



Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational
Mathematics (IMACM)

Preprint BUW-IMACM 26/07

Achraf Zinihi and Matthias Ehrhardt

Constraint-Aware Physics-Informed Neural Networks for SEIR Reaction-Diffusion Epidemic Models with Vital Dynamics

May 19, 2026

<http://www.imacm.uni-wuppertal.de>

Constraint-Aware Physics-Informed Neural Networks for SEIR Reaction-Diffusion Epidemic Models with Vital Dynamics

Achraf Zinihi^{a,b}, Matthias Ehrhardt^{a,*}

^a*University of Wuppertal, Applied and Computational Mathematics,
Gaußstrasse 20, 42119 Wuppertal, Germany*

^b*Department of Mathematics, AMNEA Group, Faculty of Sciences and Techniques,
Moulay Ismail University of Meknes, Errachidia 52000, Morocco*

Abstract

Reaction-diffusion epidemic models with vital dynamics are an important framework for describing the spatial and temporal spread of infectious diseases. In this work, we present a constraint-aware, physics-informed neural network (PINN) approach to an SEIR reaction-diffusion system with homogeneous Neumann boundary conditions. Due to the scarcity of spatial epidemiological datasets, we generate synthetic benchmark data using structure-preserving implicit-explicit nonstandard finite difference (NSFD) schemes that ensure positivity, boundedness, and numerical stability.

The PINN framework integrates PDE residuals, observational data, boundary conditions, and epidemiological constraints within a unified optimization procedure. Specifically, the loss function incorporates the non-negativity of compartment populations and the admissibility of epidemiological parameters. We apply the method to forward simulation and inverse parameter estimation in one- and two-dimensional settings. Numerical experiments demonstrate the framework's ability to accurately reconstruct spatiotemporal epidemic dynamics and reliably identify parameters, even when data is sparse or noisy. These results underscore the potential of constraint-aware PINNs as a robust, data-driven methodology for spatial epidemic modeling.

Keywords: Epidemic modeling, Spatiotemporal dynamics, NSFD scheme, Physics-informed Neural networks, Parameter estimation, Fourier feature mapping.

2020 Mathematics Subject Classification: 92D30, 35K57, 68T07, 92-10.

1. Introduction

The spatiotemporal spread of infectious diseases is a primary concern in both mathematical epidemiology and public health policy. This concern has motivated the development of models capable of predicting, explaining, and ultimately controlling outbreaks. Classical compartmental models, such as the SIR, SIQR, and SEIR systems, have provided essential insights into epidemic dynamics since the pioneering work of Kermack and McKendrick [1, 2]. These models are traditionally formulated as systems of ordinary differential equations (ODEs) and capture average population-level dynamics, but they neglect spatial heterogeneity and human mobility. Both of these factors are critical drivers of epidemic spread. To address these limitations, reaction-diffusion formulations of epidemic models have been introduced to simulate spatially structured outbreaks, wave-like infection fronts, and localized clustering phenomena [3–6].

*Corresponding author

Email addresses: a.zinihi@edu.umi.ac.ma (Achraf Zinihi), ehrhardt@uni-wuppertal.de (Matthias Ehrhardt)

Despite their conceptual appeal, reaction-diffusion epidemic partial differential equations (PDEs) present substantial challenges in practice. Classical numerical methods, such as finite difference schemes, finite element methods, and spectral solvers [7–9], provide accurate approximations, but they suffer from the curse of dimensionality and become prohibitively expensive in two or more spatial dimensions. Furthermore, inverse problems, such as estimating epidemiological parameters from noisy, sparse, or heterogeneous surveillance data, are notoriously ill-posed [10], often yielding unstable or biologically implausible estimates.

In recent years, physics-informed neural networks (PINNs) have emerged as a promising approach to data-driven partial differential equation (PDE) modeling [11, 12]. PINNs embed governing equations and boundary conditions directly into the loss function of a neural network, enabling mesh-free solution of PDEs and the simultaneous integration of data and mechanistic knowledge. This approach has grown rapidly across disciplines, including fluid mechanics, materials science, and epidemiology [13–15]. In the context of epidemic systems, PINNs have been successfully applied to ODE-based models of diseases such as COVID-19 and other respiratory viruses [16–18], as well as for forecasting hospitalizations [19, 20] and even for fractional-order epidemic dynamics [21, 22]. Recent works have emphasized bi-objective optimization [23], physics-informed graph neural ODEs [24], neural ODE–Bayesian hybrids [25], and variance-aware PINNs [26], all of which point toward the rapidly evolving frontier of physics-informed machine learning in epidemiology.

Nevertheless, the application of PINNs to spatial epidemic PDEs is still in its early stages. Most existing works focus on non-spatial ODEs or simplified 1D PDE settings [27, 28]. Furthermore, the systematic enforcement of epidemiological constraints, such as the non-negativity of populations, the boundedness of parameters, and the conservation of the total population under vital dynamics, has rarely been addressed despite being essential for biologically meaningful results. Without these constraints, PINN approximations may yield unrealistic solutions, such as negative compartment sizes or unbounded growth.

Based on the above discussion, several important gaps remain in the current literature on PINN-based epidemic modeling. Most existing studies are limited to ODE systems or 1D PDE formulations. Meanwhile, 2D reaction–diffusion models, which are better suited for realistic spatial epidemic propagation, have not been widely explored. Additionally, epidemiological constraints, such as the non-negativity of compartment populations, bounded transmission rates, and population balance under vital dynamics, are rarely explicitly incorporated into the learning framework despite their importance for biological realism.

Another major limitation is the lack of high-quality spatiotemporal benchmark datasets for reaction-diffusion epidemic PDEs. This makes systematic validation and reproducibility difficult. Consequently, synthetic data generation using structure-preserving methods, such as NSFD schemes, is essential. Finally, rigorous comparative studies of 1D and 2D epidemic PINN models with respect to predictive accuracy, computational efficiency, and parameter identifiability are lacking.

This paper presents a constraint-aware PINN framework for SEIR reaction-diffusion epidemic models with vital dynamics and homogeneous Neumann boundary conditions. This approach incorporates biologically motivated constraints into the loss function to enforce the positivity and boundedness of compartment populations, the admissibility of epidemiological parameters, and population balance under births and natural deaths. Synthetic 1D and 2D benchmark datasets are generated using structure-preserving implicit-explicit nonstandard finite difference (NSFD) schemes under various noise and sparsity conditions to enable systematic and reproducible evaluation. Furthermore, we perform comprehensive comparisons between 1D and 2D PINN formulations to assess predictive accuracy, robustness, and computational efficiency.

The remainder of this work is structured as follows. Section 2 introduces the proposed SEIR reaction-diffusion model with vital dynamics and establishes its well-posedness. Section 3 details the construction of synthetic datasets using NSFD schemes and their numerical analysis. Section 4 presents the constraint-aware PINN framework, including the network architecture, loss function design, and training strategy. Section 5 reports numerical results, including forward predictions, inverse parameter estimation, and comparisons between 1D and 2D settings. Finally, Section 6 concludes the paper with a summary of the findings.

2. Mathematical Model

We consider an SEIR-type epidemic system with *vital dynamics* and spatial diffusion, defined on a bounded spatial domain $\Omega \subset \mathbb{R}^d$ ($d = 1, 2$) with an outward unit normal vector \vec{n} on the boundary $\partial\Omega$. The temporal domain is $t \in [0, T]$ for some final time $T > 0$. The state variables are the population densities of the susceptible $S(t, x)$, exposed $E(t, x)$, infected $I(t, x)$, and recovered $R(t, x)$ classes at time t and spatial position $x \in \Omega$.

2.1. The PDE System

We propose a four-dimensional epidemiological model to describe the transmission dynamics of an epidemic. The total population is divided into four compartments: susceptible (S), exposed (E), infected (I), and recovered (R). Individuals are recruited into the susceptible class at a constant rate Λ and die at a natural rate μ . Susceptible individuals may become infected upon contact with infected individuals at a rate β , with a fraction p entering the exposed class and the complementary fraction $1 - p$ progressing directly to the infected class. Exposed individuals can either progress to the infected class at rate δ or recover without becoming infectious at rate η . Infected individuals recover at a rate γ . Recovered individuals acquire permanent immunity and typically only die through natural causes. The transitions between compartments are summarized in Figure 1.

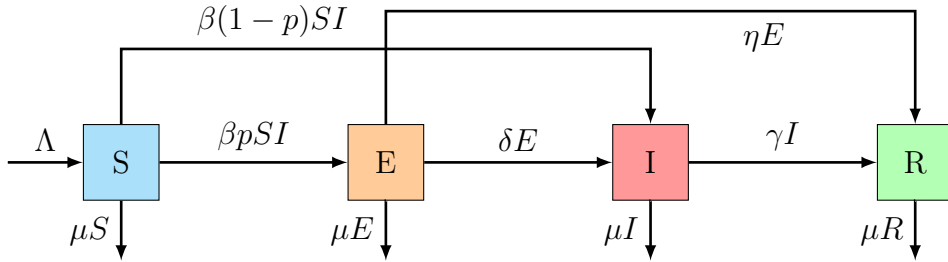


Figure 1: Flow diagram of the SEIR reaction-diffusion model with vital dynamics.

Spatial mobility is incorporated through diffusion terms, with a common diffusion coefficient $\lambda > 0$ for all compartments. This allows the model to capture the heterogeneity in disease spread across a spatial domain. The model parameters are given in Table 1.

Table 1: Epidemic parameters for the proposed SEIR reaction-diffusion model.

Symbol	Description
Λ	Recruitment rate (immigration or birth)
μ	Natural mortality rate
β	Transmission coefficient
p	Fraction of infections entering the exposed class
δ	Progression rate from exposed to infected
η	Recovery rate from the exposed class
γ	Recovery rate from the infected class
λ	Diffusion coefficient (identical for all classes)

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with smooth boundary $\partial\Omega$. The spatiotemporal dynamics of the SEIR system are described by the following system of reaction-diffusion equations

$$\begin{cases} \frac{\partial S(t, x)}{\partial t} = \lambda\Delta S(t, x) + \Lambda - \beta S(t, x)I(t, x) - \mu S(t, x), \\ \frac{\partial E(t, x)}{\partial t} = \lambda\Delta E(t, x) + \beta p S(t, x)I(t, x) - (\delta + \eta + \mu)E(t, x), \\ \frac{\partial I(t, x)}{\partial t} = \lambda\Delta I(t, x) + \beta(1 - p)S(t, x)I(t, x) + \delta E(t, x) - (\gamma + \mu)I(t, x), \\ \frac{\partial R(t, x)}{\partial t} = \lambda\Delta R(t, x) + \eta E(t, x) + \gamma I(t, x) - \mu R(t, x). \end{cases} \quad \text{in } \mathcal{U}_T, \quad (1)$$

Here, $\mathcal{U}_T = [0, T] \times \Omega$ and $\Delta = \nabla^2$ is the Laplace operator modeling spatial diffusion. We impose homogeneous Neumann boundary conditions to describe a closed system with no population flux across the boundary

$$\nabla S \cdot \vec{n} = \nabla E \cdot \vec{n} = \nabla I \cdot \vec{n} = \nabla R \cdot \vec{n} = 0, \quad \text{on } \Sigma_T = [0, T] \times \partial\Omega, \quad (2)$$

where \vec{n} is the outward unit normal to $\partial\Omega$. The initial conditions specify the initial spatial distribution of each compartment

$$S(0, x) = S_0(x), \quad E(0, x) = E_0(x), \quad I(0, x) = I_0(x), \quad R(0, x) = R_0(x), \quad \text{in } \Omega. \quad (3)$$

Unlike standard SEIR models, which assume a constant population size, the inclusion of recruitment and mortality ensures that the total population

$$N(t, x) = S(t, x) + E(t, x) + I(t, x) + R(t, x),$$

is bounded but not constant. We obtain for $N(t, x)$ the temporal ODE

$$\frac{\partial N}{\partial t} - \lambda\Delta N = \Lambda - \mu N, \quad (4)$$

i.e. N remains bounded due to the balance between births and natural deaths, see Section 2.2.

2.2. Well-Posedness Analysis

To ensure the mathematical consistency of the proposed SEIR reaction-diffusion model, we analyze the well-posedness of the system (1)–(3). Our goal is to verify that the system admits a unique global solution that is bounded and nonnegative for all time. To this end, we rewrite the model in abstract form

$$\begin{cases} \partial_t \psi(t) + \mathcal{T}\psi(t) = \mathcal{F}(\psi(t)), \\ \psi(0) = \psi_0, \end{cases}$$

where $\psi(t) = (S, E, I, R)(t, \cdot)$, the linear operator \mathcal{T} is associated with diffusion terms

$$\mathcal{T}: \begin{aligned} \mathcal{D}_{\mathcal{T}} &= \left\{ v \in (H^2(\Omega))^4 \mid \nabla v_i \cdot \vec{n} = 0, 1 \leq i \leq 4 \right\} \subset (L^2(\Omega))^4 \longrightarrow (L^2(\Omega))^4, \\ v &\longrightarrow -\lambda \Delta v, \end{aligned}$$

and \mathcal{F} contains the nonlinear reaction dynamics. First, we establish that all solutions of the system remain uniformly bounded in finite time using properties of analytic C_0 -semigroups generated by $-\Delta$ under Neumann boundary conditions [29, pp. 230-251]. Then, multiplying each equation by the negative part of the corresponding variable, applying integration by parts, and using Cauchy-Schwarz and Gronwall's inequalities proves that nonnegative initial data lead to nonnegative solutions, ensuring biological feasibility. Furthermore, the existence and uniqueness of a global strong solution follow from standard results in functional analysis since the nonlinear function \mathcal{F} is Lipschitz continuous (due to the boundedness of the solution) and the Laplace operator is strongly elliptic. Finally, the solution satisfies the regularity properties

$$\psi_i \in W^{1,2}(0, T; L^2(\Omega)) \cap L^2(0, T; H^2(\Omega)) \cap L^\infty(0, T; H^1(\Omega)),$$

for each component, which guarantees both smoothness and stability of the dynamics.

For further details and rigorous proofs of the well-posedness analysis, see [9, pp. 8-11]. This analysis shows that the proposed system retains the essential biological characteristics of epidemic dynamics. Solutions remain nonnegative when initial data is nonnegative, and the total population, $N(t, x)$, is uniformly bounded by Λ/μ in the long-term regime. Additionally, including the exposed compartment alongside vital dynamics offers a biologically accurate depiction of incubation and demographic turnover. These properties justify using the model as a reliable testing ground for further analytical and numerical investigations, including forward and inverse PINN-based studies in spatial epidemiology.

3. Generation of Synthetic Data via NSFD Schemes

As there are no publicly available real-world epidemiological datasets with the spatial resolution required for reaction-diffusion systems, we generate synthetic data by numerically solving equations (1)–(3). To ensure that the numerical solutions preserve essential epidemiological properties such as non-negativity, boundedness and numerical stability, we use an NSFD scheme.

Classical finite-difference (FD) discretizations, such as the explicit/implicit Euler methods or Runge-Kutta methods, can produce negative or unbounded compartment values when applied to nonlinear epidemic PDEs, see e.g. [30]. In contrast, NSFD schemes [31, 32] are explicitly designed to preserve the qualitative features of the underlying continuous system, including the positivity of solutions, the invariance of feasible regions, and the boundedness of the total population N . These properties are essential to epidemiological models because negative population densities or unbounded growth are biologically meaningless.

3.1. NSFD Schemes

Using the methodology outlined in a previous work of the authors [33], we develop NSFD schemes to numerically approximate the spatiotemporal SEIR model (1). Let $t_n = nk$ denote the discrete time levels with a time step size of $k > 0$. The time derivative is approximated by

$$\left. \frac{\partial u(t, \cdot)}{\partial t} \right|_{t=t_n} \approx \frac{u^{n+1} - u^n}{\phi(k)},$$

where $\phi(k) > 0$ is a *nonstandard denominator function* satisfying $\phi(k) \rightarrow k$ as $k \rightarrow 0$. First, we describe the one-dimensional NSFD discretization, then we extend the approach to two-dimensional spatial domains.

3.1.1. One-Dimensional NSFD Scheme

Discretize the spatial domain $\Omega = [0, L]$ into M grid points with spacing $h = L/M$. For $j = 1, \dots, M$ and $n \geq 0$, denote $S_j^n \approx S(t_n, x_j)$, and similarly for the other compartments E , I , and R . The term $\Delta S(t_n, x_j)$ is approximated using the following nonstandard ('skew') discretization of the Laplacian

$$\Delta_h^2 S_j^n = \frac{S_{j+1}^n - 2S_j^{n+1} + S_{j-1}^n}{h^2}.$$

to guarantee the non-negativity of the solution. Consequently, the NSFD discretization of the system (1) in 1D reads

$$\begin{cases} \frac{S_j^{n+1} - S_j^n}{\phi(k)} = \lambda \Delta_h^2 S_j^n + \Lambda - (\beta I_j^n + \mu) S_j^{n+1}, \\ \frac{E_j^{n+1} - E_j^n}{\phi(k)} = \lambda \Delta_h^2 E_j^n + \beta p S_j^{n+1} I_j^n - (\delta + \eta + \mu) E_j^{n+1}, \\ \frac{I_j^{n+1} - I_j^n}{\phi(k)} = \lambda \Delta_h^2 I_j^n + \beta(1-p) S_j^{n+1} I_j^n + \delta E_j^{n+1} - (\gamma + \mu) I_j^{n+1}, \\ \frac{R_j^{n+1} - R_j^n}{\phi(k)} = \lambda \Delta_h^2 R_j^n + \eta E_j^{n+1} + \gamma I_j^{n+1} - \mu R_j^{n+1}. \end{cases} \quad (5)$$

The NSFD scheme (5) is reformulated in a sequentially explicit form

$$\begin{cases} S_j^{n+1} = \frac{S_j^n + \lambda r(k)(S_{j-1}^n + S_{j+1}^n) + \phi(k)\Lambda}{1 + 2\lambda r(k) + \phi(k)(\beta I_j^n + \mu)}, \\ E_j^{n+1} = \frac{E_j^n + \lambda r(k)(E_{j-1}^n + E_{j+1}^n) + \phi(k)\beta p S_j^{n+1} I_j^n}{1 + 2\lambda r(k) + \phi(k)(\delta + \eta + \mu)}, \\ I_j^{n+1} = \frac{I_j^n + \lambda r(k)(I_{j-1}^n + I_{j+1}^n) + \phi(k)[\beta(1-p) S_j^{n+1} I_j^n + \delta E_j^{n+1}]}{1 + 2\lambda r(k) + \phi(k)(\gamma + \mu)}, \\ R_j^{n+1} = \frac{R_j^n + \lambda r(k)(R_{j-1}^n + R_{j+1}^n) + \phi(k)[\eta E_j^{n+1} + \gamma I_j^{n+1}]}{1 + 2\lambda r(k) + \phi(k)\mu}, \end{cases} \quad (6)$$

where we used the abbreviation $r(k) = \phi(k)/h^2$ for the generalized parabolic mesh ratio. From (5) we obtain the difference equation for the discrete total population

$$\frac{N_j^{n+1} - N_j^n}{\phi(k)} = \lambda \Delta_h^2 N_j^n + \Lambda - \mu N_j^{n+1}, \quad n = 0, 1, 2, \dots, \quad (7)$$

which is the discrete analogue to (4).

3.1.2. Two-Dimensional NSFD Scheme

For a two-dimensional domain $\Omega = [0, L_x] \times [0, L_y]$ with uniform mesh spacings $h_x = h_y = h$, let $S_{j,\ell}^n \approx S(t_n, x_j, y_\ell)$ denote the pointwise approximation, with indices j, ℓ for space and n for time. Using the same methodology as in the previous section with

$$\Delta_h^2 S_{j,\ell}^n = \frac{S_{j+1,\ell}^n - 2S_{j,\ell}^{n+1} + S_{j-1,\ell}^n}{h^2} + \frac{S_{j,\ell+1}^n - 2S_{j,\ell}^{n+1} + S_{j,\ell-1}^n}{h^2},$$

the 2D NSFDF scheme reads

$$\begin{cases} S_{j,\ell}^{n+1} = \frac{S_{j,\ell}^n + \lambda r(k) \sum_{\sigma=\pm 1} (S_{j+\sigma,\ell}^n + S_{j,\ell+\sigma}^n) + \phi(k)\Lambda}{1 + 4\lambda r(k) + \phi(k)(\beta I_{j,\ell}^n + \mu)}, \\ E_{j,\ell}^{n+1} = \frac{E_{j,\ell}^n + \lambda r(k) \sum_{\sigma=\pm 1} (E_{j+\sigma,\ell}^n + E_{j,\ell+\sigma}^n) + \phi(k)\beta p S_{j,\ell}^{n+1} I_{j,\ell}^n}{1 + 4\lambda r(k) + \phi(k)(\delta + \eta + \mu)}, \\ I_{j,\ell}^{n+1} = \frac{I_{j,\ell}^n + \lambda r(k) \sum_{\sigma=\pm 1} (I_{j+\sigma,\ell}^n + I_{j,\ell+\sigma}^n) + \phi(k)[\beta(1-p)S_{j,\ell}^{n+1} I_{j,\ell}^n + \delta E_{j,\ell}^{n+1}]}{1 + 4\lambda r(k) + \phi(k)(\gamma + \mu)}, \\ R_{j,\ell}^{n+1} = \frac{R_{j,\ell}^n + \lambda r(k) \sum_{\sigma=\pm 1} (R_{j+\sigma,\ell}^n + R_{j,\ell+\sigma}^n) + \phi(k)[\eta E_{j,\ell}^{n+1} + \gamma I_{j,\ell}^{n+1}]}{1 + 4\lambda r(k) + \phi(k)\mu}. \end{cases} \quad (8)$$

3.1.3. Qualitative Properties of the Scheme

The NSFDF schemes (6)–(8) preserve several essential qualitative properties of the continuous epidemic model. In particular, nonnegative initial conditions guarantee nonnegative numerical solutions for all compartments and for any choices of the step sizes $k, h > 0$. Moreover, the *discrete total population* $N_{j,\ell}^n = S_{j,\ell}^n + E_{j,\ell}^n + I_{j,\ell}^n + R_{j,\ell}^n$ remains uniformly bounded by Λ/μ , consistently with the continuous formulation. The implicit treatment of the reactive terms together with the nonstandard denominator function $\phi(k)$ also provides enhanced numerical stability, allowing the use of relatively large time steps k without numerical blow-up.

3.2. Numerical Properties and Rationale for the NSFDF Scheme

The construction of the above NSFDF scheme adheres to Mickens’s foundational principles for structure-preserving discretizations [31, 32]. Standard FD methods often fail to maintain qualitative consistency when applied to nonlinear epidemiological PDEs, leading to negative densities or instability. In contrast, the present formulation introduces a *nonstandard denominator function* $\phi(k)$ and a *local nonlinear stabilizer* $\nu(U_j^n)$, that collectively ensure positivity, boundedness, and improved temporal consistency.

This framework builds on advances in NSFDF schemes for reaction-diffusion systems [33–36]. Specifically, the local correction term $\nu(U_j^n)$ compensates for the order reduction arising from the positivity-preserving Laplacian discretization, as demonstrated in [33]. The resulting scheme achieves first-order accuracy in time and second-order accuracy in space, with a local truncation error $\mathcal{O}(k) + \mathcal{O}(h^2)$.

The positivity of the denominator function $\phi(k)$ for all $k > 0$, guarantees the unconditional nonnegativity of the discrete solution. Meanwhile, the implicit-explicit coupling of the diffusive and reactive terms enhances the numerical stability, even for relatively large parabolic ratios $r = k/h^2$. These results align with stability analyses reported in [33–35], confirming the reliability of the NSFDF approach in generating stable, high-fidelity datasets for data-driven modeling frameworks.

For both the 1D and 2D schemes, and in line with [33, 36], the non-standard denominator function is chosen as the identity function $\phi(k) = k$ for the sake of simplicity, in order to simplify the proposed data generation schemes. For a more advanced choice of $\phi(k)$ originating from the asymptotic behaviour of the total population for $t \rightarrow \infty$, we refer the reader to [30].

4. The Physics-Informed Neural Network Framework

In this section, we describe the PINN framework that we used to approximate the solutions of the SEIR reaction-diffusion system (1). PINNs integrate observational data, initial and boundary conditions, and the governing PDE system into a unified loss function. This enables both forward simulation and inverse parameter estimation.

4.1. The Neural Network Architecture

We approximate the state variables

$$(S, E, I, R)(t, x) \approx \mathcal{NN}_\theta^{\Theta_p}(t, x) := (\hat{S}_\theta, \hat{E}_\theta, \hat{I}_\theta, \hat{R}_\theta)(t, x),$$

where θ denotes the vector of trainable neural network parameters (so-called hyperparameters), including all network weights, biases, and, in the inverse problem setting, a subset of unknown epidemiological parameters to be identified from data. Here, the set of trainable physical parameters is chosen as

$$\Theta_p = (\beta, \delta, \gamma, \lambda)^\top \in \mathbb{R}^4.$$

These parameters in Θ_p were selected because they directly influence the dynamics of epidemic transmission, the temporal progression of the disease, and the mechanisms of spatial spreading. Furthermore, they are typically challenging to measure accurately and can vary significantly across outbreaks, geographical regions, and intervention strategies. Thus, estimating these parameters from observational data is an important inverse problem in epidemiological modeling.

In contrast, the demographic parameters Λ and μ are assumed to be known constants and are not included in the training procedure. These quantities correspond to long-term demographic processes, such as birth rates and natural mortality, which evolve on slower timescales than epidemic dynamics. They can often be estimated independently using demographic statistics.

Furthermore, the parameter p and the transition coefficient η are treated as fixed quantities to reduce the dimensionality of the inverse problem and improve parameter identifiability during training. Estimating too many parameters simultaneously can cause optimization instability and non-uniqueness issues, especially when observational data is sparse or noisy. Thus, the chosen subset of parameters in Θ_p strikes a balance between epidemiological relevance, practical identifiability, and computational stability within the PINN framework. Each compartment is approximated by a fully connected feedforward neural network. The unknown parameters in Θ_p are optimized jointly with the neural network weights using gradient-based optimization and automatic differentiation.

The inputs to the neural network are the temporal coordinate $t \in [0, T]$ and the spatial coordinates $x \in \Omega$ for $d = 1$, or $(x, y) \in \Omega$ for $d = 2$, while the outputs correspond to the four epidemiological compartments $(\hat{S}, \hat{E}, \hat{I}, \hat{R})$. To improve the representation of high-frequency spatial patterns and mitigate the spectral bias commonly encountered in standard multilayer perceptrons, *Fourier feature mappings* [37, 38] are employed for the spatial inputs. Specifically, the spatial coordinates are transformed according to

$$\gamma(x) = [\sin(2\pi\mathbf{B}x), \cos(2\pi\mathbf{B}x)],$$

where \mathbf{B} is a matrix of randomly sampled frequencies. This encoding significantly enhances the expressive power of the PINN, enabling the accurate reconstruction of sharp epidemic fronts and localized infection waves in both one- and two-dimensional domains. The network consists of L hidden layers with nonlinear activations (`tanh` or `swish`), and residual skip connections are incorporated to improve trainability and stability for long-time dynamics.

Figure 2 illustrates the architecture of the proposed SEIR-PINN. The neural network takes the time t and the spatial coordinates $x \in \Omega$ as inputs and outputs the approximate compartmental fields $(\hat{S}, \hat{E}, \hat{I}, \hat{R})$. Automatic differentiation computes the temporal and spatial derivatives, enabling the evaluation of the PDE residuals corresponding to the reaction-diffusion SEIR system (1). Together with the data misfit, boundary and initial condition losses, and epidemiological constraints, these residuals contribute to the total composite loss (9), which guides the training process. These residuals, along with the data misfit, boundary and initial condition

losses, and epidemiological constraints, contribute to the total composite loss (9), which guides the training process.

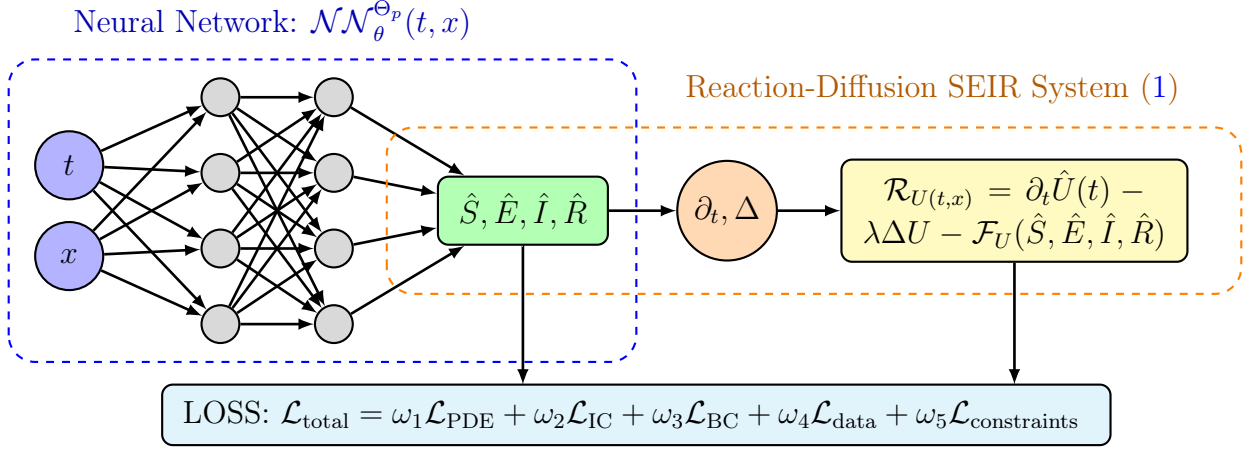


Figure 2: Schematic diagram of the proposed PINN architecture for the reaction-diffusion SEIR model (1)–(3).

4.2. The Loss Function Design

The total loss function is defined as a weighted sum of multiple components

$$\mathcal{L}_{\text{total}} = \omega_1 \mathcal{L}_{\text{PDE}} + \omega_2 \mathcal{L}_{\text{IC}} + \omega_3 \mathcal{L}_{\text{BC}} + \omega_4 \mathcal{L}_{\text{data}} + \omega_5 \mathcal{L}_{\text{constraints}}, \quad (9)$$

where each component is explained below.

(i) *PDE Residual Loss.* For each compartment, the PDE residual is computed via automatic differentiation (AD) of the neural outputs

$$\mathcal{R}_S = \partial_t \hat{S} - \lambda \Delta \hat{S} - \Lambda + \beta \hat{S} \hat{I} + \mu \hat{S},$$

with analogous definitions \mathcal{R}_E , \mathcal{R}_I , and \mathcal{R}_R . The PDE loss is

$$\mathcal{L}_{\text{PDE}} = \frac{1}{N_r} \sum_{i=1}^{N_r} (\mathcal{R}_S^2 + \mathcal{R}_E^2 + \mathcal{R}_I^2 + \mathcal{R}_R^2)(t_i, x_i),$$

evaluated at N_r randomly sampled collocation points $(t_i, x_i) \in \mathcal{U}_T$.

(ii) *Initial Condition Loss.* Enforces agreement with prescribed initial data

$$\mathcal{L}_{\text{IC}} = \frac{1}{N_{\text{IC}}} \sum_{i=1}^{N_{\text{IC}}} \left(|\hat{S}_0(x_i) - S_0(x_i)|^2 + |\hat{E}_0(x_i) - E_0(x_i)|^2 + |\hat{I}_0(x_i) - I_0(x_i)|^2 + |\hat{R}_0(x_i) - R_0(x_i)|^2 \right),$$

where N_{IC} denotes the number of sampled points at $t = 0$.

(iii) *Boundary Condition Loss.* Homogeneous Neumann conditions are enforced weakly by penalizing the normal derivative

$$\mathcal{L}_{\text{BC}} = \frac{1}{N_{\partial\Omega}} \sum_{i=1}^{N_{\partial\Omega}} \left(|\nabla \hat{S} \cdot \vec{n}|^2 + |\nabla \hat{E} \cdot \vec{n}|^2 + |\nabla \hat{I} \cdot \vec{n}|^2 + |\nabla \hat{R} \cdot \vec{n}|^2 \right)(t_i, x_i),$$

for $N_{\partial\Omega}$ sampled boundary points $(t_i, x_i) \in \Sigma_T$.

(iv) *Data Loss.* When synthetic observations are available at discrete time-space locations $\{(t_j, x_j)\}$, we minimize the misfit between the network outputs and the data values \mathcal{D}_j

$$\mathcal{L}_{\text{data}} = \frac{1}{N_d} \sum_{j=1}^{N_d} |\hat{U}(t_j, x_j) - \mathcal{D}_j|^2,$$

where N_d represents the number of observational data points and U denotes the observed compartments (e.g., I or (I, R)).

(v) *Constraint Loss.* To ensure epidemiological consistency, the PINN framework incorporates additional constraint-based penalty terms within the loss function. Non-negativity of the compartment populations is enforced through soft penalties of the form $\max(0, -\hat{U})^2$ penalties for $U \in \{S, E, I, R\}$. Furthermore, deviations from the expected population balance induced by the vital dynamics are penalized via

$$\mathcal{L}_{\text{pop}} = \frac{1}{N_r} \sum_{i=1}^{N_r} \left(\hat{S} + \hat{E} + \hat{I} + \hat{R} - \frac{\Lambda}{\mu} \right)_+^2,$$

where $(\cdot)_+$ denotes the positive part, ensuring that the total population does not exceed its carrying capacity. In addition, admissibility of learned epidemiological parameters is guaranteed through bounded transformations such as $\beta = \beta_{\max} \sigma(\tilde{\beta})$, where σ denotes the logistic sigmoid function, ensuring $0 < \beta < \beta_{\max}$. Altogether, the constraint contribution to the optimization problem is given by

$$\mathcal{L}_{\text{constraints}} = \mathcal{L}_{\text{nonneg}} + \mathcal{L}_{\text{pop}} + \mathcal{L}_{\text{param}}.$$

In the present work, we formulate the components of the composite loss function using the standard L^2 -based mean squared error (MSE). The L^2 -based MSE is widely used in the PINN literature due to its smooth optimization properties and compatibility with gradient-based training methods. Recent studies have shown that alternative losses, such as L^1 -based formulations, may improve robustness in certain situations [39]. However, we use the L^2 -loss in this work because the considered SEIR reaction-diffusion system produces sufficiently smooth solutions and stable residual distributions. This makes the L^2 formulation well-suited for our framework.

4.3. The Training Strategy

The training procedure integrates the composite loss function defined in Section 4.2 with a staged optimization strategy. This strategy is designed to enhance convergence stability and parameter identifiability. The complete workflow is summarized below in Algorithm 1.

The proposed SEIR-PINN framework enables three complementary learning modes:

Forward simulation: Approximating the full spatiotemporal evolution $(S, E, I, R)(t, x)$ given initial and boundary conditions.

Inverse identification: Estimating unknown epidemiological and diffusion parameters from sparse or noisy data.

Constraint-aware prediction: Ensuring epidemiological plausibility by embedding population conservation and non-negativity directly into the optimization.

Algorithm 1 thus provides a stable and interpretable training process capable of reconstructing epidemic dynamics and inferring key transmission parameters with high fidelity.

Algorithm 1 SEIR-PINN Training Algorithm

- 1: *Initialization*: Neural network parameters θ are initialized using Xavier initialization [40]. When the inverse problem is considered, the epidemiological and diffusion parameters $\Theta_p = (\beta, p, \delta, \eta, \gamma, \lambda)$ are initialized within biologically admissible ranges inferred from prior studies, while demographic quantities (Λ, μ) are fixed.
 - 2: *Staged Optimization*: Training is performed in three successive stages to promote robustness and accuracy:
 - i*: *Stage I (Data fitting)*: The network is pretrained by minimizing $\mathcal{L}_{\text{data}} + \mathcal{L}_{\text{IC}}$ using the Adam optimizer [41] to match observations and initial conditions.
 - ii*: *Stage II (Physics enforcement)*: The PDE residual and boundary terms are introduced, and the total composite loss $\mathcal{L}_{\text{total}}$ in (9) is minimized using Adam with a reduced learning rate to ensure the stable incorporation of physical constraints.
 - iii*: *Stage III (Refinement)*: The optimization switches to the quasi-Newton L-BFGS method [11, 21, 41] for fine-tuning until convergence, achieving high-precision satisfaction of the governing equations.
 - 3: *Adaptive Sampling*: To improve sample efficiency and numerical accuracy, collocation points are resampled dynamically at each epoch, with a higher probability in regions with large PDE residuals. Training is accelerated for two-dimensional domains via GPU parallelization and mixed-precision arithmetic.
 - 4: *Monitoring and Constraints*: Non-negativity, boundedness, and population balance constraints (Section 4.2) are softly enforced during training. Throughout the process, the total loss and its individual components are monitored, and early stopping is applied when the relative improvement falls below a predefined tolerance threshold.
-

5. Numerical Results

In this section, we present and discuss the numerical results obtained from Sections 3 and 4 applied to the one- and two-dimensional SEIR reaction-diffusion system (1)–(3). The PINN is trained to simultaneously solve the forward problem and recover the unknown epidemiological parameters $\{\beta, \delta, \gamma, \lambda\}$ via the inverse problem formulation described in Section 4.

Synthetic ground-truth data are generated using the NSFD schemes, namely (6) in the one-dimensional setting and (8) in the two-dimensional setting, with both formulations enforcing zero-flux Neumann boundary conditions.

We initialize the system with normalized initial conditions, where each compartment is expressed as a proportion of the total population. All experiments are conducted using the parameter values listed in Table 2, and the true values of the inverse parameters are used as reference targets for evaluating the recovery accuracy. Parameters marked with (*) are estimated through the PINN inverse problem, while the remaining parameters are kept fixed.

Table 2: Parameter values used in all numerical experiments.

Parameter	Symbol	Value	Reference	Role	Status
Recruitment rate	Λ	1.0000	[33, 42]	Fixed	—
Death rate	μ	0.0100	[43–45]	Fixed	—
Transmission rate	β	0.4000	[43–49]	Inverse	(*)
Direct exposure fraction	p	0.3000	Assumed	Fixed	—
Progression rate ($E \rightarrow I$)	δ	0.3000	[46, 47]	Inverse	(*)
Recovery rate ($E \rightarrow R$)	η	0.1000	[43–48]	Fixed	—
Recovery rate ($I \rightarrow R$)	γ	0.2000	[44, 45]	Inverse	(*)
Diffusion coefficient	λ	0.0500	[33, 42]	Inverse	(*)

5.1. One-Dimensional Results

We begin with the one-dimensional configuration in which the spatial domain is $\Omega = [0, L]$ with $L = 1$ (*dimensionless*) and the time horizon is $T = 5$ (*years*). The initial conditions place a Gaussian seed of exposed and infected individuals centred at $x = 0.5$, while the susceptible population is nearly uniform at $S_0 \approx 90\%$; the recovered compartment is initialized to zero.

The model is trained for 8000 epochs using the Adam optimizer, with an initial learning rate of 10^{-3} and cosine annealing scheduling down to 10^{-5} . To stabilize training, gradient clipping with a maximum norm of 1.0 is employed.

5.1.1. Space–Time Dynamics Comparison

Figure 3 displays the space–time contour maps for all compartments across the domain \mathcal{U} , comparing the reference data (true solution, left column), the PINN prediction (centre column), and the pointwise absolute error $E_A = |\hat{U} - U|$ (right column). The susceptible population decreases monotonically from its initial high value as the disease spreads outward from the center due to diffusion, while the exposed and infected compartments exhibit a wave-like propagation pattern consistent with reaction–diffusion dynamics. The absolute errors are mainly concentrated near zero, which supports that the PINN accurately reproduces the qualitative and quantitative behavior of all four compartments.

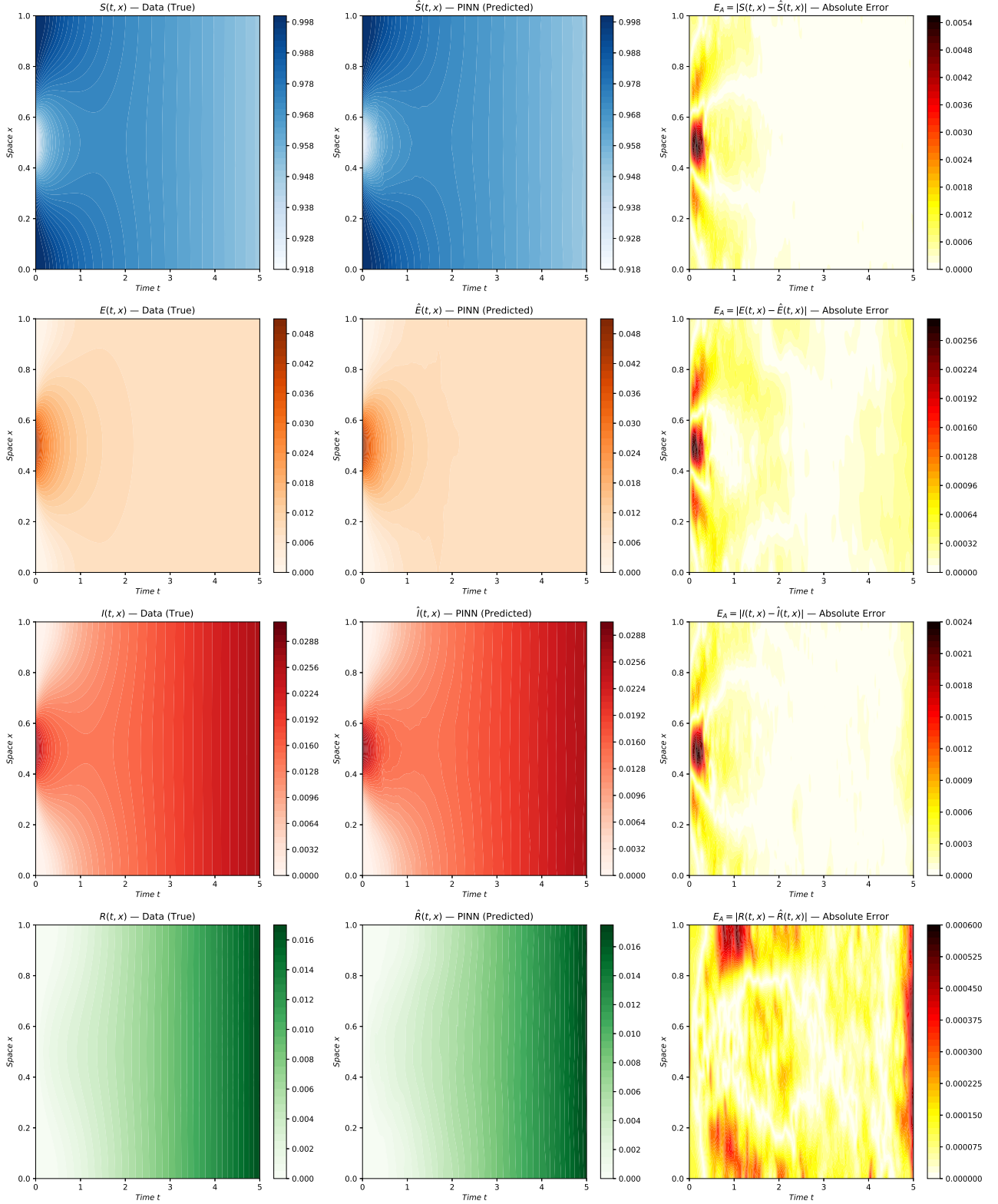


Figure 3: Space–time dynamics: Reference data vs PINN prediction; 1D case.

5.1.2. Fixed-Point Temporal Profiles

Figure 4 compares the PINN prediction against the reference data at two fixed spatial locations: the midpoint $x = 0.50$ and the boundary point $x = 1.00$. In both cases, the PINN accurately reproduces the timing and magnitude of the epidemic peaks, with minor discrepancies observed for the exposed compartment at the boundary location. Overall, the

errors remain small and are concentrated near zero.

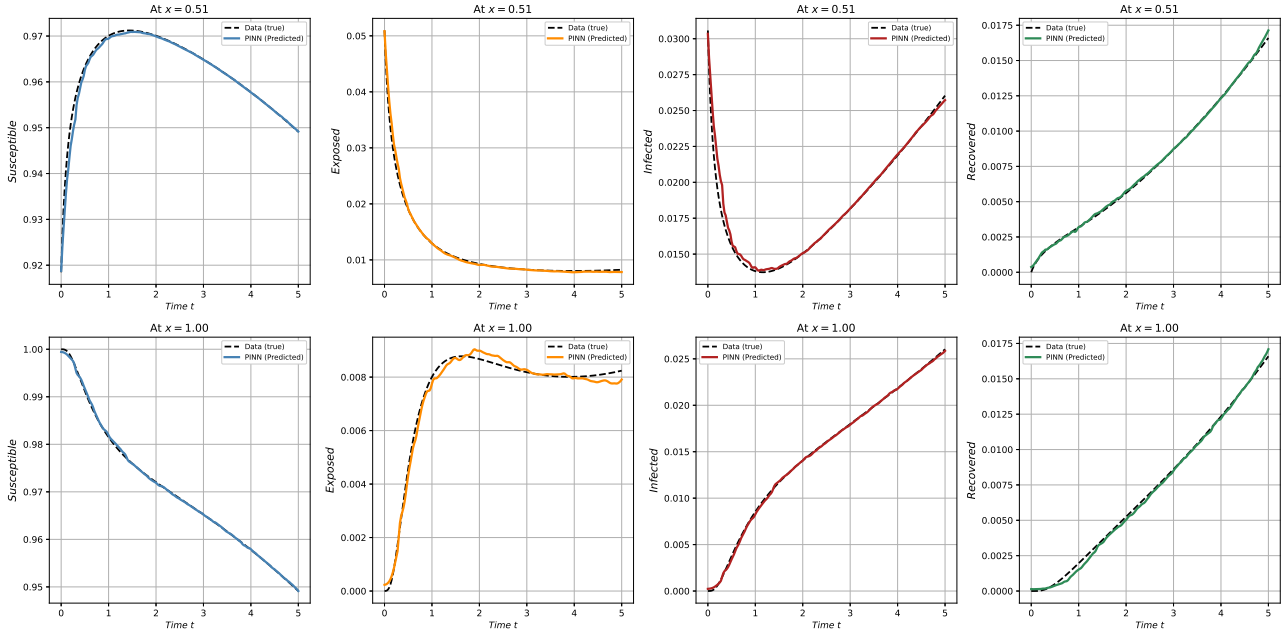


Figure 4: Temporal dynamics at fixed spatial locations: Reference data vs PINN prediction; 1D case.

5.1.3. Training Loss Convergence

Figure 5 illustrates the evolution of each loss component over 8000 training epochs. The total loss decreases by approximately two to three orders of magnitude. The data loss and initial condition loss dominate in the early epochs, guiding the network toward the correct solution manifold.

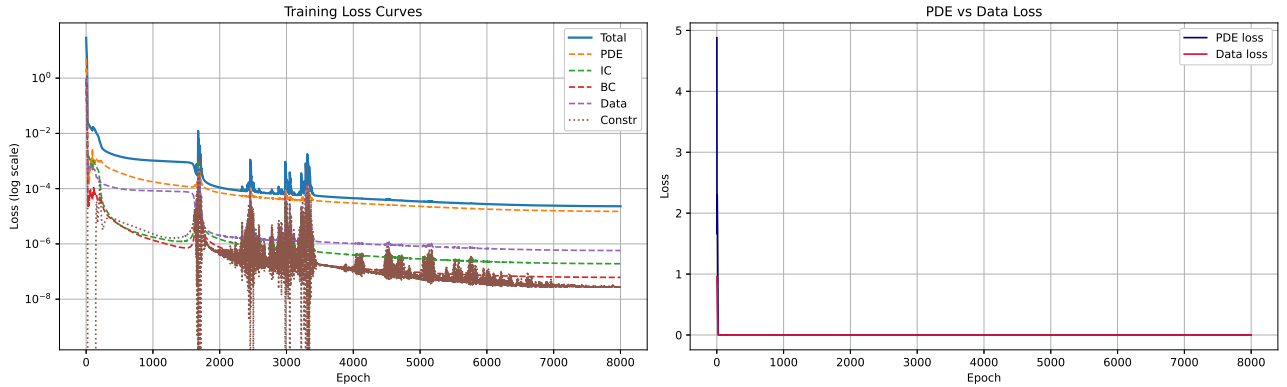


Figure 5: PINN training loss evolution over 8000 epochs; 1D case.

5.1.4. Parameter Recovery

Figure 6 complements this analysis by illustrating the convergence history of the four estimated parameters β , δ , γ , and λ .

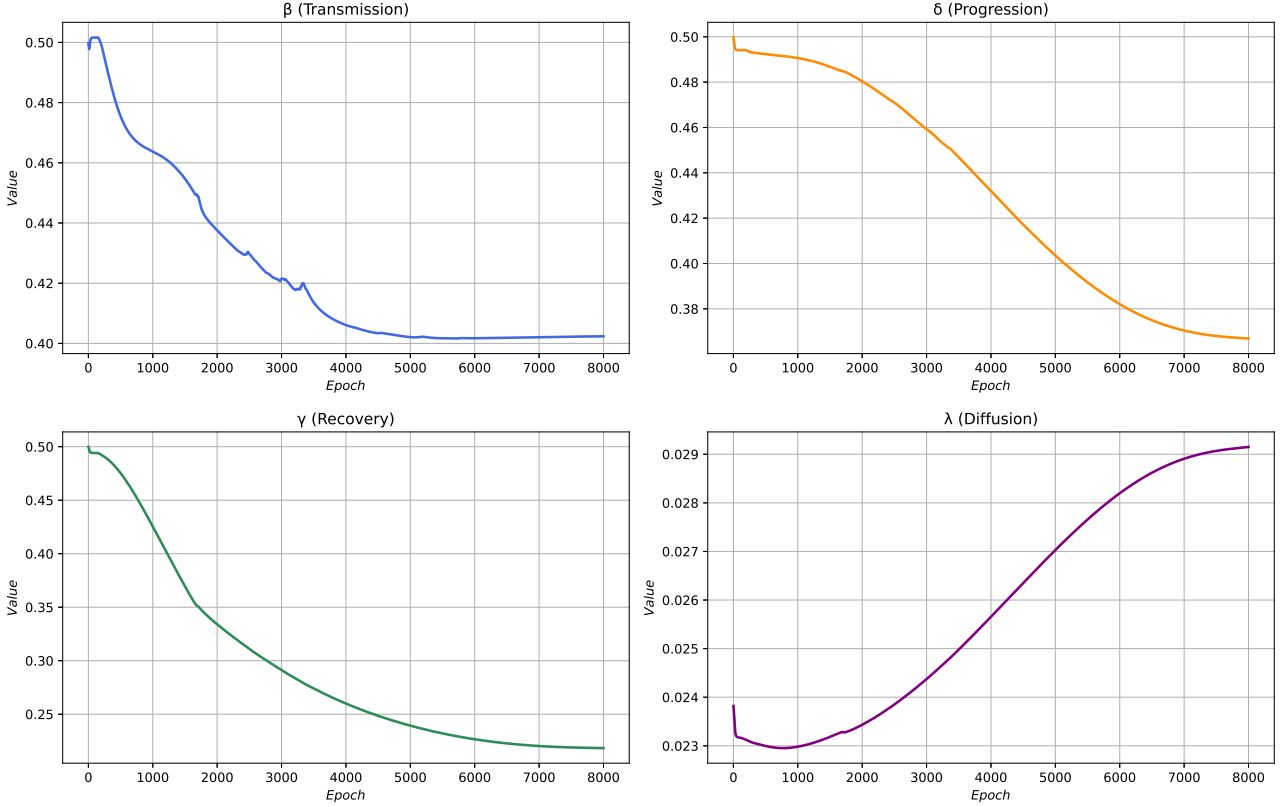


Figure 6: Convergence of estimated parameters during PINN training; 1D case.

5.1.5. Quantitative Error Analysis

Tables 3 and 4 summarize the pointwise error metrics for each compartment and the parameter recovery results, respectively.

Table 3: Quantitative error metrics between the PINN predictions and the reference data; 1D case.

Compartment	Rel. L^2 Error	MAE	RMSE	Max Error
S	6.0606 e-04	2.5171 e-04	5.8601 e-04	5.5081 e-03
E	3.5408 e-02	2.2358 e-04	3.5202 e-04	2.7316 e-03
I	1.4550 e-02	1.2322 e-04	2.5424 e-04	2.3538 e-03
R	1.6345 e-02	1.0933 e-04	1.4441 e-04	5.9132 e-04

Table 4: Comparison between the true and PINN-estimated values of the inferred parameters; 1D case.

Parameter	True Value	Estimated	Abs. Error	Rel. Error
β	0.40	0.4024	5.9132 e-04	0.59%
δ	0.30	0.3669	6.6900 e-02	22.30%
γ	0.20	0.2183	1.8300 e-02	9.15%
λ	0.05	0.0292	2.0800 e-02	41.60%

5.2. Two-Dimensional Results

We now extend the numerical investigation to the two-dimensional setting with spatial domain $\Omega = [0, L_x] \times [0, L_y]$, $L_x = L_y = L$, and the same time horizon T . The added spatial dimension substantially increases both the complexity of the forward problem and the difficulty of the inverse identification task, as the Laplacian now acts on two spatial directions simultaneously and the boundary condition must be enforced on all four walls of the domain.

Training proceeds for 10,000 epochs with the same Adam optimizer and cosine annealing schedule. The four Neumann boundary conditions are enforced independently for each wall: the left and right walls enforce $\partial_x \hat{U} = 0$ while the top and bottom walls enforce $\partial_y \hat{U} = 0$, with each wall contributing an independent term to the total boundary loss.

5.2.1. Spatial Snapshots at Five Time Instances

Figure 7 presents the spatial distribution of each compartment at five representative time instances $t = 0, T/4, T/2, 3T/4, T$. For each compartment, the reference data and the PINN prediction are shown as filled contour plots. At $t = 0$, the infectious seed is confined to a small neighborhood around the center of the domain, and both the reference data and the PINN prediction show close agreement, as indicated by the colormap scales. As time progresses, the disease diffuses radially outward and simultaneously grows due to the nonlinear reaction kinetics, producing a spreading wave visible in the I and E maps. The susceptible population S is depleted most rapidly at the centre, while the recovered compartment R accumulates from the centre outward. The PINN accurately reproduces this spatio-temporal evolution across all five snapshots, capturing both the diffusive spreading and the reaction-driven dynamics.

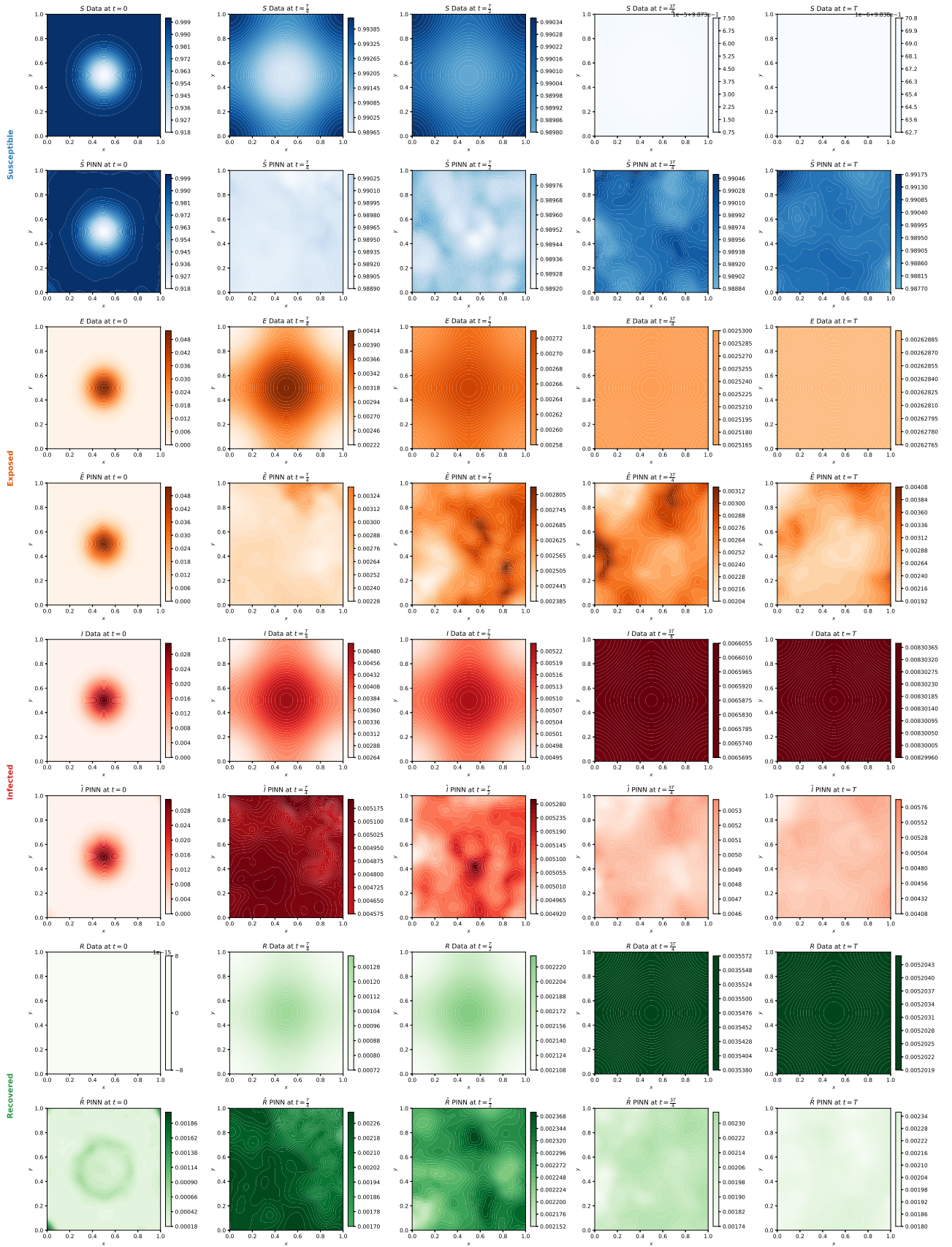


Figure 7: Spatial snapshots: Reference data vs PINN prediction; 2D case.

5.2.2. Three-Dimensional Surface Comparison

Figure 8 presents 3D surface plots of the reference data and PINN prediction for all four compartments at the representative time instances $t = 0, T/2, T$. The surface visualizations clearly highlight the spatial heterogeneity of the compartment distributions. The PINN surfaces closely follow the reference data in terms of height, shape, and boundary behavior, with only minor discrepancies, as further illustrated in Figure 9 and Table 5. No spurious oscillations are observed near the domain boundaries, confirming the effectiveness of the Neumann boundary loss in the two-dimensional setting.

Susceptible

Exposed

Infected

Recovered

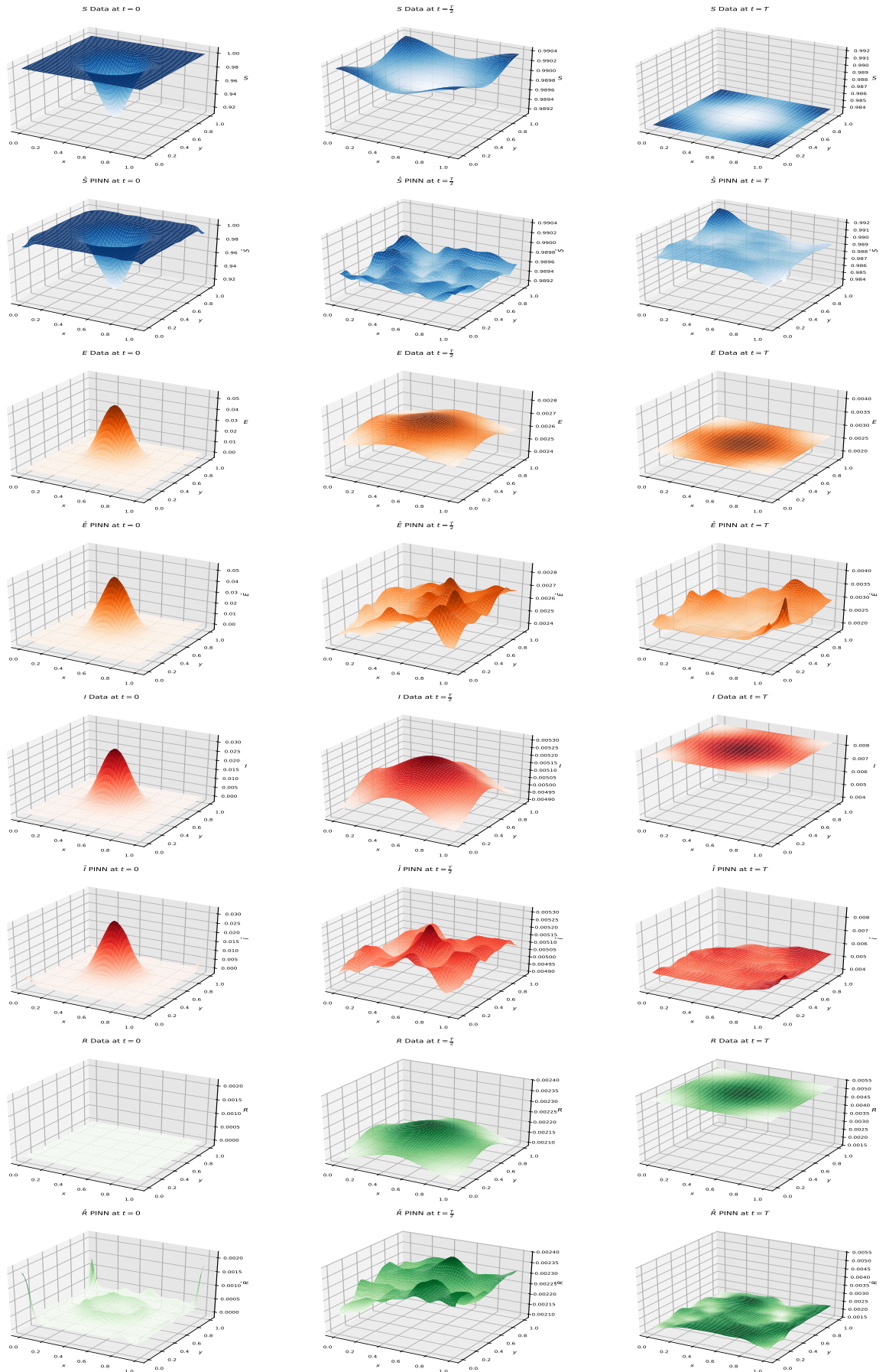


Figure 8: Three-dimensional surface: Reference data vs PINN prediction; 2D case.

5.2.3. Pointwise Error Maps

Figure 9 displays the absolute error maps $E_A = |\hat{U} - U|$ for all four compartments at five time snapshots $t = 0, T/4, T/2, 3T/4, T$. Importantly, the errors are distributed across the domain without any systematic spatial bias, confirming that the collocation strategy provides adequate coverage of the spatio-temporal domain. Table 5 reports the corresponding quantitative error metrics between the PINN predictions and the reference data.

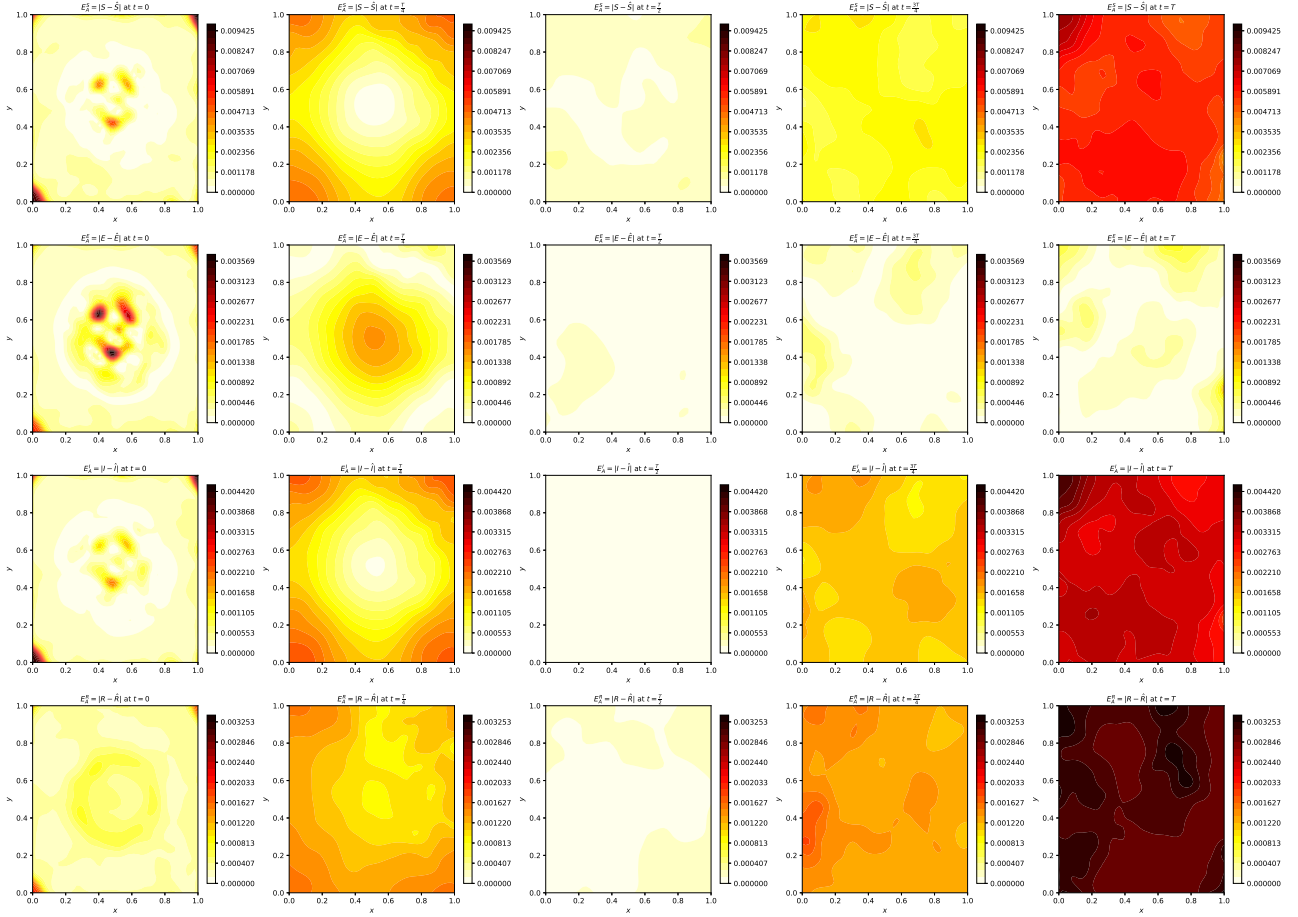


Figure 9: Pointwise error maps at five time snapshots; 2D case.

5.2.4. Training Loss Convergence

Figure 10 shows the loss component histories over the 10 000 training epochs. The overall convergence behaviour is qualitatively similar to the 1D case, Figure 5, but with a slower rate of decrease reflecting the greater approximation difficulty.

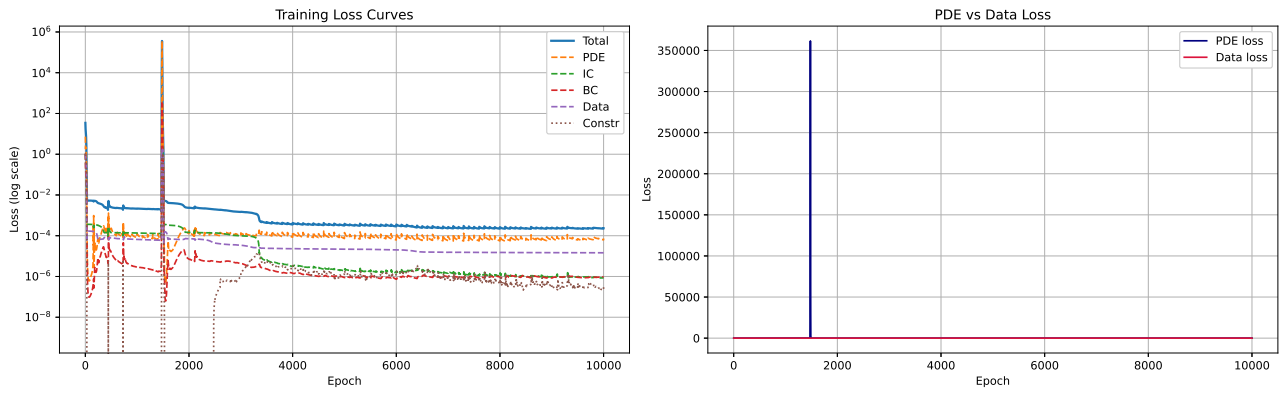


Figure 10: PINN training loss evolution over 10000 epochs; 2D case.

5.2.5. Parameter Recovery

Figure 11 complements this analysis by illustrating the convergence history of the four estimated parameters β , δ , γ , and λ in the 2D setting. The parameter recovery is slightly less accurate than in 1D case, as expected due to the larger hypothesis space and the smaller relative contribution of each data point to the parameter gradients. Table 6 reports the comparison between the true and PINN-estimated values of the inferred parameters.

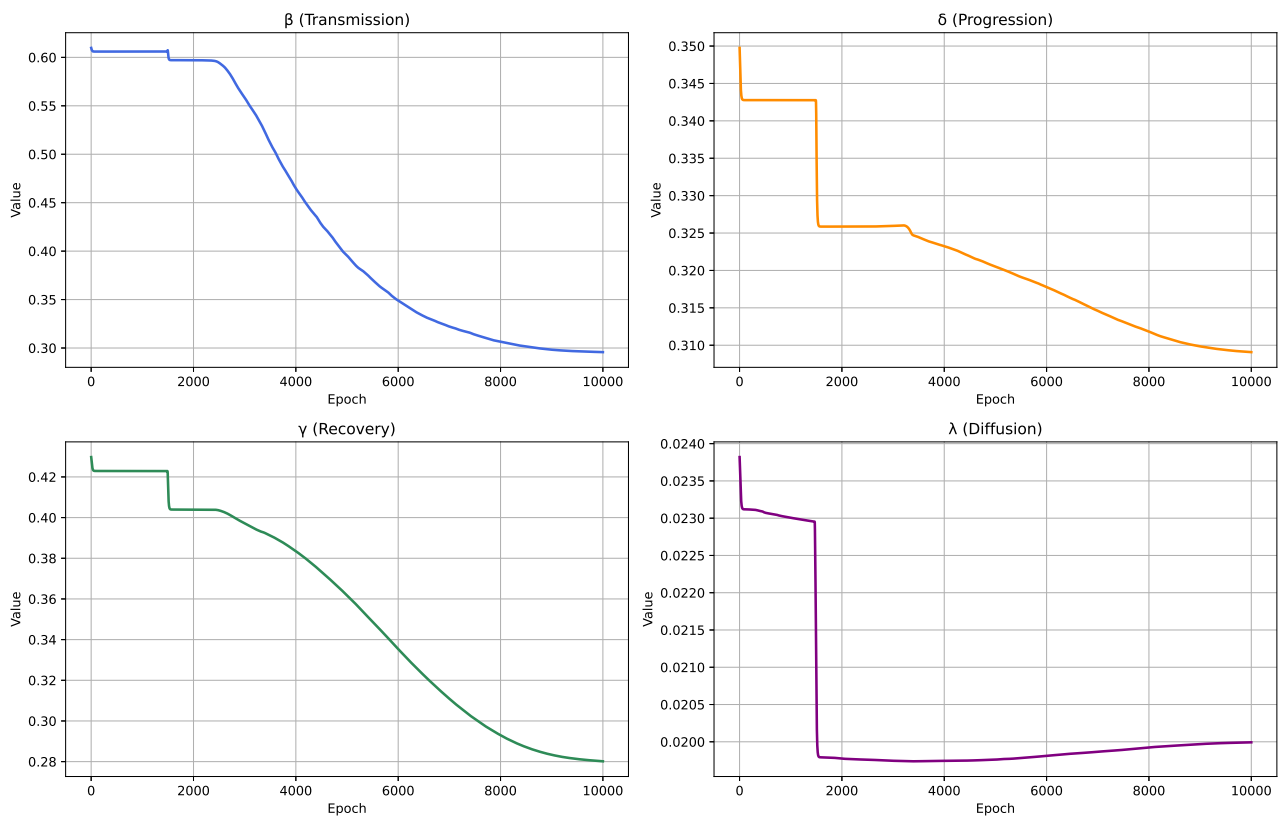


Figure 11: Convergence of estimated parameters during PINN training; 2D case.

5.2.6. Quantitative Error Analysis

Tables 5 and 6 summarize the pointwise error metrics for each compartment and the parameter recovery results, respectively.

Table 5: Quantitative error metrics between the PINN predictions and the reference data; 2D case.

Compartment	Rel. L^2 Error	MAE	RMSE	Max Error
S	3.1121 e-03	2.3606 e-03	3.0792 e-03	9.8181 e-03
E	8.3622 e-02	2.7817 e-04	4.3298 e-04	3.7180 e-03
I	2.8472 e-01	1.3021 e-03	1.7549 e-03	4.6044 e-03
R	5.2685 e-01	1.2035 e-03	1.5878 e-03	3.3886 e-03

Table 6: Comparison between the true and PINN-estimated values of the inferred parameters; 2D case.

Parameter	True Value	Estimated	Abs. Error	Rel. Error
β	0.40	0.2957	1.0426 e-01	26.06%
δ	0.30	0.4091	1.0910 e-01	36.37%
γ	0.20	0.2802	8.0200 e-02	40.10%
λ	0.05	0.0200	3.0000 e-02	60.00%

In this case, the observed estimation errors are primarily due to the initialization and optimization of the unknown parameters within the broad interval $(0, 1)$ during the training process, which corresponds to a general identification setting. If prior information about narrower admissible parameter intervals were available, the accuracy of the parameter estimation could be significantly improved. In the present work, however, we intentionally considered the general interval $(0, 1)$ since no precise prior bounds on the parameters were assumed.

6. Conclusion and Future Work

In this work, we developed a constraint-aware PINN framework for solving and identifying parameters in SEIR reaction-diffusion epidemic systems with vital dynamics and homogeneous Neumann boundary conditions. The proposed methodology combines epidemiological constraints, observational data, and governing partial differential equations within a unified optimization framework capable of addressing both forward simulation and inverse parameter estimation problems. To overcome the scarcity of high-resolution spatial epidemiological datasets, synthetic benchmark data were generated using positivity-preserving NSFD schemes. These schemes preserve essential qualitative properties of the epidemic dynamics, including non-negativity, boundedness, and numerical stability, thereby providing reliable reference solutions for training and validation. The numerical experiments conducted in one- and two-dimensional spatial domains demonstrated that the proposed PINN framework accurately reconstructs the spatiotemporal evolution of all epidemiological compartments while maintaining low approximation errors throughout the computational domain. The proposed framework establishes an effective connection between mechanistic epidemiological modeling and modern scientific machine learning, offering promising perspectives for the analysis and prediction of complex spatiotemporal disease dynamics.

Several research directions may further extend the present work. First, the proposed framework may be generalized to more complex epidemiological structures, including SEIARD, age-structured, delayed, fractional-order, stochastic, or multi-strain reaction-diffusion systems. Incorporating heterogeneous diffusion coefficients and spatially varying transmission parameters would further improve the realism of the model in practical applications. Another promising

direction concerns the integration of real epidemiological datasets. Future studies may investigate the assimilation of noisy and incomplete observations obtained from mobility networks, geographic surveillance systems, or public health databases. Such developments would allow the framework to be applied to realistic forecasting, monitoring, and epidemic control scenarios. In addition, extending the framework to stochastic reaction–diffusion epidemic models would enable the analysis of environmental variability, demographic fluctuations, random mobility patterns, and uncertainty propagation in spatiotemporal disease transmission. From a computational perspective, future work may focus on improving the efficiency and scalability of PINN training for large-scale multidimensional problems. Possible directions include adaptive residual sampling strategies, domain decomposition techniques, operator-learning architectures, and hybrid numerical–neural methodologies. Finally, incorporating optimal control strategies and uncertainty quantification within the PINN framework represents another important research direction. Such developments could provide valuable tools for evaluating intervention policies, estimating confidence intervals for inferred parameters, and supporting data-driven decision-making in epidemic management and public health planning.

Declarations

Author Contributions

A. Zinihi: Conceptualization, Software, Methodology, Validation, Formal Analysis, Investigation, Writing-Original Draft, Writing-Review and Editing, Visualization.

M. Ehrhardt: Conceptualization, Supervision, Methodology, Formal Analysis, Investigation, Writing-Review and Editing.

Computational Setup

All numerical simulations for the PINN framework were conducted on a workstation running Microsoft Windows 11 Home. The system features an Intel Core Ultra 9 275HX CPU (24 cores, 24 threads) with 32 GB of RAM.

Computations were accelerated using a hybrid GPU setup, including an integrated Intel GPU and an NVIDIA GeForce RTX 5070 Ti Laptop GPU (12 GB VRAM).

Software Environment

The simulations were implemented in *Python 3.13* using PyTorch.

Data Availability

This study uses synthetic datasets as described in Sections 3 and 5. All data analyzed or generated in this work, which support the results, are included in this article.

Conflict of Interest

The authors declare that there are no problems or conflicts of interest between them that may affect the study in this paper.

References

- [1] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. Royal Soc. A*, 115(772):700–721, 1927.
- [2] H. W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.

- [3] A. Zinihi, M. R. Sidi Ammi, and D. F. M. Torres. Fractional differential equations of a reaction-diffusion SIR model involving the Caputo-fractional time-derivative and a nonlinear diffusion operator. *Evolut. Eqs. Contr. Theo.*, 14(5):944–967, 2025.
- [4] W. Wang, Y. Cai, M. Wu, K. Wang, and Z. Li. Complex dynamics of a reaction-diffusion epidemic model. *Nonlin. Anal. Real World Appl.*, 13(5):2240–2258, 2012.
- [5] N. Wang, L. Zhang, and Z. Teng. Dynamics in a reaction-diffusion epidemic model via environmental driven infection in heterogenous space. *J. Biol. Dynam.*, 16(1):373–396, 2021.
- [6] A. Zinihi, M. R. Sidi Ammi, and M. Ehrhardt. Optimal control of a diffusive epidemiological model involving the Caputo-Fabrizio fractional time-derivative. *Part. Diff. Eqs. Appl. Math.*, 14:101188, 2025.
- [7] L. Chang, S. Gao, and Z. Wang. Optimal control of pattern formations for an SIR reaction-diffusion epidemic model. *J. Theor. Bio.*, 536:111003, 2022.
- [8] J. C. Strikwerda. *Finite difference schemes and partial differential equations*. Society for Industrial and Applied Mathematics, 2004.
- [9] A. Zinihi, M. Ehrhardt, and M. R. Sidi Ammi. Spatiotemporal SEIQR epidemic modeling with optimal control for vaccination, treatment, and social measures, 2025. arXiv preprint.
- [10] H. T. Banks and K. Kunisch. *Estimation techniques for distributed parameter systems*. Birkhäuser Boston, 1989.
- [11] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, 378:686–707, 2019.
- [12] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Rev. Phys.*, 3(6):422–440, 2021.
- [13] G. Kissas, Y. Yang, E. Hwuang, W. R. Witschey, J. A. Detre, and P. Perdikaris. Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks. *Comput. Meth. Appl. Mech. Engrg.*, 358:112623, 2020.
- [14] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli. Scientific machine learning through physics-informed neural networks: Where we are and what’s next. *J. Sci. Comput.*, 92(3), 2022.
- [15] J. D. Toscano, V. Oommen, A. J. Varghese, Z. Zou, N. Ahmadi Daryakenari, C. Wu, and G. E. Karniadakis. From PINNs to PIKANs: Recent advances in physics-informed machine learning. *Machine Learn. Comput. Sci. Engrg.*, 1(1), 2025.
- [16] A. Zinihi, M. R. Sidi Ammi, and A. Bachir. Multi-city modeling of epidemics using a topology-based SIR model: Neural network-enhanced SAIRD model. *J. Comput. Sci.*, 92:102721, 2025.
- [17] S. Berkhahn and M. Ehrhardt. A physics-informed neural network to model COVID-19 infection and hospitalization scenarios. *Adv. Contin. Discr. Mod.*, 2022(1), 2022.

- [18] C. Millevoi, D. Pasetto, and M. Ferronato. A physics-informed neural network approach for compartmental epidemiological models. *PLoS Comput. Biol.*, 20(9):e1012387, 2024.
- [19] S. P. Nelson, R. Raja, P. Eswaran, J. Alzabut, and G. Rajchakit. Modeling the dynamics of COVID-19 in Japan: Employing data-driven deep learning approach. *Int. J. Mach. Learn. Cybern.*, 2024.
- [20] E. Shamsara, F. König, and N. Pfeifer. An informed deep learning model of the Omicron wave and the impact of vaccination. *Comput. Biol. Med.*, 191:109968, 2025.
- [21] A. Zinihi. Identifying memory effects in epidemics via a fractional SEIRD model and physics-informed neural networks, 2025. arXiv preprint.
- [22] C. P. Igiri and S. Shikaa. Monkeypox transmission dynamics using fractional disease informed neural network: A global and continental analysis. *IEEE Access*, 13:77611–77641, 2025.
- [23] F. Heldmann, S. Berkhahn, M. Ehrhardt, and K. Klamroth. PINN training using biobjective optimization: The trade-off between data loss and residual loss. *J. Comput. Phys.*, 488:112211, 2023.
- [24] H. Cheng, Y. Mao, and X. Jia. A framework based on physics-informed graph neural ODE: For continuous spatial-temporal pandemic prediction. *Appl. Intell.*, 54(24):12661–12675, 2024.
- [25] D. Liu and A. Sopasakis. A combined neural ODE-Bayesian optimization approach to resolve dynamics and estimate parameters for a modified SIR model with immune memory. *Heliyon*, 10(19):e38276, 2024.
- [26] B. Shan, Y. Li, and S.-J. Huang. VI-PINNs: Variance-involved physics-informed neural networks for fast and accurate prediction of partial differential equations. *Neurocomputing*, 623:129360, 2025.
- [27] W. G. Madden, W. Jin, B. Lopman, A. Zuffe, B. Dalziel, C. Jessica E. Metcalf, B. T. Grenfell, and M. S. Y. Lau. Deep neural networks for endemic measles dynamics: Comparative analysis and integration with mechanistic models. *PLoS Comput. Biol.*, 20(11):e1012616, 2024.
- [28] X. Chen, F. Li, H. Lian, and P. Wang. A deep learning framework for predicting the spread of diffusion diseases. *Electronic Research Archive*, 33(4):2475–2502, 2025.
- [29] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Springer Science & Business Media, 2012.
- [30] M. H. Maamar, M. Ehrhardt, and L. Tabharit. A nonstandard finite difference scheme for a time-fractional model of Zika virus transmission. *Math. Biosci. Engrg.*, 21(1):924–962, 2024.
- [31] R. E. Mickens. *Nonstandard finite difference models of differential equations*. World Scientific, 1993.
- [32] R. E. Mickens. *Nonstandard finite difference schemes: Methodology and applications*. World Scientific Publishing Company, 2020.

- [33] A. Zinihi, M. Ehrhardt, and M. R. Sidi Ammi. A nonstandard finite difference scheme for an SEIQR epidemiological PDE model. *Appl. Math. Comput.*, 520:129953, 2026.
- [34] B. M. Chen-Charpentier and H. V. Kojouharov. An unconditionally positivity preserving scheme for advection-diffusion reaction equations. *Math. Comput. Model.*, 57(9–10):2177–2185, 2013.
- [35] S. A. Pasha, Y. Nawaz, and M. S. Arif. On the nonstandard finite difference method for reaction–diffusion models. *Chaos, Solitons & Fractals*, 166:112929, 2023.
- [36] M. Ehrhardt and R. E. Mickens. A nonstandard finite difference scheme for convection-diffusion equations having constant coefficients. *Appl. Math. Comput.*, 219(12):6591–6604, 2013.
- [37] S. Wang, Y. Teng, and P. Perdikaris. Understanding and mitigating gradient pathologies in physics-informed neural networks. *SIAM J. Sci. Comput.*, 43(5):A3055–A3081, 2021.
- [38] M. Tancik, P. P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. T. Barron, and R. Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, pages 7537–7547, 2020.
- [39] C. Wang, S. Li, D. He, and L. Wang. Is L^2 physics informed loss always suitable for training physics informed neural network? *Adv. Neural Inform. Process. Syst.*, 35:8278–8290, 2022.
- [40] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the 13th international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.
- [41] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Math. Programm.*, 45(1-3):503–528, 1989.
- [42] P. Liu and H. X. Li. Global behavior of a multi-group SEIR epidemic model with spatial diffusion in a heterogeneous environment. *International Journal of Applied Mathematics and Computer Science*, 32(2), 2022.
- [43] Centers for Disease Control and Prevention. CDC COVID data tracker; <https://www.cdc.gov/covid-data-tracker/index.html>, Accessed on: May 11, 2026.
- [44] Centers for Disease Control and Prevention. Seasonal influenza (Flu), <https://www.cdc.gov/flu/index.htm>, Accessed on: May 11, 2026.
- [45] A. Zinihi, M. Ehrhardt, and M. R. Sidi Ammi. A Koopman operator framework for nonlinear epidemic dynamics: Application to an SIRSD model. *Nonlinear Science*, 6:100095, April 2026.
- [46] UK Health Security Agency. Epidemiology of COVID-19 in England: January 2020 to December 2024; <https://www.gov.uk/government/publications/epidemiology-of-covid-19-in-england/epidemiology-of-covid-19-in-england-january-2020-to-december-2024>, Accessed on: May 10, 2026.

- [47] World Health Organization. COVID-19 epidemiological update, 24 December 2024; <https://www.who.int/publications/m/item/covid-19-epidemiological-update---24-december-2024>, Accessed on: May 10, 2026.
- [48] Johns Hopkins University. COVID-19 dashboard; <https://coronavirus.jhu.edu/data>, Accessed on: May 11, 2026 (2020).
- [49] European Centre for Disease Prevention and Control. Seasonal influenza data; <https://www.ecdc.europa.eu/en/seasonal-influenza>, Accessed on: May 11, 2026.