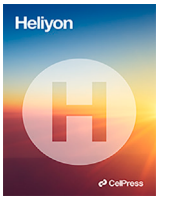
Heliyon 10 (2024) e25363

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Predicting Covid-19 pandemic waves with biologically and behaviorally informed universal differential equations

Bruce Kuwahara, Chris T. Bauch *

Department of Applied Mathematics, University of Waterloo, 200 University Ave West, Waterloo, Ontario, Canada

ARTICLE INFO

Keywords:
Machine learning
Compartmental model
Covid-19
Physics informed neural networks

ABSTRACT

During the COVID-19 pandemic, it became clear that pandemic waves and population responses were locked in a mutual feedback loop in a classic example of a coupled behavior-disease system. We demonstrate for the first time that universal differential equation (UDE) models are able to extract this interplay from data. We develop a UDE model for COVID-19 and test its ability to make predictions of second pandemic waves. We find that UDEs are capable of learning coupled behavior-disease dynamics and predicting second waves in a variety of populations, provided they are supplied with learning biases describing simple assumptions about disease transmission and population response. Though not yet suitable for deployment as a policy-guiding tool, our results demonstrate potential benefits, drawbacks, and useful techniques when applying universal differential equations to coupled systems.

1. Introduction

The COVID-19 pandemic generated an enormous demand for mathematical models. Models were developed to simulate and predict environmental effects [1,2], understand economic trade-offs, [3–5], and most prominently, to predict cases and guide policy [6–8]. Disease transmission models were often mechanistic in nature, seeking to represent known or hypothesized transmission mechanisms in a mathematical, stochastic, or agent-based framework [9–11].

The advance warning provided by these models allowed public health institutions to prepare by implementing policies to mitigate the second wave when it arrived [12,13]. Modeling efforts were widely applied to investigate the impact of public health measures such as testing [14], school and workplace closures [15,16], vaccination strategies [17–20], or to stimulate action by projecting the impacts of a worst-case 'do nothing' scenario where governments and populations did not attempt to mitigate the pandemic [13,15].

Fortunately, most governments and members of the public did respond to the pandemic by taking measures to reduce case incidence. Numerous studies have shown that non-pharmaceutical interventions such as lockdowns, school closures, and social distancing protocols reduce case notifications and health impacts of COVID-19 [21–24]. The anticipation of infection risk in the face of rising case incidence supports adherence to these measures [25]. However, these prophylactic measures are economically costly and mentally fatiguing [26,27]. So, as the risk of infection wanes, the public's willingness to abide by them wanes as well.

The ensuing relaxation of COVID mitigation efforts may potentially result in another pandemic wave. This two-way interaction – where infection spread influences behavior, which in turn influences infection spread, – suggests that the concept of coupled behavior-disease systems [28] may be useful for studying COVID-19 pandemic waves.

* Corresponding author.
E-mail address: cbauch@uwaterloo.ca (C.T. Bauch).

https://doi.org/10.1016/j.heliyon.2024.e25363
Received 28 July 2023; Received in revised form 29 December 2023; Accepted 25 January 2024

Available online 5 February 2024

As such, many mechanistic models informed by economic, social, or psychological assumptions have incorporated behavior-disease dynamics to study the impact of interventions in the context of population behavioral feedbacks [29–35], including for COVID-19 [36–39]. The movement of people within and between communities has been particularly useful for this purpose [17, 40].

Among the most valuable insights provided by these models is the occurrence of multiple pandemic waves, which are predicted under a wide range of conditions by these models due to waning stringency causing a resurgence of the infection [41,17,42,43]. With hindsight, we can confirm that these models were correct – second waves occurred virtually everywhere during the COVID-19 pandemic (and did so before the arrival of new variants).

Alongside these mechanistic models, the plethora of epidemiological, sociological, and economic data generated by the pandemic allowed machine learning models to flourish [44–48]. These models have proven adept at integrating vast quantities of data on a multitude of factors (including behavior) affecting disease spread. Consequently, they often adapt better to regional variability compared to mechanistic models [48,47]. However, machine learning models have significant drawbacks. They can fit existing data well and accurately predict days to a couple of weeks into the future, but pay for this predictive accuracy with reduced interpretability compared to traditional models [44]. Compared to mechanistic models with relatively few easily understood parameters, it is far more difficult to extract qualitative understandings of disease dynamics (such as second waves) from the hundreds or thousands of parameters in purely machine-learning models. They are also easy to over-fit (although mechanistic models also suffer from this risk), meaning their predictive value may be limited.

Recently, advances in high-performance automatic differentiation have enabled new techniques that combine the interpretability and qualitative understanding from mechanistic models with potentially higher predictive power and scalability of machine learning. Physics-informed machine learning (PIML) is one such methodology. The key idea is to create ML models that encode physical laws by inferring them from large amounts of data (observational bias), building them into the model's architecture (structural bias), or training the model to uphold them (learning bias) [49].

Of particular interest for qualitative epidemic modeling are the latter two biases, as they reduce the model's reliance on large amounts of data. By incorporating these biases, the model is prevented (in the case of structural biases) or at least discouraged (for learning biases) from making biologically impossible predictions such as negative population sizes or proportions that do not sum to unity. Learning biases can also discourage overfitting the data by introducing other objectives for the model.

Thus far, learning biases have primarily been limited to solving various forms of partial differential equations (PDEs) [50–52]. In these models, a neural network is trained to simultaneously fit data and to satisfy a PDE. In addition to physics, learning biases have been used in biologically informed machine learning (BIML) applications. These include blood flow dynamics [53], drug responses [54], and cancer detection and classification [55–57].

In terms of structural bias, universal differential equations (UDEs) have recently emerged as a method of interest. UDEs involve training neural networks embedded in differential equation models. Known dynamics can be included explicitly while leaving unknown processes to be learned by the neural network [58]. The explicit parts of the UDE can be made to retain valuable laws such as invariant quantities. UDEs have been applied successfully on predator-prey models, metabolic networks, batteries, and photonics [58,59]. For instance, recent research uses a neural network to learn the change in COVID-19 quarantine measures in a population over time, within the framework of a modified QSEIR (quarantine, susceptible, infectious, recovered) model. The trained network was then used to quantify the effectiveness of those measures for different regions [46,60].

UDEs and learning biases both have a promising track record in these contexts, but their ability to make qualitative long-term predictions about coupled behavior-disease dynamics (of the sort provided by mechanistic models) has not yet been widely tested. In fact, to our knowledge, learning biases and UDEs have yet to be combined at all.

This research gap motivated our study. Our objective was to combine observational biases (UDEs) with satisfiable learning biases in a coupled behavior-disease model for COVID-19. We trained a compartmental UDE model to fit behavioral and epidemic data while penalizing deviations from several simple socio-biological assumptions. We hypothesized that a UDE model can learn the pattern of coupled behavior-disease interactions and hence predict a second wave (either qualitatively or quantitatively), having only seen the first wave (and its learning biases). We also hypothesized that without those learning biases, the model will learn much less effectively. We note that mechanistic epidemic models [10] commonly predict a second wave of COVID-19 if the modeler imposes an increase in the contact rate parameter after the first wave, to replicate the effect of relaxing restrictions [61]. In contrast, here we are interested in the more challenging problem of endogenizing the decision to relax restrictions by using a coupled behavior-disease dynamical framework that is intended to predict decision-making regarding COVID-19 restrictions, along with the resulting changes in the contact rates.

2. Results

A complete description of our model appears in the Materials and Methods section. A complicated mathematical model can easily be made to fit an epidemic curve, but runs the risk of over-fitting the data and thus not being useful for prediction [62,63]. Simpler mathematical models allow us to test our hypotheses by incorporating aspects we understand without becoming overburdened by details that we cannot reliably describe mathematically [63].

Hence, we used a UDE framework that allows us to leave the coupled behavior-disease dynamics of a simple compartmental behavior-disease model unspecified, save for a few plausible assumptions ("learning biases"). In doing so, we can test the validity of those assumptions. Compartmental models divide the human population into mutually exclusive compartments based on infec-

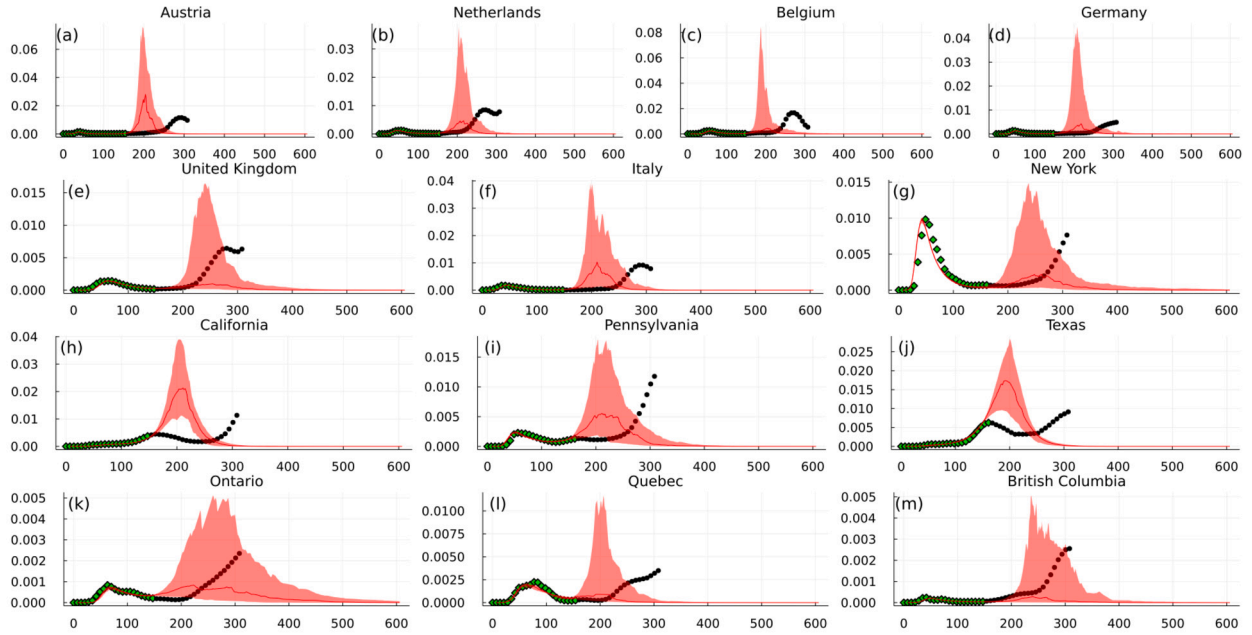


Fig. 1. Infection prevalence time series predictions for all regions produced by the model with learning biases. Infection prevalence is the proportion of the total population that is infected at any given time. Green dots represent training data (first 22 weeks) and black dots show unseen data (a further 23 weeks). Predictions are generated using the median (solid line) and interquartile range (ribbon) of 100 independently trained instances of the model per region. Note that many of these populations (e.g. Texas, California) had larger first waves than is apparent in these plots, on account of high under-reporting rates in the first wave. In all regions, the model fits training data well. It frequently predicts a second wave in all regions except Ontario and British Columbia, in which it predicts greater continuation of the first wave.

tious status, and which are generally implemented as differential equations. These models have been a mainstay of mathematical epidemiology for decades [10].

The algorithm learned the manner in which the force of infection responds to mobility and manner in which mobility responds to its current value, the number of active cases, recent new cases, and recovered cases. The learning biases inform the model with several plausible assumptions: namely, that force of infection increases with mobility, that mobility decreases with more active and recent cases, that mobility tends toward 0 (the pre-COVID average) in the absence of cases, that this tendency is stronger the more people have recovered, and that mobility cannot fall below a 100% reduction or exceed a 200% increase from the pre-COVID average (see Methods). The learning biases strongly discourage infeasible values of mobility and make data-fitting relatively less important for the optimizer. As a result, the model makes out-of-sample predictions (i.e. second waves) frequently.

To ensure consistent and repeatable results, we ran the model on each region 100 times both with and without learning biases. We trained the algorithm on the first wave and tested whether it could predict the second wave. Overall, the model with learning biases was successful in every region in which we tested it, though some more so than others. It consistently learned to fit the data and constraint losses, predicted second waves, and seldom made biologically implausible predictions.

We compared UDE models with and without learning bias. The model without learning biases, while not entirely a failure, was much less successful. Though it was generally able to fit the data, it predicted second waves much less frequently and made many more unrealistic predictions. Details are provided in the following subsections.

2.1. Model predictions

To get a sense of the model's average behavior, we plotted the median prediction of the 100 simulations for each region. An example for New York can be found in Section 2, Fig. 2(a-f) (analogs for other regions can be found in the Supplementary Appendix, figures 3-13). The model with learning biases has consistent behavior within the training region. The median prediction shows a small second wave and the interquartile range shows one of similar size to the first.

The model without learning biases fits the data comparably well, but has greatly reduced variability outside of the training region. Second wave predictions are smaller or non-existent, typically only suggested by the upper quartile rather than the median. Section 2.1 Table 1 gives a numerical summary of the biased model's second wave prediction performance across all regions (an analogous table for the unbiased model can be found in the Supplementary Appendix (Table 1). Section 2.1 Fig. 1(a-m) shows a graphical summary of the biased model's performance across all regions. The Supplementary Appendix (Figure 2) contains an analog for the unbiased model.

Table 1
Summary of true second wave and the biased model's second wave predictions.

Region	True peak time[1,2]	True size[1]	Prediction rate	Mean predicted time[1,3]	Mean predicted size
Austria	294	1.15×10^{-2}	0.86	222	5.88×10^{-2}
Belgium	273	1.67×10^{-2}	0.68	217	5.18×10^{-2}
Germany	315	4.90×10^{-3}	0.78	259	4.53×10^{-2}
Netherlands	273	8.50×10^{-3}	0.68	226	4.09×10^{-2}
Italy	287	9.18×10^{-3}	0.83	214	4.46×10^{-2}
United Kingdom	280	6.44×10^{-3}	0.63	266	2.95×10^{-2}
California, USA	336	1.85×10^{-2}	0.17	454	2.74×10^{-3}
New York, USA	350	1.26×10^{-2}	0.79	276	1.94×10^{-2}
Pennsylvania, USA	329	1.29×10^{-2}	0.80	212	2.05×10^{-2}
Texas, USA	217	3.27×10^{-3}	0.25	342	9.89×10^{-3}
British Columbia, Canada	308	2.56×10^{-3}	0.047	357	2.52×10^{-3}
Ontario, Canada	343	3.54×10^{-3}	0.098	158	4.19×10^{-3}
Quebec, Canada	336	4.84×10^{-3}	0.48	534	1.79×10^{-2}

[1] Calculated using the same mechanism as in section 4.3.
[2] Measured in days since February 18, 2020.
[3] Excluding simulations which do not predict second waves.

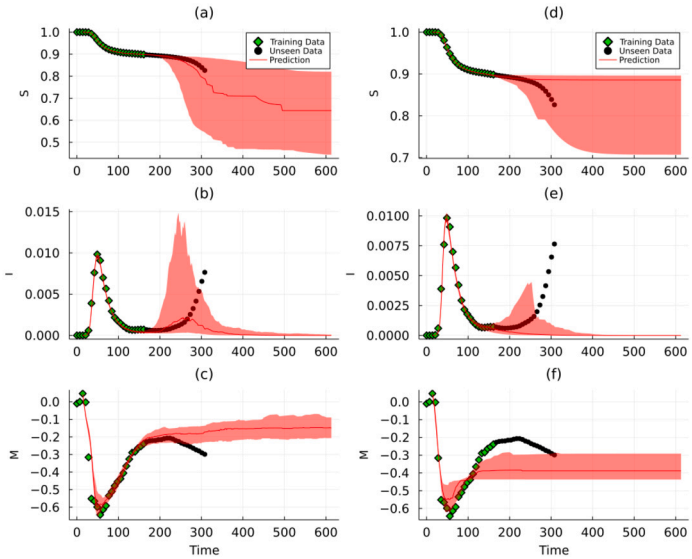


Fig. 2. Predicted time series of all model states for New York state with learning biases (a-c) and without (d-f). Panels (a) and (b) show susceptible fraction, (b) and (d) show infected, and (c) and (f) show mobility. Green dots represent training data, while black dots represent unseen data. All predictions are generated using the median (solid line) and interquartile range (ribbon) of 100 independently trained instances of the model.

2.2. Biological feasibility

Both models, with and without learning biases, tended to make feasible predictions, in the sense that all model states remained within their respective bounds. The biased model was stable 88% of the time, while the unbiased model was stable 85% of the time, across all regions.

However, when evaluating the learning bias loss functions on the trained models, it becomes clear that the model with learning biases is more reliable in this regard. The biased model achieves better losses across all loss objectives, including accuracy, compared to the unbiased model. Comparison of all loss functions can be found in the supplementary material.

The unbiased model does particularly poorly on the mobility upper and lower bounds (on the order of 10^4 times worse than biased), and the tendency for mobility to return to baseline in the absence of infection (roughly 10^3 times worse).

2.3. Second wave prediction

As a more robust metric for second wave prediction, we counted the number of local maxima exceeding at least 10^{-3} in the infected time series for each model simulation. The value 10^{-3} was tuned to exceed the size of any insignificant background fluctuations during the lulls between actual waves. With learning biases, the model predicted second waves regularly for most regions. (For example, it predicted second waves more than 63% of the time for all European regions. It performed worst on Ontario and British

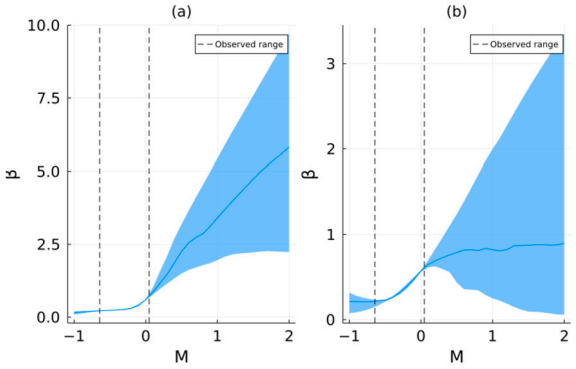


Fig. 3. Predicted force of infection based on mobility level for New York state with learning biases (a) and without (b). Dotted lines indicate values of mobility seen by the model during training. Solid line shows the median prediction of 100 model instances and ribbon shows interquartile range.

Columbia. This may be because the training data for these regions did not include the peak of the first wave, so the model predicts the first wave to increase further.

The unbiased model, meanwhile, rarely predicted second waves for any region (see Supplementary Appendix Table 1 for details). Its best performance was on Quebec, where it predicted second waves 51% of the time. This was also the only region in which it outperformed the biased model, which predicted second waves 48% for of the time. Otherwise, it predicted second waves less than 66% as often as the biased model, sometimes as little as 1.6% as often. It predicted no second waves at all for British Columbia.

Most of the time, both models predict the second wave too early (exceptions being Texas and Quebec), but the biased model's estimate is usually closer (only excepting Texas and Quebec). In terms of wave size, both models' median predictions undershoot the actual second wave size, but the biased model's upper quartile frequently exceeds it. The biased model's upper quartile only exceeds the true size for Germany and otherwise falls well short of it.

2.4. Transmissibility

One of the main uses of this model is that the trained neural network representing the force of infection can, once trained, be analyzed to examine the learned relationship between mobility and the transmission rate.

Section 2.4 Fig. 3(a,b) shows the distribution in the response of β to mobility predicted by the model with learning biases for New York. The models all converge on the same relationship within the training region and on low out-of-sample values, but they diverge for large ones. It is also worth noting that the prediction is, as expected, monotonically increasing. Once again, all regions demonstrate similar behavior (see Supplementary Appendix figures 14-24).

As with the time series predictions, the model without learning biases fits the data similarly well within the region on which it has been trained. However, outside that region, it extrapolates a flatter curve that is about equally likely to be higher or lower than the median.

For a quantitative sense of how β responds to mobility, we evaluated each trained network at the baseline value of mobility to determine the value of β, and hence R_0 ($= \beta/\gamma$), the basic reproduction number of the virus at the baseline value of mobility. We also applied Newton's method to the trained neural network to find the value of mobility (M_{crit}) at which R_0 drops below 1, the value below which the infection will die out [64]. Results for the biased model can be found in table Section 2.4 Table 2. The unbiased model results are negligibly different for the R_0. The results for M_{crit} are more variable. These unbiased model results can be found in the Supplementary Appendix Table 2.

The R_0 predictions, averaged over all simulations for a given region, range from 1.60 (British Columbia) to 2.60 (Germany). While estimates of R_0 for COVID-19 vary significantly between countries and times, this is in line with estimates of between 2.4 and 2.4 for the original COVID-19 strain [65–68]. It is also consistent with other models, which found Germany and the Netherlands to have higher values [69].

The model typically estimates that a 40-50% reduction in mobility is necessary to reduce R_0 below 1. This is consistently more extreme than other studies have found (20-40%) [70], but not entirely implausible considering the interquartile range. That said, we cannot interpret any result for Belgium, California, or the UK where the interquartile range exceeds physically realistic bounds.

3. Discussion

Our results show that socially and biologically informed machine learning models can perform qualitative prediction tasks. When supplied with learning biases, the model routinely predicted second pandemic waves similar to those that occurred in most populations during the COVID-19 pandemic. The model seldom produced implausible predictions for mobility, and where it did, this tended to result from a failure to converge during training.

The most significant result is that the biased model predicts a second wave in every region except the Canadian province of British Columbia. The biased model predicted the second wave peak more consistently and closer to the actual time than the unbiased model

Table 2
Predicted R_0 and required mobility reduction.

Region	Predicted R_0	M_{crit}	Convergence[1]
Austria	2.2 (±0.2)	−0.44 (±0.20)	0.96
Belgium	2.6 (±1.2)	−0.56 (±0.54)*	0.76
Germany	2.6 (±0.3)	−0.51 (±0.21)	0.98
Italy	2.1 (±0.5)	−0.49 (±0.19)	0.99
Netherlands	2.1 (±0.2)	−0.49 (±0.17)	0.99
United Kingdom	1.9 (±0.6)	−0.17 (±7.99)*	0.79
California, USA	1.6 (±0.9)	−0.44 (±3.96)*	0.6
New York, USA	2.4 (±0.2)	−0.52 (±0.19)	0.92
Pennsylvania, USA	2.3 (±0.7)	−0.56 (±0.14)	0.94
Texas, USA	1.7 (±0.8)	−0.57 (±0.52)	0.82
British Columbia	1.6 (±0.1)	−0.44 (±0.22)	0.99
Ontario	2.4 (±0.7)	−0.65 (±0.21)	0.80
Quebec	2.3 (±1.2)	−0.53 (±0.30)	0.92

Parentheses show 95% credible intervals.
[1] Not all neural networks showed a root. Interval calculated using only those that did.
* Interval exceeds physically realistic values.

without behavioral (mobility) feedback. The biased model also tended to predict a second wave that was much larger than the first wave, as occurred in most populations during the COVID-19 pandemic, although the predicted second wave was often larger in magnitude than what occurred in reality.

This ability to predict second waves is valuable from a public health perspective, for mitigation of population health impacts. Though our model does not explicitly include government policy, it can influence behavior, and knowing the likely trajectory of future cases under current policy can help decision-makers assess whether mandates should be tightened or loosened [43,41]. In practice, our model could be used to simulate possible outcomes by using the trained β network, but changing M to a time signal representing total lifting of restrictions, gradual reopening, or continuing heavy restriction. Such a model may need to account in some way for the costs of each policy.

Mixed machine learning models need not supplant traditional models entirely, but they can be a valuable auxiliary. As our model shows, they need not be overly complex or computationally expensive. They can interpret large amounts of data, generalize well to a variety of different regions, and given appropriate learning biases, can be relied upon to make feasible predictions.

Epidemic models are often under-determined by data [71]. UDEs allow a new approach to this problem. Since neural networks are universal approximators [72], they can represent the full range of possible functions that could fit the available data. By training multiple iterations of a UDE model and analyzing their trajectories, we can see a range of feasible outcomes for the system with just one single model. For example, two UDE predictions can fit the data and biological constraints equally well, yet one may predict a massive second wave, while the other predicts a rapid return to normalcy. A third may produce several smaller waves with corresponding mobility changes. That said, it is important to assign sufficient weight to the learning biases to avoid discouraging such a range of behaviors in favor of a single, overfitted solution.

The ability of UDEs to examine a range of data-fitting functions could be further enhanced with sparse regression methods [58,73,74]. By applying sparse regression to our trained β model's output, one could derive a multitude of symbolic equations that could be used to mathematically model the system.

The results also support our hypothesis that learning biases are effective at accelerating training and assuring socially and biologically plausible solutions while achieving superior training performance. While some attributes can be learned passably well by the unbiased model given sufficient training time, the biased model still achieves better losses on these attributes by at least two orders of magnitude. Good performance on training data should not be taken too seriously since it may be a sign of overtraining. However, this does not appear to be the case in our model. The vast majority of the average training loss comes from a few highly divergent solutions. The improved performance by the biased model indicates reduced proclivity for such solutions.

The fact that a monotonic β is learned comparably well by both models indicates that both of them are instrumentally useful for the model to learn when satisfying the loss objective. This gives a good sanity check that these features are present in the real system and assuming them in the model is reasonable. The upper and lower bounds on mobility, however, are not typically inferred by the model without explicit instruction. This is not unexpected since the observed data never nears the bounds. By fitting the data well, the model never needs to learn what happens at those bounds. However, including these boundaries as learning biases gives greater assurance that the model will not produce divergent or unstable solutions if, for example, it were used to predict what would happen in a scenario where those bounds were neared. Of course, it is preferable to ensure stability mathematically using structural biases, but this may not always be feasible. The other objective that the unbiased model tends not to learn is the tendency for M to return to the baseline value of 0. This may be because, as people become accustomed to life with the virus, the "baseline," i.e. the average societal preference in the absence of disease, shifts downward.

It is interesting that the learning biases help generate greater variability in the out-of-sample time series predictions. This is likely because, in the absence of any other objective, the model consistently converges to a single global optimum for data fitting. Since the model's extended prediction tends to remain within a small region of state space (M remaining negative, I relatively small), the greater potential variance is never realized. The model with learning biases, meanwhile, is relatively less concerned with fitting

the data and hence has more freedom to explore the parameter space. The fact that the constraint losses are evaluated according to randomly generated sample points also confers greater variability to the results of the biased model.

The biased model also has greater variability in the upper quartile of its β response but reduced variability in the lower quartile. This makes sense – the biased model has learned that for any M greater than those it has seen, the value of β must also be greater (and vice-versa for M less than what it has seen). The unbiased model, having no such information, cannot make an informed prediction, and so is equally likely to predict a continued increase or an unrealistic decrease.

These variability trade-offs favor the biased model. Greater variability in time series prediction is valuable because (assuming the predictions are biologically feasible) it shows a greater variety of possible outcomes and assigns a degree of confidence to those outcomes. The reduced variability in predicting the transmissibility is also desirable because it derives from a better understanding of the system.

Although our current model is retrospective and hence not useful as a predictive tool, it demonstrates potential for the future. Socioeconomic factors will continue to be complex, and regional and temporal variability will persist. If data-driven approaches can help overcome the challenges these factors present, we should facilitate their use by ensuring continued availability of high-quality data – both for endemic diseases and future pandemics. Universal differential equations specifically, when fine-tuned and supplied with appropriate learning biases, could be useful alongside traditional models to quickly gain perspective on the state of outbreaks across the world without having to develop specialized models for each region.

3.1. Limitations

We used a heavily simplified model of COVID-19. It is not intended to capture all details of the pandemic, nor is it meant to recommend specific health policies. We assumed the acquired immunity is permanent, which it may well not be [75]. We do not account for vaccines, which came into play around the end of 2020 [76]. Thus, the long-term predictions (i.e. beyond 300 days or so) should be taken only as evidence that the model does not produce wildly implausible behavior rather than a serious attempt to forecast cases too far in the future. The emergence of new variants, first reported in September 2020 [77] at the end of the second wave in many populations, means that predictions for the tail end of 2020 are beyond the model's intended scope. Similarly, spatial structure is important and can influence dynamics [78,24]. Even in the short term, the model is not intended to predict cases or to precisely estimate the virus's basic reproduction number. It is limited by our ability to consistently measure recovery rates and estimate under-reporting ratios, which almost certainly vary between regions and over time within regions. For simplicity, we also left out asymptomatic transmission, seasonal changes in infectiousness, age structure, and reinfection, all of which hamper the model's short-term predictive ability compared to more complex models [71,79].

In addition to epidemiological features, the UDE framework also comes with limitations. For one, each model instance provides a single prediction with no indication of confidence level. We have done our best to mitigate this by running many model instances, but the optimization process may still favor certain solutions over others. As such, the intervals we present should not be considered true prediction intervals of the sort a probabilistic model might provide, but rather a general measurement of the model's tendency. For another, while UDEs are a flexible framework, they are still deterministic differential equations. They do not account for intrinsic randomness in the system or the technically discrete nature of an epidemic. These limitations are mitigated by choosing regions with large, fairly concentrated populations and averaging data weekly. Further considerations on network structure are discussed in 4.1. None of these limitations changed our conclusions, since our goal was to show that UDEs and PIML can fit available data while making qualitatively correct our-of-sample predictions.

3.2. Future directions

Future work could improve our model by incorporating some of the aforementioned details of the pandemic. This could give insight into other behavior-disease interactions like vaccine usage [80] or allow an examination of how these dynamics changed over the course of the pandemic. In sections 4.1-4.2.1 we also provide some methodological changes that could further develop the UDE/PIML themes, particularly regarding how to use learning biases effectively.

Probably the biggest opportunity for future work is to apply this type of data-driven differential equation model to other systems. Other infectious diseases, particularly those for which vaccines are available, are also coupled behavior-disease systems [80,74] and so could be amenable to this type of model. Beyond epidemic modeling, climate systems are also known to have important behavioral components [81,82]. Ultimately, one of the greatest advantages of UDEs is that, as per their name, they can theoretically be applied to any dynamical system [58]. It is only a matter of testing them to see if they provide valuable insight.

4. Materials and methods

Our model is a universal delay differential equation (UDDE) based on the standard SIR (susceptible-infected-recovered) model:

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta(M(t-\tau_1))S(t)I(t)\\ \frac{dI(t)}{dt} &= \beta(M(t-\tau_1))S(t)I(t) - \gamma I(t)\\ \frac{dM(t)}{dt} &= e^{-\delta t} f(I(t-\tau_1), \Delta I(t-\tau_2), M(t), R(t))\end{aligned} \tag{1}$$

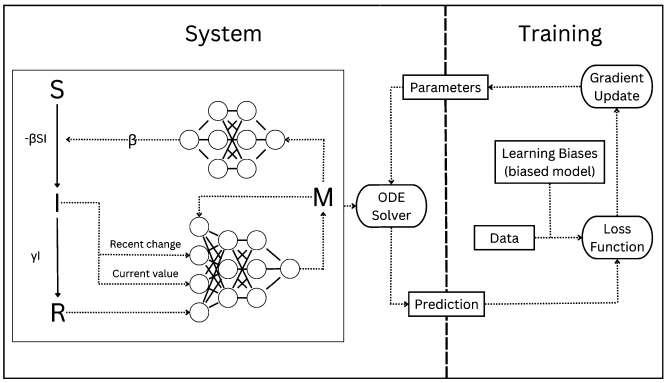


Fig. 4. Schematic of our model showing the relevant differential equations, neural networks, and training procedure. Neural networks are depicted with the actual topology used in the model. The learning biases are present only in the biased model. Otherwise, the biased and unbiased models have the same structure.

where S, I and R represent the susceptible, infected and recovered proportions of the population respectively (R can be recovered as $1 - S - I$), and M represents the relative difference in mobility compared to the baseline (i.e. 0 is the baseline, +1 is double the baseline, and -1 is complete reduction to no mobility. $\beta(M)$ represents the transmission rate as it depends on mobility, and $f(I, \Delta I, M, R)$ represents the dynamics governing social/behavioral (mobility) response to the infectious disease [42]. Both β and f were learned by the algorithm. $\Delta I(t)$ represents the change in I between the current time t and a previous time $t - \tau_2$. The $e^{-\delta t}$ factor accounts for several factors that reduce the population response to the virus over time, including pandemic fatigue, the development of medical interventions that make the infection less fatal (such as 'proning'), substituting less disruptive interventions (such as masking) for mobility reductions, and (for longer-term predictions than we study in this model) the evolution to milder virulence over time. δ is a trainable parameter. Section 4 Fig. 4 shows a schematic of the model.

The non-trainable model parameters are γ, the per-capita recovery rate, τ_1, the delay between a change in M and the corresponding change in prevalence, and τ_2, the reverse delay: the time between a change in prevalence and corresponding behavioral response [42,83]. The values we used are $\gamma = 0.25\text{day}^{-1}$ [84], $\tau_1 = 14$days, and $\tau_2 = 10$days [85,70]. This does assume these variables do not vary spatially or temporally (which they may not).

We chose the SIR model as a template for our model for two main reasons. First, relevant data in the form of case notifications suffice to reconstruct the values for all model states as shown in 4.3. Second, a simple model (like the SIR model) allows all more complex dynamics to be learned from the data by the neural network components. A model with an "exposed" category (SEIR) or even something more complex could function as well, but previous work has found that although COVID-19 may have a latency period, the SIR model performs just as well if not better than the SEIR model for estimating disease parameters [68].

Our model inherits several structural biases from the standard SIR model template. First, $S = 0$ and $I = 0$ are both invariant, preventing any infeasible negative values for these variables. Second, it retains the conservation relation $S + I + R = 1$. Thus, regardless of the functions fit by the neural network, $S(t)$ and $I(t)$ are guaranteed to be plausible. Of course, the model also inherits some biases and limitations from its SIR base. Namely, it assumes removal rate from death and recovery are constant and does not account directly for asymptomatic cases, latency period, differing severity, or potential for loss of immunity and subsequent reinfection.

4.1. Neural networks

The influence of mobility (i.e. contact rate) on the transmission rate is represented by neural networks that are used to represent $\beta(M)$ and $f(I, \Delta I, M, R)$. These networks each have linear output layers with one neuron and 2 hidden layers with 3 neurons per hidden layer and Gaussian Error Linear Unit (GELU) activation functions. This gives the β network 22 parameters and the f network 31. Accounting for the δ decay parameter, the model has 54 trainable parameters.

UDEs present an inherent tradeoff between parsimony and bias. Part of the appeal of UDEs is the ability to represent arbitrary functions – including discontinuous, non-differentiable ones – with neural networks. However, this universal approximation property relies on arbitrarily large networks [72]. However, smaller models train faster and can achieve a more favorable ratio of predictive power to number of parameters. We chose to lean more towards parsimony, so our model may struggle to learn highly discontinuous effects (such as the sudden implementation of country-wide lockdowns), being biased instead towards simpler continuous functions. The hyperparameter space was too large for us to optimize every aspect of the model, so different network parameters (size, activation, structure) may yield better results.

4.2. Training methodology

The baseline (unbiased) model, which received no social or biological feedback, was trained only to fit the data (details in section 4.3). The model's prediction is generated by solving the delay differential equation system to get its prediction for each state at each time step. We use the method of steps with the Rosenbrock23 differential equation solver to perform this process. The model's

Table 3
Biologically informed loss functions.

Biological assumption	Loss function
1. $\beta(M) \geq 0$ for all M	$\text{relu}(-\beta(M_i))$
2. Higher values of M correspond to higher values of $\beta(M)$	$\text{relu}(\beta(M_i)) - (\beta(M_i + \epsilon))$
3. M cannot go below a minimum value, i.e. a reduction of 100% from baseline	$\text{relu}(-f(M_{\min}, I_i, \Delta I_i, R))$
4. M cannot exceed a maximum value, set[1] at $M_{\max} = 2(M_{\text{baseline}} - M_{\min})$	$\text{relu}(f(M_{\max}, I_i, \Delta I_i, R))$
5. f is monotonically decreasing in I	$\text{relu}(f(M_i, I_i + \epsilon, \Delta I_i, R)) - f(M_i, I_i, \Delta I_i, R)))$
6. f is monotonically decreasing in ΔI	$\text{relu}(f(M_i, I_i, \Delta I_i + \epsilon, R)) - f(M_i, I_i, \Delta I_i, R)))$
7. In the absence of infection, M tends towards baseline	$\text{relu}[f(M, I, \Delta I, R)(M_i - M_{\text{baseline}})]$
8. M tends toward baseline more strongly when R is higher	$\text{relu}[\lvert f(M, I, \Delta I, R)\rvert - \lvert f(M, I, \Delta I, R + \epsilon)\rvert]$

[1] This upper bound is deliberately (if arbitrarily) set much higher than the maximum of the data to stop solutions diverging to infinity without excessively constraining the model.

predictions are recorded for each day. This prediction is then compared to the training data using a scaled mean-squared error loss function:

$$L(\Theta) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{y_{ij} - \bar{y}_{ij}}{m(\bar{y}_{i\max} - \bar{y}_{i\min})} \right)^2 \tag{2}$$

Here, n is the dimension of the system, m is the number of data points, y_{ij} is the true value of the ith variable's jth data point, and $\bar{y}_{ij}$ is the prediction for ith variable's jth data point. k is the size of the parameter vector Θ, and Θ_l is the lth entry in Θ. Scaling the loss function in this way helps ensure all variables are given equal importance despite having different ranges [86].

Both biased and unbiased models for all regions were trained on the first 160 days, giving $m = 22$ data points after weekly averaging (see 4.3). This time period fully encompasses the first wave for all populations studied, but does not include the beginning of the second wave.

4.2.1. Learning biases

The socially and biologically informed model was trained to minimize the same accuracy loss objective as well as 8 other objectives, each encoding a social or biological assumption. These biologically informed loss functions are deliberately constructed to give 0 loss to any functions that satisfy the relevant assumptions. This allows the model greater freedom to explore the range of biologically feasible functions.

To evaluate these additional loss functions, we generate 100 random points in the region $0 \leq I \leq 1, I - 1 \leq \Delta I \leq I, -100 \leq M \leq 100$ and evaluate each loss function at each point. The total loss at each iteration is then a weighted sum of these losses and the accuracy loss. We tried dynamically updating the weights for each loss function as in [31], but this did not significantly improve results. The biological assumptions and corresponding loss functions are displayed in Section 4.2.1 Table 3.

We only tested a few values for the learning bias weight. The optimal value for achieving tolerable performance on training data while assuring qualitatively realistic long-term predictions may be higher, lower, or vary between loss functions or throughout the training process.

Model parameters were randomly initialized. To save training time, parameter choices that gave initial errors of more than 10^4 were re-initialized. To optimize the parameters, we use the Zygote package [87] for reverse-mode automatic differentiation to obtain gradients of each loss function with respect to the model parameters (note that we only take the accuracy gradient for the unbiased model). We pass these gradient values to the Adam optimizer to update the model parameters. One training epoch thus consists of a one predict-differentiate-update cycle. The code used to implement this algorithm is available online (see section Data availability).

We found that training the model on the entire training set at once caused it to become stuck in a local optimum where I never increased. Thus, we trained the models in stages to achieve a better fit more quickly. The model trained on the first quarter of the data in the first stage (50,000 epochs at a learning rate of 0.001), the first half in the second (10,000 epochs at a learning rate of 0.001), and the entire training set in the third (20,000 epochs at a learning rate of 0.0005).

Repeating our model with more computing time and power could be informative. Although we were able to run the model with enough iterations to ensure all models converged to a good degree, some certainly converged better than others. The mobility data was a particular challenge, with fairly sharp downturns and upturns occasionally not always fully captured. This could be assisted by using collocation-based training to speed up the process [88].

4.3. Data

4.3.1. Case data

Daily case notification data was taken from the Johns Hopkins CSSE dataset [89]. We derived daily susceptible and infected proportions using the following system:

$$\begin{aligned} I_T(t_n) &= I_T(t_{n-1}) - \gamma I_T(t_{n-1}) + 5C(t_n) \\ S_T(t_n) &= S_T(t_{n-1}) - 5C(t_n) \end{aligned} \tag{3}$$

where $S_T(t_n)$ and $I_T(t_n)$ are the total number of susceptible and infected individuals (not proportions) on day t_n and $C(t_n)$ are the number of new cases on day t_n. The parameter γ is the same used in the model as discussed above. In order to account for under-reporting of cases, we multiply $C(t_n)$ by 5 (corresponding to a ratio of 4 unreported cases per reported case). This ratio is in accordance with prior estimates [90]

Finally, we divided I and S by the total population to get a proportion at each time step to ensure different regions are comparable.

4.3.2. Mobility data

Daily mobility data (M) comes from the Google Community Mobility Report [91]. We mean-normalized this data to give it a comparable range to S and I. We chose the Retail and Recreation category of mobility data as it corresponded most to what we were trying to measure: voluntary activities in indoor settings that place people at risk of becoming infected. It is also strongly correlated with infection, so it is reasonable to expect it can be a good predictor of cases. Repeating the simulations with workplace mobility would be a good test of the model's validity. Other mobility measures, such as parks, would be difficult to use due to their weaker correlation with infections [23].

Once we had data for each day, we sub-sampled it by taking a weekly moving average to reduce irregularities from weekends, holidays, and days where data was not available. We set the initial condition as the first data point for which $I \neq 0$.

We chose relatively populous regions across Western Europe, the US, and Canada for which case and mobility data were available. All these regions experienced second waves to some degree, with varying intensities and timings. Our list of regions is non-exhaustive so future work could study other regions

CRediT authorship contribution statement

Bruce Kuwahara: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Chris T. Bauch:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors of "Predicting COVID-19 pandemic waves with biologically and behaviorally informed universal differential equations" in submission to EPIDEMICS have no competing interests to declare.

Data availability

The code used to run and analyze the model is available at https://github.com/bkuwahara/predicting-2ndwaves-udes. All empirical data are publicly available (see References).

Acknowledgements

The authors thank the anonymous reviewers for their valuable suggestions.

This work was supported by an NSERC Discovery Grant to C.T.B., and an NSERC Undergraduate Student Research Assistant Award to B.K.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e25363.

References

[1] Sina Abbasi, Babek Erdebilli, Green closed-loop supply chain networks&rsquo, response to various carbon policies during covid-19, Sustainability 15 (4) (2023).
[2] Sina Abbasi, Environmental impact assessment with rapid impact assessment matrix method during the Covid-19 pandemic: a case study in Tehran, Res. Sq. (2023).
[3] Ernesto DR. Santibanez Gonzalez, Sina Abbasi, Mahsa Azhdarifard, Designing a reliable aggregate production planning problem during the disaster period, Sustain. Oper. Comput. (2023).
[4] Sina Abbasi, Maryam Daneshmand-Mehr, Armin Ghane Kanafi, The sustainable supply chain of co2 emissions during the coronavirus disease (Covid-19) pandemic, J. Ind. Eng. Int. 17 (4) (2021) 83–108.
[5] Sina Abbasi, Maryam Daneshmand-Mehr, Armin Ghane, Kanafi Green, Closed-loop supply chain network design during the coronavirus (Covid-19) pandemic: a case study in the Iranian automotive industry, Environ. Model. Assess. 28 (1) (February 2023) 69–103.
[6] Alessandro Vespignani, Huaiyu Tian, Christopher Dye, James O. Lloyd-Smith, Rosalind M. Eggo, Munik Shrestha, Samuel V. Scarpino, Bernardo Gutierrez, Moritz UG Kraemer, Joseph Wu, et al., Modelling Covid-19, Nat. Rev. Phys. 2 (6) (2020) 279–281.
[7] Julia R. Gog, How you can help with Covid-19 modelling, Nat. Rev. Phys. 2 (6) (2020) 274–275.
[8] C. Jessica E. Metcalf, Dylan H. Morris, Sang Woo Park, Mathematical models to guide pandemic response, Science 369 (6502) (2020) 368–369.
[9] Meagan C. Fitzpatrick, Chris T. Bauch, Jeffrey P. Townsend, Alison P. Galvani, Modelling microbial infection to address global health challenges, Nat. Microbiol. 4 (10) (2019) 1612–1619.
[10] Herbert W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (4) (2000) 599–653.

[11] Karen KL Hwang, Christina J. Edholm, Omar Saucedo, Linda JS Allen, Nika Shakiba, A hybrid epidemic model to explore stochasticity in Covid-19 dynamics, Bull. Math. Biol. 84 (9) (2022) 91.
[12] Estee Y. Cramer, Evan L. Ray, Velma K. Lopez, Johannes Bracher, Andrea Brennen, Alvaro J. Castro Rivadeneira, Aaron Gerding, Tilmann Gneiting, Katie H. House, Yuxin Huang, et al., Evaluation of individual and ensemble probabilistic forecasts of Covid-19 mortality in the United States, Proc. Natl. Acad. Sci. 119 (15) (2022) e2113561119.
[13] Kathyrn R. Fair, Vadim A. Karatayev, Madhur Anand, Chris T. Bauch, Estimating Covid-19 cases and deaths prevented by non-pharmaceutical interventions, and the impact of individual actions: a retrospective model-based analysis, Epidemics 39 (2022) 100557.
[14] Chad R. Wells, Jeffrey P. Townsend, Abhishek Pandey, Seyed M. Moghadas, Gary Krieger, Burton Singer, Robert H. McDonald, Meagan C. Fitzpatrick, Alison P. Galvani, Optimal Covid-19 quarantine and testing strategies, Nat. Commun. 12 (1) (2021) 356.
[15] Ashleigh R. Tuite, David N. Fisman, Amy L. Greer, Mathematical modelling of Covid-19 transmission and mitigation strategies in the population of Ontario, Canada, Can. Med. Assoc. J. 192 (19) (2020) E497–E505.
[16] Jude Bayham, Eli P. Fenichel, Impact of school closures for Covid-19 on the us health-care workforce and net mortality: a modelling study, Lancet Public Health 5 (5) (2020) e271–e278.
[17] Peter C. Jentsch, Madhur Anand, Chris T. Bauch, Prioritising Covid-19 vaccination in changing social and epidemiological landscapes: a mathematical modelling study, Lancet Infect. Dis. 21 (8) (2021) 1097–1106.
[18] Seyed M. Moghadas, Thomas N. Vilches, Kevin Zhang, Chad R. Wells, Affan Shoukat, Burton H. Singer, Lauren Ancel Meyers, Kathleen M. Neuzil, Joanne M. Langley, Meagan C. Fitzpatrick, et al., The impact of vaccination on coronavirus disease 2019 (Covid-19) outbreaks in the United States, Clin. Infect. Dis. 73 (12) (2021) 2257–2264.
[19] Caroline E. Wagner, Chadi M. Saad-Roy, Bryan T. Grenfell, Modelling vaccination strategies for Covid-19, Nat. Rev. Immunol. 22 (3) (2022) 139–141.
[20] Sina Abbasi, Sasan Zahmatkesh, Awais Bokhari, Mostafa Hajiaghaei-Keshteli, Designing a vaccine supply chain network considering environmental aspects, J. Clean. Prod. 417 (2023) 137935.
[21] Dongshan Zhu, Xikun Han, Karla Santo, Social distancing in Latin America during the Covid-19 pandemic: an analysis using the stringency index and Google community mobility reports, J. Travel Med. 27 (8) (2020).
[22] Antonio Paez, Using Google community mobility reports to investigate the incidence of Covid-19 in the United States, Find. 5 (2020).
[23] M. Sulyok, M. Walker, Community movement and Covid-19: a global study using Google's community mobility reports, Epidemiol. Infect. 148 (2020) e284.
[24] Vadim A. Karatayev, Madhur Anand, Chris T. Bauch, Local lockdowns outperform global lockdown on the far side of the Covid-19 epidemic curve, Proc. Natl. Acad. Sci. 117 (39) (2020) 24575–24580.
[25] Yehuda Pollak, Haym Dayan, Rachel Shoham, Itai Berger, Predictors of adherence to public health instructions during the Covid-19 pandemic, 2020, medRxiv.
[26] Abel Brodeur, Andrew E. Clark, Sarah Fleche, Nattavudh Powdthavee, Covid-19, lockdowns and well-being: evidence from Google trends, J. Public Econ. 193 (2021) 104346.
[27] Isaac Yen-Hao, Prima Alam, Heidi.J. Larson, Leesa Lin, Social consequences of mass quarantine during epidemics: a systematic review with implications for the Covid-19 response, J. Travel Med. 27 (7) (2020).
[28] Zhen Wang, Michael A. Andrews, Zhi-Xi Wu, Lin Wang, Chris T. Bauch, Coupled disease–behavior dynamics on complex networks: a review, Phys. Life Rev. 15 (2015) 1–29.
[29] Chris Bauch, Alberto d'Onofrio, Piero Manfredi, Behavioral Epidemiology of Infectious Diseases: an Overview, Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases, 2013, pp. 1–19.
[30] Sebastian Funk, Marcel Salathé, Vincent AA Jansen, Modelling the influence of human behaviour on the spread of infectious diseases: a review, J. R. Soc. Interface 7 (50) (2010) 1247–1256.
[31] Sifan Wang, Xinling Yu, Paris Perdikaris, When and why pinns fail to train: a neural tangent kernel perspective, J. Comput. Phys. 449 (2022) 110768.
[32] Timothy C. Reluga, Game theory of social distancing in response to an epidemic, PLoS Comput. Biol. 6 (5) (05 2010) 1–9.
[33] Alina Glaubitz, Feng Fu, Oscillatory dynamics in the dilemma of social distancing, Proc. R. Soc. A 476 (2243) (2020) 20200686.
[34] Chris T. Bauch, David JD Earn, Vaccination and the theory of games, Proc. Natl. Acad. Sci. 101 (36) (2004) 13391–13394.
[35] Alberto d'Onofrio, Piero Manfredi, Ernesto Salinelli, Vaccinating behaviour, information, and the dynamics of sir vaccine preventable diseases, Theor. Popul. Biol. 71 (3) (2007) 301–317.
[36] Iain R. Moyles, Jane M. Heffernan, Jude D. Kong, Cost and social distancing dynamics in a mathematical model of Covid-19 with application to Ontario, Canada, R. Soc. Open Sci. 8 (2) (2021) 201770.
[37] KM Ariful Kabir, Jun Tanimoto, Evolutionary game theory modelling to represent the behavioural dynamics of economic shutdowns and shield immunity in the Covid-19 pandemic, R. Soc. Open Sci. 7 (9) (2020) 201095.
[38] Joshua M. Epstein, Erez Hatna, Jennifer Crodelle, Triple contagion: a two-fears epidemic model, J. R. Soc. Interface 18 (181) (2021) 20210186.
[39] Bruno Buonomo, Rossella Della Marca, Alberto d'Onofrio, Maria Groppi, A behavioural modelling approach to assess the impact of Covid-19 vaccine hesitancy, J. Theor. Biol. 534 (2022) 110973.
[40] Nauman Ahmed, Amr Elsonbaty, Ali Raza, Muhammad Rafiq, Waleed Adel, Numerical simulation and stability analysis of a novel reaction–diffusion Covid-19 model, Nonlinear Dyn. 106 (2) (October 2021) 1293–1310.
[41] Sansao A. Pedro, Frank T. Ndjomatchoua, Peter Jentsch, Jean M. Tchuenche, Madhur Anand, Chris T. Bauch, Conditions for a second wave of Covid-19 due to interactions between disease dynamics and social processes, Front. Phys. 8 (2020).
[42] Matthew D. Johnston, Bruce Pell, A dynamical framework for modeling fear of infection and frustration with social distancing in Covid-19 spread, 2020.
[43] Ashleigh R. Tuite, Amy L. Greer, Risk for Covid-19 resurgence related to duration and effectiveness of physical distancing in Ontario, Canada, Ann. Intern. Med. 173 (8) (2020) 675–678, PMID: 32459528.
[44] Iman Rahimi, Fang Chen, Amir H. Gandomi, A review on Covid-19 forecasting models, Neural Comput. Appl. (2021).
[45] Maximilian Nickel, Levent Sagun, Mark Ibrahim, Matt Le, Timothee Lacroix, Neural relational autoregression for high-resolution Covid-19 forecasting, in: Meta AI ML Applications, 2020.
[46] Raj Dandekar, Chris Rackauckas, George Barastathis, A machine learning-aided global diagnostic and comparative tool to assess effect of quarantine control in Covid-19 spread, Patterns 1 (9) (2020).
[47] Gergo Pinter, Imre Felde, Amir Mosavi, Pedram Ghamisi, Richard Gloaguen, Covid-19 pandemic prediction for Hungary, a hybrid machine learning approach, Mathematics 8 (6) (2020).
[48] Sina F. Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R. Varkonyi-Koczy, Uwe Reuter, Timon Rabczuk, Peter M. Atkinson, Covid-19 outbreak prediction with machine learning, Algorithms 13 (10) (2020).
[49] George E. Karniadakis, Ioannidis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, Liu Yang, Physics-informed machine learning, Nat. Rev. Phys. 3 (6) (2021) 422–440.
[50] Maziar Raissi, Paris Perdikaris, George Em Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2018) 11.
[51] Sharmila Karumuri, Rohit Tripathy, Ilias Bilionis, Jitesh Panchal, Simulator-free solution of high-dimensional stochastic elliptic partial differential equations using deep neural networks, J. Comput. Phys. 404 (2020) 109120.
[52] Dongkun Zhang, Ling Guo, George Em Karniadakis, Learning in modal space: solving time-dependent stochastic pdes using physics-informed neural networks, SIAM J. Sci. Comput. 42 (2) (2020) A639–A665.

[53] Georgios Kissas, Yibo Yang, Eileen Hwuang, Walter R. Witschey, John A. Detre, Paris Perdikaris, Machine learning in cardiovascular flows modeling: predicting arterial blood pressure from non-invasive 4d flow mri data using physics-informed neural networks, Comput. Methods Appl. Mech. Eng. 358 (2020) 112623.
[54] Casey S. Greene, James C. Costello, Biologically informed neural networks predict drug responses, Cancer Cell 38 (5) (2020) 613–615.
[55] Haitham A. Elmarakeby, Justin Hwang, Rand Arafeh, Jett Crowdis, Sydney Gang, David Liu, Saud H. Al Dubayan, Keyan Salari, Steven Kregel Camden Richter, Taylor E. Arnoff, Jihya Park, William C. Hahn, Eliezer M. Van Allen, Biologically informed deep neural network for prostate cancer discovery, Nature 598 (2021) 348–352.
[56] Andrew Beers, Ken Chang, James M. Brown, Emmett Sartor, C.P. Mammen, Elizabeth R. Gerstner, Bruce R. Rosen, Jayashree Kalpathy-Cramer, Sequential 3d u-nets for biologically-informed brain tumor segmentation, CoRR, arXiv:1709.02967 [abs], 2017.
[57] John H. Lagergren, John T. Nardini, Ruth E. Baker, Matthew J. Simpson, Kevin B. Flores, Biologically-informed neural networks guide mechanistic modeling from sparse experimental data, PLoS Comput. Biol. 16 (12) (12 2020) 1–29.
[58] Christopher Rackauckas, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Supekar, Dominic Skinner, Ali Jasim Ramadhan, Universal differential equations for scientific machine learning, CoRR, arXiv:2001.04385 [abs], 2020.
[59] Momchil Minkov, Ian A.D. Williamson, Lucio C. Andreani, Dario Gerace, Beicheng Lou, Alex Y. Song, Tyler W. Hughes, Shanhui Fan, Inverse design of photonic crystals through automatic differentiation, ACS Photonics 7 (7) (2020) 1729–1741.
[60] Raj Dandekar, George Barbastathis, Neural network aided quarantine control model estimation of COVID spread in Wuhan, China, 2020.
[61] Marissa Renardy, Marisa Eisenberg, Denise Kirschner, Predicting the second wave of Covid-19 in washtenaw county, mi, J. Theor. Biol. 507 (2020) 110461.
[62] Junling Ma, Estimating epidemic exponential growth rate and basic reproduction number, Infect. Dis. Model. 5 (2020) 129–141.
[63] Robert M. May, Uses and abuses of mathematics in biology, Science 303 (5659) (2004) 790–793.
[64] Chris T. Bauch, Estimating the Covid-19 R number: a bargain with the devil? Lancet Infect. Dis. 21 (2) (February 2021) 151–153, Publisher: Elsevier.
[65] Ying Liu, Albert A. Gayle, Annelies Wilder-Smith, Joacim Rocklöv, The reproductive number of Covid-19 is higher compared to Sars coronavirus, J. Travel Med. 27 (2) (2020).
[66] Nithya C. Achaiah, Sindhu B. Subbarajasetty, Rajesh M. Shetty, R0 and re of Covid-19: can we predict when the pandemic outbreak will be contained? Indian J. Crit. Care Med. (2020).
[67] Vishal Srivastava, Smriti Srivastava, Gopal Chaudhary, Fadi Al-Turjman, A systematic approach for Covid-19 predictions and parameter estimation, Pers. Ubiquitous Comput. (2020).
[68] Weston C. Roda, Marie B. Varughese, Donglin Han, Michael Y. Li, Why is it difficult to accurately predict the Covid-19 epidemic?, Infect. Dis. Model. 5 (2020) 271–281.
[69] Kevin Linka, Mathias Peirlinck, Ellen Kuhl, The reproduction number of Covid-19 and its correlation with public health interventions, 2020.
[70] Robert B. Noland, Mobility and the effective reproduction rate of Covid-19, J. Transp. Health 20 (2021) 101016.
[71] Hyun Mo Yang, Luis Pedro Lombardi Junior, Ariana Campos Yang, Are the sir and seir models suitable to estimate the basic reproduction number for the Covid-19 epidemic? 2020, medRxiv.
[72] Sho Sonoda, Noboru Murata, Neural network with unbounded activation functions is universal approximator, Appl. Comput. Harmon. Anal. 43 (2) (2017) 233–268.
[73] Steven L. Brunton, Joshua L. Proctor, J. Nathan Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. 113 (15) (2016) 3932–3937.
[74] Jonathan Horrocks, Chris T. Bauch, Algorithmic discovery of dynamic models from infectious disease data, Sci. Rep. 10 (1) (2020) 7061.
[75] Chris Baraniuk, How long does Covid-19 immunity last? BMJ 373 (2021).
[76] AJMC Staff, A Timeline of Covid-19 Vaccine Developments in 2021, Online, June 2021.
[77] World Health Organization, Tracking Sars-Cov-2 Variants, Online, November 2021.
[78] Xingru Chen, Feng Fu, Imperfect vaccine and hysteresis, Proc. R. Soc. B 286 (1894) (2019) 20182406.
[79] Amr Elsonbaty, Zulqurnain Sabir, Rajagopalan Ramaswamy, Waleed Adel, Dynamical analysis of a novel discrete fractional sitrs model for Covid-19, Fractals 29 (08) (2021) 2140035.
[80] Chris T. Bauch, Imitation dynamics predict vaccinating behaviour, Proc. R. Soc. Lond. B 272 (1573) (2005) 1669–1675.
[81] Alison P. Galvani, Chris T. Bauch, Madhur Anand, Burton H. Singer, Simon A. Levin, Human–environment interactions in population and ecosystem health, Proc. Natl. Acad. Sci. 113 (51) (2016) 14502–14506.
[82] Thomas M. Bury, Chris T. Bauch, Madhur Anand, Charting pathways to climate change mitigation in a coupled socio-climate model, PLoS Comput. Biol. 15 (6) (2019) e1007000.
[83] Xingru Chen, Feng Fu, Highly coordinated nationwide massive travel restrictions are central to effective mitigation and control of Covid-19 outbreaks in China, Proc. R. Soc. A 478 (2260) (2022) 20220040.
[84] Hiroshi Nishiura, Natalie M. Limton, Andrei R. Akhmetzhanov, Serial interval of novel coronavirus (Covid-19) infections, Int. J. Infect. Dis. 93 (2020) 284–286.
[85] Hamada S. Badr , Hongru Du, Maximilian Marshall, Ensheng Dong, Marietta M. Squire, Lauren M. Gardner, Association between mobility patterns and Covid-19 transmission in the usa: a mathematical modelling study, Lancet Infect. Dis. 20 (11) (2020) 1247–1254.
[86] Suyong Kim, Weiqi Ji, Sili Deng, Yingbo Ma, Christopher Rackauckas, Stiff neural ordinary differential equations, Chaos, Interdiscip. J. Nonlinear Sci. 31 (9) (2021) 093122.
[87] Michael Innes, Don't unroll adjoint: differentiating ssa-form programs, CoRR, arXiv:1810.07951 [abs], 2018.
[88] Elisabeth Roesch, Christopher Rackauckas, Michael P.H. Stumpf, Collocation based training of neural ordinary differential equations, Stat. Appl. Genet. Mol. Biol. 20 (2) (2021) 37–49.
[89] Dataset* Johns Hopkins University. Covid-19 data repository by the center for systems science and engineering, (csse) at johns Hopkins university.
[90] Shelly Bolotin, Vanessa Tran, Shelley L. Deeks, Adriana Peci, Kevin A. Brown, Sarah A. Buchan, Katherene Ogbulafor, Tubani Ramoutar, Michelle Nguyen, Rakesh Thakkar, Reynato DelaCruz, Reem Mustfa, Jocelyn Maregmen, Orville Woods, Ted Krasna, Kirby Cronin, Selma Osman, Eugene Joh, Vanessa G. Allen, Assessment of Population Infection with SARS-CoV-2 in Ontario, Canada, March to June 2020, Euro Surveill. 26 (50) (December 2021), Place: Sweden.
[91] Dataset* Google. Google community mobility report.