University *of* Wisconsin
Eau Claire

Application of SIRD Model and Physics-Informed Neural Networks to Model COVID-19 Disease Dynamics

Aleya Hadenfeldt, Paige Keller

Introduction

Infectious diseases such as COVID-19 pose ongoing challenges to public health systems worldwide. Understanding how a disease spreads and changes over time is critical for informing public policy and protecting communities. Mathematical models are commonly used to describe these dynamics and provide insight into potential outcomes during an outbreak.

One widely used approach is the SIRD model, which divides a population into four groups and describes how individuals move between these groups using differential equations. Although this model offers a clear and interpretable framework, it relies on simplifying assumptions and often struggles to capture complex, time-dependent patterns influenced by real-world factors such as behavior changes, testing availability, and vaccination efforts.

To better capture these effects, we moved to data-driven methods such as neural networks which can be used to model nonlinear relationships directly from observed data. In this work, we compare a traditional SIRD model with neural network approaches using real-world COVID-19 data. This comparison motivates the use of Physics-Informed Neural Networks (PINNs), which combine epidemiological models with machine learning to produce more flexible and realistic representations of disease dynamics.

Constant SIRD Model

The **SIRD** model partitions the population into four compartments: Susceptible (S), Infected (I), Recovered (R), and Deceased (D), with a fixed total population $N = S + I + R + D$. Disease dynamics are governed by a system of ordinary differential equations assuming constant transmission and recovery rates. The SIRD model equations:

$$\frac{dS}{dt} = -\beta\, SI$$

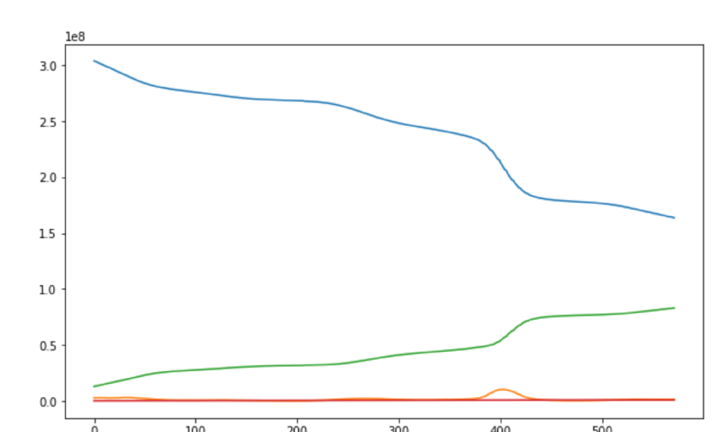
$$\frac{dI}{dt} = \beta\, SI - \gamma\, I - \mu\, I$$

$$\frac{dR}{dt} = \gamma\, I$$

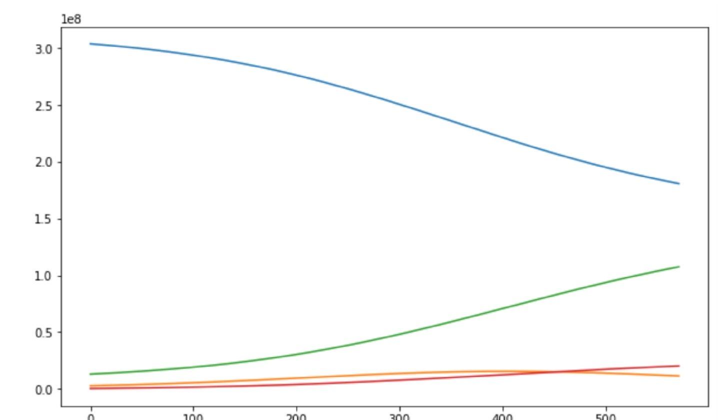
$$\frac{dD}{dt} = \mu\, I$$

Implemented a Constant SIRD model in Python to establish a baseline for comparison with the PINN model. The output of the SIRD model is shown in the figures to the right.

Observed Data



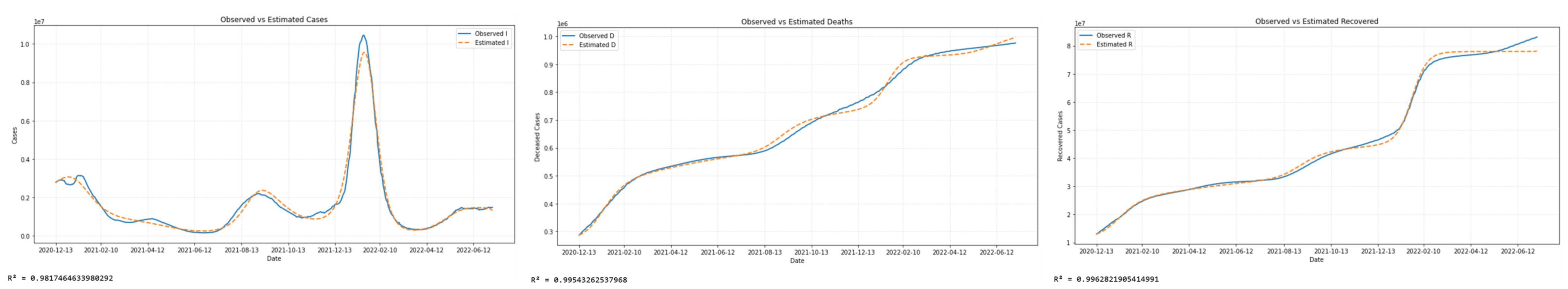
SIRD Approximation



Time Varying SIRD Model

In a **Time Varying SIRD** model, the parameters aren't fixed, but change over time. This method requires more data, but offers greater flexibility and can better reflect an outbreaks dynamics.

- Building on the framework presented in [2], we adapted their R implementation of the Time-Varying SIR model into a Time-Varying SIRD model and applied it to our dataset. This produces a new CSV containing $\beta(t)$, $\gamma(t)$, $\mu(t)$, and the estimated time series for confirmed cases, recovered cases, and deaths.



R Squared = 0.98 R Squared = 0.995 R Squared = 0.996

Data

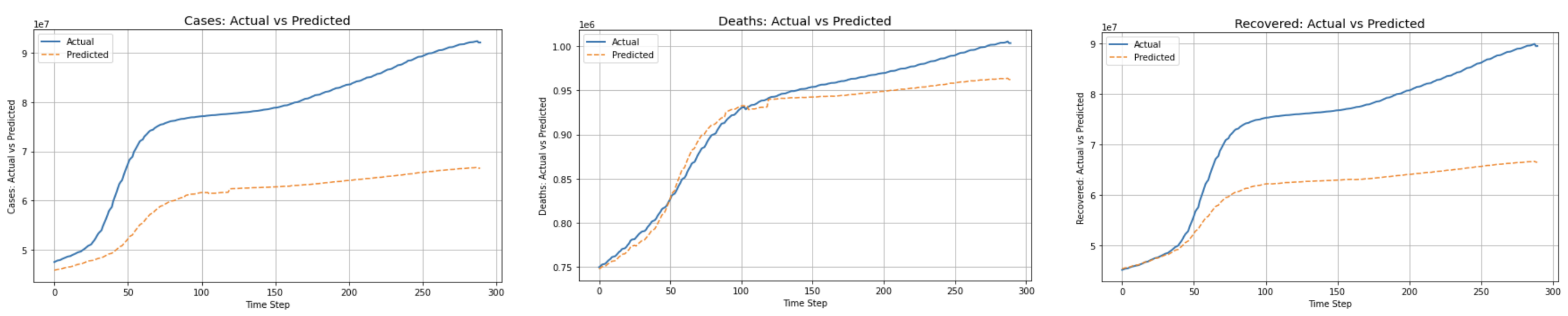
Data was pulled from the public Google COVID-19 Repository where 998 days of features were downloaded. These features included cumulative cases and deaths, daily new cases and deaths, stringency index, and cumulative fully vaccinated. Within this dataset, 5.4% of the values were missing so K-Nearest Neighbor (KNN) Imputation was applied. This imputed data was applied to both the SIRD and PINN models.

KNN Imputation: a data imputation method that fills in missing entries with the mean of the k most similar rows in the dataset.

PINN Models

Physics-Informed Neural Networks (PINNs) embed known physical laws directly into the loss function of a neural network. In this research, the SIRD differential equations act as guidelines, leading the network toward solutions that satisfy both observed data and epidemiological dynamics.

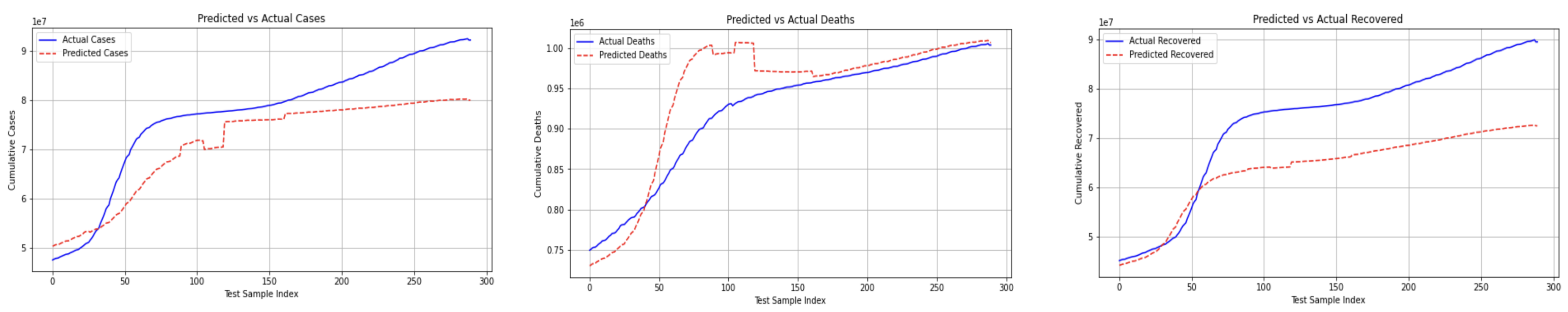
1. Scale the exogenous features: $x_{\text{scaled}} = \frac{x-\mu}{\sigma}$
2. Scale the time feature: $x_{\text{scaled}} = 2 \cdot \frac{x-x_{\min}}{x_{\max}-x_{\min}} - 1$
3. **Multi-Output PINN Model:** Predicts all 3 SIRD parameters simultaneously



Mean Absolute Percentage Error

- Cases: 20.7405% • Deaths: 1.7083% • : 16.3253%

4. **Three Separate PINN Models:** Predicts each of the SIRD parameters individually



Mean Absolute Percentage Error

- Cases: 7.98% • Deaths: 3.157% • Recovered: 12.475%

Conclusions and Future Work

The Time Varying SIRD model performed better than the PINN model. This can be due to the PINN models great sensitivity to data scaling: this includes the Data Imputation and Normalization methods used in the beginning stages. This could also be due to a lack of time for perfected hyperparameter tuning. However, disease dynamics change over time, and the results from the Time-Varying SIRD and PINNs show that.

Data/Code Availability: The code developed and used for this research is available on GitHub. https://github.com/Aleya-Hadenfeldt

References

[1] Yulia Abramova and Vasiliy Leonenko. The past helps the future: Coupling differential equations with machine learning methods to model epidemic outbreaks. In *Computational Science – ICCS 2024: 24th International Conference, Malaga, Spain, July 2–4, 2024, Proceedings, Part IV*, page 247–254, Berlin, Heidelberg, 2024. Springer-Verlag.

[2] Hyokyoung G Hong and Yi Li. Estimation of time-varying reproduction numbers underlying epidemiological processes: A new statistical tool for the covid-19 pandemic. *PloS one*, 15(7):e0236464, 2020.

[3] Seungchan Ko and Sang Hyeon Park. Vs-pinn: A fast and efficient training of physics-informed neural networks using variable-scaling methods for solving pdes with stiff behavior. *Journal of Computational Physics*, 494:112516, 2023.

[4] Sagi Shaier, Maziar Raissi, and Padmanabhan Seshaiyer. Data-driven approaches for predicting spread of infectious diseases through dinns: Disease informed neural networks, 2022.