

Brief Communication

PAN-cODE: COVID-19 forecasting using conditional latent ODEs

Ruian Shi^{1,2}, Haoran Zhang^{1,2}, and Quaid Morris^{1,2,3}

¹Department of Computer Science, University of Toronto, Toronto, Canada, ²Vector Institute for Artificial Intelligence, Toronto, Ontario, Canada and ³Computational and Systems Biology, Memorial Sloan Kettering Cancer Center, New York City, New York, USA

Corresponding Author: Quaid Morris, PhD, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, 417 East 68th Street, New York City, NY 10021, USA; morrisq@mskcc.org

Received 3 December 2021; Revised 22 July 2022; Editorial Decision 22 August 2022; Accepted 31 August 2022

ABSTRACT

The coronavirus disease 2019 (COVID-19) pandemic has caused millions of deaths around the world and revealed the need for data-driven models of pandemic spread. Accurate pandemic caseload forecasting allows informed policy decisions on the adoption of non-pharmaceutical interventions (NPIs) to reduce disease transmission. Using COVID-19 as an example, we present Pandemic conditional Ordinary Differential Equation (PAN-cODE), a deep learning method to forecast daily increases in pandemic infections and deaths. By using a deep conditional latent variable model, PAN-cODE can generate alternative caseload trajectories based on alternate adoptions of NPIs, allowing stakeholders to make policy decisions in an informed manner. PAN-cODE also allows caseload estimation for regions that are unseen during model training. We demonstrate that, despite using less detailed data and having fully automated training, PAN-cODE's performance is comparable to state-of-the-art methods on 4-week-ahead and 6-week-ahead forecasting. Finally, we highlight the ability of PAN-cODE to generate realistic alternative outcome trajectories on select US regions.

Key words: pandemic prediction, deep learning, time series forecasting, latent variable models

INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic remains a severe threat to public health, affecting 235 million individuals, and causing 4.8 million deaths as of October 2021.¹ Forecasting future disease caseload is critical for pandemic response² and enables informed policy decisions on medical resource allocation and policies to reduce disease transmission. For COVID-19, non-pharmaceutical interventions (NPIs) such as social distancing and mask mandates are critical in slowing transmission³ and remain important even as vaccines are deployed.⁴ In [Figure 1](#), we visualize the negatively correlated relationship between NPI adoption and COVID-19 transmission in the state of California. While NPIs drastically reduce COVID-19 transmission, they also incur high socio-economic costs,^{5,6} motivating the ability to model future COVID-19 caseload

as a function of NPI adoption stringency. Current approaches to caseload forecasting include statistical methods,^{7,8} compartmental models,^{9–14} deep neural networks,^{15–20} or ensembles such as the COVID-19 Forecast Hub.²¹ These methods provide reliable short-term forecasting, but as we further outline in [Supplementary Appendix A](#), many do not offer predictions past 4 weeks, cannot explicitly model the relationship between NPI adoption and future caseload, and require expert-intervention to be fit to data.

Here, we introduce the Pandemic conditional Ordinary Differential Equation (PAN-cODE), a fully-automated pandemic caseload forecasting method capable of conditioning forecasts on the stringency of NPI adoption. PAN-cODE uses deep neural networks in a Latent Ordinary Differential Equation (ODE)²² architecture to learn a conditional generative model of caseload dynamics only using past

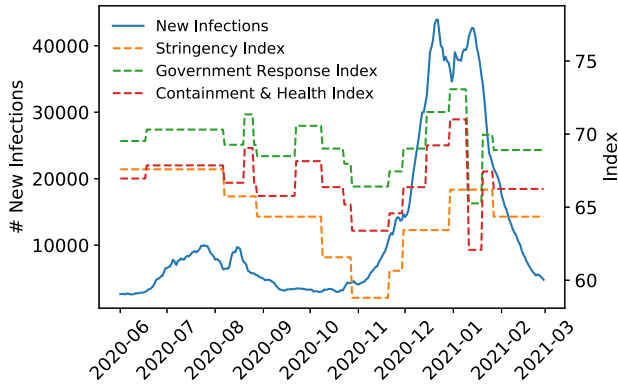


Figure 1. The number of daily new infections over time in the state of California plotted alongside indices representing the level of NPIs adopted. These indices are provided by the Oxford COVID-19 Government Response Tracker³⁴ and summarize the level of NPIs adopted into a single metric scored out of 100. NPIs have a time-delayed, negatively correlated effect on COVID-19 caseload, seen most clearly from 2020–2010 to 2021–2001.

daily infection/death counts and NPI policy. When trained on US state and county level COVID-19 caseload data,²³ PAN-cODE offers 4-week forecasts with accuracy comparable to state-of-the-art methods and beats all COVID-19 Forecast Hub methods at 6-week-ahead forecasting while using minimal data features. We show that PAN-cODE can generalize learned dynamics to accurately predict outcomes in unseen regions. Finally, we demonstrate PAN-cODE's ability to generate realistic outcomes based on alternative NPI adoptions for selected US states.

BACKGROUND

Deep epidemiological forecasting

Epidemiological forecasting is traditionally performed using compartmental SIR models^{24–26} or time series methods.^{7,8} By leveraging the non-linear representational power of deep neural networks, deep learning approaches for epidemiological forecasting aim to better capture the complex relationship between historical caseloads, high-dimensional covariate time series, and future caseload trajectories. Current deep epidemiological forecasting methods use neural networks to estimate compartmental model coefficients from input features,^{9,15} integrate spatio-temporal graph neural networks to model how geographic proximity affects viral transmission,^{20,27–30} or apply deep time series methods such as LSTMs^{14,18,19} for forecasting.

Neural and latent ODEs

PAN-cODE learns COVID-19 caseload dynamics from historical data using the Neural ODE.³¹ Where $x_{0:N}$ denotes input data and $t_{0:N}$ represents their time of observation, the Neural ODE represents time series as the solution to an ODE such that:

$$x_{0:N} = \text{ODESolve}(f_{\theta}, x_0, t_{0:N})$$

where f_{θ} represents $\frac{dx}{dt} = f(x, t, \theta)$ and is parameterized by a neural network. Alternatively, f_{θ} could be parameterized by a known ODE, such as the SIR equations, where unknown parameters are then optimized by the Neural ODE through backpropagation.

The Neural ODE can be arranged into a variational auto-encoder architecture known as the Latent ODE²² to increase its representational ability. The Latent ODE first encodes multidimensional input features using the GRU-ODE: a Gated Recurrent Unit

(GRU)³² that evolves hidden states between observations using a Neural ODE. The GRU-ODE encoder outputs the parameters for a factorized Gaussian variational distribution over the initial state of a latent trajectory. From this variational distribution, we sample a latent initial state denoted z_0 , and solve a separate Neural ODE parameterized by f_{ψ} from z_0 to obtain a latent trajectory. Finally, this latent trajectory is decoded with neural network f_{ϕ} into the output trajectory. The Latent ODE is represented by the following set of equations:

$$\begin{aligned} \mu_{z_0}, \sigma_{z_0}^2 &= \text{GRUODE}_{f_{\theta}}(x_{0:N}, t_{0:N}) \\ z_0 &\sim q(z_0 | x_{0:N}) = \mathcal{N}(\mu_{z_0}, \sigma_{z_0}^2) \\ z_{0:N} &= \text{ODESolve}(f_{\psi}, z_0, t_{0:N}) \\ x_i &\sim \mathcal{N}(f_{\phi}(z_i), \sigma^2) \quad \text{for } i = 1, \dots, N \end{aligned}$$

where σ^2 is a fixed variance term. The Latent ODE is trained by maximizing the evidence lower bound (ELBO), defined as:

$$\mathbb{E}_{z_0 \sim q(z_0 | x_{0:N})}[\log p(x_{0:N})] - \text{KL}[q(z_0 | x_{0:N}) \| p(z_0)]$$

METHOD

PAN-cODE introduces novel modifications to the Latent ODE architecture to allow generation of alternative forecasts conditioned on the stringency of NPI adoption. We visualize the PAN-cODE architecture in Figure 2. PAN-cODE first learns a representation of current caseload dynamics using observed data in the Encoding Region, which includes input data up to the forecasting date, denoted f_c . The objective is to output the caseload in the Prediction Region, which begins on the forecasting date and extends to the prediction date, denoted pd . We denote the encoding region daily infection and death count as $i_{0:f_c}$ and $d_{0:f_c}$, respectively. Other covariates, denoted $f_{0:f_c}^{1:D}$, are also included as input for the GRU-ODE encoder. We assume regularly spaced time intervals between observations and arbitrarily set $t_{0:f_c}$. However, in contrast to most other time-series prediction methods, irregularly sampled time series can be easily handled by setting $t_{0:f_c}$ to the actual times of observation. The GRU-ODE encoder outputs the parameters $(\mu_{z_0}, \sigma_{z_0}^2)$ for the variational distribution over the initial state of a latent trajectory.

PAN-cODE introduces the conditional Latent ODE, which extends work on deep conditional generative models³³ to the Latent ODE. The conditional Latent ODE allows the latent initial state to be conditioned on injected data features. For PAN-cODE, we condition on 4 metrics representing the stringency of NPI adoption at forecasting time, denoted $I_{f_c} \in \mathbb{R}^4$. These metrics are described in the experimental section. The conditional Latent ODE is implemented by concatenating the vector of conditional features to the sampled latent initial state. Letting $X = [i_{0:f_c}, d_{0:f_c}, f_{0:f_c}^{1:D}]$, the augmented latent initial state \tilde{z}_0 is obtained as:

$$\begin{aligned} \mu_{z_0}, \sigma_{z_0}^2 &= \text{GRUODE}_{f_{\theta}}(X, t_{0:f_c}) \\ z_0 &\sim \mathcal{N}(\mu_{z_0}, \sigma_{z_0}^2) \\ \tilde{z}_0 &= [z_0, I_{f_c}] \end{aligned}$$

Consequently, the ELBO objective becomes:

$$\mathbb{E}[\log p(x_{0:N} | I_{f_c})] - \text{KL}[q(z_0 | x_{0:N}, I_{f_c}) \| p(z_0 | I_{f_c})]$$

During training, the model learns a correlational relationship between the conditional variables and caseload forecasts. Afterwards,

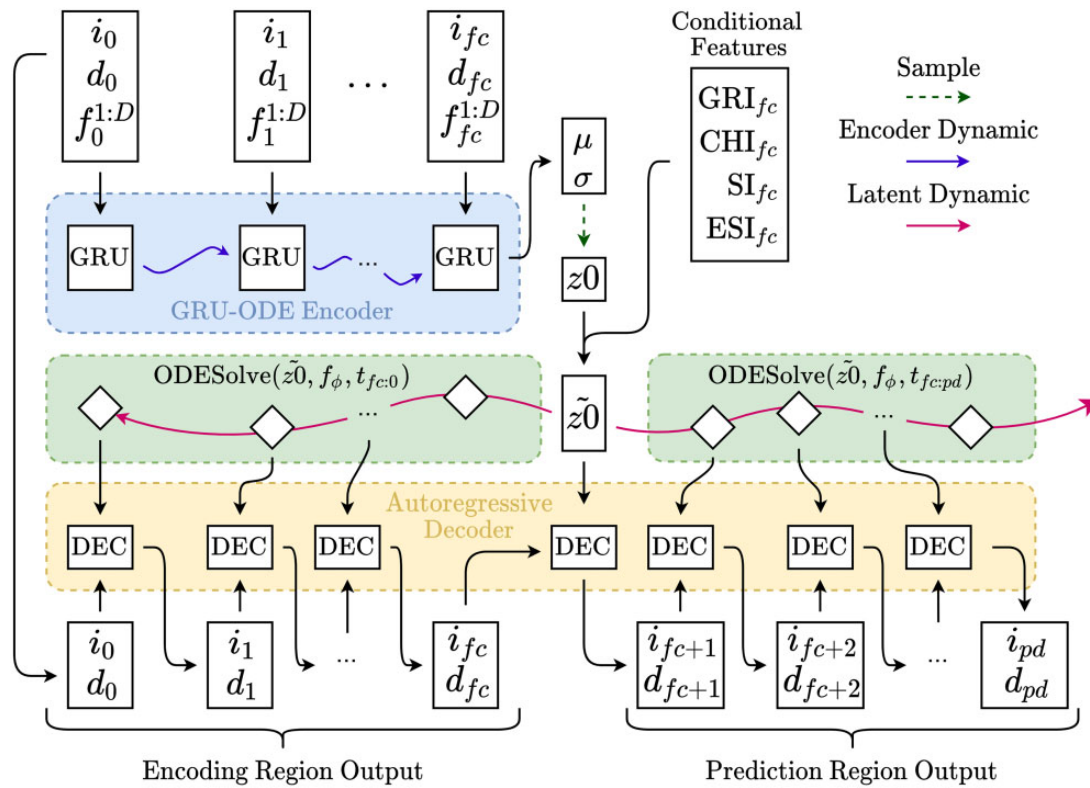


Figure 2. Architectural schematic of the PAN-cODE. We input the daily counts of infection (i), death (d), and other covariates ($f^{1:D}$) up to forecast date f_c . The encoding region time series are encoded by a GRU-ODE encoder, which outputs parameters μ_{z_0} , $\sigma_{z_0}^2$ for the variational posterior on the latent initial state. We obtain a sample latent initial state z_0 from the variational posterior, which is then concatenated with the conditional features to obtain the augmented latent initial state \tilde{z}_0 . From \tilde{z}_0 , the Neural ODE is used to solve the latent trajectory for the prediction and encoding regions, which is passed through the autoregressive decoder to output the predicted daily infection and death count, up to prediction date pd .

alternative caseload trajectories can be generated by modifying the values of I_{f_c} to correspond to modified NPI adoption stringency.

PAN-cODE uses an auto-regressive decoder (ARD) to convert latent trajectories into data space. The ARD is a simple feed-forward neural network combining the latent output at time t with the previously predicted output at time $t - 1$, thus representing predictions as a function of the previous output. This is represented in the equation:

$$(i_t, d_t) = \text{Linear}(i_{t-1}, d_{t-1}, z_t \mid \phi)$$

where parameters ϕ are shared across timepoints. The ARD serves to restrict the maximum change in caseload between timepoints when mapping to the data space.

PAN-cODE is trained to reconstruct Encoding Region observations and make predictions in the Prediction Region. After obtaining augmented latent initial state \tilde{z}_0 , PAN-cODE obtains Encoding Region reconstructions by solving a latent trajectory backwards in time from the forecast date to the initial date and decoding it using the ARD:

$$(i_{f_c:1}, d_{f_c:1}) = \text{ARD}(\text{ODESolve}(\tilde{z}_0, f_\phi, t_{f_c:1}))$$

Next, a latent trajectory is solved forward in time to obtain the caseload trajectory in the Prediction Region:

$$(i_{f_c:pd}, d_{f_c:pd}) = \text{ARD}(\text{ODESolve}(\tilde{z}_0, f_\phi, t_{f_c:pd}))$$

Forecasts for longer prediction windows can be obtained by simply modifying pd .

EXPERIMENTS

Data

We train PAN-cODE using daily COVID-19 caseload counts in US states and counties from Google Cloud Platform’s (GCP) Open Data resource.²³ Roughly 2500 trajectories are available, each with daily observations since February 2020. We apply a 7-day rolling average function to remove reporting noise and apply a shifted log transform of $\log(x + 1)$ for numerical stability.

We provide PAN-cODE with historical NPI adoption stringency as tracked by the Oxford COVID-19 Government Response Tracker.³⁴ The daily status of NPI adoption is represented by a vector of integers denoting the level of adoption for a specific NPI. When forecasting deaths, we add a 14-day delay between these features and COVID-19 caseload to account for the delayed effect of social restrictions.³⁵ We also include the daily temperate index from GCP to account for seasonality in COVID-19. The choice of data smoothing function and feature shift value was selected experimentally and is shown in [Supplementary Appendix C](#).

We use 4 metrics provided by OxGRCT as the conditional features in PAN-cODE. These metrics aggregate the level of NPI adoption in specific areas into a scalar score out of 100, where a higher score indicates higher stringency. These metrics are the Stringency Index, the Government Response Index (GRI), the Containment Health Index (CHI), and Economic Support Index. We standardize these metrics and use their values on the forecasting date as the conditional latent variables. These indices are also used as input covariates.

Hyper-parameters and training

We randomly vary the forecasting date per training epoch to expose PAN-cODE to the correlation between conditional features and future caseloads under various conditions. We first set a validation cut-off date such that data after this date are unseen during training. For each training epoch, we randomly sample an epoch forecast date from between the start date and the validation cut-off date. PAN-cODE is provided data up the epoch forecast date and is tasked with predicting caseload 4 weeks ahead. PAN-cODE is trained by maximizing the ELBO using the Adam optimizer.³⁶ We describe hyper-parameters in [Supplementary Appendix D](#).

Tasks and evaluation

We evaluate the forecasting ability of the PAN-cODE using prediction windows of 4 and 6 weeks. We evaluate against baseline methods using the metrics of median absolute error (MAE) and mean rank, which is computed as the average, across the 51-state level regions, of the rank of the method's MAE in ranked list of the MAE's of all compared methods. We include an evaluation on data from regions unseen during training to highlight the generalization performance of PAN-cODE. We generate forecasts for up to 6 weeks from 2 forecasting dates: December 28, 2020 and March 8, 2021 to evaluate the model's performance under different stages of COVID-19 progression. Data after each forecasting date are not used during training or model selection. The source code for PAN-cODE is available at: <https://github.com/morrislab/PAN-cODE>.

RESULTS

PAN-cODE outperforms baselines on 6-week forecasts

In [Table 1](#), we compare the performance of PAN-cODE on forecasting the number of deaths in US states against our baselines and

competitive methods from the COVID-19 Forecast Hub, described in [Supplementary Appendix A](#). We apply the Wilcoxon ranked-sum test using a *P*-value threshold of .05 against the best performing method for each metric and find that the error in PAN-cODE's forecasts is not significantly larger than that of the best model in every metric. Furthermore, on the March 8, 2021 6-week forecasting task, PAN-cODE provides significantly lower error than all other methods, demonstrating its reliability for long-term forecasting. Daily infection forecasting results are reported in [Supplementary Appendix B](#).

PAN-cODE generalizes to unseen regions

In [Table 2](#), we show the performance of PAN-cODE on countries entirely *unseen* during training. We find that PAN-cODE makes reasonable projections for 4- and 6-week infections forecasts, far outperforming baselines. The VAE-GRU baseline performed extremely poorly in this task and was omitted from our results.

PAN-cODE generates sensible alternative forecasts

In [Figure 3](#), we demonstrate the ability of PAN-cODE to generate alternative outcome trajectories. This ability to generate alternative forecasts can aid policy decisions on NPI adoption for the desired level of transmission reduction. We demonstrate PAN-cODE on both state and county level forecasting and generate alternative trajectories by increasing/decreasing the GRI and CHI measures of NPI stringency. We note that the features for death forecasts are offset by 14 days, meaning the alternative daily death forecasts assume that the change was made 14 days prior.

The importance of features in forecasting

We computed feature importance estimates using LIME³⁷ for the trained PAN-cODE model (see [Supplementary Appendix E](#)).

Table 1. Median absolute error (MAE) and mean rank of the forecasted total deaths for PAN-cODE versus several baseline methods for all 50 states of the United States and the District of Columbia

Model	December 28, 2020 forecast				March 8, 2021 forecast			
	4 weeks ahead		6 weeks ahead		4 weeks ahead		6 weeks ahead	
	MAE	Mean Rank	MAE	Mean Rank	MAE	Mean Rank	MAE	Mean Rank
PAN-cODE	<u>167</u>	<u>7.39</u>	<u>207</u>	<u>3.63</u>	<u>93</u>	<u>8.53</u>	<u>80</u>	<u>4.41</u>
JHU_IDD-CovidSP	210	9.76	312	4.51	95	11.06	120	5.82
Covid19Sim-Simulator	160	8.43	321	4.49	<u>67</u>	<u>9.2</u>	132	5.31
IowaStateLW-STEM	170	8.88	380	4.51	158	13.65	260	7.82
Columbia_UNC-SurvCon	210	9.62	446	4.92	108	11.86	145	6.96
UCLA-SuEIR	283	12.24	383	5.55	159	13.88	212	8.04
USC-SI_kJα	185	8.64	–	–	84	10.55	134	5.72
JHUAPL_Bucky	432	14.59	853	7.33	95	11.18	109	5.96
GRU-ODE	371	12.94	644	6.61	90	11.39	132	6.1
VAE-GRU	964	18.86	1359	9.02	158	13.53	225	7.37
Google_Harvard-CPF	<u>118</u>	6.22	–	–	86	9.49	–	–
UMass-MechBayes	329	11.1	–	–	74	8.12	–	–
CU-select	486	14.2	–	–	82	9.96	–	–
COVIDhub-ensemble	<u>154</u>	6.91	–	–	78	8.13	–	–
Caltech-CS156	<u>156</u>	6.10	–	–	93	9.35	–	–
UCSD_NEU-DeepGLEAM	225	9.91	–	–	75	8.48	–	–
GT-DeepCOVID	131	8.45	–	–	79	9.48	–	–
UA-EpiCovDA	439	14.98	–	–	116	11.69	–	–
TTU-squider	259	10.76	–	–	422	17.45	–	–
Mean COVID Hub	240	–	380	–	122	–	174	–
Baseline (PrevWeek)	218	10.02	343	4.43	120	12.5	295	7.74

Note: Mean rank is computed across the 51 ranked lists of methods ordered by increasing MAE. In each column, the best method is underlined; methods whose MAEs are not significantly worse than this model in a non-parametric paired test (ie, Wilcoxon sign-rank $P > .05$) are indicated in bold italics.

Table 2. Percentage error of the incident infections for PAN-cODE versus our baselines on several countries unseen during training, forecasted from December 28, 2020

Model	4 weeks ahead				6 weeks ahead			
	Canada	United Kingdom	India	Russia	Canada	United Kingdom	India	Russia
Baseline (PrevWeek)	39.0%	66.7%	63.1%	53.4%	103.4%	213.4%	85.9%	93.0%
GRU-ODE	474.4%	36.4%	144.8%	-5.5%	777.7%	156.5%	179.2%	15.4%
PAN-cODE	-10.4%	8.2%	7.2%	0.4%	-51.4%	-21.6%	-55.3%	-55.2%

Note: The lowest percentage error is shown in bold.

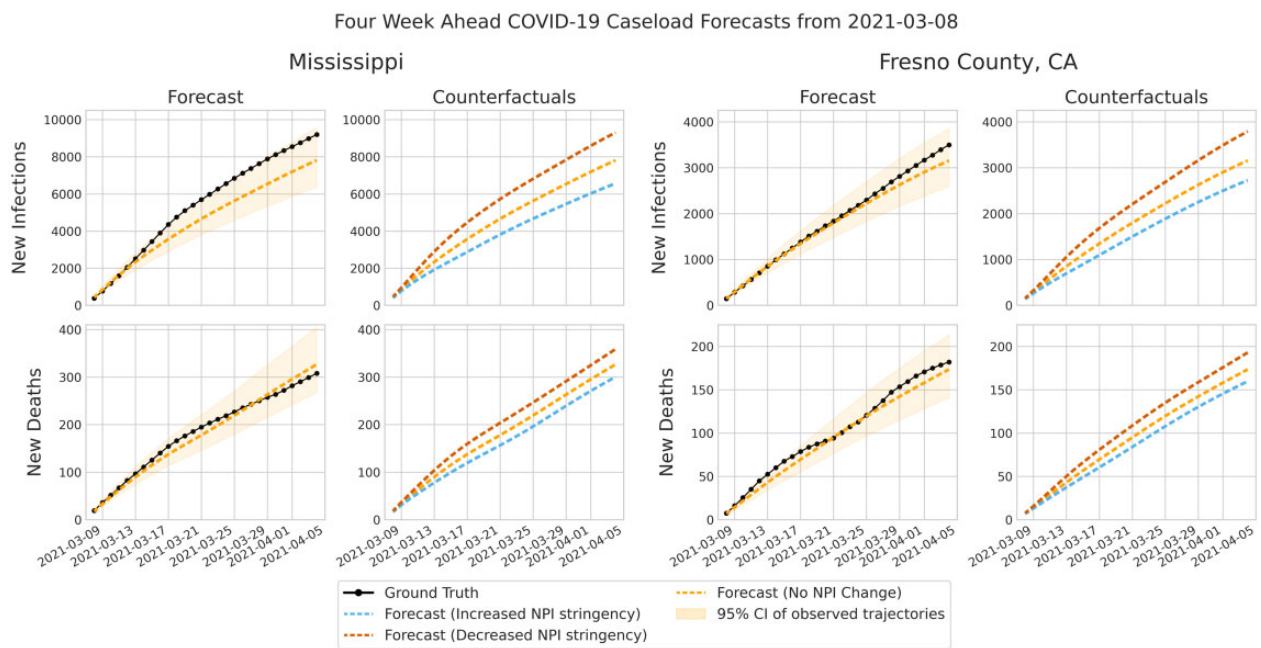


Figure 3. Four week ahead forecasts of COVID-19 infection and deaths at the state (Mississippi) and county (Fresno County, CA) level. The cumulative increase in infections and deaths from the forecast date of May 8, 2021 is shown. PAN-cODE forecasts are conditioned on the OxGCRT Government Response Index, the value of which is modified at the forecasting date to obtain alternative outcome trajectories. We visualize the effects of increasing and decreasing NPI adoption compared to the predicted trajectory, which uses the real NPI adoption magnitude at forecast time. The 95% confidence interval is shown as a shaded area associated with 2 standard deviations from the mean trajectory, as computed by sampling 100 latent trajectories with the modeled observation noise.

The model anticipates that adopting preventative measures has a high impact on infection counts 21 days before the forecast date but that nearer to the forecast date (10 days prior), the actual caseload has the greatest relevance. This behavior is consistent with the known lagged effect of NPIs. Note that different feature importance models make different assumptions about the relationship between local feature perturbations and their impact on the model’s predictions. As such, other feature importance methods, such as Boruta-SHAP³⁸ might find different relationships than LIME did.

DISCUSSION

PAN-cODE offers competitive performance with several advantages over existing methods. While PAN-cODE uses a dynamical model for disease transmission, the actual dynamical function is learned directly from data, avoiding the need to manually specify an ODE function as in traditional compartmental models like the SIR model. However, it would be straightforward to incorporate a SIR compartmental model into PAN-cODE. By using the Neural ODE, PAN-cODE would be able to fit the dynamical parameters of this

SIR model using backpropagation, avoiding the need for sampling or manual parameter specification in, eg, JHU_IDD-CovidSP.³⁹ PAN-cODE also uses fewer data features compared to other methods such as Google_Harvard-CPF.¹⁵ Notably, PAN-cODE does not require mobility or hospitalization data, enabling forecasts in regions where data is limited, or at smaller regional resolutions. Due to the continuous-time Latent ODE architecture, PAN-cODE is also capable of natively handling datasets for pandemics where observations are sparsely or irregularly observed.

PAN-cODE directly learns the correlation between NPI stringency and future caseload without expert input or other manual adjustments. Consequently, the capability for PAN-cODE to provide alternative forecasts can be useful for policy-makers when determining the appropriate level of social restrictions required to obtain the desired level of transmission reduction. In contrast, existing methods capable of alternative outcome forecasting typically rely on manual adjustment of the R0 value in their compartmental model formulation.^{39,40} Historically, the process of manual estimation of pandemic trajectory can be error-prone.⁴¹ Given this complex relationship between NPI adoption and future caseload

trajectories, PAN-cODE's data-driven nature allows our capability to model this relationship to scale with the amount of available data. However, we do note that PAN-cODE does not explicitly learn a causal model between NPI stringency and future caseload. Building a formal causal model is likely difficult due to delayed and noisy reporting, and we leave this as future work.

Ideally, the counterfactuals in Figure 3 would show the effects of individual NPIs. Although we use individual NPIs to help predict the initial latent state, besides this, PAN-cODE's predictions are only conditioned on the 4 stringency measures, because we found that the data on the individual NPIs was not rich enough to support a more precise conditioning. Richer data could be used to train a PAN-cODE model with the ability to model more counterfactuals.

CONCLUSION

We present PAN-cODE, a deep conditional generative approach to pandemic forecasting using COVID-19 as an example. PAN-cODE conditions forecasts on NPI adoption stringency and can generate alternative caseload forecasts for modified NPI adoption stringencies. We demonstrate the performance of PAN-cODE on US state and county caseload forecasting and find it performs significantly better than all existing methods^{10,15,21,25,26,29,39,42-50} on 6-week death forecasting from March 8, 2021 and never performs significantly worse than the best performing method in all other death forecasting evaluation categories. Compared to existing methods, PAN-cODE requires minimal data features, can provide longer-term forecasts, and does not require retraining to generalize forecasts to unseen regions. The fully data-driven nature of PAN-cODE offers a scalable solution to inform public health response for COVID-19 and future pandemic outbreaks.

FUNDING

Resources used in preparing this research were provided, in part, by the Ontario Institute for Cancer Research, the Memorial Sloan Kettering Cancer Center, the National Institute for Health (National Cancer Institute Cancer Center Support Grant P30 CA008748), the Province of Ontario, the Government of Canada through CIFAR, and companies sponsoring the Vector Institute www.vectorinstitute.ai/partners.

AUTHOR CONTRIBUTIONS

RS and HZ implemented the method and experimental pipelines. All authors contributed toward development of the method and writing the manuscript.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

Data on government NPI adoption were obtained from the Oxford COVID-19 Government Response Tracker³⁴ and are available at: <https://github.com/OxCGRT/covid-policy-tracker>. All other data used in this article were obtained from Google Cloud Platform's COVID-19 Open-Data²³ repository and are available at: <https://github.com/GoogleCloudPlatform/covid-19-open-data>.

REFERENCES

- Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020; 20 (5): 533-4.
- Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd SC; 2009 H1N1 Surveillance Group. Improving the evidence base for decision making during a pandemic: the example of 2009 influenza A/H1N1. *Biosecur Bioterror* 2011; 9 (2): 89-115.
- Haug N, Geyrhofer L, Londei A, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat Hum Behav* 2020; 4 (12): 1303-12.
- Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2021; 21 (6): 793-802.
- Aragona M, Barbato A, Cavani A, Costanzo G, Mirisola C. Negative impacts of COVID-19 lockdown on mental health service access and follow-up adherence for immigrants and individuals in socio-economic difficulties. *Public Health* 2020; 186: 52-6.
- Dey M, Loewenstein M. How many workers are employed in sectors directly affected by COVID-19 shutdowns, where do they work, and how much do they earn? *MLR* 2020: 1-19. doi: [10.21916/mlr.2020.6](https://doi.org/10.21916/mlr.2020.6).
- Sahai AK, Rath N, Sood V, Singh MP. ARIMA modelling & forecasting of COVID-19 in top five affected countries. *Diabetes Metab Syndr* 2020; 14 (5): 1419-27.
- IHME COVID-19 Health Service Utilization Forecasting Team; Murray CJL. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. 2020. doi: [10.1101/2020.03.27.20043752](https://doi.org/10.1101/2020.03.27.20043752).
- Nadler P, Arcucci R, Guo Y. Proceedings of the Machine Learning for Health NeurIPS Workshop. *PMLR* 2020; 136: 254-66.
- Srivastava A, Xu T, Prasanna VK. Fast and accurate forecasting of COVID-19 deaths using the SIKJx Model. *ArXiv200705180 Phys. Q-Bio* 2020. <http://arxiv.org/abs/2007.05180>. Accessed May 10, 2021.
- Srivastava A, Prasanna VK. Learning to forecast and forecasting to learn from the COVID-19 pandemic. *ArXiv200411372 Cs Q-Bio Stat* 2020. <http://arxiv.org/abs/2004.11372>. Accessed May 10, 2021.
- Zaman G, Kang YH, Jung IH. Optimal treatment of an SIR epidemic model with time delay. *Biosystems* 2009; 98 (1): 43-50.
- Peng L, Yang W, Zhang D, Zhuge C, Hong L. Epidemic analysis of COVID-19 in China by dynamical modeling. 2020. doi: [10.1101/2020.02.16.20023465](https://doi.org/10.1101/2020.02.16.20023465).
- Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J Thorac Dis* 2020; 12 (3): 165-74.
- Arik S, Li C-L, Yoon J, et al. Interpretable sequence learning for Covid-19 forecasting. *Adv Neural Inf Process Syst* 2020; 33: 18807-18.
- Shastri S, Singh K, Kumar S, Kour P, Mansotra V. Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. *Chaos Solitons Fractals* 2020; 140: 110227.
- Zeroual A, Harrou F, Dairi A, Sun Y. Deep learning methods for forecasting COVID-19 time-series data: a comparative study. *Chaos Solitons Fractals* 2020; 140: 110121.
- Shahid F, Zameer A, Muneeb M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos Solitons Fractals* 2020; 140: 110212.
- Ibrahim MR, Haworth J, Lipani A, Aslam N, Cheng T, Christie N. Variational-LSTM autoencoder to forecast the spread of coronavirus across the globe. *PLoS One* 2021; 16 (1): e0246120.
- Gao J, Sharma R, Qian C, et al. STAN: spatio-temporal attention network for pandemic prediction using real-world evidence. *J Am Med Inform Assoc* 2021; 28 (4): 733-43.
- Ray EL, Wattanachit N, Niemi J, et al. Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S. 2020. doi: [10.1101/2020.08.19.20177493](https://doi.org/10.1101/2020.08.19.20177493).

22. Rubanova Y, Chen RTQ, Duvenaud D. Latent ODEs for irregularly-sampled time series. *ArXiv190703907 Cs Stat* 2019. <http://arxiv.org/abs/1907.03907>. Accessed June 7, 2021.
23. Wahltinez O, Murphy K, Brenner M, et al. COVID-19 Open-Data: curating a fine-grained, global-scale data repository for SARS-CoV-2. 2020. <https://goo.gl/covid-19-open-data>. Accessed September 28, 2021.
24. Kermack WO, McKendrick AG, Walker GT. A contribution to the mathematical theory of epidemics. In: Proceedings of the Royal Society A, vol. 115, no. 772. The Royal Society; 1927. doi: [10.1098/rspa.1927.0118](https://doi.org/10.1098/rspa.1927.0118).
25. Yamana T, Pei S, Kandula S, Shaman J. Projection of COVID-19 cases and deaths in the US as individual states re-open May 4, 2020. *medRxiv* 2020. doi: [10.1101/2020.05.04.20090670](https://doi.org/10.1101/2020.05.04.20090670).
26. Gibson GC, Reich NG, Sheldon D. Real-time mechanistic Bayesian forecasts of Covid-19 mortality. *Epidemiology*, preprint, 2020. doi: [10.1101/2020.12.22.20248736](https://doi.org/10.1101/2020.12.22.20248736).
27. Kapoor A, Ben X, Liu L, et al. Examining COVID-19 forecasting using spatio-temporal graph neural networks. *ArXiv200703113 Cs*, 2020. <http://arxiv.org/abs/2007.03113>. Accessed: June 7, 2021.
28. Cao D, Wang Y, Duan J, et al., Spectral temporal graph neural network for multivariate time-series forecasting. *ArXiv210307719 Cs* 2021. <http://arxiv.org/abs/2103.07719>. Accessed June 7, 2021.
29. Wu D, Gao L, Xiong X, et al. DeepGLEAM: a hybrid mechanistic and deep learning model for COVID-19 forecasting. *ArXiv210206684 Cs* 2021. <http://arxiv.org/abs/2102.06684>. Accessed June 7, 2021.
30. Gatta VL, Moscato V, Postiglione M, Sperli G. An epidemiological neural network exploiting dynamic graph structured data applied to the COVID-19 outbreak. *IEEE Trans Big Data* 2021; 7 (1): 45–55.
31. Chen RTQ, Rubanova Y, Bettencourt J, Duvenaud D. Neural Ordinary Differential Equations. *ArXiv180607366 Cs Stat* 2019. <http://arxiv.org/abs/1806.07366>. Accessed June 7, 2021.
32. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *ArXiv14123555 Cs* 2014. <http://arxiv.org/abs/1412.3555>. Accessed June 7, 2021.
33. Sohn K, Lee H, Yan X. Learning structured output representation using deep conditional generative models. *Advances in Neural Information Processing Systems* 2015, 28. <https://papers.nips.cc/paper/2015/hash/8d55a249e6baa5c06772297520da2051-Abstract.html>. Accessed September 28, 2021.
34. Hale T, Angrist N, Goldszmidt R, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat Hum Behav* 2021; 5 (4): 529–38.
35. Vinceti M, Filippini T, Rothman KJ, et al. Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. *EClinicalMedicine* 2020; 25: 100457.
36. Kingma DP, Ba J. Adam: a method for stochastic optimization. *ArXiv Prepr. ArXiv14126980*, 2014. <https://arxiv.org/abs/1412.6980>.
37. Ribeiro MT, Singh S, Guestrin C. ‘Why should I trust you?’ Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. August 13–17, 2016: 1135–1144; San Francisco.
38. Eoghan K. BorutaShap: a wrapper feature selection method which combines the Boruta feature selection algorithm with Shapley values. 2020. <https://doi.org/10.5281/zenodo.4247618>.
39. Lemaitre JC, Grantz KH, Kaminsky J, et al. A scenario modeling pipeline for COVID-19 emergency planning. *Sci Rep* 2021; 11 (1): 7534.
40. Li ML, Bouardi HT, Lami OS, Trikalinos TA, Trichakis NK, Bertsimas D. Forecasting COVID-19 and analyzing the effect of government interventions. *Oper Res* 2022. doi: [10.1287/oper.2022.2306](https://doi.org/10.1287/oper.2022.2306).
41. Ioannidis JP, Cripps S, Tanner MA. Forecasting for COVID-19 has failed. *Int J Forecast* 2022; 38 (2): 423–38.
42. Bathwal R, Chitta P, Tirumala K, Varadarajan V. Ensemble machine learning methods for modeling COVID19 deaths. *ArXiv201004052 Cs Stat* 2020. <http://arxiv.org/abs/2010.04052>. Accessed June 7, 2021.
43. Rodríguez A, Tabassum A, Cui J, et al. DeepCOVID: an operational deep learning-driven framework for explainable real-time COVID-19 forecasting. *Proc Conf AAAI Artif Intell* 2021; 35 (17): 15393–400.
44. Lega J. Parameter estimation from ICC curves. *J Biol Dyn* 2021; 15 (1): 195–212.
45. Wang Q, Xie S, Wang Y, Zeng D. Survival-convolution models for predicting COVID-19 cases and assessing effects of mitigation strategies. *Front Public Health* 2020; 8: 325.
46. Zou D, Wang L, Xu P, Chen J, Zhang W, Gu Q. Epidemic model guided machine learning for COVID-19 forecasts in the United States. *Epidemiology*, preprint 2020. doi: [10.1101/2020.05.24.20111989](https://doi.org/10.1101/2020.05.24.20111989).
47. Khan ZS, Van Bussel F, Hussain F. A predictive model for Covid-19 spread applied to eight US states. *ArXiv200605955 Phys. Q-Bio* 2020. <http://arxiv.org/abs/2006.05955>. Accessed June 8, 2021.
48. Wang L, Wang G, Gao L, et al. Spatiotemporal dynamics, nowcasting and forecasting of COVID-19 in the United States. *ArXiv200414103 Stat* 2020. <http://arxiv.org/abs/2004.14103>. Accessed June 8, 2021.
49. Kinsey M, Tallaksen K, Obrecht RF, et al. ‘Bucky Model,’ Github Repository. <https://github.com/mattkinsey/bucky>. Accessed May 11, 2022.
50. Chhatwal J, Ayer T, Linas B, et al. ‘COVID-19 simulator,’ COVID-19 simulator. <https://covid19sim.org/>. Accessed May 11, 2022.