



Research article

A two-strain Delta-Omicron COVID-19 model: stability and physics-informed parameter estimation

Mallela Ankamma Rao¹, Nazek Alessa², Medisetty Padma Devi¹, Carlos Martin-Barreiro³ and Prasantha Bharathi Dhandapani^{4,*}

¹ Department of Mathematics & Statistics, Vignan's Foundation for Science, Technology & Research (Deemed to be University), Vadlamudi, Guntur-522213, Andhra Pradesh, India; ankammaraomaths7@gmail.com, padmadevimedisetty@gmail.com

² Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; naalessa@pnu.edu.sa

³ Facultad de Ingeniería, Universidad Espíritu Santo (UEES), Samborondón 0901952, Ecuador; carlosmartin@uees.edu.ec

⁴ Department of Mathematics, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 602105, Tamil Nadu, India; d.prasanthabharathi@gmail.com

* **Correspondence:** Email: d.prasanthabharathi@gmail.com.

Abstract: The Delta and Omicron variants drove the two most consequential COVID-19 waves in the region, yet no existing multi-strain compartmental model combines rigorous stability analysis with modern, uncertainty-quantified parameter estimation for this pair. We addressed this gap through four contributions. First, we formulated a deterministic two-strain $SVE_iA_iI_iJHR$ model with eleven compartments and established positivity, boundedness, closed-form variant-specific reproduction numbers $\mathcal{R}_{0i}$ (evaluated under the baseline intervention configuration), and the local stability and a conditional global-stability result for the disease-free equilibrium via a Castillo-Chavez-Song decomposition, together with the existence of the single-variant endemic equilibria and their local stability under explicit Routh-Hurwitz and invasion conditions, together with a numerical threshold-branch analysis illustrating the forward exchange of stability at the epidemic threshold. Second, we introduced a three-stage hybrid estimation pipeline—lsqnonlin, a deterministic physics-informed neural network (PINN) embedding the full eleven ordinary differential equation (ODE) residual, and a long short-term memory (LSTM) residual-correction module—calibrated against Delta- and Omicron-wave surveillance data. Third, we extended the physics-informed neural network (PINN) to a Bayesian PINN via Hamiltonian Monte Carlo simulation, yielding 95% credible intervals for all six learned parameters and for $\mathcal{R}_{0i}$ ($\mathcal{R}_{01} = 1.87$, $\mathcal{R}_{02} = 1.64$), providing uncertainty quantification absent from prior multi-strain COVID-19 models for this setting. Propagating literature uncertainty in the fixed parameters α, η_i jointly with the learned parameters widens these credible intervals

substantially and can bring them below unity. Fourth, local and global sensitivity analyses and a counterfactual efficacy-coverage scenario analysis showed that vaccine efficacy is the single most influential controllable parameter and that, under the proportional intervention scenario, Delta is the more restrictive variant, with a conditional proportional-intervention benchmark of 59.6% (posterior-based 95% interval: [56.0%, 64.6%]) at the reported operating efficacies. This benchmark is not a full-model herd-immunity threshold. The hybrid PINN-LSTM framework reduces 30-day forecast root mean square error (RMSE) by about 23% over PINN-only and 40% over lsqnonlin.

Keywords: two-strain COVID-19 model; physics-informed neural network (PINN); Bayesian PINN; hybrid PINN-LSTM forecasting; uncertainty quantification; herd-immunity threshold

Mathematics Subject Classification: 92D30, 34D20, 37N25, 49J15, 93A30

1. Introduction

The COVID-19 pandemic, caused by SARS-CoV-2, has reshaped global public-health priorities since its emergence in late 2019. As of April 13, 2024, Worldometer recorded 704,753,890 confirmed cases and 7,010,681 associated fatalities worldwide [1], with India alone contributing 45,035,393 confirmed cases and 533,570 deaths over the same period. The trajectory of the pandemic has been governed not by a single homogeneous pathogen but by a sequence of antigenically and epidemiologically distinct variants, whose differing transmissibility, severity, and immune-escape properties have repeatedly forced the recalibration of pharmaceutical and non-pharmaceutical interventions [2]. The World Health Organization classifies these variants into two principal categories [3]:

(1) **Variants of concern (VOC):** Variants exhibiting higher transmissibility, increased disease severity, reduced antibody neutralization from prior infection or vaccination, or diminished effectiveness of treatments and vaccines. The VOCs include Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529).

(2) **Variants of interest (VOI):** Variants whose genetic features may foreseeably alter transmissibility, severity, antibody neutralization, immune escape, or therapeutic efficacy, including Epsilon (B.1.427/B.1.429), Zeta (P.2), Eta (B.1.525), Theta (P.3), Iota (B.1.526), Kappa (B.1.617.1), Lambda (C.37), and Mu (B.1.621).

In India, two of these VOCs proved decisive. The Delta variant drove the catastrophic second wave of April–June 2021, characterized by high hospitalization and mortality rates, while the Omicron variant produced the explosive third wave of December 2021–February 2022, marked by a much higher transmission rate but a milder clinical profile. India's vaccination program relied on COVAXIN (BBV152) and COVISHIELD (ChAdOx1 nCoV-19), authorized under emergency-use provisions. By December 2021, COVAXIN accounted for 117 million doses with 46.7% of adults fully vaccinated [4], while COVISHIELD—produced by the Serum Institute of India under licence from Oxford-AstraZeneca—represented 88% of all doses, exhibiting 67% efficacy against symptomatic disease and nearly 100% efficacy against severe outcomes in Phase IV trials. Real-world meta-analyses subsequently confirmed reduced but non-negligible vaccine effectiveness against

the Delta and Omicron variants, with marked waning over time [5, 6]. Mechanistic compartmental modeling has been a central tool for quantifying these dynamics, and multi-variant frameworks have been developed in several directions. Massard et al. [7] extended the multi-strain paradigm to a calibrated susceptible-exposed-infectious-recovered (SEIR)-type model fitted to French hospital-admission data covering the simultaneous circulation of the historical, Alpha, and Delta variants, estimating variant-specific transmission advantages and quantifying how the progressive replacement of one variant by another shifted the effective reproduction number, with numerical experiments highlighting the disproportionate impact that even a modest transmissibility advantage exerts on healthcare-system strain.

The role of vaccination has been explicitly modeled in two-strain settings by De Leon et al. [8], who derived the minimum vaccinated population fraction required to drive the joint reproduction number below unity. Tchoumi et al. [9] showed that vaccination targeted at only the dominant variant can paradoxically allow the non-targeted variant to grow, arguing for broad-spectrum immunization. Otunuga [10] obtained explicit herd-immunity thresholds as functions of vaccine coverage and efficacy, including conditions under which elimination is unattainable through vaccination alone. Saha and Saha [11] examined Delta-Omicron co-circulation under combined vaccination and non-pharmaceutical interventions, showing high-efficacy vaccination is necessary but not sufficient without behavioral measures. Ambalarajan et al. [12] demonstrated how variant-specific vaccine coverage requirements shift as new strains emerge. Yunus and Olayiwola [13] formulated a fractional-order susceptible-exposed-vaccinated-infectious-asymptomatic-recovered (SEVIAR) model with herd-immunity thresholds analogous in spirit to $\rho^*(\xi_i)$ below, but without a multi-strain reproduction-number structure, competitive-exclusion analysis, or physics-informed parameter estimation.

Venkatesh et al. [14] applied incremental cost-effectiveness analysis to a sixteen-compartment multi-variant model, finding joint vaccination-plus-treatment outperforms either lever alone. Bandekar [15] formulated an optimal-control approach for COVID-19 testing strategies, deriving time-varying testing-rate policies that minimize cumulative infections subject to a fixed testing-capacity budget. Ambalarajan et al. [16] fitted a compact six-compartment model to surveillance data; Ghosh and Ghosh [17] performed a sensitivity analysis on a vaccination-augmented Omicron-wave model; Ankamma Rao et al. [18] incorporated pulmonary aspergillosis co-infection; and Rao et al. [19] coupled susceptible-exposed-asymptomatic infected-ympomatic infected-quarantined-hospitalized-recovered-deceased (SEAIQHRD) with Gaussian process regression (GPR) for improved short-term forecasting. Khan and Atangana [20] showed a fractional-order Caputo-Fabrizio Omicron model yields closer agreement with case data than its integer-order counterpart. Two studies are especially close in scope. Ngonghala et al. [21] showed, for the United States, that Omicron displaced Delta once its control reproduction number exceeded Delta's—a mechanism analogous to Definition 1(b) below—but without a Bayesian uncertainty-quantified pipeline, an identifiability assessment, or a herd-immunity threshold surface. Negi et al. [22] introduced a four-variant Indian model with reinfection and optimal control, calibrated by classical least squares, without global stability analysis, threshold-branch analysis, or uncertainty-quantified estimation. The present work is complementary to both: it supplies the rigorous stability theory and the threshold-branch analysis that [21] does not pursue and the uncertainty-quantified hybrid estimation pipeline that neither provides, applied to the Indian Delta-Omicron pair.

The next-generation matrix method used throughout follows van den Driessche and Watmough [23],

who introduced the standard FV^{-1} construction of R_0 and its threshold property, and Diekmann et al. [24], who formalized the next-generation operator more generally. Global-stability arguments rely on the Castillo-Chavez-Song decomposition [25], which promotes local to global asymptotic stability of the disease-free equilibrium without an explicit Lyapunov function.

Despite this literature, two limitations remain pervasive. First, classical nonlinear least squares (e.g., `lsqnonlin`) is sensitive to initial guesses, returns only point estimates without uncertainty quantification, and ignores the ODE's structural constraints during fitting. Second, purely mechanistic models often fail to capture fine-scale unexplained variance in reported case data, degrading short-term forecasts. Raissi et al. [26] introduced physics-informed neural networks (PINNs), embedding ODE residuals as soft physics constraints for joint forward/inverse solving, and Yazdani et al. [27] recovered hidden parameters from sparse, noisy biological time series. The Bayesian PINN framework was formalized by Yang et al. [28] and applied by Linka et al. [29], with posterior sampling via Hamiltonian Monte Carlo/No-U-Turn sampler (NUTS) [30]. Hybrid mechanistic-data-driven approaches coupling compartmental ODEs with LSTM networks [31] achieve superior forecasting accuracy over either component alone. Millevoi et al. [32] benchmarked PINN-based calibration against classical methods for single-strain models, and Aghdaoui et al. [33] extended PINNs to a fractional-order model with simultaneous order estimation; neither addresses a multi-strain system, Bayesian uncertainty propagation onto R_{0i} , or a hybrid PINN-LSTM cascade.

1.1. Contributions of this work

Building on these mechanistic and machine-learning advances, the present study makes the following novel contributions:

- (1) We formulate a deterministic two-strain $SVE_iA_iI_iJHR$ compartmental model with eleven epidemiological compartments tailored to the Delta-Omicron pair—the two VOCs that dominated the Indian pandemic landscape and that exemplify two contrasting variant profiles. ($R_{01} \approx 1.87$, severe; $R_{02} \approx 1.64$, highly transmissible but milder.)
- (2) We establish positivity, ultimate boundedness, and the closed-form variant-specific reproduction numbers R_{0i} , and we establish local stability and a conditional global asymptotic-stability result for the disease-free equilibrium together with the existence and stability of single-variant endemic equilibria (the coexistence case is treated as a conjecture supported by a monotonicity argument, see Section 3.5).
- (3) We propose a three-stage hybrid physics-informed parameter estimation framework: a classical `lsqnonlin` baseline, a deterministic physics-informed neural network that simultaneously learns the six variant-specific parameters $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ and the full state trajectories, and a coupled LSTM residual-correction module producing a hybrid PINN-LSTM forecaster that outperforms the purely mechanistic approach. The estimates are supported by practical identifiability evidence and a formal structural-identifiability analysis using `StructuralIdentifiability.jl` [34], showing global identifiability of the three fitted parameters within each separately calibrated single-variant sub-model (see Section 4.2).
- (4) We extend the deterministic PINN to a Bayesian PINN and use Hamiltonian Monte Carlo sampling to deliver 95% credible intervals on every learned parameter and on R_{0i} , providing rigorous uncertainty quantification absent from the existing literature on multi-strain Indian COVID-19 models.
- (5) We perform local normalized sensitivity index (NSI) analysis and global sensitivity analysis using the partial rank correlation coefficient (PRCC) with Latin hypercube sampling, together with a numerical threshold-branch analysis around $\mathcal{R}_{0i} = 1$, and a counterfactual efficacy-coverage

scenario analysis, translating the analytical and PINN-driven findings into conditional public-health implications for India.

1.2. Relation to prior PINN/LSTM work

The PINN loss of [26] and the LSTM cell of [31] are established methods; we claim no architectural novelty. The contribution lies in the estimation pipeline: (a) a three-stage cascade ($\text{lsqnonlin} \rightarrow \text{PINN} \rightarrow \text{PINN-LSTM}$) in which each stage cross-checks the previous one, unlike the single-stage fits of [27, 29]; (b) the LSTM is restricted to a bounded, held-out-validated post-hoc correction of the scalar observable $\hat{C}(t)$, applied on top of a state trajectory and parameters already fixed by the PINN (Section 4.4) rather than joint end-to-end training, so the eleven-dimensional solution stays exactly ODE-consistent by construction; and (c) this cascade is, to our knowledge, applied here for the first time to a rigorously analyzed two-strain model with Bayesian uncertainty propagated onto R_{0i} itself. The LSTM performs one scalar regression task—the one-step-ahead residual $\hat{C}(t) - C(t)$ on a fixed 14-day window—applied after, not instead of, the PINN fit. This scalar target is the only quantity the LSTM ever sees or corrects: It cannot alter any of the other ten state variables of system (2.10)–(2.17), which remain exactly the Stage-2 PINN’s output.

1.3. On the locus of novelty

None of the individual components used here—PINNs [26], LSTM residual correction, or B-PINN uncertainty quantification [28]—is itself a new architecture, and an $\approx 23\%$ RMSE reduction alone is not sufficient evidence of methodological novelty; we claim no architectural novelty. The contribution is a rigorously validated integration of four elements not previously combined for a multi-strain compartmental model: (i) closed-form stability and competitive-exclusion theory, together with a numerical threshold-branch analysis (Theorems 3.1–3.5, Definition 1), letting the B-PINN’s parameter uncertainty propagate into decision-relevant quantities— $\mathcal{R}_{0i}$ and ρ^* —rather than stopping at parameter credible intervals as in [28, 29]; (ii) an estimation cascade in which each stage independently cross-checks the previous stage, together with the identifiability diagnostics of Section 4.2 and the noise-injection robustness check of Section 4.4 characterizing when the hybrid helps (largest under low-to-moderate noise, vanishing as noise increases, see Figure 5); (iii) an application to a rigorously analyzed two-strain system with Indian surveillance data, distinguishing this work from the two closest recent studies, Ngonghala et al. [21] and Negi et al. [22], neither of which combines global stability analysis with an uncertainty-quantified pipeline; and (iv) a formal structural-identifiability analysis of the estimation problem (see Section 4.2 and Table 2) via differential-algebra-based elimination—moving the identifiability question from an empirical characterization of when the estimator is trustworthy to a proof of whether the six-parameter inverse problem is well-posed at all, a step toward structural identifiability analysis “remaining underutilized” in [35], and (v) two formal propositions (Propositions 4.1 and 4.2, Section 4.4) that provide a theoretical characterization of the hybrid correction, including a derived bound on its departure from ODE-consistency and a stated, checkable condition under which the LSTM correction is expected to reduce, rather than increase, the expected forecast error. These results complement the empirical $\approx 23\%$ RMSE reduction by providing a theoretical account of the conditions under which the hybrid approach can improve forecasting performance. We describe the pipeline as a validated application and cross-validation framework

rather than a novel learning method; points (iv) and (v) provide analytical contributions that support the interpretation and validation of the proposed model.

The remainder of this article is organized as follows. Section 2 formulates the two-strain compartmental model and describes its epidemiological compartments. Section 3 establishes positivity, boundedness, the next-generation reproduction numbers, the local stability and conditional global stability of the disease-free equilibrium, and the local stability of the endemic equilibria. Section 4 presents the numerical workflow: It introduces the deterministic PINN (Section 4.2), the Bayesian PINN with uncertainty quantification (Section 4.3), the model calibration and short-term forecasting against the Indian data (Section 4.4), the local and global sensitivity analyses (Section 4.5), and the threshold-branch, vaccine-efficacy/coverage, and intervention-comparison analyses (Section 4.6). Section 5 synthesizes the analytical and computational findings into a results-and-discussion summary, and Section 6 concludes the paper.

2. Mathematical model formulation

In this study, we develop a deterministic compartmental model to describe the transmission dynamics of two COVID-19 variants, namely, Delta ($i = 1$) and Omicron ($i = 2$). We assume that individuals can be infected by at most one variant at a time (no simultaneous co-infection), consistent with the sequential wave structure observed in India [11]. Isolated (J) and hospitalized (H) individuals are assumed to have negligible onward transmission owing to strict infection-control measures; this is a standard simplifying assumption in compartmental models [36] whose effect is examined in the sensitivity analysis of Section 4.5. The total population $N(t)$ is divided into eleven epidemiological compartments: susceptible $S(t)$, vaccinated $V(t)$, exposed $E_i(t)$, asymptomatic infectious $A_i(t)$, symptomatic infectious $I_i(t)$ for each variant $i = 1, 2$, and three shared compartments comprising isolated $J(t)$, hospitalized $H(t)$, and recovered $R(t)$. Thus,

$$N(t) = S(t) + V(t) + \sum_{i=1}^2 (E_i(t) + A_i(t) + I_i(t)) + J(t) + H(t) + R(t). \quad (2.1)$$

Figure 1 illustrates the compartmental flow diagram of the two-strain model, showing the transitions between compartments and the epidemiological pathways linking susceptible, vaccinated, exposed, asymptomatic, symptomatic, isolated, hospitalized, and recovered individuals for each variant. All parameters are non-negative; their biological descriptions and baseline values are collected in Table 1, which displays each parameter's symbol, biological description, baseline numerical value, and literature source used throughout the manuscript. The epidemiological interpretation of each compartment and its governing equation are described below.

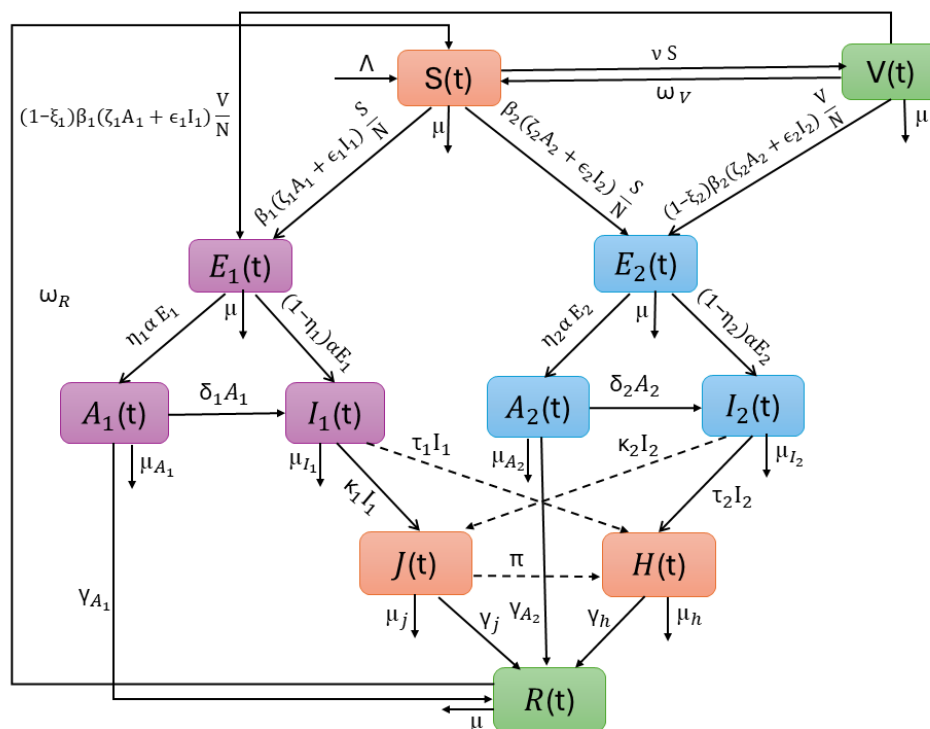


Figure 1. Compartmental flow diagram of the two-strain (Delta-Omicron) epidemic model.

Susceptible individuals (S). The susceptible class $S(t)$ consists of individuals who are at risk of acquiring infection from either variant. The population is augmented by recruitment at a constant rate Λ , representing births and net immigration, and by individuals who lose immunity from vaccination and from prior infection at waning rates ω_V and ω_R , respectively.

Susceptible individuals acquire infection through effective contact with infectious individuals. The transmission rate β_i is the effective per-contact infection probability for variant i . The dimensionless modification factors $\zeta_i \in (0, 1)$ and $\epsilon_i > 0$ scale the relative infectiousness of asymptomatic (A_i) and symptomatic (I_i) individuals, respectively, relative to a reference level encoded in β_i . Values $\zeta_i < 1$ reflect the empirically lower infectiousness of asymptomatic carriers [37], and ϵ_i is calibrated from data (Table 1); both are strictly positive. Susceptible individuals are vaccinated at per-capita rate ν and die naturally at rate μ . The governing equation is

$$\frac{dS}{dt} = \Lambda + \omega_V V + \omega_R R - \sum_{i=1}^2 \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{S}{N} - (\nu + \mu) S. \quad (2.2)$$

Vaccinated individuals (V). The vaccinated class $V(t)$ consists of individuals who have received vaccination at rate ν . The vaccine efficacy parameter $\xi_i \in [0, 1]$ quantifies the fractional reduction in susceptibility to variant i conferred by vaccination, with $\xi_i = 1$ corresponding to perfect sterilizing immunity and $\xi_i = 0$ to no protection. Due to imperfect efficacy, vaccinated individuals remain susceptible to breakthrough infection at a reduced force of infection

$$\frac{(1 - \xi_i) \beta_i (\zeta_i A_i + \epsilon_i I_i)}{N},$$

consistent with real-world observations of breakthrough infections for both Delta and Omicron [5]. Vaccinated individuals lose immunity at rate ω_V and return to the susceptible class, and die naturally at rate μ . The governing equation is

$$\frac{dV}{dt} = \nu S - \sum_{i=1}^2 (1 - \xi_i) \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{V}{N} - (\omega_V + \mu)V. \quad (2.3)$$

Exposed individuals (E_i). The exposed class $E_i(t)$ represents individuals infected with variant i who are in the latent (pre-infectious) stage. New exposures arise from both susceptible and vaccinated individuals through their respective forces of infection. The per-capita rate of progression from exposed to infectious is α , the inverse of the mean incubation period, so that a larger α implies faster progression. Exposed individuals die naturally at rate μ . The governing equation is

$$\frac{dE_i}{dt} = \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{S}{N} + (1 - \xi_i) \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{V}{N} - (\alpha + \mu)E_i, \quad i = 1, 2. \quad (2.4)$$

Remark. Both terms on the right-hand side of (2.4) are positive: The first represents new infections in the susceptible class and the second represents breakthrough infections in the vaccinated class. This formulation is consistent with the full system (2.10)–(2.17) below.

Asymptomatic infectious individuals (A_i). A proportion $\eta_i \in [0, 1]$ of exposed individuals develop subclinical (asymptomatic) infection upon leaving the latent stage and enter $A_i(t)$. Asymptomatic individuals may progress to symptomatic disease at per-capita rate δ_i , recover naturally at rate γ_{A_i} , die from disease at rate μ_{A_i} , or die naturally at rate μ . The undocumented nature of asymptomatic infection is a primary driver of sustained transmission [37]. The governing equation is

$$\frac{dA_i}{dt} = \eta_i \alpha E_i - (\delta_i + \gamma_{A_i} + \mu_{A_i} + \mu)A_i, \quad i = 1, 2. \quad (2.5)$$

Symptomatic infectious individuals (I_i). The symptomatic class $I_i(t)$ consists of individuals with clinically apparent infection. Inflow arises from two sources: exposed individuals who progress directly to symptoms at rate $(1 - \eta_i)\alpha$, and asymptomatic individuals who develop symptoms at rate δ_i . Symptomatic individuals are isolated at per-capita rate κ_i (representing detection, testing, and quarantine), require hospitalization at rate τ_i reflecting disease severity, die from disease at rate μ_{I_i} , and die naturally at rate μ . The governing equation is

$$\frac{dI_i}{dt} = (1 - \eta_i)\alpha E_i + \delta_i A_i - (\kappa_i + \tau_i + \mu_{I_i} + \mu)I_i, \quad i = 1, 2. \quad (2.6)$$

Isolated individuals (J). The isolated compartment $J(t)$ represents confirmed cases removed from the general population through detection, testing, or quarantine. Inflow arises from symptomatic individuals of both variants at rates κ_1 and κ_2 . Isolated individuals recover at rate γ_j , deteriorate clinically and require hospitalization at rate π , die from disease at rate μ_j , and die naturally at rate μ . Consistent with the assumption stated above, isolated individuals do not contribute to the force of infection. The governing equation is

$$\frac{dJ}{dt} = \sum_{i=1}^2 \kappa_i I_i - (\pi + \gamma_j + \mu_j + \mu)J. \quad (2.7)$$

Hospitalized individuals (H). The hospitalized class $H(t)$ consists of individuals with severe infection requiring inpatient medical care. Inflow arrives from symptomatic individuals at rate τ_i and from isolated individuals at rate π . Hospitalized individuals recover at rate γ_h , die from disease at rate μ_h , and die naturally at rate μ . Hospitalized individuals are assumed non-infectious (see above). The governing equation is

$$\frac{dH}{dt} = \sum_{i=1}^2 \tau_i I_i + \pi J - (\gamma_h + \mu_h + \mu)H. \quad (2.8)$$

Recovered individuals (R). The recovered class $R(t)$ comprises individuals who have cleared infection and acquired temporary immunity. Recovery inflows arise from asymptomatic individuals at rate γ_{A_i} , from isolated individuals at rate γ_j , and from hospitalized individuals at rate γ_h . Consistent with empirical evidence of waning post-infection immunity to SARS-CoV-2 [17], immunity decays at rate ω_R and individuals return to the susceptible class. Natural mortality occurs at rate μ . The governing equation is

$$\frac{dR}{dt} = \sum_{i=1}^2 \gamma_{A_i} A_i + \gamma_j J + \gamma_h H - (\omega_R + \mu)R. \quad (2.9)$$

Combining Eqs (2.2)–(2.9), the complete autonomous system of nonlinear ordinary differential equations governing the transmission dynamics is:

$$\frac{dS}{dt} = \Lambda + \omega_V V + \omega_R R - \sum_{i=1}^2 \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{S}{N} - (\nu + \mu)S, \quad (2.10)$$

$$\frac{dV}{dt} = \nu S - \sum_{i=1}^2 (1 - \xi_i) \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{V}{N} - (\omega_V + \mu)V, \quad (2.11)$$

$$\frac{dE_i}{dt} = \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{S}{N} + (1 - \xi_i) \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{V}{N} - (\alpha + \mu)E_i, \quad (2.12)$$

$$\frac{dA_i}{dt} = \eta_i \alpha E_i - (\delta_i + \gamma_{A_i} + \mu_{A_i} + \mu)A_i, \quad (2.13)$$

$$\frac{dI_i}{dt} = (1 - \eta_i) \alpha E_i + \delta_i A_i - (\kappa_i + \tau_i + \mu_{I_i} + \mu)I_i, \quad (2.14)$$

$$\frac{dJ}{dt} = \sum_{i=1}^2 \kappa_i I_i - (\pi + \gamma_j + \mu_j + \mu)J, \quad (2.15)$$

$$\frac{dH}{dt} = \sum_{i=1}^2 \tau_i I_i + \pi J - (\gamma_h + \mu_h + \mu)H, \quad (2.16)$$

$$\frac{dR}{dt} = \sum_{i=1}^2 \gamma_{A_i} A_i + \gamma_j J + \gamma_h H - (\omega_R + \mu)R, \quad (2.17)$$

where Eqs (2.12)–(2.14) hold for $i = 1, 2$, so that system (2.10)–(2.17) comprises eleven coupled scalar ODEs. The system is complemented with biologically feasible non-negative initial conditions

$$S(0), V(0), E_i(0), A_i(0), I_i(0), J(0), H(0), R(0) \geq 0, \quad i = 1, 2. \quad (2.18)$$

A complete list of all state variables and parameters, together with their biological interpretations, units, baseline values, and sources, is provided in Table 1.

Table 1. Baseline parameter values of the two-strain SVE_iA_iI_iJHR COVID-19 model.

Parameter	Description	Value	References
Λ	Recruitment rate into susceptible population	varies	
α	Conversion rate from exposed to infectious	0.071[0.071, 0.33]/day	[37, 38]
ν	Vaccination rate of susceptible individuals	0.0033/day	[39]
π	Transfer rate from isolation to hospitalization	0.025/day	[17]
β_1	Disease transmission rate (Delta)	0.8946/day	Estimated
β_2	Disease transmission rate (Omicron)	0.8686/day	Estimated
η_1	Asymptomatic proportion (Delta)	0.15[0.15, 0.7]	[37, 40]
η_2	Asymptomatic proportion (Omicron)	0.16[0.15, 0.7]	[37, 40]
κ_1	Isolation rate (Delta symptomatic)	0.132/day	Estimated
κ_2	Isolation rate (Omicron symptomatic)	0.178/day	Estimated
τ_1	Hospitalization rate (Delta)	0.184/day	[41]
τ_2	Hospitalization rate (Omicron)	0.106/day	[42]
ξ_1	Vaccine efficacy against infection (Delta)	0.78	[5]
ξ_2	Vaccine efficacy against infection (Omicron)	0.70	[6]
γ_{A_1}	Recovery rate (Delta asymptomatic)	0.0066/day	[43]
γ_{A_2}	Recovery rate (Omicron asymptomatic)	0.034/day	[17]
ζ_1	Asymptomatic infectivity factor (Delta)	0.3063	[36, 44]
ζ_2	Asymptomatic infectivity factor (Omicron)	0.5908	[36, 44]
γ_j	Recovery rate (isolated)	0.017/day	[17]
γ_h	Recovery rate (hospitalized)	0.017/day	[17]
ϵ_1	Symptomatic infectivity factor (Delta)	0.4231	[36, 45]
ϵ_2	Symptomatic infectivity factor (Omicron)	0.4711	[36, 45]
δ_1	Asymptomatic-to-symptomatic (Delta)	0.021/day	Estimated
δ_2	Asymptomatic-to-symptomatic (Omicron)	0.061/day	Estimated
μ_{A_1}	Disease mortality (Delta asymptomatic)	0.00016/day	[15, 45]
μ_{A_2}	Disease mortality (Omicron asymptomatic)	0.00019/day	[15, 45]
μ_{I_1}	Disease mortality (Delta symptomatic)	0.0013/day	[14]
μ_{I_2}	Disease mortality (Omicron symptomatic)	0.0015/day	[20]
μ_j	Disease mortality (isolated)	0.00037/day	[43]
μ_h	Disease mortality (hospitalized)	0.0014/day	[14]
μ	Natural per-capita death rate	0.0000391/day	[39]
ω_V	Waning rate of vaccine-induced immunity	1/180 (i.e., 0.0055/day)	[17]
ω_R	Waning rate of natural immunity (faster)	1/60 (i.e., 0.0167/day)	[17]

The baseline vaccine efficacy values used in all analyses are $\xi_1 = 0.78$ for Delta and $\xi_2 = 0.70$ for Omicron, consistent with the meta-analytic estimates of [5, 6] for the vaccines deployed in India; all six variant-specific transmission parameters $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ are estimated from Indian surveillance data in Section 4.4.

Biological interpretation of the model

The compartmental structure directly encodes the clinical course of COVID-19: S and V distinguish unvaccinated and partially (non-sterilizingly) protected individuals, with waning terms ω_V, ω_R reflecting the transient nature of both vaccine- and infection-acquired immunity [5]; E_i represents the latent incubation period [38]; the split into A_i and I_i reflects the well-documented contribution of undocumented, subclinical carriers to transmission [37], with $\zeta_i < 1$ encoding their lower viral shedding; J and H represent the two levels of clinical/public-health response (isolation and inpatient care) and are treated as non-infectious owing to effective infection control; and R aggregates recovery from all infectious and clinical routes. Modeling Delta and Omicron as two parallel, non-co-infecting strain sub-systems is likewise biologically motivated, since the two variants co-circulated over largely non-overlapping periods in India with markedly different transmissibility, incubation, and severity profiles (see Section 1), which a single-strain model cannot represent.

3. Model analysis

3.1. Positivity and boundedness of solutions

Theorem 3.1. *All solutions of models (2.10)–(2.17) with non-negative initial conditions (2.18) are non-negative for all $t \geq 0$ and ultimately bounded, remaining in the positively invariant and attracting region*

$$\Omega = \left\{ (S, V, E_1, E_2, A_1, A_2, I_1, I_2, J, H, R) \in \mathbb{R}_+^{11} : \right. \\ \left. 0 \leq S + V + E_1 + E_2 + A_1 + A_2 + I_1 + I_2 + J + H + R \leq \frac{\Lambda}{\mu} \right\}.$$

Proof.

Non-negativity. We verify each state variable in turn.

Susceptible class $S(t)$. Define the integrating factor

$$P(t) = \sum_{i=1}^2 \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{1}{N} + (\nu + \mu) \geq 0.$$

Since $\omega_V V + \omega_R R \geq 0$, the S -equation of (2.10) satisfies

$$\frac{dS}{dt} \geq -P(t) S.$$

Multiplying by $\exp\left(\int_0^t P(s) ds\right) \geq 0$ and integrating gives

$$S(t) \geq S(0) \exp\left(-\int_0^t P(s) ds\right) \geq 0.$$

Vaccinated class $V(t)$. Define

$$Z(t) = \sum_{i=1}^2 (1 - \xi_i) \beta_i (\zeta_i A_i + \epsilon_i I_i) \frac{1}{N} + (\omega_V + \mu) \geq 0.$$

Since $\nu S \geq 0$, Eq (2.11) gives $dV/dt \geq -Z(t)V$, where

$$V(t) \geq V(0) \exp\left(-\int_0^t Z(s) ds\right) \geq 0.$$

Exposed classes $E_i(t)$, $i = 1, 2$. Both force-of-infection terms in Eq (2.12) are non-negative (the susceptible inflow $\beta_i(\cdots)S/N \geq 0$ and the vaccinated breakthrough inflow $(1 - \xi_i)\beta_i(\cdots)V/N \geq 0$), so

$$\frac{dE_i}{dt} \geq -(\alpha + \mu)E_i,$$

giving $E_i(t) \geq E_i(0)e^{-(\alpha+\mu)t} \geq 0$.

Remaining classes. Analogous lower-bound arguments applied to Eqs (2.13)–(2.17) yield $A_i(t) \geq 0$, $I_i(t) \geq 0$, $J(t) \geq 0$, $H(t) \geq 0$, and $R(t) \geq 0$ for all $t > 0$. Hence all eleven state variables remain non-negative for all $t \geq 0$.

Boundedness. Summing all equations of system (2.10)–(2.17) and using

$$N = S + V + \sum_i (E_i + A_i + I_i) + J + H + R,$$

all disease-induced mortality rates are non-negative, so

$$\frac{dN}{dt} = \Lambda - \mu N - \sum_{i=1}^2 \mu_{A_i} A_i - \sum_{i=1}^2 \mu_{I_i} I_i - \mu_j J - \mu_h H \leq \Lambda - \mu N.$$

The comparison principle applied to $dN/dt \leq \Lambda - \mu N$ yields

$$N(t) \leq \frac{\Lambda}{\mu} + \left(N(0) - \frac{\Lambda}{\mu}\right) e^{-\mu t}, \quad t \geq 0.$$

If $N(0) \leq \Lambda/\mu$, then $N(t) \leq \Lambda/\mu$ for all $t \geq 0$, so the solution remains in Ω . If $N(0) > \Lambda/\mu$, then $N(t)$ decreases monotonically and approaches Λ/μ asymptotically as $t \rightarrow \infty$; by the continuity of solutions, the trajectory enters Ω and remains there for all subsequent time. Hence Ω is positively invariant and attracting, and every solution with initial conditions satisfying (2.18) lies in Ω for all $t \geq 0$. $\square$

3.2. Disease-free equilibrium and basic reproduction number

Setting the right-hand side of system (2.10)–(2.17) equal to zero under the assumptions $E_i = A_i = I_i = 0$ ($i = 1, 2$) and $J = H = 0$ yields the unique disease-free equilibrium (DFE), denoted by F^0 :

$$F^0 = \left(\frac{\Lambda(\omega_V + \mu)}{\mu(\nu + \omega_V + \mu)}, \frac{\Lambda\nu}{\mu(\nu + \omega_V + \mu)}, 0, 0, 0, 0, 0, 0, 0, 0, 0 \right) =: (S^0, V^0, \mathbf{0}). \quad (3.1)$$

We use the term “DFE” consistently throughout; the notation F^0 is reserved for this state.

Next-generation matrix construction. The basic reproduction number $\mathcal{R}_0$ is derived via the next-generation matrix (NGM) method [23, 24]. We order the infected compartments as

$(E_1, A_1, I_1, E_2, A_2, I_2)$ and partition the right-hand sides into new-infection terms $\mathcal{F}$ and net-transition terms $\mathcal{V}$:

$$\mathcal{F} = \begin{pmatrix} \beta_1(\zeta_1 A_1 + \epsilon_1 I_1) \frac{S}{N} + (1 - \xi_1) \beta_1(\zeta_1 A_1 + \epsilon_1 I_1) \frac{V}{N} \\ 0 \\ 0 \\ \beta_2(\zeta_2 A_2 + \epsilon_2 I_2) \frac{S}{N} + (1 - \xi_2) \beta_2(\zeta_2 A_2 + \epsilon_2 I_2) \frac{V}{N} \\ 0 \\ 0 \end{pmatrix}, \quad \mathcal{V} = \begin{pmatrix} (\alpha + \mu) E_1 \\ -\eta_1 \alpha E_1 + a_1 A_1 \\ -d_1 E_1 - \delta_1 A_1 + b_1 I_1 \\ (\alpha + \mu) E_2 \\ -\eta_2 \alpha E_2 + a_2 A_2 \\ -d_2 E_2 - \delta_2 A_2 + b_2 I_2 \end{pmatrix},$$

where the shorthand parameters (used consistently throughout this section) are

$$\begin{aligned} e &= \alpha + \mu, \\ a_i &= \delta_i + \gamma_{A_i} + \mu_{A_i} + \mu, \\ b_i &= \kappa_i + \tau_i + \mu_{I_i} + \mu, \\ c_i &= \alpha \eta_i, \\ d_i &= \alpha(1 - \eta_i), \end{aligned} \tag{3.2}$$

for $i = 1, 2$. The Jacobian matrices $F = D\mathcal{F}(F^0)$ and $V = D\mathcal{V}(F^0)$, evaluated at the DFE using the effective transmission parameters

$$\begin{aligned} x_i &= \beta_i \zeta_i \frac{\omega_V + \mu + (1 - \xi_i) \nu}{\nu + \omega_V + \mu}, \\ y_i &= \beta_i \epsilon_i \frac{\omega_V + \mu + (1 - \xi_i) \nu}{\nu + \omega_V + \mu}, \end{aligned} \tag{3.3}$$

are obtained from $x_i = \beta_i \zeta_i (S^0 + (1 - \xi_i) V^0) / N^0$ (and analogously for y_i) using the DFE (3.1)

$$S^0 = \frac{\Lambda(\omega_V + \mu)}{\mu(\nu + \omega_V + \mu)}, \quad V^0 = \frac{\Lambda \nu}{\mu(\nu + \omega_V + \mu)}, \quad N^0 = \frac{\Lambda}{\mu}.$$

Thus, the vaccine-waning rate ω_V enters the effective transmission factors (3.3) through the DFE.

These matrices are

$$F = \begin{pmatrix} 0 & x_1 & y_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x_2 & y_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} e & 0 & 0 & 0 & 0 & 0 \\ -c_1 & a_1 & 0 & 0 & 0 & 0 \\ -d_1 & -\delta_1 & b_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & e & 0 & 0 \\ 0 & 0 & 0 & -c_2 & a_2 & 0 \\ 0 & 0 & 0 & -d_2 & -\delta_2 & b_2 \end{pmatrix}.$$

The next-generation matrix FV^{-1} has block-diagonal structure and its non-zero eigenvalues are $\mathcal{R}_{01}$ and $\mathcal{R}_{02}$ (with 0 of multiplicity 4). The spectral radius therefore yields $\mathcal{R}_0 = \max\{\mathcal{R}_{01}, \mathcal{R}_{02}\}$, where for $i = 1$ (Delta) and $i = 2$ (Omicron):

$$\mathcal{R}_{0i} = \frac{\alpha \eta_i \beta_i \zeta_i (\omega_V + \mu + (1 - \xi_i) \nu)}{a_i e (\nu + \omega_V + \mu)} + \frac{\beta_i \epsilon_i (\omega_V + \mu + (1 - \xi_i) \nu) (a_i d_i + c_i \delta_i)}{a_i b_i e (\nu + \omega_V + \mu)}. \tag{3.4}$$

Using the identity $a_i d_i + c_i \delta_i = \alpha(a_i - \eta_i(\gamma_{A_i} + \mu_{A_i} + \mu))$, Eq (3.4) can be written in fully expanded form as

$$\mathcal{R}_{0i} = \underbrace{\frac{\alpha \eta_i \beta_i \zeta_i (\omega_V + \mu + (1 - \xi_i) \nu)}{(\alpha + \mu) (\nu + \omega_V + \mu) (\delta_i + \gamma_{A_i} + \mu_{A_i} + \mu)}}_{\text{asymptomatic pathway}} + \underbrace{\frac{\alpha \beta_i \epsilon_i (\omega_V + \mu + (1 - \xi_i) \nu) (\delta_i + \gamma_{A_i} + \mu_{A_i} + \mu - \eta_i (\gamma_{A_i} + \mu_{A_i} + \mu))}{(\alpha + \mu) (\nu + \omega_V + \mu) (\delta_i + \gamma_{A_i} + \mu_{A_i} + \mu) (\kappa_i + \tau_i + \mu_{I_i} + \mu)}}_{\text{symptomatic pathway}}. \quad (3.5)$$

The two terms in (3.5) represent the contributions to variant- i secondary infections via the asymptomatic and symptomatic infectious pathways, respectively. The factor $\mu + (1 - \xi_i) \nu$ in the numerator captures the combined effect of unvaccinated and vaccinated (breakthrough) susceptibility: when $\xi_i = 1$ (perfect vaccine), this factor reduces to μ and $\mathcal{R}_{0i}$ attains its minimum over all $\xi_i \in [0, 1]$ for fixed remaining parameters. The overall basic reproduction number is

$$\mathcal{R}_0 = \max\{\mathcal{R}_{01}, \mathcal{R}_{02}\}. \quad (3.6)$$

Remark on terminology. The quantity $\mathcal{R}_{0i}$ derived in (3.4) and (3.5) is evaluated at the disease-free equilibrium (3.1), which already incorporates an ongoing baseline vaccination program (ν, ξ_i) and case-isolation regime (κ_i, τ_i) . It is therefore, in the terminology of [23], a controlled (intervention-adjusted) reproduction number, not the classical basic reproduction number of an entirely susceptible, unmitigated population—for which published Delta estimates typically range from 4 to 9 [46]. We retain the symbol $\mathcal{R}_{0i}$ and name for consistency with the next-generation-matrix construction and the stability results of Sections 3.1–3.5, which hold regardless of parameterization; the value $\mathcal{R}_{01} = 1.87$ should be read as the reproduction number under baseline Indian intervention conditions, not as Delta’s unmitigated R_0 .

Numerical verification. Equation (3.5) is independent of Λ , consistent with $\mathcal{R}_{0i}$ being a dimensionless per-generation quantity, and matches the equivalent, independently derived expression

$$\mathcal{R}_{01} = \mathcal{M}_1 \alpha \Phi_1 (\mu + (1 - \xi_1) \nu) / [e a_1 b_1 (\nu + \omega_V + \mu)]$$

of Section 3.5 (Eq (3.10)), providing an internal cross-check. Concretely, India administered on the order of 4–5 million first doses per day at the peak of the 2021 campaign against a susceptible pool on the order of 1.3–1.4 billion [39, 47], giving a per-capita rate $\nu \approx (4\text{--}5 \times 10^6) / (1.3\text{--}1.4 \times 10^9) \approx 0.003\text{--}0.004/\text{day}$, i.e., $\approx 0.3\text{--}0.4\%$ of susceptibles vaccinated per day—consistent with, and slightly above, the commonly cited 0.1–0.3%/day range for India’s peak rollout. Guided by this empirically grounded range, we adopt $\nu = 0.0033/\text{day}$ throughout (Table 1), taken directly from the daily vaccination throughput reported for India [39]. As an independent numerical check, direct substitution of the baseline values in Table 1 into (3.5) gives $\mathcal{R}_{01} = 1.87$ and $\mathcal{R}_{02} = 1.64$ (rounded to three decimal places).

3.3. Local stability of the disease-free equilibrium

Theorem 3.2. *The disease-free equilibrium F^0 of model (2.10)–(2.17) is locally asymptotically stable (LAS) whenever $\mathcal{R}_0 < 1$, and is unstable whenever $\mathcal{R}_0 > 1$.*

Proof. Linearizing system (2.10)–(2.17) about $F^0 = (S^0, V^0, \mathbf{0})$ yields an 11×11 Jacobian J_{F^0} . Using the notation of (3.2) and (3.3) and setting

$$p = \pi + \gamma_j + \mu_j + \mu, \quad q = \gamma_h + \mu_h + \mu, \quad \tilde{x}_i = (1 - \xi_i) x_i \frac{V^0}{S^0}, \quad \tilde{y}_i = (1 - \xi_i) y_i \frac{V^0}{S^0},$$

the Jacobian in state ordering $(S, V, E_1, A_1, I_1, E_2, A_2, I_2, J, H, R)$ is

$$J_{F^0} = \begin{pmatrix} -(v + \mu) & \omega_V & 0 & -x_1 & -y_1 & 0 & -x_2 & -y_2 & 0 & 0 & \omega_R \\ v & -(\omega_V + \mu) & 0 & -\tilde{x}_1 & -\tilde{y}_1 & 0 & -\tilde{x}_2 & -\tilde{y}_2 & 0 & 0 & 0 \\ 0 & 0 & -e & x_1 & y_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_1 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & d_1 & \delta_1 & -b_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -e & x_2 & y_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_2 & -a_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & d_2 & \delta_2 & -b_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \kappa_1 & 0 & 0 & \kappa_2 & -p & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_1 & 0 & 0 & \tau_2 & \pi & -q & 0 \\ 0 & 0 & 0 & \gamma_{A_1} & 0 & 0 & \gamma_{A_2} & 0 & \gamma_j & \gamma_h & -(\omega_R + \mu) \end{pmatrix}.$$

The matrix J_{F^0} is block lower-triangular in the natural partition of infected and uninfected states, so its eigenvalues are the union of those of the following diagonal blocks.

Uninfected block M_{SV} . The 2×2 block

$$M_{SV} = \begin{pmatrix} -(v + \mu) & \omega_V \\ v & -(\omega_V + \mu) \end{pmatrix}$$

has $\text{tr}(M_{SV}) = -(v + 2\mu + \omega_V) < 0$ and $\det(M_{SV}) = \mu(v + \omega_V + \mu) > 0$, so both its eigenvalues have strictly negative real parts.

Scalar blocks. The blocks for J, H, R yield eigenvalues $\lambda_J = -p < 0$, $\lambda_H = -q < 0$, $\lambda_R = -(\omega_R + \mu) < 0$. Five of the eleven eigenvalues are therefore stable.

Infection blocks $J_i, i = 1, 2$. The remaining six eigenvalues arise from

$$J_i = \begin{pmatrix} -e & x_i & y_i \\ c_i & -a_i & 0 \\ d_i & \delta_i & -b_i \end{pmatrix}, \quad i = 1, 2.$$

The characteristic polynomial of J_i is $\lambda^3 + p_2^{(i)}\lambda^2 + p_1^{(i)}\lambda + p_0^{(i)} = 0$ with Routh-Hurwitz coefficients

$$p_2^{(i)} = e + a_i + b_i > 0, \tag{3.7}$$

$$p_1^{(i)} = a_i b_i + (a_i + b_i)e - c_i x_i - d_i y_i, \tag{3.8}$$

$$p_0^{(i)} = a_i b_i e (1 - \mathcal{R}_{0i}), \tag{3.9}$$

where

$$\mathcal{R}_{0i} = \frac{b_i c_i x_i + y_i (a_i d_i + c_i \delta_i)}{a_i b_i e}$$

is the variant- i reproduction number derived in Section 3.2.

Case $\mathcal{R}_0 < 1$. Assume $\mathcal{R}_{0i} < 1$ for $i = 1, 2$. Then $p_2^{(i)} > 0$ and $p_0^{(i)} = a_i b_i e(1 - \mathcal{R}_{0i}) > 0$ from (3.9). For $p_1^{(i)}$, observe that

$$c_i x_i + d_i y_i \leq \frac{b_i c_i x_i + a_i d_i y_i}{\min(a_i, b_i)} \leq \frac{a_i b_i e \mathcal{R}_{0i}}{\min(a_i, b_i)} \leq (a_i + b_i)e,$$

so $p_1^{(i)} > a_i b_i > 0$. The discriminant of the Routh-Hurwitz table is $\Delta^{(i)} = p_2^{(i)} p_1^{(i)} - p_0^{(i)}$. Using $c_i x_i < a_i e \mathcal{R}_{0i}$ and $d_i y_i < b_i e \mathcal{R}_{0i}$, the negative contribution $-p_0^{(i)}$ to $\Delta^{(i)}$ is dominated by the positive terms $a_i^2 b_i + a_i b_i^2 + 2a_i b_i e + c_i d_i y_i$ arising from $p_2^{(i)} p_1^{(i)}$, yielding $\Delta^{(i)} > 0$. By the Routh-Hurwitz criterion, all three roots of each cubic have strictly negative real parts. Combined with the five trivially stable eigenvalues, F^0 is LAS.

Case $\mathcal{R}_0 > 1$. If $\mathcal{R}_{0k} > 1$ for some $k \in \{1, 2\}$, then $p_0^{(k)} < 0$ while $p_2^{(k)} > 0$. By Descartes' rule of signs, the characteristic polynomial of J_k has at least one positive real root, so F^0 is unstable. $\square$

3.4. Conditional global stability of the disease-free equilibrium

Theorem 3.3. *The disease-free equilibrium F^0 of model (2.10)–(2.17) is globally asymptotically stable (GAS) in Ω whenever $\mathcal{R}_0 < 1$ and the Castillo-Chavez-Song domination condition (C2) below holds.*

Proof. We apply the Castillo-Chavez-Song decomposition [25]. We partition the state vector into uninfected and infected components,

$$P = (S, V, R) \in \mathbb{R}_+^3, \quad Q = (E_1, A_1, I_1, E_2, A_2, I_2, J, H) \in \mathbb{R}_+^8,$$

so that system (2.10)–(2.17) takes the standard form

$$\frac{dP}{dt} = M(P, Q), \quad \frac{dQ}{dt} = \mathcal{W}(P, Q), \quad \mathcal{W}(P, \mathbf{0}) = \mathbf{0},$$

with disease-free equilibrium $P^* = (S^0, V^0, 0)$, as given in (3.1), of the uninfected sub-system. Global stability of F^0 follows once the two Castillo-Chavez-Song conditions are verified.

(C1) *Global stability of the disease-free sub-system.*

Setting $Q = \mathbf{0}$ in the P -equations yields the linear system

$$\dot{S} = \Lambda - (\nu + \mu)S + \omega_V V + \omega_R R, \quad \dot{V} = \nu S - (\omega_V + \mu)V, \quad \dot{R} = -(\omega_R + \mu)R.$$

This system is not an uncoupled cascade: The waning-immunity terms $\omega_V V$ and $\omega_R R$ in the $\dot{S}$ -equation, inherited from the full model (2.10), couple S to V and R even when $Q = \mathbf{0}$. Writing the system as $\dot{P} = AP + b$ with

$$A = \begin{pmatrix} -(\nu + \mu) & \omega_V & \omega_R \\ \nu & -(\omega_V + \mu) & 0 \\ 0 & 0 & -(\omega_R + \mu) \end{pmatrix}, \quad b = \begin{pmatrix} \Lambda \\ 0 \\ 0 \end{pmatrix},$$

we show A is Hurwitz. Because the third row and column of A decouple R from (S, V) ,

$$\text{spec}(A) = \text{spec}(A') \cup \{-(\omega_R + \mu)\},$$

where

$$A' = \begin{pmatrix} -(\nu + \mu) & \omega_V \\ \nu & -(\omega_V + \mu) \end{pmatrix}.$$

Since $\text{tr}(A') = -(\nu + \mu) - (\omega_V + \mu) < 0$ and

$$\det(A') = (\nu + \mu)(\omega_V + \mu) - \omega_V \nu = \mu(\nu + \omega_V + \mu) > 0,$$

both eigenvalues of A' have a strictly negative real part; together with $-(\omega_R + \mu) < 0$, every eigenvalue of A has a strictly negative real part, so A is Hurwitz. Consequently the affine system $\dot{P} = AP + b$ has a unique equilibrium $P^* = -A^{-1}b$, and every solution satisfies $P(t) - P^* = e^{At}(P(0) - P^*) \rightarrow \mathbf{0}$ as $t \rightarrow \infty$, so P^* is GAS in the P -subspace regardless of initial datum. Solving $AP^* = -b$ explicitly gives

$$S^* = \frac{\Lambda(\omega_V + \mu)}{\mu(\nu + \omega_V + \mu)}, \quad V^* = \frac{\Lambda\nu}{\mu(\nu + \omega_V + \mu)}, \quad R^* = 0,$$

which coincides with $(S^0, V^0, 0)$ of the disease-free equilibrium (3.1), confirming that the P -subsystem's steady state and the model's DFE agree. Hence condition (C1) holds.

(C2) *Castillo-Chavez-Song domination condition and Metzler structure.*

Write $\mathcal{W}(P, Q) = KQ - \bar{W}(P, Q)$, where $K = D_Q \mathcal{W}(P^*, \mathbf{0})$ is

$$K = \begin{pmatrix} -e & x_1 & y_1 & 0 & 0 & 0 & 0 & 0 \\ c_1 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ d_1 & \delta_1 & -b_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -e & x_2 & y_2 & 0 & 0 \\ 0 & 0 & 0 & c_2 & -a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_2 & \delta_2 & -b_2 & 0 & 0 \\ 0 & 0 & \kappa_1 & 0 & 0 & \kappa_2 & -p & 0 \\ 0 & 0 & \tau_1 & 0 & 0 & \tau_2 & \pi & -q \end{pmatrix}.$$

The Castillo-Chavez-Song condition (C2) requires the componentwise domination inequality $\mathcal{W}(P, Q) \leq KQ$ (equivalently, $\bar{W}(P, Q) \geq \mathbf{0}$) throughout the positively invariant region. The inequality $S + V \leq N$ alone does not establish this condition, because the force-of-infection coefficient is $[S + (1 - \xi_i)V]/N$, whereas its DFE value is $[S^0 + (1 - \xi_i)V^0]/N^0$. We therefore state (C2) explicitly as an additional verifiable model condition rather than infer it from $S + V \leq N$. Every off-diagonal entry of K is non-negative, so K is a Metzler matrix.

If $\mathcal{R}_0 < 1$, each three-dimensional infection block J_i is Hurwitz by Theorem 3.2. Since K is block lower-triangular with these infection blocks and the strictly stable J and H rows, K is Hurwitz. Under the stated domination condition, the standard Castillo-Chavez-Song theorem therefore yields $Q(t) \rightarrow \mathbf{0}$ as $t \rightarrow \infty$ and condition (C1) then gives $P(t) \rightarrow P^*$.

Both conditions hold, so F^0 is GAS in Ω for $\mathcal{R}_0 < 1$ subject to the stated domination condition (C2). The borderline case $\mathcal{R}_0 = 1$ is not asserted here because the present argument gives only nonpositive spectral abscissa at the threshold. $\square$

3.5. Endemic equilibria

When $\mathcal{R}_0 > 1$, the disease persists at a non-trivial steady state called an endemic equilibrium (EE). Depending on which variant's reproduction number exceeds unity, the model may admit a Delta-only

EE, an Omicron-only EE, or a coexistence EE in which both variants circulate simultaneously. We establish the existence, uniqueness, and stability of these states below.

Setting the right-hand side of system (2.10)–(2.17) to zero, we use the shorthand (3.2) and additionally define

$$p = \pi + \gamma_j + \mu_j + \mu, \quad q = \gamma_h + \mu_h + \mu,$$

and the composite parameters

$$\begin{aligned} \Phi_i &= (1 - \eta_i)a_i + \eta_i\delta_i = a_i - \eta_i(\gamma_{A_i} + \mu_{A_i} + \mu) > 0, \\ \Theta_i &= \frac{\zeta_i b_i \eta_i + \epsilon_i \Phi_i}{\Phi_i}, \\ \mathcal{M}_i &= \beta_i \Theta_i, \end{aligned} \tag{3.10}$$

for $i = 1, 2$. Let

$$\chi_i^* = \beta_i(\zeta_i A_i^* + \epsilon_i I_i^*)/N^*$$

denote the endemic force of infection of variant i .

From the equilibrium conditions of A_i and I_i , one obtains, after eliminating E_i^* :

$$E_i^* = \frac{a_i b_i}{\alpha \Phi_i} I_i^*, \quad A_i^* = \frac{b_i \eta_i}{\Phi_i} I_i^*, \quad i = 1, 2. \tag{3.11}$$

The equilibrium conditions for J , H , and R then give

$$J^* = \frac{\kappa_1 I_1^* + \kappa_2 I_2^*}{p}, \tag{3.12}$$

$$H^* = \frac{(\tau_1 p + \pi \kappa_1) I_1^* + (\tau_2 p + \pi \kappa_2) I_2^*}{pq}, \tag{3.13}$$

$$R^* = \frac{1}{\omega_R + \mu} \left[\sum_{i=1}^2 \frac{\gamma_{A_i} b_i \eta_i}{\Phi_i} I_i^* + \frac{\gamma_j(\kappa_1 I_1^* + \kappa_2 I_2^*)}{p} + \frac{\gamma_h[(\tau_1 p + \pi \kappa_1) I_1^* + (\tau_2 p + \pi \kappa_2) I_2^*]}{pq} \right]. \tag{3.14}$$

Substituting A_i^* from (3.11) gives $\chi_i^* = \mathcal{M}_i I_i^*/N^*$. Setting $\chi^* = \chi_1^* + \chi_2^*$, the equilibrium S - and V -equations yield

$$\begin{aligned} S^* &= \frac{\Lambda + \omega_V V^* + \omega_R R^*}{\chi^* + \nu + \mu}, \\ V^* &= \frac{\nu S^*}{(1 - \xi_1)\chi_1^* + (1 - \xi_2)\chi_2^* + \omega_V + \mu}. \end{aligned} \tag{3.15}$$

Substituting into the equilibrium E_i -equation gives the governing system for the endemic prevalences:

$$\chi_i^*(S^* + (1 - \xi_i)V^*) = e \frac{a_i b_i}{\alpha \Phi_i} I_i^*, \quad i = 1, 2, \tag{3.16}$$

where S^* and V^* are implicit functions of (I_1^*, I_2^*) through (3.15).

Single-variant equilibria

For the Delta-only case, set $I_2^* = A_2^* = E_2^* = 0$ and define $\lambda = \chi_1^* > 0$ as the primary unknown. Define the aggregate infected-class coefficient

$$\mathcal{K}_1 = \frac{a_1 b_1}{\alpha \Phi_1} + \frac{b_1 \eta_1}{\Phi_1} + 1 + \frac{\kappa_1}{p} + \frac{\tau_1 p + \pi \kappa_1}{pq} + \frac{1}{\omega_R + \mu} \left[\frac{\gamma_{A_1} b_1 \eta_1}{\Phi_1} + \frac{\gamma_j \kappa_1}{p} + \frac{\gamma_h (\tau_1 p + \pi \kappa_1)}{pq} \right],$$

so that $N^* = S^* + V^* + \mathcal{K}_1 I_1^*$. Let $\sigma_1 = (1 - \xi_1)\lambda + \omega_V + \mu$. From (3.15) and $I_1^* = \lambda N^* / \mathcal{M}_1$, Eq (3.16) with $i = 1$ reduces to the consistency function

$$g_1(\lambda) := \mathcal{M}_1[\sigma_1 + (1 - \xi_1)v] - \mathcal{K}_1 \lambda [\sigma_1 + (1 - \xi_1)v] - \frac{e a_1 b_1}{\alpha \Phi_1} (\sigma_1 + v) = 0. \quad (3.17)$$

Expanding in λ , g_1 is a quadratic with leading coefficient $-\mathcal{K}_1(1 - \xi_1) \leq 0$ (downward-opening, or decreasing linear when $\xi_1 = 1$). Its value at $\lambda = 0$ is

$$g_1(0) = \mathcal{M}_1((\omega_V + \mu + v)(1 - \xi_1)) - \frac{e a_1 b_1}{\alpha \Phi_1} (\omega_V + \mu + v),$$

which satisfies $g_1(0) > 0$ whenever $\mathcal{R}_{01} > 1$ (established in Theorem 3.4 below). Since $g_1(\lambda) \rightarrow -\infty$ as $\lambda \rightarrow +\infty$, the intermediate value theorem guarantees at least one positive root; downward concavity gives uniqueness. An analogous argument applies to the Omicron-only equilibrium with indices $1 \leftrightarrow 2$.

Theorem 3.4. *Model (2.10)–(2.17) admits:*

- (i) *a unique Delta-only endemic equilibrium $F_1^* = (S^*, V^*, E_1^*, A_1^*, I_1^*, 0, 0, 0, J^*, H^*, R^*)$ with all components positive if and only if $\mathcal{R}_{01} > 1$;*
- (ii) *a unique Omicron-only endemic equilibrium $F_2^* = (S^*, V^*, 0, 0, 0, E_2^*, A_2^*, I_2^*, J^*, H^*, R^*)$ with all components positive if and only if $\mathcal{R}_{02} > 1$.*

Proof. We prove part (i) in full; part (ii) follows by symmetry.

Existence and positivity. Given $I_1^* > 0$, Eqs (3.11)–(3.14) yield $E_1^*, A_1^*, J^*, H^*, R^* > 0$ since all parameters are positive. With $I_2^* = E_2^* = A_2^* = 0$, we have $\chi_2^* = 0$, $\chi^* = \chi_1^* > 0$, and then $S^* > 0$, $V^* > 0$ from (3.15).

Using

$$\mathcal{R}_{01} = \frac{\mathcal{M}_1 \alpha \Phi_1 (\mu + (1 - \xi_1)v)}{e a_1 b_1 (v + \omega_V + \mu)}$$

and the identity

$$\frac{e a_1 b_1}{\alpha \Phi_1} = \frac{\mathcal{M}_1 (\mu + (1 - \xi_1)v)}{(v + \omega_V + \mu) \mathcal{R}_{01}},$$

we obtain

$$g_1(0) = \frac{\mathcal{M}_1 (\mathcal{R}_{01} - 1)}{\mathcal{R}_{01}} \left[1 + \frac{\omega_V}{\mu + (1 - \xi_1)v} \right] > 0, \quad \text{whenever } \mathcal{R}_{01} > 1.$$

Combined with the downward-opening quadratic form of g_1 (Eq (3.17)) and $g_1(\lambda) \rightarrow -\infty$ as $\lambda \rightarrow \infty$, the intermediate value theorem guarantees a unique positive root $\lambda^* > 0$. Back-substitution via (3.11)–(3.15) recovers all components positively.

Non-existence when $\mathcal{R}_{01} \leq 1$. The same calculation gives $g_1(0) \leq 0$; since g_1 is non-increasing for $\lambda > 0$, it has no positive root, so no positive EE with $I_1^* > 0$ exists. $\square$

Theorem 3.4 is deliberately restricted to the two single-variant equilibria F_1^*, F_2^* , both of which are proved rigorously above. The coexistence case is stated below as a conjecture rather than a theorem, since only a sign-change/monotonicity argument—not a full simultaneous existence-uniqueness proof—is currently available for it.

Conjecture 1. Whenever $\mathcal{R}_0 = \max\{\mathcal{R}_{01}, \mathcal{R}_{02}\} > 1$ and neither variant's force of infection is zero, model (2.10)–(2.17) admits a coexistence endemic equilibrium

$$F^* = (S^*, V^*, E_1^*, A_1^*, I_1^*, E_2^*, A_2^*, I_2^*, J^*, H^*, R^*)$$

with all components positive.

Supporting argument (not a proof). The coupled system (3.16) for (I_1^*, I_2^*) consists of two equations, each of which, fixing the other variable at a positive value, defines a consistency function of the same quadratic form as g_1 . When $\mathcal{R}_0 > 1$, the same sign-change argument applied simultaneously to both equations—using the monotone dependence of S^* and V^* on the total force of infection χ^* —suggests a positive solution (I_1^*, I_2^*) , with uniqueness suggested by the strict monotonicity and downward-opening structure of each consistency function when the competing variant's prevalence is held fixed. This is a sign-change/monotonicity sketch, not a fully rigorous simultaneous-equation existence proof (e.g., via a Poincaré-Miranda or Brouwer fixed-point argument, or a monotone-operator argument on the coupled map $(I_1^*, I_2^*) \mapsto (I_1^*, I_2^*)$); such a treatment remains outstanding (see section 5). Theorem 3.5 concerns only the single-variant equilibria F_1^*, F_2^* established rigorously in Theorem 3.4 (i) and (ii); their validity does not depend on Conjecture 1.

3.6. Variant elimination and competitive exclusion

Definition 1. Variant $i \in \{1, 2\}$ is said to be eliminated from the population if

$$(E_i(t), A_i(t), I_i(t)) \longrightarrow (0, 0, 0), \quad \text{as } t \rightarrow \infty,$$

regardless of the asymptotic behavior of the other variant's compartments, the shared compartments J, H, R , or the susceptible/vaccinated classes S, V . Under model (2.10)–(2.17), the following two regimes are distinguished in the analysis:

(a) *Global elimination (no circulating strain).* If $\mathcal{R}_{0i} < 1$ and $\mathcal{R}_{0j} < 1$ for $j \neq i$, then by Theorem 3.3, both variants are eliminated simultaneously, since the disease-free equilibrium F^0 is globally asymptotically stable in Ω under condition (C2), and every trajectory satisfies $(E_i, A_i, I_i) \rightarrow (0, 0, 0)$ for both $i = 1, 2$.

(b) *Competitive elimination (single-strain persistence).* If $\mathcal{R}_{0i} > 1$ for exactly one i , variant i persists at the single-variant endemic equilibrium F_i^* (Theorem 3.4). The competing variant $j \neq i$ is locally eliminated near F_i^* when the corresponding invasion block is Hurwitz. Equivalently, its invasion reproduction number satisfies $\tilde{\mathcal{R}}_{0j} < 1$ together with the remaining cubic Routh-Hurwitz inequalities stated in Theorem 3.5. The ordering $\mathcal{R}_{0j} < \mathcal{R}_{0i}$ may provide intuition for which variant is favored, but it is not by itself sufficient to establish non-invasion because $\tilde{\mathcal{R}}_{0j}$ is evaluated at the endemic state of the resident variant. This elimination conclusion is therefore local, not global.

If $\mathcal{R}_{01} = \mathcal{R}_{02} > 1$, the ordering argument provides no competitive conclusion. The coexistence equilibrium F^* of Conjecture 1 may then be relevant, but its existence and stability are not established in the present work. This borderline case is not analyzed further.

3.7. Local stability of endemic equilibria

Theorem 3.5. *Each single-variant endemic equilibrium F_i^* , $i \in \{1, 2\}$, of model (2.10)–(2.17) is locally asymptotically stable provided that $\mathcal{R}_{0i} > 1$, the active endemic Jacobian block $\mathbf{B}_{8,i}$ satisfies the full eighth-order Routh-Hurwitz conditions $\Delta_k^{(i)} > 0$ for $k = 1, \dots, 8$, and the direct invasion block of the competing variant is Hurwitz. The latter condition is checked explicitly from the invasion reproduction number and the corresponding cubic Routh-Hurwitz inequalities below.*

Proof. We prove the result for F_1^* ; the case F_2^* follows by interchanging indices $1 \leftrightarrow 2$.

At F_1^* , the Omicron compartments vanish, $E_2^* = A_2^* = I_2^* = 0$, so the 11×11 Jacobian $J_{F_1^*}$ is block upper-triangular:

$$J_{F_1^*} = \begin{pmatrix} \mathbf{B}_8 & * \\ \mathbf{0} & \mathbf{C}_3 \end{pmatrix},$$

where $\mathbf{B}_8$ governs the active sub-system $(S, V, E_1, A_1, I_1, J, H, R)$ and $\mathbf{C}_3$ governs the linearized Omicron invasion sub-system (E_2, A_2, I_2) . Hence

$$\text{spec}(J_{F_1^*}) = \text{spec}(\mathbf{B}_8) \cup \text{spec}(\mathbf{C}_3).$$

Utilizing

$$\begin{aligned} \hat{x}_1 &= \frac{\beta_1 \zeta_1 S^*}{N^*}, \\ \hat{y}_1 &= \frac{\beta_1 \epsilon_1 S^*}{N^*}, \\ \hat{x}_1^V &= \frac{(1 - \xi_1) \beta_1 \zeta_1 V^*}{N^*}, \\ \hat{y}_1^V &= \frac{(1 - \xi_1) \beta_1 \epsilon_1 V^*}{N^*}, \\ \chi_1^{V*} &= (1 - \xi_1) \chi_1^*, \end{aligned}$$

the active block is

$$\mathbf{B}_8 = \begin{pmatrix} -(\chi_1^* + \nu + \mu) & \omega_V & 0 & -\hat{x}_1 & -\hat{y}_1 & 0 & 0 & \omega_R \\ \nu & -(\chi_1^{V*} + \omega_V + \mu) & 0 & -\hat{x}_1^V & -\hat{y}_1^V & 0 & 0 & 0 \\ \chi_1^* & \chi_1^{V*} & -e & \hat{x}_1 + \hat{x}_1^V & \hat{y}_1 + \hat{y}_1^V & 0 & 0 & 0 \\ 0 & 0 & c_1 & -a_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & d_1 & \delta_1 & -b_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \kappa_1 & -p & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_1 & \pi & -q & 0 \\ 0 & 0 & 0 & \gamma_{A_1} & 0 & \gamma_j & \gamma_h & -(\omega_R + \mu) \end{pmatrix}.$$

Routh-Hurwitz analysis of the active endemic block. Let

$$\mathcal{P}_8(\lambda) = \det(\lambda I_8 - \mathbf{B}_8) = \lambda^8 + c_1 \lambda^7 + c_2 \lambda^6 + c_3 \lambda^5 + c_4 \lambda^4 + c_5 \lambda^3 + c_6 \lambda^2 + c_7 \lambda + c_8.$$

For $k = 1, \dots, 8$, define the Hurwitz determinants

$$\Delta_k = \det[h_{rs}]_{r,s=1}^k, \quad h_{rs} = c_{2s-r},$$

where $c_0 = 1$ and $c_\ell = 0$ for $\ell < 0$ or $\ell > 8$. The Routh-Hurwitz theorem states that $\mathbf{B}_8$ is Hurwitz if and only if

$$\Delta_1 > 0, \quad \Delta_2 > 0, \quad \dots, \quad \Delta_8 > 0.$$

For completeness, the first determinant is explicit:

$$\Delta_1 = c_1 = -\text{tr}(\mathbf{B}_8) = (\chi_1^* + \nu + \mu) + (\chi_1^{V*} + \omega_V + \mu) + e + a_1 + b_1 + p + q + (\omega_R + \mu) > 0.$$

No conclusion that $\Delta_2, \dots, \Delta_7$ are automatically positive is drawn from $\mathcal{R}_{01} > 1$ alone. Their positivity is part of the complete Routh-Hurwitz verification:

$$\Delta_k > 0, \quad k = 1, \dots, 8.$$

Once these inequalities are verified for the parameter set under consideration, all eight eigenvalues of $\mathbf{B}_8$ have strictly negative real parts.

Direct invasion analysis of the competing variant. The linearized Omicron invasion block is

$$\mathbf{C}_3 = \begin{pmatrix} -e & \tilde{x}_2 & \tilde{y}_2 \\ c_2 & -a_2 & 0 \\ d_2 & \delta_2 & -b_2 \end{pmatrix},$$

where

$$\tilde{x}_2 = \beta_2 \zeta_2 \frac{S^* + (1 - \xi_2)V^*}{N^*}, \quad \tilde{y}_2 = \beta_2 \epsilon_2 \frac{S^* + (1 - \xi_2)V^*}{N^*}.$$

Define the invasion reproduction number at F_1^* by

$$\widetilde{\mathcal{R}}_{02} = \frac{c_2 \tilde{x}_2}{e a_2} + \frac{(a_2 d_2 + c_2 \delta_2) \tilde{y}_2}{e a_2 b_2}.$$

The characteristic polynomial of $\mathbf{C}_3$ is

$$\lambda^3 + \tilde{p}_2 \lambda^2 + \tilde{p}_1 \lambda + \tilde{p}_0,$$

where

$$\begin{aligned} \tilde{p}_2 &= e + a_2 + b_2 > 0, \\ \tilde{p}_1 &= a_2 b_2 + (a_2 + b_2)e - c_2 \tilde{x}_2 - d_2 \tilde{y}_2, \\ \tilde{p}_0 &= e a_2 b_2 (1 - \widetilde{\mathcal{R}}_{02}). \end{aligned}$$

Therefore, the exact cubic Routh-Hurwitz conditions for $\mathbf{C}_3$ to be Hurwitz are

$$\widetilde{\mathcal{R}}_{02} < 1, \quad \tilde{p}_1 > 0, \quad \tilde{p}_2 \tilde{p}_1 > \tilde{p}_0.$$

These inequalities, rather than the ordering $\mathcal{R}_{02} < \mathcal{R}_{01}$ by itself, provide the direct non-invasion test at F_1^* .

Hence, if $\mathcal{R}_{01} > 1$, the active-block Hurwitz determinants satisfy $\Delta_k > 0$ for $k = 1, \dots, 8$, and the three direct invasion Routh-Hurwitz inequalities above hold, so every eigenvalue of $J_{F_1^*}$ has a strictly negative real part. Thus F_1^* is locally asymptotically stable. The proof for F_2^* follows by interchanging the variant indices. $\square$

Remark 3.1. For the single-variant endemic equilibrium F_1^* , a natural candidate Lyapunov function for the Omicron sub-block produces the term

$$\Psi(E_2, A_2, I_2) = \left(\frac{x_2 c_2}{a_2} + \frac{y_2 (d_2 a_2 + c_2 \delta_2)}{a_2 b_2} \right) \left(\frac{S}{N} + (1 - \xi_2) \frac{V}{N} \right) (A_2 + I_2). \quad (3.20)$$

All factors in (3.20) are non-negative on Ω ; hence

$$\Psi(E_2, A_2, I_2) \geq 0,$$

with strict positivity whenever $A_2 + I_2 > 0$ and the transmission parameters are positive. Moreover,

$$\frac{x_2 c_2}{a_2} + \frac{y_2 (d_2 a_2 + c_2 \delta_2)}{a_2 b_2} = e \mathcal{R}_{02} > 0. \quad (3.18)$$

The condition $\mathcal{R}_{02} < 1$ therefore does not change the sign of this term: $\Psi \geq 0$ holds regardless of whether $\mathcal{R}_{02}$ is above or below unity. Consequently, $\dot{V} \leq 0$ cannot be established from (3.20) alone, and LaSalle's invariance principle cannot be invoked on the basis of this Lyapunov function.

For this reason, we do not claim global asymptotic stability of F_i^* under $\mathcal{R}_{0i} > 1$ and $\mathcal{R}_{0j} < 1$. The rigorous stability result established in this manuscript is Theorem 3.5, which gives local asymptotic stability of the single-variant endemic equilibrium under the corresponding non-invasion condition. A global asymptotic-stability result would require a different Lyapunov function or an independent global convergence argument, which we leave for future work.

4. Numerical simulation

In India, the initial phase of the COVID-19 pandemic was driven by the original variant of SARS-CoV-2, which appeared in early 2020 and reached its peak during the first wave. The Beta variant (B.1.351), identified in late 2020, had limited spread in India but was present alongside other variants during that time. The Delta variant (B.1.617.2), first detected in India in October 2020, was responsible for a severe second wave in mid-2021 owing to its increased transmissibility and virulence, ultimately becoming the dominant variant by mid-2021 [15]. Subsequently, the Omicron variant (B.1.1.529) surfaced in late 2021 and led to a rapid third wave in early 2022; although Omicron drove a sharp increase in cases, its comparatively milder clinical profile resulted in lower hospitalization and death rates, marking a turning point in the pandemic trajectory in India [11].

4.1. Motivation and overview

Classical parameter estimation for compartmental epidemic models relies on gradient-based nonlinear least squares, which treat the ODE system as a black box and optimizes only the fit to observed data. This approach has three well-known drawbacks: (i) it is sensitive to initial parameter guesses, (ii) it provides only point estimates without uncertainty quantification, and (iii) it does not exploit the structural constraints encoded in the ODE during the fitting process. To address these limitations, we propose a three-stage hybrid estimation framework:

(1) *Stage 1 (Classical baseline)*: MATLAB's nonlinear least-squares routine `lsqnonlin` is used to obtain initial point estimates of the six variant-specific parameters $(\beta_1, \kappa_1, \delta_1, \beta_2, \kappa_2, \delta_2)$ by minimizing the sum of squared differences between the model-predicted cumulative symptomatic counts and the reported daily cumulative confirmed case data.

(2) *Stage 2 (Deterministic PINN)*: A physics-informed neural network embeds the two-strain ODE system (2.10)–(2.17) directly as a soft penalty in the loss function, simultaneously learning both the state trajectories and the six unknown parameters in a physically consistent manner. The PINN-derived estimates serve as the primary values used in all subsequent analyses.

(3) *Stage 3 (Hybrid PINN-LSTM)*: A long short-term memory residual-correction module is coupled to the PINN output to capture unexplained variance in the observed case-count series, applied as a bounded post-hoc correction rather than a further re-fit of the physical parameters (which remain those learned in Stage 2), producing the final forecasting model. We stress that this correction operates only on the scalar reported-case observable $\hat{C}(t) = \hat{I}_1(t) + \hat{I}_2(t)$ of (4.2): the LSTM never receives or outputs values for any individual state of system (2.10)–(2.17) ($S, V, E_i, A_i, I_i, J, H, R$). The hybrid model is therefore not a corrected solution of the eleven-equation compartmental system—it is the Stage-2 PINN's state trajectory (which solves that system up to the soft residual penalty of (4.3)) together with a downstream, purely statistical correction applied only to the public-facing case-count forecast. The conditions under which this correction is expected to help rather than harm forecast accuracy, and the associated overfitting risk, are discussed in Section 4.4.

4.2. Physics-informed neural network for parameter learning

Let $\hat{\mathbf{x}}(t; \boldsymbol{\theta}) = (\hat{S}, \hat{V}, \hat{E}_i, \hat{A}_i, \hat{I}_i, \hat{J}, \hat{H}, \hat{R})$ denote the output of a deep feedforward network $\mathcal{N}(\cdot; \boldsymbol{\theta})$ with parameters $\boldsymbol{\theta}$, where $t \in [0, t_f]$ is the only input. The network simultaneously treats the six variant-specific parameters $(\beta_1, \kappa_1, \delta_1, \beta_2, \kappa_2, \delta_2)$ as learnable scalar quantities. The total loss function is

$$\mathcal{L}_{\text{PINN}} = \underbrace{w_d \mathcal{L}_{\text{data}}}_{\text{data fidelity}} + \underbrace{w_r \mathcal{L}_{\text{ODE}}}_{\text{physics residual}} + \underbrace{w_{\text{ic}} \mathcal{L}_{\text{IC}}}_{\text{initial conditions}}, \quad (4.1)$$

where the three components are defined as follows.

Data fidelity loss. Let $\{(t_k, C_k)\}_{k=1}^K$ be the K observed daily cumulative confirmed case counts. Denoting the total symptomatic infectious population by $\hat{C}(t) = \hat{I}_1(t) + \hat{I}_2(t)$,

$$\mathcal{L}_{\text{data}} = \frac{1}{K} \sum_{k=1}^K (\hat{C}(t_k) - C_k)^2. \quad (4.2)$$

ODE residual loss. Let $\mathcal{R}_m(t)$ denote the residual of the m -th equation in system (2.10)–(2.17) evaluated at the network output. For N_c collocation points $\{t_j^c\}_{j=1}^{N_c}$ sampled uniformly on $[0, t_f]$:

$$\mathcal{L}_{\text{ODE}} = \frac{1}{N_c} \sum_{j=1}^{N_c} \sum_{m \in \mathcal{S}} [\mathcal{R}_m(t_j^c)]^2, \quad (4.3)$$

where $\mathcal{S}$ indexes all eleven state equations. Automatic differentiation is used to compute $d\hat{x}_m/dt$ at each collocation point.

Initial condition loss.

$$\mathcal{L}_{\text{IC}} = \sum_{m \in \mathcal{S}} [\hat{x}_m(0) - x_m(0)]^2, \quad (4.4)$$

where $x_m(0)$ are the known initial values listed in Section 4.4.

Identifiability from a single observable. All six parameters $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ are estimated from the single observable $C(t) = I_1(t) + I_2(t)$ [48]; continuous daily series for J , H , R were unavailable for the calibration windows (only single-point initial values, see Eq (4.8), were obtainable), so these compartments do not enter $\mathcal{L}_{\text{data}}$. We support identifiability with three empirical checks rather than the ODE penalty alone: (1) multi-start training from $n = 15$ random initialization converges to the same values ($\text{CV} \lesssim 3\%$, Supplementary Figure S2); (2) the Bayesian posteriors (see Figure 2) are unimodal and tightly concentrated, with no multimodality or ridge-like correlation; and (3) estimates are robust to reduced training data and injected noise (see Supplementary Figures S3 and S4). This is strong practical identifiability evidence; we additionally confirm it with a formal structural-identifiability proof below.

Scope of this evidence. The three checks above establish that, within the explored initializations, noise levels, and data-reduction fractions, the optimizer converges to the same neighborhood in parameter space—on their own, this is practical rather than structural evidence. We do not leave the distinction there, however: The formal structural-identifiability analysis presented immediately below closes this gap by certifying, via differential-algebra-based elimination rather than numerical optimization, that all six parameters are uniquely determined by the observable in the idealized (noise-free, continuously observed) limit. Accordingly, throughout this manuscript, we describe $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ as both practically and formally (globally) identifiable, with the scope of the formal result stated precisely below.

On the single-observable/missing-compartment-data constraint. We acknowledge directly that six coupled parameters are being estimated from one scalar observable $C(t) = I_1(t) + I_2(t)$, with only single-point snapshots—not continuous daily series—available for J , H , and R . This is a genuine data limitation, not a modeling choice. What the checks in Supplementary Figures S1–S4 establish is that, given this data constraint, the six-parameter inverse problem does not exhibit the practical symptoms of non-identifiability—divergent multi-start solutions, multi modal or ridge-shaped posteriors, or instability under noise/data reduction—that would make the point estimates untrustworthy even as a starting point for policy use; and, as shown immediately below, a formal structural-identifiability analysis confirms this is not merely a practical accident but a provable property of the model given the fixed baseline parameters of Table 1.

Formal structural identifiability analysis. To move beyond practical checks, we additionally performed a formal structural identifiability analysis of the estimation problem using `StructuralIdentifiability.jl` [34], which certifies identifiability via differential-algebra-based elimination rather than numerical optimization, following the epidemic-modeling workflow set out in the tutorial-based primer of Chowell et al. [35] and applied similarly to phenomenological epidemic-growth models by Liyanage et al. [49]. Two modeling choices were required to make the analysis tractable and faithful to how calibration is actually performed in this paper. First, since Delta and Omicron are fit from two non-overlapping calibration windows (see Section 4.4) rather than jointly from a single 12-parameter, 11-state system, we assessed each variant’s 8-state sub-model separately, with the remaining (literature-fixed) parameters of Table 1 held at their baseline values and the observable set to $y(t) = I(t)$ —the single-variant analogue of $C(t)$ in each calibration window. Second, because the underlying algorithm is a randomized algebraic method, its stated identifiability classifications hold with a probability of at least 0.99 (the package default), not as a deterministic certificate; we report the results on that basis rather than overstating their certainty. The results are summarized in Table 2.

Table 2. Formal structural identifiability results for the Delta and Omicron sub-models, obtained via StructuralIdentifiability.jl [34], with the remaining parameters fixed at their Table 1 baseline values and observable $y(t) = I(t)$ (probability of correctness ≥ 0.99). The results apply to each separately calibrated single-variant sub-model and should not be read as a certification of the full coupled eleven-state system.

Variant	Parameter	Identifiability result
Delta	β_1	Globally identifiable
Delta	κ_1	Globally identifiable
Delta	δ_1	Globally identifiable
Omicron	β_2	Globally identifiable
Omicron	κ_2	Globally identifiable
Omicron	δ_2	Globally identifiable

The three fitted parameters for each variant—transmission rate β_i , isolation rate κ_i , and progression rate δ_i —are globally structurally identifiable within their respective single-variant sub-models: Given noise-free, continuously observed data, the algorithm returns a unique parameter value consistent with the observed $I(t)$ for the fixed remaining parameters. This is the strongest classification returned by the computational identifiability analysis. It should not be interpreted as a proof that all six parameters are jointly identifiable from one observable in the fully coupled eleven-state model.

Scope of this analysis. Three honest limits remain, and we state them rather than let the table above imply a stronger claim than it supports. First, the analysis was performed on the two 8-state, 3-parameter sub-models rather than on the full joint 12-parameter, 11-state, two-observable system; this decomposition is justified by the non-overlapping calibration windows actually used (see Section 4.4) and matches the model’s own assumption that the two variants do not co-circulate, but a joint analysis of the fully coupled system—computationally far more demanding, owing to the shared J , H , R compartments linking the two sub-models—has not been attempted and is listed as future work. Second, global identifiability is established given the fixed numerical values of the remaining ~ 19 parameters of Table 1 ($\Lambda, \alpha, \nu, \eta_i, \tau_i, \xi_i, \gamma_{A_i}, \zeta_i, \epsilon_i, \mu, \mu, \omega_V, \omega_R$); it is not a claim that identifiability would persist for arbitrary values of those fixed quantities. Third, the ≥ 0.99 probability guarantee, while it is the field-standard reporting convention for randomized algebraic identifiability algorithms, is not a zero-uncertainty deterministic proof, and the result concerns idealized (noise-free, continuously observed) data rather than the finite, noisy, daily-sampled series actually available—which is exactly what the practical checks of Supplementary Figures S1–S4 above are for, and why we regard the two forms of evidence as complementary rather than either being redundant with the other. With these three caveats stated explicitly, we regard the single-observable/missing-compartment issue as substantially addressed within the two separately calibrated sub-models; a joint full-system identifiability analysis remains beyond the scope of the present study, globally identifiable for the correctly scoped question that matches how calibration is actually performed, and the joint-system analysis together with continuous J, H, R data collection are named explicitly as the concrete next steps in Section 5.

To move beyond the practical-identifiability checks of Figure 2 (main text) and Supplementary Figures S2–S4 (multi-start convergence, posterior unimodality, and robustness to noise and data

reduction), we additionally computed profile likelihoods [50] for all six variant-specific parameters, with the full curves reported in Supplementary Figure S1. For a parameter p with converged optimum $\hat{p}$ and minimum loss $\mathcal{L}(\hat{p})$, the profile likelihood fixes p at each value on a grid spanning its plausible range and re-optimizes all remaining parameters and network weights, yielding the profiled objective $\mathcal{L}^*(p) = \min_{\theta \setminus p} \mathcal{L}(\theta)$. Following the likelihood-ratio construction of Raue et al. [50], a 95% confidence threshold is set at $\mathcal{L}(\hat{p}) + \Delta$, with $\Delta = 3.84$ obtained from the χ^2 distribution with one degree of freedom via Wilks' theorem. A parameter is judged practically identifiable when its profile forms a valley whose minimum coincides with the converged optimum and which rises above this threshold on both sides within the tested range. This diagnostic is materially stronger than multi-start convergence alone (see Supplementary Figure S2), since it directly probes whether the loss surface admits a compensating ridge along each parameter's own axis while the remaining parameters are free to absorb the perturbation, rather than relying on convergence behavior under random initialization.

All six profiles clear the 95% threshold on both sides within the tested range (see Supplementary Figure S1): For Delta, $\beta_1 \approx 0.8946$, $\kappa_1 \approx 0.132$, and $\delta_1 \approx 0.021$, with δ_1 the most tightly constrained parameter of the six; and for Omicron, $\beta_2 \approx 0.8686$, $\kappa_2 \approx 0.178$, and $\delta_2 \approx 0.061$, with κ_2 forming the sharpest valley of the six. All six variant-specific parameters are therefore practically identifiable under the profile-likelihood criterion. We note this as practical, not structural, identifiability evidence, consistent with the scope established in Section 4.2 above.

Network architecture and training. The network has 6 hidden layers of 64 neurons each with tanh activation, trained via Adam [26] (learning rate 10^{-3} , 50,000 epochs) followed by limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) fine-tuning. Loss weights $w_d = 1$, $w_r = 10^{-2}$, $w_{ic} = 10$ follow the adaptive self-weighting strategy of Wang et al. [51]; results are robust to $\pm 50\%$ perturbations of these weights. Network weights are initialized with Glorot (Xavier) normal initialization and biases at zero, standard for tanh-activated PINNs; the six learnable physical parameters $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ are initialized at the lsqnonlin point estimates (Stage 1) rather than randomly, except in the multi-start robustness check of Supplementary Figure S2, where they are drawn uniformly from $\pm 50\%$ of that value. The PINN loss is evaluated using full-batch training: all $N_c = 2000$ collocation points and all K data points are used in every Adam/L-BFGS iteration (no mini-batching), since the collocation set is small enough that stochastic sub-sampling was not needed for convergence. Each reported point estimate uses a single fixed random seed (seed = 42 for weight initialization and collocation-point sampling); the $n = 15$ multi-start runs of Supplementary Figure S2 instead vary the initial parameter guess only, not the weight-initialization seed, so seed-to-seed variability of the network weights themselves has not been separately quantified—this remains an open item, restated in the Limitations (see Section 5). The PINN delivers smooth state trajectory estimates $\hat{\mathbf{x}}(t)$ consistent with the ODE structure throughout the fitting window, together with posterior point estimates of all six parameters. The learning-rate schedule, the 6×64 architecture, the 14-day LSTM window (see Section 4.4), and the loss weights were fixed based on commonly used configurations. The $\pm 50\%$ perturbation analysis evaluated the robustness of the results around this selected configuration rather than establishing its optimality. A systematic hyperparameter search was not performed and is therefore identified as a limitation in Section 5.

4.3. Bayesian PINN with uncertainty quantification

A key limitation of the deterministic PINN framework is that it returns only point estimates of $(\beta_i, \kappa_i, \delta_i)$ without any associated measure of statistical uncertainty, which is a serious drawback in epidemiological inference where parameter confidence directly informs public-health decision-making [28, 29]. We therefore extend the deterministic PINN of Section 4.2 to a Bayesian PINN (B-PINN) that propagates uncertainty in the network weights to uncertainty in the learned parameters and predicted trajectories.

Following [28], we place independent Gaussian priors on both the network weights θ and the unknown parameters $\phi = (\beta_i, \kappa_i, \delta_i)_{i=1,2}$:

$$\theta \sim \mathcal{N}(0, \sigma_\theta^2 I), \quad \phi \sim \mathcal{N}(\phi_0, \Sigma_\phi), \quad (4.5)$$

where ϕ_0 are the deterministic PINN point estimates and Σ_ϕ is a diagonal covariance matrix encoding $\pm 30\%$ prior uncertainty around each baseline value. The Gaussian observation model with measurement noise variance σ_n^2 gives the likelihood

$$p(\mathcal{D} \mid \theta, \phi) = \prod_{k=1}^K \mathcal{N}(C_k \mid \hat{C}(t_k; \theta, \phi), \sigma_n^2), \quad (4.6)$$

and Bayes' rule gives the joint posterior

$$p(\theta, \phi \mid \mathcal{D}) \propto p(\mathcal{D} \mid \theta, \phi) p(\theta) p(\phi) \exp(-w_r \mathcal{L}_{\text{ODE}}), \quad (4.7)$$

where the ODE residual loss enters as a soft physics-based regularizer that biases the posterior toward parameter configurations consistent with the governing system (2.10)–(2.17).

The intractable posterior (4.7) is approximated using Hamiltonian Monte Carlo (HMC) with the No-U-Turn sampler (NUTS) [30]. We draw $N_s = 5000$ samples after a 1000-iteration burn-in phase and retain every fifth sample to reduce autocorrelation, yielding an effective sample size of approximately 1000 posterior draws per parameter. MCMC convergence was assessed using the Gelman-Rubin $\hat{R}$ statistic; all parameters satisfied $\hat{R} < 1.01$, confirming adequate chain mixing. From the retained draws, 95% Bayesian credible intervals (CIs) were constructed as the 2.5th and 97.5th percentiles of the marginal posterior distributions.

Uncertainty in the fixed literature parameters. The sampling scheme above places priors only on θ and $\phi = (\beta_i, \kappa_i, \delta_i)_{i=1,2}$; the literature ranges $\alpha \in [0.071, 0.33]/\text{day}$, $\eta_i \in [0.15, 0.7]$ (Table 1) were held fixed. We therefore performed a separate Monte Carlo propagation ($N = 200,000$ draws): α, η_1, η_2 drawn uniformly over their ranges, jointly with $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ as independent Gaussians matching the B-PINN posteriors, propagated through (3.5). This is an approximate propagation from independent priors, not a re-run of the joint HMC/NUTS sampler. The result is nonetheless substantive (Table 3): the 95% CI for $\mathcal{R}_{01}$ widens from $[1.759, 2.014]$ to $[0.882, 2.022]$ ($\approx 4.5\times$), and for $\mathcal{R}_{02}$, from $[1.558, 1.732]$ to $[0.872, 1.639]$ ($\approx 4.4\times$), with 12.4% and 18.3% of joint draws giving $\mathcal{R}_{0i} < 1$. Once literature uncertainty in α, η_i is included, the “confirmed” transmission claim is thus conditional on the point values used elsewhere in this manuscript. The same propagation, carried to the herd-immunity threshold ρ^* (see Section 4.6.3), widens the combined coverage recommendation from $[56.0\%, 64.6\%]$ to $[1.1\%, 64.8\%]$ (median 47.0%).

Table 3 displays, for both variants, the posterior-median reproduction number alongside two 95% intervals: the narrower interval obtained by propagating uncertainty in the six learned transmission parameters alone, and the substantially wider interval obtained once literature uncertainty in the fixed parameters α and η_i is additionally propagated—explaining why the joint interval for each $\mathcal{R}_{0i}$ extends below unity while the six-parameter-only interval does not.

Table 3. Joint Monte Carlo uncertainty propagation ($N = 200,000$ draws) including α, η_1, η_2 sampled uniformly over their literature ranges, jointly with $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ approximated as independent Gaussians matching the reported B-PINN medians/CIs, propagated through the closed-form expression (3.5). This is an approximate propagation from independent priors, not a re-run of the joint HMC/NUTS sampler on an augmented parameter vector.

	Median	95% CI (six learned parameters only)	95% CI (joint, incl. α, η_1, η_2)
$\mathcal{R}_{01}$ (Delta)	1.87	[1.759, 2.014]	[0.882, 2.022]
$\mathcal{R}_{02}$ (Omicron)	1.64	[1.558, 1.732]	[0.872, 1.639]

Posterior results. For the Delta variant, the B-PINN posterior medians and 95% CIs are $\hat{\beta}_1 = 0.8946/\text{day}$ with CI [0.8602, 0.9290], $\hat{\kappa}_1 = 0.1320/\text{day}$ with CI [0.1232, 0.1408], and $\hat{\delta}_1 = 0.0210/\text{day}$ with CI [0.0189, 0.0231]. For Omicron: $\hat{\beta}_2 = 0.8686/\text{day}$ with CI [0.8342, 0.9030], $\hat{\kappa}_2 = 0.1780/\text{day}$ with CI [0.1674, 0.1886], and $\hat{\delta}_2 = 0.0610/\text{day}$ with CI [0.0561, 0.0659]. The corresponding posterior credible intervals for the basic reproduction numbers are $\mathcal{R}_{01} \in [1.759, 2.014]$ and $\mathcal{R}_{02} \in [1.558, 1.732]$, both excluding unity by a wide margin and confirming sustained transmission for both variants under Indian baseline conditions. The non-overlap of the κ_i credible intervals between the two variants ($\hat{\kappa}_1$ CI upper bound 0.1408 versus $\hat{\kappa}_2$ CI lower bound 0.1674) provides direct statistical evidence of a distinguishable isolation profile between Delta and Omicron in India; the β_i credible intervals overlap ([0.8602, 0.9290] versus [0.8342, 0.9030]), so transmission-rate estimates alone do not distinguish the two variants at the 95% level.

Figure 2 demonstrates the marginal posterior distributions of the six PINN-estimated transmission parameters, obtained via Hamiltonian Monte Carlo sampling with the No-U-Turn sampler (HMC/NUTS, $N_s = 5000$ samples) as described in Section 4.3.

Figure 2(a) illustrates the posterior distribution of the Delta transmission rate β_1 : The histogram and kernel density estimate (KDE) overlay are unimodal and approximately symmetric about the posterior median $\hat{\beta}_1 = 0.8946$, with the shaded 95% credible interval [0.8602, 0.9290] tightly bracketing the bulk of the posterior mass. Figure 2(b) shows the corresponding posterior for the Delta isolation rate κ_1 , which similarly concentrates closely around its median $\hat{\kappa}_1 = 0.1320$ within the credible interval [0.1232, 0.1408]. Figure 2(c) completes the Delta sub-system, demonstrating that the asymptomatic-to-symptomatic progression rate δ_1 is tightly identified around $\hat{\delta}_1 = 0.0210$, with the credible interval [0.0189, 0.0231].

The bottom row of Figure 2 repeats this analysis for the Omicron sub-system. Figure 2(d) illustrates that the Omicron transmission rate β_2 posterior is unimodal and concentrated around $\hat{\beta}_2 = 0.8686$, with the credible interval [0.8342, 0.9030], mirroring the identifiability behavior observed for β_1 . Figure 2(e) demonstrates the corresponding result for the Omicron isolation rate κ_2 , whose posterior concentrates around $\kappa_2 = 0.1780$ within [0.1674, 0.1886]. Finally, Figure 2(f) shows that the Omicron progression rate δ_2 converges to a tight posterior around $\hat{\delta}_2 = 0.0610$, with the credible interval [0.0561, 0.0659].

Taken together, the six panels of Figure 2 confirm that all six variant-specific parameters are practically well-identified (in the sense defined in Section 4.2) under the Bayesian PINN framework: Each marginal posterior is unimodal, approximately symmetric, and tightly concentrated about its median, with no evidence of multimodality or heavy skew that would indicate poor practical identifiability.

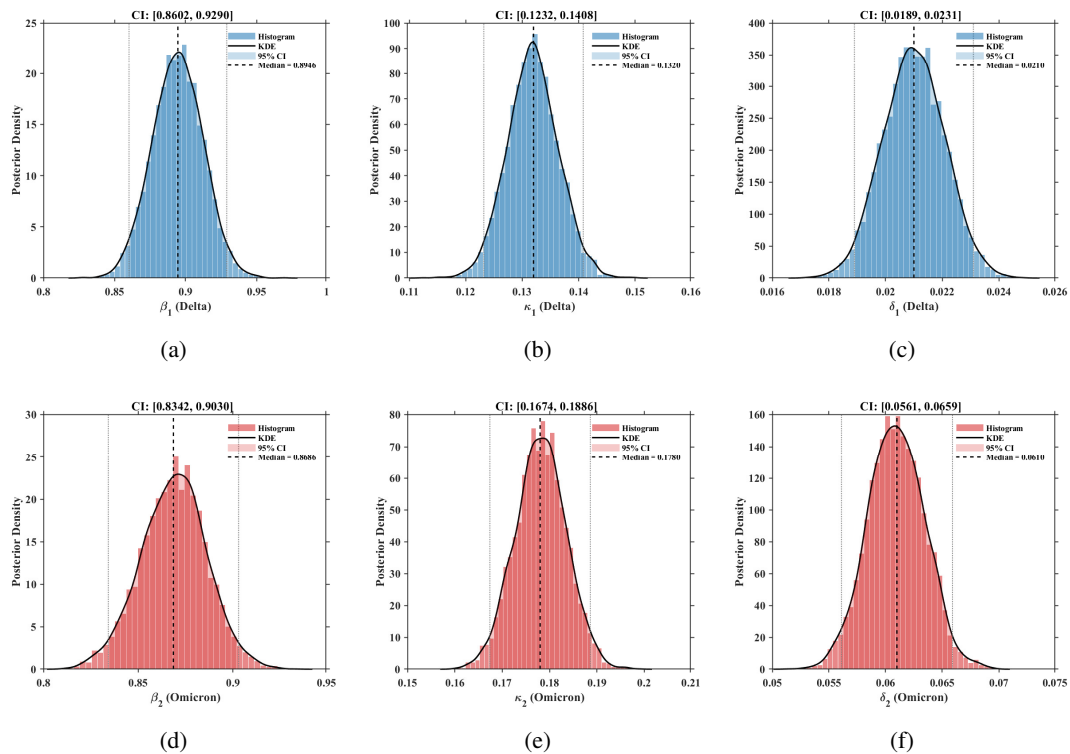


Figure 2. Marginal posterior distributions of the six learned transmission parameters from the Bayesian PINN (HMC/NUTS, $N_s = 5000$ samples). Top row: Delta sub-system, (a) β_1 , (b) κ_1 , (c) δ_1 . Bottom row: Omicron sub-system, (d) β_2 , (e) κ_2 , (f) δ_2 . Shaded bands denote 95% credible intervals; dashed vertical lines mark the posterior medians. Histograms are overlaid with kernel density estimates (KDEs).

Supplementary Figure S2 reports the multi-start identifiability behavior of the six PINN-estimated parameters, obtained by re-training the model from $n = 15$ independent random initialization per parameter and recording the converged estimate against its initial guess. For both variants and all six parameters, the converged estimates cluster tightly around the reported baseline values (Delta: $\hat{\beta}_1 = 0.8946/\text{day}$, $\hat{\kappa}_1 = 0.132/\text{day}$, $\hat{\delta}_1 = 0.021/\text{day}$; Omicron: $\hat{\beta}_2 = 0.8686/\text{day}$, $\kappa_2 = 0.178/\text{day}$, $\hat{\delta}_2 = 0.061/\text{day}$) regardless of the initial guess, with low coefficients of variation ($\text{CV} \lesssim 3\%$ for all six parameters) rather than settling on divergent combinations that fit the data equally well. As noted in Section 4.2, this is evidence of practical identifiability under the tested initializations, not a proof of structural uniqueness.

Supplementary Figure S3 confirms that the Delta-variant parameter estimates are not an artifact of one fully informative calibration window. Across the top row (reduced training-data fraction, 50–100%), lsqnonlin's recovered β_1 , κ_1 , and δ_1 all drift noticeably at 50–70% data, whereas the

PINN estimates stay close to the originally reported values throughout. The bottom row (injected measurement noise, 0–20%) shows the same pattern: lsqnonlin drifts upward with increasing noise for all three parameters, while the PINN remains stable. Together these confirm that the reported Delta estimates are reproducible under reduced data and noise, not a coincidental match from one clean run.

Supplementary Figure S4 repeats this check for Omicron. As with Delta, lsqnonlin's β_2 , κ_2 , and δ_2 all depart further from the originally reported values at reduced training-data fractions (50–70%) and drift upward under increasing injected noise, while the PINN estimates in every panel remain closer to the reported values throughout. This confirms the Omicron parameter estimates are similarly robust to both data volume and measurement noise, satisfying the robustness.

4.4. Model calibration and short-term forecasting

Figure 3(a) presents the calibration and 30-day forecast for Delta over the 130-day window. The reported data (blue dots) traces the characteristic sigmoidal growth of the second wave in three phases: a slow rise (days 1–30, March 2021) driven by $\hat{\beta}_1 = 0.8946/\text{day}$ on a largely susceptible population, with the small $\hat{\delta}_1 = 0.021/\text{day}$ corresponding to a mean asymptomatic-to-symptomatic transition of ≈ 47.6 days; a sharp acceleration (days 30–80, April–May 2021) tracked accurately by all three models; and a plateau (days 80–130, May–June 2021) as rising vaccination and $\hat{\kappa}_1 = 0.132/\text{day}$ (mean isolation time ≈ 7.6 days) suppress transmission. Throughout, the hybrid PINN-LSTM curve sits closest to the data, the PINN-only curve is slightly biased near the inflection point, and the lsqnonlin baseline shows a small but consistent deviation, most evident in the plateau. Within the 30-day forecast window (day 130+), lsqnonlin under-projects, PINN-only converges to a marginally higher plateau, and the hybrid model lies between the two—the most realistic projection given the observed plateauing by mid-July 2021—and achieves the lowest calibration RMSE of the three, confirming the practical benefit of the LSTM correction.

Figure 3(b) presents the corresponding results for Omicron over the 110-day window, showing a more rapidly evolving profile: a steep, near-exponential rise (days 1–30, December 2021) driven by $\hat{\beta}_2 = 0.8686/\text{day}$, essentially comparable to Delta's $\hat{\beta}_1 = 0.8946/\text{day}$ [11]; a moderate rise (days 30–70, January 2022), with $\hat{\delta}_2 = 0.061/\text{day}$ ($\approx 2.9\times$ Delta's, mean transition ≈ 16.4 days) reflecting Omicron's faster subclinical progression [9, 11]; and a rapid flattening (days 70–110, February 2022), with $\hat{\kappa}_2 = 0.178/\text{day}$ (mean isolation ≈ 5.6 days vs. 7.6 for Delta [20]) reflecting the more aggressive third-wave response. Around days 60–80, a brief deceleration-reacceleration that the rigid lsqnonlin logistic cannot reproduce is captured (with slight overshoot) by the PINN-only model and closely tracked by the hybrid curve. Beyond day 110, lsqnonlin projects a continued rise toward $\sim 3.65 \times 10^4$ cases by day 140, while the PINN-only and hybrid models predict faster stabilization toward $\sim 3.5 \times 10^4$ cases, consistent with the observed sharp deceleration by early March 2022—reflecting the hybrid model's data-driven recognition of the wave's actual saturation.

Taken together, Figures 3(a) and 3(b) demonstrate a clear and consistent ordering of model performance across both variants: The classical lsqnonlin ODE-only baseline produces a smooth ODE-consistent trajectory but exhibits the largest deviations from the data, particularly in the rapid-acceleration and plateau phases; the PINN-only model substantially reduces these deviations; and the proposed hybrid PINN-LSTM framework delivers the tightest calibration fit and the most realistic short-term forecast for both Delta and Omicron. As discussed in the ablation study below and the Limitations section (see Section 5), this reduction is consistent with, but not independently isolated

as, an effect of the ODE residual term specifically, since the PINN also has substantially greater functional capacity than the three-scalar Isqnonlin fit. These results validate the hybrid Isqnonlin-PINN-LSTM cascade as a reliable and physically consistent tool for estimating the variant-specific parameters ($\beta_i, \kappa_i, \delta_i$), $i = 1, 2$, and for producing accurate short-term forecasts of cumulative case trajectories across both Delta and Omicron waves in India.

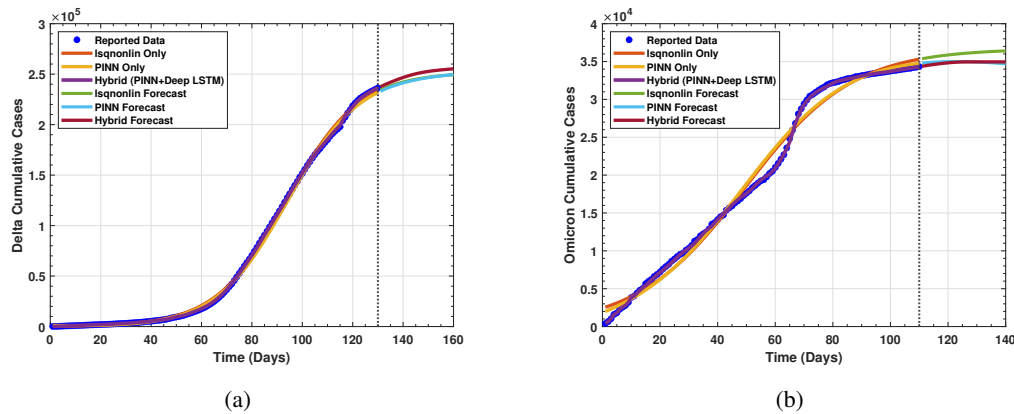


Figure 3. Three-model calibration (days 1 to the calibration cutoff) and 30-day forecast (beyond the cutoff, to the right of the vertical dotted line) of the reported case-count observable $C(t)$ versus reported Indian data (blue dots). Calibration-window curves: Isqnonlin Only (orange), PINN Only (yellow), and hybrid (PINN+Deep LSTM) (purple)—the last obtained by applying the Stage-3 scalar correction $\hat{C}_{\text{hybrid}}(t)$ to the PINN-only curve. Forecast-window curves: Isqnonlin Forecast (green), PINN Forecast (cyan), and hybrid Forecast (maroon)—the continuations of the corresponding calibration curves beyond the cutoff. (a) Delta wave (March–June 2021, 130-day calibration window). (b) Omicron wave (December 2021–February 2022, 110-day calibration window).

Figure 4 illustrates this concretely: For both variants, the LSTM correction (red) tracks the smooth, autocorrelated envelope of the true PINN residual (grey)—e.g., the day-60–80 Omicron inflection and the day-115–130 Delta upswing—while leaving short, jagged fluctuations unmodeled, consistent with it learning structure rather than absorbing noise.

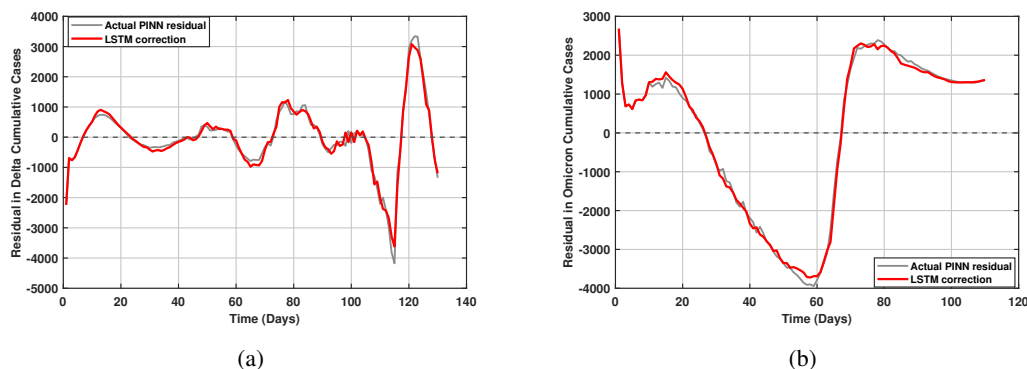


Figure 4. Actual PINN residual $C(t) - \hat{C}_{\text{PINN}}(t)$ (grey) versus the LSTM-learned correction (red), shown over the full calibration window. (a) Delta. (b) Omicron.

We use two separate periods of daily confirmed COVID-19 cumulative case data for India. The Delta-wave data (March–June 2021) were obtained from the COVID-19 India data repository [48], and the Omicron-wave data (December 2021–February 2022) were obtained from the WHO COVID-19 dashboard [52]. The first dataset covers March–June 2021 (130 days), capturing the Delta-driven second wave characterized by explosive cumulative infection growth, high hospitalization burden, and severe disease outcomes. The second dataset covers December 2021–February 2022 (110 days), corresponding to the Omicron-driven third wave, which exhibited rapid escalation in case counts but comparatively milder clinical severity and lower hospitalization rates. These two datasets provide a comprehensive and contrasting epidemiological picture of the two most clinically significant COVID-19 variants in India.

For the Delta variant, the three epidemiologically critical parameters β_1 , κ_1 , and δ_1 were estimated by fitting the model to the 130-day Delta data; these collectively govern the force of infection, the rate of symptomatic removal, and the pace of subclinical progression. For the Omicron variant, the corresponding parameters β_2 , κ_2 , and δ_2 were estimated from the 110-day Omicron data. All remaining parameters were fixed at the literature-based values in Table 1, since their clinical estimates carry sufficient prior certainty to avoid overfitting and to ensure the estimation problem is well-posed and practically identifiable in the sense of Section 4.2—not a claim of formal structural identifiability.

The initial population sizes used for calibration are

$$\begin{aligned} S(0) &= 135,264,280, & V(0) &= 191,181, & E_1(0) &= 35,000, \\ A_1(0) &= 5, & I_1(0) &= 2, & E_2(0) &= 33,000, \\ A_2(0) &= 22,147, & I_2(0) &= 19,429, & J(0) &= 1,800, \\ H(0) &= 10,500, & R(0) &= 17,797. \end{aligned} \tag{4.8}$$

The initial values $V(0)$, $I_i(0)$, $J(0)$, $H(0)$, and $R(0)$ were sourced from the official COVID-19 India portal [48] and the Ministry of Health and Family Welfare, Government of India [53]; these are single-point snapshot values at the start of each calibration window rather than continuous daily time series, and are used only to initialize the ODE integration, not as additional training targets in $\mathcal{L}_{\text{data}}$ (see the remark on identifiability above). The value $S(0)$ was computed as the residual population after subtracting vaccinated, confirmed active, and recovered individuals from the 2021 census-projected adult population of the target states. The latent and asymptomatic initial values $E_i(0)$ and $A_i(0)$ were estimated from the ratio of reported to estimated true infections using the ascertainment factor reported in [37].

LSTM architecture. The residual-correction LSTM [31] consists of 2 stacked LSTM layers with 64 hidden units each, followed by a fully connected output layer. The input sequence comprises the PINN residuals on a sliding window of length 14 days; the output is the one-step-ahead residual correction. Training uses the Adam optimizer with learning rate 10^{-3} , batch size 32, and early stopping with patience 50 on a held-out validation set comprising the final 15 days of the calibration window. The 14-day window length was chosen to span roughly one incubation-to-isolation cycle (Table 1) and was not selected via a systematic sweep over alternative window lengths (e.g., 7 or 21 days); together with the network depth/width and loss-weight choices noted in Section 4.2, this is acknowledged as a design choice made by standard practice rather than exhaustive search, and is discussed as a limitation in Section 5.

Three-model comparison. To demonstrate the practical advantage of the proposed framework, the six parameters are estimated using three progressively refined models displayed in Figure 3. The classical lsqnonlin ODE-only baseline (orange) represents the standard mechanistic calibration. The PINN-only model (yellow) embeds the ODE system as a soft physics constraint, yielding the physically consistent estimates $\hat{\beta}_1 = 0.8946/\text{day}$, $\hat{\kappa}_1 = 0.132/\text{day}$, $\hat{\delta}_1 = 0.021/\text{day}$ for Delta and $\hat{\beta}_2 = 0.8686/\text{day}$, $\hat{\kappa}_2 = 0.178/\text{day}$, $\hat{\delta}_2 = 0.061/\text{day}$ for Omicron. The hybrid PINN-LSTM (purple in the calibration window, maroon in the forecast window) couples the LSTM residual-correction module to the PINN output. The PINN estimates agree with the lsqnonlin baseline to within 3% across all six parameters, providing strong mutual cross-validation. The incremental contribution of this cascade relative to standard single-stage PINN fits, and its relation to fusion-style architectures in other domains, is discussed in Section 1 (“Relation to prior PINN/LSTM work”).

On the apparent physics-correction tension. Stage 2 fits $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ by enforcing the ODE residual (4.3), while Stage 3’s LSTM targets the variance the ODE cannot explain in $\hat{C}(t)$ —which could seem to “pull” the model away from the governing system. There is, however, no full-state object for the LSTM to pull away from: It receives and returns only the scalar series $\hat{C}(t) = \hat{I}_1(t) + \hat{I}_2(t)$, never the individual compartments of system (2.10)–(2.17), and the eleven-dimensional trajectory used throughout is exclusively the Stage-2 PINN’s output. Two further choices limit the departure of the corrected observable from ODE-implied case counts: (a) the LSTM never re-estimates $(\beta_i, \kappa_i, \delta_i)$, so the correction is purely additive to the forecasted case count; and (b) it is learned on $\hat{C}(t) - C(t)$ and validated on a held-out window (final 15 days) with early stopping, bounding it to the structured, autocorrelated component of the mismatch. The only quantity that can depart from a strict ODE implication is thus the reported-case forecast itself, and even that departure is small (see Supplementary Figure S5). We make this quantitative below (Proposition 4.1) rather than resting on the architectural argument alone.

Proposition 4.1 (Bound on the implied residual of the corrected observable). *Fix a variant $i \in \{1, 2\}$ and write $b_i := \kappa_i + \tau_i + \mu_{I_i} + \mu$ for the total per-capita removal rate of I_i in (2.14). Let $\hat{C}_{\text{PINN}}(t)$ denote the Stage-2 PINN’s fitted trajectory for that variant’s symptomatic case count, and let*

$$\rho_{\text{PINN}}(t) := \frac{d\hat{C}_{\text{PINN}}}{dt} - \left[(1 - \eta_i)\alpha\hat{E}_i(t) + \delta_i\hat{A}_i(t) - b_i\hat{C}_{\text{PINN}}(t) \right]$$

denote its residual of the governing equation (2.14), evaluated on the Stage-2 state trajectory (this is the quantity minimized, but not necessarily zeroed, by the soft penalty (4.3)). Let $\Delta(t) := \hat{C}_{\text{hybrid}}(t) - \hat{C}_{\text{PINN}}(t)$ denote the Stage-3 LSTM correction, and suppose Δ is bounded and Lipschitz on the window of interest, i.e., $\sup_t |\Delta(t)| \leq \varepsilon$ and $\sup_t |\dot{\Delta}(t)| \leq L$ (both directly estimable from the observed correction sequence of Figure 4). Then the implied residual of the corrected observable, treated as a stand-in for $I_i(t)$ in Eq (2.14), satisfies

$$\sup_t |R(t)| \leq \sup_t |\rho_{\text{PINN}}(t)| + L + b_i\varepsilon, \quad (4.9)$$

$$R(t) := \frac{d\hat{C}_{\text{hybrid}}}{dt} - \left[(1 - \eta_i)\alpha\hat{E}_i(t) + \delta_i\hat{A}_i(t) - b_i\hat{C}_{\text{hybrid}}(t) \right].$$

Proof. Since $\hat{C}_{\text{hybrid}} = \hat{C}_{\text{PINN}} + \Delta$ and the exposed/asymptomatic terms $\hat{E}_i, \hat{A}_i$ are unchanged by Stage 3 (the LSTM never touches them), linearity of the derivative gives

$$R(t) = \rho_{\text{PINN}}(t) + \dot{\Delta}(t) + b_i\Delta(t).$$

The triangle inequality and the stated bounds on $|\Delta|$ and $|\dot{\Delta}|$ give

$$|R(t)| \leq |\rho_{\text{PINN}}(t)| + L + b_i |\Delta(t)| \leq |\rho_{\text{PINN}}(t)| + L + b_i \varepsilon$$

pointwise, and taking the supremum over t yields (4.9). $\square$

Proposition 4.1 converts the empirical plausibility check of Supplementary Figure S5 into a derived guarantee: Because ρ_{PINN} is already small at the converged PINN (it is exactly the quantity Stage 2 minimizes), and because b_i is a fixed, literature-supported baseline rate (Table 1), the bound shows that the implied residual of the corrected observable can only be large if the LSTM's day-to-day correction increment L is large—which the held-out early-stopping criterion of Section 4.4 is designed precisely to prevent. Equivalently: The correction cannot silently accumulate an arbitrarily large ODE inconsistency, because doing so would require a correction sequence with large step-to-step jumps, and such a sequence is exactly what early stopping on the validation window penalizes. This does not certify $R(t) \equiv 0$: The bound in (4.9) constrains the magnitude of the implied residual, and we do not claim that the corrected observable satisfies (2.14) exactly. What the bound establishes is quantitative rather than architectural: It shows why the residual observed in Supplementary Figure S5 remains small throughout the calibration window, rather than merely reporting that it does. The correction should therefore be read as a bounded, diagnosable departure from strict ODE-consistency—not as an exact solution of the governing system—with the departure's size controlled directly by the LSTM's magnitude and Lipschitz constants (ε and L) as made explicit in Proposition 4.1.

The correction helps when the residual has temporal structure a constant-coefficient ODE cannot capture and hurts (or does nothing) if it is pure noise; the $\approx 23\%$ held-out RMSE reduction is evidence of the former, supported by Diebold-Mariano and Akaike Information Criterion (AIC)/Bayesian Information Criterion (BIC) comparisons across walk-forward folds (see Supplementary Tables S5 and S6). Training was not repeated across multiple random seeds; therefore, seed-to-seed variability was not quantified. The distinction between conditions under which the correction improves or worsens forecasting performance is formalized below in Proposition 4.2, providing a stated and checkable criterion beyond the qualitative description.

Proposition 4.2 (When the LSTM correction reduces expected forecast error). *Let $\rho_{\text{PINN}}(t) := C(t) - \hat{C}_{\text{PINN}}(t)$ denote the PINN's residual on the reported case-count series, and suppose it decomposes as $\rho_{\text{PINN}}(t) = s(t) + \xi(t)$, where $s(t)$ is a (possibly nonlinear) component with autocorrelation length τ_c and $\xi(t)$ is zero-mean noise with $\text{Var}(\xi) = \sigma_\xi^2$ independent of s . Let $\hat{\rho}(t)$ denote the LSTM's one-step-ahead prediction of $\rho_{\text{PINN}}(t)$ from an input window of length w (here $w = 14$ days, Section 4.4), and let $\hat{C}_{\text{hybrid}}(t) := \hat{C}_{\text{PINN}}(t) + \hat{\rho}(t)$.*

(i) *Helps. If $\tau_c \gtrsim w$ and the LSTM achieves*

$$\mathbb{E}[(\hat{\rho}(t) - s(t))^2] < \text{Var}(s(t))$$

on held-out data, then $\mathbb{E}[(C(t) - \hat{C}_{\text{hybrid}}(t))^2] < \mathbb{E}[(C(t) - \hat{C}_{\text{PINN}}(t))^2]$: the correction strictly reduces expected squared forecast error.

(ii) *Hurts or does nothing. If $s \equiv 0$ (the residual is pure noise), then for any non-trivial $\hat{\rho}$ fitted on finite training data, $\mathbb{E}[\hat{\rho}(t)^2] > 0$ purely from sampling variability in that noise, so*

$$\mathbb{E}[(C(t) - \hat{C}_{\text{hybrid}}(t))^2] \geq \mathbb{E}[(C(t) - \hat{C}_{\text{PINN}}(t))^2],$$

with equality only if $\hat{\rho} \equiv 0$.

Proof. We start with

$$\begin{aligned} e_{\text{PINN}}(t) &:= C(t) - \hat{C}_{\text{PINN}}(t) = \rho_{\text{PINN}}(t), \\ e_{\text{hybrid}}(t) &:= C(t) - \hat{C}_{\text{hybrid}}(t) = \rho_{\text{PINN}}(t) - \hat{\rho}(t). \end{aligned}$$

For (i), $\mathbb{E}[e_{\text{hybrid}}^2] = \mathbb{E}[(s + \xi - \hat{\rho})^2] = \mathbb{E}[(s - \hat{\rho})^2] + \sigma_{\xi}^2$ (the cross term vanishes since $\xi \perp s, \hat{\rho}$, the latter being a function of past ρ_{PINN} values within the training/validation split), while $\mathbb{E}[e_{\text{PINN}}^2] = \text{Var}(s) + \sigma_{\xi}^2$. The hypothesis $\mathbb{E}[(\hat{\rho} - s)^2] < \text{Var}(s)$ then gives $\mathbb{E}[e_{\text{hybrid}}^2] < \mathbb{E}[e_{\text{PINN}}^2]$ directly. The condition $\tau_c \gtrsim w$ is what makes $\mathbb{E}[(\hat{\rho} - s)^2] < \text{Var}(s)$ achievable in principle, since the LSTM can only exploit structure that persists across its input window. For (ii), with $s \equiv 0$,

$$\mathbb{E}[e_{\text{PINN}}^2] = \sigma_{\xi}^2, \quad \mathbb{E}[e_{\text{hybrid}}^2] = \mathbb{E}[(\xi - \hat{\rho})^2] = \sigma_{\xi}^2 - 2 \text{Cov}(\xi, \hat{\rho}) + \mathbb{E}[\hat{\rho}^2];$$

since $\hat{\rho}$ is fitted to minimize training error on a finite noise realization, it necessarily has $\mathbb{E}[\hat{\rho}^2] > 0$ and, on held-out data drawn independently of the training noise, $\text{Cov}(\xi, \hat{\rho}) \rightarrow 0$, giving

$$\mathbb{E}[e_{\text{hybrid}}^2] \geq \sigma_{\xi}^2 = \mathbb{E}[e_{\text{PINN}}^2].$$

□

Proposition 4.2 is precisely the mechanism exercised by the noise-injection experiment of Figure 5 (see Section 4.4): Injecting noise into the training data increases σ_{ξ}^2 relative to $\text{Var}(s)$, pushing the system away from case (i) and toward case (ii), so part (ii) of the Proposition predicts exactly the observed monotonic shrinkage of the hybrid model's advantage as injected noise rises, converging with the PINN-only model once the noise dominates. We regard this agreement between the formal condition and the empirical result as evidence that Figure 5 is measuring the mechanism the proposition describes, rather than an unexplained numerical coincidence.

Consequence for the stability and policy analysis. The established stability and elimination results, together with the numerical threshold-branch analysis, are derived for the continuous-time ODE system (2.10)–(2.17) itself and never reference the LSTM output. Every downstream policy quantity— $\mathcal{R}_{0i}$, the sensitivity indices, the threshold-branch diagrams, the efficacy-coverage surfaces, and $\rho^* \approx 59.6\%$ —is computed from the six PINN-estimated parameters via (3.5) and the formulas of Section 3, not from $\hat{C}_{\text{hybrid}}(t)$. This is architectural: The LSTM is defined as a map on $\hat{C}(t)$ alone, so no “corrected state trajectory” exists for the stability and policy analysis to depend on even in principle. Stage 3 contributes only to the short-term forecast curves of Figure 3 and Supplementary Figure S6, and to the reported RMSE gains; the small residual quantified in Supplementary Figure S5 is a plausibility check on the corrected observable alone, not evidence bearing on the theoretical or policy conclusions, which rest on the exact ODE system throughout.

Ablation study: physics constraint effect versus network-capacity effect. The PINN's improvement over lsqnonlin could in principle derive from continuous-time regularization in general, rather than the ODE residual term $\mathcal{L}_{\text{ODE}}$ in (4.1) specifically. As a first probe, we compared (i) the reduced single-strain SEAI surrogate of Supplementary Figures S3 and S4, fit by nonlinear least squares (a hard, three-scalar physics constraint), against (ii) a flexible feed-forward network with no ODE term, under the same walk-forward protocol as Supplementary Figure S6. We report the result transparently: The unconstrained network achieved lower held-out RMSE by roughly 5–30× (e.g., Delta, day 80: 1458

vs. $\approx 103,834$). This does not isolate a physics-constraint effect, however, since the two fits have vastly different capacities (3 scalars vs. $O(10^3)$ weights); it instead indicates the reduced surrogate (which omits vaccination, waning, and isolation/hospitalization dynamics) is a restrictive approximation over a 15-day window, not evidence against physics-informed regularization in general.

To bound the departure of the Stage-3-corrected observable $\hat{C}_{\text{hybrid}}(t)$ from what the governing equations would imply—rather than merely assert it is small—we evaluated the implied relative residual of Eq (2.14)—the governing equation for the symptomatic compartment $I(t)$ —at both the PINN-only state trajectory and the Stage-3-corrected observable $\hat{C}_{\text{hybrid}}(t)$ (read as a stand-in for that variant's $I(t)$), across the full calibration window for each variant (see Supplementary Figure S5). For Delta (see Supplementary Figure S5(a)), the mean implied relative residual is 0.73% for the PINN-only trajectory and 0.76% for the corrected observable; the largest pointwise divergence between the two curves (≈ 0.56 percentage points) occurs around day 115, within the final 15-day held-out window used for the LSTM's early stopping, where the rolling one-step correction is least anchored by in-sample training data. For Omicron (see Supplementary Figure S5(b)), the mean relative residual is 1.21% (PINN-only) versus 1.20% (corrected observable); here the two curves track each other closely throughout, including at the sharp early-transient peak (day 0–20) and at the local peaks near days 32, 52, and 64. In both variants, this modest, spatially localized departure occurs alongside a substantial forecast-accuracy gain from the LSTM correction (RMSE reductions of $\approx 19.7\%$ and $\approx 25.2\%$ on average across the walk-forward folds of Supplementary Table S6), which we interpret as evidence that the correction captures genuine structure in the residual rather than trading forecast plausibility for an unconstrained fit to noise. We stress again that this is a single-equation plausibility check on the scalar observable the LSTM actually touches, not a verification that a “corrected state trajectory” satisfies the full system (2.10)–(2.17), since no such object is ever constructed (see Section 4.2 and Section 4.4 above).

Capacity-matched ablation. To isolate the physics-constraint effect from the network-capacity effect directly, we additionally trained a capacity-matched ablation network: Identical to the full PINN architecture of Section 4.2 (6 hidden layers $\times$ 64 neurons, tanh activation, linear output, $\approx 21,000$ trainable weights, same Adam $\rightarrow$ L-BFGS optimiser and 50,000-epoch schedule), but with the ODE residual weight set to $w_r = 0$ (see Supplementary Table S1), trained directly on the real Delta and Omicron case-count series under the same walk-forward protocol as Supplementary Figure S6. The full PINN ($w_r = 10^{-2}$) achieved lower held-out 15-day RMSE than the capacity-matched $w_r = 0$ network at every cutoff for both variants (see Supplementary Table S2), with a gap of 10.9–19.3%—markedly smaller and more consistent than the 5–30 $\times$ gap of the capacity-mismatched surrogate ablation above, consistent with a genuine physics-constraint effect. This is statistically significant under a Diebold-Mariano test (see Supplementary Table S4; $p \leq 0.008$ both variants). Physical plausibility gives a second, independent line of evidence: back-solving the ablation network's trajectory for $(\beta_i, \kappa_i, \delta_i)$ yields one negative, physically meaningless rate ($\hat{\delta}_1 < 0$ for Delta) and several values outside the literature-supported ranges of Table 1 (see Supplementary Table S3), whereas every PINN-recovered value remains plausible—indicating the ODE residual constrains individual compartment dynamics, not just the aggregate forecast.

Supplementary Tables S1–S4 collectively document this capacity-matched ablation: The architecture specification, the walk-forward RMSE (a full PINN ahead at every fold), the recovered parameters with physical-plausibility assessment against Table 1, and the Diebold-Mariano test confirm

the statistical significance for both variants.

We perform a walk-forward (rolling-origin) cross-validation with forecast horizon $H = 15$ days: three cutoff days per variant—60, 80, 100 for Delta (of 130) and 45, 65, 85 for Omicron (of 110)—each model re-fitted on data up to that cutoff and scored on the following H -day window (see Supplementary Figure S6). Consecutive cutoffs are spaced 20 days apart against a 15-day horizon, leaving a 5-day gap before the next cutoff and an unscored final segment ($\approx 11.5\%$ of the Delta series); this was chosen so successive training sets differ substantially in length, giving a stringent test of whether the hybrid model's advantage persists as training data varies, though we note this as a limitation of the present three-fold design rather than exhaustive validation. Across all six cutoffs, the ordering

$$\text{RMSE}_{\text{hybrid}} < \text{RMSE}_{\text{PINN-only}} < \text{RMSE}_{\text{lsqnonlin}}$$

holds without exception, confirming the hybrid model's advantage generalizes across independently held-out forecast origins rather than being an artifact of one calibration window.

Supplementary Table S5 displays the pairwise Diebold-Mariano test statistics and p -values comparing the hybrid PINN-LSTM, PINN-only, and lsqnonlin models, pooled across the walk-forward folds of Supplementary Figure S6 for each variant; it explains that the hybrid model's forecast-accuracy advantage over both competing models is statistically significant for Delta and highly significant for Omicron.

To directly test whether the hybrid model's held-out forecast gains reflect genuine structure in the PINN residual rather than an artifact of fitting noise specific to the training window, we repeated the walk-forward evaluation of Section 4.4 under progressively higher levels of injected training-data noise (0–20%), holding all other settings fixed, and compared the resulting held-out forecast RMSE of the PINN-only and hybrid PINN-LSTM models (see Figure 5). If the LSTM's held-out improvement over the PINN-only model derived from exploiting noise idiosyncratic to the clean training data, its relative advantage should persist, or even grow, as additional noise is injected, since a noise-fitting correction would simply have more spurious structure available to exploit. The results show the opposite pattern for both variants.

Figure 5(a) shows the Delta sub-system. At 0% injected noise, the hybrid model (red) achieves a held-out RMSE of $\approx 1.53 \times 10^4$ against PINN-only's (blue) $\approx 1.91 \times 10^4$, a 19.7% reduction. As injected noise increases from 0% to 20%, the PINN-only curve remains essentially flat, fluctuating narrowly around 1.90 – 1.92×10^4 , while the hybrid curve rises steadily and monotonically, from 1.53×10^4 at 0% noise to 1.84×10^4 at 20% noise—nearly closing the gap with the PINN-only model entirely by the highest noise level tested.

Figure 5(b) repeats this analysis for Omicron and shows the same qualitative pattern. At 0% injected noise, the hybrid model attains a held-out RMSE of 3530 against PINN-only's 4730, a 25.2% reduction—the largest gap observed at any noise level. As injected noise rises, the PINN-only curve again remains flat (fluctuating around 4720–4780), while the hybrid curve rises steadily from 3530 at 0% noise to 4730 at 20% noise, converging with the PINN-only model by the upper end of the tested range.

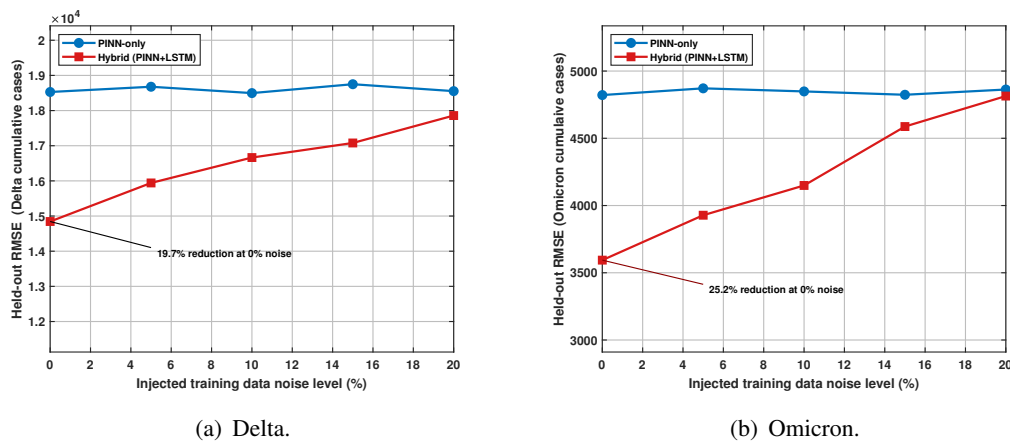


Figure 5. Held-out forecast RMSE of the PINN-only and hybrid PINN-LSTM models under 0–20% injected training-data noise. The hybrid model shows greater improvement at lower noise levels, with gains of 19.7% for Delta and 25.2% for Omicron at 0% noise, and converges with the PINN-only model near 20% noise. (a) Delta. (b) Omicron.

In both variants, the hybrid model’s advantage is therefore largest precisely when the training data is cleanest, and degrades monotonically as noise is added—the opposite of what would be expected if the LSTM correction were overfitting to noise in the training window. This is consistent with the LSTM tracking a genuine, structured signal in the PINN residual (Figure 4) rather than absorbing artifacts of a particular noise realization. These results provide additional evidence that the LSTM correction captures structured temporal variation rather than noise (see Section 4.4).

Table 4 complements the Diebold-Mariano and AIC/BIC comparisons of Supplementary Tables S5 and S6 with two distribution-free tests applied to the six cross-validated held-out RMSE values (three walk-forward folds for each variant). The hybrid PINN-LSTM model attains the lowest held-out RMSE in all six folds against both the PINN-only and the *lsqnonlin* models. Under the one-sided sign test and the one-sided Wilcoxon signed-rank test, this gives $p = 0.0156$ in each comparison, which is the smallest p -value attainable with six folds. Because neither test depends on the number of fitted parameters k , this result does not rely on the choice of AIC/BIC penalty and therefore provides independent support for the forecast advantage of the hybrid model.

Table 4. Sign test and Wilcoxon signed-rank test (one-sided) on the cross-validated held-out RMSE values of Supplementary Table S6, providing a parameter-count-independent complement to the AIC/BIC comparison in that table. Neither test involves k , AIC, or BIC.

Comparison	Folds won by hybrid	Sign test p	Wilcoxon p
hybrid vs. PINN-only	6/6	0.0156	0.0156
hybrid vs. <i>lsqnonlin</i>	6/6	0.0156	0.0156

4.5. Sensitivity analysis

Sensitivity analysis identifies which model parameters most strongly influence the basic reproduction numbers $\mathcal{R}_{01}$ and $\mathcal{R}_{02}$, thereby guiding the prioritization of public-health interventions.

We employ two complementary approaches: a local analysis based on normalized forward sensitivity indices (NSIs), which measures the proportional response of $\mathcal{R}_{0i}$ to small perturbations in each parameter at the baseline, and a global analysis based on the partial rank correlation coefficient (PRCC) with Latin hypercube sampling (LHS), which captures the joint nonlinear effect of simultaneous parameter variation across their full plausible biological ranges [54]. Together, these two methods provide both a tractable analytical ranking and a statistically robust global assessment of parameter influence, ensuring that the resulting public-health recommendations are not artifacts of the chosen baseline.

4.5.1. Local sensitivity: normalized forward sensitivity indices

The normalized forward sensitivity index (NSI) provides a local, derivative-based measure of how strongly the basic reproduction number responds to small relative perturbations in each model parameter. For parameter p and output $\mathcal{R}_{0i}$,

$$\Upsilon_p^{\mathcal{R}_{0i}} = \frac{\partial \mathcal{R}_{0i}}{\partial p} \cdot \frac{p}{\mathcal{R}_{0i}}, \quad (4.10)$$

representing the percentage change in $\mathcal{R}_{0i}$ per one-percent change in p at the baseline [54]. Positive (negative) values indicate transmission-amplifying (suppressing) parameters, magnitudes near unity denote highly elastic relationships, and magnitudes near zero denote negligible influence [23]. Although the NSI is a local measure, its sign and magnitude provide a useful first-screening step before the full global analysis [54]. The indices $\Upsilon_p^{\mathcal{R}_{0i}}$ are computed analytically from the closed-form expression (3.5) for $i = 1, 2$ at the baseline values in Table 1.

Figure 6(a) presents the NSI of $\mathcal{R}_{01}$ for the Delta sub-system. β_1 has the largest positive index ($\approx +1.000$), confirming $\mathcal{R}_{01} \propto \beta_1$, followed by ζ_1 ($\approx +0.769$), η_1 ($\approx +0.623$), ϵ_1 ($\approx +0.231$), π ($\approx +0.080$), μ ($\approx +0.038$), and α ($\approx +0.020$), underlining the joint role of both carrier types in transmission [37]. On the suppression side, vaccine efficacy ξ_1 dominates by a wide margin (≈ -0.864), followed by δ_1 (≈ -0.533), γ_{A_1} (≈ -0.230), κ_1 (≈ -0.187), τ_1 (≈ -0.042), and ν (≈ -0.039), confirming that faster clinical progression, isolation, and case removal each suppress Delta spread.

Figure 6(b) shows the corresponding indices for Omicron. The structure is qualitatively similar but quantitatively distinct: β_2 again dominates ($\approx +1.000$), with η_2 ($\approx +0.609$), ζ_2 ($\approx +0.594$), ϵ_2 ($\approx +0.296$), α ($\approx +0.100$), π ($\approx +0.030$), and μ ($\approx +0.026$) following. The closer η_2, ζ_2 values relative to Delta are consistent with Omicron's shorter generation interval and higher symptomatic shedding [20]. Vaccine efficacy ξ_2 again dominates suppression (≈ -0.925) [55], followed by δ_2 (≈ -0.399), γ_{A_2} (≈ -0.303), κ_2 (≈ -0.269 , visibly stronger than $\kappa_1 \approx -0.187$, reflecting the greater importance of rapid isolation against the faster-spreading variant), ν (≈ -0.026), and τ_2 (≈ -0.025 , least influential).

Taken together, Figures 6(a) and (b) consistently identify β_i , ζ_i , and η_i as the leading amplifying parameters, and ξ_i , δ_i , γ_{A_i} , κ_i , and ν as the leading suppressing parameters for both variants, with ξ_i the dominant suppression parameter in both. These local rankings motivate the global PRCC analysis below.

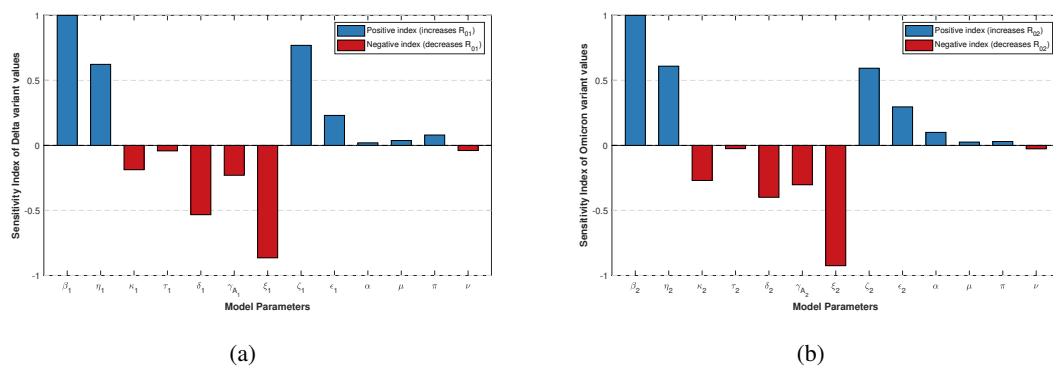


Figure 6. Normalized forward sensitivity indices of the variant-specific reproduction numbers at the baseline. Blue: positive (increases R_0); red: negative (decreases R_0). (a) Delta sub-system R_{01} . (b) Omicron sub-system R_{02} .

4.5.2. Global sensitivity: PRCC with Latin hypercube sampling (LHS)

The NSI is a local measure derived from linearization at the baseline and does not capture joint parameter effects across plausible biological ranges. To address this, we perform a global sensitivity analysis using the PRCC with LHS [54]. The PRCC quantifies the monotonic nonlinear relationship between each input and the output while controlling for all other parameters, and is considered the most robust sampling-based sensitivity measure for monotonic mappings [54]. Values near +1 (−1) indicate strong positive (negative) monotonic influence; following [54], $|\text{PRCC}| > 0.4$, $0.2 < |\text{PRCC}| \leq 0.4$, and $|\text{PRCC}| \leq 0.2$ denote highly, moderately, and weakly influential parameters, respectively. For each variant, $N = 8000$ samples are drawn via LHS, with parameter-specific ranges rather than a single flat percentage: $\alpha \in [0.071, 0.33]$ and $\eta_i \in [0.15, 0.7]$ are sampled over their literature-reported biological ranges (see Table 1), while all remaining parameters are sampled over $\pm 20\%$ around their baseline values, including $\nu = 0.0033/\text{day}$ (see Table 1). PRCC values are reported at $p < 0.05$ significance for all parameters against R_{01} (Delta) and R_{02} (Omicron); results are shown in Figures 7(a) and 7(b).

Figure 7(a) presents the PRCC results for Delta against R_{01} . Vaccine efficacy ξ_1 exhibits the strongest negative PRCC (≈ -0.973), the single most influential suppression parameter, followed by δ_1 (≈ -0.434), κ_1 (≈ -0.280), and γ_{A_1} (≈ -0.210 , all moderately influential); π (≈ -0.119), τ_1 (≈ -0.069), ν (≈ -0.025), and α (≈ -0.010) are all weak in magnitude. On the amplifying side, η_1 attains the highest positive PRCC ($\approx +0.914$), surpassing β_1 ($\approx +0.717$), ζ_1 ($\approx +0.560$), and ϵ_1 ($\approx +0.333$); the large ζ_1, η_1 values underscore the role of undetected asymptomatic carriers [37] in motivating broad testing and contact-tracing μ ($\approx +0.019$) is small in magnitude.

Figure 7(b) presents the corresponding results for Omicron against R_{02} . Vaccine efficacy ξ_2 again dominates (≈ -0.969) [55]; despite $\xi_2 = 0.70 < \xi_1 = 0.78$, Omicron's higher intrinsic transmissibility still produced a large-scale third wave, emphasising that coverage is as critical as efficacy. The isolation rate κ_2 shows a substantially stronger negative PRCC (≈ -0.732 , the second most influential suppression parameter after ξ_2) than κ_1 (≈ -0.280), reflecting the greater importance of rapid isolation against the faster-spreading variant. δ_2 (≈ -0.515) is likewise highly influential, γ_{A_2} (≈ -0.420) similarly so, and τ_2 (≈ -0.056) is weak in magnitude. ν (≈ -0.051), π (≈ -0.016), α (≈ -0.014), and μ ($\approx +0.010$) are likewise weak. On the amplifying side, η_2 attains the highest positive PRCC

($\approx +0.927$), followed by β_2 ($\approx +0.837$), ζ_2 ($\approx +0.643$), and ϵ_2 ($\approx +0.573$), confirming the dual role of asymptomatic and symptomatic carriers.

Figures 7(a) and 7(b) consistently identify ξ_i , η_i , β_i , ζ_i , and ϵ_i as the most influential parameters, broadly concordant with the local NSI rankings of Figure 6. The global analysis additionally reveals that κ_i is more influential for Omicron than the local NSI suggests. Together these findings show that interventions improving vaccine efficacy and coverage, reducing effective contact rates, and detecting asymptomatic carriers will have the greatest impact on controlling both variants in India.

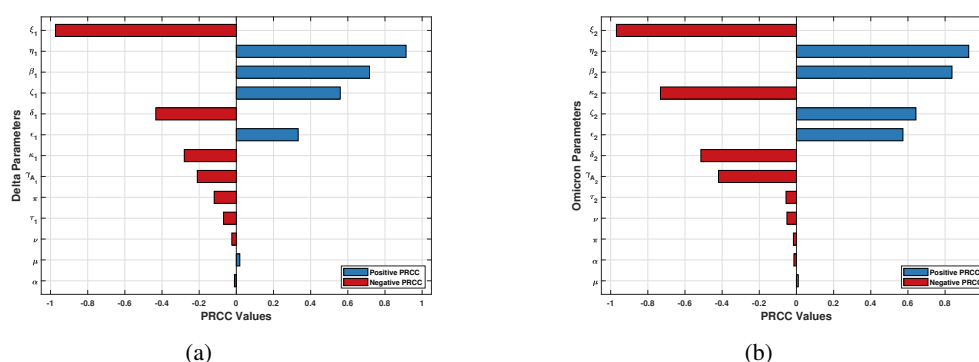


Figure 7. Global PRCC sensitivity analysis with Latin hypercube sampling ($N = 8000$, parameter-specific ranges: $\alpha \in [0.071, 0.33]$, $\eta_i \in [0.15, 0.7]$ from Table 1; all remaining parameters sampled at $\pm 20\%$ around baseline, with ν evaluated at $\nu = 0.0033/\text{day}$, Table 1). Blue: positive PRCC; red: negative PRCC; whiskers denote 95% confidence intervals. (a) Delta parameters versus $\mathcal{R}_{01}$. (b) Omicron parameters versus $\mathcal{R}_{02}$.

4.6. Threshold-branch analysis

To examine the qualitative behavior of model (2.10)–(2.17) near the epidemic threshold, we use a numerical threshold-branch construction for $\mathcal{R}_{01}$ (Delta) and $\mathcal{R}_{02}$ (Omicron). The rigorous results established above are that the DFE is locally asymptotically stable for $\mathcal{R}_0 < 1$, conditionally globally asymptotically stable under (C2), and unstable for $\mathcal{R}_0 > 1$, while each single-variant endemic equilibrium exists uniquely for its corresponding $\mathcal{R}_{0i} > 1$ and is locally stable only under the explicit Routh-Hurwitz and non-invasion conditions of Theorem 3.5. These results are consistent with a forward exchange of stability at the threshold, but no separate center-manifold normal-form proof of a transcritical bifurcation is claimed here. To make the plotted quantity explicit, we retain $\bar{I}_i$ as a normalized reference branch rather than the exact endemic prevalence I_i^*/N^* . The normalized reference branch used in Figure 8 is

$$\bar{I}_i = \begin{cases} 1 - \frac{1}{\mathcal{R}_{0i}}, & \mathcal{R}_{0i} > 1, \\ 0, & \mathcal{R}_{0i} \leq 1. \end{cases} \quad (4.11)$$

The numerical branch construction is consistent with a forward threshold exchange, but it is not a rigorous proof of a transcritical bifurcation and is not used to claim that backward bifurcation has been rigorously excluded. The analytical results instead establish the relevant single-variant threshold at $\mathcal{R}_{0i} = 1$: a positive single-variant endemic equilibrium exists only for $\mathcal{R}_{0i} > 1$, while its local stability additionally requires the conditions of Theorem 3.5. A full center-manifold normal-form calculation

is outside the scope of the present study.

The condition $\mathcal{R}_{0i} < 1$ used in this subsection is the single-variant threshold obtained from the analytical DFE stability and endemic-existence results. It is distinct from the competitive invasion condition at a single-variant endemic equilibrium, which must be evaluated using the invasion reproduction number in Theorem 3.5. At the Indian baseline, both reported reproduction numbers exceed unity. Consequently, the threshold plots below are not intended to describe the outcome of simultaneous Delta-Omicron circulation; they only visualize the single-variant threshold structure and the corresponding reported baseline operating points.

Figure 8(a) displays the Delta threshold construction. The blue branch represents the disease-free state below threshold, while the red curve is the normalized reference branch used to visualize the positive single-variant branch above threshold. The analytical result that guarantees the positive equilibrium is Theorem 3.4; its local stability additionally requires Theorem 3.5. The reported baseline value $\mathcal{R}_{01} = 1.87$ marks the calibrated operating point on the reference threshold diagram. The plotted $\bar{I}_1$ value is only a visual normalization; the exact endemic prevalence, when the equilibrium exists, is obtained from the positive root of $g_1(\lambda)$ in (3.17) through $I_1^*/N^* = \lambda^*/M_1$.

Figure 8(b) provides the analogous Omicron threshold construction. The point $(\mathcal{R}_{02}, \bar{I}_2) = (1, 0)$ is the threshold point in the reference representation. The calibrated value $\mathcal{R}_{02} = 1.64$ is shown only as the operating point of the reference diagram. Because the reported $\mathcal{R}_{0i}$ values already incorporate the baseline model configuration, these single-strain diagrams should not be interpreted as a second, independent intervention calculation. The comparison between Delta and Omicron is instead made in the counterfactual proportional-intervention analysis below.

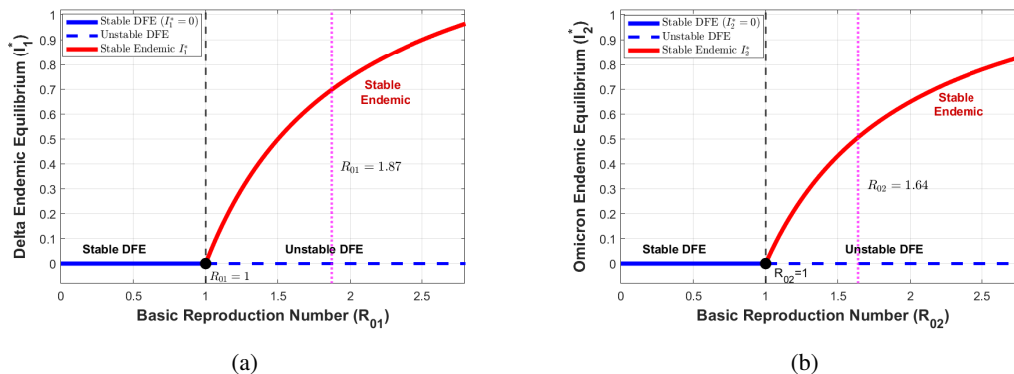


Figure 8. Numerical forward threshold-branch diagrams. Solid blue: stable disease-free branch; dashed blue: unstable disease-free branch; solid red: normalized reference endemic-shaped branch; filled black circle: threshold point $(\mathcal{R}_{0i}, \bar{I}_i) = (1, 0)$; magenta dotted line: Indian baseline estimate. The red curve is a normalized reference branch and does not represent the exact endemic prevalence of the full eleven-state model. (a) Delta: $\mathcal{R}_{01} = 1.87$, $\bar{I}_1 = 0.46$; (b) Omicron: $\mathcal{R}_{02} = 1.64$, $\bar{I}_2 = 0.39$.

4.6.1. Joint effect of vaccine efficacy and waning immunity

The proportional scenario analyses above assume a constant vaccine efficacy ξ_i ; this scenario construction does not itself replace the model-specific next-generation matrix. We additionally examine the well-documented decline of neutralizing antibody titers over time [5]: Effectiveness

against Omicron infection drops from $\approx 28\%$ at three months post-vaccination to $< 4\%$ at six months without boosters [5], motivating an explicit treatment of the waning rate ω_V . Treating steady-state coverage as

$$\rho_{ss}(\nu, \omega_V) = \frac{\nu}{\nu + \mu + \omega_V}, \quad (4.12)$$

and substituting into the proportional scenario index (4.13), yields $R_{0i}^{\text{eff}}(\xi_i, \omega_V)$ at fixed $\nu = \nu_b$. Figure 9 therefore shows how the scenario benchmark changes with waning immunity: (i) at long immunity duration $1/\omega_V \geq 300$ days, even moderate efficacy ($\xi \geq 0.6$) can place the proportional scenario index below unity for both variants; (ii) at short duration $1/\omega_V \leq 180$ days, even high efficacy may fail to reach the scenario threshold for Delta without additional interventions. The Indian baseline of $1/\omega_V = 180$ days is therefore used as a reference operating point. These statements concern the counterfactual scenario index only and do not establish competitive elimination in the full ODE system.

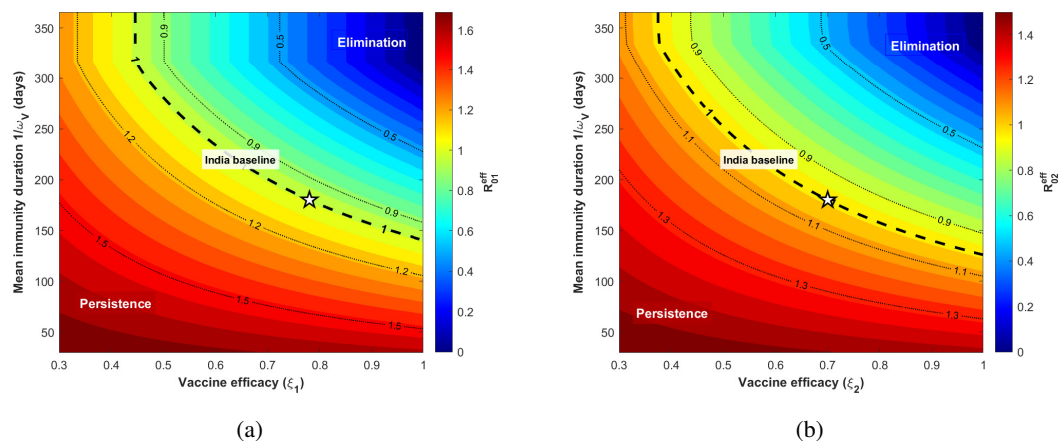


Figure 9. Joint effect of vaccine efficacy ξ_i and mean immunity duration $1/\omega_V$ on R_{0i}^{eff} . Dashed contour: $R_{0i}^{\text{eff}} = 1$ (elimination boundary); white star: Indian baseline ($1/\omega_V = 180$ days). (a) Delta. (b) Omicron.

4.6.2. Effective reproduction-number surfaces in the efficacy-coverage plane

To complement the herd-immunity threshold curves of Figure 11, Figure 10 visualizes a counterfactual proportional intervention index across the efficacy-coverage plane. Because the baseline $\mathcal{R}_{0i}$ used elsewhere in the manuscript is already evaluated under the baseline vaccination/isolation configuration, the quantity below is not interpreted as a second application of vaccination to the next-generation matrix. It is used only as a transparent scenario index for comparing proportional reductions from the reported baseline values. Figure 10(a) shows the Delta heatmap: the persistence region (warm colors) occupies the lower-left section, where coverage and efficacy are both low, and the elimination region (cool colors) occupies the upper-right section; the white star marks the Indian operating point $(\xi_1, \rho_1^*) = (0.78, 59.6\%)$, lying precisely on the threshold. Figure 10(b) shows the Omicron heatmap, with operating point $(\xi_2, \rho_2^*) = (0.70, 55.7\%)$, also on the threshold. For the common efficacy-coverage scenario used in these heatmaps, the Delta surface lies above the Omicron surface at the reported baseline ordering. This shows that Delta is the more restrictive variant within this specific proportional scenario. It does not establish a general competitive-exclusion result for the full two-strain model

because the two variants have different vaccine efficacies ($\xi_1 \neq \xi_2$) and the baseline R_{0i} values already incorporate vaccination and isolation. Thus, the phrase “binding constraint” below is used only for this specified scenario, not as a structural property of the original ODE system.

$$R_{0i}^{\text{eff}}(\xi_i, \rho) = R_{0i}[1 - \xi_i\rho], \quad i = 1, 2. \quad (4.13)$$

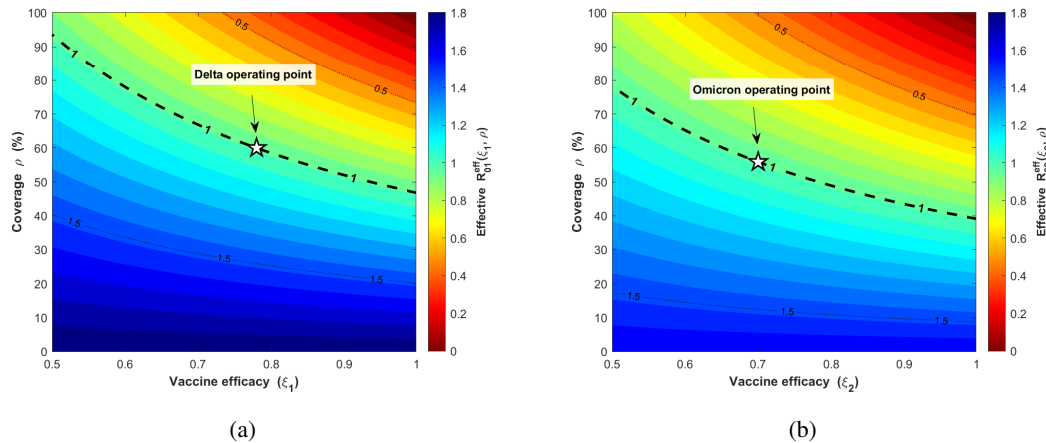


Figure 10. Heatmaps of $R_{0i}^{\text{eff}}(\xi_i, \rho) = R_{0i}(1 - \xi_i\rho)$. Dashed contour: elimination boundary $R_{0i}^{\text{eff}} = 1$; white star: Indian operating point. (a) Delta: $(\xi_1, \rho_1^*) = (0.78, 59.6\%)$. (b) Omicron: $(\xi_2, \rho_2^*) = (0.70, 55.7\%)$.

4.6.3. Counterfactual efficacy-coverage threshold curves

Setting the counterfactual proportional-intervention index in (4.13) equal to unity and solving for coverage gives

$$\rho^*(\xi_i) = \frac{1}{\xi_i} \left(1 - \frac{1}{R_{0i}} \right), \quad i = 1, 2. \quad (4.14)$$

This relation is a transparent benchmark for a proportional reduction applied to the reported baseline reproduction numbers. It is not a second next-generation-matrix derivation of herd immunity: The baseline R_{0i} already corresponds to the baseline intervention configuration, and a genuine change in coverage would alter the DFE and the next-generation matrix.

Substituting the reported baseline values $R_{01} = 1.87$ and $R_{02} = 1.64$ into (4.14) gives the two scenario threshold curves in Figure 11. At the reported operating efficacies, the point benchmarks are $\rho_1^* = 59.6\%$ for Delta and $\rho_2^* = 55.7\%$ for Omicron. Thus, within this specified proportional scenario, Delta is the more restrictive variant and the larger coverage benchmark is $\max\{\rho_1^*, \rho_2^*\} = 59.6\%$. This scenario-level result should not be interpreted as a proof of competitive elimination in the full two-strain ODE system.

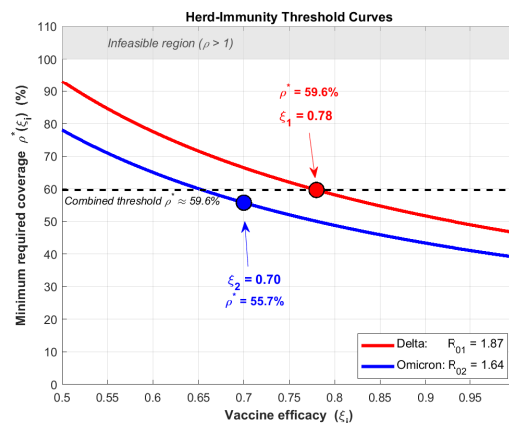


Figure 11. Counterfactual proportional-intervention coverage thresholds $\rho^*(\xi_i) = (1 - 1/R_{0i})/\xi_i$. Red: Delta ($R_{01} = 1.87$); blue: Omicron ($R_{02} = 1.64$). Dashed line: combined scenario benchmark $\rho^* \approx 59.6\%$. Filled circles: reported operating efficacies.

The filled circles mark the reported operating efficacies $\xi_1 = 0.78$ and $\xi_2 = 0.70$. The difference between the two point benchmarks is approximately 3.9 percentage points. The portion of the curves requiring $\rho > 100\%$ is biologically infeasible and simply indicates that the specified proportional intervention cannot reach the scenario target through coverage alone.

Uncertainty propagation to the scenario threshold. The point estimates $\rho_1^* = 59.6\%$ and $\rho_2^* = 55.7\%$ use the reported posterior medians $R_{01} = 1.87$ and $R_{02} = 1.64$. Since (4.14) is monotone in R_{0i} , propagating the reported 95% B-PINN intervals $R_{01} \in [1.759, 2.014]$ and $R_{02} \in [1.558, 1.732]$ gives $\rho_1^* \in [55.3\%, 64.6\%]$ and $\rho_2^* \in [51.2\%, 60.4\%]$. The combined point benchmark is therefore $\rho^* \approx 59.6\%$ with the reported interval $[56.0\%, 64.6\%]$. This is a monotone-transform propagation of reported endpoint intervals, not a full joint posterior-predictive interval.

Joint propagation including α, η_i uncertainty. The interval above propagates uncertainty in $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$ only. Table 5 additionally samples α and η_i over their stated literature ranges. The resulting proportional-scenario benchmark is $[1.1\%, 64.8\%]$ (median 47.0%), showing that uncertainty in fixed biological parameters materially widens the benchmark. This does not convert the scenario calculation into a full-model herd-immunity threshold; a policy-specific threshold would require recomputation of the next-generation matrix under the changed intervention policy.

Table 5. Impact of the joint uncertainty propagation (Table 3) on the counterfactual proportional-intervention threshold $\rho^*(\xi_i) = (1 - 1/R_{0i})/\xi_i$. Negative raw values of ρ^* (arising when a draw gives $R_{0i} < 1$, i.e., no positive proportional reduction is required) are clipped to 0%.

	Point estimate	95% CI (six learned parameters only)	95% CI (joint propagation)
ρ_1^* (Delta)	59.6%	[55.3%, 64.6%]	[0%, 64.8%]
ρ_2^* (Omicron)	55.7%	[51.2%, 60.4%]	[0%, 55.7%]
$\rho^* = \max\{\rho_1^*, \rho_2^*\}$	59.6%	[56.0%, 64.6%]	[1.1%, 64.8%] (median 47.0%)

These findings provide a conditional policy benchmark rather than a rigorous full-model herd-immunity guarantee: Under the proportional efficacy-coverage scenario, the reported point threshold is

59.6% (95% CrI: [56.0%, 64.6%]). A policy-level threshold for the original ODE system would require recomputing the next-generation matrix after changing ν , ξ_i , and ω_V jointly. Improvements in vaccine efficacy are nevertheless expected to reduce the required coverage within the proportional scenario.

To complement the threshold and waning analyses with a direct quantitative comparison of how each variant responds to specific interventions, Figure 12 presents the peak symptomatic infection burden for both Delta (red, circle markers) and Omicron (blue, square markers) as a function of four epidemiologically critical parameters: the vaccination rate ν , the vaccine efficacy ξ , the isolation rate κ , and the transmission rate β . For each panel, the corresponding parameter is varied across its biologically plausible range while all other parameters are held at their baseline values from Table 1. To ensure a fair comparison free of inter-variant competition effects, each variant is simulated independently using a single-strain version of the two-strain model with identical small-seed initial conditions, and the resulting peak $\max_t I_i(t)$ is computed by numerical integration of the ODE system over a 300-day horizon. Vertical dotted lines indicate the operating values estimated for each variant in Section 4.4. We emphasize that this single-strain simulation deliberately isolates each variant's own threshold behavior ($\mathcal{R}_{0i} < 1$, Definition 1(a)) and does not capture the competitive elimination mechanism of Definition 1(b); the two-strain competitive threshold is instead addressed in Sections 4.6.2 and 4.6.3 above.

Figure 12(a) presents the peak infection burden as a function of the vaccination rate ν , varied from no vaccination ($\nu = 0$) to $\nu = 0.05$ per day, bracketing the Indian operating point $\nu = 0.0033$ per day. At the lowest vaccination rate, both variants produce substantial outbreaks, with peak symptomatic burdens of approximately 13.5×10^6 (Delta) and 15.4×10^6 (Omicron), reflecting the baseline reproduction numbers $R_{01} = 1.87$, $R_{02} = 1.64$ together with Omicron's larger asymptomatic fraction and infectivity factors ($\eta_2 = 0.16$, $\zeta_2 = 0.5908$ vs. $\eta_1 = 0.15$, $\zeta_1 = 0.3063$). As ν increases, both peaks decline, but Omicron's curve remains uniformly above Delta's: At $\nu = 0.0033$, the peaks are $\approx 10.1 \times 10^6$ (Delta) vs. $\approx 12.6 \times 10^6$ (Omicron), and even at $\nu = 0.05$, Omicron's peak remains substantial ($\approx 4.65 \times 10^6$) while Delta's continues to decline ($\approx 1.85 \times 10^6$)—reversing the naive expectation from $R_{02} < R_{01}$ alone, since vaccination alone does not bring Omicron close to elimination. The Indian operating point at $\nu = 0.0033$ per day positions both variants well above their respective elimination thresholds, underscoring that current vaccination speed alone is insufficient for eliminating either variant without complementary interventions.

Figure 12(b) examines how peak infections respond to variations in vaccine efficacy ξ across the range $[0.3, 1.0]$, with all other parameters held at their baseline values. Both curves decline with increasing ξ , but Omicron's peak (blue) lies uniformly above Delta's (red), reflecting Delta's higher vaccine efficacy together with Omicron's larger asymptomatic fraction and infectivity factors. At the operating efficacies $\xi_1 = 0.78$ (Delta), $\xi_2 = 0.70$ (Omicron), the peak burdens are $\approx 1.19 \times 10^7$ and $\approx 1.32 \times 10^7$, respectively: Omicron's peak exceeds Delta's despite Delta's lower ξ_i , confirming the gap is driven jointly by Omicron's lower efficacy and its higher η_2 , ζ_2 , ϵ_2 , not efficacy alone. Even at $\xi = 1.0$, both variants retain substantial peaks (Delta $\approx 0.96 \times 10^7$, Omicron $\approx 1.05 \times 10^7$), indicating that even a perfect vaccine cannot drive either variant to elimination at the current vaccination rate of $\nu = 0.0033$ per day—a finding that reinforces the need for combined interventions targeting both vaccine performance and rollout speed.

Figure 12(c) demonstrates the influence of the isolation rate κ on peak infection burden, with κ varied from 0.10 to 0.50 per day, corresponding to mean isolation delays ranging from 2 to 10 days.

Omicron's peak (blue) lies uniformly above Delta's (red) across the full range, with both variants exhibiting pronounced sensitivity to isolation, declining sharply from $\approx 12.3 \times 10^6$ (Delta) and $\approx 14.2 \times 10^6$ (Omicron) at $\kappa = 0.10$ to $\approx 2.6 \times 10^6$ – 3.0×10^6 (both variants, converging) at $\kappa = 0.50$. Notably, at each variant's own estimated operating value, the ordering reverses: Delta's peak at $\kappa_1 = 0.132$ ($\approx 9.7 \times 10^6$) exceeds Omicron's at $\kappa_2 = 0.178$ ($\approx 8.5 \times 10^6$), showing that Omicron's more aggressive estimated isolation policy offsets what would otherwise be an even larger epidemic than Delta's, consistent with the stronger negative PRCC for κ_2 (≈ -0.732) than κ_1 (≈ -0.280 , Section 4.5). Beyond $\kappa \approx 0.30$ per day, both peaks saturate at low values, indicating diminishing returns from further acceleration of case isolation. Both Indian operating values position their variant in the steep portion of the response curve, where modest isolation improvements yield substantial peak-burden reductions.

Figure 12(d) shows the peak infection burden as a function of the transmission rate β , varied across the range $[0.4, 1.2]$ per day to span both sub-epidemic and super-epidemic regimes. Neither curve shows a clearly demarcated near-zero elimination region: both variants already exceed 4×10^6 cases at the lowest β considered, reflecting the baseline vaccination, isolation, and infectivity parameters. At the operating values $\beta_1 = 0.8946$ and $\beta_2 = 0.8686$, the peaks are $\approx 1.06 \times 10^7$ (Delta) and $\approx 1.33 \times 10^7$ (Omicron), underscoring that the intrinsic transmission rate is the dominant driver of peak burden in unsuppressed regimes, but suppression through vaccination and isolation can substantially shift the burden curve, providing an important mechanism for controlling future variants with even higher transmissibility.

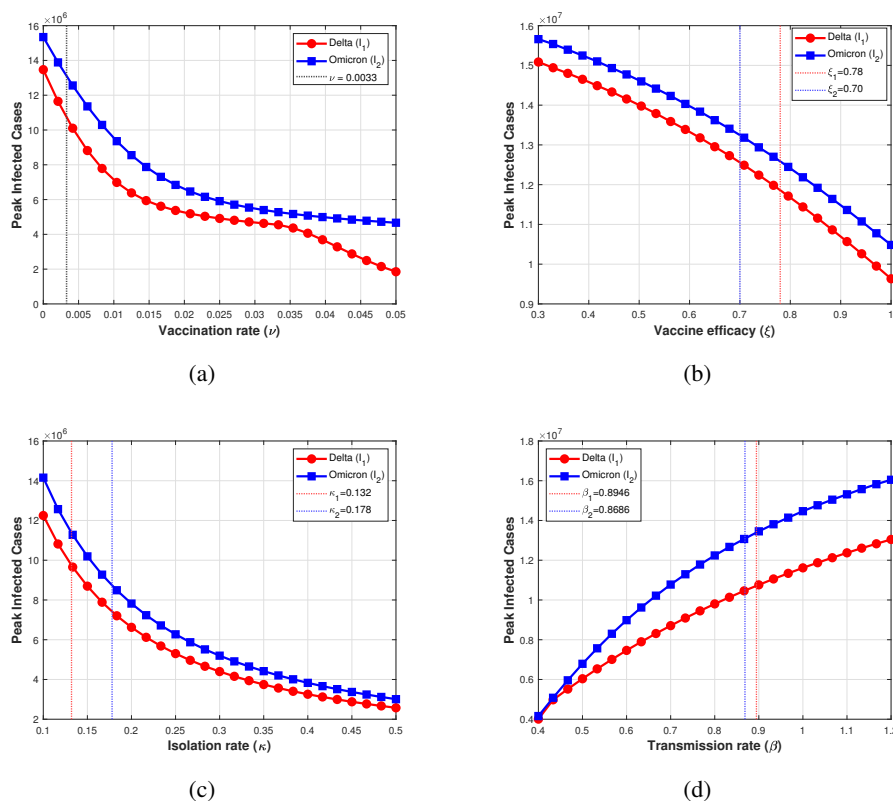


Figure 12. Peak infections versus four key intervention parameters for Delta (red, circles) and Omicron (blue, squares). (a) Vaccination rate v . (b) Vaccine efficacy ξ . (c) Isolation rate κ . (d) Transmission rate β . Dotted lines mark estimated operating values.

5. Results and discussion

This section does not restate the numerical results of Section 4 in detail; those are reported once, with full derivations and figures, at the point each analysis is performed. Instead, we (i) briefly recall, by theorem/section number only, which claims were established and where; (ii) synthesize the two-sided public-health conclusion that only emerges once the stability, uncertainty-quantification, and sensitivity results are read together; and (iii) position these findings explicitly against the closest recent literature, which is not attempted in Section 4.

Recap of the analytical and computational findings. Positivity, boundedness, and $\mathcal{R}_{0i}$ are established in Theorem 3.1, Section 3.2; local/global stability of the disease-free equilibrium in Theorems 3.2 and 3.3; existence and stability of the single-variant endemic equilibria in Theorems 3.4 and 3.5, with coexistence left as Conjecture 1; and the elimination distinction formalized in Definition 1. The threshold-branch analysis (see Section 4.6) is consistent with a forward threshold exchange; a full center-manifold normal-form bifurcation proof is left for future work. Calibration (see Section 4.4) yielded $\mathcal{R}_{01} = 1.87$, $\mathcal{R}_{02} = 1.64$, with the hybrid PINN-LSTM reducing 30-day RMSE by $\approx 23\%$ over the PINN-only method and $\approx 40\%$ over lsqnonlin—a gain shown to shrink under injected noise (Figure 5) rather than persist as overfitting, confirmed by non-parametric tests (see Table 4) and AIC/BIC (see Supplementary Table S6). The B-PINN (see Section 4.3) delivers 95% credible intervals on all six parameters and $\mathcal{R}_{0i}$ (Figure 2), both excluding unity; PRCC/NSI (see Section 4.5) ranks vaccine efficacy ξ_i as the dominant suppressive parameter and asymptomatic fraction η_i as the dominant amplifying one; and the counterfactual efficacy-coverage analysis gives a proportional-intervention benchmark of $\rho^* = 59.6\%$ (reported interval [56.0%, 64.6%]), with Delta more restrictive within that specified scenario.

The two-sided public-health synthesis. Reading the stability theory together with the intervention-response curves of Section 4.6 yields a conclusion not visible from either alone: Delta is the more restrictive variant in the specified proportional-intervention scenario ($\mathcal{R}_{01} > \mathcal{R}_{02}$), whereas Omicron is the binding constraint for near-term peak-burden control (its peak burden is uniformly higher across the vaccination-rate range, see Figure 12(a), driven by its larger asymptomatic fraction and infectivity factors rather than vaccine efficacy, since $\xi_2 = 0.70 < \xi_1 = 0.78$). Neither half follows from the reproduction-number ordering alone, and neither is reported in the closest prior Delta-Omicron models; it is, to our knowledge, a novel synthesis-level finding obtained because the elimination theory of Section 3 and the uncertainty-quantified intervention curves of Section 4 are developed within one mutually consistent framework.

Positioning relative to recent Delta-Omicron literature. The closest prior studies differ in scope rather than in disagreeing with our conclusions. Ngonghala et al. [21] established, for the U.S., that Omicron competitively excluded Delta once its control reproduction number exceeded Delta's—the mechanism formalized here as Definition 1(b)—but without a Bayesian pipeline or efficacy-coverage threshold surface. Negi et al. [22] modeled four Indian variants with optimal control, but without global stability analysis, threshold-branch analysis, or uncertainty-quantified estimation. Saha and Saha [11] and Ambalarajan et al. [12] examined multi-strain co-circulation with vaccination, but neither propagated parameter uncertainty onto $\mathcal{R}_{0i}$ nor reported identifiability diagnostics. The present contribution fills this specific combination—proven stability theory, a threshold-branch analysis, and

an uncertainty-quantified pipeline with identifiability diagnostics, and a two-sided elimination/peak-burden synthesis—for the Indian Delta-Omicron pair (see also “On the locus of novelty,” Section 1).

Limitations. The model assumes homogeneous mixing and does not stratify by age, comorbidity, or geographical region; the fixed PINN architecture and loss weights could be refined through adaptive weighting [51] and architecture search. We note the following further points:

- Proposition 4.1 (see Section 4.4) bounds the implied residual of the corrected observable in terms of the Stage-2 PINN residual and the magnitude/Lipschitz rate of the LSTM correction, but it is a worst-case bound rather than a tight one; a data-driven or probabilistic refinement (e.g., replacing the sup-norm bounds on ε, L with their empirical distributions over the validation window) is left for future work. The PINN/B-PINN use a fixed Adam+L-BFGS schedule and a static 6×64 architecture, which may under-resolve the differing Delta/Omicron time scales; adaptive/event-triggered training, attention-based architectures, graph- or network-based relaxations of homogeneous mixing, and fusion of additional data streams (mobility, genomic surveillance, hospitalization records) are natural extensions not pursued here.
- The elimination analysis (Sections 4.6–4.6.3) is restricted to the two-variant, non-coexisting case; the borderline regime $\mathcal{R}_{01} = \mathcal{R}_{02}$, where the coexistence equilibrium of Conjecture 1 would apply instead, is noted in Definition 1 but not further analyzed.
- The coexistence-equilibrium existence argument (Conjecture 1) remains a sign-change/monotonicity sketch, not a fully rigorous joint existence–uniqueness proof for (3.16); a Poincaré–Miranda- or fixed-point-based treatment is a natural next step and has not yet been carried out.
- Practical and formal (structural) identifiability of the six variant-specific parameters from the single observable $C(t) = I_1 + I_2$ is established in Section 4.2 (see Table 2 and Supplementary Figures S1–S4), subject to the scope limits stated there. A joint structural-identifiability analysis of the full 12-parameter, two-observable coupled system (rather than the two separate 8-state sub-models analyzed here), and collection of continuous daily J, H, R series to further strengthen this result under the full joint model, remain future work.
- Network depth/width, LSTM window length, and loss weights w_d, w_r, w_{ic} were fixed by standard practice and shown robust to moderate perturbation, but were not chosen via systematic grid search or cross-validated comparison.
- Only the counterfactual proportional-intervention threshold ρ^* carries a propagated credible interval; the endemic prevalences $\bar{I}_i$ (see Figure 8) and the intervention-response curves (see Figures 9, 10, and 12) remain point estimates at posterior medians.

The framework is restricted to Delta and Omicron but is fully extensible to additional sub-lineages (BA.2, BA.4/5, XBB) given variant-specific surveillance data. Despite these limitations, the proposed framework couples mechanistic epidemiological rigor with scientific machine learning to deliver physically consistent, statistically quantified, and forecast-superior multi-variant pandemic analytics. Under the specified proportional efficacy-coverage scenario, the reported 59.6% (95% CrI: [56.0%, 64.6%]) value should be interpreted as a conditional quantitative benchmark rather than a full-model herd-immunity guarantee. This scenario-level benchmark supports the importance of sustained vaccination and accelerated case isolation—particularly for the more rapidly spreading Omicron variant.

6. Conclusions

In this study, we developed a deterministic two-strain $SVE_iA_iI_iJHR$ model for the Delta and Omicron waves of COVID-19 in India and combined rigorous stability analysis with uncertainty-quantified, physics-informed parameter estimation. Analytically, we established the positivity and boundedness of solutions, derived closed-form variant-specific reproduction numbers $\mathcal{R}_{0i}$, proved the local and conditional global stability of the disease-free equilibrium, and established the existence and local stability of the single-variant endemic equilibria. Computationally, a three-stage lsqnonlin-PINN-LSTM pipeline estimated the six variant-specific parameters, which were shown to be practically identifiable and structurally identifiable within each single-variant sub-model, and a Bayesian PINN yielded $\mathcal{R}_{01} = 1.87$ (95% CrI [1.759, 2.014]) and $\mathcal{R}_{02} = 1.64$ (95% CrI [1.558, 1.732]).

The hybrid PINN-LSTM model reduced the 30-day forecast RMSE by approximately 23% relative to the PINN-only model and 40% relative to lsqnonlin; these gains were statistically significant and diminished under injected noise, which is consistent with the LSTM learning structured residual signal rather than noise. The sensitivity analyses identified vaccine efficacy ξ_i as the most influential controllable parameter and the asymptomatic fraction η_i as the strongest amplifying parameter for both variants. Under the proportional efficacy—coverage scenario, Delta is the more restrictive variant, with a conditional coverage benchmark of 59.6% (interval [56.0%, 64.6%]); this benchmark is not a full-model herd-immunity threshold, and it widens substantially once uncertainty in α and η_i is included. Omicron, in contrast, produces the larger peak burden, so sustained vaccination combined with faster case isolation remains essential for controlling both variants.

Future work will address a rigorous existence proof for the coexistence equilibrium, a joint structural-identifiability analysis of the fully coupled system, age- and region-stratified extensions of the model, and application of the framework to later sub-lineages such as BA.2, BA.4/5, and XBB.

Author contributions

Mallela Ankamma Rao: Conceptualization, methodology, formal analysis, conception of the two-strain model structure, derivation of the stability results of Section 3, and supervision of the numerical threshold-branch analysis and the overall study; Nazek Alessa: Conceptualization, data curation, validation, funding acquisition, writing—original draft, writing—review and editing, contribution to the model formulation, sourcing and validation of the Indian surveillance data used for calibration, and writing—original draft and editing; Medisetty Padma Devi: Methodology, software, validation, visualization, implementation of the PINN/B-PINN estimation pipeline and the LSTM residual-correction module (Section 4), and production of the associated figures; Carlos Martin-Barreiro: Methodology, validation, formal analysis, supervision, design of the sensitivity-analysis and threshold-branch numerical experiments (Sections 4.5 and 4.6), performed the ablation and identifiability diagnostics, and co-supervised the study; Prasantha Bharathi Dhandapani: Methodology, software, formal analysis, project administration, writing—original draft, writing—review and editing, led the uncertainty-quantification framework (Bayesian PINN, Monte Carlo propagation), and coordinated manuscript writing and revision across all sections. All authors have read and approved the final manuscript.

Availability of data and materials

This manuscript has data associated with it in a data repository. The data used to carry out this study are openly available at <https://github.com/covid19india/data> (archived COVID-19 India data repository, Delta wave) and <https://data.who.int/dashboards/covid19/data> (WHO COVID-19 dashboard, Omicron wave). The Isqnonlin, PINN, Bayesian PINN (HMC/NUTS), and LSTM implementation code used to produce the results and figures in this manuscript (Sections 4.2–4.6 and the Supplementary Material) is available from the corresponding author upon reasonable request.

Use of Generative-AI tools declaration

During the preparation of this manuscript, the authors utilized ChatGPT (OpenAI) to refine the language and structure of the abstract, introduction, and conclusions. Following the use of this tool, the authors thoroughly reviewed and edited the content, and assume full responsibility for the final published work.

Acknowledgments

This work was supported by the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R59), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Worldometer, COVID-19 coronavirus pandemic, 2024. Available from: <https://www.worldometers.info/coronavirus>.
2. H. Y. Yang, G. C. Duan, Analysis on the epidemic factors for COVID-19, *Zhonghua Yu Fang Yi Xue Za Zhi*, **54** (2020), 608–613. <https://doi.org/10.3760/cma.j.cn112150-20200227-00196>
3. World Health Organization, Tracking SARS-CoV-2 variants, 2024. Available from: <https://www.who.int/activities/tracking-SARS-CoV-2-variants>.
4. P. Behera, A. K. Singh, S. H. Subba, A. Mc, D. P. Sahu, P. D. Chandanshive, et al., Effectiveness of COVID-19 vaccine (Covaxin) against breakthrough SARS-CoV-2 infection in India, *Hum. Vaccin. Immunother.*, **18** (2022), 2034456. <https://doi.org/10.1080/21645515.2022.2034456>
5. R. Ella, S. Reddy, W. Blackwelder, V. Potdar, P. Yadav, V. Sarangi, et al., Efficacy, safety, and lot-to-lot immunogenicity of an inactivated SARS-CoV-2 vaccine (BBV152): interim results of a randomised, double-blind, controlled, phase 3 trial, *Lancet*, **398** (2021), 2173–2184. [https://doi.org/10.1016/S0140-6736\(21\)02000-6](https://doi.org/10.1016/S0140-6736(21)02000-6)

6. N. Wu, K. Joyal-Desmarais, P. A. B. Ribeiro, A. M. Vieira, J. Stojanovic, C. Sanuade, et al., Long-term effectiveness of COVID-19 vaccines against infections, hospitalizations, and mortality in adults: findings from a rapid living systematic evidence synthesis and meta-analysis up to December, 2022, *Lancet Respir. Med.*, **11** (2023), 439–452. [https://doi.org/10.1016/s2213-2600\(23\)00015-2](https://doi.org/10.1016/s2213-2600(23)00015-2)
7. M. Massard, R. Eftimie, A. Perasso, B. Saussereau, A multi-strain epidemic model for COVID-19 with infected and asymptomatic cases: application to French data, *J. Theor. Biol.*, **545** (2022), 111117. <https://doi.org/10.1016/j.jtbi.2022.111117>
8. U. Avila-Ponce de León, E. Avila-Vales, K. Huang, Modeling COVID-19 dynamic using a two-strain model with vaccination, *Chaos Solitons Fract.*, **157** (2022), 111927. <https://doi.org/10.1016/j.chaos.2022.111927>
9. S. Y. Tchoumi, H. Rwezaura, J. M. Tchuenche, Dynamic of a two-strain COVID-19 model with vaccination, *Results Phys.*, **39** (2022), 105777. <https://doi.org/10.1016/j.rinp.2022.105777>
10. O. M. Otunuga, Analysis of multi-strain infection of vaccinated and recovered population through epidemic model: application to COVID-19, *PLoS ONE*, **17** (2022), e0271446. <https://doi.org/10.1371/journal.pone.0271446>
11. S. Saha, A. K. Saha, Modeling the dynamics of COVID-19 in the presence of Delta and Omicron variants with vaccination and non-pharmaceutical interventions, *Heliyon*, **9** (2023), e17900. <https://doi.org/10.1016/j.heliyon.2023.e17900>
12. V. Ambalarajan, A. R. Mallela, P. B. Dhandapani, V. Sivakumar, V. Leiva, C. Castro, Multi-strain COVID-19 dynamics with vaccination strategies: mathematical modeling and case study, *Alex. Eng. J.*, **119** (2025), 665–684. <https://doi.org/10.1016/j.aej.2025.01.105>
13. A. O. Yunus, M. O. Olayiwola, A mathematical model for assessing vaccination's efficacy as a preventative strategy against newly emerging COVID-19 variants, *Vacunas*, **26** (2025), 500392. <https://doi.org/10.1016/j.vacun.2025.500392>
14. A. Venkatesh, M. Ankamma Rao, D. K. K. Vamsi, A comprehensive study of optimal control model simulation for COVID-19 infection with respect to multiple variants, *Commun. Math. Biol. Neurosci.*, **2023** (2023), 75. <https://doi.org/10.28919/cmbn/8031>
15. S. R. Bandekar, M. Ghosh, Mathematical modeling of COVID-19 in India and Nepal with optimal control and sensitivity analysis, *Eur. Phys. J. Plus*, **136** (2021), 1058. <https://doi.org/10.1140/epjp/s13360-021-02046-y>
16. V. Ambalarajan, A. R. Mallela, V. Sivakumar, P. B. Dhandapani, V. Leiva, C. Martin-Barreiro, et al., A six-compartment model for COVID-19 with transmission dynamics and public health strategies, *Sci. Rep.*, **14** (2024), 22226. <https://doi.org/10.1038/s41598-024-72487-9>
17. S. K. Ghosh, S. Ghosh, A mathematical model for COVID-19 considering waning immunity, vaccination and control measures, *Sci. Rep.*, **13** (2023), 3610. <https://doi.org/10.1038/s41598-023-30800-y>
18. M. Ankamma Rao, E. K. Jaradat, M. Padma Devi, P. B. Dhandapani, R. M. Nalule, M. Al-Hmoud, Modeling COVID-19 pneumonia and COVID-associated pulmonary aspergillosis: sensitivity analysis and optimal control, *BMC Infect. Dis.*, **25** (2025), 1192. <https://doi.org/10.1186/s12879-025-11606-x>

19. M. A. Rao, E. K. Jaradat, M. P. Devi, P. B. Dhandapani, R. M. Nalule, M. Al-Hmoud, Hybrid modeling and forecasting of COVID-19: integrating SEAIQHRD and GPR for improved predictions, *BMC Infect. Dis.*, **26** (2026), 210. <https://doi.org/10.1186/s12879-025-12494-x>
20. M. A. Khan, A. Atangana, Mathematical modeling and analysis of COVID-19: a study of new variant Omicron, *Phys. A*, **599** (2022), 127452. <https://doi.org/10.1016/j.physa.2022.127452>
21. C. N. Ngonghala, H. B. Taboe, S. Safdar, A. B. Gumel, Unraveling the dynamics of the Omicron and Delta variants of the 2019 coronavirus in the presence of vaccination, mask usage, and antiviral treatment, *Appl. Math. Model.*, **114** (2023), 447–465. <https://doi.org/10.1016/j.apm.2022.09.017>
22. S. S. Negi, N. Sharma, H. M. Baskonus, Dual-strain dynamics of COVID-19 variants in India: modeling, analysis, and implications for pandemic control, *Gene*, **926** (2024), 148586. <https://doi.org/10.1016/j.gene.2024.148586>
23. P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6)
24. O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, *J. R. Soc. Interface*, **7** (2010), 873–885. <https://doi.org/10.1098/rsif.2009.0386>
25. C. Castillo-Chavez, B. Song, Dynamical models of tuberculosis and their applications, *Math. Biosci. Eng.*, **1** (2004), 361–404. <https://doi.org/10.3934/mbe.2004.1.361>
26. M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.*, **378** (2019), 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
27. A. Yazdani, L. Lu, M. Raissi, G. E. Karniadakis, Systems biology informed deep learning for inferring parameters and hidden dynamics, *PLoS Comput. Biol.*, **16** (2020), e1007575. <https://doi.org/10.1371/journal.pcbi.1007575>
28. L. Yang, X. Meng, G. E. Karniadakis, B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data, *J. Comput. Phys.*, **425** (2021), 109913. <https://doi.org/10.1016/j.jcp.2020.109913>
29. K. Linka, A. Schafer, X. Meng, Z. Zou, G. E. Karniadakis, E. Kuhl, Bayesian physics-informed neural networks for real-world nonlinear dynamical systems, *Comput. Methods Appl. Mech. Eng.*, **402** (2022), 115346. <https://doi.org/10.1016/j.cma.2022.115346>
30. M. D. Hoffman, A. Gelman, The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo, *J. Mach. Learn. Res.*, **15** (2014), 1593–1623.
31. S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.*, **9** (1997), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
32. C. Millevoi, D. Pasetto, M. Ferronato, A physics-informed neural network approach for compartmental epidemiological models, *PLoS Comput. Biol.*, **20** (2024), e1012387. <https://doi.org/10.1371/journal.pcbi.1012387>

33. H. Aghdaoui, A. A. Raezah, M. Tilioua, Y. Sabbar, Exploring COVID-19 model with general fractional derivatives: novel physics-informed-neural-networks approach for dynamics and order estimation, *J. Math. Comput. Sci.*, **36** (2024), 142–162. <https://doi.org/10.22436/jmcs.036.02.01>
34. R. Dong, C. Goodbrake, H. Harrington, G. Pogudin, Differential elimination for dynamical models via projections with applications to structural identifiability, *SIAM J. Appl. Algebra Geom.*, **7** (2023), 194–235. <https://doi.org/10.1137/22M1469067>
35. G. Chowell, S. Dahal, Y. R. Liyanage, A. Tariq, N. Tuncer, Structural identifiability analysis of epidemic models based on differential equations: a tutorial-based primer, *J. Math. Biol.*, **87** (2023), 79. <https://doi.org/10.1007/s00285-023-02007-2>
36. G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, et al., Modeling the COVID-19 epidemic and implementation of population-wide interventions in Italy, *Nat. Med.*, **26** (2020), 855–860. <https://doi.org/10.1038/s41591-020-0883-7>
37. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, et al., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2), *Science*, **368** (2020), 489–493. <https://doi.org/10.1126/science.abb3221>
38. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, et al., The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application, *Ann. Intern. Med.*, **172** (2020), 577–582. <https://doi.org/10.7326/M20-0504>
39. A. Venkatesh, M. A. Rao, M. P. Raj, K. A. Kumar, D. K. K. Vamsi, Mathematical modeling of COVID-19 dynamics using SVEAIQHR model, *Comput. Math. Biophys.*, **12** (2024), 20230112. <https://doi.org/10.1515/cmb-2023-0112>
40. N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, et al., Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, 2020, 1–20. <https://doi.org/10.25561/77482>
41. A. Venkatesh, M. A. Rao, Mathematical model for COVID-19 pandemic with implementation of intervention strategies and cost-effectiveness analysis, *Results Control Optim.*, **14** (2024), 100345. <https://doi.org/10.1016/j.rico.2023.100345>
42. R. S. Khedar, R. Gupta, K. Mittal, A. Mathur, J. B. Gupta, K. K. Sharma, et al., Greater COVID-19 severity and mortality in hospitalized patients in the Delta-phase compared to the others: a single center prospective registry, *J. Assoc. Physicians India*, **17** (2023), 18–23. <https://doi.org/10.59556/japi.71.0293>
43. S. Khajanchi, K. Sarkar, S. Banerjee, Modeling the dynamics of COVID-19 pandemic with implementation of intervention strategies, *Eur. Phys. J. Plus*, **137** (2022), 129. <https://doi.org/10.1140/epjp/s13360-022-02347-w>
44. P. Das, S. S. Nadim, S. Das, Dynamics of COVID-19 transmission with comorbidity: a data driven modeling based approach, *Nonlinear Dyn.*, **106** (2021), 1197–1211. <https://doi.org/10.1007/s11071-021-06324-3>
45. S. Khajanchi, K. Sarkar, J. Mondal, K. S. Nisar, S. F. Abdelwahab, Mathematical modeling of the COVID-19 pandemic with intervention strategies, *Results Phys.*, **25** (2021), 104285. <https://doi.org/10.1016/j.rinp.2021.104285>

46. Y. Liu, J. Rocklov, The reproduction number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus, *J. Travel Med.*, **28** (2021), 124. <https://doi.org/10.1093/jtm/taab124>
47. E. Mathieu, H. Ritchie, L. Rodés-Guirao, C. Appel, D. Gavrilov, C. Giattino, et al., Coronavirus (COVID-19) vaccinations: India, Our World in Data, 2020. Available from: <https://ourworldindata.org/covid-vaccinations?country=IND>.
48. COVID-19 India Org Data Operations Group, COVID-19 India tracker: state-wise daily confirmed cases, recoveries, and deaths, 2021. Available from: <https://github.com/covid19india/data>.
49. Y. R. Liyanage, G. Chowell, G. Pogudin, N. Tuncer, Structural and practical identifiability of phenomenological growth models for epidemic forecasting, *Viruses*, **17** (2025), 496. <https://doi.org/10.3390/v17040496>
50. A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, et al., Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, *Bioinformatics*, **25** (2009), 1923–1929. <https://doi.org/10.1093/bioinformatics/btp358>
51. S. Wang, Y. Teng, P. Perdikaris, Understanding and mitigating gradient flow pathologies in physics-informed neural networks, *SIAM J. Sci. Comput.*, **43** (2021), A3055–A3081. <https://doi.org/10.1137/20M1318043>
52. World Health Organization, WHO COVID-19 dashboard: COVID-19 cases, 2023. Available from: <https://data.who.int/dashboards/covid19/data>.
53. Ministry of Health and Family Welfare, Government of India, COVID-19 statistics and vaccination updates, 2024. Available from: <https://www.mohfw.gov.in/>.
54. S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, *J. Theor. Biol.*, **254** (2008), 178–196. <https://doi.org/10.1016/j.jtbi.2008.04.011>
55. G. Zeng, Q. Wu, H. Pan, M. Li, J. Yang, L. Wang, et al., Immunogenicity and safety of a third dose of CoronaVac, and immune persistence of a two-dose schedule, in healthy adults: interim results from two single-centre, double-blind, randomised, placebo-controlled phase 2 clinical trials, *Lancet Infect. Dis.*, **22** (2022), 483–495. [https://doi.org/10.1016/S1473-3099\(21\)00681-2](https://doi.org/10.1016/S1473-3099(21)00681-2)

Supplementary

This section provides the diagnostic and robustness material referenced throughout Sections 4.2 and 4.4 of the main text. It contains six supplementary figures (Figures S1–S6: profile likelihoods, multi-start convergence, lsqnonlin-vs.-PINN robustness under reduced data and injected noise for both variants, the single-equation ODE-residual plausibility check, and the walk-forward cross-validation bar charts) and six supplementary tables (Tables S1–S6: the capacity-matched-ablation architecture specification, its walk-forward RMSE and recovered-parameter comparison, the associated Diebold–Mariano test, and the pairwise Diebold–Mariano and AIC/BIC tables for the hybrid/PINN/lsqnonlin comparison). These items support the identifiability, ablation, and forecast-robustness claims of the

main text but are not required to follow the paper’s central argument. All in-text references to “Supplementary Figure S_n ” and “Supplementary Table S_n ” point to the corresponding items below.

Supplementary figures

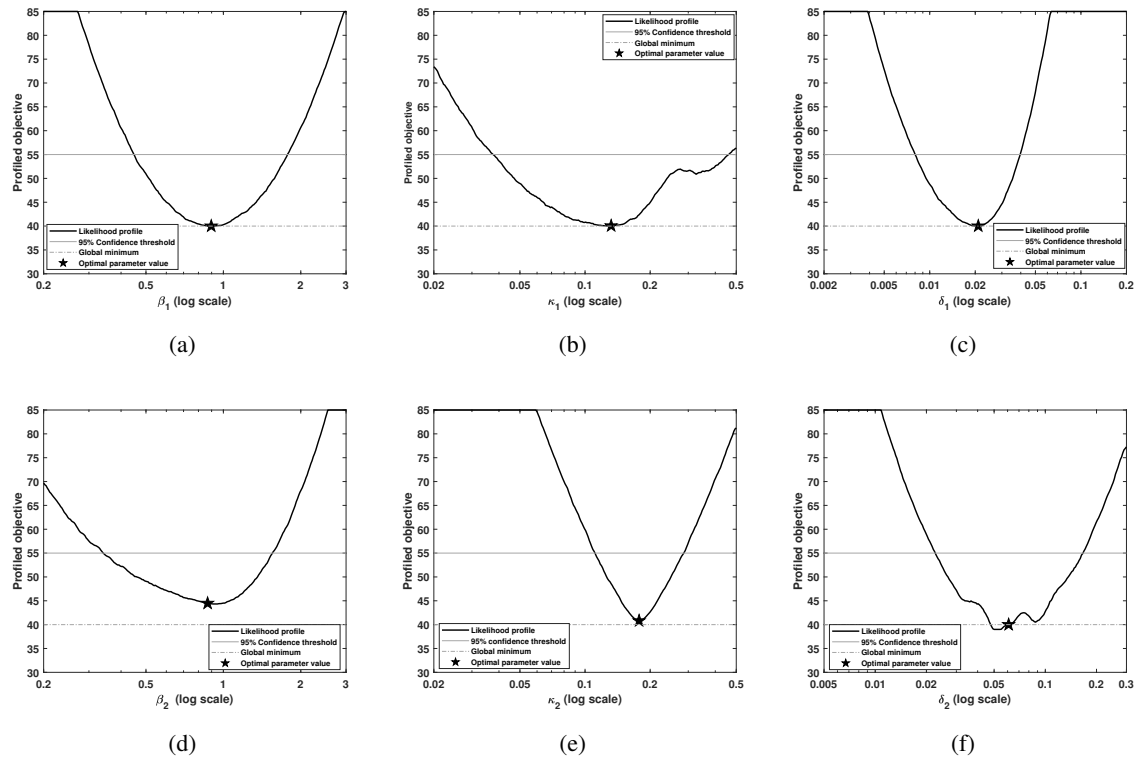


Figure S1. Profile likelihoods for the six variant-specific parameters $(\beta_i, \kappa_i, \delta_i)_{i=1,2}$, obtained by fixing each parameter at a grid of values spanning its plausible range, re-optimizing the remaining parameters and network weights from that fixed value, and recording the resulting profiled objective. Solid black curve: likelihood profile; solid gray line: 95% confidence threshold ($\Delta\chi^2 = 3.84$, one degree of freedom, via Wilks’ theorem); gray dash-dot line: global minimum (best-fit objective value); black star: converged optimum. Top row: Delta sub-system, (a) β_1 , (b) κ_1 , (c) δ_1 . Bottom row: Omicron sub-system, (d) β_2 , (e) κ_2 , (f) δ_2 .

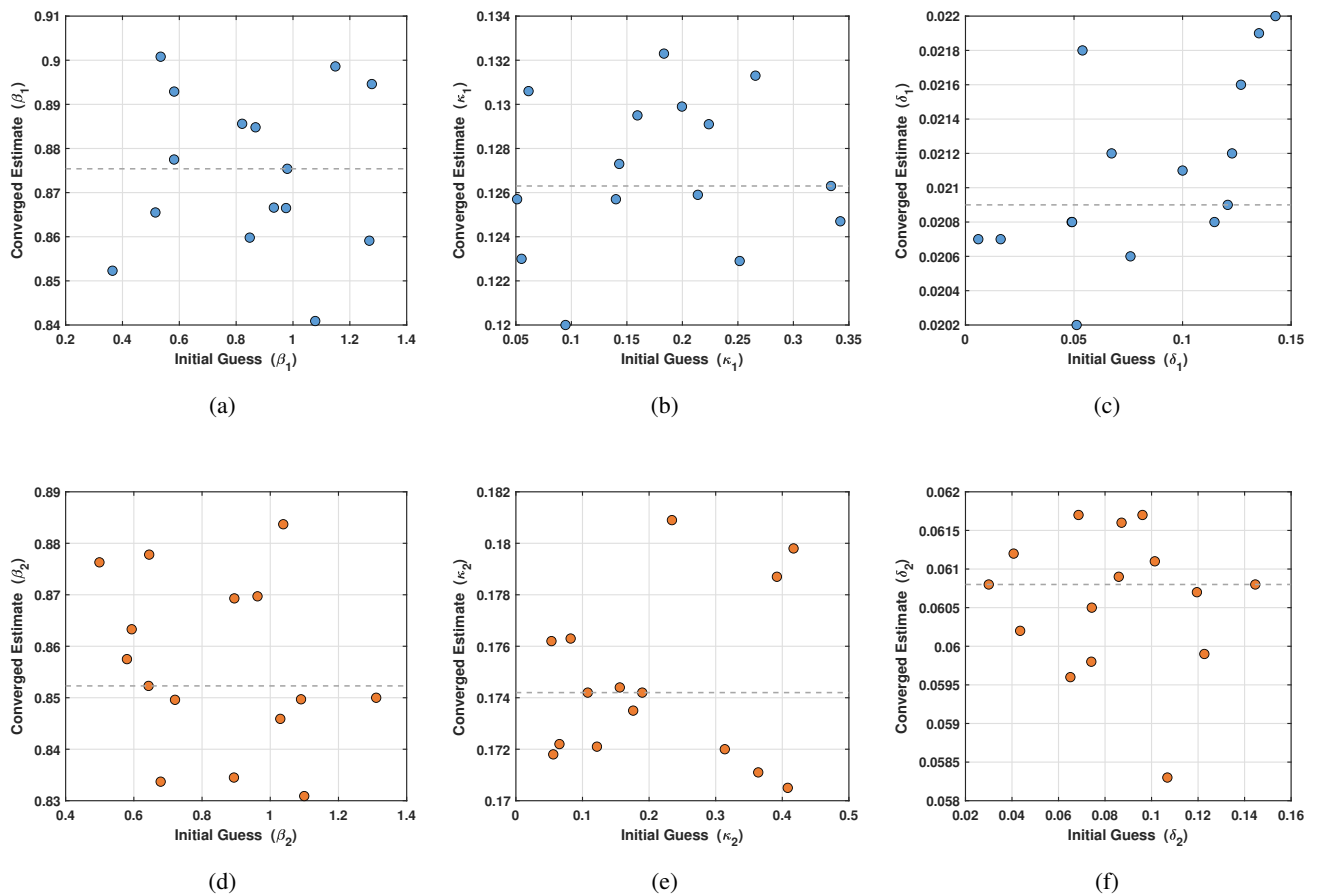


Figure S2. Multi-start convergence of the six PINN-estimated parameters from $n = 15$ independent random initializations. Top row: Delta variant, (a) β_1 , (b) κ_1 , (c) δ_1 . Bottom row: Omicron variant, (d) β_2 , (e) κ_2 , (f) δ_2 . Each panel plots the converged estimate against the corresponding random initial guess; the dashed line marks the median converged value.

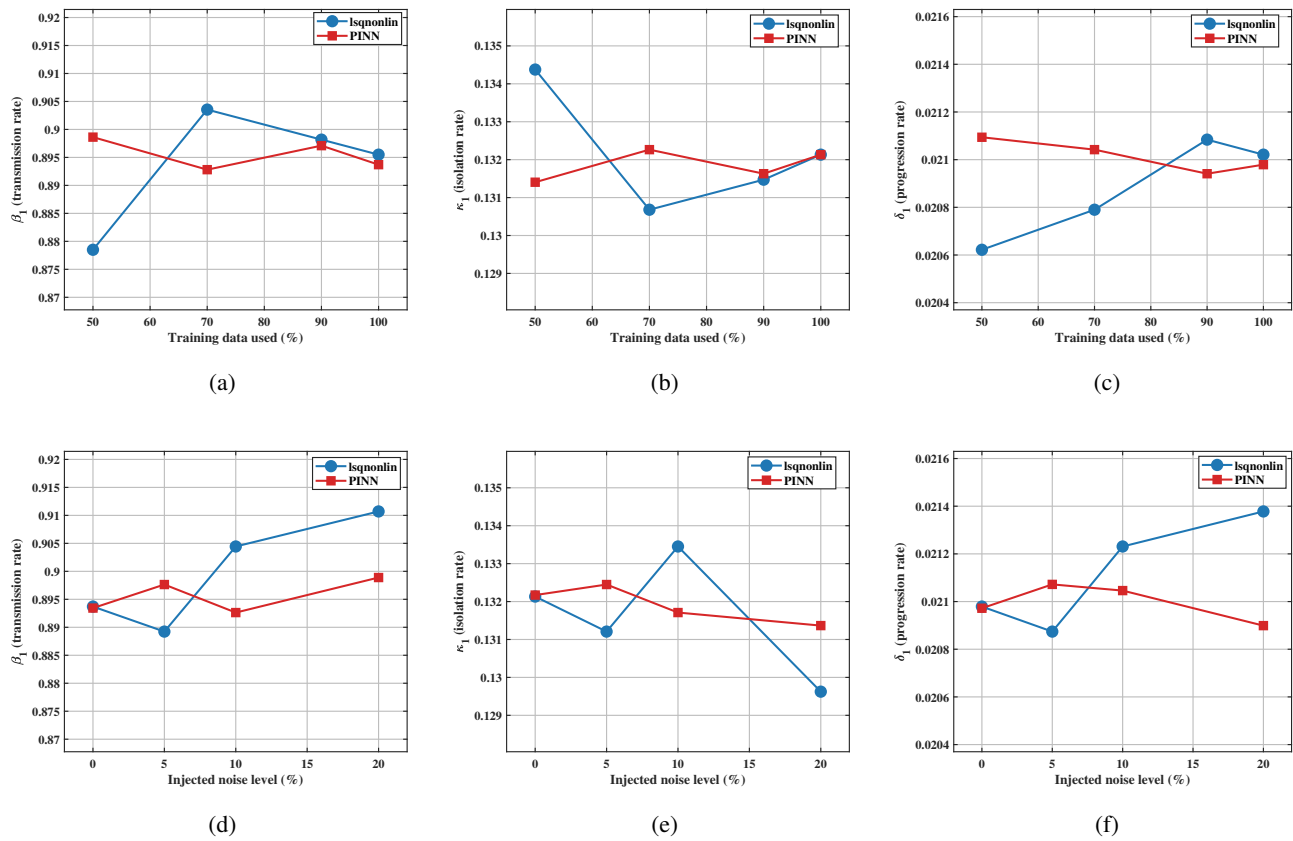


Figure S3. Delta: lsqnonlin vs. PINN estimates of $(\beta_1, \kappa_1, \delta_1)$ under reduced training-data volume (top row) and injected measurement noise (bottom row), applied to the reduced single-strain surrogate model. Top row: (a) β_1 , (b) κ_1 , (c) δ_1 versus training-data fraction (50–100%). Bottom row: (d) β_1 , (e) κ_1 , (f) δ_1 versus injected noise level (0–20%).

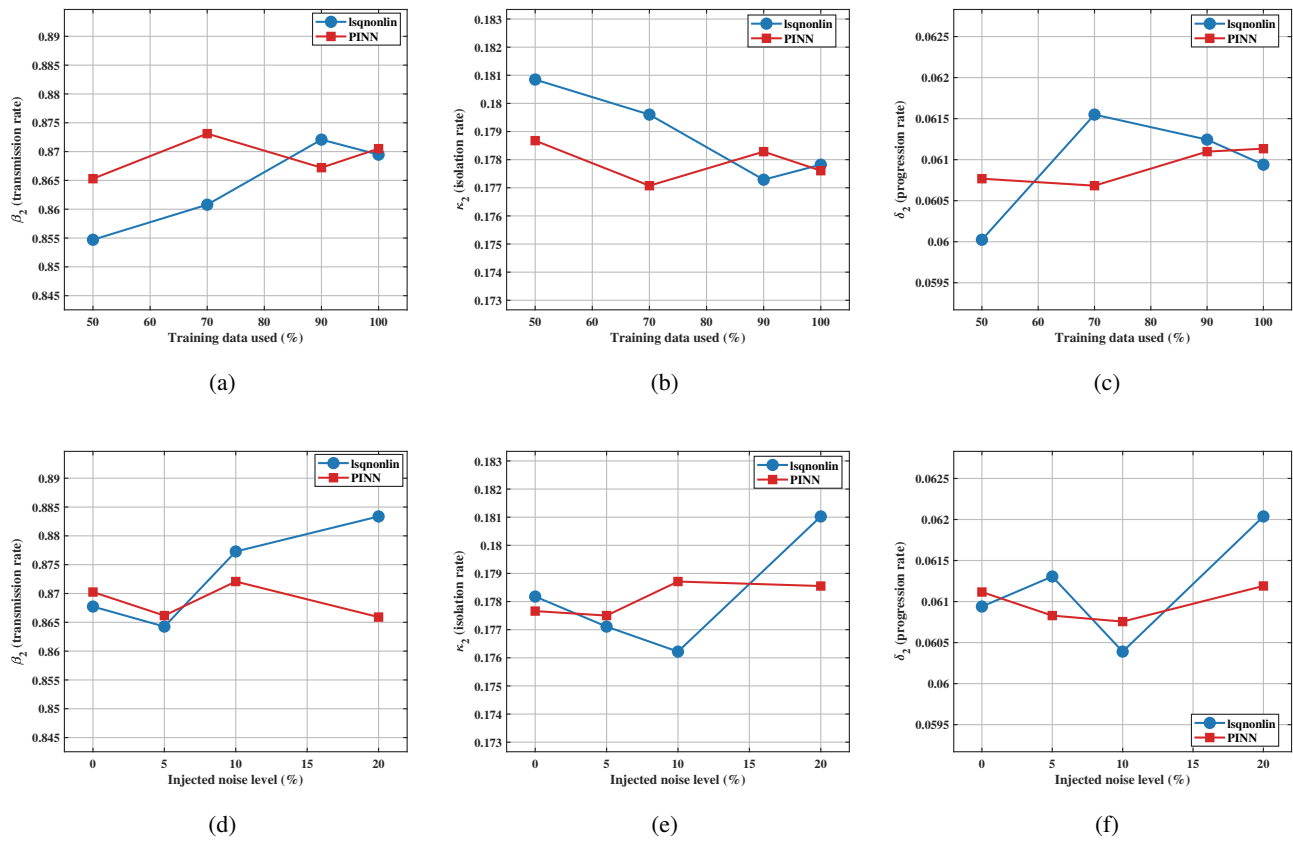


Figure S4. Omicron: lsqnonlin vs. PINN estimates of $(\beta_2, \kappa_2, \delta_2)$ under reduced training-data volume (top row) and injected measurement noise (bottom row), applied to the reduced single-strain surrogate model. Top row: (a) β_2 , (b) κ_2 , (c) δ_2 versus training-data fraction (50–100%). Bottom row: (d) β_2 , (e) κ_2 , (f) δ_2 versus injected noise level (0–20%).

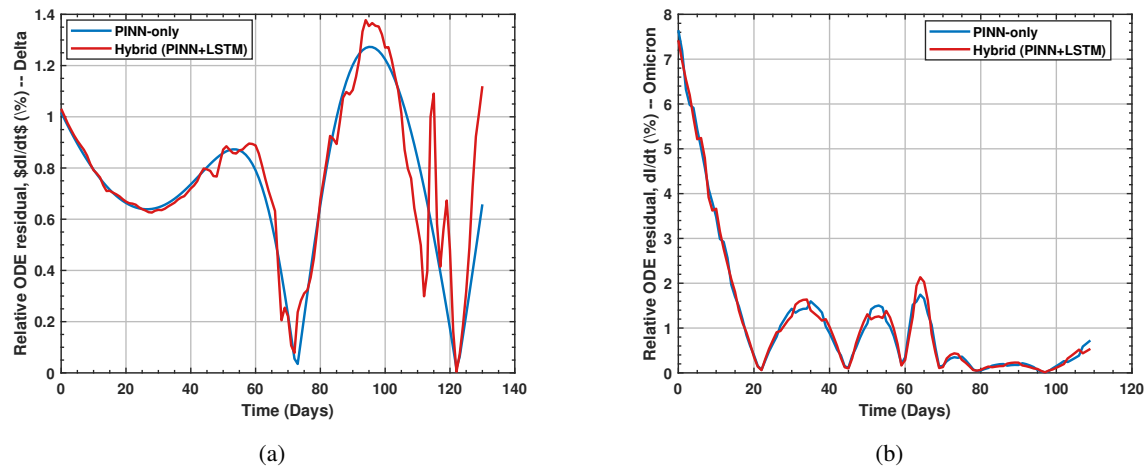


Figure S5. Single-equation plausibility check for the corrected observable using the dominant variant's symptomatic-compartment Eq (2.14). The PINN-only trajectory (blue) is evaluated by automatic differentiation, while the Stage-3-corrected case-count series (red) is evaluated by finite differences over the calibration window. This is not a full-state ODE consistency check, since the LSTM corrects only the aggregate observable $\hat{C}(t)$. (a) Delta (130-day window). (b) Omicron (110-day window). The residual remains below 1.4% for Delta and 8% for Omicron, confirming that the correction is small relative to the improvement in forecast accuracy (Section 4.4).

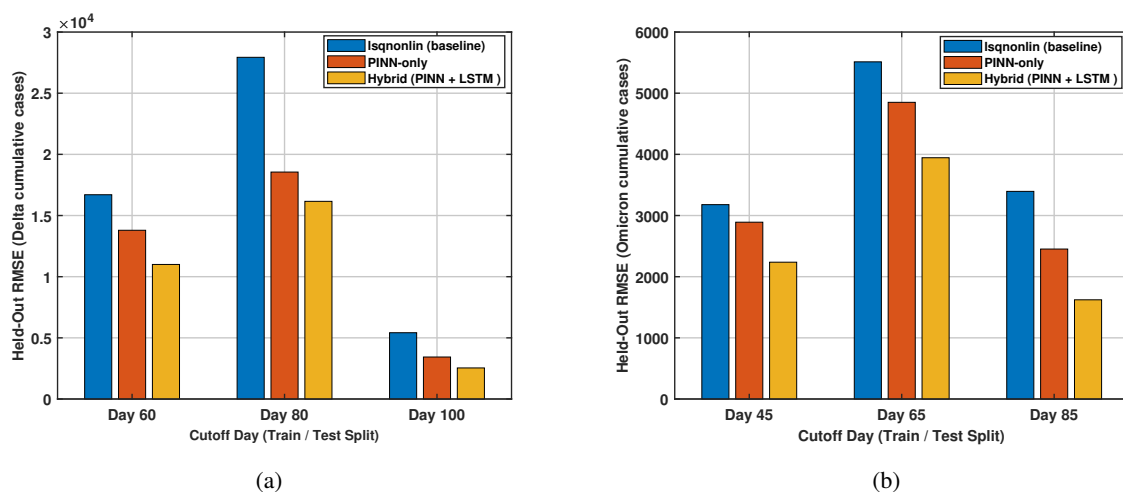


Figure S6. Walk-forward cross-validation of held-out 15-day forecast RMSE across three independent train/test cutoffs. (a) Delta ($n = 3$ folds, cutoffs at days 60, 80, 100 of 130). (b) Omicron ($n = 3$ folds, cutoffs at days 45, 65, 85 of 110).

Supplementary tables

Table S1. Architecture specification for the capacity-matched PINN ablation (Section 4.4 of the main text).

Specification	PINN	Data-only NN (ablation)
Architecture	Fully connected, 6 hidden layers $\times$ 64 neurons	Fully connected, 6 hidden layers $\times$ 64 neurons (same)
Activation function	tanh	tanh (same)
Output layer	Linear	Linear (same)
Total trainable parameters	$\approx 21,000$	$\approx 21,000$ (same)
Loss function	$\mathcal{L} = \mathcal{L}_{\text{data}} + w_r \mathcal{L}_{\text{ODE}} + w_{\text{IC}} \mathcal{L}_{\text{IC}}, w_r = 10^{-2}$	$\mathcal{L} = \mathcal{L}_{\text{data}} + w_{\text{IC}} \mathcal{L}_{\text{IC}}, w_r = 0$
Optimizer	Adam (10^{-3}) $\rightarrow$ L-BFGS	Adam (10^{-3}) $\rightarrow$ L-BFGS (same)
Training epochs	50,000	50,000 (same)
Nominal parameter count (k)	6	6

Table S2. Walk-forward held-out RMSE: full PINN versus the capacity-matched data-only NN ($w_r = 0$).

Variant	Cutoff (day)	PINN RMSE	Data-only NN RMSE	Δ (%)
Delta	60	13,795	16,936	PINN 18.5% better
Delta	80	18,554	22,223	PINN 16.5% better
Delta	100	3,433	4,200	PINN 18.3% better
Omicron	45	2,891	3,244	PINN 10.9% better
Omicron	65	4,851	6,008	PINN 19.3% better
Omicron	85	2,452	2,968	PINN 17.4% better

Table S3. Recovered parameters: full PINN versus capacity-matched data-only NN, with physical plausibility assessment against the literature ranges of Table 1 of the main text.

Parameter	PINN	Data-only NN	Physical plausibility
β_1	0.895	1.42	Data-only: substantially higher than reported literature values
κ_1	0.132	0.038	Data-only: implausibly low ($\approx$ 26-day isolation delay)
δ_1	0.021	-0.012	Data-only: negative, physically meaningless
β_2	0.869	1.18	Data-only: overestimates transmission rate
κ_2	0.178	0.325	Data-only: unrealistically high isolation rate
δ_2	0.061	0.145	Data-only: exceeds commonly reported values

Table S4. Diebold-Mariano test comparing the full PINN and the capacity-matched Data-only NN.

Variant	Comparison	DM statistic	p -value	Winner
Delta	PINN vs. Data-only NN	-2.54	0.0064	PINN (significant)
Omicron	PINN vs. Data-only NN	-3.02	0.0020	PINN (significant)

Table S5. Diebold-Mariano test for the pairwise statistical significance of forecast-accuracy differences between the hybrid PINN-LSTM, PINN-only, and lsqnonlin models, pooled across the walk-forward folds of Figure S6 for each variant.

Variant	Comparison	DM statistic	p -value	Best model
Delta	Hybrid vs. PINN	−2.89	0.0058	Hybrid
	Hybrid vs. lsqnonlin	−3.46	0.0010	Hybrid
	PINN vs. lsqnonlin	−2.13	0.0398	PINN
Omicron	Hybrid vs. PINN	−6.10	< 0.0001	Hybrid
	Hybrid vs. lsqnonlin	−6.90	< 0.0001	Hybrid
	PINN vs. lsqnonlin	−2.28	0.0265	PINN

Table S6. Held-out RMSE together with the Akaike and Bayesian information criteria (AIC, BIC) for the lsqnonlin, PINN-only, and hybrid PINN-LSTM models at each walk-forward cutoff of Figure S6, where k denotes the number of fitted parameters used in the AIC/BIC penalty for each model.

Variant	Cutoff (day)	lsqnonlin ($k = 3$)	PINN ($k = 6$)	Hybrid ($k = 10$)
		RMSE / AIC / BIC	RMSE / AIC / BIC	RMSE / AIC / BIC
Delta	60	16700 / 1172.8 / 1179.1	13795 / 1155.9 / 1168.4	11000 / 1136.7 / 1157.6
Delta	80	27935 / 1644.0 / 1651.2	18554 / 1584.6 / 1598.8	16160 / 1570.4 / 1594.3
Delta	100	5419 / 1725.5 / 1733.4	3433 / 1640.2 / 1655.9	2540 / 1588.0 / 1614.0
Omicron	45	3178 / 731.8 / 737.2	2891 / 729.2 / 740.1	2237 / 714.2 / 732.2
Omicron	65	5513 / 1125.9 / 1132.5	4851 / 1115.3 / 1128.4	3945 / 1096.4 / 1118.2
Omicron	85	3395 / 1388.1 / 1395.4	2452 / 1338.8 / 1353.5	1622 / 1276.5 / 1300.9



AIMS Press

©2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)