SOLVING THE DISEASE SPREAD INVERSE PROBLEM USING SYNTHETIC TRAJECTORY DATA AND NEURAL APPROXIMATION

Anonymous authors
Paper under double-blind review

ABSTRACT

The spread of many infectious diseases is modeled using variants of the SIR compartmental model, which is a coupled differential equation. The coefficients of the SIR model determine the spread trajectories of disease, on whose basis proactive measures can be taken. Hence, the coefficient estimates must be both fast and accurate. Shaier et al. in (1) used Physics Informed Neural Networks (PINNs) to estimate the parameters of the SIR model. There are two drawbacks to this approach. First, the training time for PINNs is high, with certain diseases taking close to 90 hrs to train. Second, PINNs don't generalize for a new SIDR trajectory, and learning its corresponding SIR parameters requires retraining the PINN from scratch. In this work, we aim to eliminate both of these drawbacks. We generate a dataset between the parameters of ODE and the spread trajectories by solving the forward problem for a large distribution of parameters using the LSODA algorithm. We then use a neural network to learn the mapping between spread trajectories and coefficients of SIDR in an offline manner. This allows us to learn the parameters of a new spread trajectory without having to retrain, enabling generalization at test time. We observe a speed-up of 3-4 orders of magnitude with accuracy comparable to that of PINNs for 11 highly infectious diseases. Further finetuning of neural network inferred ODE coefficients using PINN further leads to 2-3 orders improvement of estimated coefficients.

1 INTRODUCTION

In the past century alone, mankind has been suffering from various deadly diseases, with Spanish flu (1918-1920), Hong Kong Flu (1968-1969), Ebola Virus (1976-ongoing), Middle East Respiratory Syndrome (MERS) (2012-ongoing), Ebola Virus (2014), Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) just to name a few. The recent outbreak of SARS-CoV-2 has led to a few million infections across the globe, and the importance of mathematical modeling of these infectious diseases has never been more important. In the early stages of the pandemic, mathematical models helped public health officials understand the potential impact of the disease and predict how quickly it could spread. This information was used to make informed decisions about which mitigation measures were necessary and which were most effective. Additionally, mathematical models were also used to evaluate the effectiveness of different vaccination strategies and make predictions about the future course of the pandemic. This information was critical in helping countries prioritize vaccine distribution and allocate resources.

One of the earliest methods for predicting the spread of infectious disease was formulated by the Kermack–McKendrick epidemic model (2). In this model, the population under consideration is divided into three classes, susceptible (S), infective (I), and removed (R). Over the years, the SIR model has been improved to account for hospitalizations, latent dynamics, posthumous transmissions, etc to improve the predictions of mathematical models like COVID (3), Anthrax(4), HIV(5), Zika(6), Smallpox(7), Tuberculosis(8), Pneumonia(9), Ebola(10), Dengue(11), Polio(12) and Measles(13). Shaier et al. in (1) used Physics-informed Neural Networks (14) to predict the coefficients of mathematical models in (3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13) for one specific SIR trajectory instance. The PINNs are trained in an unsupervised manner and use a physics-informed loss to solve the inverse problem of coefficient estimation of differential equations. The physics-informed loss, however, has

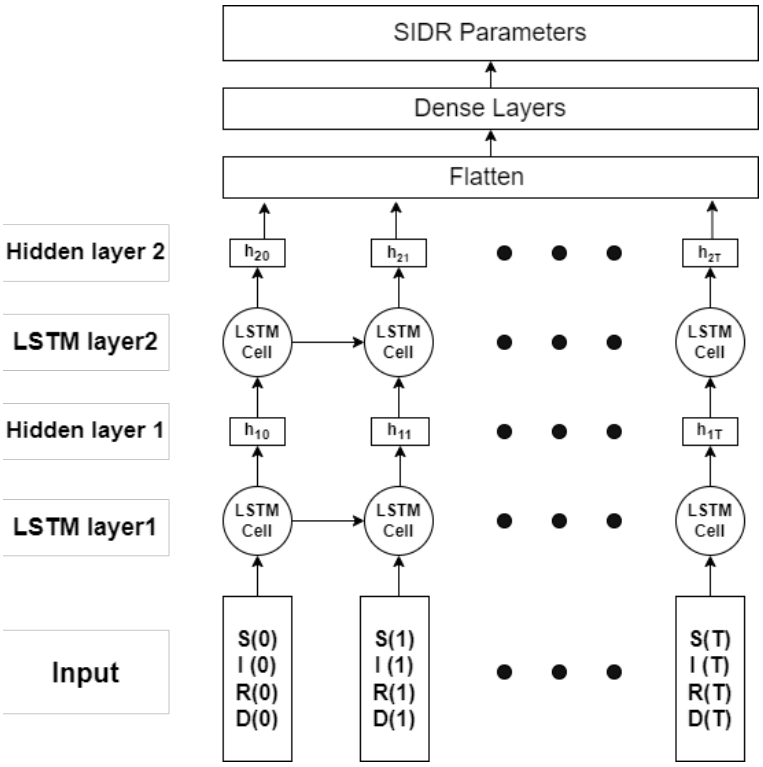

Figure 1: Two layered LSTM architecture

a complex loss landscape (15) and struggles to converge quickly, leading to larger training durations, with some diseases requiring around 90 hrs to train. Furthermore, PINNs must be retrained for a new trajectory estimation, as the underlying parameters governing the trajectory differ. In this work, we circumvent both these problems. We generate a dataset wherein we solve the computationally cheaper forward problem multiple times to generate spread trajectories, given the parameters of the SIR model. Then we use a neural network to learn the inverse mapping between trajectories and parameters. We use this trained neural network to estimate the parameters of the SIR model for a new test task, avoiding retraining. We observe a speed-up of 3-4 orders of magnitude, while parameter estimations are on par with PINNs. Further finetuning of neural network inferred parameters using PINNs outperform PINNs from scratch by 2-3 orders of magnitude at 25 times lower train time.

2 METHODOLOGY

We first generate the dataset. We solve the forward problem of generating spread trajectories, given the parameters of the disease ODEs using the LSODA (16) algorithm. The LSODA algorithm is computationally very cheap and just takes 2-3 seconds for simulating the trajectories per example. In the next step, we use a neural network to learn the mapping between spread trajectory and parameters. As the spread trajectories have temporal variations, a multi-layered LSTM (17) is the most natural choice as our choice of Neural Network. Figure 1 represents our multi-layered LSTM architecture. At every time step, we pass the spread population $[S(t),I(t),R(t),D(t)]$, where $t \in [0, T]$ as input, which is fed to the multilayered LSTM. The final hidden output layers of LSTM are flattened and passed through dense layers to obtain the final estimations of the parameters of the ODE.

3 EXPERIMENTS

Table 1 consists of the information on the differential equations governing the diseases. We normalize our inputs (trajectories) before passing through the neural network. We use a multi-layer LSTM architecture with 2 layers and 256 hidden neurons to learn the mapping between the trajectories

and the parameters, followed by flattening and 2 dense layers with 64 neurons. We consider 9000 training examples, 500 validation, and 500 test examples for every disease. Table 3 consists of the ranges of each parameter to be estimated for every disease. We use an Adam Optimizer for 60k epochs with multiplicative decay of 0.1 every 20k epochs with an initial learning rate of $1e^{-3}$. We use relative percentage-L2-error as our metric for parameters with non-zero ground-truth value, and mean-average-error for parameters having zero as their true value. All experiments were conducted on NVIDIA P100 GPU with 16 GB GPU Memory and 1.32 GHz GPU Memory clock using Pytorch framework.

4 COVID-19

In this section, we elaborate upon the mathematical model of Covid-19 and discuss the impact of its coefficients. The mathematical model governing the spread of Covid-19 (3) is modeled as follows:

$$dS/dt = -(\alpha/N)SI \tag{1a}$$
$$dI/dt = (\alpha/N)SI - \beta I - \gamma I \tag{1b}$$
$$dD/dt = \gamma I \tag{1c}$$
$$dR/dt = \beta I \tag{1d}$$

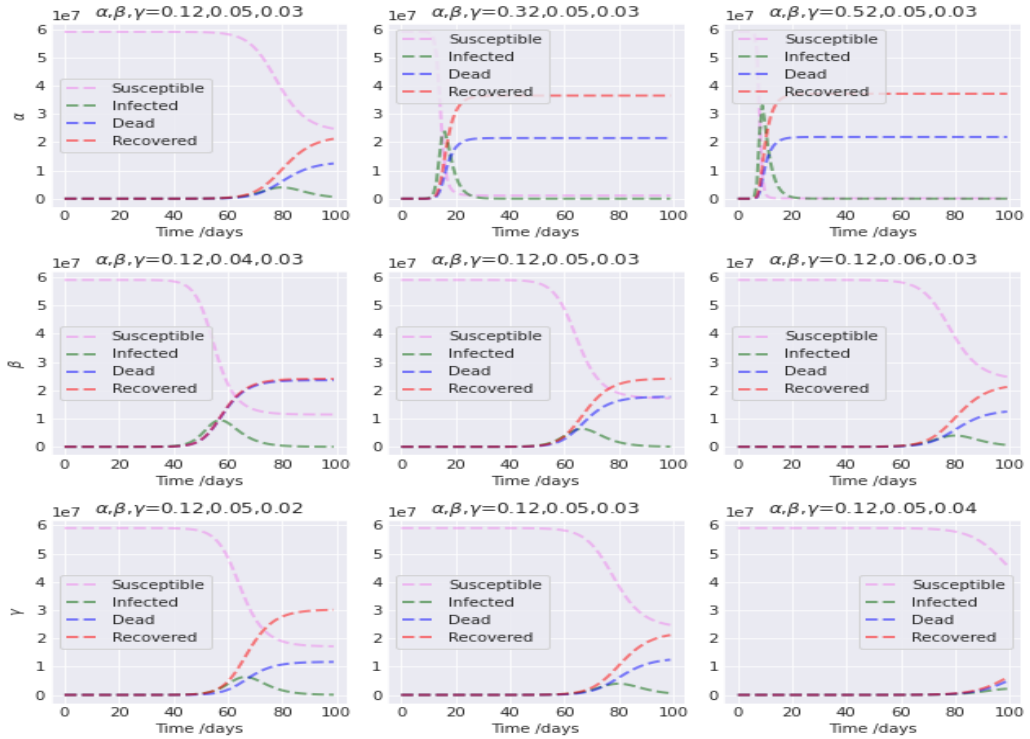

Figure 2: Influence of SIDR parameters on Covid trajectories

Here S, I, D, and R stand for susceptible, infected, dead, and recovered populations respectively. $\alpha \in [0.12, 0.52]$,$\beta \in [0.04, 0.06]$ and $\gamma \in [0.02, 0.04]$ are the parameters that control the trajectories. Figure 2 provides a visualization of how the trajectories are influenced by changes in the range of parameters, with the first, second, and third rows representing changes w.r.t α,β,γ respectively. We observe, as the value of α increases, the infection peaks in a lesser number of days, and reaches a higher peak. An increase in β and γ leads to lowering and delaying the infection and death trajectory peak. β and γ has an opposite effect as α.

	% Relative L2-error				Time statistics			
	LSTM		PINN-Finetune		LSTM		PINN	
	Mean	Std-dev	Mean	Std-dev	Training	Test	Training	Finetuning
Covid	0.7366	0.067	$1.71e^{-3}$	$3.59e^{-5}$	0.85 hrs	2 s	20 mins	11 mins
HIV	0.1691	0.0124	$1.19e^{-3}$	$6.67e^{-5}$	1.14 hrs	2 s	22 hrs	1.25 hrs
Small Pox	3.2057	1.2611	$2.44e^{-3}$	$4.47e^{-5}$	1.25 hrs	2 s	14 hrs	0.75 hrs
Tuberculosis	2.9033	0.2271	$2.62e^{-3}$	$2.76e^{-5}$	1.16 hrs	2 s	12 hrs	0.75 hrs
Pneumonia	2.2597	0.3917	$1.05e^{-3}$	$8.17e^{-5}$	1.34 hrs	2 s	41 hrs	1.25 hrs
Dengue	1.5690	0.1609	$8.41e^{-4}$	$6.24e^{-5}$	1.62 hrs	2 s	33 hrs	1.45 hrs
Ebola	2.4076	0.1741	$2.69e^{-3}$	$6.38e^{-5}$	1.46 hrs	2 s	58 hrs	1.5 hrs
Anthrax	2.0815	0.7015	$5.18e^{-4}$	$1.43e^{-5}$	2.21 hrs	2 s	91 hrs	1.5 hrs
Polio	0.2601	0.025	$3.88e^{-3}$	2.15^{-5}	1.75 hrs	2 s	66 hrs	1.5 hrs
Measles	2.4145	0.355	$5.57e^{-3}$	$1.09e^{-5}$	1.82 hrs	2 s	28 hrs	2 hrs
Zika	0.2416	0.0375	$4.29e^{-3}$	$8.33e^{-5}$	2.06 hrs	2 s	13 hrs	2.25 hrs

5 OBSERVATIONS

Table 5 summarizes the results of our experiments. We provide two results: LSTM-inferred parameters (Column 2,3) and fine-tuned parameters by PINNs using LSTM-inferred parameters (Column 4,5) as initialization. The statistics of mean and standard deviation of relative-L2-error for every disease are across all test tasks and all parameter coefficients of interest. We observe that PINN-based finetuning improves the error estimates by 2-3 orders, almost learning the true parameters perfectly, and also reduces the standard deviation by 4 orders. HIV, Small Pox has the smallest and largest LSTM mean error, while Anthrax and Measles have the smallest and largest PINN-finetuned mean errors. Columns 6-9 provide the time statistics. We observe the training time for LSTM to generalize for 9000 train tasks is approximately 18.64 times faster than the train time for PINN for just 1 task. Furthermore, as LSTMs generalize the mapping between trajectories and SIR parameters, they can directly infer the parameters of the SIR differential equation for a new set of spread trajectories without retraining, unlike PINNs. Thus, LSTMs take just 2 s to infer for a new task. Additionally, the convergence time of PINN finetuning is on average 25 times lower than PINNs from scratch, indicating the importance of better initialization to faster convergence of PINNs.

In table 3, for every disease, we provide the range of the parameter of the differential equation and its corresponding parameter-wise $\%$ relative-L2-errors for LSTM inference and PINN finetuning. For LSTM, we observe across all diseases, the parameters which higher magnitude values (in the range of $1e^0$ or higher) have a lower mean and standard deviation in contrast to parameters with low magnitude values ($1e^{-1}$ and lower). For a fair comparison with existing literature, we compare the performance of LSTM at test-time viz-a-viz with PINNs for the spread trajectories in (1), and compile its results in table 2 and Figure 3. We observe, apart from Covid, HIV, and Zika, LSTMs provide better estimates than PINNs for a majority of the parameters. Even on the parameters where LSTMs are worse (all diseases including Covid, HIV, and Zika), the estimates are not too far off from PINNs and have a negligible difference from the true trajectories, as observed in figure 3. Finetuning the parameters using PINNs on top of LSTM inference further improves the parameter estimates, as observed by 2-3 orders lower error over both LSTM inferred parameters and PINNs trained from scratch.

6 CONCLUSION

In this work, we show NNs are capable of learning the mapping between disease spread trajectories and the coefficients of the differential equation governing them. NNs can generalize for a new system of trajectories and predict the ODE coefficients accurately without requiring re-training. This eliminates the two drawbacks suffered by PINNs: Longer time and inability to generalize for a new task. Finetuning ODE parameters using PINNs leads to an even better estimate of ODE parameters at 25 times lesser train time than PINNs with the random initialization, further improving the training cost. Quicker and more accurate estimation of governing parameter coefficients allows for proactive measures like hospitalization, vaccination, or lockdowns, which can significantly help reduce human casualties.

REFERENCES

[1] Sagi Shaier and Maziar Raissi. Disease informed neural networks. *CoRR*, abs/2110.05445, 2021.

[2] William Ogilvy Kermack, A. G. McKendrick, and Gilbert Thomas Walker. 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):700–721, 1927.

[3] Jesús Fernández-Villaverde and Charles I Jones. Estimating and simulating a sird model of covid-19 for many countries, states, and cities. Working Paper 27128, National Bureau of Economic Research, May 2020.

[4] Chadi Saad-Roy, P. Driessche, and Abdul-Aziz Yakubu. A mathematical model of anthrax transmission in animal populations. *Bulletin of Mathematical Biology*, 79, 12 2016.

[5] Alan S. Perelson, Denise E. Kirschner, and Rob De Boer. Dynamics of hiv infection of cd4+ t cells. *Mathematical Biosciences*, 114(1):81–125, 1993.

[6] Daozhou Gao, Yijun Lou, Daihai He, Travis C. Porco, Yang Kuang, Gerardo Chowell, and Shigui Ruan. Prevention and control of zika as a mosquito-borne and sexually transmitted disease: A mathematical modeling analysis. *Scientific Reports*, 6(1), jun 2016.

[7] Raymond Gani and Steve Leach. Transmission potential of smallpox in contemporary populations. *Nature*, 414:748–751, 2001.

[8] Carlos Castillo-Chávez and Zhilan Feng. To treat or not to treat: The case of tuberculosis. *Journal of mathematical biology*, 35:629–56, 07 1997.

[9] Getachew Tilahun, Oluwole Makinde, and David Malonza. Modelling and optimal control of pneumonia disease with cost-effective strategies. *Journal of Biological Dynamics*, 11:1–27, 06 2017.

[10] J. Legrand, R. F. Grais, P. Y. Boelle, A. J. Valleron, and A. Flahault. Understanding the dynamics of ebola epidemics. *Epidemiology and Infection*, 135(4):610–621, 2007.

[11] Salisu Garba, Abba Gumel, and Mohd Bakar. Backward bifurcation in dengue transmission dynamics. *Mathematical biosciences*, 215:11–25, 06 2008.

[12] Svetlana Bunimovich-Mendrazitsky and Lewi Stone. Modeling polio as a disease of development. *Journal of Theoretical Biology*, 237(3):302–315, 2005.

[13] Benjamin M. Bolker and Bryan T. Grenfell. Space, persistence and dynamics of measles epidemics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 348 1325:309–20, 1995.

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

[15] Sifan Wang, Xinling Yu, and Paris Perdikaris. When and why pinns fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, page 110768, 10 2021.

[16] A. C. Hindmarsh and L. R. Petzold. Lsoda, ordinary differential equation solver for stiff or non-stiff system, Sep 2005. MATHEMATICAL METHODS AND COMPUTING.

[17] Libin Yang, Yu Zheng, Xiaoyan Cai, Hang Dai, Dejun Mu, Lantian Guo, and Tao Dai. A lstm based model for personalized context-aware citation recommendation. *IEEE Access*, 6:59618–59627, 2018.

A APPENDIX

Comparison of LSTM with PINN in literature task

			Predicted Parameter			% Relative L2-Error		
		Actual	PINN	LSTM	Finetune	PINN	LSTM	Finetune
Covid	α	0.191	0.193	0.193	0.191	1.151	0.837	$1.1e^{-4}$
	β	0.05	0.0501	0.050	0.049	0.2	0.400	$2.4e^{-4}$
	γ	0.029	0.029	0.029	0.029	1.02	1.701	$4.7e^{-4}$
HIV	s	10	10.001	10.05	10.001	0.007	0.500	$2.26e^{-3}$
	μ_T	0.02	0.021	0.0198	0.0200	3.812	4.633	$1.59e^{-2}$
	μ_I	0.26	0.261	0.260	0.257	0.489	1.154	$3.31e^{-2}$
	μ_b	0.24	0.242	0.239	0.240	0.727	0.417	$2.23e^{-3}$
	μ_V	2.4	2.419	2.407	2.400	0.829	0.292	1.15^{-3}
	r	0.03	0.031	0.030	0.03	2.016	1.333	$7.74e^{-3}$
	N	250	249.703	248.85	250.001	0.118	0.460	$2.56e^{-3}$
	$T_{\max}$	1500	$1.506e^{3}$	$1.491e^{3}$	1499.96	0.436	0.563	$1.49e^{-3}$
	k_1	$2.4e^{-5}$	$2.46e^{-5}$	2.19e-5	$2.4e^{-5}$	2.448	8.75	$9.17e^{-3}$
	k_1'	$2e^{-5}$	$2.03e^{-5}$	1.94e-5	1.99e-5	1.599	3.00	$3.75e^{-3}$
Small Pox	χ_1	0.06	0.0554	0.0539	0.5994	7.7222	11.31	$1.16e^{-2}$
	χ_2	0.04	0.0380	0.0407	0.040	4.924	1.749	$5.95e^{-3}$
	1	0.975	0.9839	0.9625	0.975	0.909	1.282	$2.69e^{-3}$
	2	0.3	0.2841	0.2951	0.3	5.285	1.633	$7.71e^{-3}$
	ρ	0.975	0.9759	0.971	0.975	0.088	0.574	$1.88e^{-2}$
	θ	0.95	0.9050	0.981	0.95	4.737	3.263	$4.61e^{-2}$
	α	0.068	0.0626	0.065	0.068	8.549	4.558	$3.85e^{-3}$
	γ	0.11	0.1034	0.111	0.11	10.859	1.181	7.24^{-2}
Tuberculosis	δ	500	509.4698	483.48	500.011	1.8587	3.3039	$1.55e^{-3}$
	β	13	12.5441	12.87	13.0004	3.6341	2.514	$1.24e^{-3}$
	c	1	1.0405	0.984	1.0003	3.8952	2.16	$1.11e^{-3}$
	μ	0.143	0.1474	0.141	0.143	3.0142	1.398	$7.84e^{-4}$
	k	0.5	0.5396	0.4985	0.5001	7.3433	0.300	$1.62e^{-4}$
	r_1	2	1.9892	2.034	1.9996	0.5424	1.69	$5.16e^{-3}$
	r_2	1	1.1243	1.0431	1.0013	11.0583	4.31	$2.97e^{-3}$
	β'	13	13.7384	12.99	13.001	5.3746	0.076	$8.68e^{-5}$
	d	0	-0.0421	0.012	$2.4e^{-3}$	0.0421	0.012	$2.4e^{-3}$
Pneumonia	π	0.01	0.0098	0.0101	0.010004	2.0032	1.000	$2.22e^{-4}$
	λ	0.1	0.0990	0.0994	0.1	0.9622	0.600	$2.22e^{-4}$
	k	0.5	0.5025	0.5086	0.5	0.5083	1.72	$5.19e^{-3}$
	χ	0.002	0.0022	0.00195	0.00201	11.6847	2.500	$4.76e^{-3}$
	τ	0.89	0.8912	0.8858	0.0890	0.1309	0.4917	$5.19e^{-3}$
	ϕ	0.0025	0.0027	0.0024	0.00250	7.4859	4.00	$3.56e^{-3}$
	χ	0.001	0.0011	0.0009	0.001	6.7374	10.00	$2.29e^{-2}$
	p	0.2	0.2033	0.1985	0.2001	1.6372	0.7500	$7.77e^{-5}$
	θ	0.008	0.0084	0.0081	0.008	14.8891	1.249	$5.95e^{-3}$
	μ	0.01	0.0092	0.0103	0.010	8.4471	2.999	$3.88e^{-3}$
	α	0.057	0.0570	0.0556	0.057	0.0005	2.456	$1.07e^{-2}$
	ρ	0.05	0.0508	0.0489	0.0499	1.5242	2.200	$1.07e^{-2}$
	β	0.0115	0.0122	0.0113	0.011501	5.8243	1.739	$1.55e^{-3}$
	η	0.2	0.2023	0.1958	0.200	1.1407	2.100	$3.55e^{-3}$
	q	0.5	0.4960	0.5037	0.5004	0.8003	0.7400	$1.15e^{-4}$
	δ	0.1	0.1038	0.1106	0.1003	3.7502	10.59	$1.35e^{-2}$
Dengue	π_b	10	9.9317	9.6514	9.9995	0.6832	0.3538	$3.19e^{-4}$
	π_v	30	29.8542	30.042	30.001	0.4859	0.1400	$1.71e^{-4}$
	λ_b	0.055	0.0552	0.0551	0.055	0.2696	0.18	$1.96e^{-5}$
	λ_v	0.05	0.0506	0.0495	00499	1.2876	1.00	$5.56e^{-4}$

	δ_b	0.99	0.9643	1.003	0.9902	2.5967	1.31	$1.02e^{-3}$
	δ_v	0.057	0.0567	0.0558	0.057	0.5294	2.1052	$4.28e^{-3}$
	μ_b	0.0195	0.0194	0.01952	0.0195	0.3835	0.1025	$6.24e^{-5}$
	μ_v	0.016	0.0159	0.0163	0.016	0.8796	1.874	$4.74e^{-3}$
	σ_b	0.53	0.5372	0.5219	0.5301	1.3567	1.528	$2.65e^{-4}$
	σ_v	0.2	0.1989	0.2004	0.2	0.5483	0.199	$1.96e^{-5}$
	τ_b	0.1	0.0902	0.1013	0.1001	9.7723	1.299	$5.27e^{-4}$
Ebola	β_1	3.532	3.5589	3.5474	2.532	0.7622	0.6317	$5.59e^{-4}$
	β_b	0.012	0.0129	0.0118	0.0120	7.8143	5.5134	$6.31e^{-4}$
	β_f	0.462	0.4638	0.4629	0.46201	0.3976	0.2784	$4.41e^{-3}$
	α	$1/12$	0.0866	0.0866	0.0833	3.9320	3.9320	$2.66e^{-3}$
	γ_b	$1/4.2$	0.2471	0.2392	0.2380	3.7853	3.6124	$7.71e^{-4}$
	θ_1	0.65	0.6523	0.6534	0.6497	0.3477	0.5624	$5.81e^{-5}$
	γ_i	0.1	0.0904	0.0906	0.0999	9.6403	9.400	$1.41e^{-3}$
	δ_1	0.47	0.4712	0.4750	0.4699	0.2565	1.0638	$8.58e^{-3}$
	γ_d	$1/8$	0.1205	0.1199	0.12499	3.6124	4.079	$2.36e^{-3}$
	δ_2	0.42	0.4247	0.4169	0.42003	1.1244	0.7380	$6.74e^{-4}$
	γ_f	0.5	0.5196	0.5074	0.4998	3.9246	1.479	$5.71e^{-4}$
	γ_{ib}	0.082	0.0811	0.0809	0.08207	1.0932	1.3416	$2.19e^{-4}$
	γ_{db}	0.07	0.0710	0.0704	0.070	0.7563	0.5714	$2.00e^{-4}$
Anthrax	r	$1/300$	0.0034	0.0032	0.00332	1.2043	1.2043	$3.27e^{-3}$
	μ	$1/600$	0.0017	0.0017	0.001667	0.5754	0.5754	$3.27e^{-3}$
	κ	0.1	0.1025	0.0999	0.100	2.5423	0.100	$4.16e^{-5}$
	η_a	0.5	0.5035	0.4979	0.5001	0.7022	0.420	$2.19e^{-4}$
	η_c	0.1	0.1024	0.1014	0.1003	2.4492	1.400	$3.55e^{-3}$
	η_i	0.01	0.0106	0.0101	0.0100	6.0459	0.999	$4.39e^{-3}$
	τ	0.1	0.0976	0.0995	0.0999	2.4492	0.500	$7.72e^{-4}$
	γ	$1/7$	0.1444	0.1429	0.14285	1.0542	0.030	$6.16e^{-3}$
	δ	$1/20$	0.0512	0.0505	0.0500	2.3508	1.002	$8.04e^{-3}$
	K	100	100.6391	98.73	100.005	0.6391	1.269	$5.18e^{-3}$
	β	0.02	0.0021	0.0199	0.0199	6.5466	0.499	$6.14e^{-3}$
	σ	0.1	0.1051	0.1006	0.1004	5.1029	0.6014	$7.44e^{-4}$
Polio	μ	0.02	0.0199	0.0200	0.020	0.0	0.4995	$5.21e^{-4}$
	α	0.5	0.5018	0.5007	0.5002	0.36	0.140	$2.41e^{-4}$
	γ_a	18	18.0246	17.9751	18.003	0.4168	0.135	$1.29e^{-4}$
	γ_c	36	36.0701	35.9541	36.007	0.3587	0.1388	$1.01e^{-4}$
	β_{aa}	40	40.2510	39.9814	40.005	0.6275	0.050	$1.37e^{-5}$
	β_{cc}	90	90.6050	89.4736	89.994	0.6722	0.5888	$8.61e^{-5}$
	β_{ac}	0	0.0002	0.0019	$1.7e^{-5}$	0.0002	0.0019	$1.7e^{-5}$
	β_{ca}	0	0.0004	-0.0004	$2.1e^{-6}$	0.0004	0.0004	$2.1e^{-6}$
Measles	μ	0.02	0.0225	0.0198	0.0200	12.704	0.999	$3.65e^{-3}$
	β_1	0.28	0.2700	0.2633	0.2801	3.5704	5.9642	$4.79e^{-3}$
	γ	100	97.0001	100.59	100.003	2.9999	0.5900	$7.87e^{-4}$
	σ	35.84	34.7127	36.18	36.007	3.1453	0.9486	$4.53e^{-3}$
Zika	a	0.5	0.5116	0.4997	0.5001	0.0588	2.32	$3.89e^{-2}$
	b	0.4	0.4033	0.3965	0.4	0.8297	1.6865	$7.36e^{-3}$
	c	0.5	0.5015	0.4986	0.5001	0.3086	0.2800	$7.66e^{-3}$
	η	0.1	0.0999	0.0963	0.1007	0.0687	3.6036	$2.98e^{-2}$
	β	0.05	0.0498	0.0502	0.05	0.4098	0.4098	$7.31e^{-3}$
	κ	0.6	0.6033	0.5967	0.6004	0.5486	0.5486	$2.99e^{-2}$
	τ	0.3	0.2902	0.2969	0.3	3.2565	2.3087	$1.72e^{-2}$
	θ	18	17.9669	18.821	18.006	0.1838	4.5611	$1.04e^{-2}$
	m	5	5.2937	4.9621	5.0009	5.8748	0.7600	$5.99e^{-3}$
	V_b	$1/5$	0.1996	0.2014	0.2003	0.1798	0.6999	$4.85e^{-3}$
	V_v	10	10.0170	9.965	10.001	0.1700	0.3049	$2.77e^{-4}$
	γ_{b1}	$1/5$	0.1991	0.2009	0.2006	0.4651	0.449	$1.96e^{-3}$

γ_{b2}	$1/20$	0.0504	0.0497	0.05	0.7261	0.6003	$2.17e^{-3}$
γ_b	$1/7$	0.1406	0.1412	0.1427	1.5967	1.16	$8.94e^{-3}$
μ_v	$1/14$	0.0723	0.0725	0.0714	1.1806	1.499	$5.25e^{-3}$

Parameter-wise statistics for inferred-LSTM and finetuned-PINN

			LSTM-inference		finetuned-PINN	
	Parameter	Range	Mean	Std. Dev	Mean	Std. Dev
Covid	α	$(0.12, 0.52)$	1.04	0.15	$1.69e^{-3}$	$3.64e^{-5}$
	β	$(0.04, 0.06)$	0.18	0.03	$1.65e^{-3}$	$3.45e^{-5}$
	γ	$(0.02, 0.04)$	0.99	0.02	$1.77e^{-3}$	$3.68e^{-5}$
HIV	s	$(5, 15)$	0.0109	0.008	$1.17e^{-3}$	$6.67e^{-5}$
	μ_T	$(0.005, 0.045)$	2.5614	0.3917	$1.18e^{-3}$	$6.64e^{-5}$
	μ_I	$(0.05, 0.45)$	0.5564	0.0318	$1.25e^{-3}$	$6.64e^{-5}$
	μ_b	$(0.05, 0.45)$	0.7548	0.0147	$1.22e^{-3}$	$6.68e^{-5}$
	μ_V	$(2, 3)$	0.8767	0.0435	$1.20e^{-3}$	$6.75e^{-5}$
	r	$(0.01, 0.05)$	2.1614	0.1947	$1.16e^{-3}$	$6.71e^{-5}$
	N	$(225, 275)$	0.0878	0.0314	$1.18e^{-3}$	$6.60e^{-5}$
	$T_{\max}$	$(1400, 1600)$	0.4519	0.041	$1.23e^{-3}$	$6.80e^{-5}$
	k_1	$\left(2e^{-5}, 3e^{-5}\right)$	2.031	0.414	$1.24e^{-3}$	$6.59e^{-5}$
	k_1'	$\left(1.5e^{-5}, 2.5e^{-5}\right)$	1.6514	0.0056	$1.11e^{-3}$	$6.61e^{-5}$
Small Pox	χ_1	$(0.04, 0.08)$	4.4817	1.0814	$2.44e^{-3}$	$4.46e^{-5}$
	χ_2	$(0.02, 0.06)$	2.136	0.465	$2.49e^{-3}$	$4.58e^{-5}$
	1	$(0.75, 1.25)$	0.775	0.064	$2.39e^{-3}$	$4.39e^{-5}$
	2	$(0.1, 0.5)$	3.3954	1.0314	$2.41e^{-3}$	$4.35e^{-5}$
	ρ	$(0.9, 1.1)$	0.619	0.155	$2.47e^{-3}$	$4.52e^{-5}$
	θ	$(0.5, 1.5)$	2.235	1.16	$2.44e^{-3}$	$4.46e^{-5}$
	α	$(0.02, 0.1)$	4.175	2.19	$2.43e^{-3}$	$4.39e^{-5}$
	γ	$(0.05, 0.15)$	4.523	2.74	$2.45e^{-3}$	$4.58e^{-5}$
Tuberculosis	δ	$(450, 550)$	1.1936	0.043	$2.65e^{-3}$	$2.75e^{-5}$
	β	$(5, 25)$	2.5617	0.0716	$2.60e^{-3}$	$2.77e^{-5}$
	c	$(0.3, 1.7)$	2.5514	0.3926	$2.63e^{-3}$	$2.76e^{-5}$
	μ	$(0.08, 0.22)$	2.9147	0.1007	$2.63e^{-3}$	$2.77e^{-5}$
	k	$(0.05, 1.2)$	7.0414	0.3591	$2.64e^{-3}$	$2.74e^{-5}$
	r_1	$(0.5, 5.5)$	0.5115	0.0321	$2.62e^{-3}$	$2.79e^{-5}$
	r_2	$(0.4, 2.4)$	5.3147	0.3628	$2.62e^{-3}$	$2.75e^{-5}$
	β'	$(5, 20)$	4.0225	0.7762	$2.61e^{-3}$	$2.78e^{-5}$
	d	$(-1, 1)$	0.0184	0.0065	$2.64e^{-3}$	$2.70e^{-5}$
Pneumonia	π	$(0.005, 0.015)$	1.165	0.134	$1.05e^{-3}$	$8.18e^{-5}$
	λ	$(0.05, 0.15)$	0.6514	0.2014	$1.04e^{-3}$	$8.14e^{-5}$
	k	$(0.3, 0.7)$	0.5017	0.0091	$1.04e^{-3}$	$8.11e^{-5}$
	χ	$(0.0005, 0.0045)$	6.6911	1.5614	$1.04e^{-3}$	$8.18e^{-5}$
	τ	$(0.6, 1.2)$	0.0514	0.0412	$1.04e^{-3}$	$8.16e^{-5}$
	ϕ	$(0.0005, 0.0045)$	5.218	0.631	$1.07e^{-3}$	$8.15e^{-5}$
	χ	$(0.0005, 0.0015)$	1.054	0.365	$1.06e^{-3}$	$8.21e^{-5}$
	p	$(0.1, 0.3)$	1.555	0.008	$1.06e^{-3}$	$8.14e^{-5}$
	θ	$(0.005, 0.0011)$	3.871	0.514	$1.07e^{-3}$	$8.20e^{-5}$
	μ	$(0.005, 0.015)$	5.591	1.135	$1.04e^{-3}$	$8.13e^{-5}$
	α	$(0.04, 0.08)$	0.0004	0.0001	$1.05e^{-3}$	$8.11e^{-5}$
	ρ	$(0.01, 0.09)$	1.835	0.3217	$1.04e^{-3}$	$8.17e^{-5}$
	β	$(0.0085, 0.0145)$	3.162	0.821	$1.08e^{-3}$	$8.19e^{-5}$
	η	$(0.1, 0.3)$	0.756	0.026	$1.07e^{-3}$	$8.17e^{-5}$
	q	$(0.3, 0.7)$	0.642	0.115	$1.05e^{-3}$	$8.22e^{-5}$
	δ	$(0.05, 0.15)$	3.311	0.515	$1.08e^{-3}$	$8.09e^{-5}$
	π_b	$(5, 15)$	0.4334	0.1124	$8.40e^{-4}$	$6.24e^{-5}$

Dengue	π_v	$(25, 35)$	0.4251	0.0514	$8.44e^{-4}$	$6.16e^{-5}$
	λ_b	$(0.01, 0.09)$	0.3014	0.0314	$8.34e^{-4}$	$6.17e^{-5}$
	λ_v	$(0.01, 0.09)$	1.1129	0.128	$8.38e^{-4}$	$6.26e^{-5}$
	δ_b	$(0.75, 1.25)$	2.336	0.217	$8.48e^{-4}$	$6.33e^{-5}$
	δ_v	$(0.04, 0.08)$	0.497	0.025	$8.37e^{-4}$	$6.29e^{-5}$
	μ_b	$(0.0175, 0.0225)$	0.333	0.117	$8.51e^{-4}$	$6.27e^{-5}$
	μ_v	$(0.006, 0.026)$	0.8325	0.0469	$8.43e^{-4}$	$6.22e^{-5}$
	σ_b	$(0.4, 0.8)$	1.265	0.4112	$8.36e^{-4}$	$6.26e^{-5}$
	σ_v	$(0.05, 0.45)$	0.6317	0.072	$8.39e^{-4}$	$6.20e^{-5}$
	τ_b	$(0.05, 0.15)$	9.092	0.464	$8.22e^{-4}$	$6.31e^{-5}$
Ebola	β_1	$(3, 4)$	0.6525	0.0759	$2.69e^{-3}$	$6.45e^{-5}$
	β_b	$(0.005, 0.015)$	4.165	1.114	$2.65e^{-3}$	$6.39e^{-5}$
	β_f	$(0.3, 0.7)$	0.2256	0.1035	$2.71e^{-3}$	$6.41e^{-5}$
	α	$(0.05, 0.15)$	3.265	0.551	$2.74e^{-3}$	$6.37e^{-5}$
	γ_b	$(0.05, 0.45)$	3.157	0.556	$2.67e^{-3}$	$6.32e^{-5}$
	θ_1	$(0.4, 0.8)$	0.3911	0.393	$2.67e^{-3}$	$6.33e^{-5}$
	γ_i	$(0.05, 0.15)$	8.815	0.45	$2.69e^{-3}$	$6.38e^{-5}$
	δ_1	$(0.3, 0.9)$	0.3161	0.1014	$2.70e^{-3}$	$6.38e^{-5}$
	γ_d	$(0.1, 0.3)$	3.321	0.469	$2.75e^{-3}$	$6.47e^{-5}$
	δ_2	$(0.2, 0.8)$	1.078	0.1015	$2.64e^{-3}$	$6.46e^{-5}$
	γ_f	$(0.2, 0.8)$	3.963	0.062	$2.66e^{-3}$	$6.38e^{-5}$
	γ_{ib}	$(0.05, 0.15)$	1.1374	0.0521	$2.72e^{-3}$	$6.35e^{-5}$
	γ_{db}	$(0.03, 0.12)$	0.8122	0.081	$2.70e^{-3}$	$6.33e^{-5}$
Anthrax	r	$(0.002, 0.004)$	1.016	0.085	$5.18e^{-4}$	$1.49e^{-5}$
	μ	$(0.001, 0.002)$	0.6951	0.1524	$5.16e^{-4}$	$1.44e^{-5}$
	κ	$(0.75, 1.25)$	2.2813	0.3621	$5.21e^{-4}$	$1.36e^{-5}$
	η_a	$(0.25, 0.75)$	0.8561	0.2003	$5.19e^{-4}$	$1.34e^{-5}$
	η_c	$(0.05, 0.15)$	2.1536	0.2718	$5.16e^{-4}$	$1.42e^{-5}$
	η_i	$(0.005, 0.015)$	3.9365	1.5216	$5.28e^{-4}$	$1.38e^{-5}$
	τ	$(0.05, 0.15)$	1.1954	0.2008	$5.05e^{-4}$	$1.35e^{-5}$
	γ	$(0.05, 0.25)$	1.773	0.616	$5.11e^{-4}$	$1.46e^{-5}$
	δ	$(0.01, 0.11)$	2.2856	0.1062	$5.31e^{-4}$	$1.38e^{-5}$
	K	$(75, 125)$	0.6117	0.0352	$5.26e^{-4}$	$1.43e^{-5}$
	β	$(0.001, 0.003)$	2.9114	2.2214	$5.17e^{-4}$	$1.49e^{-5}$
	σ	$(0.05, 0.15)$	3.1815	1.774	$5.22e^{-4}$	$1.49e^{-5}$
Polio	μ	$(0.01, 0.03)$	0.0051	0.0018	$3.76e^{-3}$	$2.15e^{-5}$
	α	$(0.25, 0.75)$	0.242	0.071	$3.55e^{-3}$	$2.05e^{-5}$
	γ_a	$(10, 30)$	0.3915	0.018	$4.04e^{-3}$	$2.25e^{-5}$
	γ_c	$(20, 60)$	0.3714	0.033	$4.15e^{-3}$	$2.13e^{-5}$
	β_{aa}	$(25, 65)$	0.6116	0.0254	$3.72e^{-3}$	$2.12e^{-5}$
	β_{cc}	$(75, 125)$	0.6115	0.0515	$3.85e^{-3}$	$2.14e^{-5}$
	β_{ac}	$(-0.5, 0.5)$	0.0003	0.0001	$3.91e^{-3}$	$2.11e^{-5}$
	β_{ca}	$(-0.5, 0.5)$	0.0003	0.0002	$3.86e^{-3}$	$2.19e^{-5}$
Measles	μ	$(0.01, 0.03)$	1.514	0.158	$5.51e^{-3}$	$1.14e^{-5}$
	β_1	$(0.15, 0.55)$	2.614	0.723	$5.61e^{-3}$	$1.07e^{-5}$
	γ	$(75, 125)$	2.515	0.365	$5.40e^{-3}$	$1.17e^{-5}$
	σ	$(20, 50)$	3.015	0.176	$5.52e^{-3}$	$1.04e^{-5}$
Zika	a	$(0.3, 0.7)$	0.0471	0.0115	$4.29e^{-3}$	$8.02e^{-5}$
	b	$(0.2, 0.6)$	0.621	0.153	$4.31e^{-3}$	$8.40e^{-5}$
	c	$(0.3, 0.7)$	0.1757	0.106	$4.31e^{-3}$	$8.47e^{-5}$
	η	$(0.05, 0.15)$	0.0555	0.012	$4.33e^{-3}$	$8.51e^{-5}$
	β	$(0.045, 0.055)$	0.3516	0.0514	$4.21e^{-3}$	$8.17e^{-5}$
	κ	$(0.2, 1.0)$	0.5115	0.0341	$4.22e^{-3}$	$8.30e^{-5}$
	τ	$(0.1, 0.5)$	2.222	0.695	$4.25e^{-3}$	$8.35e^{-5}$
	θ	$(10, 30)$	0.3014	0.1112	$4.37e^{-3}$	$8.37e^{-5}$
	m	$(3, 7)$	4.935	0.685	$4.24e^{-3}$	$8.15e^{-5}$

V_b	$(0.175, 0.225)$	0.1762	0.085	$4.26e^{-3}$	$8.23e^{-5}$
V_v	$(7.5, 12.5)$	0.159	0.014	$4.39e^{-3}$	$8.33e^{-5}$
γ_{b1}	$(0.1, 0.3)$	0.4425	0.0319	$4.28e^{-3}$	$8.23e^{-5}$
γ_{b2}	$(0.03, 0.07)$	0.358	0.264	$4.26e^{-3}$	$8.36e^{-5}$
γ_b	$(0.12, 0.18)$	1.081	0.362	$4.26e^{-3}$	$8.41e^{-5}$
μ_v	$(0.05, 0.09)$	1.355	0.415	$4.33e^{-3}$	$8.43e^{-5}$

HIV	$dT/dt = s - \mu_T T + rT\left(1 - ((T+I)/T_{\max}) - k_1 VT\right)$ $dI/dt = k_1' VT - \mu_I I$ $dV/dt = N\mu_b I - k_1 VT - \mu_V V$
Small-Pox	$dS/dt = \chi_1\left(1 - {}_1\right) Ci - \beta(\phi + \rho - \phi\rho)SI$ $dEn/dt = \beta\phi(1-\rho)SI - \alpha En$ $dEi/dt = \beta\phi\rho SI - \left(\chi_{12} + \alpha\left(1 - {}_2\right)\right) Ei$ $dCi/dt = \beta\rho(1-\phi)SI - \chi_1 Ci$ $dI/dt = \alpha(1-\theta)En - (\theta+\gamma)I$ $dQ/dt = \alpha\left(1 - {}_2\right) Ei + \theta(\alpha En + I) - \chi_2 Q$ $dU/dt = \gamma I + \chi_2 Q$ $dV/dt = \chi_1\left({}_2Ei + {}_1Ci\right)$
Tuberculosis	$dS/dt = \delta - \beta cSI/N - \mu S$ $dL/dt = \beta cSI/N - \left(\mu + k + r_1\right) L + \beta' cT/N$ $dI/dt = kL - (\mu + d)I - r_2 I$ $dT/dt = r_1 L + r_2 I - \beta' cT/N - \mu T$
Pneumonia	$dS/dt = (1-p)\pi + \phi V + \delta R - (\mu + \lambda + \theta)S$ $dV/dt = p\pi + \theta S - (\mu + \lambda + \phi)V$ $dC/dt = \rho\lambda S + \rho\lambda V + (1-q)\eta I - (\mu + \beta + \chi)C$ $dI/dt = (1-\rho)\lambda S + (1-\rho)\lambda V + \chi C - (\mu + \alpha + \eta)I$ $dR/dt = \beta C + q\eta I - (\mu + \delta)R$
Dengue	$dSb/dt = \pi_b - \lambda_b Sb - \mu_b Sb$ $dEb/dt = \lambda_b Sb - \left(\sigma_b \mu_b\right) Eb$ $dIb/dt = \sigma_b Eb - \left(\tau_b + \mu_b + \delta_b\right) Ib$ $dRb/dt = \tau_b Ib - \mu_b Rb$ $dSv/dt = \pi_v - \delta_v Sv - \mu_v Sv$ $dEv/dt = \delta_v Sv - \left(\sigma_v + \mu_v\right) Ev$ $dIv/dt = \sigma_v Ev - \left(\mu_v + \delta_v\right) Iv$
Ebola	$dS/dt = -1/N\left(\beta_1 SI + \beta_b SH + \beta_f SF\right)$ $dE/dt = 1/N\left(\beta_1 SI + \beta_b SH + \beta_f SF\right) - \alpha E$ $dI/dt = \alpha E - \left(\gamma_b\theta_1 + \gamma_i\left(1-\theta_1\right)\left(1-\delta_1\right) + \gamma_d\left(1-\theta_1\right)\delta_1\right) I$ $dH/dt = \gamma_b\theta_1 I - \left(\gamma_d b\delta_2 + \gamma_i b\left(1-\delta_2\right)\right) H$ $dF/dt = \gamma_d\left(1-\theta_1\right)\delta_1 I + \gamma_d b\delta_2 H - \gamma_f F$ $dR/dt = \gamma_i\left(1-\theta_1\right)\left(1-\delta_1\right) I + \gamma_i b\left(1-\delta_2\right) H + \gamma_f F$
Anthrax	$dS/dt = r(S+I)(1-(S+I)/K) - \eta_a AS - \eta_c SC - \eta_i(SI)/(S+I) - \mu S + \tau I$ $dI/dt = \eta_a AS + \eta_c SC + \left(\eta_i(SI)/(S+I) - (\gamma + \mu + \tau)\right) I$ $dA/dt = -\sigma A + \beta C$ $dC/dt = (\gamma + \mu)I - \delta(S+I)C - \kappa C$
Polio	$dSc/dt = \mu N - \left(\alpha + \mu + \left(\beta_{cc}/Nc\right) Ic + \left(\beta_{ca}/Nc\right) Ia\right) Sc$ $dSa/dt = \alpha Sc - \left(\mu + \left(\beta_{aa}/Na\right) Ia + \left(\beta_{ac}/Na\right) Ic\right) Sa$ $dIc/dt = \left(\left(\beta_{cc}/Nc\right) Ic + \left(\beta_{ca}/Nc\right) Ia\right) Sc - \left(\gamma_c + \alpha + \mu\right) Ic$ $dIa/dt = \left(\left(\beta_{ac}/Na\right) Ic + \left(\beta_{aa}/Na\right) Ia\right) Sa - \left(\gamma_a + \mu\right) Ia + \alpha Ic$ $dRc/dt = \gamma_c Ic - \mu Rc - \alpha Rc$ $dRa/dt = \gamma_a Ia - \mu Ra + \alpha Rc$
Measles	$dS/dt = \mu(N-S) - (\beta SI)/N$ $dE/dt = (\beta SI)/N - (\mu\sigma)E$ $dI/dt = \sigma E - (\mu + \gamma)I$
Zika	$dSb/dt = -ab(Iv/Nb)Sb - \beta\left(\left(\kappa Eb + Ib_1 + \tau Ib_2\right)/Nb\right) Sb$ $dEb/dt = \theta\left(-ab(Iv/Nb)Sb - \beta\left(\left(\kappa Eb + Ib_1 + \tau Ib_2\right)/Nb\right) Sb\right) - V_b Eb$ $dIb_1/dt = V_b Eb - \gamma_{b1} Ib_1$ $dIb_2/dt = \gamma_{b1} Ib_1 - \gamma_{b2} Ib_2$ $dAb/dt = (1-\theta)\left(ab(Iv/Nb)Sb - \beta\left(\left(\kappa Eb + Ib_1 + \tau Ib_2\right)/Nb\right) Sb\right) - \gamma_b Ab$ $dRb/dt = \gamma_{b2} Ib_2 + \gamma_b Ab$ $dSv/dt = \mu_v Nv - ac\left(\left(\eta Eb + Ib_1\right)/Nb\right) Sv - \mu_v Sv$ $dEv/dt = ac\left(\left(\eta Eb + Ib_1\right)/Nb\right) - \left(V_v + \mu_v\right) Ev$ $dIv/dt = V_v Ev - \mu_v Iv$

Table 1: Differential Equations governing the diseases

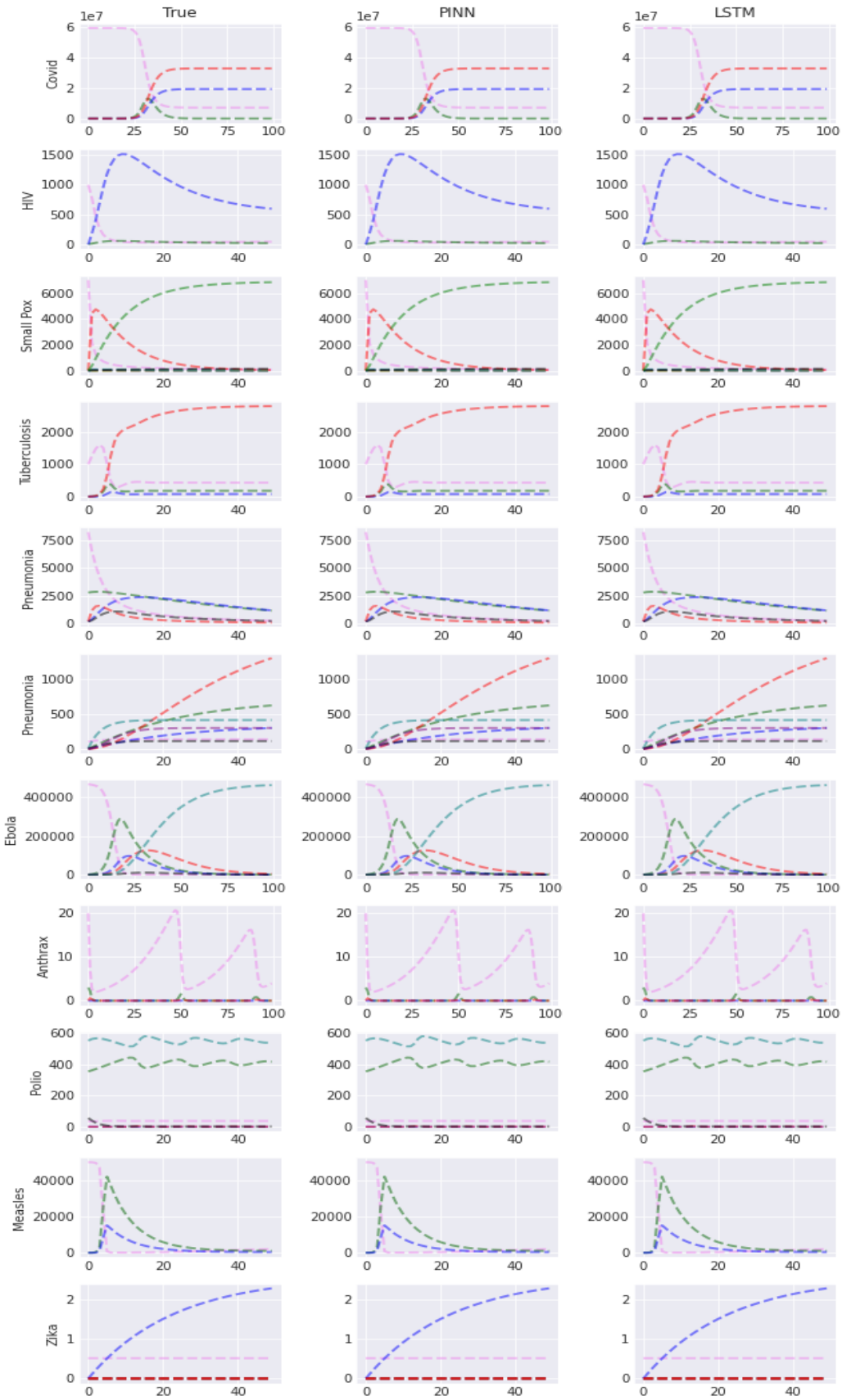

Figure 3: Comparison of True, PINN, LSTM trajectories on benchmark task