

ARTICLE

Data-Driven Modeling and Prediction of a Fractional-Order Epidemic System: Dynamics, Discrete Physics-Informed Inverse Learning, and Application to Mpox

Saleh Albadran¹, Younes Brahim Oumedjber², Omar Kahouli^{3,*}, Imed Bouzida⁴, Moawia Farah⁵ and Ali Aloui³

¹Department of Electrical Engineering, College of Engineering, University of Ha'il, Ha'il, Saudi Arabia

²Laboratory of Materials Science for Applications and the Environment (LaSMAE), Department of Sciences, Teacher Education College of Sétif-Messaoud Zeghar, El Eulma, Sétif, Algeria

³Department of Electronics Engineering, Applied College, University of Ha'il, Ha'il, Saudi Arabia

⁴Department of Mathematics and Statistics, College of Engineering, Abu Dhabi University, Abu Dhabi, United Arab Emirates

⁵Department of Management Information Systems, Applied College, University of Ha'il, Ha'il, Saudi Arabia

*Corresponding Author: Omar Kahouli. Email: a.Kahouli@uoh.edu.sa

Received: 20 May 2026; Accepted: 30 July 2026; Published: 21 October 2026

ABSTRACT: Mpox outbreaks continue to challenge public-health surveillance because reported cases and deaths provide only a partial view of the underlying transmission process. In particular, latent infection, asymptomatic transmission, quarantine, hospitalization, and reporting effects cannot be read directly from routine aggregate surveillance data. A fractional-order epidemic model with susceptible, exposed, vaccinated, infected, asymptomatic, quarantined, hospitalized, recovered, and deceased compartments (SEVIAQHRD) is proposed for the global mpox outbreak. The resulting framework combines mathematical analysis, observation-level calibration, and a discrete physics-informed neural network (PINN) inverse-learning stage. Vaccination, asymptomatic transmission, quarantine, hospitalization, waning immunity, and disease-induced mortality are incorporated into the model. The main analytical properties of the system are established, and the latent epidemic dynamics are linked to reported daily and cumulative cases and deaths through an observation layer. A mechanistic calibration procedure is first employed to infer observation-compatible epidemic trajectories, after which a neural inference stage constrained by the same discrete fractional dynamics is used for reduced-parameter estimation, latent-state reconstruction, validation, and short-horizon forecasting. Over the highlighted 20-day segment within the shared 120-day observation window, the calibrated fractional configuration yields lower observation-level errors than the controlled integer-order limit obtained by setting $\vartheta = 1$ while retaining the same parameterization. The PINN stage is evaluated using the first 100 days for training and the final 20 days as a held-out short-horizon test segment. An observation-level structural identifiability assessment further shows that the aggregate surveillance outputs constrain composite parameter-state combinations rather than all calibrated quantities independently.

KEYWORDS: Mpox; fractional-order epidemic model; SEVIAQHRD model; observation layer; physics-informed inverse learning; short-horizon forecasting

1 Introduction

Mpox has re-emerged as a major public-health concern because of its cross-regional spread, heterogeneous clinical presentation, and time-varying transmission patterns [1–5]. A recent review provides a

broader synthesis of mpox modeling studies [6]. Routine surveillance systems mainly provide daily and cumulative counts of reported cases and deaths, whereas several epidemiologically relevant states, including latent, asymptomatic, quarantined, and hospitalized populations, remain unobserved [7,8]. This partial observability motivates the use of mechanistic compartmental models to reconstruct hidden transmission pathways, assess intervention effects, and estimate epidemic burden from limited surveillance outputs [9,10].

Classical compartmental models provide an interpretable description of transmission dynamics and threshold behavior, but integer-order formulations do not explicitly represent nonlocal temporal dependence. Fractional-order models extend this framework by incorporating memory and hereditary effects and have recently been applied to mpox [11–13], Ebola [14], and generalized susceptible–exposed–infectious–recovered (SEIR) dynamics [15,16]. The Caputo derivative is adopted here because it accommodates standard initial conditions with direct epidemiological meaning while retaining a nonlocal representation of the epidemic history. Given the relatively short observation window considered in this study, the fractional formulation is interpreted as a memory-capable extension of the classical model rather than as evidence of fully resolved long-range memory in the surveillance data.

The use of a fractional epidemic model in an inverse problem raises two complementary challenges. First, the model must satisfy the basic analytical requirements of well-posedness, positivity, boundedness, global continuation, and threshold consistency. Second, the latent state variables must be connected to the reported surveillance quantities through an epidemiologically interpretable observation map. Physics-informed neural networks provide a useful framework for combining incomplete observations with governing equations [17–20]. In a fractional setting, however, the discrete representation of the nonlocal operator must remain consistent across the mechanistic and neural stages. For this reason, the present neural inverse-learning procedure enforces the same discrete L1 approximation of the Caputo derivative as the forward solver.

Against this background, we develop a nonautonomous fractional SEVIAQHRD model for the first global mpox wave. The model incorporates vaccination, asymptomatic infection, quarantine, hospitalization, waning immunity, and disease-induced mortality. The complete framework combines mathematical analysis, observation-level calibration, and discrete physics-informed inverse learning for latent-state reconstruction, reduced-parameter estimation, validation, and short-horizon forecasting. The sequence and interaction of these components are summarized in Fig. 1.

The contributions of this study are threefold. First, we formulate and analyze a nonautonomous fractional SEVIAQHRD model and establish its main well-posedness, positivity, boundedness, continuation, and threshold properties. Second, we construct an observation-level calibration framework that connects the latent model dynamics to world-level reported cases and deaths, and we assess the resulting parameter interpretation through an observation-level structural identifiability map. Third, we compare the calibrated fractional configuration with its controlled integer-order limit and couple the mechanistic system to a discrete PINN-based inverse-learning stage that uses the same L1 fractional operator for reduced-parameter estimation, latent-state inference, validation, and short-horizon forecasting. These components provide a unified and interpretable framework for studying partially observed mpox dynamics without treating the reported surveillance series as direct measurements of all epidemic compartments.

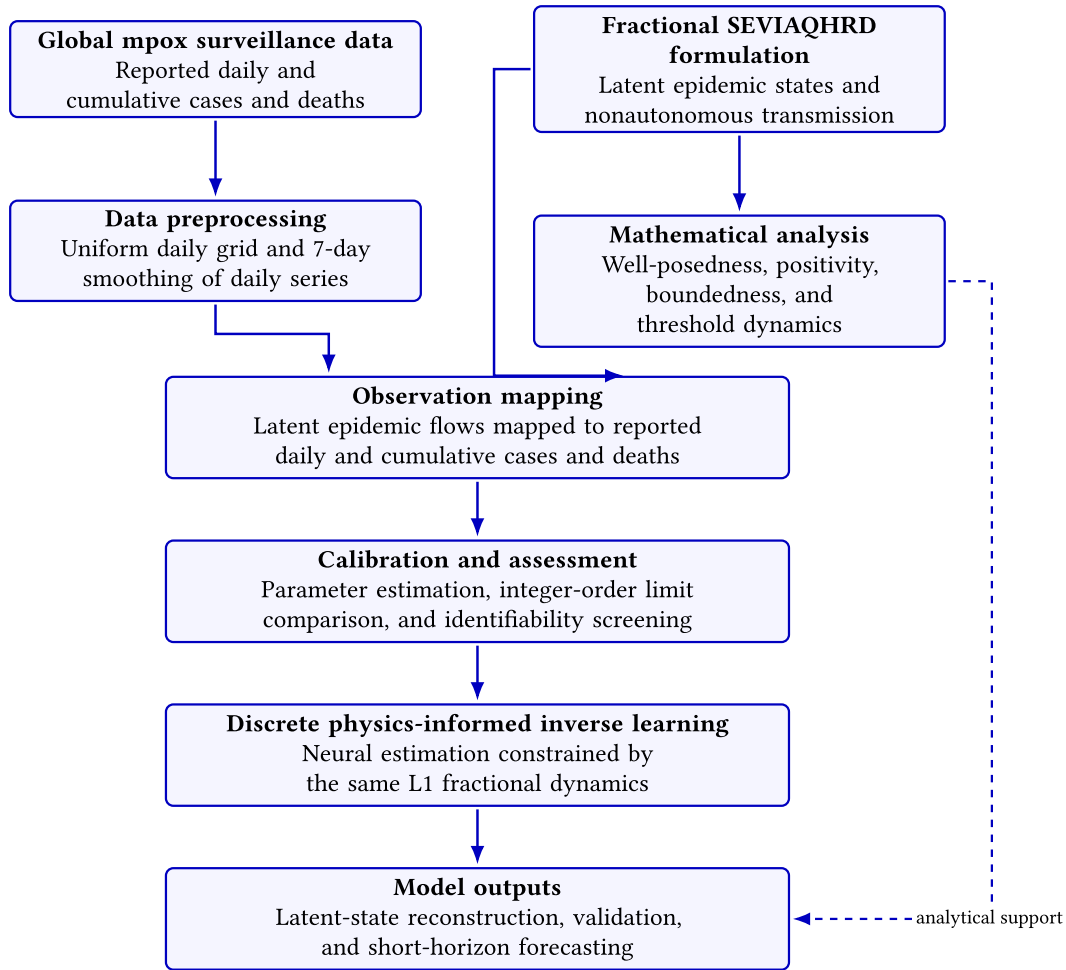


Figure 1: Schematic pipeline of the proposed data-driven fractional epidemic framework. Global mpox surveillance data are preprocessed and linked to the latent SEVIAQHRD dynamics through an explicit observation map. Mechanistic calibration provides observation-compatible trajectories and initial parameter information, while a controlled integer-order limit and an observation-level identifiability assessment clarify model comparison and parameter interpretation. The discrete physics-informed neural stage then enforces the same L1 fractional dynamics for inverse estimation, latent-state reconstruction, validation, and short-horizon forecasting.

2 Mathematical Model and Analysis

In this section, we formulate the fractional SEVIAQHRD model and summarize its main analytical properties, including well-posedness, positivity, boundedness, global continuation, and the basic threshold dynamics.

2.1 Fractional SEVIAQHRD Model

Let

$$Y(t) = (S, V, E, I, A, Q, H, R)^\top, \quad X(t) = (Y(t), D(t))^\top,$$

where S, V, E, I, A, Q, H, R , and D denote the susceptible, vaccinated, exposed, symptomatic infectious, asymptomatic infectious, quarantined, hospitalized, recovered, and disease-induced mortality compartments, respectively. For $x \in C^1([0, T])$ and $0 < \vartheta < 1$, the Caputo fractional derivative is defined by

$${}_0^C D_t^\vartheta x(t) = \frac{1}{\Gamma(1-\vartheta)} \int_0^t (t-\tau)^{-\vartheta} x'(\tau) d\tau,$$

see [21]. For $\vartheta = 1$, it reduces to the classical first-order derivative $x'(t)$. For brevity, the time argument is omitted in the fractional SEVIAQHRD model (1) when no confusion arises.

The fractional SEVIAQHRD model is

$$\begin{cases} {}_0^C D_t^\vartheta S = \Lambda_\vartheta + \omega V + \xi_\vartheta R - (\lambda_S + \tau + \delta_\vartheta) S, \\ {}_0^C D_t^\vartheta V = \tau S - (\lambda_V + \omega + \delta_\vartheta) V, \\ {}_0^C D_t^\vartheta E = \lambda_S S + \lambda_V V - (\sigma_\vartheta + \delta_\vartheta) E, \\ {}_0^C D_t^\vartheta I = (1-\omega)\sigma_\vartheta E - (\gamma_{I,\vartheta} + q_\vartheta + h_{I,\vartheta} + \nu_{I,\vartheta} + \delta_\vartheta) I, \\ {}_0^C D_t^\vartheta A = \omega\sigma_\vartheta E - (\gamma_{A,\vartheta} + \delta_\vartheta) A, \\ {}_0^C D_t^\vartheta Q = q_\vartheta I - (\gamma_{Q,\vartheta} + h_{Q,\vartheta} + \nu_{Q,\vartheta} + \delta_\vartheta) Q, \\ {}_0^C D_t^\vartheta H = h_{I,\vartheta} I + h_{Q,\vartheta} Q - (\gamma_{H,\vartheta} + \nu_{H,\vartheta} + \delta_\vartheta) H, \\ {}_0^C D_t^\vartheta R = \gamma_{I,\vartheta} I + \gamma_{A,\vartheta} A + \gamma_{Q,\vartheta} Q + \gamma_{H,\vartheta} H - (\xi_\vartheta + \delta_\vartheta) R, \\ {}_0^C D_t^\vartheta D = \nu_{I,\vartheta} I + \nu_{Q,\vartheta} Q + \nu_{H,\vartheta} H. \end{cases} \quad (1)$$

$$\lambda_S(t) = (1-\chi)\beta(t) \frac{\beta_I I(t) + \beta_A A(t)}{N(t)}, \quad \lambda_V(t) = (1-\rho)\lambda_S(t), \quad (2)$$

$$N(t) = S(t) + V(t) + E(t) + I(t) + A(t) + Q(t) + H(t) + R(t), \quad \beta(t) = e^{-kt}. \quad (3)$$

Here χ represents population-wide mitigation, ρ the additional protection due to vaccination, and $k > 0$ the rate of transmission decay associated with behavioral adaptation, awareness, and intervention effects. The exponential transmission profile is adopted as a parsimonious representation of rapid attenuation during the early global mpox wave. Epidemiologically, the decay parameter k summarizes the combined effects of behavioral adaptation, increased awareness, targeted isolation, contact reduction, and progressive public-health response. This choice is not intended to assert that mpox transmission must follow an exponential law over longer periods. Rather, it provides a low-dimensional time-varying transmission mechanism suited to the short first-wave window considered here. More flexible alternatives, such as piecewise-constant, logistic, spline-based, or fully data-driven transmission profiles, can be incorporated in future extensions when longer or more resolved data are available.

The main epidemiological parameters have the following biological interpretations. The recruitment term Λ_ϑ represents inflow into the susceptible class, τ is the vaccination rate, and ω is the loss of vaccine-induced protection. The quantities β_I and β_A denote transmission contributions from symptomatic and asymptomatic infectious individuals, respectively, while χ models population-wide mitigation and ρ the additional protection associated with vaccination. The progression rate from exposure is governed by σ_ϑ , with ω denoting the fraction entering the asymptomatic class. The parameters q_ϑ , $h_{I,\vartheta}$, and $h_{Q,\vartheta}$ describe quarantine and hospitalization flows; $\gamma_{I,\vartheta}$, $\gamma_{A,\vartheta}$, $\gamma_{Q,\vartheta}$, $\gamma_{H,\vartheta}$ are recovery rates; $\nu_{I,\vartheta}$, $\nu_{Q,\vartheta}$, $\nu_{H,\vartheta}$ are disease-induced mortality rates; δ_ϑ is the natural mortality rate; and ξ_ϑ represents waning immunity from the recovered class back to susceptibility. The model is restricted to human-level mpox transmission

compartments and their observation-level surveillance outputs. No ecological predator–prey mechanism, predator intervention, or rodent-consumption process is included in the present formulation.

The initial data are

$$X(0) = X^0 = (Y^0, D^0) \in \mathbb{R}_+^9, \quad S^0 > 0, \quad N^0 := \sum_{j=1}^8 Y_j^0 > 0.$$

Fig. 2 summarizes the transitions $S, V \rightarrow E \rightarrow \{I, A\} \rightarrow \{Q, H, R, D\}$ together with waning immunity $R \rightarrow S$; natural mortality is represented as a common loss acting on all living compartments.

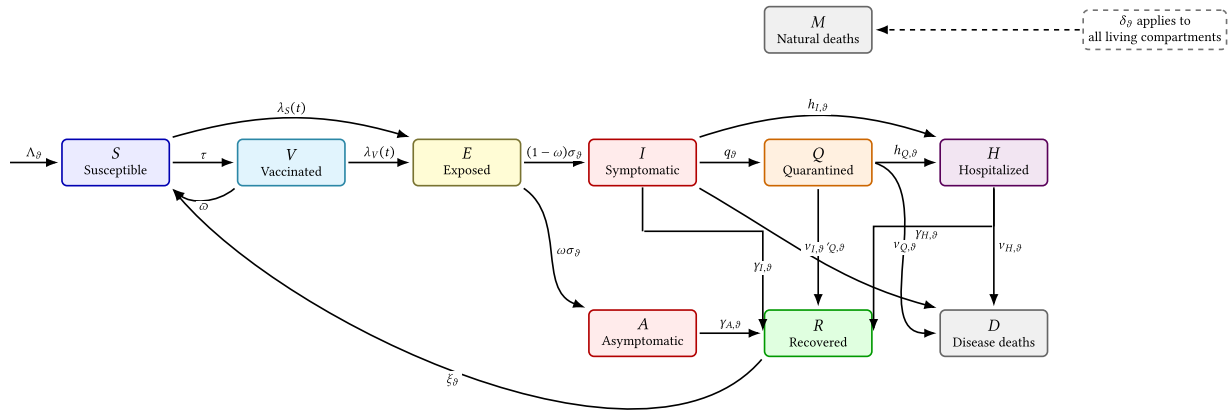


Figure 2: Compartmental flow diagram of the proposed fractional SEVIAQHRD model. The diagram summarizes the human-level transitions among susceptible, vaccinated, exposed, symptomatic, asymptomatic, quarantined, hospitalized, recovered, and disease-induced death compartments. The node M denotes natural mortality acting on all living compartments; no predator–prey or ecological intervention component is included.

Remark 1: All transition and removal rates in (1) and (2) have unit $\text{day}^{-\vartheta}$. In particular, β_I and β_A are transmission-rate parameters, whereas $\beta(t) = e^{-kt}$ is dimensionless. Also, Λ_g has unit $\text{persons day}^{-\vartheta}$, k has unit day^{-1} , and χ, ρ, ω are dimensionless. For $\vartheta = 1$, the model reduces to its classical integer-order counterpart.

Remark 2: Because $\beta(t) = e^{-kt}$, the system is nonautonomous. Accordingly, the baseline threshold $\mathcal{R}_0$ is derived from the frozen invasion system at $t = 0$, while $\mathcal{R}_e(t) = e^{-kt}\mathcal{R}_0$ is used only as a time-dependent transmission indicator.

Remark 3: For $0 < \vartheta < 1$, $D(t)$ is a memory-driven mortality stock rather than a classical instantaneous cumulative count. Hence the cumulative reported deaths introduced later are reconstructed in observation space and are not identified literally with the state variable $D(t)$.

Since D does not feed back into the first eight equations, the analysis focuses on the living subsystem Y . Summing the living equations yields

$${}_0^C D_t^\vartheta N(t) = \Lambda_g - \delta_g N(t) - v_{I,\vartheta} I(t) - v_{Q,\vartheta} Q(t) - v_{H,\vartheta} H(t), \quad (4)$$

$${}_0^C D_t^\vartheta (N(t) + D(t)) = \Lambda_g - \delta_g N(t). \quad (5)$$

These identities are used in the invariant-region and boundedness analysis.

2.2 Well-Posedness and Positivity

With $Y = (S, V, E, I, A, Q, H, R)^\top$, the living subsystem of the model (1) is

$${}_0^C D_t^\vartheta Y(t) = \mathcal{G}(t, Y(t)), \quad Y(0) = Y^0, \quad (6)$$

where $\mathcal{G}$ denotes the first eight components of the full vector field. Let

$$\mathcal{O} := \left\{ Y \in \mathbb{R}^8 : N(Y) := \sum_{j=1}^8 Y_j > 0 \right\}.$$

Since $N(Y^0) > 0$, we have $Y^0 \in \mathcal{O}$, and (6) is equivalent to

$$Y(t) = Y^0 + \frac{1}{\Gamma(\vartheta)} \int_0^t (t - \tau)^{\vartheta-1} \mathcal{G}(\tau, Y(\tau)) d\tau. \quad (7)$$

Theorem 1: Assume that all parameters of the system (1) are nonnegative and that $Y^0 \in \mathbb{R}_+^8$ satisfies $N(Y^0) > 0$. Then there exists $T^* > 0$ such that the system (1) admits a unique local solution

$$(Y, D) \in C([0, T^*], \mathcal{O} \times \mathbb{R}).$$

Moreover, the solution depends continuously on the initial data.

Proof: Fix $T > 0$. Since $N(Y^0) > 0$, there exists $r > 0$ such that $\overline{B_r(Y^0)} \subset \mathcal{O}$. On each compact subset of $[0, T] \times \overline{B_r(Y^0)}$, the vector field $\mathcal{G}$ is continuous in t and locally Lipschitz in Y . Hence, by the standard local theory for Caputo initial-value problems applied to (7), there exists a unique $Y \in C([0, T^*], \mathcal{O})$ for some $T^* > 0$ [21,22].

Once Y is known, D is uniquely recovered from

$$D(t) = D^0 + \frac{1}{\Gamma(\vartheta)} \int_0^t (t - \tau)^{\vartheta-1} (\nu_{I,\vartheta} I + \nu_{Q,\vartheta} Q + \nu_{H,\vartheta} H)(\tau) d\tau. \quad (8)$$

Continuous dependence follows from the same theory. $\square$

Proposition 1: If $X^0 = (Y^0, D^0) \in \mathbb{R}_+^9$ and $N(Y^0) > 0$, then the corresponding local solution satisfies

$$X(t) \in \mathbb{R}_+^9, \quad t \in [0, T^*].$$

Proof: Since $\mathcal{G}$ is quasi-positive on $\mathbb{R}_+^8 \cap \mathcal{O}$, standard positivity results for Caputo systems imply $Y(t) \in \mathbb{R}_+^8$ on $[0, T^*]$ [23]. Then (8), together with $D^0 \geq 0$, gives $D(t) \geq D^0 \geq 0$. $\square$

For $0 < \vartheta < 1$, $D(t)$ is nondecreasing by (8), but it follows a memory-dependent fractional law rather than a classical cumulative variable generated by an ordinary differential equation.

2.3 Boundedness and Global Continuation

Assume

$$\Lambda_\vartheta > 0, \quad \delta_\vartheta > 0.$$

Since the incidence terms involve $1/N(t)$, it is enough to show that $N(t)$ remains bounded above and strictly away from zero.

Lemma 1: Let (Y, D) be the local nonnegative solution of the system (1) on $[0, T^*]$. Then, for all $t \in [0, T^*]$,

$$N(t) \leq N(0)E_{\vartheta}(-\delta_{\vartheta}t^{\vartheta}) + \frac{\Lambda_{\vartheta}}{\delta_{\vartheta}}\left(1 - E_{\vartheta}(-\delta_{\vartheta}t^{\vartheta})\right), \quad (9)$$

$$N(t) \geq N(0)E_{\vartheta}(-c_{\vartheta}t^{\vartheta}) + \frac{\Lambda_{\vartheta}}{c_{\vartheta}}\left(1 - E_{\vartheta}(-c_{\vartheta}t^{\vartheta})\right), \quad (10)$$

where

$$c_{\vartheta} := \delta_{\vartheta} + \max\{v_{I,\vartheta}, v_{Q,\vartheta}, v_{H,\vartheta}\}.$$

Consequently,

$$m_{\vartheta} \leq N(t) \leq M_{\vartheta}, \quad t \in [0, T^*] \quad (11)$$

where

$$m_{\vartheta} := \min\left\{N(0), \frac{\Lambda_{\vartheta}}{c_{\vartheta}}\right\} > 0, \quad M_{\vartheta} := \max\left\{N(0), \frac{\Lambda_{\vartheta}}{\delta_{\vartheta}}\right\}.$$

Proof: From (4),

$${}_0^C D_t^{\vartheta} N = \Lambda_{\vartheta} - \delta_{\vartheta} N - v_{I,\vartheta} I - v_{Q,\vartheta} Q - v_{H,\vartheta} H.$$

Since $I, Q, H \geq 0$,

$${}_0^C D_t^{\vartheta} N \leq \Lambda_{\vartheta} - \delta_{\vartheta} N.$$

Comparing with

$${}_0^C D_t^{\vartheta} y = \Lambda_{\vartheta} - \delta_{\vartheta} y, \quad y(0) = N(0),$$

and using the fractional comparison principle yields (9); see [23].

Also, because $I + Q + H \leq N$,

$$v_{I,\vartheta} I + v_{Q,\vartheta} Q + v_{H,\vartheta} H \leq \max\{v_{I,\vartheta}, v_{Q,\vartheta}, v_{H,\vartheta}\}(I + Q + H) \leq (c_{\vartheta} - \delta_{\vartheta})N.$$

Hence

$${}_0^C D_t^{\vartheta} N \geq \Lambda_{\vartheta} - c_{\vartheta} N.$$

Comparison with

$${}_0^C D_t^{\vartheta} z = \Lambda_{\vartheta} - c_{\vartheta} z, \quad z(0) = N(0),$$

gives (10); again see [23].

Finally, for $a > 0$, $E_{\vartheta}(-at^{\vartheta}) \in (0, 1]$. Thus the right-hand sides of (9) and (10) lie between $\{N(0), \Lambda_{\vartheta}/\delta_{\vartheta}\}$ and $\{N(0), \Lambda_{\vartheta}/c_{\vartheta}\}$, respectively, which gives (11). $\square$

Proposition 2: *The set*

$$\Omega_{\vartheta} := \left\{ Y \in \mathbb{R}_+^8 : N(Y) \leq M_{\vartheta} \right\}$$

is positively invariant for (6).

Proof: By Proposition 1 and Lemma 1,

$$Y(t) \in \mathbb{R}_+^8 \quad \text{and} \quad N(t) \leq M_{\vartheta} \quad \text{for all } t \in [0, T^*].$$

Hence $Y(t) \in \Omega_{\vartheta}$ on $[0, T^*]$. $\square$

Theorem 2: *Under the assumptions of Theorem 1, together with $\Lambda_{\vartheta} > 0$ and $\delta_{\vartheta} > 0$, the solution of the system (1) extends uniquely to all $t \geq 0$. More precisely,*

$$Y \in C([0, \infty), \mathbb{R}_+^8 \cap \mathcal{O}), \quad D \in C([0, \infty), \mathbb{R}_+).$$

Proof: Let $[0, T_{\max})$ be the maximal existence interval and suppose $T_{\max} < \infty$. By Lemma 1,

$$m_{\vartheta} \leq N(t) \leq M_{\vartheta}, \quad 0 \leq t < T_{\max}.$$

Therefore,

$$Y(t) \in K_{\vartheta} := \left\{ Y \in \mathbb{R}_+^8 : m_{\vartheta} \leq N(Y) \leq M_{\vartheta} \right\} \subset \mathcal{O}, \quad 0 \leq t < T_{\max}.$$

Since $Y(t) \in \mathbb{R}_+^8$ and $N(Y(t)) \leq M_{\vartheta}$, each component satisfies $0 \leq Y_j(t) \leq M_{\vartheta}$; hence K_{ϑ} is compact.

Moreover, $\mathcal{G}$ is continuous on $[0, T_{\max}] \times K_{\vartheta}$, so it is bounded there. By (7), Y is uniformly Hölder continuous of order ϑ on $[0, T_{\max})$, hence uniformly continuous. Since $T_{\max} < \infty$, the limit

$$Y^* := \lim_{t \rightarrow T_{\max}^-} Y(t)$$

exists and belongs to $K_{\vartheta} \subset \mathcal{O}$.

Thus the trajectory stays in a compact subset of $\mathcal{O}$ up to $T_{\max}$. The continuation theorem for Caputo–Volterra problems [22] then contradicts the maximality of $T_{\max}$. Hence $T_{\max} = \infty$.

Once Y is globally defined, D is globally recovered from (8), and $D(t) \geq 0$ follows from Proposition 1. $\square$

Therefore, the model is globally well posed on $[0, \infty)$ and evolves in the positively invariant region Ω_{ϑ} .

2.4 Disease-Free State, Baseline Reproduction Number, and Stability

Since the system (1) is nonautonomous through $\beta(t) = e^{-kt}$, we first study the frozen system obtained by setting $e^{-kt} \equiv 1$, then derive a conservative extinction criterion for the original model.

Proposition 3: *The living subsystem of the model (1) admits a unique disease-free equilibrium*

$$\mathcal{E}^0 = \left(\frac{\Lambda_{\vartheta}(\bar{\omega} + \delta_{\vartheta})}{\delta_{\vartheta}(\tau + \bar{\omega} + \delta_{\vartheta})}, \frac{\Lambda_{\vartheta}\tau}{\delta_{\vartheta}(\tau + \bar{\omega} + \delta_{\vartheta})}, 0, 0, 0, 0, 0, 0 \right). \quad (12)$$

Consequently, the full system admits the disease-free manifold

$$\mathcal{M}^0 = \left\{ (\mathcal{E}^0, D^*) : D^* \geq 0 \right\}. \quad (13)$$

Proof: At the disease-free state,

$$E = I = A = Q = H = R = 0,$$

so (S^*, V^*) satisfies

$$0 = \Lambda_{\vartheta} + \omega V - (\tau + \delta_{\vartheta})S, \quad 0 = \tau S - (\omega + \delta_{\vartheta})V.$$

The determinant is

$$(\tau + \delta_{\vartheta})(\omega + \delta_{\vartheta}) - \tau\omega = \delta_{\vartheta}(\tau + \omega + \delta_{\vartheta}) > 0,$$

so the solution is unique and given by (12). At the disease-free state,

$${}_0^C D_t^{\vartheta} D = 0,$$

hence $D(t) \equiv D^* \geq 0$, which gives (13). $\square$

For later use, define the total removal rates

$$\begin{aligned} m_I &:= \gamma_{I,\vartheta} + q_{\vartheta} + h_{I,\vartheta} + \nu_{I,\vartheta} + \delta_{\vartheta}, & m_A &:= \gamma_{A,\vartheta} + \delta_{\vartheta}, \\ m_Q &:= \gamma_{Q,\vartheta} + h_{Q,\vartheta} + \nu_{Q,\vartheta} + \delta_{\vartheta}, & m_H &:= \gamma_{H,\vartheta} + \nu_{H,\vartheta} + \delta_{\vartheta}. \end{aligned} \quad (14)$$

Proposition 4: The baseline reproduction number $\mathcal{R}_0$ of the frozen system is given by

$$\mathcal{R}_0 = (1 - \chi) \frac{\omega + \delta_{\vartheta} + (1 - \rho)\tau}{\tau + \omega + \delta_{\vartheta}} \frac{\sigma_{\vartheta}}{\sigma_{\vartheta} + \delta_{\vartheta}} \left((1 - \omega) \frac{\beta_I}{m_I} + \omega \frac{\beta_A}{m_A} \right). \quad (15)$$

Proof: At $\mathcal{E}^0$,

$$N^* = S^* + V^* = \frac{\Lambda_{\vartheta}(\omega + \delta_{\vartheta}) + \Lambda_{\vartheta}\tau}{\delta_{\vartheta}(\tau + \omega + \delta_{\vartheta})} = \frac{\Lambda_{\vartheta}}{\delta_{\vartheta}}.$$

Hence

$$\alpha_{\vartheta} := (1 - \chi) \frac{S^* + (1 - \rho)V^*}{N^*} = (1 - \chi) \frac{\omega + \delta_{\vartheta} + (1 - \rho)\tau}{\tau + \omega + \delta_{\vartheta}}.$$

Applying the next-generation method to the transmission block $Z = (E, I, A)^{\top}$ gives

$$\tilde{F} = \begin{pmatrix} 0 & \alpha_{\vartheta}\beta_I & \alpha_{\vartheta}\beta_A \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tilde{V} = \begin{pmatrix} \sigma_{\vartheta} + \delta_{\vartheta} & 0 & 0 \\ -(1 - \omega)\sigma_{\vartheta} & m_I & 0 \\ -\omega\sigma_{\vartheta} & 0 & m_A \end{pmatrix},$$

since Q and H do not generate new infections. A direct computation yields

$$\rho(\tilde{F}\tilde{V}^{-1}) = \alpha_{\vartheta} \frac{\sigma_{\vartheta}}{\sigma_{\vartheta} + \delta_{\vartheta}} \left((1 - \omega) \frac{\beta_I}{m_I} + \omega \frac{\beta_A}{m_A} \right),$$

which is (15). Biologically, Eq. (15) decomposes the initial invasion potential into four factors: the population-wide mitigation factor $(1 - \chi)$, the susceptible/vaccinated availability factor, the probability of leaving the exposed class before natural removal, and the weighted infectious contribution of symptomatic and asymptomatic individuals. The quarantined and hospitalized classes do not appear in the new-infection block

because they do not generate new infections in the present model; instead, they influence downstream burden and removal dynamics. $\square$

Remark 4: For the original nonautonomous system, we use

$$\mathcal{R}_e(t) = e^{-kt} \mathcal{R}_0$$

as a transmission indicator. It is not the classical effective reproduction number, since it does not include the current susceptible fraction $S(t)/N(t)$.

Theorem 3: For the frozen system, the disease-free equilibrium $\mathcal{E}^0$ is locally asymptotically stable if $\mathcal{R}_0 < 1$, and unstable if $\mathcal{R}_0 > 1$.

Proof: At $\mathcal{E}^0$, the linearization splits into the (S, V) -block

$$A_0 = \begin{pmatrix} -(\tau + \delta_\vartheta) & \bar{\omega} \\ \tau & -(\bar{\omega} + \delta_\vartheta) \end{pmatrix},$$

the infected block J_{inf} on (E, I, A, Q, H) , and the scalar block $-(\xi_\vartheta + \delta_\vartheta)$ for R . Thus the spectrum is

$$\sigma(A_0) \cup \sigma(J_{\text{inf}}) \cup \{-(\xi_\vartheta + \delta_\vartheta)\}.$$

The matrix A_0 has negative trace and positive determinant, so its eigenvalues have negative real parts. Also,

$$-(\xi_\vartheta + \delta_\vartheta) < 0.$$

By Proposition 4 and the next-generation theorem,

$$\mathcal{R}_0 < 1 \iff s(J_{\text{inf}}) < 0, \quad \mathcal{R}_0 > 1 \iff s(J_{\text{inf}}) > 0,$$

see [22]. Therefore, if $\mathcal{R}_0 < 1$, all eigenvalues of the Jacobian have negative real parts, and local asymptotic stability follows from Matignon's criterion [24] when $0 < \vartheta < 1$, and from classical linearization when $\vartheta = 1$. If $\mathcal{R}_0 > 1$, then J_{inf} is unstable, hence so is $\mathcal{E}^0$. $\square$

For the original system, define the conservative threshold

$$\mathcal{R}_c = (1 - \chi) \frac{\sigma_\vartheta}{\sigma_\vartheta + \delta_\vartheta} \left((1 - \omega) \frac{\beta_I}{m_I} + \omega \frac{\beta_A}{m_A} \right). \quad (16)$$

This is obtained by replacing $e^{-kt}[S(t) + (1 - \rho)V(t)]/N(t)$ with its upper bound 1. Moreover,

$$\mathcal{R}_0 = \frac{\bar{\omega} + \delta_\vartheta + (1 - \rho)\tau}{\tau + \bar{\omega} + \delta_\vartheta} \mathcal{R}_c \leq \mathcal{R}_c,$$

so $\mathcal{R}_c < 1$ is more restrictive than $\mathcal{R}_0 < 1$.

Theorem 4: If $\mathcal{R}_c < 1$, then every global solution of the system (1) satisfies

$$E(t), I(t), A(t), Q(t), H(t), R(t) \rightarrow 0, \quad (S(t), V(t)) \rightarrow (S^*, V^*),$$

where (S^*, V^*) is given by (12). Consequently,

$$\text{dist}(X(t), \mathcal{M}^0) \rightarrow 0.$$

Proof: Let $Z = (E, I, A)^\top$. Since $e^{-kt} \leq 1$ and

$$0 \leq \frac{S(t) + (1-\rho)V(t)}{N(t)} \leq 1,$$

the E -equation implies

$${}_0^C D_t^\vartheta E \leq (1-\chi)(\beta_I I + \beta_A A) - (\sigma_\vartheta + \delta_\vartheta)E.$$

Together with the exact equations for I and A ,

$${}_0^C D_t^\vartheta Z(t) \leq BZ(t),$$

where

$$B = \begin{pmatrix} -(\sigma_\vartheta + \delta_\vartheta) & (1-\chi)\beta_I & (1-\chi)\beta_A \\ (1-\omega)\sigma_\vartheta & -m_I & 0 \\ \omega\sigma_\vartheta & 0 & -m_A \end{pmatrix}.$$

Let $\tilde{Z}$ solve

$${}_0^C D_t^\vartheta \tilde{Z} = B\tilde{Z}, \quad \tilde{Z}(0) = Z(0).$$

Write $B = F_c - V_c$, where

$$F_c = \begin{pmatrix} 0 & (1-\chi)\beta_I & (1-\chi)\beta_A \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad V_c = \begin{pmatrix} \sigma_\vartheta + \delta_\vartheta & 0 & 0 \\ -(1-\omega)\sigma_\vartheta & m_I & 0 \\ -\omega\sigma_\vartheta & 0 & m_A \end{pmatrix}.$$

This sign convention is consistent with the decomposition $B = F_c - V_c$ and with the associated next-generation operator. Then $\rho(F_c V_c^{-1}) = \mathcal{R}_c$, so $\mathcal{R}_c < 1$ implies

$$s(B) < 0$$

by the next-generation theorem [25]. Since B is Metzler, the comparison system is cooperative, and the vector-valued fractional comparison principle yields

$$0 \leq Z(t) \leq \tilde{Z}(t), \quad t \geq 0,$$

see, e.g., [21]. The comparison system is asymptotically stable (Matignon for $0 < \vartheta < 1$, classical for $\vartheta = 1$), hence

$$\tilde{Z}(t) \rightarrow 0,$$

and therefore

$$E(t), I(t), A(t) \rightarrow 0.$$

Next,

$${}_0^C D_t^\vartheta Q = q_\vartheta I - m_Q Q, \quad {}_0^C D_t^\vartheta H = h_{I,\vartheta} I + h_{Q,\vartheta} Q - m_H H,$$

and

$${}_0^C D_t^\vartheta R = \gamma_{I,\vartheta} I + \gamma_{A,\vartheta} A + \gamma_{Q,\vartheta} Q + \gamma_{H,\vartheta} H - (\xi_\vartheta + \delta_\vartheta) R.$$

Their homogeneous parts are asymptotically stable and their inputs vanish, so

$$Q(t), H(t), R(t) \rightarrow 0,$$

see, e.g., [21].

Finally, with

$$U(t) := \begin{pmatrix} S(t) - S^* \\ V(t) - V^* \end{pmatrix},$$

the first two equations of the system (1) and the equilibrium identities at (S^*, V^*) give

$${}_0^C D_t^\vartheta U(t) = A_0 U(t) + g(t),$$

where

$$g(t) = \begin{pmatrix} \xi_\vartheta R(t) - \lambda_S(t) S(t) \\ -\lambda_V(t) V(t) \end{pmatrix}.$$

By Proposition 2, $0 \leq S(t), V(t) \leq M_\vartheta$. Since $R(t) \rightarrow 0$ and $\lambda_S(t), \lambda_V(t) \rightarrow 0$ because $I(t), A(t) \rightarrow 0$, we obtain $g(t) \rightarrow 0$. As A_0 has eigenvalues with negative real parts, the same linear fractional asymptotic result gives $U(t) \rightarrow 0$. Thus

$$(S(t), V(t)) \rightarrow (S^*, V^*).$$

$$\text{Since } \mathcal{M}^0 = \{(\mathcal{E}^0, D^*) : D^* \geq 0\},$$

$$\text{dist}(X(t), \mathcal{M}^0) = \|Y(t) - \mathcal{E}^0\| \rightarrow 0,$$

which completes the proof. In plain terms, the proof first bounds the exposed, symptomatic, and asymptomatic subsystem by a cooperative linear fractional comparison system. The condition $\mathcal{R}_c < 1$ makes this comparison system stable. Once E , I , and A decay, the quarantined, hospitalized, and recovered compartments decay through stable linear equations with vanishing inputs, and the susceptible–vaccinated subsystem converges back to the disease-free balance. $\square$

Remark 5: The condition $\mathcal{R}_c < 1$ is conservative, but it yields global extinction for the nonautonomous system.

2.5 Data Window and Observation Map

The numerical study uses the global mpox surveillance dataset from the Our World in Data repository, aggregated at the *World* level [26]. In the present world-level calibration, the total population size is taken as

$$N_{\text{pop}} = 7.95 \times 10^9,$$

which corresponds to the approximate world population over the retained study period and is used only to initialize the aggregate susceptible compartment in a consistent large-population setting. Calibration is restricted to the first global epidemic wave,

$$2022-05-01 \leq t \leq 2022-08-28,$$

giving $N_{\text{obs}} = 120$ daily observations. This window is used as a controlled first-wave mechanistic calibration setting. For the neural validation stage, the same 120-day interval is partitioned into the first 100 days for training and the final 20 days for held-out short-horizon testing; the held-out observations are not used in PINN fitting. Later waves are not pooled into the same calibration window because behavioral response, surveillance intensity, vaccination coverage, and reporting practices may change substantially across epidemic phases; including them without regime-specific recalibration would mix distinct transmission regimes and could obscure the interpretation of the calibrated parameters.

The retained observables are

$$\tilde{C}_{\text{day}}^{\text{obs}}(t_n), \quad \tilde{D}_{\text{day}}^{\text{obs}}(t_n), \quad C_{\text{cum}}^{\text{obs}}(t_n), \quad D_{\text{cum}}^{\text{obs}}(t_n), \quad n = 0, \dots, N_{\text{obs}} - 1,$$

on the uniform grid

$$t_n = n\Delta t, \quad \Delta t = 1 \text{ day}.$$

Here $\tilde{C}_{\text{day}}^{\text{obs}}$ and $\tilde{D}_{\text{day}}^{\text{obs}}$ are 7-day moving-average daily series, while the cumulative case and death series are kept unsmoothed. Thus, the observation layer is designed to connect the latent epidemic states to two incidence-type quantities (daily reported cases and daily reported deaths) and to their cumulative surveillance counterparts.

Since the state vector of the system (1) is only partially observed, calibration is performed through the observation map

$$C_{\text{day}}^{\text{raw}}(t_n) = r_C(1 - \omega)\sigma_{\vartheta}E(t_n), \quad 0 \leq r_C \leq 1,$$

and

$$D_{\text{day}}^{\text{raw}}(t_n) = r_D(v_{I,\vartheta}I(t_n) + v_{Q,\vartheta}Q(t_n) + v_{H,\vartheta}H(t_n)), \quad 0 \leq r_D \leq 1. \quad (17)$$

Here r_C and r_D are effective reporting coefficients that account, in an aggregated way, for under-reporting and partial observation in the reported case and death streams. In particular, r_C links the symptomatic-emergence flow $(1 - \omega)\sigma_{\vartheta}E$ to reported daily cases, whereas r_D links the mortality flows from the infectious, quarantined, and hospitalized compartments to reported daily deaths. These coefficients should therefore be interpreted as phenomenological observation parameters rather than direct administrative reporting rates.

Remark 6: For $0 < \vartheta < 1$, $(1 - \omega)\sigma_{\vartheta}E(t_n)$ has units $\text{persons} \cdot \text{day}^{-\vartheta}$, and the same applies to the modeled daily death flow in (17). Since surveillance data are recorded as daily counts, we introduce a dimensional conversion factor s_{ϑ} with units $\text{day}^{\vartheta-1}$, and define

$$C_{\text{day}}^{\text{mod}}(t_n) = s_{\vartheta} C_{\text{day}}^{\text{raw}}(t_n), \quad D_{\text{day}}^{\text{mod}}(t_n) = s_{\vartheta} D_{\text{day}}^{\text{raw}}(t_n).$$

In the present implementation, we take

$$s_{\vartheta} = (\Delta t)^{\vartheta-1}.$$

Since the observation grid uses $\Delta t = 1$ day, this gives $s_9 = 1$ numerically throughout the reported experiments. Accordingly, s_9 is not included in the calibration vector, and r_C and r_D retain the interpretation of dimensionless effective reporting coefficients.

Because the observed daily series are 7-day moving averages, the modeled daily series used in calibration are smoothed in the same way. For $n = 6, \dots, N_{\text{obs}} - 1$, we define

$$\tilde{C}_{\text{day}}^{\text{mod}}(t_n) = \frac{1}{7} \sum_{j=0}^6 C_{\text{day}}^{\text{mod}}(t_{n-j}),$$

and

$$\tilde{D}_{\text{day}}^{\text{mod}}(t_n) = \frac{1}{7} \sum_{j=0}^6 D_{\text{day}}^{\text{mod}}(t_{n-j}).$$

This smoothing is introduced solely to match the preprocessed daily surveillance series used in calibration.

The cumulative modeled outputs are reconstructed at observation level by

$$C_{\text{cum}}^{\text{mod}}(t_0) = C_{\text{cum}}^{\text{obs}}(t_0), \quad D_{\text{cum}}^{\text{mod}}(t_0) = D_{\text{cum}}^{\text{obs}}(t_0),$$

and, for $n \geq 1$,

$$C_{\text{cum}}^{\text{mod}}(t_n) = C_{\text{cum}}^{\text{mod}}(t_{n-1}) + \frac{\Delta t}{2} (C_{\text{day}}^{\text{mod}}(t_n) + C_{\text{day}}^{\text{mod}}(t_{n-1})), \quad (18)$$

$$D_{\text{cum}}^{\text{mod}}(t_n) = D_{\text{cum}}^{\text{mod}}(t_{n-1}) + \frac{\Delta t}{2} (D_{\text{day}}^{\text{mod}}(t_n) + D_{\text{day}}^{\text{mod}}(t_{n-1})). \quad (19)$$

Remark 7: Formulas (18) and (19) are observation-level numerical reconstructions used only to match cumulative surveillance series; they are not fractional integrals of the state trajectory and should not be confused with the internal state variable $D(t)$. In particular, the cumulative observation variables are reconstructed from modeled daily outputs in observation space rather than read directly from the latent state vector.

The inverse problem is therefore posed on

$$\tilde{C}_{\text{day}}^{\text{obs}}, \quad \tilde{D}_{\text{day}}^{\text{obs}}, \quad C_{\text{cum}}^{\text{obs}}, \quad D_{\text{cum}}^{\text{obs}},$$

through

$$\tilde{C}_{\text{day}}^{\text{mod}}, \quad \tilde{D}_{\text{day}}^{\text{mod}}, \quad C_{\text{cum}}^{\text{mod}}, \quad D_{\text{cum}}^{\text{mod}}.$$

No explicit reporting-delay mechanism is introduced in the present observation layer. Thus, the calibration assumes an effective same-grid linkage between latent epidemic flows and reported daily quantities. This is a simplifying assumption that is reasonable for the present aggregated short-horizon study, but it may contribute to residual mismatch in daily counts and should be regarded as a limitation of the current framework.

2.6 Fractional Discretization and Forward Solver

We solve the system (1) on the uniform grid

$$t_n = n\Delta t, \quad n = 0, \dots, N_T, \quad \Delta t = 1 \text{ day},$$

with

$$X_n \approx X(t_n) = (S_n, V_n, E_n, I_n, A_n, Q_n, H_n, R_n, D_n)^\top.$$

For $0 < \vartheta \leq 1$, we use the classical L1 approximation of the Caputo derivative:

$${}_0^C D_t^\vartheta x(t_n) \approx \frac{1}{\Gamma(2-\vartheta)\Delta t^\vartheta} \sum_{j=0}^{n-1} a_j (x_{n-j} - x_{n-j-1}), \quad a_j = (j+1)^{1-\vartheta} - j^{1-\vartheta}.$$

Define

$$\kappa_\vartheta := \frac{1}{\Gamma(2-\vartheta)\Delta t^\vartheta}, \quad \mathcal{M}_n[x] := \sum_{j=1}^{n-1} a_j (x_{n-j} - x_{n-j-1}).$$

Since $a_0 = 1$,

$${}_0^C D_t^\vartheta x(t_n) \approx \kappa_\vartheta [(x_n - x_{n-1}) + \mathcal{M}_n[x]],$$

with $\mathcal{M}_1[x] = 0$.

We adopt the explicit lagged L1 scheme

$$\kappa_\vartheta [(X_n - X_{n-1}) + \mathcal{M}_n[X]] = \mathcal{H}(t_{n-1}, X_{n-1}), \quad n \geq 1,$$

that is,

$$X_n = X_{n-1} - \mathcal{M}_n[X] + \kappa_\vartheta^{-1} \mathcal{H}(t_{n-1}, X_{n-1}),$$

so each step depends only on previously computed values.

The discrete incidence terms are

$$\lambda_{S,n-1} = (1-\chi) e^{-k t_{n-1}} \frac{\beta_I I_{n-1} + \beta_A A_{n-1}}{\max\{N_{n-1}, \varepsilon\}}, \quad \lambda_{V,n-1} = (1-\rho) \lambda_{S,n-1},$$

where

$$N_{n-1} = S_{n-1} + V_{n-1} + E_{n-1} + I_{n-1} + A_{n-1} + Q_{n-1} + H_{n-1} + R_{n-1}, \quad \varepsilon = 10^{-8}.$$

The safeguard ε is used only in implementation. The coefficients m_I, m_A, m_Q, m_H are those in (14).

For $\vartheta = 1$, $a_j = 0$ for $j \geq 1$ and $\kappa_1^{-1} = \Delta t$, so the scheme reduces to forward Euler.

For sufficiently smooth solutions, the L1 approximation of the Caputo derivative has local consistency order [27].

$$O(\Delta t^{2-\vartheta})$$

For fractional problems with weak initial singularities, the convergence of L1-type discretizations may deteriorate, often to

$$O(\Delta t^{\min(2-\vartheta, \vartheta)})$$

under limited regularity assumptions [28]. Since the present solver uses an explicit lagged evaluation of the nonlinear right-hand side, a sharper global convergence order for the full nonlinear scheme is not required

for the present data-assimilation objective; instead, its numerical robustness is assessed empirically by mesh refinement over the finite observation window.

The explicit lagged scheme is used here as a practical forward solver over a short time horizon. In the present computations, with $\Delta t = 1$ day, the quantity

$$\kappa_g^{-1} m_{\max}, \quad m_{\max} := \max\{m_I, m_A, m_Q, m_H\},$$

remains below unity for the retained parameter sets. This is taken as a useful a posteriori numerical indicator, rather than as a general stability theorem for the full memory-dependent scheme.

A direct evaluation of the memory term costs $O(N_T^2)$ operations and $O(N_T)$ memory per state component. Since the observation windows considered here are short, this cost is negligible. The same discrete L1 operator is used in both calibration and PINN construction.

2.7 Calibration, Hidden-State Initialization, and Forecasting

Let

$$\Theta = (\theta_1, \dots, \theta_p)$$

denote the vector of calibrated quantities. In the baseline mechanistic calibration,

$$\Theta = (\beta_{I,\vartheta}, \beta_{A,\vartheta}, \sigma_{\vartheta}, q_{\vartheta}, k, r_C, r_D, \eta_E, \eta_I, \eta_A, \eta_Q, \eta_H, \eta_R),$$

where $\eta_E, \eta_I, \eta_A, \eta_Q, \eta_H, \eta_R$ are initialization multipliers for the hidden compartments. Thus, the estimated quantities consist of transmission, progression, decay, and observation parameters together with a small set of latent-state initialization factors. The remaining model parameters are kept fixed in the mechanistic calibration stage, either because they are structurally prescribed by the model formulation or because they are treated as externally specified inputs in the present short-window study.

To avoid ambiguity between calibrated, fixed, and neural-stage quantities, [Table 1](#) summarizes the role of the main parameter groups used in the computational pipeline. This table makes explicit how the quantities entering the calibration, threshold formulas, and PINN stage are related.

Table 1: Classification and role of the main parameter groups in the calibration and inverse-learning pipeline.

Parameter Group	Symbols	Status	Role in the Pipeline
Transmission and attenuation	$\beta_{I,\vartheta}, \beta_{A,\vartheta}, k$	Mechanistically calibrated	Determine the force of infection and the early-wave time-varying decay $\beta(t) = e^{-kt}$.
Progression and quarantine	$\sigma_{\vartheta}, q_{\vartheta}$	Mechanistically calibrated; selected quantities relearned in PINN	Control the movement from exposure to infection and from infection to quarantine.
Observation/reporting layer	r_C, r_D	Mechanistically calibrated	Map latent epidemic flows to reported daily cases and deaths in the forward calibration.

(Continued)

Table 1 (continued)

Parameter Group	Symbols	Status	Role in the Pipeline
Neural observation layer	$\widehat{c}_C, \widehat{c}_D$	Learned in the PINN stage	Stage-specific effective observation coefficients; analogous to r_C, r_D but not identified with them.
Hidden-state initialization	$\eta_E, \eta_I, \eta_A, \eta_Q, \eta_H, \eta_R$	Mechanistically calibrated auxiliary variables	Provide feasible nonnegative initial values for latent compartments that are not directly observed.
Fixed biological rates	$\gamma_{I,\vartheta}, \gamma_{A,\vartheta}, \gamma_{Q,\vartheta}, \gamma_{H,\vartheta}, h_{I,\vartheta}, h_{Q,\vartheta}, \nu_{I,\vartheta}, \nu_{Q,\vartheta}, \nu_{H,\vartheta}, \delta_\vartheta, \tau, \varpi, \rho, \chi, \omega$	Fixed or externally specified in the retained experiment	Define recovery, hospitalization, mortality, vaccination, protection, mitigation, and asymptomatic fractions used consistently in the model and threshold formulas.

Calibration uses the first $N_{\text{tr}} = 120$ daily observations and the four observed series from [Section 2.5](#). The objective function is

$$J(\Theta) = w_1 E_{\text{dc}} + w_2 E_{\text{dd}} + w_3 E_{\text{cc}} + w_4 E_{\text{cd}}, \quad (20)$$

with $(w_1, w_2, w_3, w_4) = (1.0, 1.0, 0.3, 0.2)$. The larger weights assigned to the daily components emphasize local wave-shape reconstruction, while the cumulative components provide a global consistency constraint over the retained observation window. The four mismatch terms are

$$E_{\text{dc}} = \frac{1}{N_{\text{tr}} - 6} \sum_{n=6}^{N_{\text{tr}}-1} \left[\log(1 + \widetilde{C}_{\text{day}}^{\text{obs}}(t_n)) - \log(1 + \widetilde{C}_{\text{day}}^{\text{mod}}(t_n; \Theta)) \right]^2,$$

$$E_{\text{dd}} = \frac{1}{N_{\text{tr}} - 6} \sum_{n=6}^{N_{\text{tr}}-1} \left[\log(1 + \widetilde{D}_{\text{day}}^{\text{obs}}(t_n)) - \log(1 + \widetilde{D}_{\text{day}}^{\text{mod}}(t_n; \Theta)) \right]^2,$$

$$E_{\text{cc}} = \frac{1}{N_{\text{tr}}} \sum_{n=0}^{N_{\text{tr}}-1} \left[\log(1 + C_{\text{cum}}^{\text{obs}}(t_n)) - \log(1 + C_{\text{cum}}^{\text{mod}}(t_n; \Theta)) \right]^2,$$

and

$$E_{\text{cd}} = \frac{1}{N_{\text{tr}}} \sum_{n=0}^{N_{\text{tr}}-1} \left[\log(1 + D_{\text{cum}}^{\text{obs}}(t_n)) - \log(1 + D_{\text{cum}}^{\text{mod}}(t_n; \Theta)) \right]^2.$$

This calibration objective should be interpreted as a practical inverse-learning criterion on partially observed epidemic outputs rather than as a claim of full structural identifiability of all estimated quantities. In particular, because the observation window is relatively short and several epidemiological compartments are latent, correlations among some fitted parameters are possible, especially between transmission-related coefficients, reporting coefficients, and hidden-state initialization factors.

The latent initial state is initialized heuristically from the first daily observations. Let

$$\widetilde{C}_0 := \max\{\widetilde{C}_{\text{day}}^{\text{obs}}(t_0), 1\}, \quad \widetilde{D}_0 := \max\{\widetilde{D}_{\text{day}}^{\text{obs}}(t_0), 0\}, \quad D_0 := D_{\text{cum}}^{\text{obs}}(t_0).$$

The baseline initialization is

$$\begin{aligned} V_0 &= 0, & E_0 &= \eta_E \tilde{C}_0, & I_0 &= \eta_I \tilde{C}_0, & A_0 &= \eta_A \tilde{C}_0, \\ Q_0 &= \eta_Q \tilde{C}_0, & H_0 &= \max\{1, \eta_H \tilde{D}_0 + 1\}, & R_0 &= \eta_R \tilde{C}_0, \end{aligned}$$

and

$$S_0 = N_{\text{pop}} - (V_0 + E_0 + I_0 + A_0 + Q_0 + H_0 + R_0 + D_0), \quad S_0 > 0,$$

where N_{pop} denotes the total population size used in the world-level calibration.

Remark 8: *This initialization is a feasible nonnegative heuristic. It is not identifiable from surveillance data alone and is not interpreted as a direct epidemiological estimate of the hidden compartments. Rather, the multipliers $\eta_E, \eta_I, \eta_A, \eta_Q, \eta_H, \eta_R$ are introduced to provide sufficient flexibility for initializing the latent epidemic state in a partially observed setting. They should therefore be viewed as auxiliary inverse-calibration variables, not as independently validated epidemiological observables.*

Accordingly, the present calibration protocol is designed to recover a dynamically consistent and observation-compatible latent trajectory on the retained study window. It is not intended to imply that all fitted parameters are uniquely recoverable from the available surveillance series, nor that the inferred hidden initial state is epidemiologically identifiable in isolation.

Integer-order benchmark protocol.

To isolate the contribution of the fractional operator, the calibrated fractional configuration is compared with the classical limit obtained by setting $\vartheta = 1$ in the same SEVIAQHRD equations and observation map. All remaining quantities are held at the calibrated fractional values. The retained trajectories generated by the same numerical solver are used for both configurations. The comparison therefore changes only the order of the governing operator and uses the same 120-day observation window and the same four reported outputs. Days 100–119 are highlighted to coincide with the held-out segment used in the PINN diagnostics. This is a controlled operator-limit comparison, not an independently re-optimized integer-order forecast, and it is interpreted accordingly.

Observation-level structural identifiability assessment.

Structural identifiability depends on the selected outputs and on which states are observed [29,30]. In the present model, the surveillance streams observe neither the infectious states nor the hidden initial conditions directly. The parameter dependence that is explicit in the observation and transmission layers can be summarized by

$$\Psi_\beta(t) = e^{-kt}(\beta_I I(t) + \beta_A A(t)), \quad \Psi_C(t) = r_C(1 - \omega)\sigma_\vartheta E(t),$$

and

$$\Psi_D(t) = r_D(v_{I,\vartheta} I(t) + v_{Q,\vartheta} Q(t) + v_{H,\vartheta} H(t)).$$

Thus, aggregate case and death observations directly constrain these composite quantities, whereas separation of their individual factors requires information about the latent trajectories. The six initialization multipliers enter only through the unobserved initial state and remain nuisance quantities without additional compartment-level measurements.

Proposition 5: *At the observation-map level, the reporting coefficients cannot be separated from their corresponding latent burdens using the four retained surveillance outputs alone.*

Proof: Let $M_D(t) = \nu_{I,\vartheta}I(t) + \nu_{Q,\vartheta}Q(t) + \nu_{H,\vartheta}H(t)$. For arbitrary constants $a_C, a_D > 0$, the transformations

$$(r_C, E) \mapsto (a_C r_C, E/a_C), \quad (r_D, M_D) \mapsto (a_D r_D, M_D/a_D)$$

leave $\Psi_C(t)$ and $\Psi_D(t)$ unchanged. The cumulative outputs are deterministic accumulations of the corresponding daily outputs and therefore do not remove these observation-level invariances. Hence the observation map is non-injective in these factors without additional latent-state measurements. The coupled state equations may restrict the admissible trajectories, but they do not convert the four aggregate streams into direct measurements of the individual factors. $\square$

Fig. 3 and Table 2 summarize the resulting observation-level assessment.

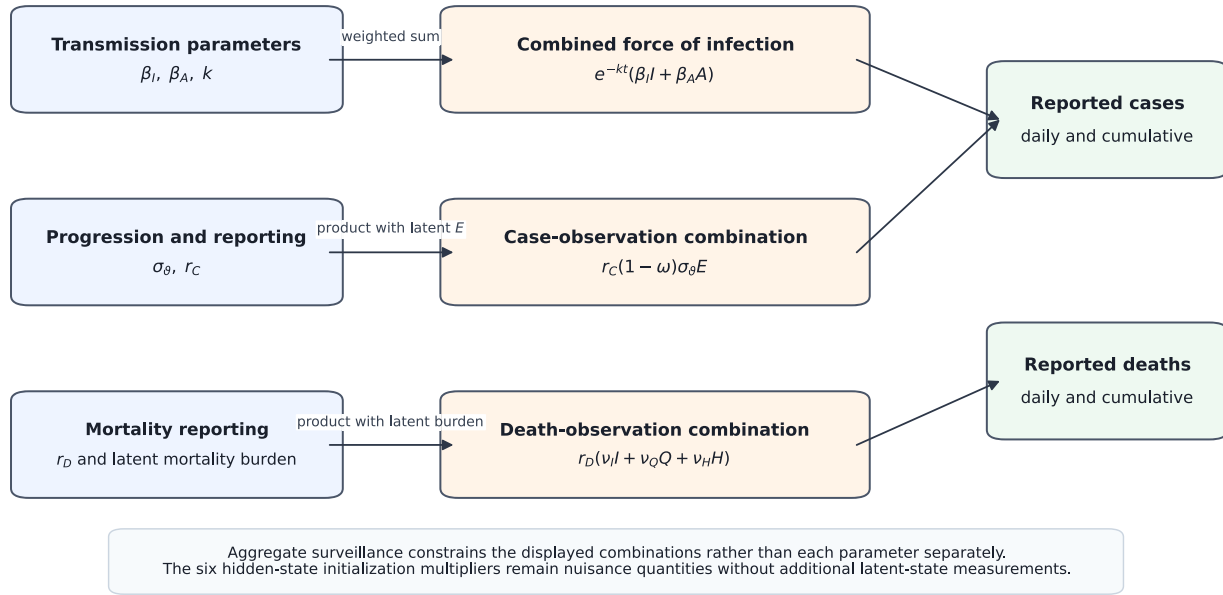


Figure 3: Observation-level structural identifiability map for the calibrated quantities. The available daily and cumulative surveillance outputs constrain the displayed parameter–state combinations rather than every individual factor. The diagram is an exact assessment of the explicit observation map and transmission combination; it is not presented as a global differential-algebra proof for the full nonautonomous fractional system.

Table 2: Observation-level identifiability assessment for the principal calibrated groups.

Quantity or Group	Directly Constrained Combination	Interpretation
β_I, β_A, k	$\Psi_\beta(t) = e^{-kt}(\beta_I I + \beta_A A)$	Conditional on the latent infectious trajectories; the three factors are not separately determined by aggregate outputs alone.
σ_ϑ, r_C	$\Psi_C(t) = r_C(1 - \omega)\sigma_\vartheta E$	Progression and case reporting are coupled with the unobserved exposed state.
r_D and mortality burden	$\Psi_D(t) = r_D(\nu_I I + \nu_Q Q + \nu_H H)$	The data determine a scaled latent mortality flow rather than r_D independently.

(Continued)

Table 2 (continued)

Quantity or Group	Directly Constrained Combination	Interpretation
q_θ	Indirectly through I, Q, H dynamics	Recoverability is conditional on the assumed compartmental dynamics and fixed removal parameters.
$\eta_E, \eta_I, \eta_A, \eta_Q, \eta_H, \eta_R$	Hidden initial state	Auxiliary nuisance variables; not separately observable from the four surveillance streams.

For each trial Θ , the forward solver of [Section 2.6](#) produces

$$\tilde{C}_{\text{day}}^{\text{mod}}, \quad \tilde{D}_{\text{day}}^{\text{mod}}, \quad C_{\text{cum}}^{\text{mod}}, \quad D_{\text{cum}}^{\text{mod}},$$

which are compared with the observed series on the calibration window.

Optimization is performed in two stages: differential evolution followed by L-BFGS-B. In the baseline implementation, differential evolution uses the `best1bin` strategy with population size 12, mutation range (0.5, 1.0), recombination 0.7, tolerance 10^{-3} , and 50 outer iterations; the resulting point is then refined by L-BFGS-B with 600 maximum iterations. This two-stage strategy is used to balance global exploration and local refinement in a nonconvex calibration landscape.

Forecasting is evaluated out of sample by propagating the calibrated mechanistic model forward with the calibrated parameter vector held fixed. We use the root mean square error (RMSE), mean absolute error (MAE), and relative ℓ_2 error as observation-level accuracy indicators on a given index set $\mathcal{I}$. Prediction errors are measured by

$$\text{RMSE}(\mathcal{I}) = \left(\frac{1}{|\mathcal{I}|} \sum_{n \in \mathcal{I}} (Y_n^{\text{mod}} - Y_n^{\text{obs}})^2 \right)^{1/2},$$

$$\text{MAE}(\mathcal{I}) = \frac{1}{|\mathcal{I}|} \sum_{n \in \mathcal{I}} |Y_n^{\text{mod}} - Y_n^{\text{obs}}|,$$

and

$$\text{Rel-}\ell_2(\mathcal{I}) = \frac{(\sum_{n \in \mathcal{I}} (Y_n^{\text{mod}} - Y_n^{\text{obs}})^2)^{1/2}}{(\sum_{n \in \mathcal{I}} (Y_n^{\text{obs}})^2)^{1/2}}.$$

These metrics are computed separately for smoothed daily cases, smoothed daily deaths, cumulative cases, and cumulative deaths. Because the same observation window is also relatively short, these forecasting diagnostics should be interpreted as short-horizon validation indicators rather than as evidence of long-range predictive certainty.

Finally, solver consistency is checked a posteriori by re-integrating the calibrated system on a refined time grid and measuring the discrepancy with the reference run on the observation window. This provides a numerical robustness check for the retained discretization over the calibration horizon, although it does not constitute a full identifiability or global convergence theorem for the inverse problem.

3 Mechanistic Numerical Results and Sensitivity Analysis

3.1 Baseline Fit and Latent-Trajectory Reconstruction

The baseline mechanistic calibration uses the forward solver of [Section 2.6](#) and the observation map of [Section 2.5](#) over the 120-day first-wave window

$$2022-05-01 \leq t \leq 2022-08-28, \quad \vartheta = 0.90.$$

This baseline calibration serves as the reference mechanistic configuration for the subsequent sensitivity and inverse-learning analyses. In particular, it provides the nominal observation-compatible trajectory against which the effects of varying selected parameters and the fractional order are examined. [Fig. 4](#) displays the fit to daily and cumulative observations, while [Fig. 5](#) shows the corresponding latent trajectories. This calibration serves as the reference configuration for the sensitivity analyses below.

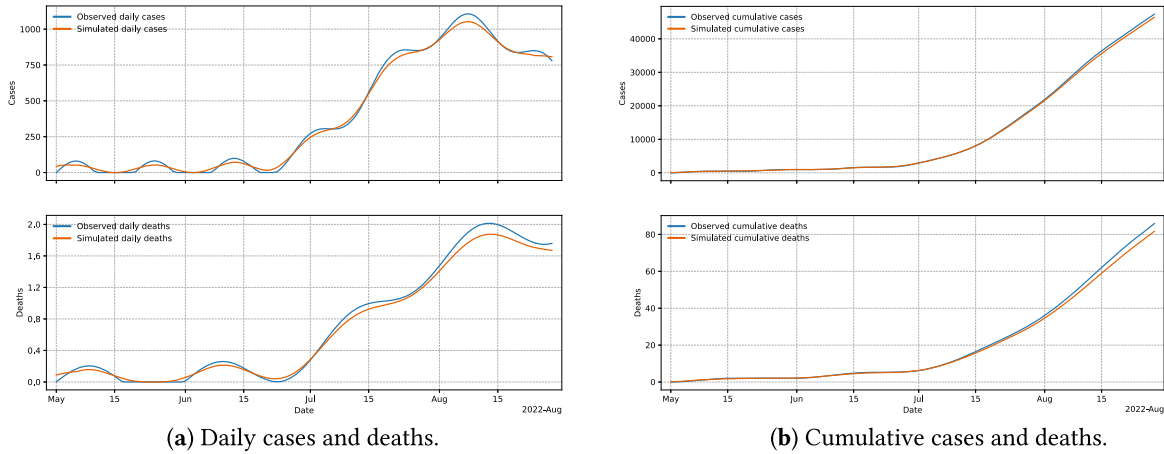


Figure 4: Baseline mechanistic fit on the first-wave window (2022-05-01 to 2022-08-28). Solid curves denote model outputs reconstructed through the observation layer, while the observed curves denote the preprocessed surveillance series. The panels show that the calibrated system captures the main daily-wave profile and cumulative growth trend used for subsequent sensitivity and PINN analyses.

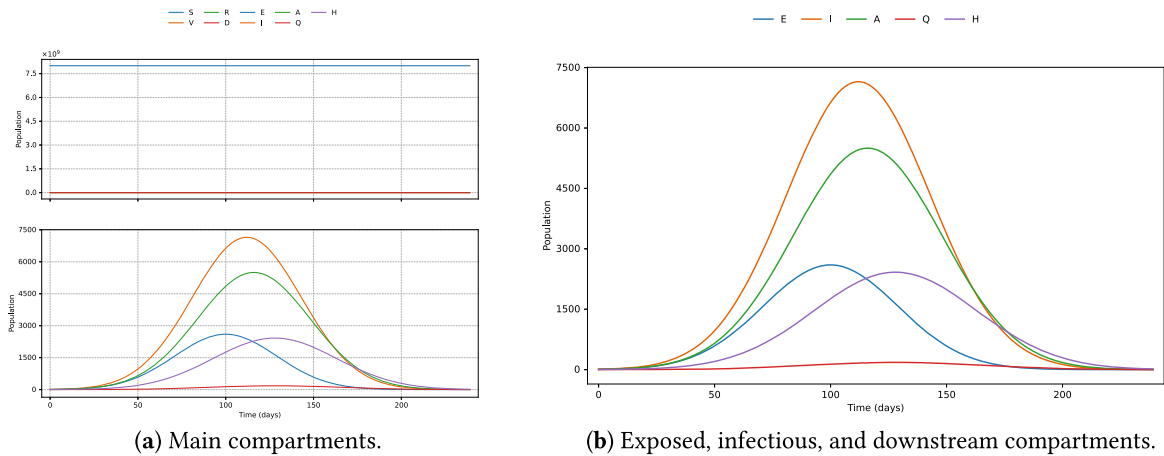


Figure 5: Latent trajectories for the baseline mechanistic calibration. These curves represent model-inferred hidden states rather than direct observations. They illustrate the hidden exposed, infectious, quarantine, hospitalization, recovery, and mortality pathways that are dynamically compatible with the reported case and death observations.

For this calibrated baseline, the threshold quantities introduced in [Section 2.4](#) should be interpreted as summary indicators of the fitted transmission regime. More precisely, $\mathcal{R}_0$ corresponds to the frozen invasion system at the initial time, whereas $\mathcal{R}_c$ is a more restrictive sufficient threshold for extinction in the original nonautonomous model. In the present data-driven setting, these quantities are used primarily to connect the analytical threshold structure with the calibrated mechanistic regime, rather than as standalone empirical descriptors of the full observed wave.

The fit in [Fig. 4](#) shows that the mechanistic calibration reproduces the main cumulative growth pattern and captures the dominant shape of the smoothed daily reported series over the retained first-wave window. At the same time, the latent trajectories in [Fig. 5](#) should be interpreted as reconstructed hidden epidemic states that are dynamically consistent with the calibrated model and observation layer, not as directly observed epidemiological time series.

Compared with simpler epidemic formulations that primarily target aggregate incidence fitting, the present SEVIAQHRD structure is designed to support a broader inverse-learning task for partially observed mpox dynamics. In particular, the inclusion of vaccination, asymptomatic infection, quarantine, hospitalization, waning immunity, and an explicit observation layer makes it possible to reconstruct multiple hidden epidemic pathways while preserving consistency with reported daily and cumulative surveillance quantities. Accordingly, the purpose of the present model is not to claim universal predictive superiority over every lower-dimensional baseline, but rather to provide an interpretable mechanistic setting in which threshold analysis, latent-state reconstruction, observation-level calibration, and discrete physics-informed inference can be carried out simultaneously. This is the main added value of the richer compartmental structure retained here. Simpler epidemic baselines may remain useful for compact description and low-dimensional trend fitting, but they do not generally offer the same level of compartmental interpretability for hidden-state recovery under the observation layer adopted in the present study.

3.2 Controlled Fractional vs. Integer-Order Comparison

[Fig. 6](#) compares the calibrated fractional configuration ($\vartheta = 0.90$) with the classical limit ($\vartheta = 1$) under the benchmark protocol defined above. The same parameterization and observation map are retained, so the comparison isolates the effect of replacing the L1 fractional dynamics by their integer-order limit. The shaded interval corresponds to days 100–119, which are also the PINN held-out segment.

The observation-level errors for the controlled comparison are reported in [Table 3](#).

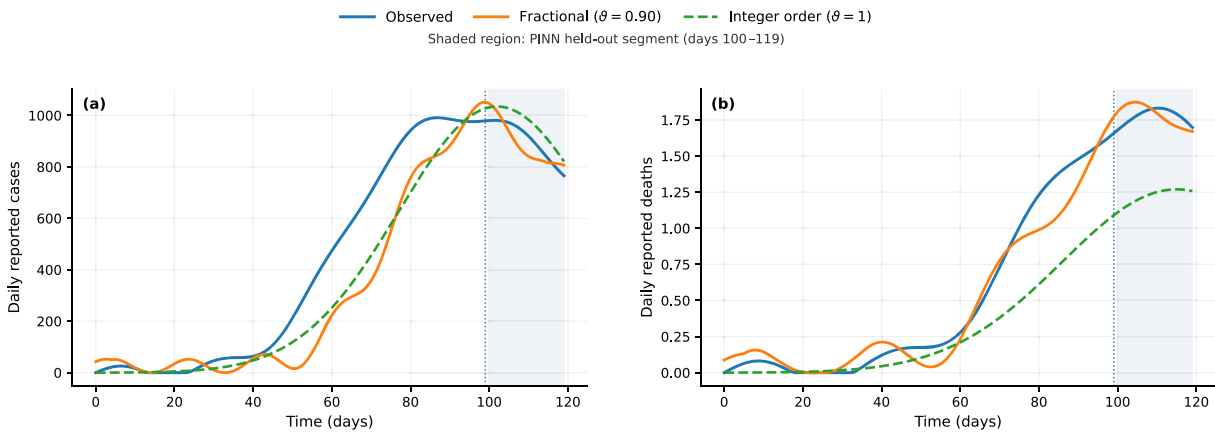


Figure 6: (Continued)

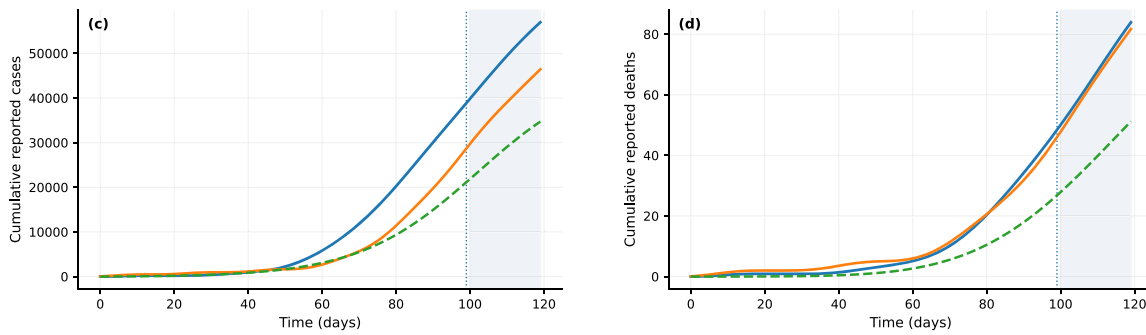


Figure 6: Controlled comparison of the calibrated fractional configuration and its integer-order limit. Panels show (a) daily reported cases, (b) daily reported deaths, (c) cumulative reported cases, and (d) cumulative reported deaths. The shaded interval marks days 100–119, matching the PINN held-out segment. The integer-order curve is obtained by setting $\vartheta = 1$ while retaining the calibrated parameterization; the figure therefore evaluates the operator limit rather than an independently re-optimized model.

Table 3: Observation-level errors over days 100–119 for the controlled fractional and integer-order configurations.

Configuration	Output	RMSE	MAE	Rel.- ℓ_2	R^2
Fractional, $\vartheta = 0.90$	Daily cases	50.2657	43.8942	0.0554	0.5271
Fractional, $\vartheta = 0.90$	Daily deaths	0.0835	0.0751	0.0470	-2.0207
Fractional, $\vartheta = 0.90$	Cumulative cases	10,341.7921	10,339.0110	0.2108	-2.9110
Fractional, $\vartheta = 0.90$	Cumulative deaths	1.8982	1.8833	0.0280	0.9665
Integer order, $\vartheta = 1$	Daily cases	61.8509	61.4441	0.0682	0.2839
Integer order, $\vartheta = 1$	Daily deaths	0.5541	0.5524	0.3119	-131.9652
Integer order, $\vartheta = 1$	Cumulative cases	20,268.2129	20,226.8780	0.4132	-14.0221
Integer order, $\vartheta = 1$	Cumulative deaths	27.8561	27.6650	0.4109	-6.2216

As summarized in Table 3, the fractional configuration has lower RMSE, MAE, and relative- ℓ_2 values for all four outputs in this controlled comparison. The coefficient of determination is unstable for daily deaths because the highlighted segment has very small magnitude and limited variation; absolute and relative errors are therefore more informative for that output. Since the integer-order configuration was not independently re-optimized, these results support the role of the fractional operator within the retained calibration but do not establish universal forecasting superiority over all integer-order models.

3.3 Numerical Support for Threshold Interpretation

To connect the analytical thresholds of Section 2.4 with the numerical experiments, we interpret the baseline and sensitivity simulations as consistency checks for the threshold structure. The quantities $\mathcal{R}_0$ and $\mathcal{R}_c$ are not fitted independently from the observation series; rather, they summarize the transmission regime implied by the calibrated parameters through Eqs. (15) and (16). The sensitivity analyses presented below provide numerical support for this interpretation: increasing transmission-related coefficients increases the reconstructed epidemic burden, whereas increasing mitigation, vaccination protection, or removal-related effects reduces the infectious load and reported burden. These observations are consistent with the threshold formulas, in which symptomatic and asymptomatic transmission terms increase $\mathcal{R}_0$ and $\mathcal{R}_c$, while mitigation and removal mechanisms act in the opposite direction.

This numerical evidence should be read as a threshold-consistency validation rather than as a separate proof. The proof of local threshold behavior for the frozen system and the conservative global extinction

condition for the nonautonomous system are given analytically in Theorems 3 and 4. The numerical contribution of this subsection is to show that the observed parametric response of the calibrated model follows the same qualitative transition logic: transmission amplification pushes the system toward larger epidemic burden, while sufficient control and removal effects move the reconstructed dynamics toward decay.

To make the transition across the threshold explicit without introducing an additional recalibration layer, we also use a controlled transmission-rescaling check. Keeping all non-transmission parameters fixed at their calibrated values, we replace

$$\beta_I \mapsto \eta \beta_I, \quad \beta_A \mapsto \eta \beta_A, \quad \eta > 0.$$

Under this perturbation, the conservative threshold satisfies

$$\mathcal{R}_c(\eta) = \eta \mathcal{R}_c(1),$$

while the denominator terms m_I and m_A , the progression factor $\sigma_{\mathfrak{g}}/(\sigma_{\mathfrak{g}} + \delta_{\mathfrak{g}})$, and the mitigation factor $(1 - \chi)$ remain unchanged. Thus, the following three normalized threshold regimes are obtained directly from Eq. (16) and are summarized in Table 4.

Table 4: Normalized threshold-crossing check obtained by rescaling the symptomatic and asymptomatic transmission contributions while keeping all other calibrated quantities fixed.

Regime	Transmission Scale	Resulting Threshold	Expected Dynamical Interpretation
Subcritical	$\eta = 0.75/\mathcal{R}_c(1)$	$\mathcal{R}_c(\eta) = 0.75 < 1$	Decay of the infected subsystem and movement toward the disease-free manifold.
Near-threshold	$\eta = 1/\mathcal{R}_c(1)$	$\mathcal{R}_c(\eta) = 1$	Boundary regime separating decay from amplification in the conservative comparison system.
Supercritical	$\eta = 1.25/\mathcal{R}_c(1)$	$\mathcal{R}_c(\eta) = 1.25 > 1$	Amplification of transmission-driven burden relative to the subcritical regime.

This rescaling check isolates the threshold mechanism from the parameter-estimation step. It is therefore a direct numerical-design counterpart of the analytical comparison argument: changing only the transmission scale moves the system from a conservative subcritical regime to a supercritical regime, while preserving the same observation map, fractional order, and removal structure.

3.4 Effects of the Fractional Order

In the baseline mechanistic experiments, the fractional order is fixed at $\mathfrak{g} = 0.90$ as a representative memory-capable configuration on the retained observation window. This value is used as a central profile point between the stronger-memory cases $\mathfrak{g} < 0.90$ and the integer-order limit $\mathfrak{g} = 1$, allowing the sensitivity analysis to assess both directions around the baseline. This choice is not treated as a uniquely identifiable empirical constant. Rather, it is used as a central baseline within a local fractional-order profile designed to examine how the reconstructed wave responds to changes in memory strength. Jointly estimating $\mathfrak{g}$ together with the full epidemiological parameter vector would substantially enlarge the inverse problem and increase parameter-correlation effects, especially because only the first 120 daily observations are used. Accordingly, the present analysis treats $\mathfrak{g} = 0.90$ as a justified baseline supported by sensitivity diagnostics, while explicitly avoiding the stronger claim that the fractional order is structurally identifiable from the available surveillance data alone.

The fractional order ϑ is varied while all other parameters remain fixed at their baseline values. Fig. 7 shows how this variation affects cumulative reported cases, daily reported cases, daily reported deaths, and the timing and magnitude of the infectious peak. Over the retained time horizon, the results indicate that ϑ influences both the shape and timing of the reconstructed epidemic wave, which supports the practical relevance of the fractional formulation in the present modeling framework. At the same time, because the calibration window is relatively short, these numerical experiments are intended to demonstrate sensitivity to hereditary dynamics on the available horizon rather than to establish a definitive empirical characterization of long-range memory effects.

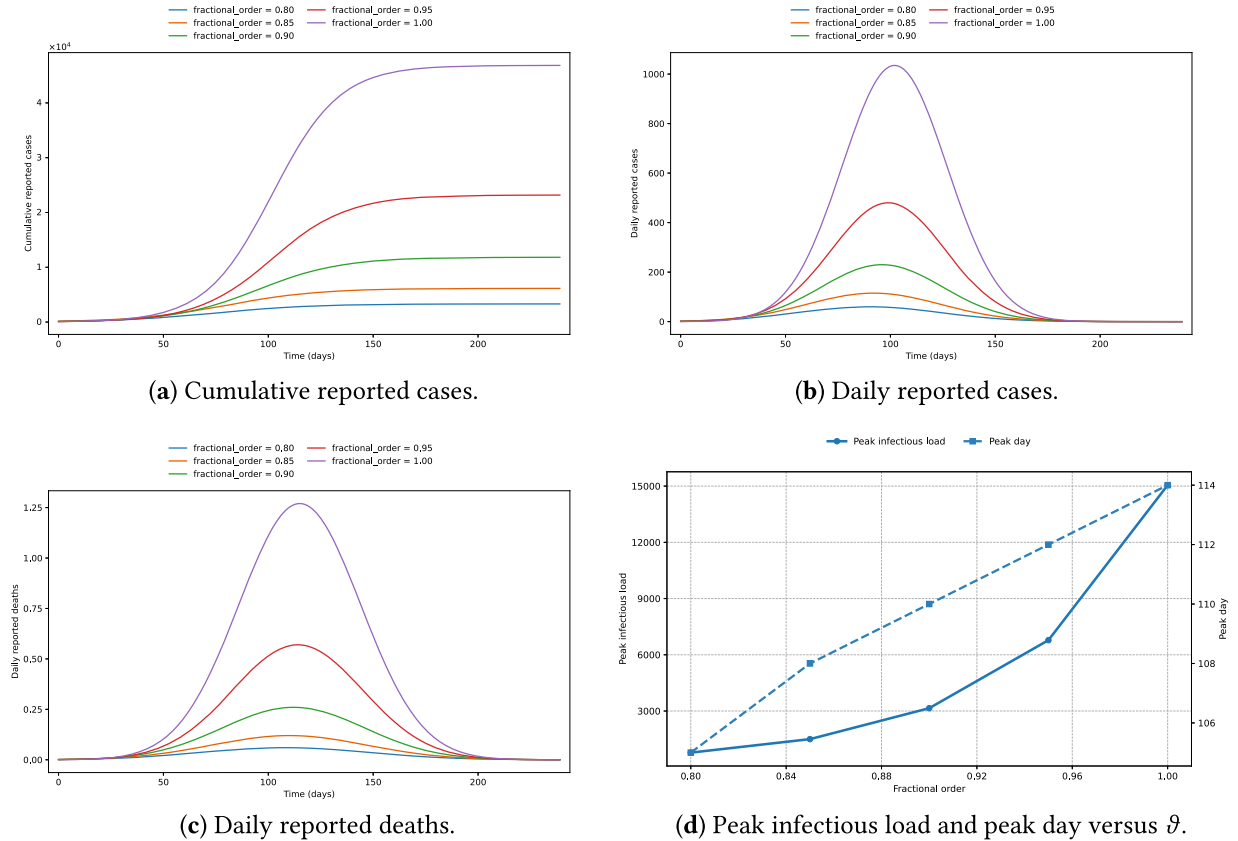


Figure 7: Sensitivity of the epidemic wave to the fractional order ϑ . The panels illustrate how changing the memory parameter affects cumulative burden, daily cases, daily deaths, and the timing and magnitude of the infectious peak. This supports the use of $\vartheta = 0.90$ as a baseline memory-capable configuration while showing that the fractional order influences the reconstructed wave shape.

3.5 Sensitivity to Transmission, Control, and Secondary Parameters

We vary one parameter at a time around the baseline calibration and then examine two selected pairwise interactions. This sensitivity analysis also complements the threshold discussion in Section 2.4. In particular, several of the parameters varied below enter directly into the frozen-system reproduction number R_0 or the conservative threshold R_c , while others affect the epidemic burden indirectly through the nonautonomous transmission dynamics and the observation layer. Thus, although no separate symbolic sensitivity index for R_0 or R_c is introduced here, Figs. 8–10 illustrate, in sequential order, how transmission, progression, quarantine, hospitalization, waning, and pairwise interaction mechanisms influence the fitted epidemic regime around the calibrated baseline.

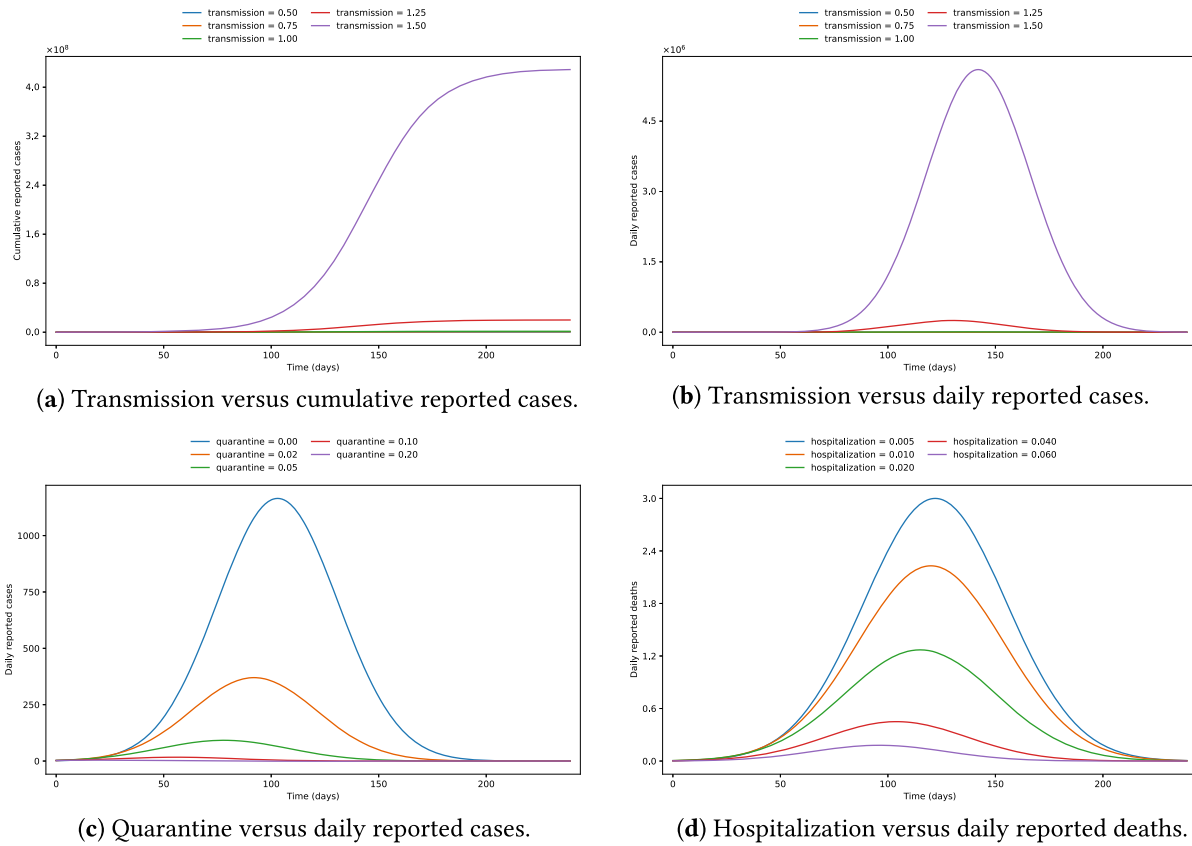


Figure 8: One-parameter sensitivity with respect to transmission and control parameters. Parameters entering the transmission and threshold expressions are varied around the calibrated baseline to show how amplification of transmission increases the epidemic burden, whereas mitigation and vaccination-related effects reduce the reconstructed reported burden.

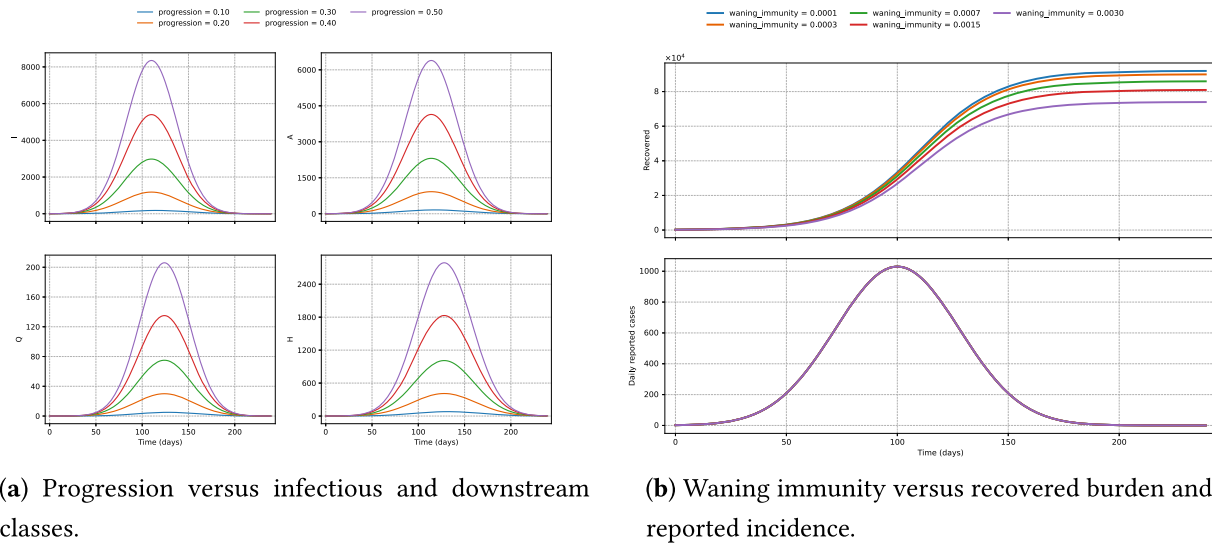


Figure 9: Secondary one-parameter sensitivity analyses. These panels quantify how removal, hospitalization, mortality, and waning-immunity mechanisms reshape the hidden epidemic burden and observation-level outputs, thereby complementing the threshold and calibration interpretations.

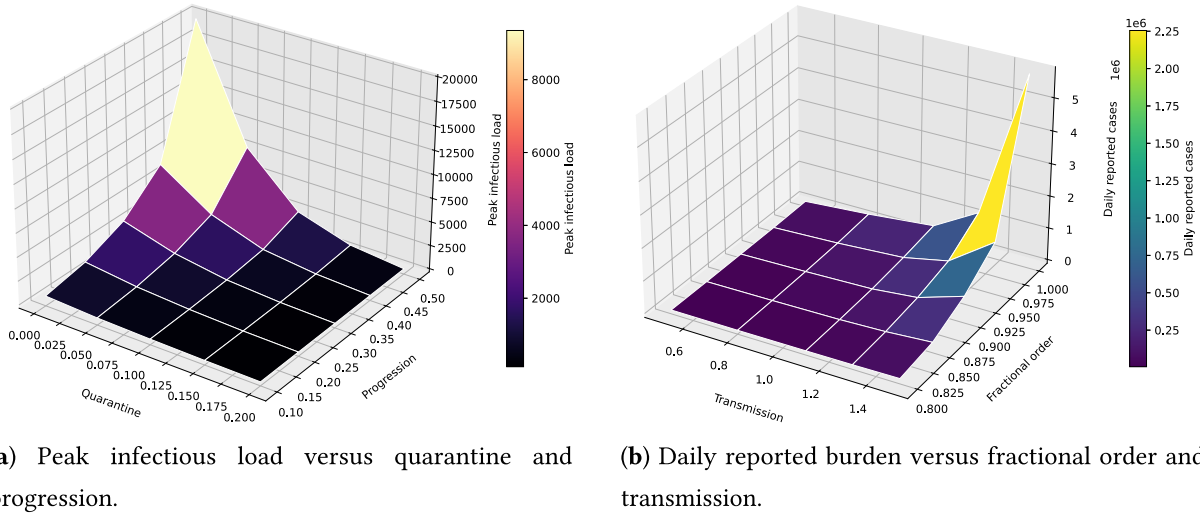


Figure 10: Selected pairwise interaction surfaces. Panel (a) shows the joint effect of quarantine and progression on peak infectious load, while panel (b) shows how fractional memory and transmission interact to shape the daily reported burden. These surfaces provide numerical support for the combined influence of transmission, removal, and memory effects.

4 PINN-Based Parameter Estimation and Forecasting

4.1 PINN Formulation and Inverse Learning Setup

To complement the mechanistic calibration stage, we employ a discrete physics-informed neural inverse-learning framework to infer latent SEVIAQHRD trajectories and selected trainable coefficients from the reported mpox time series. Throughout this section, we retain the term PINN in the broad physics-informed sense, while emphasizing that the governing fractional dynamics are enforced through a discrete residual based on the L1 approximation of the Caputo derivative rather than through direct automatic differentiation of the Caputo operator. The same processed world-level dataset and observation layer as in the mechanistic stage are used.

Let

$$\widehat{X}(t; \psi) = (\widehat{S}, \widehat{V}, \widehat{E}, \widehat{I}, \widehat{A}, \widehat{Q}, \widehat{H}, \widehat{R}, \widehat{D})^\top(t; \psi)$$

denote the neural approximation of the state vector, where ψ collects the network weights and biases. The state network is a fully connected multilayer perceptron (MLP) with one scalar input, three hidden layers of width 128, tanh activations, and positivity-preserving output transforms. Consistently with the diagnostic panels reported below, the trainable coefficient vector is

$$\widehat{\Theta} = (\widehat{\beta}_{I,\vartheta}, \widehat{\sigma}_\vartheta, \widehat{q}_\vartheta, \widehat{\xi}_\vartheta, \widehat{c}_C, \widehat{c}_D).$$

The remaining epidemiological parameters, including the asymptomatic-transmission contribution, are kept fixed from the mechanistic baseline in the neural stage. Thus, the neural stage does not attempt to retrain the full mechanistic parameter set; instead, it focuses on a reduced set of influential transmission, transition, waning-immunity, and observation coefficients while learning a latent state trajectory constrained by the observation model and the governing fractional dynamics.

Fig. 11 shows the state network.

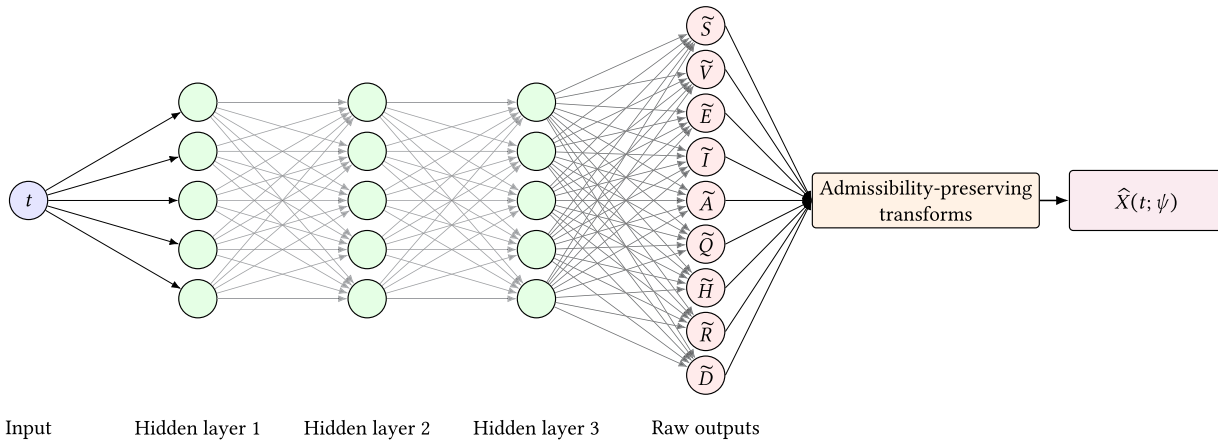


Figure 11: Neural architecture used to approximate the latent SEVIAQHRD state vector. A scalar time input is propagated through three fully connected hidden layers and mapped to the nine state components through admissibility-preserving output transforms.

The neural model is trained on the same daily grid

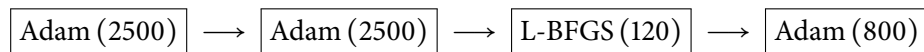
$$t_n = n\Delta t, \quad \Delta t = 1 \text{ day},$$

using the same discrete L1 approximation of the fractional derivative as the mechanistic solver. More precisely, the physics term in the loss is constructed from the discrete residual obtained by substituting the neural trajectory into the L1-based fractional time-stepping relation. Hence the mechanistic and neural stages are numerically aligned through a common discrete fractional operator.

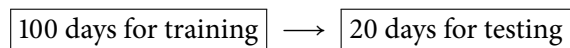
The observation layer follows the same computational model as in [Section 2.5](#): daily reported cases are reconstructed from the symptomatic emergence flow, daily reported deaths from the mortality flows of the infectious, quarantined, and hospitalized classes, and cumulative outputs by discrete accumulation. Daily observations use 7-day moving averages, whereas cumulative cases and deaths are unsmoothed. This shared observation layer makes the neural stage directly comparable with the mechanistic calibration and ensures that both stages are fitted against the same reported quantities.

The loss is hybrid and stage-wise. Besides the physics and initial-condition losses, it includes a solver-tracking loss against the mechanistic reference trajectory, a focus-state loss on selected compartments, a monotonicity penalty for the mortality stock, and a smoothness penalty on key trajectories. The data misfit is evaluated in logarithmic scale on daily and cumulative cases and deaths. The role of this composite construction is to balance four objectives simultaneously: agreement with reported surveillance outputs, consistency with the discrete fractional dynamics, numerical regularity of the inferred latent states, and controlled proximity to the mechanistic reference solution during training.

Training proceeds in four stages:



The two Adam stages provide initial stochastic exploration and coarse stabilization, the L-BFGS phase improves deterministic refinement in the constrained loss landscape, and the final Adam stage is used as a short post-refinement adjustment. Forecasting follows a 120-day train–test protocol:



The first 100 days are used for PINN fitting, and the final 20 days are reserved for out-of-sample validation. Relative to mechanistic calibration alone, the added value of the neural stage is that it produces a single constrained inference framework in which latent-state reconstruction, reduced-parameter estimation, data fitting, and short-horizon validation are carried out simultaneously under the same discrete fractional epidemic structure.

4.2 PINN Results, Validation, and Forecasting

We report the PINN results through train–test validation, full-window reconstruction, and training diagnostics. The purpose of this section is not only to demonstrate goodness of fit, but also to clarify how the discrete physics-informed neural stage complements the baseline mechanistic calibration on the same observation layer.

Fig. 12 shows the train–test comparison for cumulative and daily outputs at the observation level. The cumulative panels indicate preservation of the long-scale epidemic trend on the test segment, while the daily panels retain the dominant peak structure and subsequent decay after the train–test split. The daily comparisons are performed on the smoothed (7-day moving-average) series, consistently with the observation model. Taken together, these panels indicate that the learned neural representation preserves the main epidemic geometry over the held-out segment without introducing visible high-frequency artifacts. In view of the relatively short retained horizon, this validation should be interpreted as short-horizon predictive support rather than as evidence of long-range forecasting reliability.

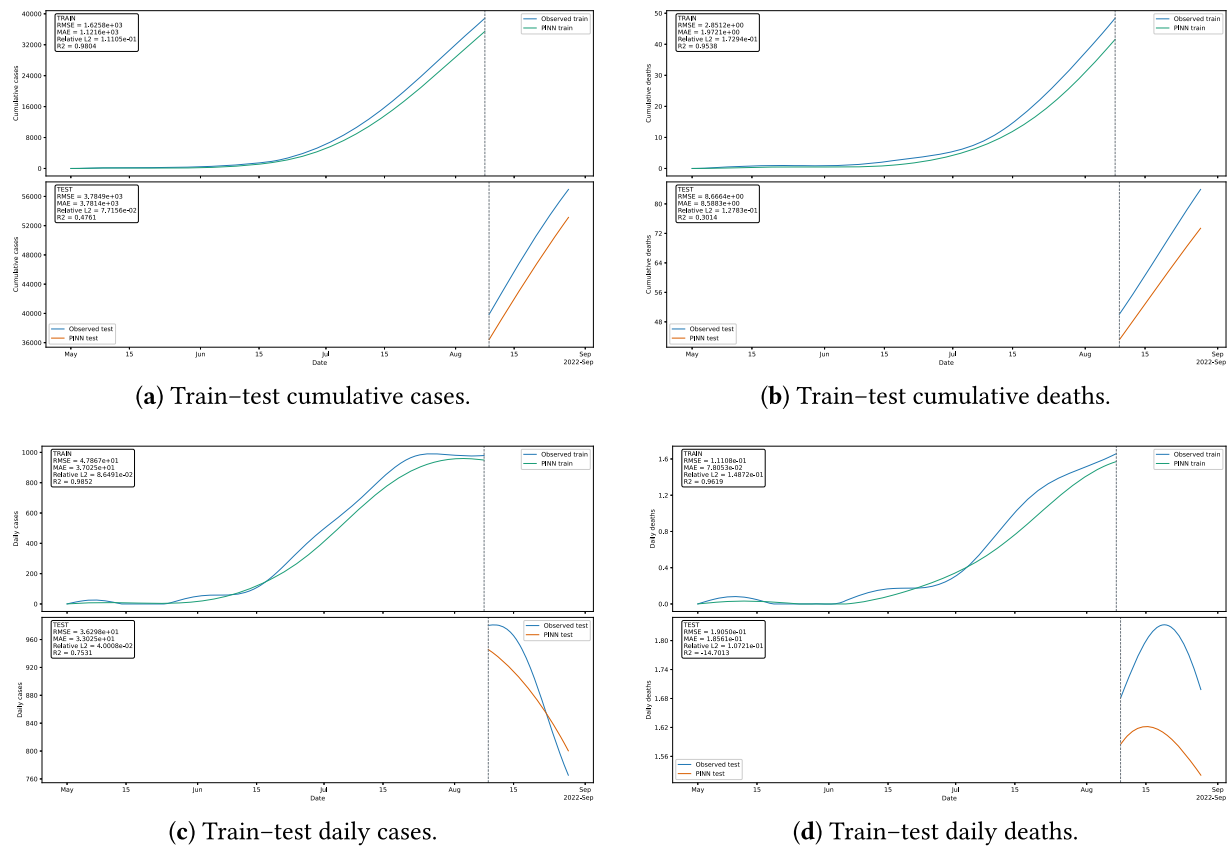


Figure 12: PINN train–test validation on cumulative and daily epidemic outputs at the observation level. The first 100 days are used for training and the final 20 days for out-of-sample testing within the retained 120-day window. The panels report RMSE, MAE, relative- ℓ_2 , and R^2 metrics for cumulative and smoothed daily outputs.

Fig. 13 shows the corresponding full-window reconstruction. The cumulative trajectories are closely reproduced over the whole horizon, while the daily trajectories recover the main wave profile without spurious oscillations. Again, daily outputs correspond to the smoothed observation-level series. This confirms that the neural stage remains numerically aligned with the observation-processing pipeline adopted in the mechanistic calibration and does not rely on a different data representation.

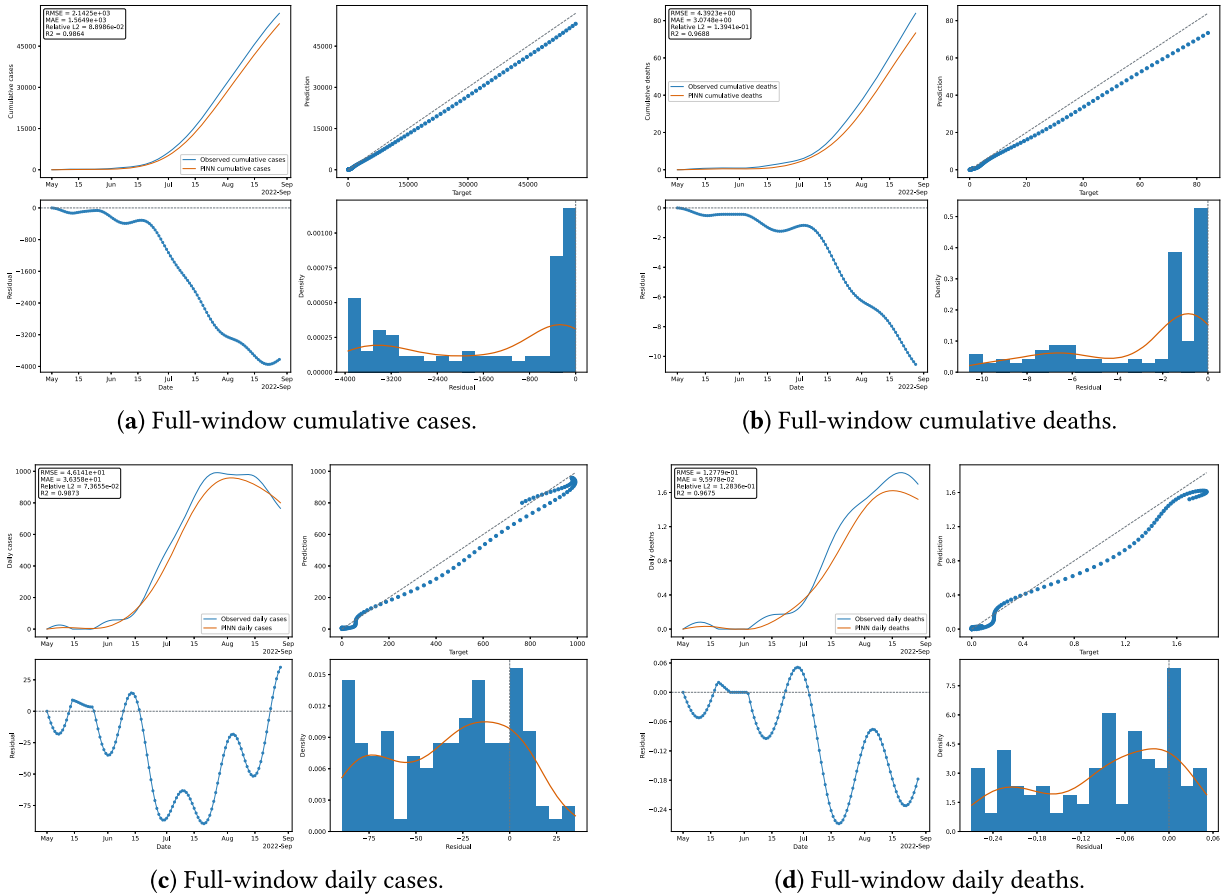


Figure 13: Full-window PINN reconstruction of cumulative and daily epidemic outputs at the observation level. The panels compare reported surveillance outputs with the constrained neural reconstruction and include residual and target–prediction diagnostics to assess observation-level agreement.

To make the short-horizon validation more explicit, Table 5 summarizes the main numerical indicators read from Figs. 12 and 13. These values are not introduced as a long-range forecasting benchmark; they are observation-level diagnostics for the first 100 training days, the final 20 held-out days, and the complete 120-day window.

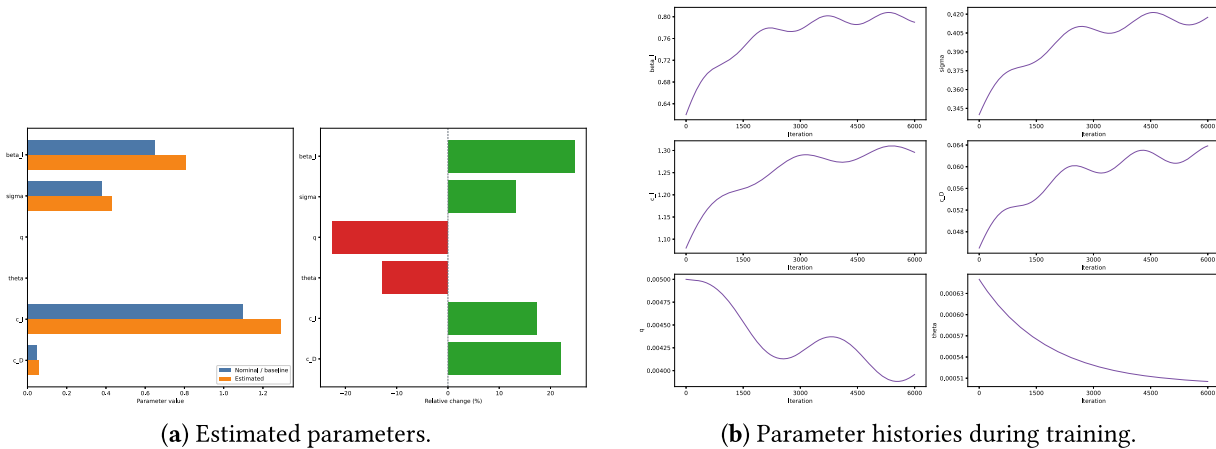
The explicit external comparison requested for short-horizon validation is provided by the controlled integer-order benchmark in Section 3.2. The PINN table is retained here as an internal train–test diagnostic of the proposed inverse-learning stage.

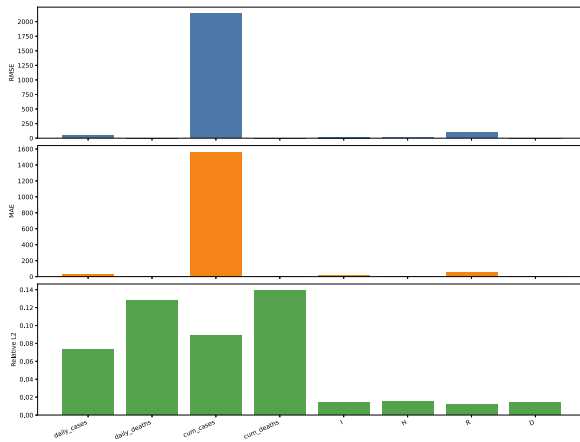
The negative R^2 for the 20-day daily-death test segment should not be overinterpreted as a failure of the epidemic reconstruction. It reflects the extremely small scale and low variability of the smoothed daily-death series over the held-out segment, where small absolute deviations can produce unstable coefficient-of-determination values. For this reason, RMSE, MAE, and relative- ℓ_2 are reported together with R^2 .

Table 5: PINN observation-level accuracy indicators extracted from the train–test and full-window diagnostic panels.

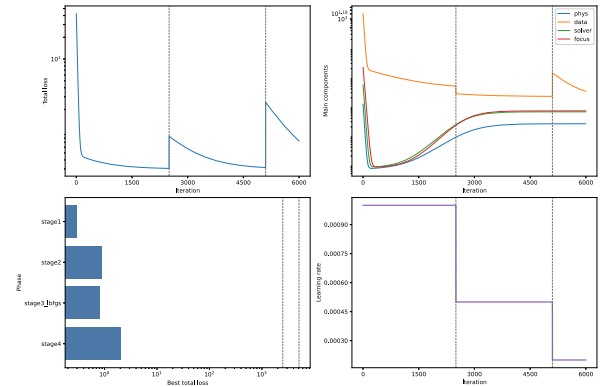
Output	Segment	RMSE	MAE	Rel.- ℓ_2	R^2
Cumulative cases	Train	1.6258×10^3	1.1216×10^3	1.1105×10^{-1}	0.9804
Cumulative cases	Test	3.7849×10^3	3.7814×10^3	7.7156×10^{-2}	0.4761
Cumulative deaths	Train	2.8512	1.9721	1.7294×10^{-1}	0.9538
Cumulative deaths	Test	8.6664	8.5883	1.2783×10^{-1}	0.3014
Daily cases	Train	4.7867×10^1	3.7025×10^1	8.6491×10^{-2}	0.9852
Daily cases	Test	3.6298×10^1	3.3025×10^1	4.0008×10^{-2}	0.7531
Daily deaths	Train	1.1108×10^{-1}	7.8053×10^{-2}	1.4872×10^{-1}	0.9619
Daily deaths	Test	1.9050×10^{-1}	1.8561×10^{-1}	1.0721×10^{-1}	-14.7013
Cumulative cases	Full window	2.1425×10^3	1.5649×10^3	8.8986×10^{-2}	0.9864
Cumulative deaths	Full window	4.3923	3.0748	1.3941×10^{-1}	0.9688
Daily cases	Full window	4.6141×10^1	3.6358×10^1	7.3655×10^{-2}	0.9873
Daily deaths	Full window	1.2779×10^{-1}	9.5978×10^{-2}	1.2836×10^{-1}	0.9675

Fig. 14 collects the main learning diagnostics, including the learned parameter values, their training histories, the summary error indicators on observation-level quantities, and the training/forecast summary panel. The diagnostics indicate stable convergence of the reduced trainable coefficients toward a consistent regime over the optimization stages. From the parameter panels, the learned values are approximately $\widehat{\beta}_{I,\vartheta} \approx 0.79$, $\widehat{\sigma}_{\vartheta} \approx 0.41$, $\widehat{q}_{\vartheta} \approx 0.0040$, $\widehat{\xi}_{\vartheta} \approx 5.0 \times 10^{-4}$, $\widehat{c}_C \approx 1.29$, and $\widehat{c}_D \approx 0.064$. These quantities should be understood as approximate learned coefficients read from the diagnostic panels for the retained observation window rather than as independently validated epidemiological constants. The observation coefficients in the neural stage are denoted by $\widehat{c}_C$ and $\widehat{c}_D$ to distinguish them from the mechanistic reporting parameters r_C and r_D . The pairs have analogous observation-level roles, but they are not assumed to be identical: r_C, r_D are calibrated within the forward mechanistic observation map, whereas $\widehat{c}_C, \widehat{c}_D$ are learned inside the constrained neural inverse problem together with the neural latent state. Thus, they should be interpreted as stage-specific effective observation coefficients, not as direct administrative reporting rates. In the diagnostic plot, the label c_I refers to the case-observation coefficient c_C , and the plotted label “theta” refers to the waning-immunity coefficient ξ_{ϑ} , not to the fractional order ϑ .

**Figure 14:** (Continued)



(c) Summary accuracy metrics (observation level).



(d) Training and forecast summary.

Figure 14: PINN diagnostics: learned parameters, training histories, and summary metrics evaluated at the observation level. The panels summarize convergence of the reduced trainable coefficients, loss components, and accuracy indicators, and should be interpreted as diagnostics for the retained short-horizon inverse-learning experiment.

Overall, the numerical evidence in Figs. 12–14 suggests that the neural stage does not merely reproduce the mechanistic baseline, but provides a reduced-parameter, dynamically constrained inference layer that preserves observation-level agreement while yielding internally consistent latent trajectories and short-horizon predictive validation.

4.3 Latent-State Inference

The PINN also provides inferred trajectories for the unobserved compartments. These are not direct measurements, but latent states jointly determined by the observation-level data constraints and the discrete fractional dynamics. Accordingly, the inferred states should be interpreted as model-based reconstructions that are admissible with respect to both the reported surveillance outputs and the retained discrete fractional epidemic structure.

Fig. 15 shows the exposed, symptomatic infectious, asymptomatic infectious, and quarantined trajectories. These inferred latent states are consistent with the learned epidemic timing and burden encoded by the observation-level fit. In particular, the exposed and infectious classes govern the timing and amplitude of the epidemic wave, while the quarantined class remains smaller and follows the infectious burden as a downstream controlled compartment. This qualitative ordering is epidemiologically coherent within the proposed model architecture and supports the interpretation of the PINN stage as a latent-state reconstruction tool rather than a purely black-box predictor.

Fig. 16 completes the inferred state picture with the hospitalized, recovered, and mortality trajectories. These downstream states are induced by the learned infectious structure and the governing fractional dynamics, and therefore should be interpreted as internally reconstructed latent trajectories rather than directly observed quantities. In particular, these compartments inherit their behavior from the coupled epidemic flow and observation constraints, so they should not be read as independently identified data series.

Overall, the PINN stage complements the mechanistic calibration by enforcing observation-level data agreement, discrete fractional dynamics, and latent-state regularity within a single constrained framework. Its main added value in the present study is therefore not the replacement of the mechanistic model, but the construction of a unified inverse-learning layer in which reduced-parameter estimation, observation-level

fitting, latent-state reconstruction, and short-horizon validation are carried out simultaneously under the same discrete fractional epidemic dynamics.

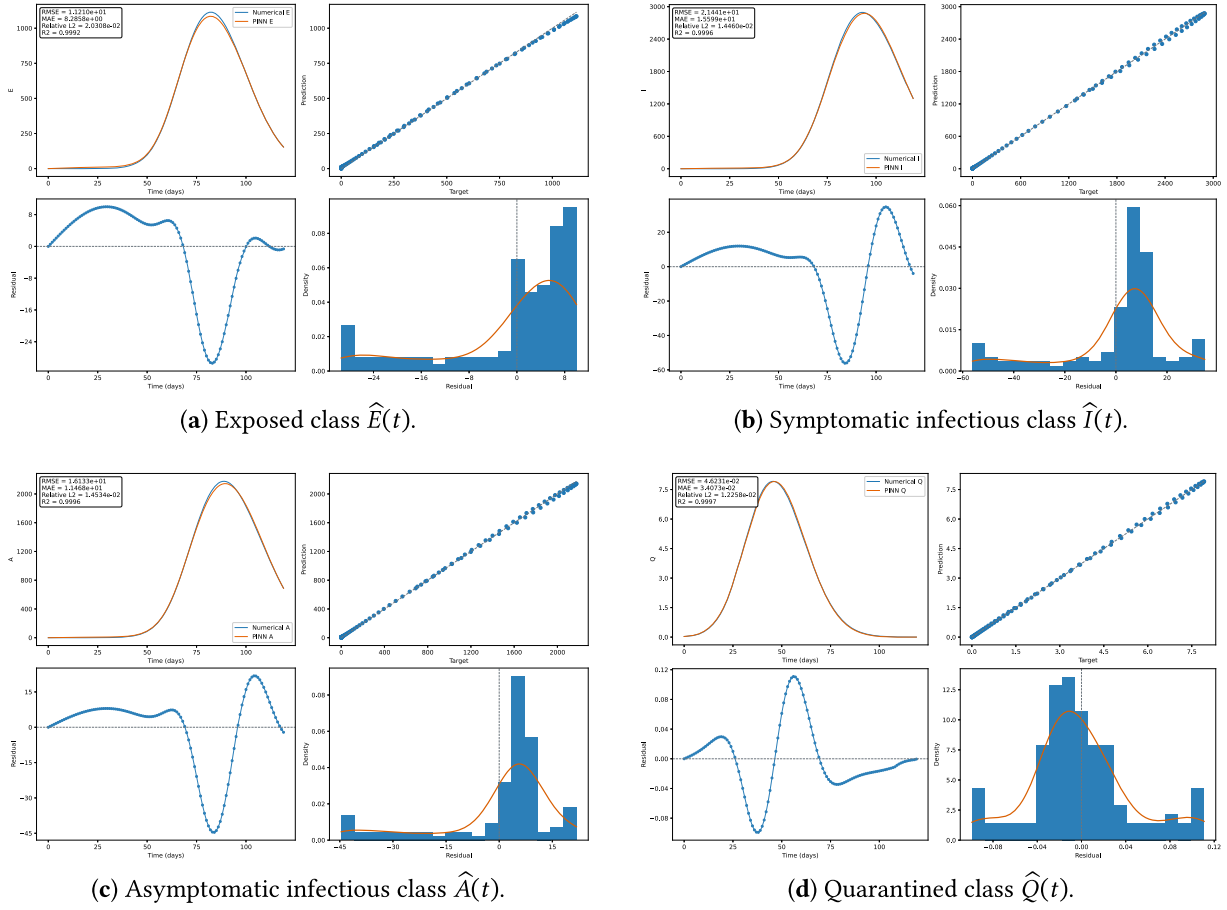


Figure 15: PINN-inferred latent trajectories: exposed, infectious, and quarantined classes. These unobserved compartments are reconstructed from the combined observation-level data loss and discrete fractional physics residual, and should be interpreted as admissible model-based latent states rather than direct measurements.

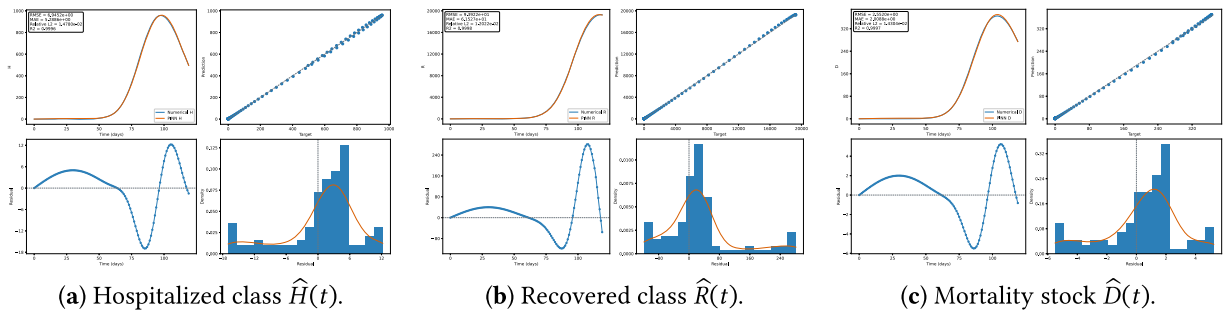


Figure 16: PINN-inferred downstream latent trajectories. The hospitalized, recovered, and mortality-stock reconstructions illustrate downstream consequences of the learned exposed and infectious dynamics under the same constrained fractional model.

5 Discussion and Public-Health Interpretation

The results should be interpreted as a joint mathematical and data-driven reconstruction of the first global mpox wave rather than as a direct measurement of all epidemic compartments. From a public-health perspective, the main value of the observation layer is that it separates the latent transmission process from the reported surveillance outputs. This distinction is important because reported daily cases and deaths are affected by under-reporting, smoothing, aggregation, and partial observation, whereas the model states describe hidden epidemic pathways such as exposure, asymptomatic infection, quarantine, hospitalization, and recovery.

The threshold quantities $\mathcal{R}_0$ and $\mathcal{R}_c$ provide interpretable summary indicators for the transmission regime. They are useful for reasoning about the relative roles of transmission, vaccination protection, mitigation, and removal mechanisms. In particular, the sensitivity experiments indicate that reducing effective transmission or increasing removal and isolation effects lowers the reconstructed infectious burden. This supports the qualitative public-health interpretation that early contact reduction, case isolation, targeted quarantine, and vaccination-related protection can shift the system toward lower epidemic burden.

The inferred latent trajectories can also support outbreak-management interpretation. For example, the exposed and infectious classes describe the hidden upstream burden driving reported cases, while the quarantined and hospitalized compartments provide model-based indications of potential isolation and clinical-care demand. These quantities are not direct administrative counts, but they offer a dynamically consistent way to interpret what types of hidden epidemic pressure may underlie the observed surveillance curves.

The controlled $\vartheta = 1$ comparison provides a direct check against the classical limit of the same compartmental model. Under the retained fractional calibration, the $\vartheta = 0.90$ configuration gives smaller observation-level errors on the highlighted segment for all four outputs. This result isolates the effect of the operator order, but it is not an independently re-optimized comparison and therefore should not be interpreted as a universal advantage over every integer-order calibration. The observation-level structural assessment also limits parameter interpretation: the surveillance data constrain the composite quantities Ψ_β , Ψ_C , and Ψ_D , whereas the individual transmission, reporting, and hidden-initialization factors are not uniquely separated by the four aggregate outputs.

Several limitations remain. First, the data are aggregated at the World level, so regional heterogeneity, local interventions, demographic structure, and reporting differences are not resolved. Second, the retained window contains only 120 daily observations; the PINN split uses 100 training days and 20 held-out days, and the model is not claimed to be validated across later waves without renewed calibration. Third, the reporting layer does not include an explicit reporting-delay mechanism. Fourth, the integer-order benchmark holds the calibrated parameterization fixed and does not replace an independent re-optimization study. Finally, the identifiability result is exact for the explicit observation-level combinations but is not a global structural-identifiability proof for the full nonautonomous fractional system. These limitations motivate regional validation, explicit delay models, uncertainty quantification, and a full-system identifiability analysis.

6 Conclusion and Future Work

This work developed and analyzed a fractional-order SEVIAQHRD model for mpox transmission, combining mathematical analysis, observation-level mechanistic calibration, and discrete physics-informed inverse learning. The model links latent epidemic states to daily and cumulative case and death observations through an explicit reporting map, while the PINN stage uses the same L1 fractional discretization for reduced-parameter estimation, latent-state reconstruction, and 100-day/20-day train-test validation.

The additional assessment adds two specific results. First, the calibrated fractional configuration produces smaller observation-level errors than the controlled $\vartheta = 1$ limit on the highlighted 20-day segment when the remaining parameterization is held fixed. Second, the observation-level structural analysis shows that aggregate surveillance constrains composite transmission and reporting quantities rather than every calibrated factor separately; the hidden-state initialization multipliers therefore remain nuisance quantities. These findings strengthen the model comparison while restricting the interpretation of individual inverse estimates to what the available outputs support.

The conclusions remain limited to the retained first-wave global window. Future work should examine regional or multi-wave data, explicit reporting delays, uncertainty quantification, joint estimation of ϑ , independent recalibration of integer-order alternatives, and a full-system structural or practical identifiability analysis. Optimal-control extensions are also relevant, but they remain outside the reconstruction and inverse-learning scope of the present paper.

Acknowledgement: The authors gratefully acknowledge the institutional facilities and academic environment that facilitated the completion of this work.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Saleh Albadran: supervision, project administration, validation, and writing—review and editing. Younes Brahim Oumedjber: conceptualization, methodology, mathematical analysis, numerical implementation, data curation, writing—original draft, and writing—review and editing. Omar Kahouli: supervision, correspondence, validation, resources, and writing—review and editing. Imed Bouzida: methodology, mathematical verification, validation, and writing—review and editing. Moawia Farah: data interpretation, software support, validation, and writing—review and editing. Ali Aloui: methodology, validation, formal analysis, and writing—review and editing. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The mpox surveillance data used in this study are publicly available from the Our World in Data repository at <https://ourworldindata.org/monkeypox>. The processed observation-level series used for calibration consist of daily and cumulative cases and deaths aggregated at the World level over the retained study window.

Ethics Approval: Not applicable. This study uses publicly available, aggregated surveillance data and does not involve individual-level human subjects data.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Laurenson-Schafer H, McMenamin M, Ebbarnesh LK, Karagiannis I, Biaukula V, Lewis R, et al. Characteristics and transmission dynamics of global travel-related mpox cases caused by clade Ib monkeypox virus. *Emerg Infect Dis.* 2026;32(2):225–8. doi:10.3201/eid3202.251530.
2. Titanji BK, Hazra A, Zucker J. Mpox clinical presentation, diagnostic approaches, and treatment strategies: a review. *JAMA.* 2024;332(19):1652–62. doi:10.1001/jama.2024.21091.
3. Lin YC, Wen TH, Shih WL, Vermund SH, Fang CT. Impact of vaccination and high-risk group awareness on the mpox epidemic in the United States, 2022–2023: a modelling study. *EClinicalMedicine.* 2024;68(13):102407. doi:10.1016/j.eclinm.2023.102407.
4. Onifade AA, Akindele OA, Ahmad I, Khan MA, Mat Isa N, Alzahrani E. Modeling monkeypox transmission with a compartmental framework to evaluate testing, isolation and public awareness strategies. *Sci Rep.* 2025;15(1):27236. doi:10.1038/s41598-025-10852-y.

5. Yadav R, Chaudhary AA, Srivastava U, Gupta S, Rustagi S, Rudayni HA, et al. Mpox 2022 to 2025 update: a comprehensive review on its complications, transmission, diagnosis, and treatment. *Viruses*. 2025;17(6):753.
6. Li J, Luo Q, Beltran E, Chen Q, Perez S, Frutos R, et al. A review of epidemiological modeling studies on monkeypox. *China CDC Wkly*. 2026;8(20):587–96. doi:10.46234/ccdcw2026.097.
7. Whitehouse M. Accelerated inference for stochastic compartmental models with over-dispersed partial observations. *Stat Comput*. 2026;36(3):109. doi:10.1007/s11222-026-10865-1.
8. Wang X, Wang H. Robust inverse reconstruction of time-varying transmission rates across model structures and incidence forms. *Bull Math Biol*. 2026;88(3):32. doi:10.1007/s11538-026-01597-4.
9. Okyere E, Asamoah Afful B, Nana-Kyere S, Agyemang Safo G, Asante C, Marri D, et al. Modeling the transmission dynamics and optimal control analysis of mpox infection incorporating vaccination control strategy. *Nonlinear Dyn*. 2026;114(3):184. doi:10.1007/s11071-025-12047-6.
10. Omame A, Iyaniwura SA, Ebenezer A, Han Q, Wang X, Bragazzi NL, et al. Pre-exposure vaccination in the high-risk population is crucial in controlling mpox resurgence in Canada. *BMC Infect Dis*. 2026;26(1):14. doi:10.1186/s12879-025-12172-y.
11. Khan MA, DarAssi MH, Tasqeeruddin S, Alzahrani AK, Jebran S, Mat Isa N. A fractional-order vaccination model to analyze the dynamics of mpox Clade I and II with real data. *Sci Rep*. 2026;16(1):11093. doi:10.1038/s41598-026-41453-y.
12. Agbata BC, Cenaj E, Dervishi R, Danjuma YJ, Shior MMA, Abah E, et al. Fractional-order mathematical model for monkeypox transmission dynamics using the Atangana-Baleanu Caputo operator. *BMC Infect Dis*. 2025;25(1):1000. doi:10.1186/s12879-025-11383-7.
13. Ramaswamy R, Mani G, Kumar D, Ege O. Mathematical model of the monkeypox virus disease via fractional order derivative. *Comput Model Eng Sci*. 2025;143(2):1843–94. doi:10.32604/cmesci.2025.063672.
14. Malathy R, Krishnan GSS, Loganathan K. A Caputo fractional-order SEIHRD model for Ebola: theoretical analysis, sensitivity, bifurcation, and numerical simulations. *Sci Rep*. 2026;16(1):13661.
15. Agarwal R, Airan P, Agarwal RP. Exploring the landscape of fractional-order models in epidemiology: a comparative simulation study. *Axioms*. 2024;13(8):545. doi:10.3390/axioms13080545.
16. Xu S, Hu Y. Dynamic analysis of a Caputo fractional-order SEIR model with a general incidence rate. *Sci Rep*. 2025;15(1):17561. doi:10.1038/s41598-025-01400-9.
17. Rahman JU, Arshad M, Mustafa N. A boundary-aware PINN framework for monkeypox transmission dynamics using Chebyshev-distributed collocation. *Int J Appl Comput Math*. 2025;11(6):251. doi:10.1007/s40819-025-02071-w.
18. Farea A, Yli-Harja O, Emmert-Streib F. Using physics-informed neural networks for modeling biological and epidemiological dynamical systems. *Mathematics*. 2025;13(10):1664. doi:10.3390/math13101664.
19. Jangir P, Agarwal G, Nisar KS. Fractional-order epidemic modeling with a deep neural network framework. *Sci Rep*. 2026;16(1):15025. doi:10.1038/s41598-026-37775-6.
20. Shafqat R, Imran, Al-Quran A, Djaouti AM. Enhancing rabies epidemic modeling with neural networks and fractional calculus. *Sci Rep*. 2026;16(1):10409. doi:10.1038/s41598-026-40853-4.
21. Diethelm K. The analysis of fractional differential equations: an application-oriented exposition using differential operators of Caputo type. Vol. 2004 of lecture notes in mathematics. Berlin/Heidelberg, Germany: Springer; 2010.
22. Li X, Li X, Wen C. Control for fractional-order systems and fractional-order control: theories and applications. Boca Raton, FL, USA: CRC Press; 2025.
23. Wu C. Comparison principles for systems of Caputo fractional-order ordinary differential equations. *Chaos Solitons Fractals*. 2023;171(4):113437. doi:10.1016/j.chaos.2023.113437.
24. Matignon D. Stability results for fractional differential equations with applications to control processing. *Comput Eng Syst Appl*. 1996;2:963–8.
25. van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci*. 2002;180(1–2):29–48. doi:10.1016/s0025-5564(02)00108-6.

26. Our World in Data. Mpox (monkeypox) data. 2026 [cited 2026 Mar 27]. Available from: <https://ourworldindata.org/monkeypox>.
27. Yang Z, Zeng F. A corrected L1 method for a time-fractional subdiffusion equation. *J Sci Comput*. 2023;95(3):85. doi:10.1007/s10915-023-02204-7.
28. Jiang Y, Chen H. Local convergence analysis of an L1-ADI scheme for a two-dimensional reaction-subdiffusion equation. *J Appl Math Comput*. 2024;70(3):1953–64. doi:10.1007/s12190-024-02037-z.
29. Liyanage YR, Saucedo O, Tuncer N, Chowell G. A tutorial on structural identifiability of epidemic models using StructuralIdentifiability.jl. *Bull Math Biol*. 2026;88(6):102. doi:10.1007/s11538-026-01648-w.
30. Saucedo O, Laubmeier A, Tang T, Levy B, Asik L, Pollington T, et al. Comparative analysis of practical identifiability methods for an SEIR model. *AIMS Math*. 2024;9(9):24722–61. doi:10.3934/math.20241204.