Article

Physics-Informed Neural Networks and Functional Interpolation for Data-Driven Parameters Discovery of Epidemiological Compartmental Models

Enrico Schiassi [1], Mario De Florio [1], Andrea D'Ambrosio [1,2], Daniele Mortari [3] and Roberto Furfaro [1,4,*]

1 Systems & Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA; e.schiassi@email.arizona.edu (E.S.); mariodf@email.arizona.edu (M.D.F.); andrea.dambrosio@uniroma1.it (A.D.)

2 School of Aerospace Engineering, Sapienza University of Rome, 00138 Rome, Italy

3 Aerospace Engineering, Texas A&M University, College Station, TX 77843-3141, USA; mortari@tamu.edu

4 Aerospace & Mechanical Engineering, University of Arizona, Tucson, AZ 85721, USA

* Correspondence: robertof@arizona.edu; Tel.: +1-520-621-2525

Abstract: In this work, we apply a novel and accurate Physics-Informed Neural Network Theory of Functional Connections (PINN-TFC) based framework, called Extreme Theory of Functional Connections (X-TFC), for data-physics-driven parameters' discovery of problems modeled via Ordinary Differential Equations (ODEs). The proposed method merges the standard PINNs with a functional interpolation technique named Theory of Functional Connections (TFC). In particular, this work focuses on the capability of X-TFC in solving inverse problems to estimate the parameters governing the epidemiological compartmental models via a deterministic approach. The epidemiological compartmental models treated in this work are Susceptible-Infectious-Recovered (SIR), Susceptible-Exposed-Infectious-Recovered (SEIR), and Susceptible-Exposed-Infectious-Recovered-Susceptible (SEIRS). The results show the low computational times, the high accuracy, and effectiveness of the X-TFC method in performing data-driven parameters' discovery systems modeled via parametric ODEs using unperturbed and perturbed data.

1. Introduction

The concern for viruses' spread has been in the researchers' spotlight for many years [1–4]. Particularly, in the last year and a half, due to the COVID-19 pandemic, this concern has become a hot topic in many research fields [5–13]. Many models exist to study the spread of viruses. The first categorization between these models can be made for deterministic and stochastic models [14–17].

Deterministic models are the simplest, with fixed input variables. They are also known as compartmental models because the individuals in the population are assigned to different subgroups, or compartments, each of which represents a specific condition of the individual in the epidemic situation [18]. Derivatives in time are used to express the transition rates of individuals from a compartment to another. Thus, the model is constructed as a system of Ordinary Differential Equations (ODEs).

Stochastic models take into account variations in input variables and provide results in terms of probability. Unlike a deterministic one, a stochastic model allows random variations in one or more inputs over time. Therefore, an estimation of the probability distributions of the outcomes can be carried out. Specifically, the variables changing in time can be the exposure risk, recovery rate, and other disease dynamics. Being able to insert the variability of the input data, the stochastic models have a more complex structure than the deterministic ones but manage to be more adherent to reality [19]. A second categorization

between the models can be made by taking into account (or not taking into account) the vital dynamics. The vital dynamics represent the demography dynamics, in which the naturally occurring births and deaths are included [20].

In this paper, deterministic models with vital dynamics are studied. Precisely, the Susceptible-Infectious-Recovered (SIR), Susceptible-Exposed-Infectious-Recovered (SEIR), and Susceptible-Exposed-Infectious-Recovered-Susceptible (SEIR) models are considered, with the vaccination factor for one of those [21,22].

This work aims to estimate various epidemiological model parameters by using the newly developed framework called *Extreme Theory of Functional Connections* (X-TFC) [23], which merges the Physics-Informed Neural Networks (PINNs) method, introduced by Raissi et al. [24], and the Theory of Functional Connections (TFC), proposed by Mortari [25]. This method aims to solve forward problems and inverse problems (*data-driven parameters discovery*) involving DEs in different perturbation scenarios. A typical field where solving inverse problems is of interest is remote sensing [26–29]. For instance, in Reference [30], the authors combine radiative and heat transfer equations to create a set of parametric DEs. The solutions of this system of equations are compared with real data to retrieve the grain size and the thermal inertia of planetary regoliths, which are the parameters governing the physics of the problem. There are two main approaches to solving mathematical and physical inverse problems: deterministic and probabilistic. The deterministic approach tackles inverse problems using standard optimization techniques. According to these techniques, a set of optimal parameters is found, which minimizes the difference between simulated and real data. However, inverse problems are known to be ill-posed [31] and hence, it becomes challenging to determine the uncertainty in the retrieved quantities mainly due to the noise in the observed data and the uncertainty in the real values of the input parameter that are not tuned.

As stated in [32], inverse problems to parameters' estimation are in general ill-posed because the problem is non-unique due to the higher number of unknowns than data/measurements and the stability of the solution to noise in the data and modeling errors is generally not guaranteed. Standard optimization techniques consider the tuning quantities to be deterministic. Therefore, the inverse problems' outputs are fixed quantities. However, these quantities are affected by uncertainties that need to be estimated. The issue is that uncertainty quantification (usually done via regularization techniques) is not trivial to perform, and it can lead to poor results, mainly when the problem is ill-posed. Moreover, nonlinear or non-convex inverse problems have more local minimum solutions. Thus, more than one acceptable solution can be computed, and it becomes challenging to select the best one via the classical optimization framework [33].

To overcome this issue, the probabilistic approach can be used, in particular Bayesian inversion techniques. In the Bayesian inversion framework, the quantities to be estimated are considered random variables. Thus, the output of the inverse modeling is the probability distribution for each of those parameters. Therefore, with the probabilistic approach, the degree of uncertainty of the quantities' values to be retrieved is included in their probability distributions [26].

Nevertheless, in this work, we tackle the inverse problem for data-driven parameters' discovery of epidemiological models via a deterministic approach. We show that solving these problems via Physics-Informed Neural Network (PINN) methods, such as X-TFC, mitigates the ill-posedness of the inverse problems toward modeling errors and noisy data. This is possible because the physics of the problem, modeled via a DE, acts as a regulator during the search of the optimal parameters (i.e., the NN training). That is, the network training is carried in a data-physics-driven fashion.

This manuscript is organized as follows. In Section 2, PINNs are introduced with particular regard to the X-TFC framework. In Section 3, the application of the X-TFC for the data-driven discovery of the parameters governing a few of the most common epidemiological compartmental models is presented. In Section 3, the results are presented and discussed. Finally, the concluding remarks are given in the last section.

2. Physics-Informed Neural Network and Functional Interpolation

PINNs are machine learning methods that include physics into a data-driven functional representation of input–output pairing collections. As defined by Raissi et al. [24], the term PINN describes NNs that embed the physics as a regularization term in the loss function. For instance, suppose that one aims to do a regression of an experimental dataset employing an NN and that the collected data represents some physical events modeled via a set of DEs. In conventional regression, one would approximate the data utilizing a NN trained to minimize a Mean Squared Error (MSE) as a loss function. Nevertheless, there is no guarantee that the physics phenomena governing the dataset would not be violated. To mitigate this issue, PINNs are introduced to ensure that the physics, modeled via DEs, is added as a penalty to the loss function. This extra term serves as a regularizator that penalizes the training when the DE and its constraints (e.g., Boundary Conditions BCs, and eventually Initial Conditions ICs) are violated. Thus, one can guarantee that the physics underlying the process is not violated. This method is defined as a data-physics-driven solution of DEs. More specifically, from the physics perspective, this approach enables the training of NNs to learn the solution of DEs in a data-physics-driven fashion. This becomes critical if the DEs do not precisely model the physics of the problem, for example, when uncertain dynamical systems are considered and/or when perturbations are nonnegligible. Conversely, when the purpose is to retrieve parameters governing some physical phenomena modeled via DEs (e.g., single scattering albedo in the radiative transfer equation), one usually refers to data-physics-driven parameters discovery of DEs (i.e., inverse problems) [24]. When data is not available, and consequently the loss function contains the residual of the DEs and its constraints solely, PINNs learn the solutions of problems involving DEs only in a physics-driven fashion.

The major shortcoming of the standard PINNs, as presented by Raissi et al. [24], is that the DE constraints are not analytically satisfied, and consequently, they need to be concurrently learned with the DE solution within the domain. Hence, during the PINN training, we deal with competing objectives: learning the DE hidden solution within the domain and the DE hidden solution on the boundaries. This drives to unbalanced gradients during the network training via gradient-based techniques that prompt PINNs to struggle frequently to accurately learn the underlying DE solution [34]. Gradient-based optimization methods may get stuck in limit cycles or diverge if several competing objectives are present [35,36]. In [34], to surmount this issue, the authors proposed a learning rate annealing algorithm that uses gradient statistics to adaptive assign proper weights to different terms (e.g., DE residuals within the domain and DE residuals on the boundaries) in the PINNs loss function during the training. In this work, we propose to employ a different and more robust PINN model, the Extreme Theory of Functional Connections (X-TFC), that merges NNs and the Theory of Functional Connections (TFC) [23,37]. X-TFC exploits the constrained expressions (CEs) introduced within the TFC to satisfy the boundary constraints analytically.

TFC, elaborated by Mortari [25], is a mathematical framework for functional interpolation where functions are approximated using these CEs. A CE is a functional that is a sum of a free function and a functional that analytically satisfies the constraints despite the choice of the free function [25,38]. TFC has multiple applications. Primarily, TFC is used for the solution of DEs because the CEs eliminate the "curse of the equation constraints" [39–41]. Moreover, TFC has already been used to solve different classes of optimal control space guidance problems such as energy optimal landing on large and small planetary bodies [42,43], fuel optimal landing on large planetary bodies [44], energy optimal relative motion problems subject to Clohessy-Wiltshire dynamics [45], and classes of transport theory problems, such as radiative transfer [46] and rarefied-gas dynamics [47]. For tackling DEs, the standard (or Vanilla as defined in [48,49]) TFC method employs a linear combination of orthogonal polynomials, such as Legendre or Chebyshev polynomials [39,40], as a free function. However, using orthogonal polynomials as a free function makes the standard TFC framework suffer from the curse of dimensionality, particularly

when solving large-scale PDEs. To overcome this limitation, X-TFC employs a shallow NN trained via the Extreme Learning Machine (ELM) algorithm [50] to represent the free function.

Being a PINN method, X-TFC can solve forward and inverse problems involving parametric DEs with high precision and low computational time. The method for solving direct problems involving parametric DEs is introduced and presented by Schiassi et al. [23]. As previously stated, the focus of this work is to apply the X-TFC for data-driven parameters discovery of compartmental epidemiological models such as SIR, SEIR, and SEIRS. In the remainder of this section, we will explain how the X-TFC is applied to tackle these problems. Such models are systems of ODEs, where the constraints are on the initial values of these systems' solutions. That is, these problems are initial value problems (IVPs). Therefore, in this section, we will also present the step-by-step derivation of the constrained expressions for these problems.

2.1. Generality on Neural Networks

Neural Networks (NNs) are one of the key components of the X-TFC framework that will be used to tackle the problems considered in this work. Therefore, for the convenience of the reader, before diving into the detailed explanation on how X-TFC works, we will give some generalities about NNs.

NNs are powerful mathematical tools, inspired by the biological neurosystems of the human brain, originally developed as function approximators for machine learning applications [51–53].

NNs are made by artificial neurons and their mutual connections. The output of every neuron is a non-linear function of the weighted sum of its inputs. The neurons are typically arranged into layers, whose number defines the type of NN. NNs with only a single layer of neurons are known as single-layer or shallow NNs. NNs with more than one layer of neurons are called Deep NNs (DNNs). Neurons and layers are not necessarily all connected among them. When all neurons and layers are connected, we generally talk about fully connected NNs (shallow or deep depending on the number of layers). Every layer of a fully connected NN can be mathematically represented as follows

$$\boldsymbol{z} = \sigma(W\boldsymbol{x} + \boldsymbol{b}) \tag{1}$$

where W is the weight matrix, $\boldsymbol{x}$ and $\boldsymbol{z}$ are the input and output vectors, respectively, $\boldsymbol{b}$ is the bias vector, and $\sigma(\cdot)$ is the activation function, which can be either different or the same for every layer.

As previously stated, NNs were originally introduced as function approximators thanks to their ability in interpolation and fitting. An NN function approximator works in a supervised manner, where the training set $\mathcal{T} = \{(x_i, y_i)\}_{i=1,\ldots,N}$ consists in x_i input points and y_i output points (which can be affected or not by noise). First, a trial function $\hat{y}(x)$ is randomly initialized and the loss function can be defined as

$$\mathcal{L} = \sum_{i=1}^{N} |\hat{y}(x_i) - y_i|^2 \tag{2}$$

The training process consists in solving an optimization problem, where the loss function ($\mathcal{L}$) is the objective function that needs to be minimized, and the decision variables are the weights and biases of each layer. Usually, the training is performed via stochastic gradient based methods such as Adam Optimizer [54].

2.2. Extreme Theory of Functional Connections (X-TFC)

In this work, we will focus on systems of ODEs (SODEs) used to describe epidemiological compartmental models. In general, we can express ODEs, in their implicit form as,

$$\mathcal{N}[f; \lambda] + \varepsilon - \mathcal{R} = 0 \tag{3}$$

subject to constraints given by initial conditions (IC) and/or boundary conditions (BC). In Equation (3), $f = f(x; \lambda(x))$ is the unknown (or latent) solution, with $x \in \mathbb{D} \subseteq \mathbb{R}$, $\lambda = \lambda(x) \in \mathbb{L} \subseteq \mathbb{R}^m$ are the parameters governing the ODE (In general, even if it is not reported in the notation, f is a function of x, and it is parametrized by λ, that in general can be x dependent as well.), $\mathcal{N}[\cdot; \lambda]$ is a linear or non-linear operator acting on f and parametrized by λ, ε is the modeling error that is negligible when solving problems where the physics is exactly modeled by the underlying DE, and $\mathcal{R}$ is a known term that in general can be x dependent and parametrized by λ as well.

The first step in our PINN-TFC based framework is to approximate the latent solution f with a constrained expression, defined within the TFC [25],

$$f(x; \lambda) = f_{CE}(x, g(x); \lambda) = A(x; \lambda) + B(x, g(x); \lambda) \tag{4}$$

where $A(x; \lambda)$ analytically satisfies the DE constraints, and $B(x, g(x); \lambda)$ projects the free function $g(x)$, which is a real valued function, onto the space of functions that vanish at the constraints [37]. In the X-TFC method we chose the free function, $g(x)$, to be a shallow NN, trained via ELM algorithm [50]. That is,

$$g(x) = \sum_{j=1}^{L} \xi_j \sigma\left(w_j x + b_j\right) = \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_L \end{bmatrix}^{\mathsf{T}} \xi = \sigma^{\mathsf{T}} \xi \tag{5}$$

where L is the number of hidden neurons, $w_j \in \mathbb{R}$ is the input weights vector connecting the jth hidden neuron and the input nodes, $\xi_j \in \mathbb{R}$ with $j = 1, ..., L$ is the jth output weight connecting the jth hidden neuron and the output node, and b_j is the bias of the jth hidden neuron, $\sigma(\cdot)$ are activation functions, and $\sigma = [\sigma_1, \cdots, \sigma_L]^{\mathsf{T}}$. According to the ELM algorithm [50], biases and input weights are randomly selected and not tuned during the training, thus they are known hyperparameters. The activation functions, $\sigma(\cdot)$, are also known, as they are user selected. Thus, the only unknown NN hyperparameters to compute are the output weights $\xi = [\xi_1, \cdots, \xi_L]^{\mathsf{T}}$. Hence we can write,

$$f(x; \lambda) = f_{CE}(x, g(x); \lambda) = f_{CE}(x, \xi; \lambda).$$

The step-by-step process to derive the constrained expression is provided in Section 2.2.1. Once f is approximated with a NN, the second step of the X-TFC method is to define the loss functions,

$$\begin{aligned} \mathcal{L}_{\text{DATA}} &= f_{\text{DATA}} - f_{CE} \tag{6} \\ \mathcal{L}_{\text{DE}} &= \mathcal{N}[f_{CE}; \lambda] + \varepsilon - \mathcal{R} \tag{7} \end{aligned}$$

where f_{DATA} are the real data that eventually can be perturbed. Once the losses are defined, we need to define the vectors with all the unknowns, that are the ξ coefficients and the parameters governing the equations λ,

$$\Xi = \{\xi, \quad \lambda\}$$

Now, by combining the losses, an augmented loss function vector is formed as follows,

$$\mathbb{L} = \{\mathcal{L}_{\text{DATA}}, \quad \mathcal{L}_{\text{DE}}\}^{\mathsf{T}} \tag{8}$$

and enforcing that for a true solution, this vector should be equal to 0. This allows the unknowns to be solved via different optimization schemes, e.g., least-square for linear problems [39] and iterative least-squares for non-linear problems [40]. When solving

inverse problems for parameter estimation, the iterative least-square method is required. Thus, the estimation of the unknowns is updated at each iteration as follows,

$$\Xi_{k+1} = \Xi_k + \Delta\Xi_k \tag{9}$$

where the *k* subscript refers to the current iteration. In general, the $\Delta\Xi_k$ term can be defined by performing classic linear least-square at each iteration of the iterative least-square procedure as follows,

$$\Delta\Xi_k = -\left(\mathbb{J}^\mathrm{T}(\Xi_k)\mathbb{J}(\Xi_k)\right)^{-1}\mathbb{J}(\Xi_k)^\mathrm{T}\mathbb{L}(\Xi_k) \tag{10}$$

where $\mathbb{J}$ is the Jacobian matrix containing the derivatives of the losses with respect to all the unknowns. One can consider computing the Jacobian either by hand or by means of computing tools, such as Symbolic or Automatic Differentiation Toolboxes. The iterative process is repeated until either of the following conditions are met,

$$L_2[\mathbb{L}(\Xi_k)] < \epsilon \qquad \text{or} \qquad L_2[\mathbb{L}(\Xi_{k+1})] > L_2[\mathbb{L}(\Xi_k)]. \tag{11}$$

where ϵ defines some user prescribed tolerance.

In Figure 1, a schematic that summarizes how the X-TFC algorithm works for solving inverse problems is shown. The main steps are also reported here:

1. Approximate the latent solution(s) with the CE;
2. Analytically satisfy the ICs/BCs;
3. Expand with the single layer NN (trained via ELM);
4. Substitute into the DE (that can be also a system of DEs);
5. Build the DE losses (that drive the training of the network, informing it with the physics of the problem);
6. Build the data losses (the data can be provided on the solutions and/or on their derivatives);
7. Train the network;
8. Build the approximate solution (with the estimated optimal parameters).
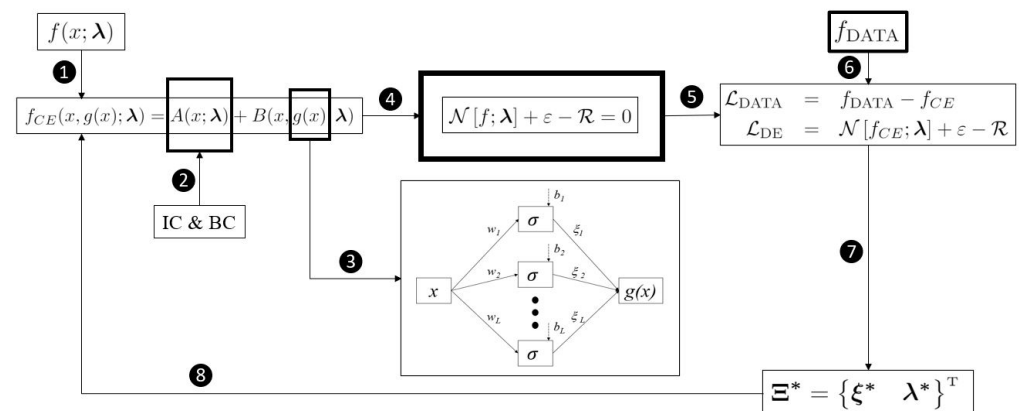
Figure 1. Schematic of the X-TFC framework for solving inverse problems.

2.2.1. Constrained Expression Derivation

Since this paper focuses on IVPs, for the convenience of the reader, we will present the step-by-step derivation of the constrained expression for these kinds of problems. The interested reader can find the general derivation for a $n + 1$ dimensional constrained expression either in [37] or [23]. Given a parametric ODE where we have a constraint on

the initial value of the solution (i.e., $f(0) = f_0$), the constrained expression for f is the following [25],

$$f_{CE}(x, g(x)) = g(x) + \eta = \sigma^{\mathrm{T}} \xi + \eta \tag{12}$$

By imposing the constraint f_0 into Equation (12) we get the following,

$$\eta = f_0 - g_0 = f_0 - \sigma_0^{\mathrm{T}} \xi \tag{13}$$

Now by plugging this results back into Equation (12) we get,

$$f_{CE} = \left[\sigma - \Omega_1 \sigma_0 \right]^{\mathrm{T}} \xi + \Omega_1 f_0, \tag{14}$$

where Ω_1 is called switching function. For an IVP with one constraint on f, we have $\Omega_1(z) = \Omega_1 = 1$.

In general f is defined in $x \in [x_0, x_f]$, which can be inconsistent with the domain where the activation function are defined. Thus we need to map it into the $z \in [z_0, z_f]$ domain as follows,

$$z = z_0 + c(x - x_0) \quad \leftrightarrow \quad x = x_0 + \frac{1}{c}(z - z_0), \tag{15}$$

where the mapping coefficient c is,

$$c = \frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\Delta z}{x_f - x_0} \tag{16}$$

According to the chain rule of the derivative we then have,

$$\frac{\mathrm{d}^n f}{\mathrm{d}x^n} = c^n \frac{\mathrm{d}^n f}{\mathrm{d}z^n} \tag{17}$$

3. Epidemiological Models Formulation

In this section, the X-TFC formulation for the data-driven parameters discovery of a series of epidemiological compartmental models is explained in detail. The presented models are the SIR [55], SEIR [56], and SEIRS [57], taking into account the vital dynamics and the vaccination (for the SEIR model) [58]. As already mentioned, the goal is to estimate the parameters of our interest through solving inverse problems via a deterministic approach.

Given fixed parameters, by integration, we solve the systems of ODEs to create a synthetic dataset (with and without noise), through which the parameters that govern the physics of the problem can be retrieved. After building the constrained expressions and the loss functions, the Jacobian matrix (the matrix containing the derivatives of the losses with respect to the unknowns) is computed in order to perform the iterative least-squares and estimate the unknowns.

3.1. SIR Model

As first problem, we consider the system of differential equations that govern the classic deterministic **SIR** (**S**usceptible-**I**nfectious-**R**ecovered) compartmental model, in which individuals in the recovered state gain total immunity to the pathogen, with vital dynamics to take into account the births (that can provide an increase in susceptible individuals) and natural death rates. The DEs governing the SIR model are the following,

$$\begin{cases} \dfrac{dS}{dt} = \mu N - \beta \dfrac{SI}{N} - \mu S \\ \dfrac{dI}{dt} = \beta \dfrac{SI}{N} - \gamma I - \mu I \\ \dfrac{dR}{dt} = \gamma I - \mu R \end{cases} \quad \text{subject to} \quad \begin{cases} S(t_0) = S_0 \\ I(t_0) = I_0 \\ R(t_0) = R_0 \end{cases} \tag{18}$$

where $N = S + I + R$ is the total population, μ is the birth and natural death rate (considered equal to maintain a constant population), β is the infectious rate, and γ is the recovery rate. An important parameter to consider is the basic reproduction number $\mathcal{R}_0$, which represents the ratio between β and γ. If $\mathcal{R}_0 > 1$, an outbreak is going to occur.

According to the TFC framework, the latent solutions are approximated with the constrained expressions. That is,

$$
\begin{cases}
S &= (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_1 + \Omega_1 S_0 \\
I &= (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_2 + \Omega_1 I_0 \\
R &= (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_3 + \Omega_1 R_0
\end{cases}
\tag{19}
$$

The first three loss functions we present take into account the regression over the data. The last three losses drive the training of the NN, informing it with the physics governing the problem. The Loss functions are reported below,

$$
\begin{cases}
\mathcal{L}_1 &= \tilde{S} - S \\
\mathcal{L}_2 &= \tilde{I} - I \\
\mathcal{L}_3 &= \tilde{R} - R \\
\mathcal{L}_4 &= \dot{S} - \mu N + \beta \dfrac{SI}{N} + \mu S \\
\mathcal{L}_5 &= \dot{I} - \beta \dfrac{SI}{N} + \gamma I + \mu I \\
\mathcal{L}_6 &= \dot{R} - \gamma I + \mu R
\end{cases}
\tag{20}
$$

To construct the Jacobian matrix $\mathcal{J}$ we need to compute the derivatives of the losses with respect to the $\boldsymbol{\xi}$ to compute the approximate solutions of the state variables, whereas the other derivatives are essential to estimate the parameters (in this case β and γ) appearing in the system of Equation (18). The resultant Jacobian matrix has the following form,

$$
\mathcal{J} =
\begin{bmatrix}
\dfrac{\partial \mathcal{L}_1}{\partial \boldsymbol{\xi}_1} & 0 & 0 & 0 & 0 \\
0 & \dfrac{\partial \mathcal{L}_2}{\partial \boldsymbol{\xi}_2} & 0 & 0 & 0 \\
0 & 0 & \dfrac{\partial \mathcal{L}_3}{\partial \boldsymbol{\xi}_3} & 0 & 0 \\
\dfrac{\partial \mathcal{L}_4}{\partial \boldsymbol{\xi}_1} & \dfrac{\partial \mathcal{L}_4}{\partial \boldsymbol{\xi}_2} & \dfrac{\partial \mathcal{L}_4}{\partial \boldsymbol{\xi}_3} & \dfrac{\partial \mathcal{L}_4}{\partial \beta} & 0 \\
\dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_1} & \dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_2} & \dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_3} & \dfrac{\partial \mathcal{L}_5}{\partial \beta} & \dfrac{\partial \mathcal{L}_5}{\partial \gamma} \\
0 & \dfrac{\partial \mathcal{L}_6}{\partial \boldsymbol{\xi}_2} & \dfrac{\partial \mathcal{L}_6}{\partial \boldsymbol{\xi}_3} & 0 & \dfrac{\partial \mathcal{L}_6}{\partial \gamma}
\end{bmatrix}
\tag{21}
$$

3.2. SEIR Model

The second problem that we aim to solve is the **SEIR** (**S**usceptible-**E**xposed-**I**nfectious-**R**ecovered) compartmental model. This model, compared to the previous one, takes into account the incubation period of a virus, i.e., the time in which a subject comes into contact with the virus but still does not develop its symptoms. Therefore, the subject is infected but is not yet considered among the infectious. In addition, a vaccination parameter which

moves people from the Susceptible to Recovered directly is added. The following is the ODEs system describing the model,

$$
\begin{cases}
\dfrac{dS}{dt} & = \mu(N - S) - \beta\dfrac{SI}{N} - \nu S \\
\dfrac{dE}{dt} & = \beta\dfrac{SI}{N} - (\mu + \phi)E \\
\dfrac{dI}{dt} & = \phi E - \gamma I - \mu I \\
\dfrac{dR}{dt} & = \gamma I - \mu R + \nu S
\end{cases}
\quad \text{subject to} \quad
\begin{cases}
S(t_0) & = S_0 \\
E(t_0) & = E_0 \\
I(t_0) & = I_0 \\
R(t_0) & = R_0
\end{cases}
\tag{22}
$$

where $N = S + E + I + R$ is the total population, μ is the birth and natural death rate (considered equal to maintain a constant population), ν is the vaccination rate, β is the infectious rate, ϕ is the rate at which an Exposed person becomes Infectious, and γ is the recovery rate.

According to the TFC framework, the latent solutions are approximated with the constrained expressions. That is,

$$
\begin{cases}
S & = (\sigma - \Omega_1\sigma_0)^{\mathsf{T}}\boldsymbol{\xi}_1 + \Omega_1 S_0 \\
E & = (\sigma - \Omega_1\sigma_0)^{\mathsf{T}}\boldsymbol{\xi}_2 + \Omega_1 E_0 \\
I & = (\sigma - \Omega_1\sigma_0)^{\mathsf{T}}\boldsymbol{\xi}_3 + \Omega_1 I_0 \\
R & = (\sigma - \Omega_1\sigma_0)^{\mathsf{T}}\boldsymbol{\xi}_4 + \Omega_1 R_0
\end{cases}
\tag{23}
$$

The first four loss functions we present take into account the regression over the data. The last four losses drive the NN, informing it with the physics governing the problem. The loss functions are reported below,

$$
\begin{cases}
\mathcal{L}_1 & = \tilde{S} - S \\
\mathcal{L}_2 & = \tilde{E} - E \\
\mathcal{L}_3 & = \tilde{I} - I \\
\mathcal{L}_4 & = \tilde{R} - R \\
\mathcal{L}_5 & = \dot{S} - \mu(N - S) + \beta\dfrac{SI}{N} + \nu S \\
\mathcal{L}_6 & = \dot{E} - \beta\dfrac{SI}{N} + (\mu + \phi)E \\
\mathcal{L}_7 & = \dot{I} - \phi E + (\gamma + \mu)I \\
\mathcal{L}_8 & = \dot{R} - \gamma I + \mu R - \nu S
\end{cases}
\tag{24}
$$

To construct the Jacobian matrix $\mathcal{J}$ we need to compute the derivatives of the losses in respect of the $\boldsymbol{\xi}$ to compute the approximate solutions of the state variables, whereas the other derivatives are essential to estimate the parameters (in this case β, γ, and ϕ) appearing in the system of Equation (18). The resultant Jacobian matrix has the following form,

$$
\mathcal{J} = \begin{bmatrix}
\dfrac{\partial \mathcal{L}_1}{\partial \boldsymbol{\xi}_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\[2mm]
\mathbf{0} & \dfrac{\partial \mathcal{L}_2}{\partial \boldsymbol{\xi}_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\[2mm]
\mathbf{0} & \mathbf{0} & \dfrac{\partial \mathcal{L}_3}{\partial \boldsymbol{\xi}_3} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\[2mm]
\mathbf{0} & \mathbf{0} & \mathbf{0} & \dfrac{\partial \mathcal{L}_4}{\partial \boldsymbol{\xi}_4} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\[2mm]
\dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_1} & \dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_2} & \dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_3} & \dfrac{\partial \mathcal{L}_5}{\partial \boldsymbol{\xi}_4} & \dfrac{\partial \mathcal{L}_5}{\partial \beta} & \mathbf{0} & \mathbf{0} \\[2mm]
\dfrac{\partial \mathcal{L}_6}{\partial \boldsymbol{\xi}_1} & \dfrac{\partial \mathcal{L}_6}{\partial \boldsymbol{\xi}_2} & \dfrac{\partial \mathcal{L}_6}{\partial \boldsymbol{\xi}_3} & \dfrac{\partial \mathcal{L}_6}{\partial \boldsymbol{\xi}_4} & \dfrac{\partial \mathcal{L}_6}{\partial \beta} & \mathbf{0} & \dfrac{\partial \mathcal{L}_6}{\partial \phi} \\[2mm]
\mathbf{0} & \dfrac{\partial \mathcal{L}_7}{\partial \boldsymbol{\xi}_2} & \dfrac{\partial \mathcal{L}_7}{\partial \boldsymbol{\xi}_3} & \mathbf{0} & \mathbf{0} & \dfrac{\partial \mathcal{L}_7}{\partial \gamma} & \dfrac{\partial \mathcal{L}_7}{\partial \phi} \\[2mm]
\dfrac{\partial \mathcal{L}_8}{\partial \boldsymbol{\xi}_1} & \mathbf{0} & \dfrac{\partial \mathcal{L}_8}{\partial \boldsymbol{\xi}_3} & \dfrac{\partial \mathcal{L}_8}{\partial \boldsymbol{\xi}_4} & \mathbf{0} & \dfrac{\partial \mathcal{L}_8}{\partial \gamma} & \mathbf{0}
\end{bmatrix}
\tag{25}
$$

3.3. SEIRS Model

The last problem we present here, is the **SEIRS** (**S**usceptible-**E**xposed-**I**nfectious-**R**ecovered-**S**usceptible) compartmental model. This model is used in the case when the immunity of recovered individuals wane, and they return to exist in the category of Susceptibles. No vaccination is considered here. This model is governed by the following system of ODEs:

$$
\begin{cases}
\dfrac{dS}{dt} & = \mu N - \beta \dfrac{SI}{N} + \zeta R - \nu S \\[2mm]
\dfrac{dE}{dt} & = \beta \dfrac{SI}{N} - \phi E - \nu E \\[2mm]
\dfrac{dI}{dt} & = \phi E - \gamma I - \mu I \\[2mm]
\dfrac{dR}{dt} & = \gamma I - \nu R - \zeta R
\end{cases}
\qquad \text{subject to} \qquad
\begin{cases}
S(t_0) & = S_0 \\
E(t_0) & = E_0 \\
I(t_0) & = I_0 \\
R(t_0) & = R_0
\end{cases}
\tag{26}
$$

where $N = S + E + I + R$ is the total population, μ is the natural deaths rate, ν is the new births rate, β is the infectious rate, ϕ is the rate at which an Exposed person becomes Infectious, ζ is the rate which Recovered individuals return to the Susceptible statue due to loss of immunity, and γ is the recovery rate.

According to the TFC framework, the latent solutions are approximated with the constrained expressions. That is,

$$
\begin{cases}
S & = (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_1 + \Omega_1 S_0 \\
E & = (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_2 + \Omega_1 E_0 \\
I & = (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_3 + \Omega_1 I_0 \\
R & = (\sigma - \Omega_1 \sigma_0)^{\mathsf{T}} \boldsymbol{\xi}_4 + \Omega_1 R_0
\end{cases}
\tag{27}
$$

The first four loss functions we present take into account the regression over the data. The last four losses drive the NN informing it with the physics governing the problem. The loss functions are reported below,

$$
\begin{cases}
\mathcal{L}_1 &= \tilde{S} - S \\
\mathcal{L}_2 &= \tilde{E} - E \\
\mathcal{L}_3 &= \tilde{I} - I \\
\mathcal{L}_4 &= \tilde{R} - R \\
\mathcal{L}_5 &= \dot{S} - \mu N + \beta \dfrac{SI}{N} - \zeta R + \nu S \\
\mathcal{L}_6 &= \dot{E} - \beta \dfrac{SI}{N} + (\phi + \nu) E \\
\mathcal{L}_7 &= \dot{I} - \phi E + (\gamma + \mu) I \\
\mathcal{L}_8 &= \dot{R} - \gamma I + (\nu + \zeta) R
\end{cases}
\tag{28}
$$

To construct the Jacobian matrix $\mathcal{J}$ we need to compute the derivatives of the losses in respect of the $\boldsymbol{\xi}$ to compute the approximate solutions of the state variables, whereas the other derivatives are essential to estimates the parameters (in this case β, γ, and ϕ) appearing in the system of Equation (18). The resultant Jacobian matrix has the following form:

$$
\mathcal{J} = \begin{bmatrix}
\dfrac{\partial \mathcal{L}_1}{\partial \xi_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \dfrac{\partial \mathcal{L}_2}{\partial \xi_2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \dfrac{\partial \mathcal{L}_3}{\partial \xi_3} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \dfrac{\partial \mathcal{L}_4}{\partial \xi_4} & 0 & 0 & 0 & 0 \\
\dfrac{\partial \mathcal{L}_5}{\partial \xi_1} & \dfrac{\partial \mathcal{L}_5}{\partial \xi_2} & \dfrac{\partial \mathcal{L}_5}{\partial \xi_3} & \dfrac{\partial \mathcal{L}_5}{\partial \xi_4} & \dfrac{\partial \mathcal{L}_5}{\partial \beta} & 0 & 0 & \dfrac{\partial \mathcal{L}_5}{\partial \zeta} \\
\dfrac{\partial \mathcal{L}_6}{\partial \xi_1} & \dfrac{\partial \mathcal{L}_6}{\partial \xi_2} & \dfrac{\partial \mathcal{L}_6}{\partial \xi_3} & \dfrac{\partial \mathcal{L}_6}{\partial \xi_4} & \dfrac{\partial \mathcal{L}_6}{\partial \beta} & 0 & \dfrac{\partial \mathcal{L}_6}{\partial \phi} & 0 \\
0 & \dfrac{\partial \mathcal{L}_7}{\partial \xi_2} & \dfrac{\partial \mathcal{L}_7}{\partial \xi_3} & 0 & 0 & \dfrac{\partial \mathcal{L}_7}{\partial \gamma} & \dfrac{\partial \mathcal{L}_7}{\partial \phi} & 0 \\
0 & 0 & \dfrac{\partial \mathcal{L}_8}{\partial \xi_3} & \dfrac{\partial \mathcal{L}_8}{\partial \xi_4} & 0 & \dfrac{\partial \mathcal{L}_8}{\partial \gamma} & 0 & \dfrac{\partial \mathcal{L}_9}{\partial \zeta}
\end{bmatrix}
\tag{29}
$$

4. Results and Discussion

To test the ability of the X-TFC in performing data-driven parameters discovery of epidemiological compartmental models, we have created synthetic datasets according to the three models presented above (SIR, SEIR, and SEIRS). In particular, for each model, a no-noisy synthetic dataset (here called original dataset $\tilde{f}_{orig}$) has been generated by simply propagating the dynamics equations of the model using the MatLab function ODE113. In addition, to simulate a more realistic example, perturbed synthetic datasets ($\tilde{f}_{pert}$) have been created by adding noise to the original dataset. That is,

$$
\tilde{f}_{pert} = \tilde{f}_{orig} + \delta \mathcal{U}(-1, -1)
\tag{30}
$$

where δ is the perturbation coefficient (equal to 0 for the original dataset) and $\mathcal{U}(\cdot, \cdot)$ represents a uniform distribution. The real values of the parameters governing the synthetic dataset are known, so that the accuracy of the results is measured by the absolute error between the real and estimated values of the parameters.

Additionally, the X-TFC method involves several hyperparameters that can be modified to obtain accurate solutions. These hyperparameters are the number of training points, n, the number of neurons, L, the type of activation function, and the probability distribution where input weights and bias are sampled from. Therefore, a sensitivity analysis

has been performed to study the behavior of the X-TFC method as these hyperparameters vary. The sensitivity analysis is only shown for the SIR model with no-noisy data, as a similar behavior has been encountered for all the other models considered. First of all, the sensitivity analysis has demonstrated that, for the models analyzed, the solution accuracy is not as sensitive to the type of activation function used or to the probability distribution used to sample the inputs weights and bias as it is to the number of training points and the number of neurons, confirming the results found from the sensitivity analysis reported in [23]. Hence, the two parameters that strongly influence the performances of the X-TFC are n and L. Figure 2a,b refer to the analysis with the original dataset ($\delta = 0$). As illustrated, high values of L ($L > 150$), with a fixed n, do not lead to an improvement of the accuracy, since Figure 2a presents an asymptotic-like behavior. The same considerations are valid varying n and keeping fixed L (Figure 2b). Indeed, the solution does not significantly improve increasing the number of discretization points. This result is also obtained if a perturbed dataset is considered (see Figure 2d). On the other hand, Figure 2c shows an interesting behavior. The accuracy of the solution gets worse by increasing the number of neurons L. This trend is probably due to the fact that X-TFC tries to overfit the perturbed data so that to diverge too much from the real curves and thus obtaining an inaccurate estimation of the parameters. The rest of this section focuses on the results obtained for each model presented previously. For these problems, the *ArcTan* activation function and a uniform random distribution ranging within [-10,10] are employed for the ELM.
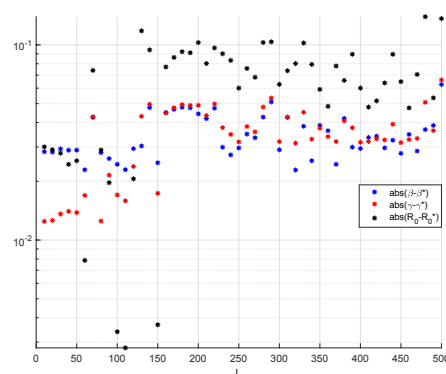
All the models tackled in this manuscript have been coded in MATLAB R2020a and ran with an Intel Core i7 - 9700 CPU PC with 64 GB of RAM.
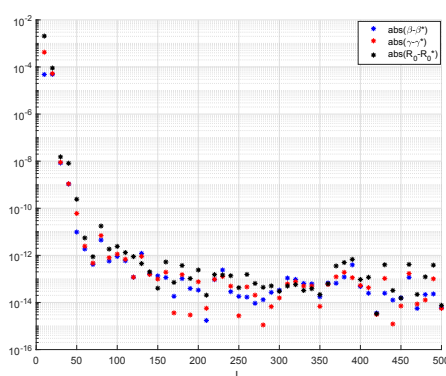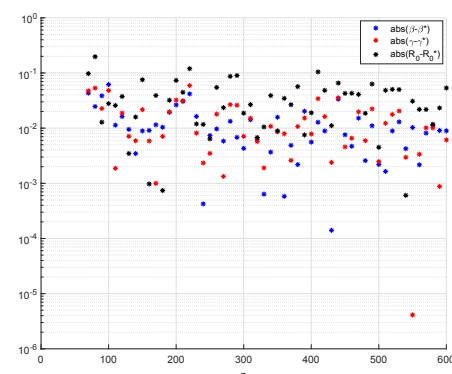

(**a**) $\delta = 0$, $n = 100$, and L varies.
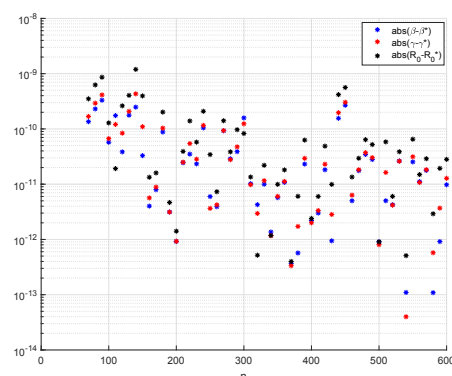
(**b**) $\delta = 0$, $L = 100$, and n varies.

(**c**) $\delta = 1$, $n = 100$, and L varies.

(**d**) $\delta = 1$, $L = 100$, and n varies.

Figure 2. Monte Carlo simulations for SIR model with an *ArcTan* activation function.

4.1. SIR Model

Here, the results and the performances for SIR problem are shown. The outputs are obtained by setting the following parameters:

- Natural mortality rate: $\mu = 0.1$ (set equal to the birth rate, to simulate a constant number of population);
- Effective contact rate (possibility to be infected): $\beta = \frac{1}{2}$;
- Removal rate (how often infected people become recovered): $\gamma = \frac{1}{3}$;
- Initial conditions: $S_0 = 100$; $I_0 = 5$; $R_0 = 0$;
- Analysis time: 15 days.

Several simulations are carried out by varying the intensity of the noise, and the outputs are reported in Table 1. While we could find the exact values of parameters with the original dataset, a slight deviation of these values occurs by increasing the perturbation coefficient δ. However, the absolute errors on the parameters result to have at least two digits of accuracy. Figure 3a,b report the perturbed and real dataset and the solution of the problem for the case of $\delta = 5$, respectively. As it can be seen, the X-TFC is able to obtain an accurate solution avoiding the overfitting on the data, as it could be expected by a simple regression on the perturbed dataset. This is due to the information about the physics of the problem, which acts as a regulator, that are embedded in the physics-informed training framework. The accuracy of the inversion with perturbed dataset is also proved by the constant value of the population N, as it has to be from the theory.

Table 1. Performances of the proposed physics-informed framework in the data-driven discovery of the SIR model with different noise on the data, with $n = 100$ and $L = 50$.

Noise	Iterations	CPU Time [s]	β	$|\mathrm{err}(\beta)|$	γ	$|\mathrm{err}(\gamma)|$	$\mathcal{R}_0$	$|\mathrm{err}(\mathcal{R}_0)|$
0	2	0.002	0.5000	0	0.3333	0	1.500	0
0.1	4	0.036	0.4999	4.2×10^{-5}	0.3334	4.2×10^{-5}	1.4997	3.2×10^{-4}
1	7	0.049	0.4996	4.4×10^{-4}	0.3338	4.2×10^{-4}	1.4968	3.2×10^{-3}
5	8	0.051	0.4978	2.2×10^{-3}	0.3353	2.1×10^{-3}	1.4884	1.56×10^{-2}

(**a**) Real and perturbed data.　　　　　(**b**) Results with perturbed data.
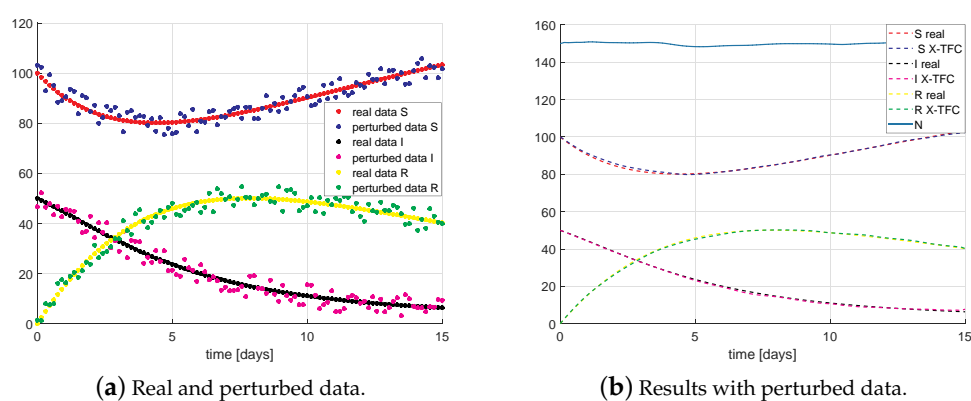
Figure 3. Results for SIR model with $\delta = 5$, $n = 100$, and $L = 50$.

4.2. SEIR Model

Here, the results and the performances for SEIR problem are shown. The outputs were obtained by setting the following parameters:

- Natural mortality rate $\mu = 0.5$ (set equal to the birth rate, to simulate a constant number of population);
- Vaccine rate $\nu = 0.5$;
- Effective contact rate (possibility to be infected) $\beta = 0.3$;
- Removal rate (how often infected people become recovered) $\gamma = 0.6$;
- Progression rate from exposed to infected $\phi = 0.9$;

- Initial conditions: $S_0 = 70$; $E_0 = 30$; $I_0 = 10$; $R_0 = 0$;
- Days = 15.

Several simulations are carried out by varying the intensity of the noise, and the outputs are reported in Table 2. While we could find the exact values of parameters with the original dataset, a slight deviation of these values occurs by increasing the perturbation coefficient δ. However, the absolute errors on the parameters result to have at least one digit of accuracy. Figure 4a,b report the perturbed and real dataset and the solution of the problem for the case of $\delta = 3$, respectively. Again, the X-TFC is able to obtain an accurate solution avoiding the overfitting on the data, as it could be expected by a simple regression on the perturbed dataset.

Table 2. Performances of the proposed physics-informed framework in the data-driven discovery of the SEIR model with different noise on the data, with $n = 100$ and $L = 80$, $\mu = \nu = 0.5$ (the training time is on the order of milliseconds).

Noise	β	$|\text{err}(\beta)|$	γ	$|\text{err}(\gamma)|$	ϕ	$|\text{err}(\phi)|$	$\mathcal{R}_0$	$|\text{err}(\mathcal{R}_0)|$
0	0.3	0	0.6	0	0.9	0	0.5	0
0.1	0.2971	2.9×10^{-3}	0.5996	3.7×10^{-4}	0.9005	4.9×10^{-4}	0.4955	4.5×10^{-3}
1	0.2711	2.9×10^{-2}	0.5962	3.8×10^{-3}	0.9048	4.8×10^{-3}	0.4547	4.5×10^{-2}
3	0.2130	8.7×10^{-2}	0.5878	1.2×10^{-2}	0.9141	1.4×10^{-2}	0.3624	1.4×10^{-1}

(**a**) Real and perturbed data.

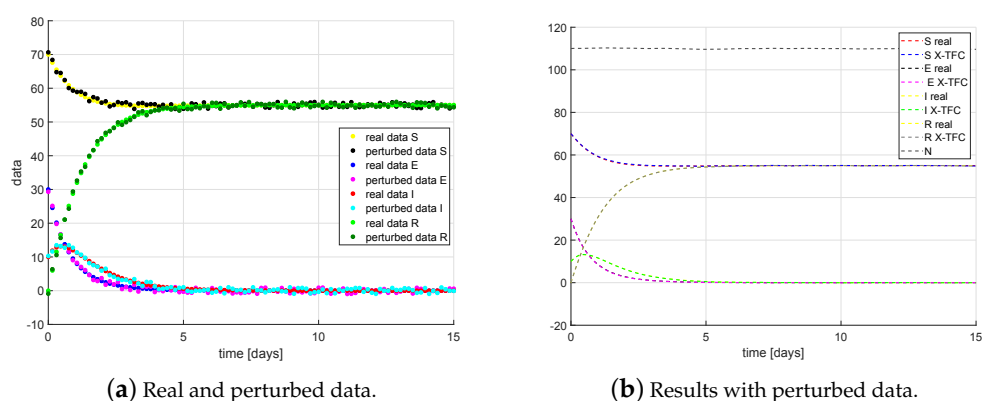(**b**) Results with perturbed data.

Figure 4. Results for SEIR model with $\delta = 3$, $n = 100$, and $L = 80$.

4.3. SEIRS Model

Here, the results and the performances for SEIRS problem are shown. The outputs were obtained by setting the following parameters:

- Natural mortality rate $\mu = 0.5$ (set equal to the birth rate, to simulate a constant number of population);
- Effective contact rate (possibility to be infected) $\beta = 0.3$;
- Removal rate (how often infected people become recovered) $\gamma = 0.6$;
- Progression rate from exposed to infected $\phi = 0.9$;
- Rate which recovered individuals return to the susceptible statue (due to loss of immunity) $\zeta = 0.5$;
- initial conditions: $S_0 = 70$; $E_0 = 30$; $I_0 = 10$; $R_0 = 0$;
- days = 15.

Several simulations are carried out by varying the intensity of the noise, and the outputs are reported in Table 3. While we could find the exact values of parameters with the original dataset, a slight deviation of these values occurs by increasing the perturbation coefficient δ. However, the absolute errors on the parameters result in having at least two digits of accuracy. Figure 5a,b report the perturbed and real dataset and the solution of the problem for the case of $\delta = 3$, respectively. Again, the X-TFC is able to obtain an accurate solution avoiding the overfitting on the data, as it could be expected by a simple regression on the perturbed dataset.

Table 3. Performances of the proposed physics-informed framework in the data-driven discovery of the SEIRS model with different noise on the data, with $n = 100$ and $L = 80$, $\mu = 0.5$ (the training time is on the order of milliseconds).

Noise	β	$|\text{err}(\beta)|$	γ	$|\text{err}(\gamma)|$	ϕ	$|\text{err}(\phi)|$	ζ	$|\text{err}(\zeta)|$	$\mathcal{R}_0$	$|\text{err}(\mathcal{R}_0)|$
0	0.3	0	0.6	0	0.9	0	0.5	0	0.5	0
0.1	0.3011	1.1×10^{-3}	0.6020	2.0×10^{-3}	0.9026	2.6×10^{-3}	0.5028	2.8×10^{-3}	0.500	1.7×10^{-4}
1	0.3093	9.3×10^{-3}	0.6183	1.8×10^{-2}	0.9249	2.5×10^{-2}	0.5251	2.5×10^{-2}	0.5002	2.1×10^{-4}
3	0.3174	1.7×10^{-2}	0.6465	4.7×10^{-2}	0.9680	6.8×10^{-2}	0.5576	5.8×10^{-2}	0.4909	9.1×10^{-3}

(**a**) Real and perturbed data.
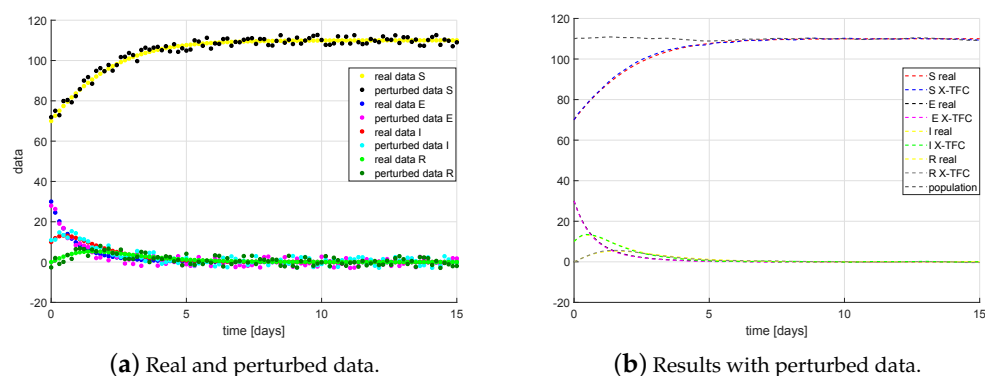

(**b**) Results with perturbed data.

Figure 5. Results for SEIRS model with $\delta = 3$, $n = 100$, and $L = 80$.

5. Conclusions

In this work, the new PINN framework X-TFC has been employed to solve data driven discovery of DEs, also called inverse problems, via a deterministic approach. In particular, compartmental epidemiological models (SIR, SEIR, SEIRS) have been taken into account as test problems. The goal was to retrieve the parameters governing the dynamics equations considering unperturbed and perturbed data, to better simulate the reality. The tests have shown fairly accurate results even when a significant noise was added to the data. Furthermore, the information about the physics of the problem (considered for the training of the X-TFC) has allowed to avoid the overfitting and thus to obtain good estimations of parameters with noisy data. The low computational times obtained are extremely important to process data as soon as they are acquired, so that the results can be updated in real time. Moreover, the good estimations of parameters allow one to make predictions about the imminent future: this makes it possible to take actions in the short term (as it should be in emergency scenarios, like the COVID-19 pandemic). Future works involve the inversion of models with non-constant parameters (i.e., parameters that follow mathematical laws) as well as probabilistic parameters estimation (via Bayesian inversion) in different research fields, such as business, biology, space, and nuclear engineering.

Author Contributions: Conceptualization: E.S., A.D., M.D.F.; methodology: A.D., M.D.F. and E.S.; software: A.D., M.D.F. and E.S.; validation: A.D., E.S. and M.D.F.; formal analysis: A.D., E.S. and M.D.F.; investigation: A.D., M.D.F. and E.S.; resources: M.D.F., A.D. and E.S.; writing—original draft preparation: M.D.F., A.D. and E.S.; writing—review and editing: D.M. and R.F.; visualization: A.D., E.S. and M.D.F.; supervision: D.M. and R.F. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. Sharp, P.M.; Bailes, E.; Robertson, D.L.; Gao, F.; Hahn, B.H. Origins and evolution of AIDS viruses. *Biol. Bull.* **1999**, *196*, 338–342. [CrossRef]
2. Lowen, A.C.; Mubareka, S.; Tumpey, T.M.; García-Sastre, A.; Palese, P. The guinea pig as a transmission model for human influenza viruses. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9988–9992. [CrossRef]
3. Geoghegan, J.L.; Senior, A.M.; Di Giallonardo, F.; Holmes, E.C. Virological factors that increase the transmissibility of emerging human viruses. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 4170–4175. [CrossRef]
4. Nelson, M.I.; Gramer, M.R.; Vincent, A.L.; Holmes, E.C. Global transmission of influenza viruses from humans to swine. *J. Gen. Virol.* **2012**, *93*, 2195. [CrossRef] [PubMed]
5. Kıymet, E.; Böncüoğlu, E.; Şahinkaya, Ş.; Cem, E.; Çelebi, M.Y.; Düzgöl, M.; Kara, A.A.; Arıkan, K.Ö.; Aydın, T.; İşgüder, R.; et al. Distribution of spreading viruses during COVID-19 pandemic: Effect of mitigation strategies. *Am. J. Infect. Control* **2021**. [CrossRef] [PubMed]
6. Galbadage, T.; Peterson, B.M.; Gunasekera, R.S. Does COVID-19 spread through droplets alone? *Front. Public Health* **2020**, *8*, 163. [CrossRef] [PubMed]
7. Arti, M.; Bhatnagar, K. Modeling and predictions for COVID 19 spread in India. *ResearchGate* **2020**. [CrossRef]
8. Castro, M.C.; Kim, S.; Barberia, L.; Ribeiro, A.F.; Gurzenda, S.; Ribeiro, K.B.; Abbott, E.; Blossom, J.; Rache, B.; Singer, B.H. Spatiotemporal pattern of COVID-19 spread in Brazil. *Science* **2021**, *372*, 821–826. [CrossRef] [PubMed]
9. Varotsos, C.A.; Krapivin, V.F. A new model for the spread of COVID-19 and the improvement of safety. *Saf. Sci.* **2020**, *132*, 104962. [CrossRef] [PubMed]
10. Caspi, G.; Shalit, U.; Kristensen, S.L.; Aronson, D.; Caspi, L.; Rossenberg, O.; Shina, A.; Caspi, O. Climate effect on COVID-19 spread rate: An online surveillance tool. *MedRxiv* **2020**. [CrossRef]
11. Aabed, K.; Lashin, M.M. An analytical study of the factors that influence COVID-19 spread. *Saudi J. Biol. Sci.* **2021**, *28*, 1177–1195. [CrossRef]
12. Piccolomiini, E.L.; Zama, F. Monitoring Italian COVID-19 spread by an adaptive SEIRD model. *MedRxiv* **2020**. [CrossRef]
13. Al-Kindi, K.M.; Alkharusi, A.; Alshukaili, D.; Al Nasiri, N.; Al-Awadhi, T.; Charabi, Y.; El Kenawy, A.M. Spatiotemporal assessment of COVID-19 spread over Oman using GIS techniques. *Earth Syst. Environ.* **2020**, *4*, 797–811. [CrossRef]
14. Adak, D.; Majumder, A.; Bairagi, N. Mathematical perspective of COVID-19 pandemic: Disease extinction criteria in deterministic and stochastic models. *Chaos Solitons Fractals* **2021**, *142*, 110381. [CrossRef] [PubMed]
15. Petrovskii, S.V.; Malchow, H.; Hilker, F.M.; Venturino, E. Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control. *Biol. Invasions* **2005**, *7*, 771–793. [CrossRef]
16. Perera, N.C. Deterministic and Stochastic Models of Virus Dynamics. Ph.D. Thesis, Texas Tech University, Lubbock, TX, USA, 2003.
17. Sazonov, I.; Grebennikov, D.; Kelbert, M.; Bocharov, G. Modelling stochastic and deterministic behaviours in virus infection dynamics. *Math. Model. Nat. Phenom.* **2017**, *12*, 63–77. [CrossRef]
18. Breda, D.; Diekmann, O.; De Graaf, W.; Pugliese, A.; Vermiglio, R. On the formulation of epidemic models (an appraisal of Kermack and McKendrick). *J. Biol. Dyn.* **2012**, *6*, 103–117. [CrossRef]
19. Britton, T. Stochastic epidemic models: A survey. *Math. Biosci.* **2010**, *225*, 24–35. [CrossRef]
20. Hethcote, H.W. Three basic epidemiological models. In *Applied Mathematical Ecology*; Springer: Berlin/Heidelberg, Germany, 1989; pp. 119–144.
21. Huang, G.; Takeuchi, Y.; Ma, W.; Wei, D. Global stability for delay SIR and SEIR epidemic models with nonlinear incidence rate. *Bull. Math. Biol.* **2010**, *72*, 1192–1207. [CrossRef] [PubMed]
22. Trawicki, M.B. Deterministic SEIRs epidemic model for modeling vital dynamics, vaccinations, and temporary immunity. *Mathematics* **2017**, *5*, 7. [CrossRef]
23. Schiassi, E.; Furfaro, R.; Leake, C.; De Florio, M.; Johnston, H.; Mortari, D. Extreme Theory of Functional Connections: A Fast Physics-Informed Neural Network Method for Solving Ordinary and Partial Differential Equations. *Neurocomputing* **2021**, *457*, 334–356. [CrossRef]
24. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
25. Mortari, D. The Theory of Connections: Connecting Points. *Mathematics* **2017**, *5*, 57. [CrossRef]
26. Schiassi, E.; Furfaro, R.; Kargel, J.S.; Watson, C.S.; Shugar, D.H.; Haritashya, U.K. GLAM Bio-Lith RT: A Tool for Remote Sensing Reflectance Simulation and Water Components Concentration Retrieval in Glacial Lakes. *Front. Earth Sci.* **2019**, *7*, 267. [CrossRef]
27. Schiassi, E.; Furfaro, R.; Mostacci, D. Bayesian inversion of coupled radiative and heat transfer models for asteroid regoliths and lakes. *Radiat. Eff. Defects Solids* **2016**, *171*, 736–745. [CrossRef]
28. Hapke, B. Bidirectional reflectance spectroscopy: 1. Theory. *J. Geophys. Res. Solid Earth* **1981**, *86*, 3039–3054. [CrossRef]
29. Hapke, B. A model of radiative and conductive energy transfer in planetary regoliths. *J. Geophys. Res. Planets* **1996**, *101*, 16817–16831. [CrossRef]
30. Hale, A.S.; Hapke, B. A time-dependent model of radiative and conductive thermal energy transport in planetary regoliths with applications to the Moon and Mercury. *Icarus* **2002**, *156*, 318–334. [CrossRef]

31. Kimes, D.S.; Knyazikhin, Y.; Privette, J.; Abuelgasim, A.; Gao, F. Inversion methods for physically-based models. *Remote Sens. Rev.* **2000**, *18*, 381–439. [CrossRef]

32. Kolehmainen, V. *Introduction to Bayesian Methods in Inverse Problems*; Department of Applied Physics, University of Eastern Finland: Kuopio, Finland, 2013.

33. Aster, R.C.; Borchers, B.; Thurber, C.H. *Parameter Estimation and Inverse Problems*; Elsevier: Amsterdam, The Netherlands, 2018.

34. Wang, S.; Teng, Y.; Perdikaris, P. Understanding and mitigating gradient pathologies in physics-informed neural networks. *arXiv* **2020**, arXiv:2001.04536.

35. Mertikopoulos, P.; Papadimitriou, C.; Piliouras, G. Cycles in adversarial regularized learning. In Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, New Orleans, LA, USA, 7–10 January 2018; pp. 2703–2717.

36. Balduzzi, D.; Racaniere, S.; Martens, J.; Foerster, J.; Tuyls, K.; Graepel, T. The mechanics of n-player differentiable games. In Proceedings of the International Conference on Machine Learning, PMLR, Stockholm, Sweden, 10–15 July 2018; pp. 354–363.

37. Leake, C.; Mortari, D. Deep theory of functional connections: A new method for estimating the solutions of partial differential equations. *Mach. Learn. Knowl. Extr.* **2020**, *2*, 37–55. [CrossRef]

38. Mortari, D.; Leake, C. The Multivariate Theory of Connections. *Mathematics* **2019**, *7*, 296. [CrossRef] [PubMed]

39. Mortari, D. Least-squares solution of linear differential equations. *Mathematics* **2017**, *5*, 48. [CrossRef]

40. Mortari, D.; Johnston, H.; Smith, L. High accuracy least-squares solutions of nonlinear differential equations. *J. Comput. Appl. Math.* **2019**, *352*, 293–307. [CrossRef]

41. Leake, C.; Johnston, H.; Mortari, D. The Multivariate Theory of Functional Connections: Theory, Proofs, and Application in Partial Differential Equations. *Mathematics* **2020**, *8*, 1303. [CrossRef]

42. Furfaro, R.; Mortari, D. Least-squares solution of a class of optimal space guidance problems via Theory of Connections. *Acta Astronaut.* **2019**. [CrossRef]

43. Schiassi, E.; D'Ambrosio, A.; Johnston, H.; Furfaro, R.; Curti, F.; Mortari, D. Complete Energy Optimal Landing on Small and Large Planetary Bodies via Theory Of Functional Connections. In Proceedings of the Astrodynamics Specialist Conference, AAS, South Lake Tahoe, CA, USA, 9–12 August 2020.

44. Johnston, H.; Schiassi, E.; Furfaro, R.; Mortari, D. Fuel-Efficient Powered Descent Guidance on Large Planetary Bodies via Theory of Functional Connections. *J. Astronaut. Sci.* under review. [CrossRef] [PubMed]

45. Drozd, K.; Furfaro, R.; Schiassi, E.; Johnston, H.; Mortari, D. Energy-optimal trajectory problems in relative motion solved via Theory of Functional Connections. *Acta Astronaut.* **2021**, *182*, 361–382. [CrossRef]

46. De Florio, M.; Schiassi, E.; Furfaro, R.; Ganapol, B.D.; Mostacci, D. Solutions of Chandrasekhar's basic problem in radiative transfer via theory of functional connections. *J. Quant. Spectrosc. Radiat. Transf.* **2021**, *259*, 107384. [CrossRef]

47. De Florio, M.; Schiassi, E.; Ganapol, B.; Furfaro, R. Physics-Informed Neural Networks for Rarefied-Gas Dynamics: Poiseuille Flow in the Bhatnagar–Gross–Krook approximation. *Phys. Fluids* **2021**, *33*, 047110. [CrossRef]

48. Johnston, H. The Theory of Functional Connections: A journey from theory to application. *arXiv* **2021**, arXiv:2105.08034.

49. Leake, C. The Multivariate Theory of Functional Connections: An n-Dimensional Constraint Embedding Technique Applied to Partial Differential Equations. *arXiv* **2021**, arXiv:2105.07070.

50. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]

51. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]

52. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386. [CrossRef]

53. Rosenblatt, F. *Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms*; Technical Report; Cornell Aeronautical Lab Inc.: Buffalo, NY, USA, 1961.

54. Zhang, Z. Improved adam optimizer for deep neural networks. In Proceedings of the 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), Banff, AB, Canada, 4–6 June 2018; pp. 1–2.

55. Brauer, F.; Driessche, P.; Wu, J. *Lecture Notes in Mathematical Epidemiology*; Springer: Berlin, Germany, 2008; Volume 75, pp. 3–22.

56. Röst, G. SEIR epidemiological model with varying infectivity and infinite delay. *Math. Biosci. Eng.* **2008**, *5*, 389–402.

57. Nakata, Y.; Kuniya, T. Global dynamics of a class of SEIRS epidemic models in a periodic environment. *J. Math. Anal. Appl.* **2010**, *363*, 230–237. [CrossRef]

58. Martcheva, M. *An Introduction to Mathematical Epidemiology*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 61.