. *IEEE Trans. Control Syst. Technol.* **2024**, *32*, 2037–2049. [CrossRef]
15. Yuan, M.; Xu, W.; Liu, F.; Wang, L.; Lu, Y. Identification method for a fractional-order system in terms of equivalent dynamic properties. *Chaos* **2024**, *34*, 073119. [CrossRef] [PubMed]
16. Ben Brahim, H.; El Alaoui, F.Z.; Zguaid, K. Some results regarding observability and initial state reconstruction for time-fractional systems. *Int. J. Optim. Control. Theor. Appl. (IJOCTA)* **2024**, *14*, 99–112. [CrossRef]
17. Du, M.; Wang, Z.; Hu, H. Measuring memory with the order of fractional derivative. *Sci. Rep.* **2013**, *3*, 3431. [CrossRef] [PubMed]
18. Tarasov, V.E. On history of mathematical economics: Application of fractional calculus. *Mathematics* **2019**, *7*, 509. [CrossRef]
19. Wei, Z.; Yuan, N. Identifying climate memory impacts on climate network analysis: A new approach based on fractional integral techniques. *Clim. Dyn.* **2024**, *62*, 3465–3476. [CrossRef]

20. Velichko, A.; Belyaev, M.; Boriskov, P. A novel approach for estimating largest Lyapunov exponents in one-dimensional chaotic time series using machine learning. *Chaos* **2025**, *35*, 101101. [CrossRef] [PubMed]
21. Garrappa, R. Numerical evaluation of two and three parameter Mittag-Leffler functions. *SIAM J. Numer. Anal.* **2015**, *53*, 1350–1369. [CrossRef]
22. Sene, N. Analysis of a fractional-order chaotic system in the context of the Caputo fractional derivative via bifurcation and Lyapunov exponents. *J. King Saud Univ.-Sci.* **2021**, *33*, 101275. [CrossRef]
23. Sene, N. Theory and applications of new fractional-order chaotic system under Caputo operator. *Int. J. Optim. Control. Theor. Appl. (IJOCTA)* **2022**, *12*, 20–38. [CrossRef]
24. Diethelm, K.; Ford, N.J.; Freed, A.D.; Luchko, Y. Algorithms for the fractional calculus: A selection of numerical methods. *Comput. Methods Appl. Mech. Eng.* **2005**, *194*, 743–773. [CrossRef]
25. Hristov, J. The logistic models and sigmoid functions: A variety of models and solution perspectives. *Trans. Comput. Model. Intell. Syst.* **2025**, *1*, 10002. [CrossRef]
26. Hristov, J. Sigmoids based on Mittag-Leffler functions: Ideas, modeling, and computational experiments. *Trans. Comput. Model. Intell. Syst.* **2026**, *3*, 10073. [CrossRef]
27. Apelblat, A. *Integral Transforms and Volterra Functions*; Nova Science Publishers: New York, NY, USA, 2010.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.