# Approaching epidemiological dynamics of COVID-19 with physics-informed neural networks

Shuai Han[a,b,c,e], Lukas Stelz[a,c], Horst Stoecker[a,c,d], Lingxiao Wang[a,b,*], Kai Zhou[a,b,*]

[a]*Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany*

[b]*Xidian-FIAS International Joint Research Center, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany*

[c]*Institut für Theoretische Physik, Goethe Universität Frankfurt, Max-von-Laue-Str. 1, 60438 Frankfurt am Main, Germany*

[d]*GSI Helmholtzzentrum für Schwerionenforschung GmbH, 64291 Darmstadt, Germany*

[e]*Xidian University, 710071 Xi'an, China*

## Abstract

A physics-informed neural network (PINN) embedded with the susceptible-infected-removed (SIR) model is devised to understand the temporal evolution dynamics of infectious diseases. Firstly, the effectiveness of this approach is demonstrated on synthetic data as generated from the numerical solution of the susceptible-asymptomatic-infected-recovered-dead (SAIRD) model. Then, the method is applied to COVID-19 data reported for Germany and shows that it can accurately identify and predict virus spread trends. The results indicate that an incomplete physics-informed model can approach more complicated dynamics efficiently. Thus, the present work demonstrates the high potential of using machine learning methods, e.g., PINNs, to study and predict epidemic dynamics in combination with compartmental models.

*Keywords:* COVID-19, Epidemiological Dynamics, Physics-informed machine learning.

## 1. Introduction

The coronavirus SARS-CoV-2 was discovered in Wuhan, China, in December 2019. The virus spread quickly around the world. It was declared a pandemic by the World Health Organization (WHO)[1] in March 2020. By January 2023, there had been 733 million confirmed cases of the resulting disease from COVID-19 and 6.69 million fatalities[1]. The spread of the virus and the impact of policy decisions on containing the disease has been studied in compartmental models[2, 3]. Non-pharmaceutical interventions, such as social distancing, were found to be effective[4, 5, 6, 7]. The role of vaccination has been explored in recent studies[8].

The unknown numbers of infectious individuals have been a major challenge for obtaining precise real-time data on the spatiotemporal spread of COVID-19[9]. Based on ambigu-

---

*Corresponding author

ous reports and predicted cases[10, 11], governments had difficulties in implementing effective intervention policies, such as the allocation of detection resources[12], the mobilization and delivery of protective and therapeutic materials[13], and the stringency of lockdown measures[14, 15, 16, 17]. Also the early shortage of detection supplies and the massive number of asymptomatic or mildly symptomatic cases also make forecasting difficult[18]. The alternative approach to the tracking of the spread of any infectious diseases is an estimate of parameters of epidemic models, such as the basic reproduction rate ($R_0$) and infection rate[19]. Studies presently explore various data-driven methods to infer those parameters from the available but limited data[20].

Epidemiological models, such as the Susceptible-Infected-Recovered (SIR) model[21], have been helpful in understanding the spread of infectious diseases. The SIR model is one of the first used compartmental models. SIR divides the population into the three SIR compartments. Various models have been derived and developed, based on the SIR model, including the Susceptible-Exposed-Infected-Removed (SEIR)[22, 23] and the Susceptible-Exposed-Asymptomatic-Infectious-Recovered (SEAIR) model[24]. Mathematical modeling based on numerical solutions of systems of Ordinary Differential Equations (ODEs) has also been used to study COVID-19 spread. These studies have provided valuable insights into the spread of the disease, including, more recently, disease prevalence curves with machine learning assistance[25, 26, 27], the impact of asymptomatic infected individuals[28], the effectiveness of wearing masks[29], and the effectiveness of prevention and control measures[30, 31, 32].

Deep Learning (DL) models, such as recurrent neural networks (RNNs), have been used to analyze the patterns in COVID-19 time series data[33, 34]. Chimmula et al.[35] used RNNs and its variant long short-term memory (LSTM) for predicting COVID-19 prevalence trends in Canada and Italy, which show reasonable predictive capabilities. Zeroual et al.[36] applied LSTM, bi-directional LSTM (BiLSTM), and gated recurrent units (GRUs) to different countries' COVID-19 data for data-driven simulations. Zhang et al.[37] used a residual neural network (ResNet) to account for external factors for model uncertainties, parameters, and for other factors which affect prediction accuracy for trend analysis of COVID-19. Chen et al.[38] showed that generalized ResNet can learn the structure of complex unknown dynamical systems. These predictions are more accurate than standard ResNet structures. However, these models require large sets of training data, while the current datasets for COVID-19 are relatively small. This leads to a lack of robustness of the models[39]. In addition, these DL models are able only to identify the dynamics of the virus based on available data; they might not be stable or accurate enough in predicting future trends[40] due to the variability of the virus and the influence of external factors like weather[41]. Frameworks that can accurately tackle the epidemic dynamics, which are governed by systems of ordinary differential equations (ODEs) or systems of partial differential equations (PDEs)[42], should be developed to effectively handle the predicament induced by the limits of recorded data. It is crucial to incorporate the details of the necessarily known full domain knowledge, e.g., the laws governing the physical system, rigorously into the machine learning treatment used[43]. Physics-Informed Neural Networks (PINNs)[44] were proposed to address this need. PINNs can introduce physical constraints into the training explicitly. The PINNs can make accurate predictions based on tiny datasets when Combined with epidemiological models[42]. They may even help to identify the underlying epidemiological dynamics[45].

The present study introduces a SIR-dynamics-informed machine learning method to explore the approach towards complicated epidemiological dynamics, as often hidden in the study of the spread of COVID-19, by incorporating prior knowledge, in the form of the ODEs, from the SIR model into the loss functions of Deep Neural Networks (DNNs) as physical regularizer. This new method is first confronted with generated synthetic data of the ODE system, by using a SAIRD model to simulate different scenarios and test the effectiveness of the new approach. The method is then validated by reported COVID-19 data from Germany between March 1, 2021 and July 1, 2021. The data from the COVID-19 Data Repository at the Johns Hopkins University are used. Such simple, SIR-physics informed models, do can apparently approach more complicated dynamics efficiently.

This paper is organized into three sections: The methodology section details the general concepts of the SIR and SAIRD models and gives the definition of the parameters. The basics of the PINN framework, and how PINNs can be used to solve systems of ODEs and their optimization problems, is introduced in the second section. Here, the SIR-dynamics-informed neural networks, the definition and calculation of loss functions and physical residuals, which will be applied to estimate the dynamics of an infectious system are detailed. The third section contains the numerical synthetical data experiments and results for real, recorded data. It also covers the set-ups of the neural network, the pre-processing of synthetic and reported data, and the analysis of the results. The conclusions summarize the findings and yield recommendations for future research.

## 2. Methodology

This section introduces two mathematical epidemiological models used the present paper. Furthermore, physics-informed neural networks are described in conjunction with the epidemiological models and their optimization approaches.

### 2.1. Mathematical Epidemiology

Mathematical modelling of epidemiological dynamics is a popular area of research in applied mathematics. Conventionally, compartmental models are developed for modelling the dynamics of epidemics within a population.

### 2.1.1. The SIR Model

A plain, but yet powerful and well-known compartmental model is the original SIR model[18]. It assumes that the size of the population $N$ remains constant. $N$ is divided into three separate groups or compartments: susceptible (S), infectious (I) and removed (R). Individuals are transferred between compartments as shown in Figure 1 with certain rates, called transition rates, $\beta I/N$ and $\gamma$:
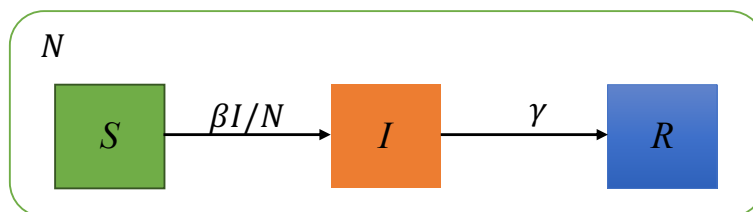


Figure 1: Schematic illustration of the interactions between the compartments in the SIR model.

Epidemiological compartmental models treat all individuals in the same compartment as sharing identical features. Therefore each compartment is homogeneous. The SIR model is described by the following set of differential equations:

$$\frac{dS}{dt} = -\frac{\beta I}{N}S \ ,$$
$$\frac{dI}{dt} = \frac{\beta I}{N}S - \gamma I \ ,$$
$$\frac{dR}{dt} = \gamma I \ . \tag{1}$$

The parameter $\beta$ is the effective contact or transmission rate. It denotes the number of effective contacts made by one infectious and one susceptible individual leading to one infection per unit of time. The removal rate $\gamma$ indicates the fraction of infectious individuals who recover or die per unit of time. Thus, $\gamma$ can be calculated using $1/D$, with D the average time duration that an infected individual can carry and transmit the virus. Equation (1) is subject to the initial conditions, $S(t_0) > 0, I(t_0) \geq 0,$ and $R(t_0) \geq 0$ at the initial time $t_0$. The model assumes to conserve the total number of individuals, thus $S(t) + I(t) + R(t) = N$ holds at any time $t$. In general, the time scale of the epidemic dynamics is assumed to be short as compared to the length of the life of individuals in the population: the effects of births and deaths on the population are simply ignored.

*2.1.2. The SAIRD Model*

The SAIRD model is an extension of the SIR model: two more compartments as shown in Figure 2 are introduced. The compartment $I$ is split into two. Here compartment $A$ represents the asymptomatic or unidentified, but infectious individuals, i.e., A is the number of individuals who, despite being infected, are either not identified or not detected. A consists of infected individuals, which are not from symptoms. The new compartment $I$ here contains those individuals which have been detected as infectious. The compartment $R$ is also split into the number of recovered, $R$, and the number of deceased, $D$, individuals.
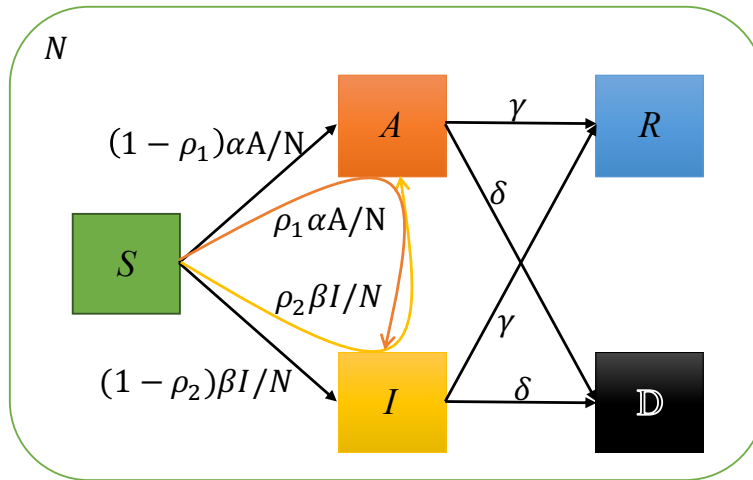


Figure 2: Diagram for the SAIRD model which is inspired from [46] illustrating the interactions of compartment.

The SAIRD model is given by the following set of differential equations:

$$\frac{dS}{dt} = -\beta I \frac{S}{N} - \alpha A \frac{S}{N} \,,$$
$$\frac{dA}{dt} = (1 - \rho_1)\alpha A \frac{S}{N} + \rho_2 \beta I \frac{S}{N} - \gamma A - \delta A \,,$$
$$\frac{dI}{dt} = (1 - \rho_2)\beta I \frac{S}{N} + \rho_1 \alpha A \frac{S}{N} - \gamma I - \delta I \,, \tag{2}$$
$$\frac{dR}{dt} = \gamma I + \gamma A \,,$$
$$\frac{dD}{dt} = \delta I + \delta A \,.$$

The transition rates and flow between the compartments are shown in Figure 2. The infections happen at a rate of $\alpha$ and $\beta$ is due to contact of a susceptible individual with an asymptomatic or symptomatic infectious individual, respectively. The probability of a susceptible individual to become a symptomatically infected individual, through contact with an asymptomatic individual, is $\rho_1$. The probability of a susceptible individual to become an asymptomatically infected individual, through contact with a symptomatically infected individual, is $\rho_2$. Infected individuals recover at a rate of $\gamma$. They will pass away at a rate of $\delta$, independently of their symptoms. This system (2) is subject to the initial conditions $S(t_0) > 0, A(t_0) \geq 0, I(t_0) \geq 0, R(t_0) \geq 0$ and $D(t_0) \geq 0$, with an initial time $t_0$. The SAIRD model assumes, just as, that the total population, including the deceased, remains constant, $N$. Hence, this model satisfies $S(t) + A(t) + I(t) + R(t) + D(t) = N$, at any time $t$.

## 2.2. Physics-Informed Neural Network

The basic idea of PINNs is to integrate any a-priori knowledge of the system into the learning process of the deep neural network. This knowledge, e.g., about basic physical laws or domain know-how, is usually given in the form of ordinary- or partial differential equations (ODEs/PDEs), and is incorporated in PINNs by the loss functions. The training is performed to optimize the network weights and biases, and also to optimize the model parameters (e.g., those inside the physical laws). The loss term consists of two terms, the data loss and the residual loss, representing the regularization from the physics prior, like the generally obeyed differential equations. Figure 3 illustrates the training process for an ODE-dynamics-informed neural network.

Figure 3 shows how a PINN can be used to tackle a solution of the system of ODEs by training the neural network using the combined losses. The system of first-order ordinary differential equations of the general form is usually written as:

$$\frac{\partial U}{\partial t}(t) + F(U(t); \lambda) = 0, \quad t \in [t_0, T] \tag{3}$$

with

$$U(t) = [u^1(t), ..., u^n(t)], \quad F(U) = [f^1(U), ..., f^n(U)] \tag{4}$$

where $u^i \in \mathbb{R}$ and $f^i : \mathbb{R} \to \mathbb{R}, i = 1, ..., n$. $t_0$ and $T$ are the initial and final time, respectively; $F$ is the function and $U$ the solution. The $\lambda \in \mathbb{R}^k$ are the unknown parameters of the system (3). $U_s$ are the observed data at times $t_1, ..., t_m$, which determine $\lambda$. The data loss is, naturally,
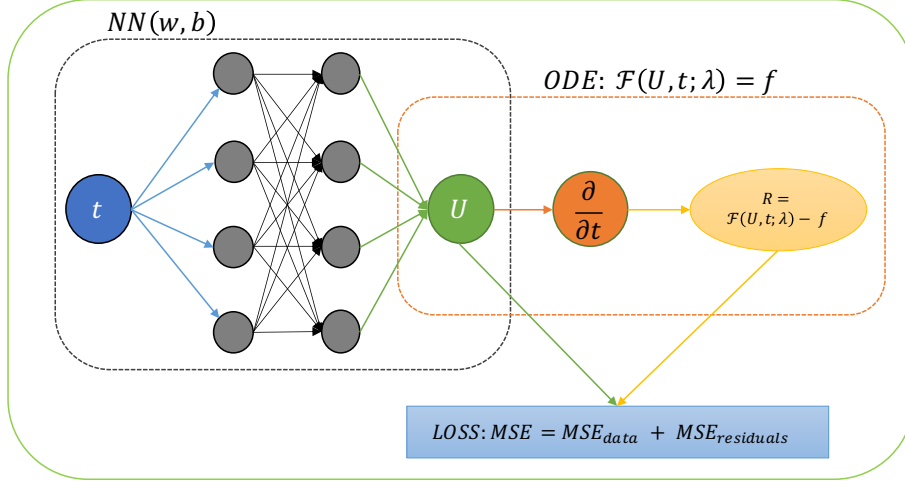
Figure 3: A schematic diagram of physics-informed neural networks (PINNs). The black dashed-line block is a common neural network that takes a time $t$ as input and the output is $U$, $\beta$ and $\gamma$ are weights and biases, respectively. The orange dashed-line block stands for the calculation of residual loss. The loss function consists of a mismatch of boundaries and initial conditions for the observed data (data loss). The residuals of the ODE is a set of random points in the spatial-temporal domain (residuals). The parameters of the PINNs can be optimized by minimising the loss $MSE = MSE_{data} + MSE_{residuals}$.

$$\mathcal{L}_{data} = \sum_{s=1}^{m} \|U(t_s) - U_s\|^2 \ . \tag{5}$$

For a conventional fitting, other than with PINNs, one identifies the optimum vector of the model parameters $\lambda$ by minimizing eq.(5). Thus, a solution $U(t)$ is obtained which is best suitable for the observed data, in the sense of the least squares deviations. The PINNs are based on a general neural network, as shown in Figure 3 (black dashed frame). Its form can be represented as follows:

$$NN^{\omega,b}(t) : \mathbb{R} \to \mathbb{R}^n \tag{6}$$

which approximates the solution

$$U(t) : \mathbb{R} \to \mathbb{R}^n \tag{7}$$

of the system of first-order ODEs. The weights $\omega$ and biases $b$ of $NN^{\omega,b}$ are trainable parameters of the neural network. For the purpose of solving ODEs like (3) with neural networks (6), the weights $\omega$ and biases $b$ can be optimized and that the neural network (6) offers the best fits of the observed data $U_s, s = 1, ..., m$, in the sense of the least squares differences,

$$\arg\min_{\omega,b}(\mathrm{MSE}_U^{\omega,b}) \tag{8}$$

Thus the loss function with respect to the observed data is

$$\mathrm{MSE}_U^{\omega,b} := \frac{1}{m} \sum_{s=1}^{m} \left\| NN^{\omega,b}(t_s) - U_s \right\|^2 . \tag{9}$$

6

The residual loss of the set of ODEs in (3), which can be expressed as

$$\mathcal{F}\left(NN^{\omega,b}, t; \lambda\right) = \frac{\partial NN^{\omega,b}}{\partial t}(t) + F\left(NN^{\omega,b}(t)\right), \tag{10}$$

allows to extend the $NN^{\omega,b}$ to a PINN by putting the residual term (10) into the loss function (9). The automatic differentiation technique of the neural networks can be used to compute the derivatives $\left(\frac{\partial NN^{\omega,b}}{\partial t}\right)$ of the output of the network with respect to the input (see Figure 3). Thus, letting $\mathcal{F}\left(NN^{\omega,b}, t; \lambda\right) = 0, \forall t \in [t_0, T]$ is equivalent to forcing a neural network (6) to fulfill the ODE dynamics (3).

In other words, the standard neural network can be turned into a PINN by adding a mean squared residual error $\text{MSE}_{\mathcal{F}}^{\omega,b,\lambda}$ to the loss function (8). Thus, PINNs can be trained to identify the optimum neural network parameters, $\omega$ and $b$, as well as the parameters $\lambda$ for the ODEs. In that way, the following ODE-dynamics-regularized optimization is solved:

$$\arg\min_{\omega,b,\lambda}(\text{MSE}_U^{\omega,b} + \text{MSE}_{\mathcal{F}}^{\omega,b,\lambda}). \tag{11}$$

### 2.3. SIR Model-Informed Machine Learning

This section explicitly introduces how to incorporate the SIR model into the PINN formalism, as a prior information. The dynamic parameters of the SIR model are estimated. This includes, both, the definition and the calculation of the data loss function and of the physical residuals.

### 2.4. Architecture

Figure 4 utilize, for the SIR model, a fully connected neural network (marked by the black-dashed frame), to evaluate $(S(t), I(t), R(t))^\top$ as defined in (6). Here, $S(t)$, $I(t)$ and $R(t)$ obey the SIR model at a given input time $t$. The residual term of the ODEs can be minimized for the SIR model as defined in (10) to enforce eq.(6). Thus, here

$$\mathcal{F}\left(NN^{\omega,b}, t; \beta, \gamma\right) = \begin{bmatrix} \frac{dS(t)}{dt} + \frac{\beta S(t)I(t)}{N} \\ \frac{dI(t)}{dt} - \frac{\beta S(t)I(t)}{N} + \gamma I(t) \\ \frac{dR(t)}{dt} - \gamma I(t) \end{bmatrix}. \tag{12}$$

Hence, the mean residual squared error of the present work is

$$\text{MSE}_{SIR} = \text{MSE}_{S_{residual}} + \text{MSE}_{I_{residual}} + \text{MSE}_{R_{residual}} \tag{13}$$

with

$$\text{MSE}_{S_{residual}} = \frac{1}{q}\sum_{i=1}^{q}\left|\frac{dS(t_i)}{dt_i} + \frac{\beta S(t_i)I(t_i)}{N}\right|^2,$$

$$\text{MSE}_{I_{residual}} = \frac{1}{q}\sum_{i=1}^{q}\left|\frac{dI(t_i)}{dt_i} - \frac{\beta S(t_i)I(t_i)}{N} + \gamma I(t_i)\right|^2, \tag{14}$$

$$\text{MSE}_{R_{residual}} = \frac{1}{q}\sum_{i=1}^{q}\left|\frac{dR(t_i)}{dt_i} - \gamma I(t_i)\right|^2.$$
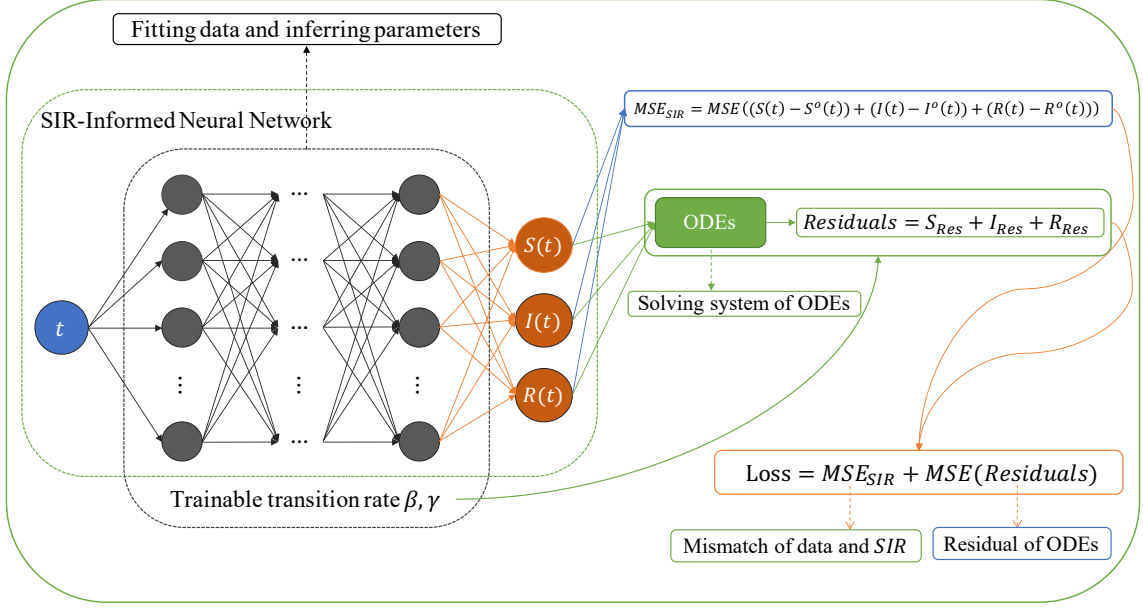
Figure 4: Schematic diagram of the SIR-dynamics informed neural network. The black-dashed frame represents the dense neural network used here. The green-dashed frame, on the other hand, represents the SIR-informed neural network, which takes time $t$ as input and outputs the susceptible-$(S)$, infected-$(I)$ and removed $(R)$ populations. The box labeled 'ODEs' represents the computation of the residual, with respect to the SIR model. The label 'Loss' is comprised of two parts: the mismatch between the available data and the network output, on one hand and the physical residual, on the other hand. The NN fits simultaneously both, the data and infers the dynamic parameters $\beta$ and $\gamma$, by satisfying the ODE dynamics, by minimizing loss function.

Here $q$ is the total number of discrete time points. Note that, the discrete time points are chosen to be consistent with the observed time step, chosen in units of one natural day. In other words, the time step between two time points is $\Delta t = 1$.

Simultaneously, the mean squared error of the data is, for the SIR model,

$$\text{MSE}_{data} = \text{MSE}_{S_{data}} + \text{MSE}_{I_{data}} + \text{MSE}_{R_{data}}. \tag{15}$$

Here

$$\text{MSE}_{S_{data}} = \frac{1}{s} \sum_{i=1}^{s} |S(t_i) - S_i^o|^2,$$

$$\text{MSE}_{I_{data}} = \frac{1}{s} \sum_{i=1}^{s} |I(t_i) - I_i^o|^2, \tag{16}$$

$$\text{MSE}_{R_{data}} = \frac{1}{s} \sum_{i=1}^{s} |R(t_i) - R_i^o|^2.$$

$S_i^o, I_i^o$ and $R_i^o$ are the observed data at time points $t_i$, and $s$ is the number of observed data points.

The total loss of the present study consists of both, the data loss and the residual loss, according to which, both, the weights and biases, as well as the trainable dynamic parameters, are optimized,

$$\underset{\omega,b,\beta,\gamma}{\arg\min}(\text{MSE}_{data} + \text{MSE}_{SIR}), \tag{17}$$

## 2.5. Parameter Identifications

Algorithm 1 shows how PINNs can be used to determine trainable parameters, including the NN- and SIR-model parameters. The input is the time point $t$, the output is the value of each compartment of the SIR model at a given $t$ value. Both, the weights $\omega$ and biases $b$, as well as the model parameters $\beta$ and $\gamma$, are initialized randomly ($\omega$ and $b$ use the PyTorch default initialization function, and $\beta$ and $\gamma$ were chosen randomly from (0,1). To ensure reproducibility, they were fixed during the practical experiments).

---

**Algorithm 1:** PINNs used to determine simultaneously the parameters of the neural network and the embedded SIR model.

---

**Data**: $t, S^o, I^o, R^o$

Randomly initialize weights $\omega$, biases $b$, and dynamics parameters $\beta$, $\gamma$ ;

**for** *epoch* ***in*** *epochs* **do**

    The values of each compartment of the SIR model can be obtained from the forward propagation of the neural network with the input as $t$

$$S, I, R = NN(t).$$

    Evaluate the composed loss function, including the data loss (with $s$ to be the number of observations in each compartment, thus the number of time points collected):

$$\text{MSE}_{SIR} = \frac{1}{s}\sum_{i=1}^{s}(|S_i - S_i^o|^2 + |I_i - I_i^o|^2 + |R_i - R_i^o|^2),$$

    denoting the mismatch of the output of the neural network and observation data. Here the residual loss:

$$\text{MSE}_{Residuals} = \frac{1}{q}\sum_{i=1}^{q}\left(\left|\frac{dS_i}{dt_i} + \frac{\beta S_i I_i}{N}\right|^2 + \left|\frac{dI_i}{dt_i} - \frac{\beta S_i I_i}{N} + \gamma I_i\right|^2 + \left|\frac{dR_i}{dt_i} - \gamma I_i\right|^2\right)$$

    stands for the sum of the residual errors for each compartment of the SIR model. Here, the residuals and data loss are calculated using the same time step $\Delta t = 1$. Thus, the total loss function can be obtained:

$$Loss = \text{MSE}_{SIR} + \text{MSE}_{Residuals}$$

    The Adam Optimizer[47] toolkit in Pytorch is utilized to update the weights $\omega$ and biases $b$, as well as $\beta$ and $\gamma$ by minimizing the loss function.

**end**

---

## 3. Experiments and Results

This section presents first the setup for generating the synthetic data from the SAIRD model, and then the pre-processing for the data reported for Germany. The details of the here devised physics-informed neural networks for the above two situations follow. Finally, the performance of proposed the PINNs are demonstrated for the testing data sets.

### 3.1. Data Preparation

#### 3.1.1. Synthetic data generation and processing

To test and validate the PINN frameworks, whose physical laws are derived from the simple SIR model, a complex mathematical model, the SAIRD model (2) is used, to generate the mock data. The advantage of using this type of data is that here the dynamics is clean, little noisy, verify that the approach does work. The data shown in Figure 5 starting the model evolution with the given model parameters and the initial conditions as shown in Table 1.

Table 1: The value of parameters and initial conditions for the SAIRD model.

| Parameter | Description | Value |
|---|---|---|
| $\rho_1$ | Probability that one becoming symptomatically by exposure to an asymptomatic carrier | 0.80 |
| $\rho_2$ | Probability that one becoming asymptomatically by exposure to a symptomatically carrier | 0.29 |
| $\alpha$ | Infection rate of individual exposed to an asymptomatic carrier | 0.1 |
| $\beta$ | Infection rate of individual exposed to symptomatic carrier | 0.17 |
| $\gamma$ | Recovery rate | 1/16 |
| $\theta$ | Death rate | 0.001 |
| $N$ | Population | 1000 |
| $A_0$ | Initial number of Asymptomatic infectious individuals | 10 |
| $I_0$ | Initial number of Infectious individuals | 20 |
| $R_0$ | Initial number of Recovered individuals | 0 |
| $D_0$ | Initial number of Dead individuals | 0 |
| $S_0$ | Initial number of Susceptible individuals | 970 |

Figure 5 shows the time evolution of the classical SIR model. Here the number of susceptibles in the population decreases as the number of actively infected people increases. The population size for recovered and death (removal) then increases until it hits a maximum or stabilizes.

The epidemiological characteristics of the SAIRD model, both its S (Susceptible) and A (Asymptomatic) compartments are derived from the S (Susceptible) compartment of the SIR model. The SAIRD's R (Recovered) and D (Dead) compartments are obtained from the R (Removed) compartment in the SIR model.

Hence, the values of the S and R compartments of the SIR model are obtained by adding the values of A to S in the SAIRD model, and the values of R to D, respectively. The compartment I have been defined identically in both models. So, the values of the three compartments in the SIR model, shown in Figure 6 are used as simulation data for the straightforward tests and validations to follow here.

All values of all compartments stabilize after approximately 100 days, the first 100 days, where compartment I exhibits a broad, 30 day wide peak, are chosen as the synthetic dataset for the calculations and testing shown in Figure 7.
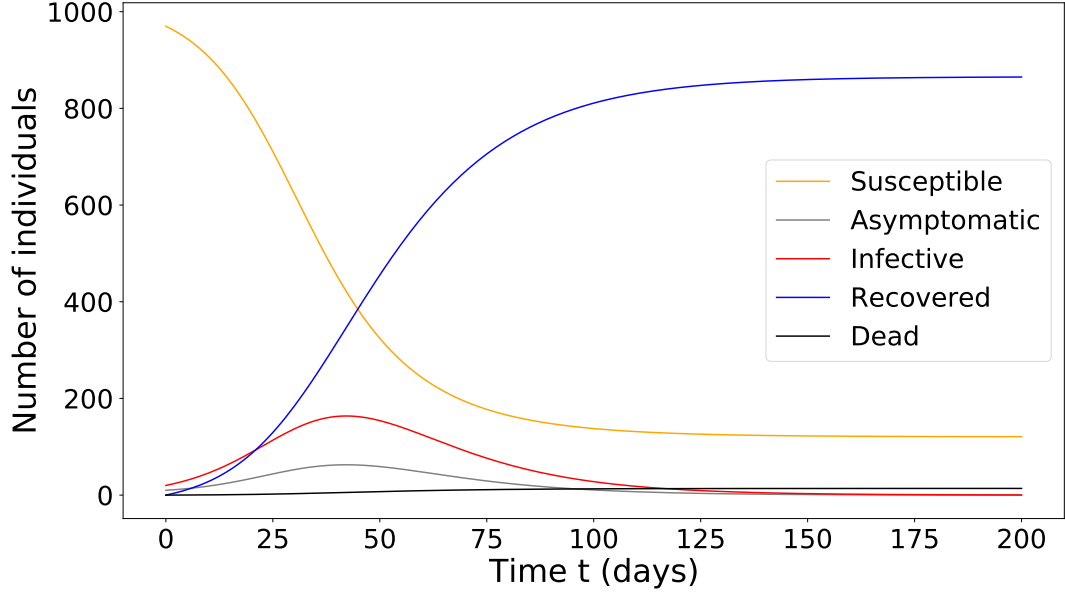
Figure 5: A Mathematical model generated time evolution of an example SAIRD model. The solid yellow line represents the number of susceptible people in the population, the solid grey line represents the number of asymptomatic infected people, the solid blue line represents the number of recovered people, and the solid red line stand for the number of active infected person. The solid black line is the death population, and The population is assumed to be constant (N=1000).
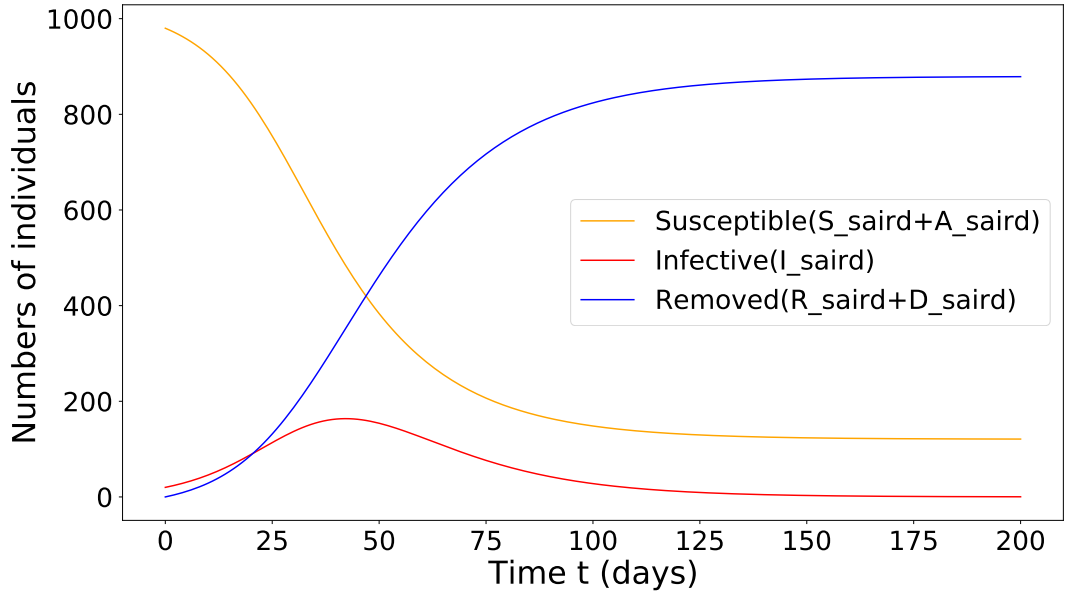


Figure 6: SIR mock data derived from SAIRD model generation data.

### 3.1.2. Pre-processing of reported data

The COVID-19 epidemic data of Germany are used for our model analysis. they are based on the cumulative number of infected individuals, the cumulative recovered individuals, and the cumulative number of deaths, as reported by Robert Koch-Institut (RKI) COVID-19 Data for the period from March 1 to July 1, 2021. The SIR dynamics integrated into our method. Hence, the first step of the preprocessing of the reported data is to extract the values for the different compartment from the records. The data for Germany track the cumulative
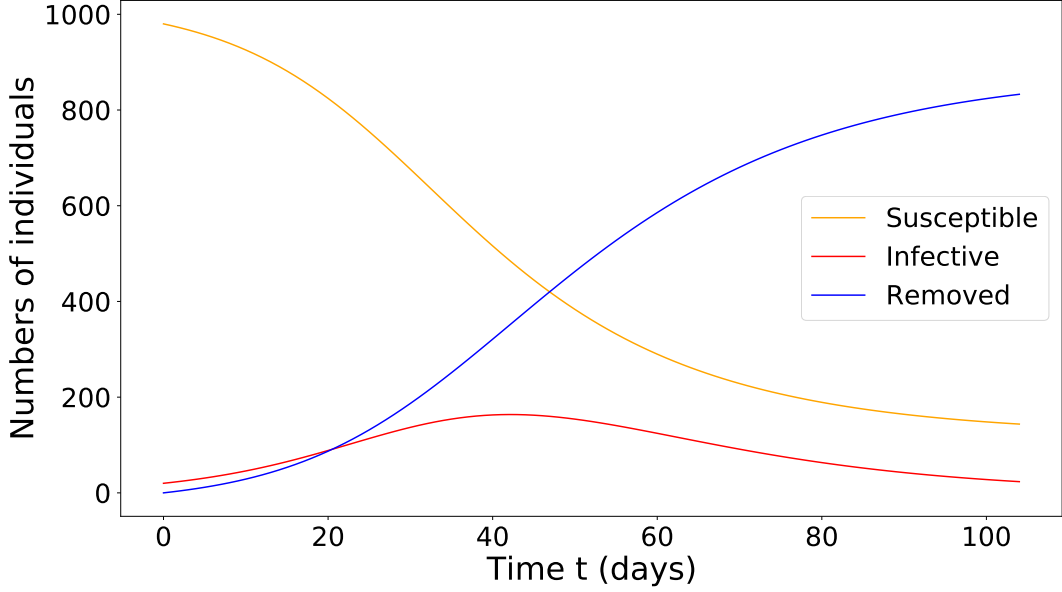
Figure 7: Selected SIR mock data derived from the SAIRD model generation data are used for the synthetic data simulations.

infectious, $I_c$, recovered, $R_c$, and deceased cases, $D_c$. In accord with the definition of the compartment $R$, in the SIR model, $R$ contains both the recovered as well as the virus-induced dead individuals, $R = R_{recovered} + D_{death}$. As the compartment $R$ does not have any outflow, $R = Re_c$. Here, $Re_c$ is the cumulative value of the removed. Given the reported data's values, the theoretical values for removed, infectious, susceptible cases are obtained by $R = R_{recovered} + D_{death}$, $I = I_c - R$, and $S = N - I - R$. Here, $N$ is the total population. All values of all compartments are normalized by dividing by the total population. Thus, N is set to 1.

Reliability and cleanliness of the reported data matter a lot for the predictability of compartmental models, in particular for the sense of parameter estimates. Reported data are, however, quite noisy, both due to misreporting, late reporting and other reasons. The reported data from Germany vary greatly substantially between weekdays and weekends. There is due to the significant reduction in detection at weekends. Therefore, the ability to make valuable estimates from the available reported data may be limited. Thus, before the data was analysed, we pre-process the dataset by applying a 7-day moving average window to smooth out weekday-weekend zigzags in the outbreak reports. As shown in the Figure 8, the reported data from the different compartments are significantly smoother and less noisy than the raw data after the sliding window pre-processing.

*3.2. Setup of SIR-Infomed Neural Networks*

The neural network structure we use in our experiments takes a single value, the time $t$, as input. There're weights $W_{i,j}$ being associated for each hidden layer, where $i$ is the position of the start node and $j$ is the position of the end node. The non-linear activation function *tanh* is applied at each node in the hidden layer,

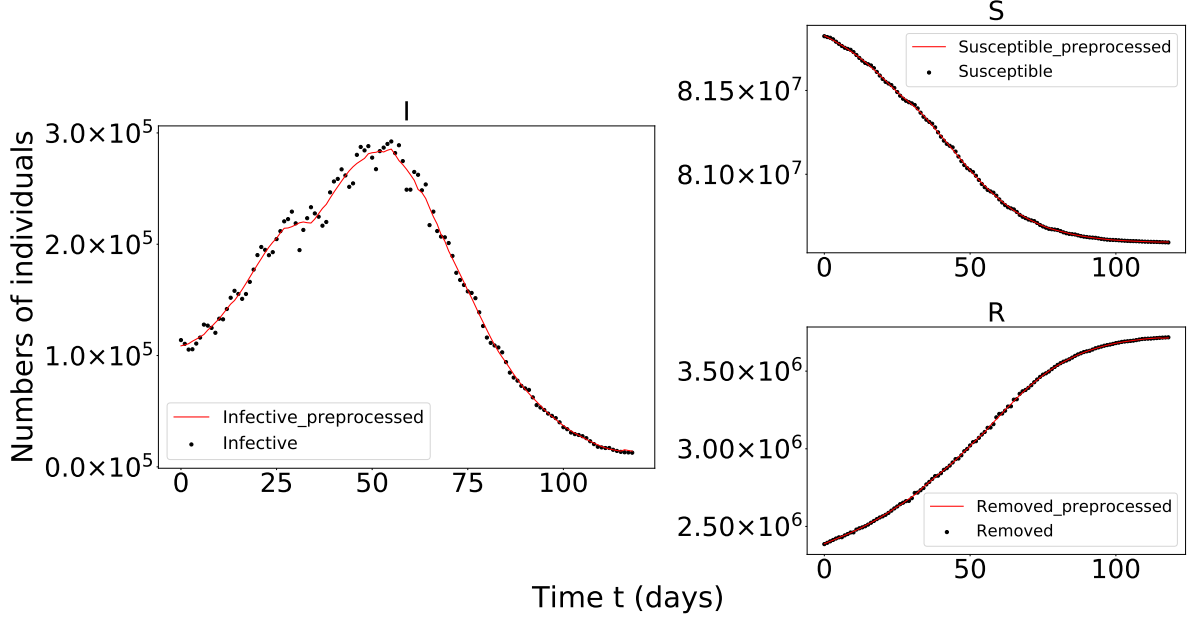$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{18}$$

12

Figure 8: The figure compares the results before and after pre-processing of the reported data. I, S and R correspond to the trends in the number of infectious, susceptible, and removed individuals, respectively, where the black scatter shows the raw, unprocessed data, and the solid red line is a curve of the data processed through a sliding window. The graphs show that the pre-processed data curves are smoother, and the fluctuations are reduced effectively.

For the output nodes of the network, *sigmoid* activation function is applied,

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \tag{19}$$

to account for the normalization applied to $S(t)$, $I(t)$ and $R(t)$ respectively. The neural network contains 4 hidden layers and 64 neurons in each layer. The Adam algorithm of the PyTorch package was selected as the optimizer. The used learning rate is 0.0001 and the training epochs is 200k. To prevent overfitting and improve the generalization ability of the model in confronting new data, we applied early stopping, to which the epoch limit is set to be 300.

*3.3. Performance on synthetic data*

We investigated using respectively 15, 25 and 37 days of data as the training set and, to ensure reproducibility, the same initial values of $\beta = 0.15$ and $\gamma = 0.15$ to test the identifiability and predictability of our proposed approach. In Figures 9, 10 and 11, we present identification and prediction results of our proposed method for using different sizes of training sets on mock data. Clearly we see that for a complex system including compartments S and R, the PINNs integrated with simple SIR dynamics can fit and predict well on the training and test sets. The numerical solutions for the ODE with the PINN learned parameters also perform well which basically would overlap the PINN results. However, for data with peaks like compartment I, the SIR dynamics informed neural network model gives slightly better results than the numerical solution of the ODE, especially at the peaks. This can also be seen in the loss value shown in the Table 2 of the results.
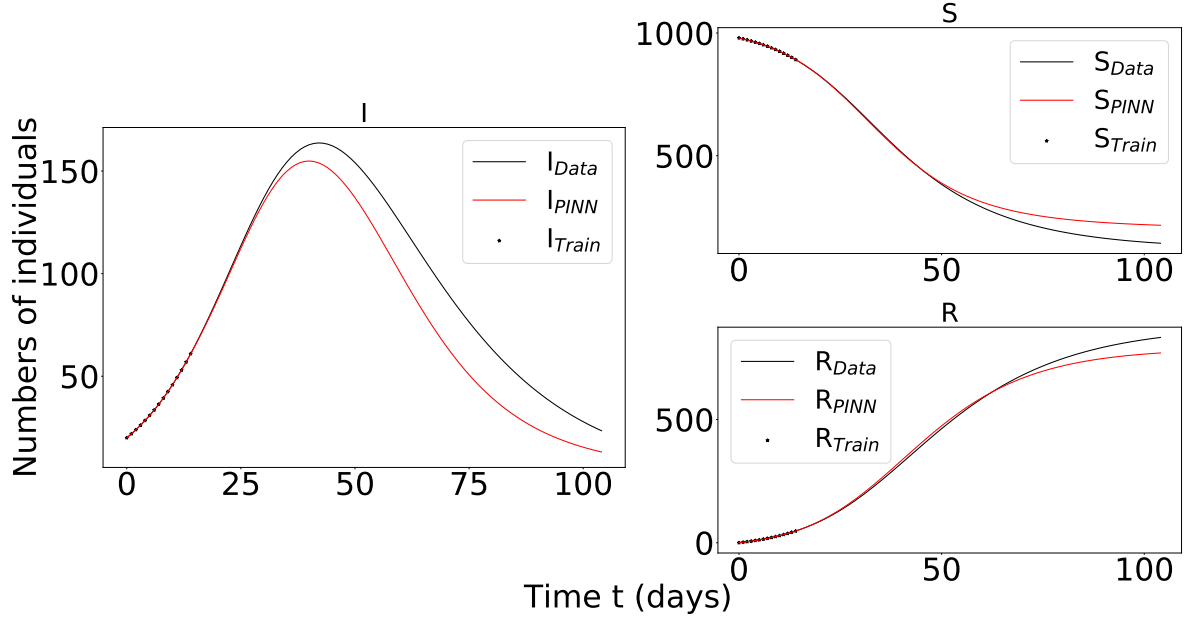
Figure 9: The graph shows results identified and predicted using the first 15 days as the training set. I, S and R diagrams of compartments I, S and R, respectively. The black star indicates the training set, the solid black line is the synthetic data, the solid red line is the result from PINNs.
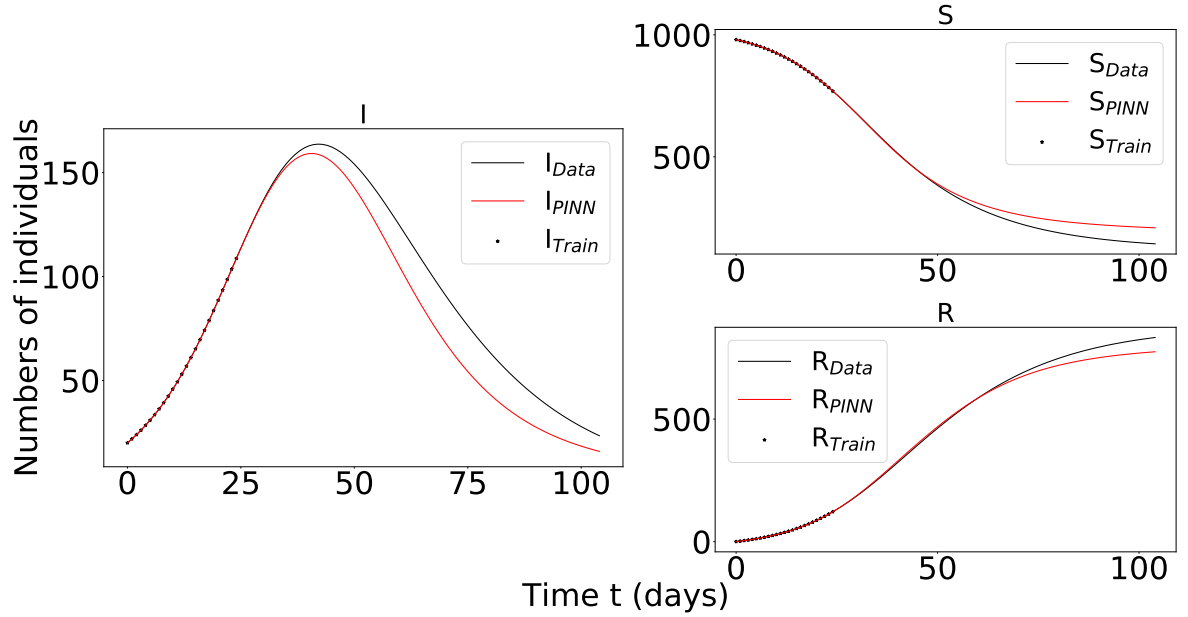


Figure 10: The graph shows results identified and predicted using the first 25 days as the training set. I, S and R diagrams of compartments I, S and R, respectively. The black star indicates the training set, the solid black line is the synthetic data, the solid red line is the result from PINNs.

## 3.4. Performance on Reported Germany Data

After validation of the method on synthetic data, we move to test it on the realistic reported data. With pre-processing performed on the raw reported German data from March 1 to July 1, 2021, we selected the first 30, 40, and 50 days record of the compartment I (infected) as the training set and, to ensure reproductively, the same initial values of $\beta = 0.25$ and $\gamma = 0.15$ to
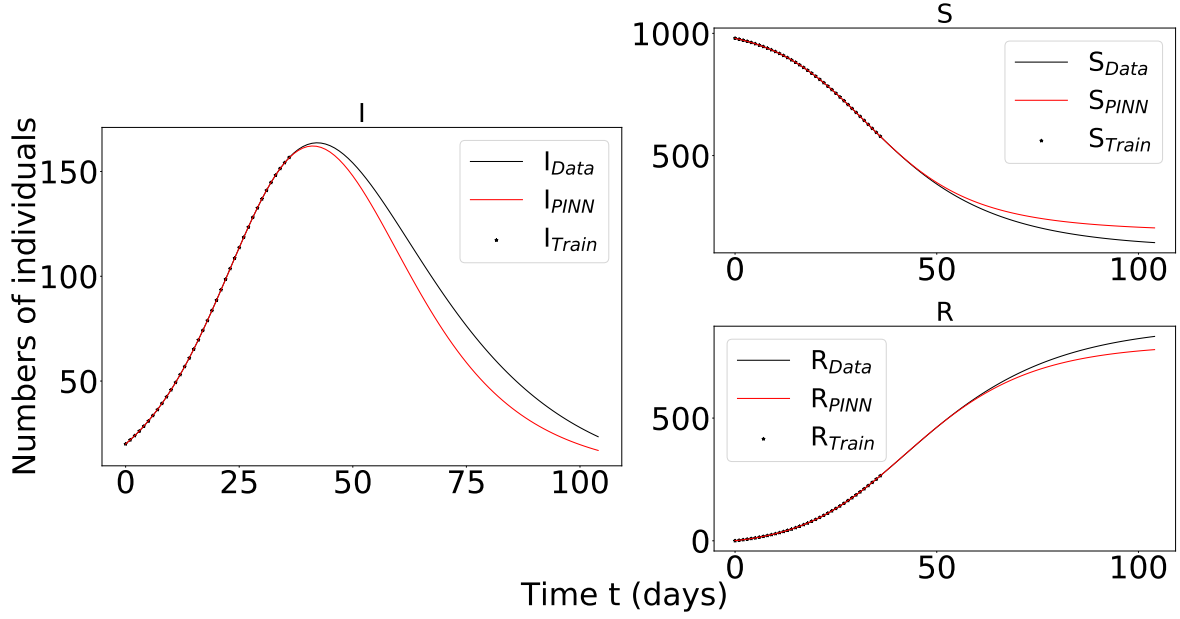
14

Figure 11: The graph shows results identified and predicted using the first 37 days as the training set. I, S and R diagrams of compartments I, S and R, respectively. The black star indicates the training set, the solid black line is the synthetic data, the solid red line is the result from PINNs.

Table 2: Error Metrics for PINNs and SIR-ODEs on different training set size.

|    |          | MSE    | MAE    |
|----|----------|--------|--------|
| 15 | PINNs    | 2202.6 | 53.889 |
|    | SIR-ODEs | 2946.3 | 63.172 |
| 25 | PINNs    | 1790.6 | 46.215 |
|    | SIR-ODEs | 2504.0 | 57.002 |
| 37 | PINNs    | 1461.8 | 39.919 |
|    | SIR-ODEs | 2025.6 | 51.145 |

Table 3: Parameter learned for PINNs on different training set size.

|          | Parameters learned | | | |
|----------|---------|---------|---------|---------|
|          | initial | 15      | 25      | 37      |
| $\beta$  | 0.15    | 0.18266 | 0.17948 | 0.17689 |
| $\gamma$ | 0.15    | 0.09486 | 0.09208 | 0.08948 |

the method for analysis and prediction testing. Figures 12, 13 and 14 show the identification and prediction results for using first 30, 40, 50 days as training set, respectively, including results from PINNs, solutions from the ODEs system (1) with PINN learned parameters, and the reported data (after pre-processing). In which, we adopt the LSODA algorithm [48] to solve the SIR model numerically with PINNs-learned parameters.

We can see that the proposed SIR-dynamics-informed neural network fits the training set

15

well, and the trend of the predictions is also reasonably consistent with the true data especially at the peak appearance for I. Compared to the solutions of the ODE with parameters learned from the PINN, the results from SIR-informed NN have better a match on the training set e.g. on the I compartment while the prediction trend maintains, because of the balance introduced between the data loss and physical residual loss. In particular, the predictions of the proposed method are closer to the actual values around the peaks. It can be seen in the Table 4. Compared to purely data-driven machine learning method on the same problem, the residual loss contained in our method serves as a physics informed regularizer. On the other hand, the data-loss can also be viewed as a regularizer for the conventional physics dynamics fitting (e.g., using SIR or any other compartmental model).

Then, the PINNs learned parameters are used to produce the numerical solution of SIR model as Eq.(1) shows, which is obtained by using the LSODA algorithm [48], which is used to solve the numerical solution problem for rigid or non-rigid systems of first-order ordinary differential equations. Then, the solution of the ODEs system is compared with the results from the PINNs.
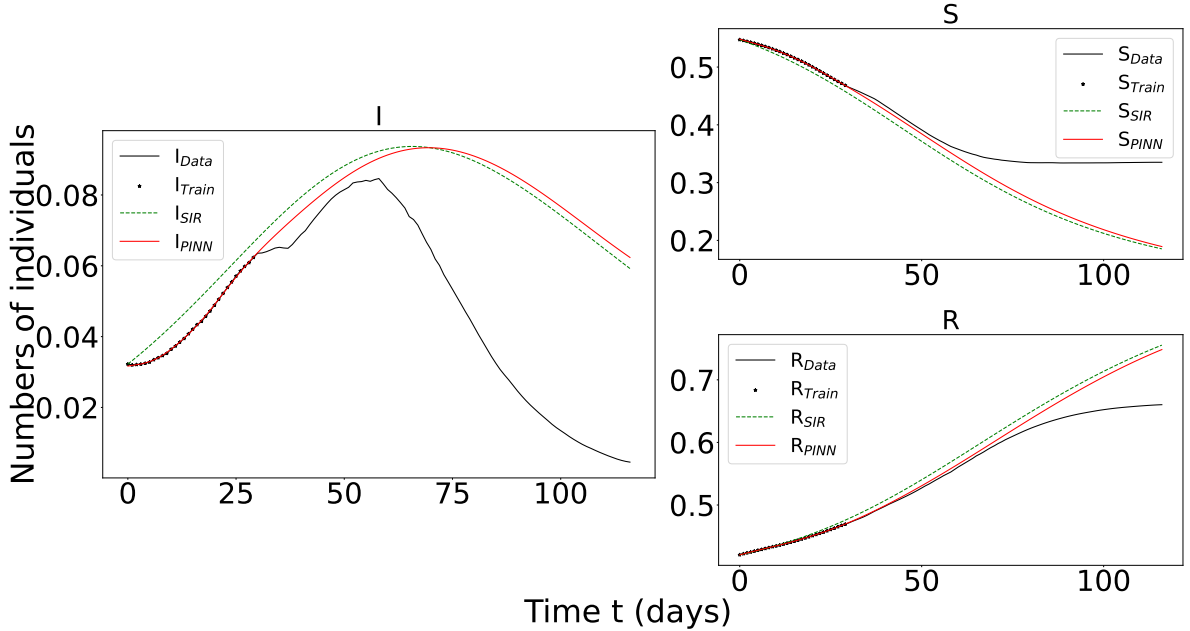


Figure 12: The graph shows the compartment results identified and predicted using the first 30 days as the training set. I, S and R diagrams of compartments I, S and R, respectively. The black star indicates the training set, the solid black line is the reported data, the red solid line is the result from PINNs, and the dashed green line indicates the result obtained by solving the SIR system using the parameters learned from PINNs.

We found that in the process of identifying dynamics and makeing forecasting on reported German data using PINN, several factors have an impact on the training convergence speed and final results: the initial value setting of the trainable parameters, the different assignment of the data loss, and residual loss weights in the loss function. However, this effect is negligible for synthetic data, which is less noisy and with cleaner dynamics. In Part A of the Appendix, we analyzed the impact of different weight assignments for the loss function's data loss and residual loss on the error and the prediction and identification of the number of infections. The goal is to analyze which weighting assignment of the data and residual loss in the loss
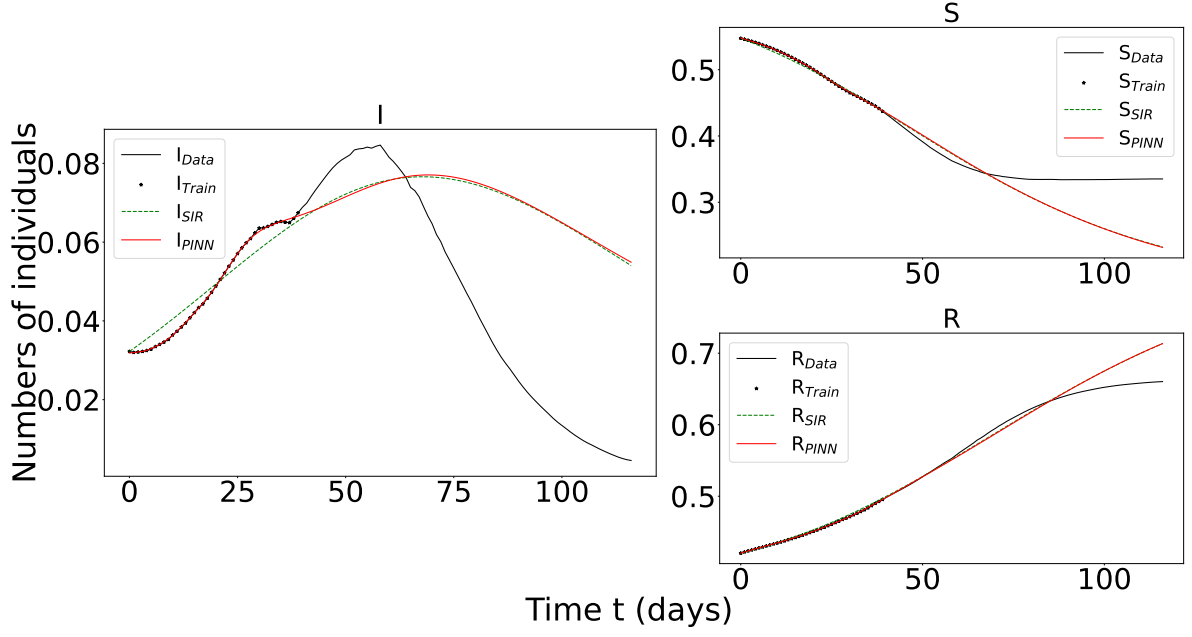
Figure 13: The graph shows the compartment results identified and predicted using the first 40 days as the training set. I, S and R diagrams of compartments I, S and R, respectively. The black star indicates the training set, the solid black line is the reported data, the red solid line is the result from PINNs, and the dashed green line indicates the result obtained by solving the SIR system using the parameters learned from PINNs.

Table 4: Error metrics for PINNs and SIR-ODEs on Germany reported data.

|    |          | MSE     | MAE     |
|----|----------|---------|---------|
| 30 | PINNs    | 0.00644 | 0.08468 |
|    | SIR-ODEs | 0.00745 | 0.10278 |
| 40 | PINNs    | 0.00248 | 0.04984 |
|    | SIR-ODEs | 0.00244 | 0.05179 |
| 50 | PINNs    | 0.00499 | 0.06934 |
|    | SIR-ODEs | 0.00568 | 0.08474 |

Table 5: The result of parameter learned from PINNs on Germany reported data.

|          | Parameters learned | | | |
|----------|---------|--------|--------|--------|
|          | initial | 30     | 40     | 50     |
| $\beta$  | 0.25    | 0.1268 | 0.1167 | 0.1233 |
| $\gamma$ | 0.15    | 0.0392 | 0.0404 | 0.0391 |

function can be optimal for the error impact of the results, i.e., lead to the lowest error and more accurate identification and prediction in compartment I.
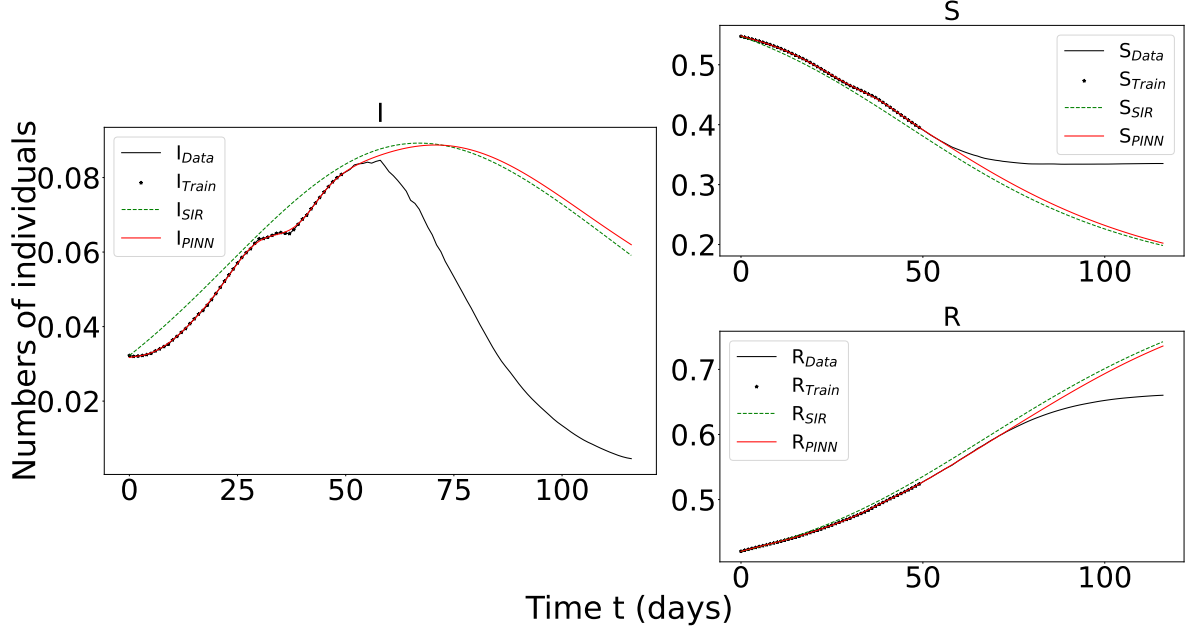
Figure 14: The graph shows the compartment results identified and predicted using the first 50 days as the training set. I, S and R diagrams of compartments I, S and R, respectively. The black star indicates the training set, the solid black line is the reported data, the red solid line is the result from PINNs, and the dashed green line indicates the result obtained by solving the SIR system using the parameters learned from PINNs.

## 4. Conclusions

A machine learning approach based on physics-informed neural networks is considered to identify, predict and estimate, the parameters for the evolution of dynamic systems. We combine it with a simple SIR compartmental model from mathematical epidemiology to learn and explore complex epidemiological dynamics. We first demonstrate the proposed method on synthetic data from a SAIRD model in multiple numerical experiments, with different-sized training sets generated by solving a system of ODEs using fixed parameters and initial conditions. It's found that the PINNs with SIR model performs better on the training set and provided a better trend of the prediction than solutions from ODEs just with model parameters estimated using PINNs.

Then, we extended our tests to realistic reported data about the COVID-19 epidemic in Germany. Specifically we picked a period in which compartment I had a peak for demonstration. The experimental results show that the PINNs combined with the SIR model also gives better results in the training set and trend prediction for the test set than solving the ODE system using the parameters estimated by the PINNs. Experimental results from both simulations and reported data show that it is feasible to put physical information from simple epidemic models into PINNs to study more complex epidemiological dynamics. Moreover, with certain finetuning, better results can be achieved.

However, in order to obtain more accurate results in actual epidemiological data, we should extend our study to more detailed models in mathematical epidemiology while explore other architectures of PINNs. This will be the subject of future work.

## Data Availability Statement

The reported COVID-19 data for Germany used for the present work are available at the COVID-19 Data Repository (https://github.com/robert-koch-institut) and provided by the German Robert Koch Institute (RKI Berlin, Germany) data.

## Conflicts of Interest

The authors declare that they do not see any conflict of interest.

## References

[1] WHO, Who coronavirus for covid-19 on 2022., World Health Organization (2022). URL https://covid19.who.int/

[2] A. Sedaghat, S. Band, A. Mosavi, L. Nadai, Predicting trends of coronavirus disease (covid-19) using sird and gaussian-sird models, in: 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), IEEE, 2020, pp. 000267–000274.

[3] A. F. Siegenfeld, N. N. Taleb, Y. Bar-Yam, What models can and cannot tell us about covid-19, Proceedings of the National Academy of Sciences 117 (28) (2020) 16092–16095.

[4] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. Pastore y Piontti, K. Mu, L. Rossi, K. Sun, et al., The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak, Science 368 (6489) (2020) 395–400.

[5] A. Wilder-Smith, D. O. Freedman, Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-ncov) outbreak., Journal of travel medicine (2020).

[6] S. Zhang, Z. Wang, R. Chang, H. Wang, C. Xu, X. Yu, L. Tsamlag, Y. Dong, H. Wang, Y. Cai, Covid-19 containment: China provides important lessons for global response, Frontiers of Medicine 14 (2) (2020) 215–219.

[7] N. Banholzer, E. Van Weenen, A. Lison, A. Cenedese, A. Seeliger, B. Kratzwald, D. Tschernutter, J. P. Salles, P. Bottrighi, S. Lehtinen, et al., Estimating the effects of non-pharmaceutical interventions on the number of new infections with covid-19 during the first epidemic wave, PLoS one 16 (6) (2021) e0252827.

[8] N. Thiel, C. Selwyn, G. Murphy, S. Simpson, A. C. Chakrabarti, Recommendations for acceleration of vaccine development and emergency use filings for covid-19 leveraging lessons from the novel oral polio vaccine, npj Vaccines 6 (1) (2021) 1–8.

[9] L. Cao, Q. Liu, Covid-19 modeling: A review, medRxiv (2022).

[10] S. W. Park, D. M. Cornforth, J. Dushoff, J. S. Weitz, The time scale of asymptomatic transmission affects estimates of epidemic potential in the covid-19 outbreak, Epidemics 31 (2020) 100392.

[11] F. Dorn, S. Khailaie, M. Stoeckli, S. C. Binder, T. Mitra, B. Lange, S. Lautenbacher, A. Peichl, P. Vanella, T. Wollmershäuser, et al., The common interests of health protection and the economy: evidence from scenario calculations of covid-19 containment policies, The European Journal of Health Economics 24 (1) (2023) 67–74.

[12] S. Tanwar, A. Kumari, D. Vekaria, N. Kumar, R. Sharma, An ai-based disease detection and prevention scheme for covid-19, Computers and Electrical Engineering 103 (2022) 108352.

[13] Y. N. Ertas, M. Mahmoodi, F. Shahabipour, V. Jahed, S. E. Diltemiz, R. Tutar, N. Ashammakhi, Role of biomaterials in the diagnosis, prevention, treatment, and study of corona virus disease 2019 (covid-19), Emergent materials 4 (1) (2021) 35–55.

[14] S. Kumar, S. Managi, Does stringency of lockdown affect air quality? evidence from indian cities, Economics of Disasters and Climate Change 4 (3) (2020) 481–502.

[15] D. I. Papadopoulos, I. Donkov, K. Charitopoulos, S. Bishara, The impact of lockdown measures on covid-19: a worldwide comparison, MedRxiv (2020).

[16] M. V. Barbarossa, J. Fuhrmann, J. Heidecke, H. V. Varma, N. Castelletti, J. H. Meinke, S. Krieg, T. Lippert, A first study on the impact of current and future control measures on the spread of covid-19 in germany, medRxiv (2020) 2020–04.

[17] M. V. Barbarossa, J. Fuhrmann, Germany's next shutdown—possible scenarios and outcomes, Influenza and other respiratory viruses 15 (3) (2021) 326–330.

[18] N. C. Peeri, N. Shrestha, M. S. Rahman, R. Zaki, Z. Tan, S. Bibi, M. Baghbanzadeh, N. Aghamohammadi, W. Zhang, U. Haque, The sars, mers and novel coronavirus (covid-19) epidemics, the newest and biggest global health threats: what lessons have we learned?, International journal of epidemiology 49 (3) (2020) 717–726.

[19] Y. Xiang, Y. Jia, L. Chen, L. Guo, B. Shu, E. Long, Covid-19 epidemic prediction and the impact of public health interventions: A review of covid-19 epidemic models, Infectious Disease Modelling 6 (2021) 324–342.

[20] I. Cooper, A. Mondal, C. G. Antonopoulos, Dynamic tracking with model-based forecasting for the spread of the covid-19 pandemic, Chaos, Solitons & Fractals 139 (2020) 110298.

[21] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character 115 (772) (1927) 700–721.

[22] J. L. Aron, I. B. Schwartz, Seasonality and period-doubling bifurcations in an epidemic model, Journal of theoretical biology 110 (4) (1984) 665–679.

[23] P. Sun, K. Li, An seir model for assessment of current covid-19 pandemic situation in the uk, medRxiv (2020).

[24] L. Basnarkov, Seair epidemic spreading model of covid-19, Chaos, Solitons & Fractals 142 (2021) 110394.

[25] R. Dandekar, G. Barbastathis, Quantifying the effect of quarantine control in covid-19 infectious spread using machine learning, MedRxiv (2020).

[26] M. Jamshidi, A. Lalbakhsh, J. Talla, Z. Peroutka, F. Hadjilooei, P. Lalbakhsh, M. Jamshidi, L. La Spada, M. Mirmozafari, M. Dehghani, et al., Artificial intelligence and covid-19: deep learning approaches for diagnosis and treatment, Ieee Access 8 (2020) 109581–109595.

[27] S. Vaid, R. Kalantar, M. Bhandari, Deep learning covid-19 detection bias: accuracy through artificial intelligence, International Orthopaedics 44 (8) (2020) 1539–1542.

[28] R. L. Stout, S. J. Rigatti, The silent pandemic covid-19 in the asymptomatic population, medRxiv (2021) 2020–12.

[29] C. J. Worby, H.-H. Chang, Face mask use in the general population and optimal resource allocation during the covid-19 pandemic, Nature communications 11 (1) (2020) 1–9.

[30] T. Oraby, M. G. Tyshenko, J. C. Maldonado, K. Vatcheva, S. Elsaadany, W. Q. Alali, J. C. Longenecker, M. Al-Zoughool, Modeling the effect of lockdown timing as a covid-19 control measure in countries with differing social contacts, Scientific reports 11 (1) (2021) 1–13.

[31] W. Choi, E. Shim, Optimal strategies for social distancing and testing to control covid-19, Journal of theoretical biology 512 (2021) 110568.

[32] M. V. Barbarossa, J. Fuhrmann, J. H. Meinke, S. Krieg, H. V. Varma, N. Castelletti, T. Lippert, Modeling the spread of covid-19 in germany: Early assessment and possible scenarios, Plos one 15 (9) (2020) e0238559.

[33] A. Hassan, I. Shahin, M. B. Alsabek, Covid-19 detection system using recurrent neural networks, in: 2020 International conference on communications, computing, cybersecurity, and informatics (CCCI), IEEE, 2020, pp. 1–5.

[34] L. Wang, T. Xu, T. Stoecker, H. Stoecker, Y. Jiang, K. Zhou, Machine learning spatio-temporal epidemiological model to evaluate germany-county-level covid-19 risk, Machine Learning: Science and Technology 2 (3) (2021) 035031.

[35] V. K. R. Chimmula, L. Zhang, Time series forecasting of covid-19 transmission in canada using lstm networks, Chaos, Solitons & Fractals 135 (2020) 109864.

[36] A. Zeroual, F. Harrou, A. Dairi, Y. Sun, Deep learning methods for forecasting covid-19 time-series data: A comparative study, Chaos, Solitons & Fractals 140 (2020) 110121.

[37] T. Qin, K. Wu, D. Xiu, Data driven governing equations approximation using deep neural networks, Journal of Computational Physics 395 (2019) 620–635.

[38] Z. Chen, D. Xiu, On generalized residual network for deep learning of unknown dynamical systems, Journal of Computational Physics 438 (2021) 110362.

[39] T. T. Nguyen, Q. V. H. Nguyen, D. T. Nguyen, E. B. Hsu, S. Yang, P. Eklund, Artificial intelligence in the battle against coronavirus (covid-19): a survey and future research directions, arXiv preprint arXiv:2008.07343 (2020).

[40] J. Chen, K. Li, Z. Zhang, K. Li, P. S. Yu, A survey on applications of artificial intelligence in fighting against covid-19, ACM Computing Surveys (CSUR) 54 (8) (2021) 1–32.

[41] M. Ganslmeier, D. Furceri, J. D. Ostry, The impact of weather on covid-19 pandemic, Scientific reports 11 (1) (2021) 1–7.

[42] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational physics 378 (2019) 686–707.

[43] R. Wang, R. Yu, Physics-guided deep learning for dynamical systems: A survey, arXiv preprint arXiv:2107.01272 (2021).

[44] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nature Reviews Physics 3 (6) (2021) 422–440.

[45] E. Kharazmi, M. Cai, X. Zheng, Z. Zhang, G. Lin, G. E. Karniadakis, Identifiability and predictability of integer-and fractional-order epidemiological models using physics-informed neural networks, Nature Computational Science 1 (11) (2021) 744–753.

[46] M. Angeli, G. Neofotistos, M. Mattheakis, E. Kaxiras, Modeling the effect of the vaccination campaign on the covid-19 pandemic, Chaos, Solitons & Fractals 154 (2022) 111621.

[47] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization (2014). `doi:10.48550/ARXIV.1412.6980`.

[48] A. C. Hindmarsh, Odepack: A systemized collection of ode solvers, Scientific computing (1983) 55–64.

## Appendix A: Sensitivity analysis for weight setup

The PINN's loss function consists of two parts, the data loss and the residual loss.

$$LOSS = \alpha_1 MSE_{data} + \alpha_2 MSE_{residual}. \qquad (20)$$

Here $\alpha_1$ and $\alpha_2$ are the weights for data loss and residual loss, respectively.

Table 6: The different allocations of data loss and residual loss are compared with the Mean Average Error(MAE) and the Mean Squared Error(MSE) of the respective results. Here $\alpha_1$ is the weight of the data loss and $\alpha_2$ is the weight of the residual loss.

| $(\alpha_1, \alpha_2)$ | (1,1) | (5,3) | (10,5) | (15,7) | (20,9) | (25,11) | (30,13) |
|---|---|---|---|---|---|---|---|
| MAE | 0.04240 | 0.03557 | 0.03551 | 0.03476 | 0.03444 | 0.03376 | 0.03423 |
| MSE | 0.00145 | 0.00107 | 0.00103 | 0.00098 | 0.00097 | 0.00092 | 0.00095 |