



An Atangana-Baleanu-Caputo Fractal-Fractional Model for Alcohol-Addiction Dynamics: Qualitative Analysis, Numerical Simulation and a Physics-Informed Neural Network Surrogate

Sahib Yar^{1,*}, Khaleel Ahmad^{2,3}, Nida Najeeb¹

¹ Department of Mathematics and Statistics, University of Southern Punjab, Multan 60000, Pakistan

² Department of Mathematics and Statistics, The University of Lahore, Pakistan

³ Center for Theoretical Physics, Khazar University, 41 Mehseti Str., Baku, AZ1096, Azerbaijan

ARTICLE INFO

Article history:

Received 21 May 2026

Received in revised form 16 September 2026

Accepted 25 September 2026

Available online 26 September 2026

Keywords:

Physics-Informed Neural Network; Alcohol-Addiction Model; Atangana-Baleanu-Caputo Fractal-Fractional Derivative; Mittag-Leffler Kernel; Basic Reproduction Number; Global Stability

ABSTRACT

Alcohol dependence is a persistent public-health problem shaped by social contact and by long behavioural memory that local integer-order models cannot capture. This paper proposes a four-compartment alcohol-addiction model with susceptible, occasional-user, regular-user, and recovered classes in the Atangana-Baleanu-Caputo fractal-fractional framework, whose Mittag-Leffler kernel encodes nonlocal memory and whose fractal order encodes irregular temporal scaling. We prove existence, uniqueness, positivity, and boundedness, as well as derive the equilibria and the basic reproduction number. An Atangana-Toufik product-integration scheme and a Levenberg-Marquardt-trained physics-informed neural network confirm the analysis with sub-percent error, giving a consistent and flexible tool for studying memory-driven addiction dynamics and for informing intervention policy.

*Corresponding author.

E-mail address: sahibyar770@gmail.com

<https://doi.org/10.66972/iscis202619>

1. Introduction

Alcohol misuse remains one of the heaviest burdens on global public health, and its consequences reach far beyond the individual to families, workplaces and entire communities. Sustained heavy drinking is associated with liver disease, cardiovascular and neurological damage, mental-health deterioration, traffic injury, domestic violence and premature death, so that understanding how drinking habits spread, intensify and remit is a pressing social question. Because the onset and persistence of drinking are strongly mediated by peer influence and social environment, the problem is naturally amenable to the same compartmental reasoning that underlies mathematical epidemiology, a viewpoint developed at length in the standard accounts of Brauer *et al.* [1] and of Hethcote [2].

Compartmental models partition a population into classes and track the flows between them, and they have proven remarkably effective for infectious diseases as well as for behavioural epidemics such as smoking, drug use and alcoholism. Two ideas from this tradition are central to the present work. The first is the threshold concept of the basic reproduction number, whose general definition and computation were placed on a firm footing by Diekmann *et al.* [3]. The second is the next-generation matrix construction of Van den Driessche & Watmough [4], which provides a systematic route to that threshold for multi-compartment systems and which we adopt below. Within this lineage, a number of alcohol-specific models have appeared, Li *et al.* [5] analysed the global dynamics of alcohol consumption under treatment intervention. Bhosale *et al.* [6] studied a fractional-order consumption model that explicitly distinguishes binge Wanassi & Baleanu [7] together with Prakash *et al.* [8] modelled blood alcohol concentration through fractional operators. These studies confirm that drinking behaviour carries hallmarks of an epidemic process. However, most of them rest on classical integer order derivatives, which are local in time and therefore cannot represent the cumulative, history-dependent nature of habit formation.

Fractional calculus supplies precisely the missing ingredient, because fractional derivatives are nonlocal and weight the entire past of a process when determining its present rate of change. The analytical foundations are by now classical and are documented in the monographs of Podlubny [9], Kilbas *et al.* [10], and Diethelm [11]. The computational aspects are surveyed by Li & Zeng [12] and by Garrappa [13]. The broad reach of the subject across particles, fields and control is described by Tarasov [14]. Its recent history is reviewed by Machado *et al.* [15], and its role in systems and control is treated by Chen *et al.* [16]. This memory property makes fractional models a natural language for social and physiological phenomena in which behaviour is shaped by accumulated experience.

A decisive refinement came with the operator of Atangana & Baleanu [17], whose kernel is the non-singular Mittag-Leffler function. By replacing the singular power-law kernel of earlier definitions, this construction removes the artificial singularity at the origin and yields a more faithful description of fading memory, as emphasised by Abdeljawad [18]. Efficient quadratures and integral representations for the operator were developed by Djida *et al.* [19]. The careful comparisons among competing definitions were carried out by Algahtani [20], by Martinez-Fuentes & Delfin-Prieto [21], who also addressed observer design, and by Sabatier [22], who scrutinised the modelling assumptions behind continuous kernels. The Mittag-Leffler kernel has since become a workhorse for biological and physical systems with memory, with stability theory and applications [23]. We also note the parallel development around the Caputo-Fabrizio operator, which was analysed for nonlinear systems by Kebede & Lakoud [24] and applied to disease spread and to breast-cancer dynamics by Kumar *et al.* [25] and by Chavada & Pathak [26]. These works clarify the relative merits of the two non-singular kernels and motivate our choice of the Mittag-Leffler kernel for its slower, more flexible memory decay.

A further generalisation, the fractal-fractional operator introduced by Atangana [27], fuses fractional differentiation with fractal geometry by combining a memory order with a fractal order, so that nonlocal memory and irregular temporal scaling are represented simultaneously. The mathematical

theory of such operators under the Mittag-Leffler kernel was developed by Rahman [28]. Their modelling power was demonstrated by Chithra *et al.* [29] for social-media addiction and by Ahmed *et al.* [30] for epidemic spread. This double flexibility is well suited to addiction, where the build-up of dependence is neither memoryless nor uniformly paced in time.

On the computational side, physics-informed neural networks have emerged as a flexible, mesh-free alternative for differential models. The framework of Raissi *et al.* [31] and the broader programme of physics-informed machine learning summarised by Karniadakis *et al.* [32] embed the governing equations directly in the training loss, producing continuous surrogates that are attractive for inverse problems and real-time prediction. Because classical formulations target local operators, the nonlocal fractional case requires a dedicated treatment, supplied by the fractional physics-informed networks of Pang *et al.* [33]. We use exactly this idea to obtain an independent, structurally different check on our deterministic solution.

Despite this active and growing literature, alcohol addiction models built on the Atangana-Baleanu fractal-fractional operator are still scarce, and the joint influence of memory order and fractal order on the route to persistence or extinction of drinking has not been systematically analysed. The few existing alcohol models that use fractional ideas [5–8] stop short of the fractal-fractional setting, do not develop a global stability theory tailored to the Mittag-Leffler kernel, and are not cross-validated by an independent learning-based solver. Closing this gap calls for a single framework that formulates a behaviourally meaningful four compartment model, proves its well-posedness and stability rigorously, simulates it accurately, and verifies the simulation by a method that shares no discretisation philosophy with it.

The present paper provides such a framework, and our contribution is fourfold. First, we formulate a four-compartment alcohol-addiction model in the Atangana-Baleanu-Caputo fractal-fractional sense, with susceptible, occasional-user, regular user and recovered classes. Second, using the Volterra representation of the operator and the Banach fixed point theorem [34], we prove existence and uniqueness of solutions and establish positivity and boundedness, so that the model is biologically admissible. Third, we derive the equilibria and the basic reproduction number through the next-generation matrix method [3, 4] and obtain a local and global stability theory, the local analysis resting on the fractional stability criterion of Matignon [35] and the global analysis on Lyapunov functionals together with the invariance principle of LaSalle [36] in the spirit of the nonlinear-systems theory of Khalil [37] and the fractional dynamic theory of Lakshmikantham *et al.* [38]. Fourth, we realise the model numerically through the Atangana-Toufik product-integration scheme [39] and cross-validate it with a physics-informed neural network trained by the Levenberg-Marquardt algorithm of Hagan & Menhaj [40] and Marquardt [41], before closing with a quantitative comparison against existing alcohol models.

A graphical overview of the modelling framework, from the compartmental structure through the qualitative theory to the neural-network surrogate, is given in Figure 1.

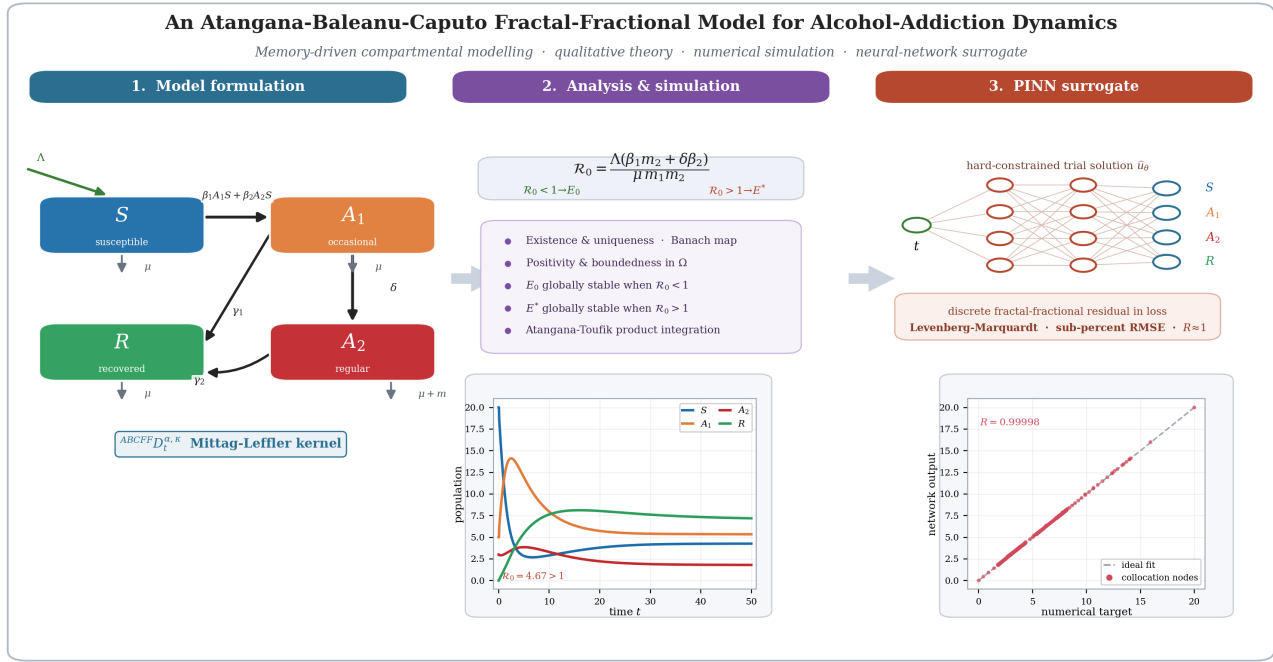


Fig. 1. Graphical overview of the modelling framework

2. Methodology

2.1 Preliminaries

We collect the definitions and auxiliary results used throughout the paper. Unless stated otherwise, $\alpha \in (0, 1)$ denotes the fractional (memory) order and $\kappa \in (0, 1]$ the fractal order, $\Gamma(\cdot)$ is the Euler Gamma function, and $B(\alpha)$ is the Atangana-Baleanu normalisation satisfying $B(0) = B(1) = 1$. We work on the Banach space:

$$X = C([0, T], \mathbb{R}), \quad \|x\|_{\infty} = \sup_{t \in [0, T]} |x(t)| \quad (1)$$

of continuous real valued functions on $[0, T]$ equipped with the supremum norm.

Definition 2.1. [9] For $\alpha > 0$ and $z \in \mathbb{C}$, the one-parameter Mittag-Leffler function is:

$$E_{\alpha}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)} \quad (2)$$

It generalises the exponential function, to which it reduces when $\alpha = 1$, and it is the kernel that governs non-singular fractional operators.

Definition 2.2. [17] For $x \in X$ and $\alpha \in (0, 1)$, the Atangana-Baleanu fractional integral is:

$${}^{AB}I_t^{\alpha} x(t) = \frac{1 - \alpha}{B(\alpha)} x(t) + \frac{\alpha}{B(\alpha) \Gamma(\alpha)} \int_0^t (t - \tau)^{\alpha-1} x(\tau) d\tau \quad (3)$$

Definition 2.3. [17] For $x \in H^1(0, T)$ and $\alpha \in (0, 1)$, the Atangana-Baleanu derivative in the Caputo sense is:

$${}^{ABC}D_t^{\alpha} x(t) = \frac{B(\alpha)}{1 - \alpha} \int_0^t x'(\tau) E_{\alpha} \left(-\frac{\alpha}{1 - \alpha} (t - \tau)^{\alpha} \right) d\tau \quad (4)$$

The Mittag-Leffler kernel is nonlocal and free of singularities, which distinguishes this operator from the classical Caputo derivative.

Definition 2.4. [27] Let $\alpha \in (0, 1)$ and $\kappa \in (0, 1]$. For a function x differentiable in the fractal sense, the Atangana-Baleanu-Caputo fractal-fractional derivative is:

$${}^{ABCF}D_t^{\alpha, \kappa} x(t) = \frac{B(\alpha)}{1 - \alpha} \int_0^t \frac{d}{d\tau^\kappa} x(\tau) E_\alpha \left(-\frac{\alpha}{1 - \alpha} (t - \tau)^\alpha \right) d\tau \quad (5)$$

where the fractal derivative satisfies $\frac{dx(\tau)}{d\tau^\kappa} = \frac{1}{\kappa \tau^{\kappa-1}} \frac{dx(\tau)}{d\tau}$, so that the order α controls the memory and the order κ controls the fractal temporal scaling.

Remark 2.5. [27] In the limit $\alpha \rightarrow 1$ and $\kappa \rightarrow 1$ the operator of Definition 2.4 reduces to the ordinary first derivative, and the fractal-fractional model formulated below collapses to its classical integer-order counterpart. The two orders therefore interpolate continuously between the memoryless, uniformly paced description and a memory laden, fractal time description.

Lemma 2.6. [11] Let $f : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be continuous and satisfy:

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad (6)$$

for some constant $L > 0$ and all admissible x, y . Then the associated fractal-fractional initial-value problem possesses a unique continuous solution in X , obtained as the limit of the Picard sequence generated by the integral operator of Lemma 2.7.

Lemma 2.7. [17] The fractal-fractional initial value problem:

$${}^{ABCF}D_t^{\alpha, \kappa} x(t) = f(t, x(t)), \quad x(0) = x_0 \quad (7)$$

is equivalent to the Volterra integral equation:

$$x(t) = x_0 + \frac{1 - \alpha}{B(\alpha)} \kappa t^{\kappa-1} f(t, x(t)) + \frac{\alpha \kappa}{B(\alpha) \Gamma(\alpha)} \int_0^t (t - \tau)^{\alpha-1} \tau^{\kappa-1} f(\tau, x(\tau)) d\tau \quad (8)$$

This representation is the starting point both for the existence theory and for the numerical scheme used later.

Theorem 2.8. [34] Let (X, d) be a complete metric space and let $\mathcal{T} : X \rightarrow X$ satisfy $d(\mathcal{T}x, \mathcal{T}y) \leq k d(x, y)$ for some $k \in (0, 1)$ and all $x, y \in X$. Then $\mathcal{T}$ has a unique fixed point $x^* \in X$, and the iterates $x_{n+1} = \mathcal{T}(x_n)$ converge to x^* from any starting point $x_0 \in X$.

Theorem 2.9. [35] For a fractional order system of order $\alpha \in (0, 1)$, an equilibrium is asymptotically stable provided every eigenvalue λ_i of the Jacobian evaluated there satisfies $|\arg(\lambda_i)| > \frac{\alpha\pi}{2}$. Since this region contains the open left half-plane, eigenvalues with negative real part always meet the criterion.

The global stability analysis rests on two further tools. The first transfers the classical Lyapunov calculus to the present operator, and the second is the elementary inequality that makes the resulting expressions sign-definite.

Lemma 2.10. [Baleanu2021, 38] Let $x(t) > 0$ be a continuously differentiable function with a positive constant reference value x^* , and let $h(u) = u - 1 - \ln u$. Then, for every $\alpha \in (0, 1)$ and $\kappa \in (0, 1]$:

$${}^{ABCF}D_t^{\alpha, \kappa} \left[x^* h\left(\frac{x(t)}{x^*}\right) \right] \leq \left(1 - \frac{x^*}{x(t)}\right) {}^{ABCF}D_t^{\alpha, \kappa} x(t) \quad (9)$$

Consequently, for a Lyapunov functional $W = \sum_i c_i x_i^* h(x_i/x_i^*)$ with constants $c_i > 0$, one has:

$${}^{ABCF}D_t^{\alpha,\kappa} W \leq \sum_i c_i \left(1 - \frac{x_i^*}{x_i}\right) {}^{ABCF}D_t^{\alpha,\kappa} x_i = \sum_i c_i \left(1 - \frac{x_i^*}{x_i}\right) \mathcal{G}_i \quad (10)$$

where $\mathcal{G}_i$ is the corresponding component of the vector field. The last equality lets the sign of the fractal-fractional derivative of W be decided by the algebraic expression on the right, which involves only the reaction terms of the model.

Lemma 2.11. For positive numbers $y_1, \dots, y_n$ whose product satisfies $y_1 y_2 \cdots y_n = 1$, the arithmetic-geometric-mean inequality gives:

$$y_1 + y_2 + \cdots + y_n \geq n, \quad \text{equivalently} \quad n - (y_1 + \cdots + y_n) \leq 0$$

with equality if and only if $y_1 = \cdots = y_n = 1$. In particular $2 - \frac{a}{b} - \frac{b}{a} \leq 0$ for all $a, b > 0$, with equality only at $a = b$.

Definitions 2.1-2.4, the results of Lemmas 2.6-2.7, 2.10-2.11 and Theorems 2.8-2.9 supply the analytical toolkit used for the formulation, the well-posedness theory and the stability analysis that follow.

2.2 Model Formulation

We now construct the alcohol-addiction model that is studied throughout the paper. At time t the total population $N(t)$ is divided into four mutually exclusive compartments:

$$N(t) = S(t) + A_1(t) + A_2(t) + R(t) \quad (11)$$

whose epidemiological meaning is collected in Table 1. The susceptible class S gathers individuals who do not yet drink in a problematic way. The occasional-user class A_1 contains irregular drinkers with no settled pattern. The regular-user class A_2 contains habitual drinkers who face a markedly higher risk of alcohol-related harm. Also, the recovered class R holds those who have stopped drinking. The transfer mechanism between the four compartments is displayed in Figure 2. Recruitment feeds the susceptible class, social contact with occasional and regular drinkers drives susceptibles into occasional use, occasional users progress to regular use, both user classes recover, and every class is subject to natural mortality while regular users carry an additional alcohol induced mortality (Figure 2).

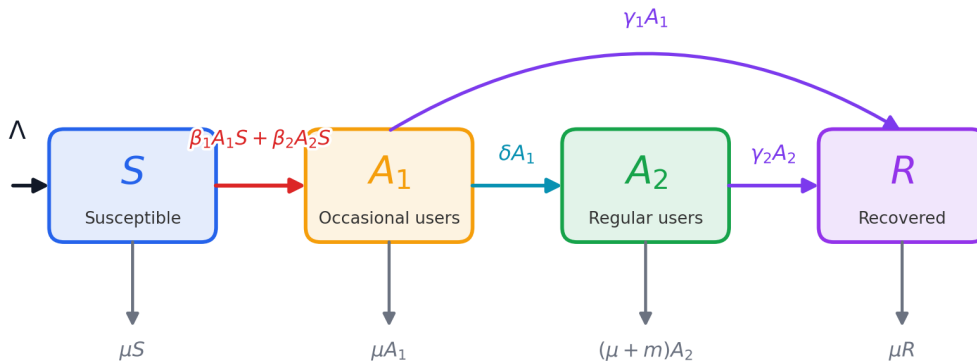


Fig. 2. Compartmental transfer diagram of the alcohol addiction model

Table 1
State variables of the alcohol-addiction model

Variable	Interpretation
$S(t)$	Susceptible individuals who are not problematic drinkers
$A_1(t)$	Occasional or irregular alcohol users
$A_2(t)$	Regular or habitual alcohol users at elevated risk
$R(t)$	Recovered individuals who have ceased drinking

The population is recruited into the susceptible class at a constant rate Λ . Through social contact with the two drinking classes, susceptible individuals begin to drink and move into the occasional-user class at the transmission rates β_1 (contact with occasional users) and β_2 (contact with regular users). Occasional users escalate to regular use at the progression rate δ . Occasional and regular users abandon drinking at the recovery rates γ_1 and γ_2 and enter the recovered class. Every compartment experiences natural mortality at rate μ , and regular users are additionally subject to an alcohol-induced mortality at rate m . The biological meaning and baseline value of each parameter are listed in Table 2. Where two values appear, the first drives the alcohol-free regime and the second the alcohol-present regime. Throughout we abbreviate $m_1 = \mu + \delta + \gamma_1$ and $m_2 = \mu + m + \gamma_2$ for the total outflow rates from the occasional-user and regular-user classes.

Table 2
Parameters of the alcohol-addiction model, with the baseline values used in the simulations

Parameter	Interpretation	Baseline
Λ	Recruitment rate into the susceptible class	2
β_1	Transmission rate through contact with occasional users	0.005 / 0.05
β_2	Transmission rate through contact with regular users	0.005 / 0.06
δ	Progression rate from occasional to regular use	0.1
γ_1	Recovery rate of occasional users	0.1
γ_2	Recovery rate of regular users	0.1
μ	Natural mortality rate	0.1
m	Alcohol-induced mortality rate of regular users	0.1

To incorporate the long behavioural memory of drinking together with irregular temporal scaling, we cast the dynamics in the Atangana-Baleanu-Caputo fractal-fractional sense of Definition 2.4. The resulting model is

$$\begin{aligned}
 {}^{ABCF}D_t^{\alpha,\kappa} S(t) &= \Lambda - \beta_1 A_1(t)S(t) - \beta_2 A_2(t)S(t) - \mu S(t), \\
 {}^{ABCF}D_t^{\alpha,\kappa} A_1(t) &= \beta_1 A_1(t)S(t) + \beta_2 A_2(t)S(t) - m_1 A_1(t), \\
 {}^{ABCF}D_t^{\alpha,\kappa} A_2(t) &= \delta A_1(t) - m_2 A_2(t), \\
 {}^{ABCF}D_t^{\alpha,\kappa} R(t) &= \gamma_1 A_1(t) + \gamma_2 A_2(t) - \mu R(t).
 \end{aligned} \tag{12}$$

The fractional order $\alpha \in (0, 1)$ tunes the depth of memory carried by the Mittag-Leffler kernel, while the fractal order $\kappa \in (0, 1]$ tunes the temporal scaling; by Remark 2.5, the choice $\alpha = \kappa = 1$ returns

the classical integer-order model. System (12) is augmented with the nonnegative initial conditions

$$S(0) = S_0 > 0, \quad A_1(0) = A_{10} \geq 0, \quad A_2(0) = A_{20} \geq 0, \quad R(0) = R_0 \geq 0 \quad (13)$$

and is studied on the biologically feasible region $\Omega = \left\{ (S, A_1, A_2, R) \in \mathbb{R}_+^4 : 0 < N(t) \leq \frac{\Lambda}{\mu} \right\}$, whose invariance is established in the next section.

2.3 Numerical Scheme

To realise the model numerically we use the Atangana-Toufik product-integration scheme, a two-step Lagrange method tailored to the non-singular Mittag-Leffler kernel. Starting from the Volterra form of Lemma 2.7, we write the fractal-modified vector field as $g(t, \Phi) = \kappa t^{\kappa-1} F(t, \Phi)$, where F is the right-hand side of Eq. (12), so that the representation reads:

$$\Phi(t) = \Phi_0 + \frac{1-\alpha}{B(\alpha)} g(t, \Phi(t)) + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} g(\tau, \Phi(\tau)) d\tau \quad (14)$$

We discretise on the uniform grid $t_n = nh$ with $h = T/N$, and at $t = t_{n+1}$ split the memory integral over the subintervals:

$$\int_0^{t_{n+1}} (t_{n+1} - \tau)^{\alpha-1} g(\tau, \Phi(\tau)) d\tau = \sum_{j=0}^n \int_{t_j}^{t_{j+1}} (t_{n+1} - \tau)^{\alpha-1} g(\tau, \Phi(\tau)) d\tau \quad (15)$$

On each subinterval g is replaced by the two-point Lagrange interpolant through the nodes t_{j-1} and t_j :

$$g(\tau, \Phi(\tau)) \approx \frac{\tau - t_{j-1}}{h} g_j - \frac{\tau - t_j}{h} g_{j-1}, \quad g_j = g(t_j, \Phi_j) \quad (16)$$

and the two elementary integrals $\int_{t_j}^{t_{j+1}} (t_{n+1} - \tau)^{\alpha-1} (\tau - t_{j-1}) d\tau$ and $\int_{t_j}^{t_{j+1}} (t_{n+1} - \tau)^{\alpha-1} (\tau - t_j) d\tau$ are evaluated in closed form by the substitution $\tau = t_j + sh$. Collecting the results and dividing the non-local part of the operator accordingly yields, with $g_j = g(t_j, \Phi_j)$, the update

$$\Phi_{n+1} = \Phi_0 + \frac{1-\alpha}{B(\alpha)} g_n + \frac{\alpha h^\alpha}{B(\alpha)\Gamma(\alpha+2)} \sum_{j=0}^n \left(A_{j,n} g_j - B_{j,n} g_{j-1} \right) \quad (17)$$

with Lagrange weights:

$$A_{j,n} = (n+1-j)^\alpha (n-j+2+\alpha) - (n-j)^\alpha (n-j+2+2\alpha) \quad (18)$$

$$B_{j,n} = (n+1-j)^{\alpha+1} - (n-j)^\alpha (n-j+1+\alpha) \quad (19)$$

and the normalisation $B(\alpha) = 1 - \alpha + \alpha/\Gamma(\alpha)$. The factor $\kappa t^{\kappa-1}$ is regularised to zero at $t = 0$, consistent with the vanishing fractal measure of the initial instant. As $\alpha \rightarrow 1$ and $\kappa \rightarrow 1$ the iteration (17) reproduces the integer order trajectory, which we verified against a high-accuracy Runge-Kutta reference, the largest discrepancy in the terminal state remaining below 5×10^{-3} across both scenarios.

2.4 Physics-Informed Neural Network

The deterministic scheme of the previous section establishes the dynamics through a dedicated marching rule. We now complement it with a mesh-free, learning-based approximation built on a

physics-informed neural network, which serves two purposes. It yields a continuous surrogate of the state trajectories that can be queried at any instant without re-running an integrator, and it provides an independent check on the numerical solution, since agreement between two methods that share no discretisation philosophy is strong evidence that the computed trajectories are trustworthy. Classical networks of this kind were devised for local operators whose residual is assembled pointwise by automatic differentiation; because the Atangana-Baleanu operator is nonlocal, its residual is expressed instead through the Volterra representation and a quadrature consistent with the memory kernel, following the fractional formulation of Pang *et al.* [33].

Let $u(t) = (S(t), A_1(t), A_2(t), R(t))^T$ be the state vector, approximated by a densely connected feed-forward network $\mathcal{N}_\theta : \mathbb{R} \rightarrow \mathbb{R}^4$ with trainable parameters θ , two hidden layers of twenty-four neurons each and hyperbolic tangent activations, the scalar time input being normalised through $t \mapsto t/T$. To enforce the initial conditions exactly rather than approximately, the raw output is embedded in the trial solution

$$\hat{u}_\theta(t) = u_0 + (1 - e^{-t/\tau}) [\mathbf{s} \odot \mathcal{N}_\theta(t/T)], \quad u_0 = (S_0, A_{10}, A_{20}, R_0)^T \quad (20)$$

where $\odot$ is the Hadamard product, $\mathbf{s} \in \mathbb{R}^4$ is a fixed output scaling vector taken from the magnitude of the reference data and $\tau > 0$ is a ramp constant. The gate $1 - e^{-t/\tau}$ vanishes at $t = 0$, so $\hat{u}_\theta(0) = u_0$ holds identically for every θ , which removes the initial condition from the optimisation and stabilises training.

The training objective combines a data-fidelity term with a physics-residual term. For the data term the high resolution Atangana-Toufik solution u_{num} of system (12) is sampled at the collocation nodes $t_n = nh$, $h = T/M$. For the physics term the discrete fractal-fractional relation is enforced through the same two step Lagrange rule used in Eq. (17). Writing the fractal-modified field as:

$$g(t, \hat{u}_\theta) = \kappa t^{\kappa-1} F(t, \hat{u}_\theta) \quad (21)$$

with F the right hand side of Eq. (12), the predicted update at node $n + 1$ is:

$$\mathcal{A}[\hat{u}_\theta]_{n+1} = u_0 + \frac{1 - \alpha}{B(\alpha)} g_n + \frac{\alpha h^\alpha}{B(\alpha) \Gamma(\alpha + 2)} \sum_{j=0}^n (A_{j,n} g_j - B_{j,n} g_{j-1}) \quad (22)$$

where $g_j = g(t_j, \hat{u}_\theta(t_j))$ and the Lagrange weights are

$$\begin{aligned} A_{j,n} &= (n + 1 - j)^\alpha (n - j + 2 + \alpha) - (n - j)^\alpha (n - j + 2 + 2\alpha) \\ B_{j,n} &= (n + 1 - j)^{\alpha+1} - (n - j)^\alpha (n - j + 1 + \alpha) \end{aligned} \quad (23)$$

The physics residual at node n is $\mathcal{R}_n^{\text{phys}} = \hat{u}_\theta(t_{n+1}) - \mathcal{A}[\hat{u}_\theta]_{n+1}$, which vanishes exactly when the network satisfies the discretised governing system. The composite loss is:

$$\mathcal{L}(\theta) = \frac{w_d}{|\mathcal{T}|} \sum_{t_i \in \mathcal{T}} \|\hat{u}_\theta(t_i) - u_{\text{num}}(t_i)\|^2 + \frac{w_p}{M} \sum_{n=1}^M \|\mathcal{R}_n^{\text{phys}}\|^2 \quad (24)$$

where $\mathcal{T}$ is the set of training nodes and the weights $w_d, w_p > 0$ balance fidelity against physics.

Since the loss (24) is a sum of squared residuals, it is minimised by the Levenberg-Marquardt algorithm, the standard choice for least-squares network training. Stack the weighted data and physics residuals into a single vector $r(\theta) \in \mathbb{R}^P$, so that $\mathcal{L}(\theta) = \|r(\theta)\|^2$, and let $J = \partial r / \partial \theta \in \mathbb{R}^{P \times Q}$ be its Jacobian, obtained by reverse-mode differentiation through both the network and the discrete operator (22). The gradient and the Gauss-Newton approximation of the Hessian are:

$$\nabla \mathcal{L}(\theta) = 2J^\top r, \quad \nabla^2 \mathcal{L}(\theta) = 2(J^\top J + \sum_k r_k \nabla^2 r_k) \approx 2J^\top J \quad (25)$$

the second-derivative term being dropped because the residuals are small near a good fit. A pure Gauss-Newton step $J^\top J \delta\theta = -J^\top r$ can be poorly conditioned, so Levenberg-Marquardt adds an adaptive multiple of $\text{diag}(J^\top J)$ to the curvature matrix, and each iteration solves the damped normal equations:

$$(J^\top J + \mu \text{diag}(J^\top J)) \delta\theta = -J^\top r, \quad \theta \leftarrow \theta + \delta\theta, \quad (26)$$

with μ the damping parameter. Small μ recovers the fast Gauss-Newton step near the minimum, while large μ tends to a short gradient-descent step in the ill-conditioned regime; a step is accepted only when it lowers $\|r\|^2$, otherwise μ is raised and the step recomputed. The collocation nodes are split into training, validation and test subsets in the ratio 70:15:15, training is monitored on the validation subset, and an early-stopping rule halts the run once the validation error fails to improve for a fixed number of epochs, so that all reported test errors are computed on nodes never used to drive the optimisation. The configuration is listed in Table 3.

Table 3
Architecture, discretisation and training hyperparameters of the physics-informed neural network

Setting	Value
Fractional order α	0.95
Fractal order κ	0.95
Time horizon T	50
Collocation nodes $M + 1$	121
Hidden layers (activation)	2×24 (tanh)
Trainable parameters	748
Train / validation / test nodes	84 / 18 / 19
Data weight w_d / physics weight w_p	1.0 / 0.04
Optimiser	Levenberg-Marquardt
Maximum epochs / early-stop patience	120 / 8
Initial-condition enforcement	hard (trial solution)

3. Results and Discussion

3.1 Positivity, Boundedness and Well-Posedness

Because the state variables of system (12) represent sub-populations, a meaningful model must keep every trajectory nonnegative and bounded, and must admit one and only one solution for each admissible initial state. We establish these three properties in turn. The vector field on the right of (12) is a polynomial in the state variables, hence continuous and continuously differentiable, and therefore locally Lipschitz on $\mathbb{R}_+^4$.

Theorem 3.1. *For every nonnegative initial condition (13), the solution of system (12) remains nonnegative for all $t \geq 0$, that is, $S(t) \geq 0$, $A_1(t) \geq 0$, $A_2(t) \geq 0$ and $R(t) \geq 0$.*

Proof. We evaluate the vector field on each face of the boundary of the nonnegative orthant $\mathbb{R}_+^4$ and show that it never points outward. On the face $S = 0$ the first equation of (12) reduces to

${}^{ABCF}D_t^{\alpha,\kappa}S|_{S=0} = \Lambda \geq 0$, so S cannot become negative. On the face $A_1 = 0$, ${}^{ABCF}D_t^{\alpha,\kappa}A_1|_{A_1=0} = \beta_2 A_2 S \geq 0$, since A_2 and S are nonnegative there. On the face $A_2 = 0$, ${}^{ABCF}D_t^{\alpha,\kappa}A_2|_{A_2=0} = \delta A_1 \geq 0$, and on the face $R = 0$, ${}^{ABCF}D_t^{\alpha,\kappa}R|_{R=0} = \gamma_1 A_1 + \gamma_2 A_2 \geq 0$. On every boundary face the fractal-fractional derivative of the vanishing variable is nonnegative, so the vector field is directed into $\mathbb{R}_+^4$ or is tangent to its boundary. Consequently a trajectory that starts in $\mathbb{R}_+^4$ cannot leave it, and all four components stay nonnegative for $t \geq 0$. $\square$

Theorem 3.2. *The total population $N(t)$ satisfies $\limsup_{t \rightarrow \infty} N(t) \leq \Lambda/\mu$, and the region Ω is positively invariant for system (12). In particular every solution is globally bounded and exists for all $t \geq 0$.*

Proof. Adding the four equations of (12) and using the linearity of the operator gives ${}^{ABCF}D_t^{\alpha,\kappa}N(t) = \Lambda - \mu(S + A_1 + A_2 + R) - mA_2 = \Lambda - \mu N(t) - mA_2(t)$. Since $A_2(t) \geq 0$ by Theorem 3.1, the loss term mA_2 is nonnegative and ${}^{ABCF}D_t^{\alpha,\kappa}N(t) \leq \Lambda - \mu N(t)$. Comparison with the linear equation ${}^{ABCF}D_t^{\alpha,\kappa}Y(t) = \Lambda - \mu Y(t)$, whose solution decreases monotonically towards Λ/μ whenever $Y(0) > \Lambda/\mu$, yields $\limsup_{t \rightarrow \infty} N(t) \leq \Lambda/\mu$. Hence any trajectory that begins in Ω remains there, so Ω is positively invariant. Nonnegativity together with the uniform bound rules out finite time blow up, and the solution is therefore defined for all $t \geq 0$. $\square$

We now turn to the existence and uniqueness of solutions. The argument rests on the Volterra representation of Lemma 2.7 and the Banach fixed point principle of Theorem 2.8, and makes explicit the Lipschitz constant that controls the contraction.

Theorem 3.3. *Let $\Phi = (S, A_1, A_2, R)^\top$ and write system (12) as ${}^{ABCF}D_t^{\alpha,\kappa}\Phi(t) = \mathcal{G}(t, \Phi(t))$ with the initial condition (13). On the bounded invariant region Ω the vector field $\mathcal{G}$ is Lipschitz in Φ , and if the contraction constant:*

$$\Theta = \mathcal{L} \left(\frac{1-\alpha}{B(\alpha)} \kappa T^{\kappa-1} + \frac{\alpha \kappa T^{\alpha+\kappa-1}}{B(\alpha) \Gamma(\alpha+1)} \right) \quad (27)$$

satisfies $\Theta < 1$, where $\mathcal{L}$ is the Lipschitz constant of $\mathcal{G}$ on Ω , then system (12)-(13) possesses a unique continuous solution on $[0, T]$.

Proof. On Ω every component is bounded by Λ/μ , so set $k_0 = \Lambda/\mu$. For two states $\Phi = (S, A_1, A_2, R)$ and $\tilde{\Phi} = (\tilde{S}, \tilde{A}_1, \tilde{A}_2, \tilde{R})$ in Ω we estimate each component of $\mathcal{G}$. For the susceptible component $|\mathcal{G}_S(t, \Phi) - \mathcal{G}_S(t, \tilde{\Phi})| = |-\beta_1(A_1 S - \tilde{A}_1 \tilde{S}) - \beta_2(A_2 S - \tilde{A}_2 \tilde{S}) - \mu(S - \tilde{S})|$. Writing $A_1 S - \tilde{A}_1 \tilde{S} = A_1(S - \tilde{S}) + \tilde{S}(A_1 - \tilde{A}_1)$ and $A_2 S - \tilde{A}_2 \tilde{S} = A_2(S - \tilde{S}) + \tilde{S}(A_2 - \tilde{A}_2)$, and using the bound $|A_i|, |\tilde{S}| \leq k_0$ on each variable, gives $|\mathcal{G}_S(t, \Phi) - \mathcal{G}_S(t, \tilde{\Phi})| \leq (\beta_1 k_0 + \beta_2 k_0 + \mu)|S - \tilde{S}| + \beta_1 k_0 |A_1 - \tilde{A}_1| + \beta_2 k_0 |A_2 - \tilde{A}_2|$. The occasional-user component differs from this only in sign, since $\mathcal{G}_{A_1} = \beta_1 A_1 S + \beta_2 A_2 S - m_1 A_1$, so the same splitting yields $|\mathcal{G}_{A_1}(t, \Phi) - \mathcal{G}_{A_1}(t, \tilde{\Phi})| \leq (\beta_1 k_0 + \beta_2 k_0)|S - \tilde{S}| + (\beta_1 k_0 + m_1)|A_1 - \tilde{A}_1| + \beta_2 k_0 |A_2 - \tilde{A}_2|$. The last two components are affine, hence trivially Lipschitz:

$$\begin{aligned} |\mathcal{G}_{A_2}(t, \Phi) - \mathcal{G}_{A_2}(t, \tilde{\Phi})| &\leq \delta |A_1 - \tilde{A}_1| + m_2 |A_2 - \tilde{A}_2|, \\ |\mathcal{G}_R(t, \Phi) - \mathcal{G}_R(t, \tilde{\Phi})| &\leq \gamma_1 |A_1 - \tilde{A}_1| + \gamma_2 |A_2 - \tilde{A}_2| + \mu |R - \tilde{R}|. \end{aligned}$$

Adding the four estimates and grouping the coefficients of each component distance, every difference $|\mathcal{G}_i(t, \Phi) - \mathcal{G}_i(t, \tilde{\Phi})|$ is bounded by a linear combination of $|S - \tilde{S}|, \dots, |R - \tilde{R}|$. Taking $\mathcal{L} = \max \{ \beta_1 k_0 + \beta_2 k_0 + \mu, \beta_1 k_0 + m_1 + \delta + \gamma_1, \beta_2 k_0 + m_2 + \gamma_2, \mu \}$, which depends only on the parameters and on $k_0 = \Lambda/\mu$, and using $\|\cdot\|_\infty$ to dominate each component distance by $\|\Phi - \tilde{\Phi}\|_\infty$, we obtain the constant $\mathcal{L} > 0$ for which $\|\mathcal{G}(t, \Phi) - \mathcal{G}(t, \tilde{\Phi})\|_\infty \leq \mathcal{L} \|\Phi - \tilde{\Phi}\|_\infty$. Thus, $\mathcal{G}$ is Lipschitz on

Ω . By Lemma 2.7 the initial value problem is equivalent to the fixed-point equation $\Phi = \mathcal{T}\Phi$, where $(\mathcal{T}\Phi)(t) = \Phi(0) + \frac{1-\alpha}{B(\alpha)} \kappa t^{\kappa-1} \mathcal{G}(t, \Phi(t)) + \frac{\alpha \kappa}{B(\alpha) \Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \tau^{\kappa-1} \mathcal{G}(\tau, \Phi(\tau)) d\tau$. The operator $\mathcal{T}$ maps the Banach space $X = C([0, T], \mathbb{R}^4)$ into itself. For $\Phi, \tilde{\Phi} \in X$, we have $\|\mathcal{T}\Phi - \mathcal{T}\tilde{\Phi}\|_\infty \leq \frac{1-\alpha}{B(\alpha)} \kappa T^{\kappa-1} \mathcal{L} \|\Phi - \tilde{\Phi}\|_\infty + \frac{\alpha \kappa}{B(\alpha) \Gamma(\alpha)} \mathcal{L} \|\Phi - \tilde{\Phi}\|_\infty \sup_{t \in [0, T]} \int_0^t (t-\tau)^{\alpha-1} \tau^{\kappa-1} d\tau$. Bounding $\tau^{\kappa-1} \leq T^{\kappa-1}$ inside the integral and using $\int_0^t (t-\tau)^{\alpha-1} d\tau = t^\alpha / \alpha \leq T^\alpha / \alpha$, together with $\Gamma(\alpha+1) = \alpha \Gamma(\alpha)$, gives $\|\mathcal{T}\Phi - \mathcal{T}\tilde{\Phi}\|_\infty \leq \mathcal{L} \left(\frac{1-\alpha}{B(\alpha)} \kappa T^{\kappa-1} + \frac{\alpha \kappa T^{\alpha+\kappa-1}}{B(\alpha) \Gamma(\alpha+1)} \right) \|\Phi - \tilde{\Phi}\|_\infty = \Theta \|\Phi - \tilde{\Phi}\|_\infty$. When $\Theta < 1$ the map $\mathcal{T}$ is a contraction on X , and Theorem 2.8 provides a unique fixed point, which is the unique continuous solution of system (12)-(13) on $[0, T]$. $\square$

Remark 3.4. The memory integral in the proof has a closed form. The substitution $\tau = t\sigma$ turns it into a Beta integral:

$$\int_0^t (t-\tau)^{\alpha-1} \tau^{\kappa-1} d\tau = t^{\alpha+\kappa-1} \int_0^1 (1-\sigma)^{\alpha-1} \sigma^{\kappa-1} d\sigma = t^{\alpha+\kappa-1} B(\alpha, \kappa) = t^{\alpha+\kappa-1} \frac{\Gamma(\alpha) \Gamma(\kappa)}{\Gamma(\alpha+\kappa)} \quad (28)$$

which is finite for every $\alpha, \kappa \in (0, 1]$ since both exponents exceed -1 . Using this value in place of the coarser bound puts the contraction constant in the equivalent form:

$$\Theta = \mathcal{L} \left(\frac{1-\alpha}{B(\alpha)} \kappa T^{\kappa-1} + \frac{\alpha \kappa \Gamma(\kappa)}{B(\alpha) \Gamma(\alpha+\kappa)} T^{\alpha+\kappa-1} \right) \quad (29)$$

and on the operating range α, κ close to one, where $\alpha + \kappa > 1$, both time factors are bounded on $[0, T]$, so Θ can be driven below one by taking the horizon T small and the solution then continued step by step to any finite time.

3.2 Equilibria and the Basic Reproduction Number

The steady states of system (12) are obtained by setting every fractal-fractional derivative to zero, which gives the algebraic system

$$\begin{aligned} 0 &= \Lambda - \beta_1 A_1 S - \beta_2 A_2 S - \mu S, \\ 0 &= \beta_1 A_1 S + \beta_2 A_2 S - m_1 A_1, \\ 0 &= \delta A_1 - m_2 A_2, \\ 0 &= \gamma_1 A_1 + \gamma_2 A_2 - \mu R \end{aligned} \quad (30)$$

Setting $A_1 = A_2 = R = 0$ in Eq. (30) leaves $\Lambda - \mu S = 0$, so the model always admits the alcohol free equilibrium $E_0 = \left(\frac{\Lambda}{\mu}, 0, 0, 0\right)$. For an alcohol-present equilibrium $E^* = (S^*, A_1^*, A_2^*, R^*)$ with $A_1^* > 0$, the third equation of (30) gives $A_2^* = (\delta/m_2) A_1^*$. Substituting this into the second equation and dividing by A_1^* yields $S^* (\beta_1 + \delta \beta_2 / m_2) = m_1$, hence $S^* = \frac{m_1 m_2}{\beta_1 m_2 + \delta \beta_2}$. The basic reproduction number, derived below, is:

$$\mathcal{R}_0 = \frac{\beta_1 S_0}{m_1} + \frac{\delta \beta_2 S_0}{m_1 m_2} = \frac{\Lambda (\beta_1 m_2 + \delta \beta_2)}{\mu m_1 m_2}, \quad S_0 = \frac{\Lambda}{\mu} \quad (31)$$

so that $S^* = \Lambda / (\mu \mathcal{R}_0)$. The first equation of (30), together with the relation $\beta_1 A_1^* S^* + \beta_2 A_2^* S^* = m_1 A_1^*$, gives $\Lambda - \mu S^* - m_1 A_1^* = 0$, where $A_1^* = \frac{\Lambda}{m_1} \left(1 - \frac{1}{\mathcal{R}_0}\right)$, $A_2^* = \frac{\delta \Lambda}{m_1 m_2} \left(1 - \frac{1}{\mathcal{R}_0}\right)$, and $R^* = \frac{\Lambda}{\mu m_1} \left(1 - \frac{1}{\mathcal{R}_0}\right) \left(\gamma_1 + \frac{\delta \gamma_2}{m_2}\right)$. The components A_1^* , A_2^* and R^* are positive precisely when $\mathcal{R}_0 > 1$, so a unique alcohol present equilibrium exists if and only if $\mathcal{R}_0 > 1$.

To obtain $\mathcal{R}_0$ we apply the next-generation matrix method to the drinking classes A_1 and A_2 , which alone drive transmission. The new-drinker and transfer vectors are:

$$\mathcal{F} = \begin{pmatrix} \beta_1 A_1 S + \beta_2 A_2 S \\ 0 \end{pmatrix}, \quad \mathcal{V} = \begin{pmatrix} m_1 A_1 \\ m_2 A_2 - \delta A_1 \end{pmatrix}.$$

Linearising at E_0 , where $S = S_0 = \Lambda/\mu$, gives the matrices:

$$F = \begin{pmatrix} \beta_1 S_0 & \beta_2 S_0 \\ 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} m_1 & 0 \\ -\delta & m_2 \end{pmatrix}, \quad V^{-1} = \frac{1}{m_1 m_2} \begin{pmatrix} m_2 & 0 \\ \delta & m_1 \end{pmatrix}.$$

To invert V we note that it is lower triangular with determinant $\det V = m_1 m_2$, so its inverse is obtained by transposing the cofactor matrix and dividing by $m_1 m_2$, which produces the expression displayed above. Forming the product row by row, the next-generation matrix is:

$$FV^{-1} = \frac{1}{m_1 m_2} \begin{pmatrix} \beta_1 S_0 & \beta_2 S_0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} m_2 & 0 \\ \delta & m_1 \end{pmatrix} = \frac{1}{m_1 m_2} \begin{pmatrix} \beta_1 S_0 m_2 + \delta \beta_2 S_0 & \beta_2 S_0 m_1 \\ 0 & 0 \end{pmatrix} \quad (32)$$

Its characteristic equation is:

$$\det(FV^{-1} - \lambda I) = \lambda \left(\lambda - \frac{\beta_1 S_0 m_2 + \delta \beta_2 S_0}{m_1 m_2} \right) = 0 \quad (33)$$

so the eigenvalues are $\lambda = 0$ and $\lambda = (\beta_1 S_0 m_2 + \delta \beta_2 S_0)/(m_1 m_2)$. The basic reproduction number is the spectral radius, the larger of the two:

$$\mathcal{R}_0 = \rho(FV^{-1}) = \frac{\beta_1 S_0 m_2 + \delta \beta_2 S_0}{m_1 m_2} = \frac{\beta_1 S_0}{m_1} + \frac{\delta \beta_2 S_0}{m_1 m_2} \quad (34)$$

which, with $S_0 = \Lambda/\mu$, reproduces expression (31). The decomposition $\mathcal{R}_0 = \mathcal{R}_{01} + \mathcal{R}_{02}$ with $\mathcal{R}_{01} = \beta_1 S_0/m_1$ and $\mathcal{R}_{02} = \delta \beta_2 S_0/(m_1 m_2)$ separates the contribution of occasional users from that of regular users, the latter reached only after progression at rate δ ; the factor δ/m_2 in $\mathcal{R}_{02}$ is exactly the fraction of occasional users who advance to the regular class before leaving it.

Theorem 3.5. *The alcohol free equilibrium E_0 exists for all admissible parameters. If $\mathcal{R}_0 < 1$ then E_0 is the only equilibrium in Ω and alcohol use cannot invade the population, whereas if $\mathcal{R}_0 > 1$ a unique alcohol present equilibrium E^* appears in the interior of Ω and drinking persists. Lowering the transmission rates β_1, β_2 or raising the recovery rates γ_1, γ_2 decreases $\mathcal{R}_0$ and so favours elimination.*

The geometry behind Theorem 3.5 is displayed in Figure 3. Panels (a) and (b) render $\mathcal{R}_0$ as a surface over the (β_1, β_2) and (β_1, δ) planes, each cut by the translucent level $\mathcal{R}_0 = 1$; the solid curve where the two meet is the critical set that separates elimination from persistence. Panels (c) to (f) trace the coordinates of the interior equilibrium over the (β_1, β_2) plane and show that A_1^*, A_2^* and R^* stay pinned at zero below the threshold and then rise from it, while S^* falls away from the carrying level Λ/μ , so that the branch of persistent drinking bifurcates forward as $\mathcal{R}_0$ crosses one. Panel (a) shows the basic reproduction number $\mathcal{R}_0$ over the (β_1, β_2) plane and panel (b) over the (β_1, δ) plane, each pierced by the translucent unit level whose intersection with the surface (solid curve) marks the critical set $\mathcal{R}_0 = 1$. Panels (c) to (f) show the coordinates of the alcohol-present equilibrium over the (β_1, β_2) plane, namely the susceptible level S^* (c), the occasional-user level A_1^* (d), the regular-user level A_2^* (e) and the recovered level R^* (f). Each user and recovered level vanishes below the threshold

and grows continuously once $\mathcal{R}_0$ exceeds one, tracing the forward transcritical bifurcation, while the susceptible level drops from the alcohol-free carrying value.

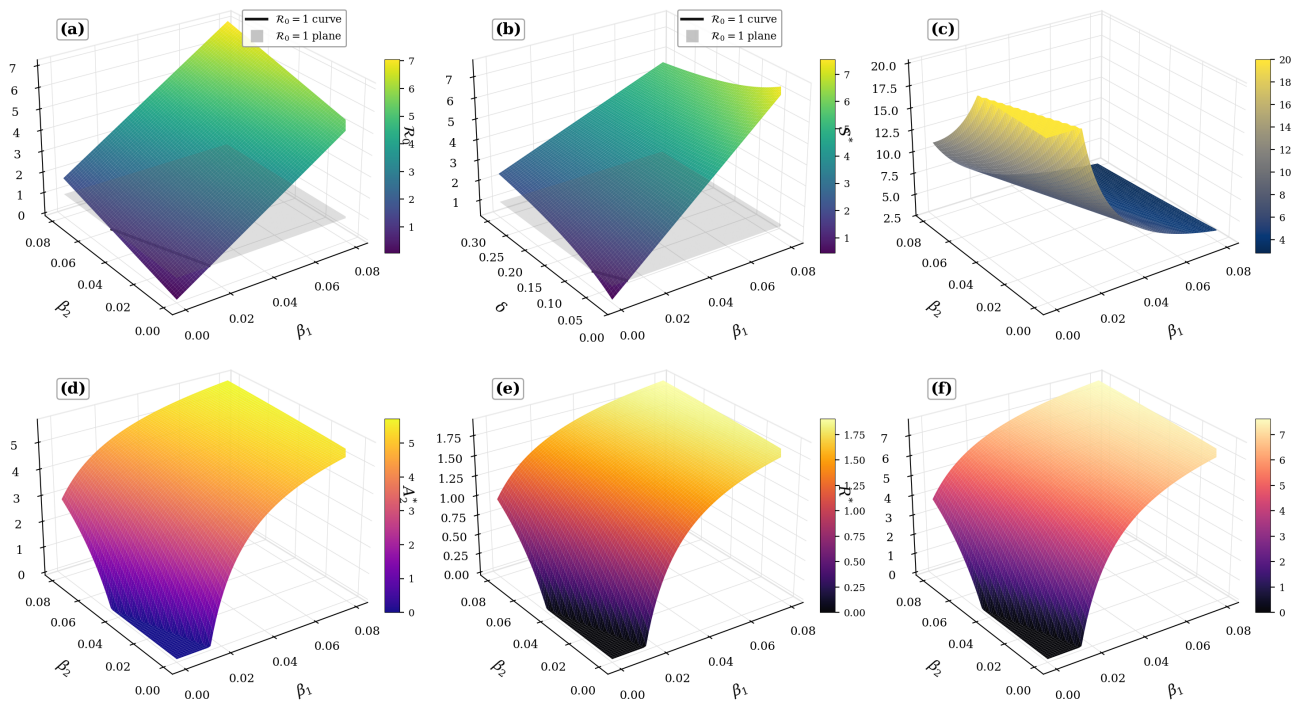


Fig. 3. Threshold and equilibrium geometry of the model

3.3 Stability Analysis

Because the vector field of the fractal-fractional model coincides with that of the underlying integer-order system, the Jacobian is unchanged and only the stability test is read in the fractional sense of Theorem 2.9. As the recovered class does not appear in the first three equations, it suffices to analyse the reduced system in (A_1, A_2, S) , whose Jacobian is:

$$J = \begin{pmatrix} \beta_1 S - m_1 & \beta_2 S & \beta_1 A_1 + \beta_2 A_2 \\ \delta & -m_2 & 0 \\ -\beta_1 S & -\beta_2 S & -\beta_1 A_1 - \beta_2 A_2 - \mu \end{pmatrix} \quad (35)$$

An equilibrium is locally asymptotically stable whenever every eigenvalue λ_i of (35) evaluated there satisfies $|\arg(\lambda_i)| > \alpha\pi/2$, a condition met in particular by any eigenvalue with negative real part. We treat the two equilibria and their global counterparts in succession.

Theorem 3.6. *If $\mathcal{R}_0 < 1$, the alcohol free equilibrium E_0 is locally asymptotically stable; if $\mathcal{R}_0 > 1$ it is unstable.*

Proof. At E_0 one has $A_1 = A_2 = 0$, so the entry $\beta_1 A_1 + \beta_2 A_2$ in the third column of (35) vanishes and the Jacobian takes the block-triangular form:

$$J(E_0) = \begin{pmatrix} \beta_1 S_0 - m_1 & \beta_2 S_0 & 0 \\ \delta & -m_2 & 0 \\ -\beta_1 S_0 & -\beta_2 S_0 & -\mu \end{pmatrix}$$

Because the top-right 2×1 block is zero, the characteristic polynomial factorises as $\det (J(E_0) - \lambda I) = (-\mu - \lambda) \det \begin{pmatrix} \beta_1 S_0 - m_1 - \lambda & \beta_2 S_0 \\ \delta & -m_2 - \lambda \end{pmatrix}$, so one eigenvalue is $\lambda_3 = -\mu < 0$ and the other two are the roots of the quadratic determined by the leading 2×2 block M . Its trace and determinant are

$$\text{tr } M = \beta_1 S_0 - m_1 - m_2, \quad \det M = (\beta_1 S_0 - m_1)(-m_2) - \beta_2 S_0 \delta = m_1 m_2 - (\beta_1 S_0 m_2 + \delta \beta_2 S_0),$$

and since $\mathcal{R}_0 = (\beta_1 S_0 m_2 + \delta \beta_2 S_0) / (m_1 m_2)$ the determinant is $\det M = m_1 m_2 (1 - \mathcal{R}_0)$. The two remaining eigenvalues therefore solve $\lambda^2 - (\text{tr } M)\lambda + \det M = \lambda^2 + (m_1 + m_2 - \beta_1 S_0)\lambda + m_1 m_2 (1 - \mathcal{R}_0) = 0$. When $\mathcal{R}_0 < 1$ the constant term $m_1 m_2 (1 - \mathcal{R}_0)$ is positive, and the relation $\beta_1 S_0 = m_1 \mathcal{R}_{01} \leq m_1 \mathcal{R}_0 < m_1$ makes the linear coefficient $m_1 + m_2 - \beta_1 S_0$ positive as well; by the Routh-Hurwitz test for a quadratic both roots then have negative real part, so all three eigenvalues satisfy the Matignon condition $|\arg \lambda_i| > \alpha\pi/2$ and E_0 is locally asymptotically stable. When $\mathcal{R}_0 > 1$ the constant term is negative, so the quadratic has one positive and one negative real root; the positive root violates the Matignon condition and E_0 is unstable. $\square$

Theorem 3.7. *Let $\mathcal{R}_0 > 1$ and let $q^* = \beta_1 A_1^* + \beta_2 A_2^*$. If the Routh-Hurwitz condition $a_1 a_2 > a_3$ holds for the characteristic coefficients given below, then the alcohol present equilibrium E^* is locally asymptotically stable.*

Proof. Writing $q^* = \beta_1 A_1^* + \beta_2 A_2^*$, the Jacobian (35) evaluated at E^* is:

$$J(E^*) = \begin{pmatrix} \beta_1 S^* - m_1 & \beta_2 S^* & q^* \\ \delta & -m_2 & 0 \\ -\beta_1 S^* & -\beta_2 S^* & -q^* - \mu \end{pmatrix}$$

and its characteristic polynomial is $\lambda^3 + a_1 \lambda^2 + a_2 \lambda + a_3 = 0$ with $a_1 = -\text{tr } J(E^*)$, $a_2 = \sum$ (principal 2×2 minors), and $a_3 = -\det J(E^*)$. Throughout we use the interior-equilibrium identity obtained from the second and third equations of (30):

$$\beta_1 S^* + \frac{\delta \beta_2 S^*}{m_2} = m_1, \quad \text{equivalently} \quad m_1 - \beta_1 S^* = \frac{\delta \beta_2 S^*}{m_2} \quad (36)$$

For the first invariant, $a_1 = -(\beta_1 S^* - m_1 - m_2 - q^* - \mu) = m_2 + \mu + q^* + (m_1 - \beta_1 S^*) = m_2 + \mu + q^* + \frac{\delta \beta_2 S^*}{m_2}$, where the last step used (36). For the constant term we expand $\det J(E^*)$ along the third column; the cofactor multiplying $-q^* - \mu$ is the minor:

$$\det \begin{pmatrix} \beta_1 S^* - m_1 & \beta_2 S^* \\ \delta & -m_2 \end{pmatrix} = m_1 m_2 - (\beta_1 S^* m_2 + \delta \beta_2 S^*) = 0$$

which vanishes by (36), so only the q^* entry contributes and a short computation gives $\det J(E^*) = -m_1 m_2 q^*$, whence $a_3 = m_1 m_2 q^*$. The middle coefficient is the sum of the three principal 2×2 minors; the minor on rows and columns $\{1, 2\}$ vanishes by the same identity, while the other two give $m_1 q^* + \mu \delta \beta_2 S^* / m_2$ and $m_2 q^* + m_2 \mu$. Adding them and clearing the denominator $a_2 = m_1 q^* + m_2 q^* + m_2 \mu + \frac{\mu \delta \beta_2 S^*}{m_2} = \frac{\delta \mu \beta_2 S^* + \delta q^* \beta_2 S^* + m_2^2 \mu + m_2^2 q^* + m_2 q^* \beta_1 S^*}{m_2}$, the second form following from $m_1 m_2 = \beta_1 S^* m_2 + \delta \beta_2 S^*$. For $\mathcal{R}_0 > 1$ the equilibrium components A_1^*, A_2^*, R^* are positive, so $q^* > 0$ and each of a_1, a_2, a_3 is a sum of positive terms; in particular $a_1, a_3 > 0$. The Routh-Hurwitz criterion for a cubic then guarantees that all three roots have negative real part precisely when $a_1 a_2 > a_3$, in which case $|\arg(\lambda_i)| > \alpha\pi/2$ for every eigenvalue and E^* is locally asymptotically stable. $\square$

Local analysis is now strengthened to global statements by means of Lyapunov functionals built from the function $h(x) = x - 1 - \ln x$, which is nonnegative on $(0, \infty)$ and vanishes only at $x = 1$. The invariance principle of LaSalle then upgrades a nonpositive derivative to global attraction.

Theorem 3.8. *If $\mathcal{R}_0 < 1$, the alcohol-free equilibrium E_0 is globally asymptotically stable in Ω .*

Proof. Write $D := {}^{ABCF}D_t^{\alpha, \kappa}$ and consider $W_1 = S_0 h\left(\frac{S}{S_0}\right) + A_1 + LA_2$ and $L = \frac{\beta_2 S_0}{m_2}$, which is positive definite with respect to E_0 because $h \geq 0$ and $A_1, A_2 \geq 0$. Applying Lemma 2.10 to the logarithmic term and using that A_1 and A_2 enter linearly:

$$DW_1 \leq \left(1 - \frac{S_0}{S}\right) DS + DA_1 + L DA_2 = \left(1 - \frac{S_0}{S}\right) \mathcal{G}_S + \mathcal{G}_{A_1} + L \mathcal{G}_{A_2}$$

where $\mathcal{G}_S, \mathcal{G}_{A_1}, \mathcal{G}_{A_2}$ are the corresponding right-hand sides of (12). Inserting $\Lambda = \mu S_0$, the susceptible contribution splits as:

$$\left(1 - \frac{S_0}{S}\right) (\mu S_0 - \mu S - (\beta_1 A_1 + \beta_2 A_2)S) = -\frac{\mu(S - S_0)^2}{S} + (\beta_1 A_1 + \beta_2 A_2)(S_0 - S)$$

where the first piece used $\mu(1 - S_0/S)(S_0 - S) = -\mu(S - S_0)^2/S$. Adding $\mathcal{G}_{A_1} = (\beta_1 A_1 + \beta_2 A_2)S - m_1 A_1$ and $L \mathcal{G}_{A_2} = L(\delta A_1 - m_2 A_2)$, the two terms $\pm(\beta_1 A_1 + \beta_2 A_2)S$ cancel and $DW_1 \leq -\frac{\mu(S - S_0)^2}{S} + A_1[\beta_1 S_0 - m_1 + L\delta] + A_2[\beta_2 S_0 - Lm_2]$. The choice $L = \beta_2 S_0/m_2$ annihilates the A_2 bracket, while $\beta_1 S_0 + \delta\beta_2 S_0/m_2 = m_1 \mathcal{R}_0$ gives $\beta_1 S_0 - m_1 + L\delta = m_1(\mathcal{R}_0 - 1)$, so that, rewriting the quadratic piece in symmetric form $DW_1 \leq \mu S_0 \left(2 - \frac{S_0}{S} - \frac{S}{S_0}\right) + m_1 A_1(\mathcal{R}_0 - 1)$. The first term is nonpositive by Lemma 2.11 applied to $a = S$, $b = S_0$, and the second is nonpositive because $\mathcal{R}_0 < 1$ and $A_1 \geq 0$. Hence $DW_1 \leq 0$, with equality only when $S = S_0$ and $A_1 = A_2 = 0$. The largest invariant set on which $DW_1 = 0$ is therefore the singleton $\{E_0\}$, and by the LaSalle invariance principle every trajectory in Ω converges to E_0 , which is globally asymptotically stable. $\square$

Theorem 3.9. *If $\mathcal{R}_0 > 1$, the alcohol present equilibrium E^* is globally asymptotically stable in the interior of Ω .*

Proof. Again write $D := {}^{ABCF}D_t^{\alpha, \kappa}$ and take the Lyapunov functional $W = S^* h\left(\frac{S}{S^*}\right) + A_1^* h\left(\frac{A_1}{A_1^*}\right) + LA_2^* h\left(\frac{A_2}{A_2^*}\right)$ and $L = \frac{\beta_2 A_2^* S^*}{\delta A_1^*}$, which is positive definite about E^* for $S, A_1, A_2 > 0$, since $h(u) \geq 0$ with equality only at $u = 1$. Applying Lemma 2.10 to each logarithmic term bounds the fractal-fractional derivative by $DW \leq \left(1 - \frac{S^*}{S}\right) \mathcal{G}_S + \left(1 - \frac{A_1^*}{A_1}\right) \mathcal{G}_{A_1} + L \left(1 - \frac{A_2^*}{A_2}\right) \mathcal{G}_{A_2}$, with $\mathcal{G}_S, \mathcal{G}_{A_1}, \mathcal{G}_{A_2}$ the right-hand sides of Eq. (12). We now eliminate the parameters in favour of the equilibrium coordinates through the relations read off from Eq. (30):

$$\Lambda = \beta_1 A_1^* S^* + \beta_2 A_2^* S^* + \mu S^*, \quad m_1 A_1^* = \beta_1 A_1^* S^* + \beta_2 A_2^* S^*, \quad m_2 A_2^* = \delta A_1^*,$$

the last of which turns the weight into $L = \beta_2 A_2^* S^*/(m_2 A_2^*) = \beta_2 S^*/m_2$. Substituting these into the three products and expanding, the terms linear in S, A_1, A_2 that carry the raw parameters cancel in pairs, and the surviving expression regroups exactly into $DW \leq \left(2 - \frac{S}{S^*} - \frac{S^*}{S}\right) (\mu S^* + \beta_1 A_1^* S^*) + \beta_2 A_2^* S^* \left(3 - \frac{S^*}{S} - \frac{A_1^* A_2 S}{A_1 A_2^* S^*} - \frac{A_1 A_2^*}{A_1^* A_2}\right)$. In the first bracket $\frac{S}{S^*} \cdot \frac{S^*}{S} = 1$, and in the second bracket the three fractions multiply to one as well; both are therefore of the form $n - \sum_i y_i$ with $y_i > 0$ and $\prod_i y_i = 1$, so Lemma 2.11 makes each nonpositive. Since the prefactors $\mu S^* + \beta_1 A_1^* S^*$ and $\beta_2 A_2^* S^*$ are positive, $DW \leq 0$, with equality only at $S = S^*, A_1 = A_1^*$ and $A_2 = A_2^*$. The only invariant set on which $DW = 0$ is $\{E^*\}$, so the LaSalle invariance principle shows that E^* attracts every interior trajectory of Ω and is globally asymptotically stable. $\square$

3.4 Numerical Simulations

All simulations use the baseline parameters of Table 2 and the initial state $S(0) = 20$, $A_1(0) = 5$, $A_2(0) = 3$, $R(0) = 0$. The alcohol-free regime takes $\beta_1 = \beta_2 = 0.005$, for which $\mathcal{R}_0 = 0.444 < 1$, while the alcohol-present regime takes $\beta_1 = 0.05$, $\beta_2 = 0.06$, for which $\mathcal{R}_0 = 4.667 > 1$. The occasional user and regular user classes decay to zero while the susceptible class approaches Λ/μ , so the trajectory converges to the alcohol free equilibrium. Figure 4 shows that in the alcohol free regime both drinking classes decay to zero and the population settles at E_0 , in agreement with Theorem 3.6 and Theorem 3.8.

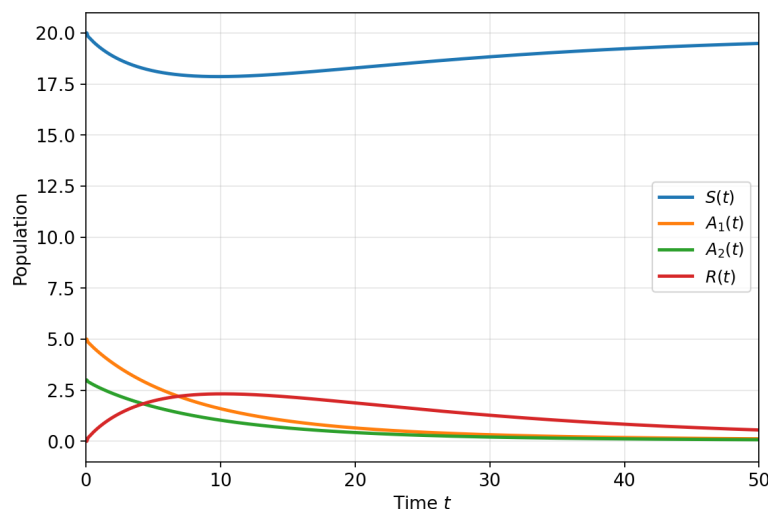


Fig. 4. Alcohol free regime with $\mathcal{R}_0 = 0.444 < 1$

Figure 5 shows the alcohol-present regime, in which the occasional user class overshoots before the trajectory relaxes onto the interior equilibrium $E^* \approx (4.29, 5.24, 1.75, 6.98)$, matching Theorem 3.7 and Theorem 3.9. After an initial transient the four compartments stabilise at the interior alcohol present equilibrium, confirming persistence of drinking above the threshold.

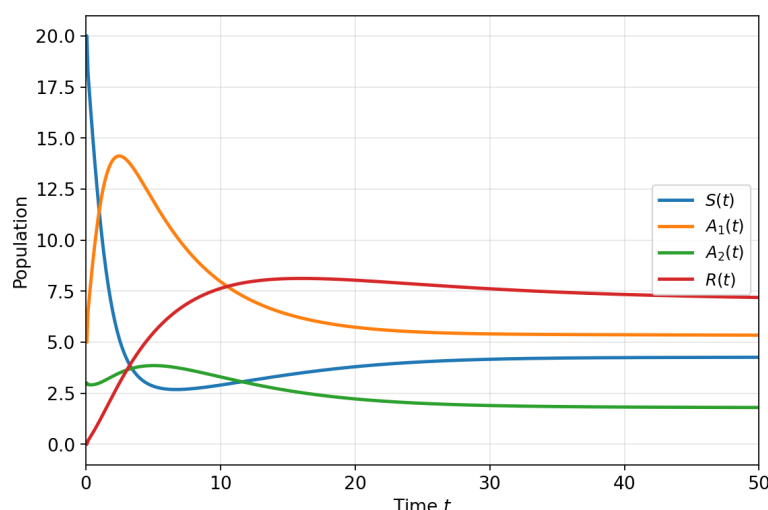


Fig. 5. Alcohol-present regime with $\mathcal{R}_0 = 4.667 > 1$

The memory and fractal time effects are isolated in Figure 6 and Figure 7. Lowering the fractional order α lengthens the memory carried by the kernel and slows the approach of the occasional-user

class to its steady level, as seen in Figure 6. Smaller values of α deepen the memory and delay convergence, illustrating the nonlocal character of the Atangana-Baleanu kernel.

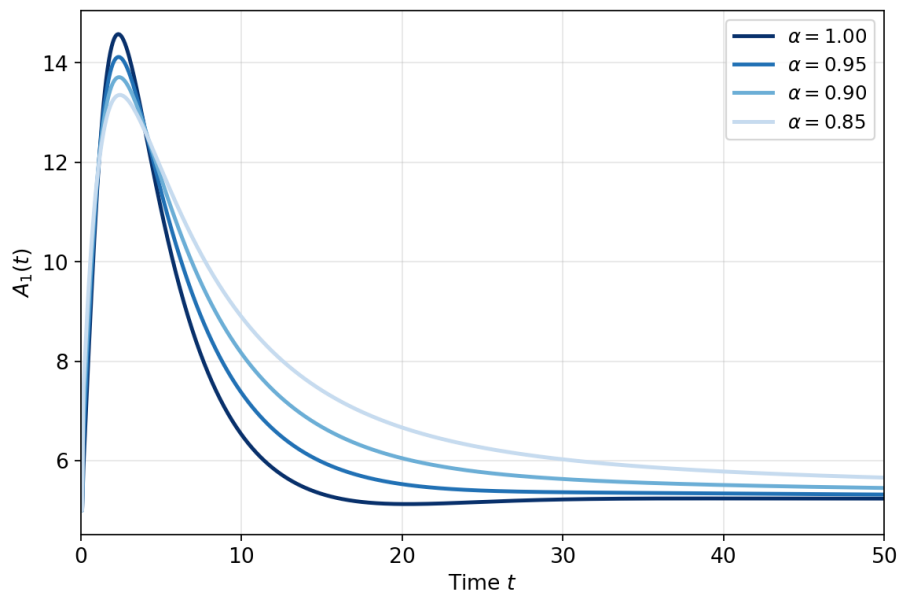


Fig. 6. Effect of the fractional order α on the occasional-user class $A_1(t)$

Varying the fractal order κ reshapes the transient of the regular-user class through the fractal time scale, as seen in Figure 7. Changes in κ alter the transient through the fractal time structure while leaving the eventual equilibrium unchanged.

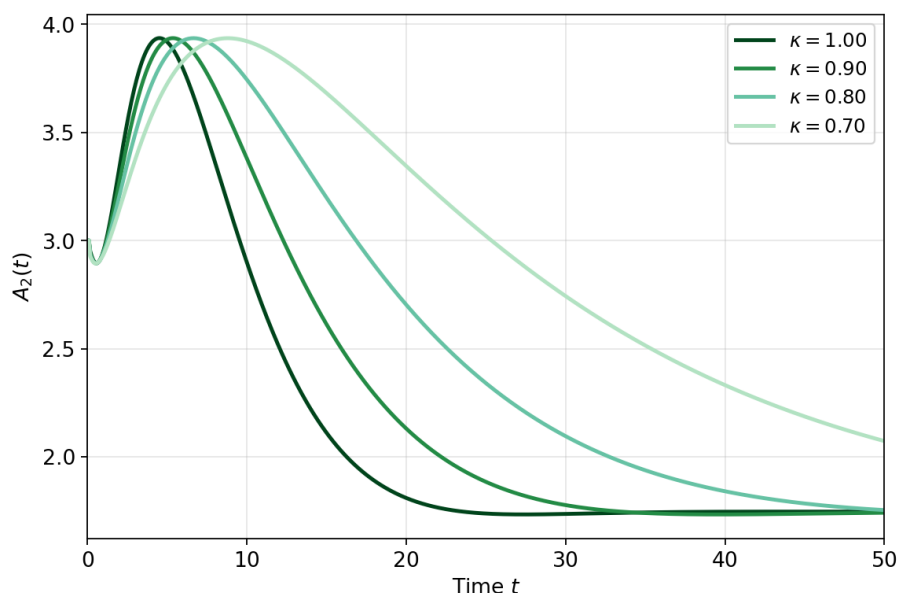


Fig. 7. Effect of the fractal order κ on the regular user class $A_2(t)$

The joint dynamical behaviour is summarised by the three-dimensional phase portrait of Figure 8, in which trajectories from the chosen initial state spiral onto the attracting equilibrium. The trajectory issuing from the initial state is attracted to the equilibrium, visualising the stability established analytically.

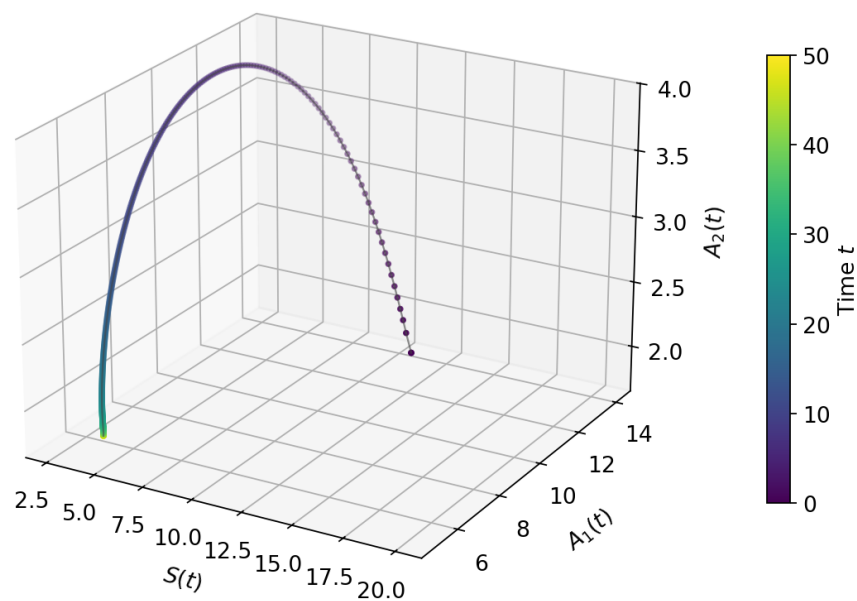


Fig. 8. Three dimensional phase portrait in the (S, A_1, A_2) space

Finally, the combined dependence on time and on the two orders is rendered as surfaces in Figure 9 and Figure 10. Figure 9 traces the occasional-user class over the time α plane.

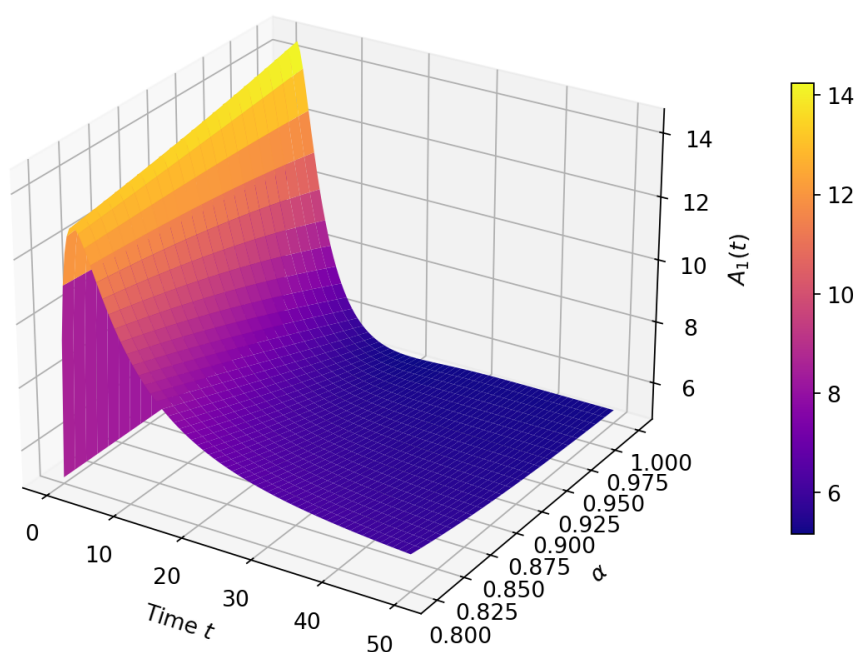


Fig. 9. Surface of the occasional user class $A_1(t)$ over the time α plane, showing the monotone influence of the fractional order on the transient

Figure 10 traces the regular-user class over the time κ plane. Figure 9 and Figure 10 confirm that the

orders modulate the transient amplitude and pace without altering the qualitative outcome dictated by $\mathcal{R}_0$.

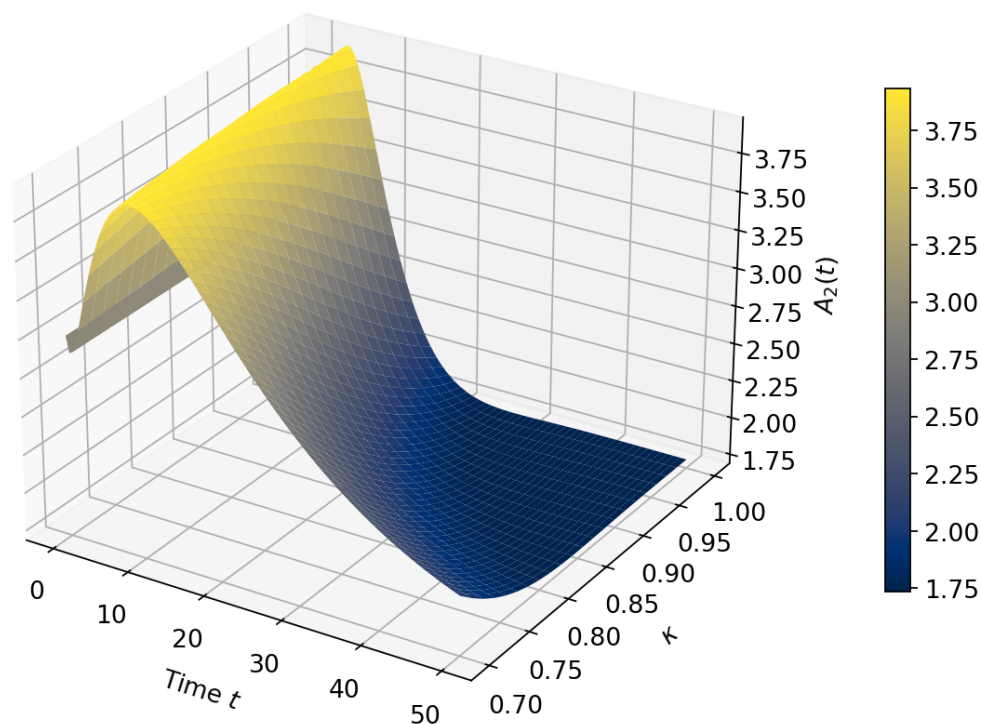


Fig. 10. Surface of the regular-user class $A_2(t)$ over the time κ plane, showing how the fractal order reshapes the transient evolution

A wider view of the same mechanism is assembled in Figure 11, which places the response of every compartment beside the two orders on one plate.

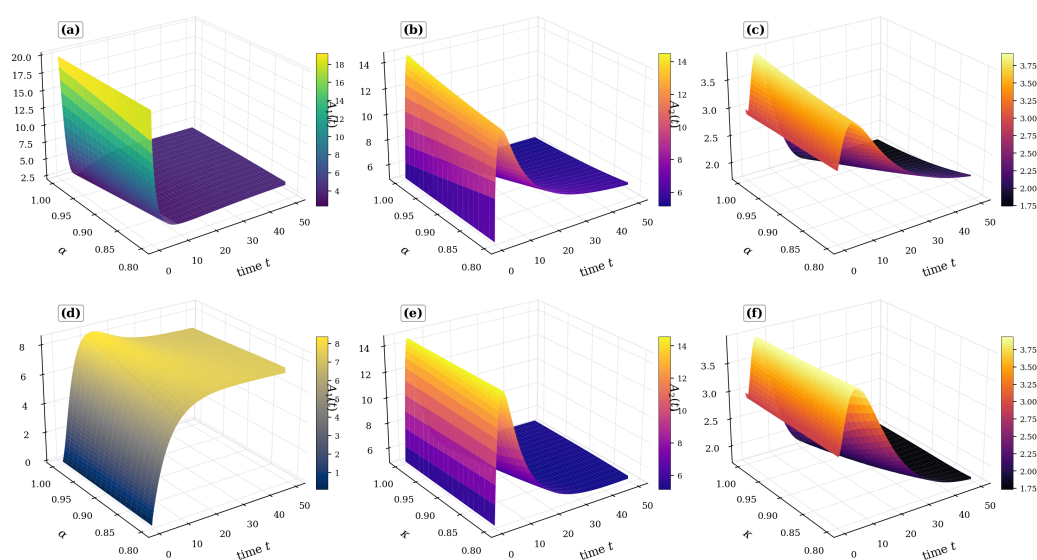


Fig. 11. Response surfaces of the compartmental model over the memory order α and the fractal order κ in the alcohol-present regime

In Figure 11, reading the memory panels (a) to (d), a smaller α deepens the tail of the Mittag-Leffler kernel and drags the susceptible, occasional-user, regular-user and recovered curves more slowly toward their limits; reading the fractal panels (e) and (f), a smaller κ compresses or stretches the early time scale and reshapes the height and timing of the occasional-user overshoot and the regular-user rise. In every panel the ridge that the trajectory settles onto at large t is the same equilibrium fixed by $\mathcal{R}_0$, so the two orders govern the route rather than the destination. Panels (a) to (d) hold the fractal order at unity and sweep the memory order $\alpha \in [0.80, 1]$ for the susceptible class $S(t)$ (a), the occasional-user class $A_1(t)$ (b), the regular-user class $A_2(t)$ (c) and the recovered class $R(t)$ (d). Panels (e) and (f) hold the memory order at unity and sweep the fractal order $\kappa \in [0.80, 1]$ for the occasional-user class $A_1(t)$ (e) and the regular-user class $A_2(t)$ (f). Lowering either order lengthens the transient and delays the settling of each class, whereas the terminal level dictated by $\mathcal{R}_0$ is preserved.

3.5 Neural-Surrogate Results

Two regimes are studied, mirroring the threshold behaviour of $\mathcal{R}_0$. The first uses $\beta_1 = \beta_2 = 0.005$, for which $\mathcal{R}_0 = 0.444$ and the trajectories relax to the alcohol-free state, and the second uses $\beta_1 = 0.05, \beta_2 = 0.06$, for which $\mathcal{R}_0 = 4.667$ and the trajectories develop a pronounced transient before settling at $E^* \approx (4.29, 5.24, 1.75, 6.98)$. The remaining parameters and the initial state match the simulations above. The converged accuracy for both regimes is reported in Table 4. The mean-squared error is reported at the best-validation epoch, while the remaining statistics are computed over all collocation nodes against the numerical target.

Table 4
Converged accuracy of the physics-informed neural network

Metric	Regime 1	Regime 2
Basic reproduction number $\mathcal{R}_0$	0.444	4.667
Best-validation epoch	119	36
Training MSE	2.85×10^{-6}	3.72×10^{-4}
Validation MSE	2.60×10^{-6}	6.64×10^{-4}
Test MSE	1.74×10^{-6}	2.91×10^{-4}
Physics-residual MSE	8.94×10^{-3}	1.92×10^{-1}
Overall RMSE	1.62×10^{-3}	2.01×10^{-2}
Overall MAE	1.14×10^{-3}	1.12×10^{-2}
Maximum (L^∞) error	9.77×10^{-3}	1.06×10^{-1}
Regression coefficient R	1.000000	0.999978
Regression slope	1.00000	0.99885

The training is documented through a standard diagnostic suite. The training-state histories in Figure 12 and Figure 13 show the gradient norm falling by several orders of magnitude while the damping parameter adapts and the validation check counter rises only near the end of the run.

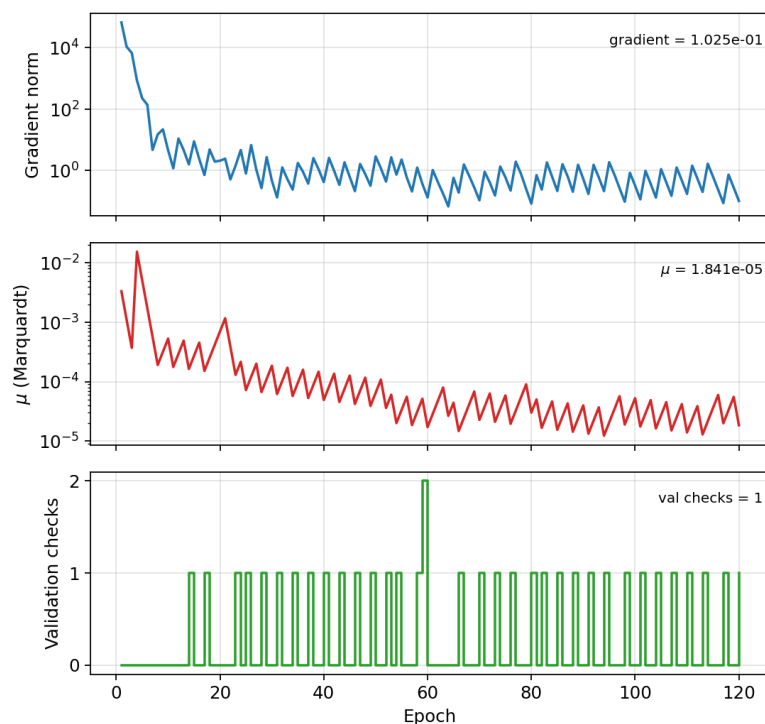


Fig. 12. Training state history for the alcohol free regime, showing the gradient norm, the Levenberg-Marquardt damping parameter and the validation-check counter against epoch

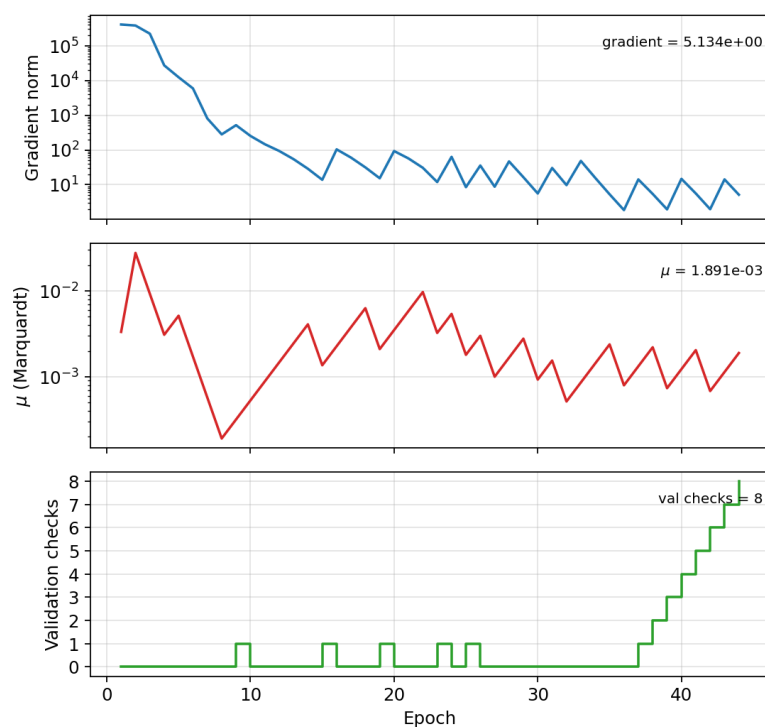


Fig. 13. Training-state history for the alcohol present regime, with early stopping triggered once the validation error stagnates

The performance curves in Figure 14 and Figure 15 confirm that the training, validation and test errors decay together with no sign of overfitting, the best-validation epoch being marked on each

curve.

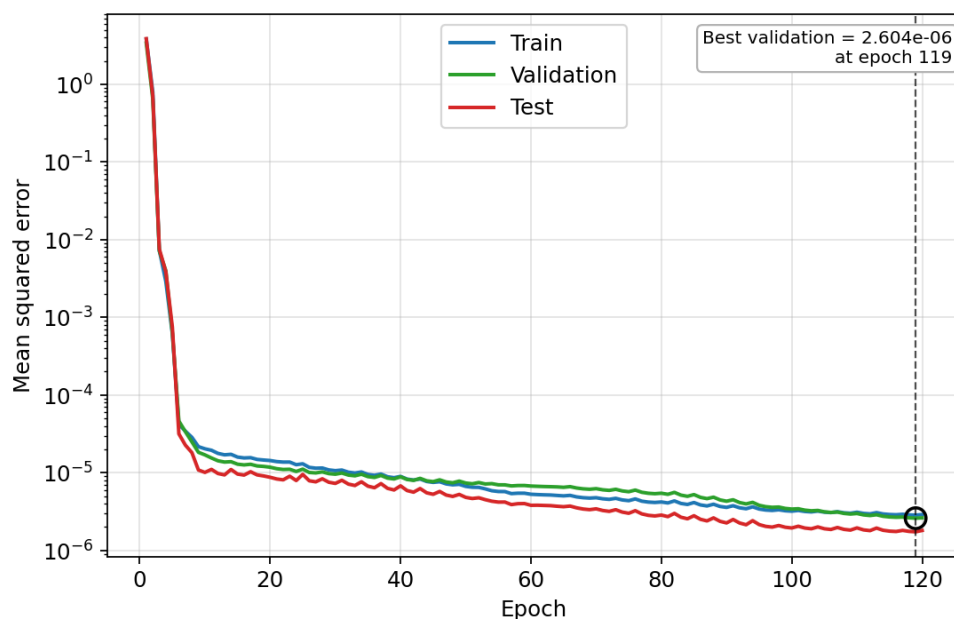


Fig. 14. Mean-squared-error performance for the alcohol-free regime; training, validation and test errors against epoch with the best-validation epoch marked

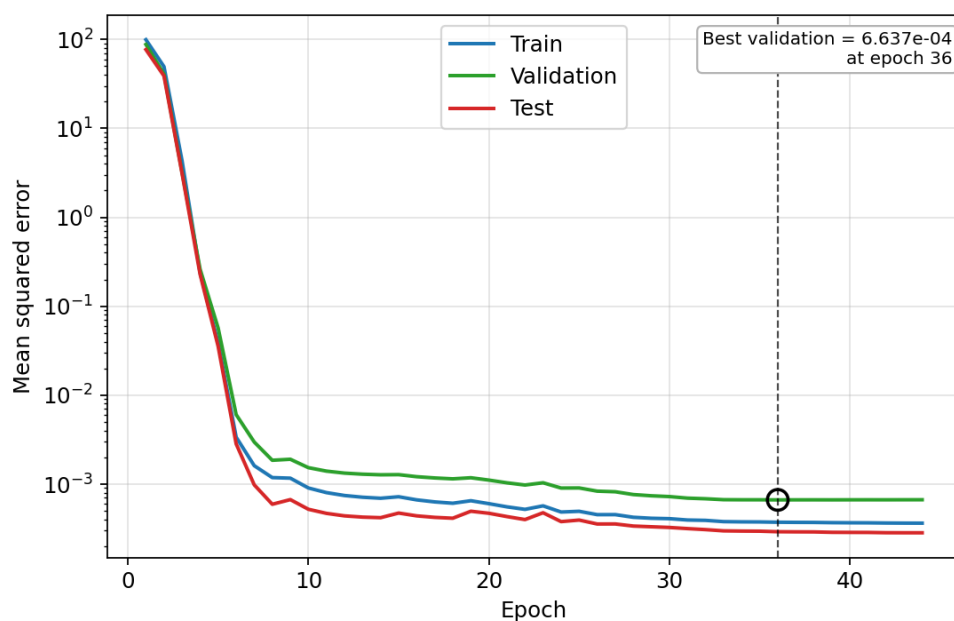


Fig. 15. Mean-squared-error performance for the alcohol-present regime; training, validation and test errors against epoch with the best-validation epoch marked

The statistical quality of the residuals is examined next. The error autocorrelations in Figure 16 and Figure 17 are sharply peaked at zero lag and lie within or close to the ninety-five percent confidence band elsewhere, reflecting small, smoothly distributed residuals.

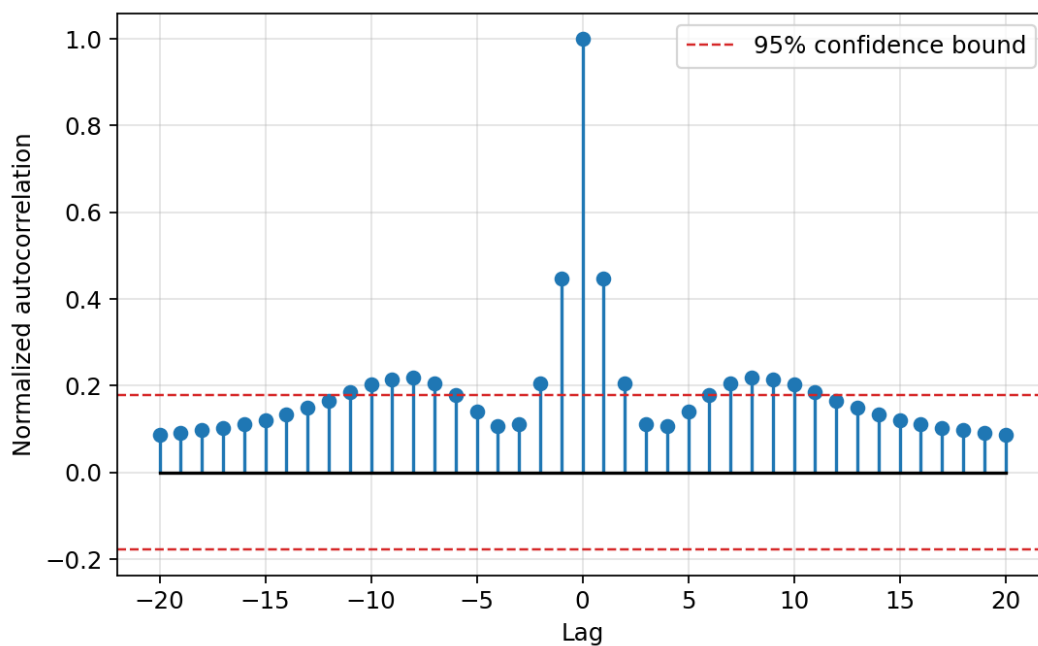


Fig. 16. Error autocorrelation for the alcohol-free regime, shown with the ninety-five percent confidence bounds

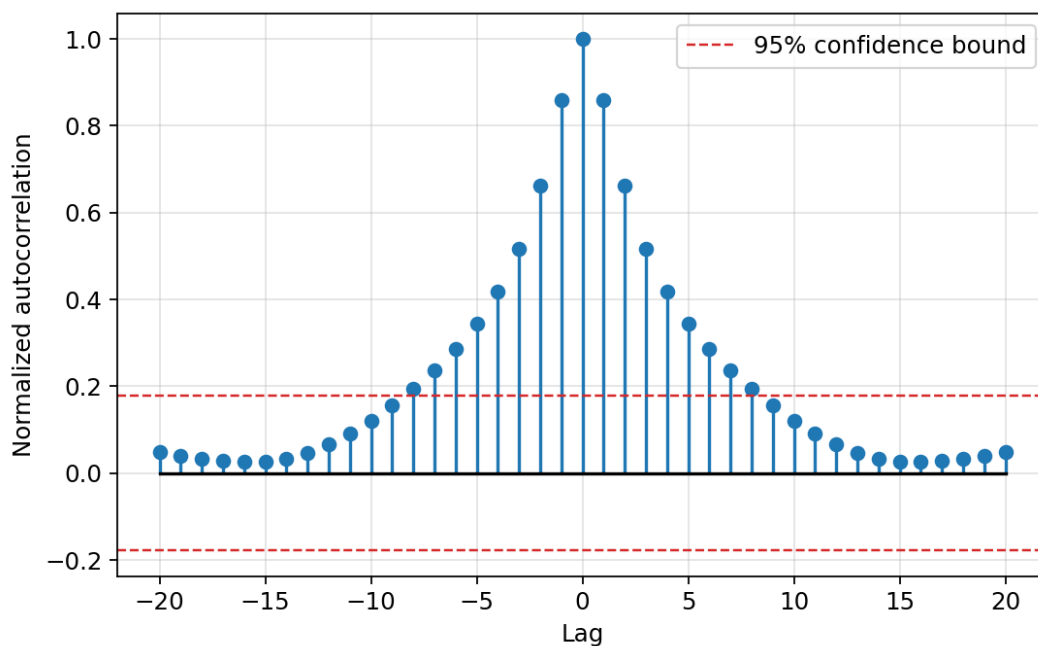


Fig. 17. Error autocorrelation for the alcohol-present regime, shown with the ninety-five percent confidence bounds

The input error cross-correlations in Figure 18 and Figure 19 remain near the confidence band, indicating little systematic dependence of the residual on the time input.

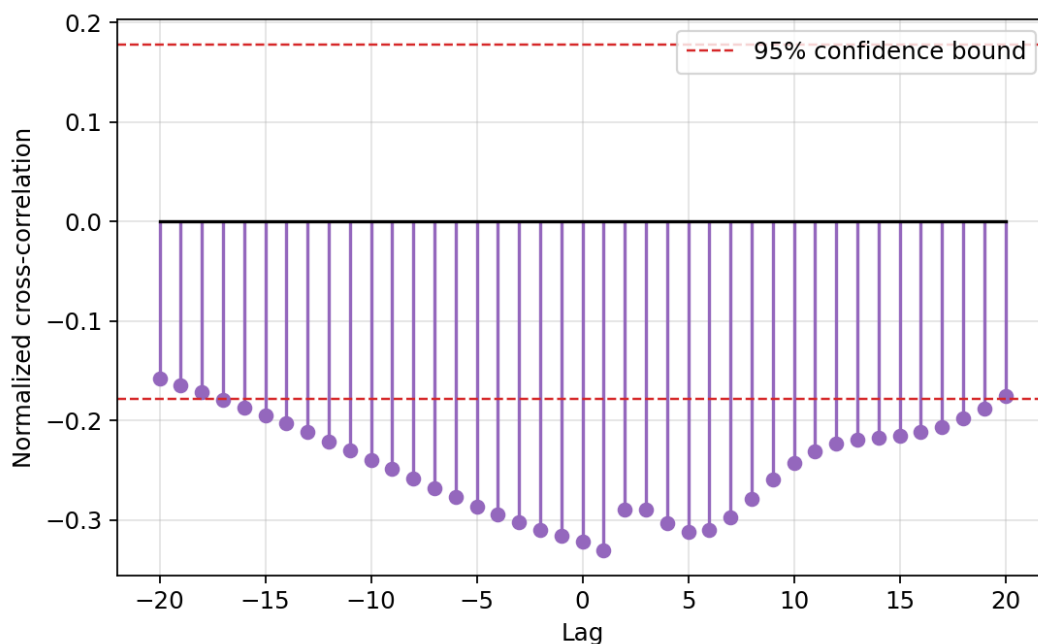


Fig. 18. Input-error cross-correlation for the alcohol-free regime, shown with the ninety-five percent confidence bounds

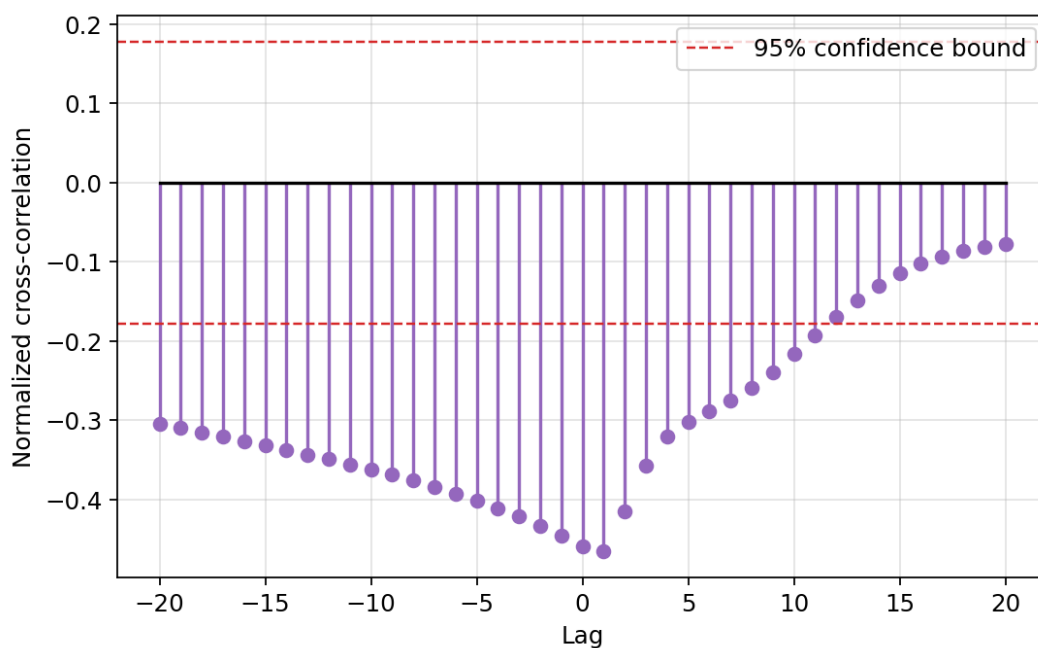


Fig. 19. Input error cross-correlation for the alcohol-present regime, shown with the ninety-five percent confidence bounds

The error histograms in Figure 20 and Figure 21 are tightly concentrated about zero for the training, validation and test partitions alike.

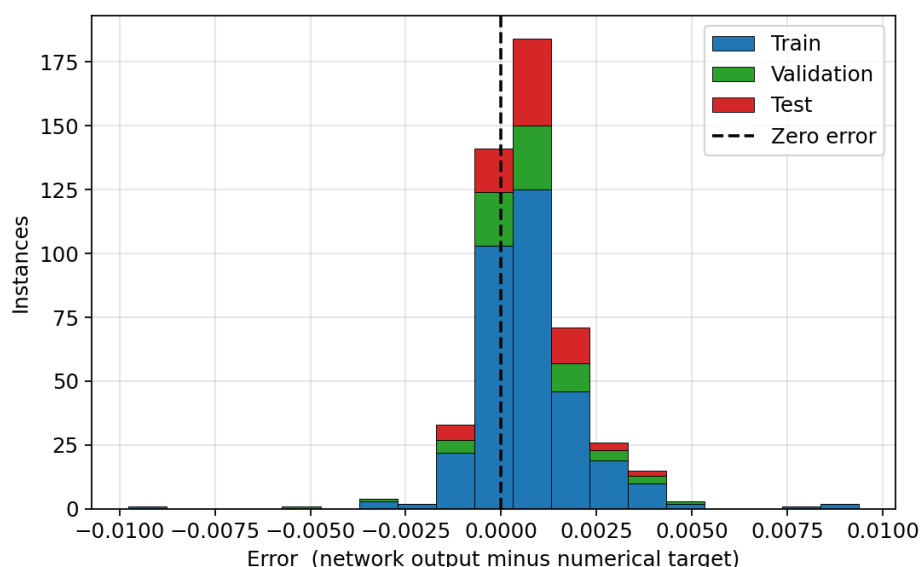


Fig. 20. Error histogram for the alcohol-free regime over the training, validation and test partitions, with the zero-error line indicated

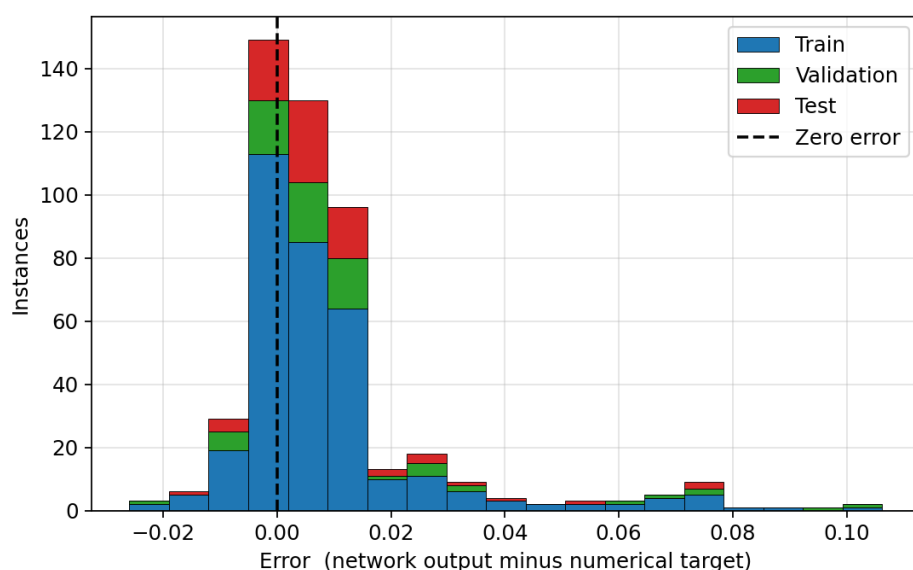


Fig. 21. Error histogram for the alcohol-present regime over the training, validation and test partitions, with the zero-error line indicated

The agreement between the surrogate and the numerical solution is quantified by the regression plots in Figure 22 and Figure 23 and visualised in the time-series reconstructions of Figure 24 and Figure 25. Network output superimposed on the numerical target for all four compartments, with the pointwise error in the lower panel (Figure 24 and Figure 25). The regression coefficients are essentially unity, and the data collapse onto the line of perfect fit across all partitions. The reconstructed trajectories are indistinguishable from the target within the plotted error bands, both for the monotone relaxation of the alcohol-free regime and for the sharp overshoot of the occasional-user class in the alcohol-present regime. The network therefore captures the memory-driven, fractal-time dynamics faithfully and acts as an accurate, continuous and independently validated solver.

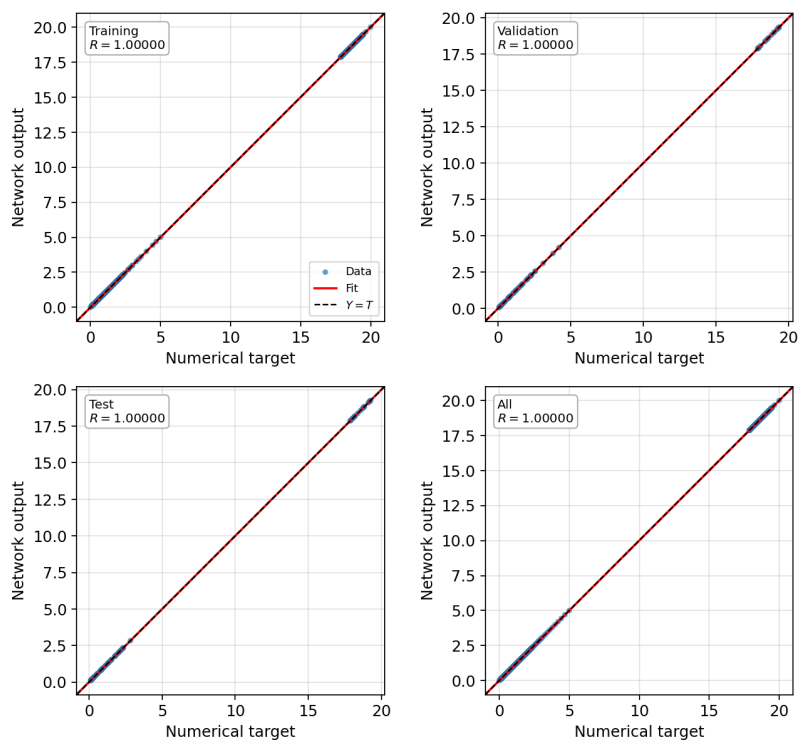


Fig. 22. Regression of network output against numerical target for the alcohol-free regime, for the training, validation, test and combined sets with the best-fit and reference lines

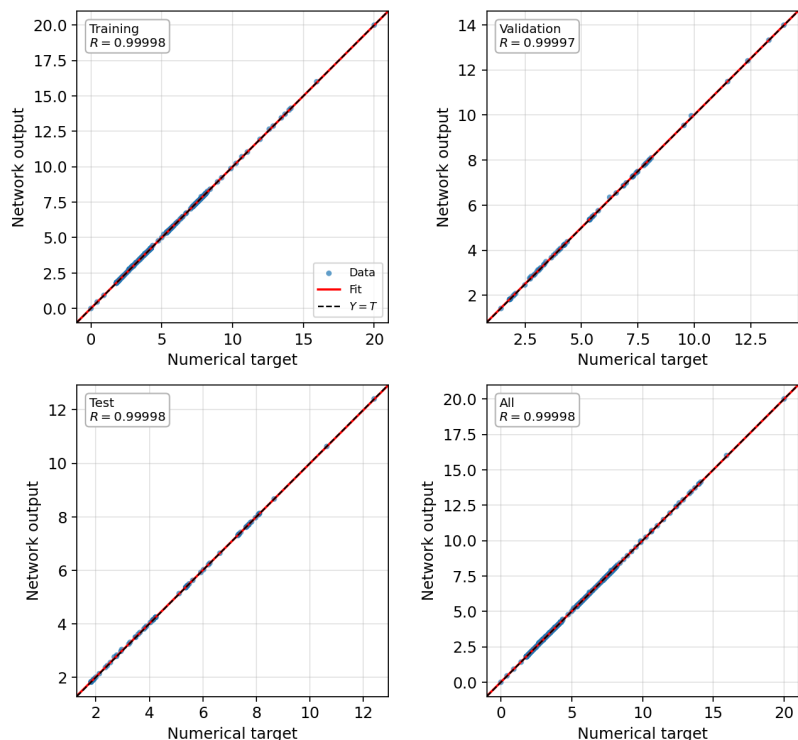


Fig. 23. Regression of network output against numerical target for the alcohol-present regime, for the training, validation, test and combined sets with the best-fit and reference lines

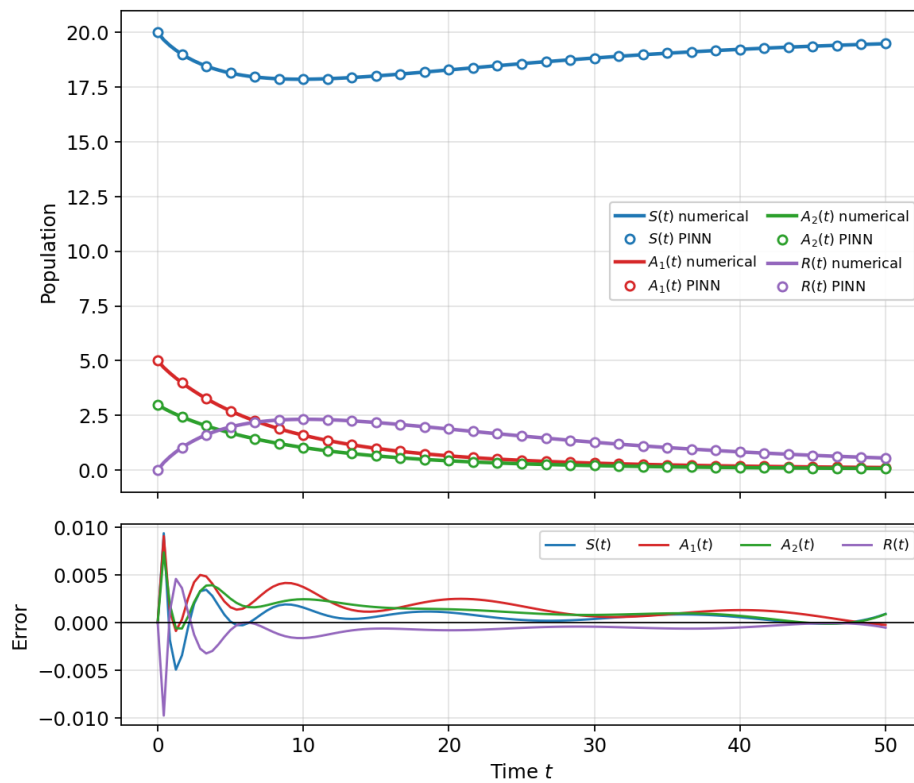


Fig. 24. Time-series reconstruction for the alcohol-free regime

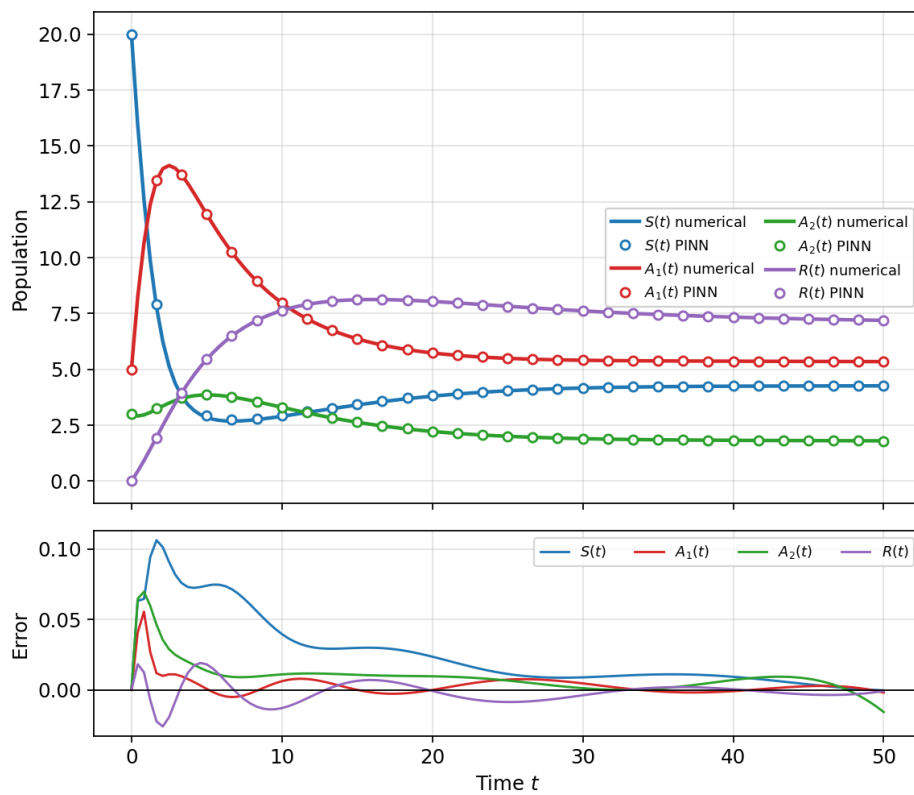


Fig. 25. Time-series reconstruction for the alcohol-present regime

The dynamical and learning pictures are brought together in Figure 26. Its first two panels launch

orbits from a spread of admissible initial states and project them onto the (S, A_1, A_2) subspace: above threshold every orbit is captured by the interior equilibrium E^* , and below threshold every orbit is drawn to the alcohol-free state E_0 , so the global stability proved in Theorems 3.8 and 3.9 is seen to hold from many starting points rather than a single one. The third panel lifts the surrogate error into three dimensions, plotting the absolute deviation of the network from the numerical target for each compartment across the horizon; the ribbons stay close to the floor throughout and swell only during the early transient, which is where the fractal-fractional memory is most active. Panels (a) and (b) project trajectories launched from a spread of initial states onto the (S, A_1, A_2) subspace: in the alcohol-present regime (a) every orbit is attracted to the interior equilibrium E^* (star), while in the alcohol-free regime (b) every orbit collapses onto E_0 (star), so both attractors are reached from all admissible data. Panel (c) plots the absolute error of the network surrogate against time for each compartment in the alcohol-present regime, the ribbons remaining small over the whole horizon and rising only through the initial transient.

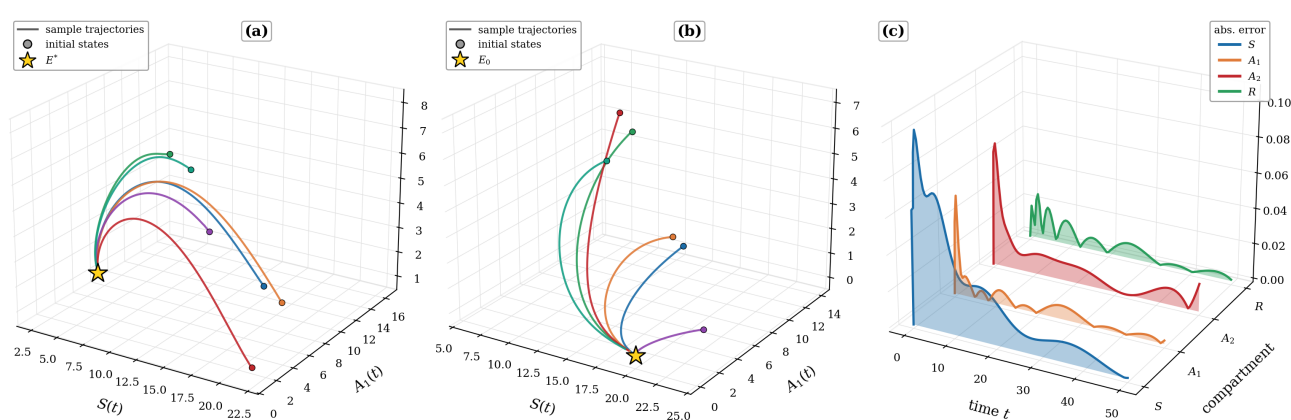


Fig. 26. Phase-space geometry and the three-dimensional distribution of the surrogate error

3.6 Comparison with Existing Approaches

It is useful to position the present framework against the main lines of work on alcohol and related behavioural-epidemic models. Classical formulations in the spirit of the epidemiological texts of Brauer *et al.* [1] and Hethcote [2] use integer-order derivatives, which are local in time and therefore treat the present rate of change as independent of the drinking history. They capture threshold behaviour through $\mathcal{R}_0$ but cannot represent the cumulative nature of habit formation. Fractional alcohol models based on the singular power-law Caputo kernel, such as the blood-alcohol studies of Prakash *et al.* [8] and of Wanassi & Baleanu [7], introduce true memory but inherit a kernel singularity at the origin. Models built on the Caputo-Fabrizio operator, analysed for nonlinear systems by Kebede & Lakoud [24] and applied to disease and cancer dynamics by Kumar *et al.* [25] and by Chavada & Pathak [26], remove that singularity through an exponential kernel but provide a faster, less flexible memory decay than the Mittag-Leffler kernel and do not encode fractal temporal scaling. Atangana-Baleanu models with the Mittag-Leffler kernel, including the freelance dynamics of Khan *et al.* [23], supply non-singular memory but, in their purely fractional form, still lack a fractal order. The fractal-fractional operator of Atangana [27], used for social-media addiction [29] and for epidemic spread [30], restores that missing degree of freedom. The recent binge-drinking model of Bhosale *et al.* [6] confirms the growing interest in fractional descriptions of drinking. Table 5 places these strands side by side.

Three features distinguish the present study within this body of work. First, it couples non-singular Mittag-Leffler memory with a fractal order in a single alcohol-addiction model, so that the depth of

Table 5

Qualitative comparison of the present framework with representative modelling approaches for alcohol and related behavioural dynamics

Approach	Non-singular kernel	Memory	Fractal order	Global stability	Neural check
Integer-order models [1, 2]	no	no	no	partial	no
Caputo fractional [7, 8]	no	yes	no	partial	no
Caputo-Fabrizio [24–26]	yes	yes	no	partial	no
ABC fractional [23]	yes	yes	no	partial	no
Fractal-fractional [29, 30]	yes	yes	yes	partial	no
This work	yes	yes	yes	yes	yes

memory and the pace of fractal time are tuned independently through α and κ . Second, it develops a detailed qualitative theory for that model, comprising positivity, boundedness, an explicit existence and uniqueness result, the next-generation derivation of $\mathcal{R}_0$, and both local and global stability of the two equilibria, whereas many fractal-fractional studies stop at well-posedness and simulation. Third, it cross-validates the deterministic solver with a physics-informed neural network that embeds the discrete fractal-fractional residual in its loss, an independent corroboration that, to our knowledge, has not previously accompanied alcohol-addiction models of this type. Taken together these elements give a framework that is at once more general in its operator, wider in the analysis it develops and more thoroughly verified than the approaches summarised in Table 5.

4. Conclusions

This research proposed and analysed a four-compartment alcohol-addiction model formulated in the Atangana-Baleanu-Caputo fractal-fractional sense, in which the non-singular Mittag-Leffler kernel carries the behavioural memory of drinking and a fractal order captures irregular temporal scaling. The classical integer-order description is recovered as the orders tend to one, so the model is a true generalisation rather than a reparametrisation.

The qualitative theory of the model has been worked out in detail. A contraction argument based on the Volterra representation and the Banach fixed-point principle yields existence and uniqueness of solutions, and positivity together with boundedness confine every trajectory to a biologically meaningful invariant region. The alcohol-free and alcohol-present equilibria are obtained in closed form, the basic reproduction number $\mathcal{R}_0$ follows from the next-generation matrix, and the stability analysis shows that the alcohol-free state is locally and globally asymptotically stable when $\mathcal{R}_0 < 1$, while a unique alcohol-present state emerges and is locally and globally asymptotically stable when $\mathcal{R}_0 > 1$. The threshold $\mathcal{R}_0$ thus governs the elimination or persistence of drinking, and the expression derived here identifies the transmission and recovery rates as the natural levers for intervention.

The analytical predictions were confirmed numerically through an Atangana-Toufik product-integration scheme, which also revealed how the fractional and fractal orders reshape the transient without altering the eventual outcome dictated by $\mathcal{R}_0$. As an independent, mesh-free corroboration, a physics-informed neural network was trained for both regimes; embedding the discrete fractal-fractional operator in its loss and optimising with a Levenberg-Marquardt scheme, it reproduced the compartmental trajectories with sub-percent root-mean-square error and near-perfect regression,

supported by a broad suite of training diagnostics. The agreement between two methods that share no discretisation philosophy reinforces confidence in the computed dynamics.

Several extensions are natural. The framework can be enriched with optimal control of prevention and treatment, with stochastic perturbations, time delays, age structure or spatial diffusion, and it can be calibrated against epidemiological and sociological data so that the orders α and κ are estimated rather than prescribed. Each of these directions would sharpen the model as a practical instrument for designing and evaluating alcohol-intervention policy.

Acknowledgement

This research was not funded by any grant.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Brauer, F., Castillo-Chavez, C., & Feng, Z. (2012). *Mathematical models in epidemiology*. Springer.
- [2] Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42(4), 599–653. <https://doi.org/10.1137/S0036144500371907>
- [3] Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28(4), 365–382. <https://doi.org/10.1007/BF00178324>
- [4] Van den Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1–2), 29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6)
- [5] Li, S., Zhang, Y., Wang, H., & Chen, X. (2024). Global dynamics and computational modeling approach for alcohol addiction with treatment intervention. *Mathematical Biosciences and Engineering*, 21(8), 14231–14258. <https://doi.org/10.1038/s41598-024-54578-9>
- [6] Bhosale, S., Lande, C., Jafari, H., & Raut, D. K. (2026). Mathematical analysis and numerical simulation of a fractional-order alcohol consumption model incorporating binge drinkers. *Boundary Value Problems*, 2026, 69. <https://doi.org/10.1186/s13661-026-02269-x>
- [7] Wanassi, O. K., & Baleanu, D. (2024). Modeling blood alcohol concentration using fractional differential equations based on the -caputo derivative. *Mathematical Methods in the Applied Sciences*, 47(12), 9451–9470. <https://doi.org/10.1002/mma.10002>
- [8] Prakash, A., Kumar, S., & Singh, R. (2024). Analysis and numerical simulation of fractional-order blood alcohol models. *Computational and Mathematical Biophysics*, 12(1), 1–20. <https://doi.org/10.1515/cmb-2024-0001>
- [9] Podlubny, I. (1999). *Fractional differential equations*. Academic Press.
- [10] Kilbas, A. A., Srivastava, H. M., & Trujillo, J. J. (2006). *Theory and applications of fractional differential equations*. Elsevier.
- [11] Diethelm, K. (2010). *The analysis of fractional differential equations*. Springer.
- [12] Li, C., & Zeng, F. (2020). *Numerical methods for fractional calculus*. CRC Press.
- [13] Garrappa, R. (2018). Numerical solution of fractional differential equations: A survey and software tutorial. *Mathematics*, 6(2), 16. <https://doi.org/10.3390/math6020016>
- [14] Tarasov, V. E. (2019). *Fractional dynamics: Applications of fractional calculus to dynamics of particles, fields and media*. Springer. <https://doi.org/10.1007/978-3-642-14003-7>

- [15] Machado, J. A. T., Kiryakova, V., & Mainardi, F. (2022). Recent history of fractional calculus. *Communications in Nonlinear Science and Numerical Simulation*, 102, 105939. <https://doi.org/10.1016/j.cnsns.2010.05.027>
- [16] Chen, Y. Q., Ding, H., & Li, C. (2021). Fractional-order systems and controls: Fundamentals and applications. *Annual Reviews in Control*, 51, 1–15. <https://doi.org/10.1007/978-1-84996-335-0>
- [17] Atangana, A., & Baleanu, D. (2016). New fractional derivatives with nonlocal and non-singular kernel: Theory and application to heat transfer model. *Thermal Science*, 20(2), 763–769. <https://doi.org/10.2298/TSCI16011018A>
- [18] Abdeljawad, T. (2020). Fractional operators with generalized mittag-leffler kernels and their applications. *Advances in Difference Equations*, 2020, 468. <https://doi.org/10.1063/1.5085726>
- [19] Djida, J. D., Area, I., & Atangana, A. (2017). New numerical scheme of atangana-baleanu fractional integral: Application to groundwater flow within leaky aquifers. *European Physical Journal Plus*, 132(5), 1–17.
- [20] Algahtani, O. J. J. (2021). Comparing the atangana-baleanu and caputo-fabrizio derivatives with fractional order. *Chaos, Solitons & Fractals*, 89, 552–559. <https://doi.org/10.1016/j.chaos.2016.03.026>
- [21] Martínez-Fuentes, O., & Delfín-Prieto, S. M. (2020). Stability of fractional nonlinear systems with mittag-leffler kernels and design of state observers. *Chaos, Solitons & Fractals*, 140, 110220.
- [22] Sabatier, J. (2020). Fractional-order derivatives defined by continuous kernels: Are they really too restrictive? *Fractal and Fractional*, 4(3), 40. <https://doi.org/10.3390/fractalfract4030040>
- [23] Khan, F. S., Khan, A., & Baleanu, D. (2022). Freelance model with atangana-baleanu-caputo fractional derivatives. *Symmetry*, 14(11), 2424. <https://doi.org/10.3390/sym14112424>
- [24] Kebede, S. G., & Lakoud, A. G. (2023). Analysis of mathematical models involving nonlinear systems of caputo-fabrizio fractional differential equations. *Boundary Value Problems*, 2023(1), 44. <https://doi.org/10.1186/s13661-023-01730-5>
- [25] Kumar, P., Kumar, S., Alkahtani, B. S. T., & Alzaid, S. S. (2024). A mathematical model for simulating the spread of infectious disease using the caputo-fabrizio fractional-order operator. *AIMS Mathematics*, 9(11), 30864–30897. <https://doi.org/10.3934/math.20241490>
- [26] Chavada, A., & Pathak, N. (2024). Transmission dynamics of breast cancer through caputo-fabrizio fractional derivative operator with real data. *Mathematical Modelling and Control*, 4(1), 119–132. <https://doi.org/10.3934/mmc.2024011>
- [27] Atangana, A. (2021). Fractal-fractional differentiation and integration: Connecting fractal geometry with memory effects. *Chaos, Solitons & Fractals*, 102, 396–406. <https://doi.org/10.1016/j.chaos.2017.04.027>
- [28] Rahman, M. (2022). Generalized fractal-fractional order problems under non-singular mittag-leffler kernel. *Results in Physics*, 35, 105346. <https://doi.org/10.1016/j.rinp.2022.105346>
- [29] Chithra, S. M., Kumar, R., & Devi, M. (2025). Fractal-fractional modelling of social media addiction with atangana-baleanu operators. *European Journal of Pure and Applied Mathematics*, 18(1), 1–25. <https://doi.org/10.29020/nybg.ejpam.v18i4.6737>
- [30] Ahmed, S., Jahan, S., & Nisar, K. S. (2025). A fractional model for the dynamics of covid-19 using atangana-baleanu fractional operators. *Journal of Mathematics and Computer Science*, 39(2), 233–248.
- [31] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>

- [32] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440. <https://doi.org/10.1038/s42254-021-00314-5>
- [33] Pang, G., Lu, L., & Karniadakis, G. E. (2019). fPINNs: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing*, 41(4), A2603–A2626. <https://doi.org/10.1137/18M1229845>
- [34] Banach, S. (1922). Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3(1), 133–181.
- [35] Matignon, D. (1996). Stability results for fractional differential equations with applications to control processing. *Proceedings of the Computational Engineering in Systems Applications Conference*, 963–968.
- [36] LaSalle, J. P. (1976). *The stability of dynamical systems*. SIAM.
- [37] Khalil, H. K. (2002). *Nonlinear systems* (3rd ed.). Prentice Hall.
- [38] Lakshmikantham, V., Leela, S., & Devi, J. V. (2009). *Theory of fractional dynamic systems*. Cambridge Scientific Publishers.
- [39] Toufik, M., & Atangana, A. (2017). New numerical approximation of fractional derivative with non-local and non-singular kernel: Application to chaotic models. *European Physical Journal Plus*, 132(10), 444. <https://doi.org/10.1140/epjp/i2017-11717-0>
- [40] Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989–993. <https://doi.org/10.1109/72.329697>
- [41] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431–441. <https://doi.org/10.1137/0111030>