

Globally Local: Hyper-local Modeling for Accurate Forecast of COVID-19

Vishrawas Gopalakrishnan, Sayali Pethe, Sarah Kefayati, Raman Srinivasan,
Paul Hake, Ajay Deshpande, Xuan Liu

IBM Watson Health, 75 Binney St, Cambridge, MA, USA 02142

Etter Hoang, Marbelly Davila

Tampa General Hospital, 1 Tampa General Circle, Tampa FL 33605

Simone Bianco and James H. Kaufman

IBM Research, Almaden Research Center, 650 Harry Rd, San Jose CA, 95120

Abstract

Multiple efforts to model the epidemiology of SARS-CoV-2 have recently been launched in support of public health response at the national, state, and county levels. While the pandemic is global, the dynamics of this infectious disease varies with geography, local policies, and local variations in demographics. An underlying assumption of most infectious disease compartment modeling is that of a well mixed population at the resolution of the areas being modeled. The implicit need to model at fine spatial resolution is impeded by the quality of ground truth data for fine scale administrative subdivisions. To understand the trade-offs and benefits of such modeling as a function of scale, we compare the predictive performance of a SARS-CoV-2 modeling at the county, county cluster, and state level for the entire United States. Our results demonstrate that accurate prediction at the county level requires hyper-local modeling with county resolution. State level modeling does not accurately predict community spread in smaller sub-regions because state populations are not well mixed, resulting in large prediction errors. As an important use case, leveraging high resolution modeling with public health data and admissions data from Hillsborough County Florida, we performed weekly forecasts of both hospital admission and ICU bed demand for the county. The repeated forecasts between March and August 2020 were used to develop accurate resource allocation plans for Tampa General Hospital.

Keywords: Epidemiological modeling, SARS-CoV-2, COVID-19, hyper-local modeling

2010 MSC: 92-D30, 91-C20

Introduction

The COVID-19 pandemic caused by the SARS-CoV-2 virus reveals the important contributions that can be made by accurate mathematical modeling of infectious disease. A variety of modeling approaches can be used to predict the course of an infection in a population. While the pandemic is global, the underlying disease dynamics (and the epidemiological parameters themselves) varies with specific local policies, non-pharmaceutical interventions (NPI), variations in demographics, healthcare infrastructure, etc. To realize the full potential of any modeling effort, forecasts and predictions must be accurate and meaningful to public health officials and governing authorities at a resolution relevant to their areas of responsibility¹. In most cases, this requires predictions be made at a hyper-local level. Models can only be used to guide closing or opening of business, predict and allocate Intensive Care Units (ICU) resources, and establish social distancing guidance if their predictions are meaningful at a scale relevant to these businesses, hospitals, and communities.

In practice, the ability to model at high resolution is limited by the granularity and accuracy of ground truth data, as well as the underlying modeling assumptions^{2,3}. Precision of public health SARS-CoV-2 case reports, including incidence and death⁴, is impeded by multiple noise sources⁵. Clinical encounters may take place adjacent to or outside a patients residential district, testing capacity, public health reporting practice, and many other factors may vary dramatically even between nearby districts. Some of these noise sources may be reduced by clustering administrative divisions or modeling at lower resolution, but this approach may violate an assumption used by most compartmental models, that the population being modeled is well mixed⁶.

In this paper, we compare the predictive performance of a SARS-CoV-2 model as a function of spatial scale, applying the model to county, county cluster, and state level divisions. For this purpose, we selected an epidemiological compartment model, as opposed to a regression model, to investigate changes in important disease parameters obtained at each resolution, and to avoid an approach where important parameters are under determined. Predictive power and scaling behavior are expected to depend on model selection. An open source model and modeling framework were chosen so others can reproduce the results and test future extensions that may improve model performance. To evaluate the performance of the selected model we also compare the accuracy to other models with predictions publicly available through the US Center for Disease Control and Prevention (CDC)^{7,8}. These other models include both epidemiological compartment models, regression models, and hybrid approaches. Our results reveal that, for the United States, even for fairly small states, different counties make dominant contributions to daily incidence at different times. The mean average percentage error (MAPE) in predicted incidence can be 30% higher for modeling at the state level compared to aggregating model results at the scale of counties or clusters of counties. Furthermore, county level predictions made by partitioning state level predictions (for a model performing well at the state level) are 3x worse for predicted incidence and 20x worse for predicted deaths

relative to the same model at the county level. These results demonstrate that modeling resolution must be guided by the requirements of the consumer of the predictions. State level modeling strongly violates the assumption of well mixed populations and will lead to inaccurate predictions at a county level even
50 while performing well at the state level. In order to test the performance of the COVID-19 models at the county level, IBM partnered with Tampa General Hospital (TGH) to forecast COVID-19 incidences within Hillsborough County. Hillsborough County accounts for a large portion of the Tampa Florida population. Tampa General Hospital (TGH) is the only Level 1 Trauma Center
55 within the Tampa region. As COVID cases rose in early March, TGH has treated an average of 20.4 percent of the Hillsborough COVID-19 cases with a max distribution of 45.0 percent of all COVID-19 cases in Hillsborough County. The forecasting models would become imperative to ensure adequate strategic planning for hospital capacity, staffing, and supplies as the virus spread.

60 **Methods**

Data Description

USAFacts COVID-19 datasets⁹ including daily COVID-19 incidence and deaths, were obtained directly from CDC, state and local agencies and used for model optimization. This data is further curated; for instance, the post-processing
65 ensures that the number of confirmed cases is a monotonically increasing curve. The data is refreshed every day around 9AM PST to reflect updates of the previous day. More information on the reference data^{10,11,12}, and on the CDC challenge models⁷ used for model comparison is available in the supplement.

Base Model

70 For this study we chose a modeling framework, the SpatioTemporal Epidemiological Modeler (STEM)^{13,14}, available through the Eclipse Foundation¹⁵. The framework and model are open source and available under the Eclipse Public License (EPL2)¹⁶. A number of models for SARS-CoV-2 from multiple authors were evaluated^{17,18,19}. For this study we use the model shown in figure 1, which
75 includes transmission from both asymptomatic (A), pre-symptomatic (C), and infectious individuals (I). Moreover, we capture infectious individuals who experience a worsening of symptoms (W) before either dying of the disease (D) or recovering (R). This model is a simple extension to the SACIR model available on the Eclipse site.
80 Model parameters and initial conditions were selected by a combination of literature references, grid search, and data-driven estimation through a Nelder-Mead AI simplex algorithm available in the STEM framework. The flow diagram of figure 1 translates into a system of ODE's reported in the supplemental Eq. 2. The definition of each compartment, all model parameters, literature values, and
85 methods of parameter estimation are discussed in the supplement. The basic epidemiological model was used to forecast at three spatial resolutions (counties, county clusters, and states). Definition of county clusters is described in the supplement.

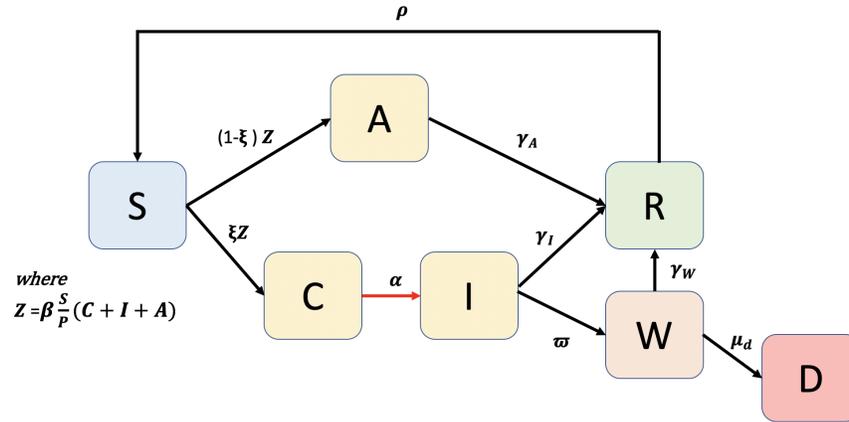


Figure 1: Compartment model use to capture the disease dynamics of SARS-CoV-2 (see text).

Model performance was measured using the normalized root mean square error (nrmse) of daily incidence, cumulative incidence, and cumulative deaths with respect to the corresponding ground truth data one to four weeks into the future. Since public health reporting usually takes place after a clinical encounter, the transition between the pre-symptomatic compartment (C) and the infectious compartment (I) was used to log daily incidence (red arrow, figure 1). Here we introduce a pre-symptomatic compartment C in place of the usual exposed compartment E based on documented evidence for replication-competent virus obtained from patients before the onset of symptoms^{20,21,22,23,24}. Thus, we describe the rate of individuals leaving the C compartment as a symptom appearance rate as opposed to an incubation rate. The processes are similar in that they both would contribute to a period of latency between exposure and clinical encounter. We chose the nomenclature to indicate that individuals in the C compartment contribute to the disease force of infection.

The steady state solution to the differential equations 2 shown in the supplement provides the following expression for the basic reproductive number, R_0 as a function of the epidemiological parameters. Background birth and death rates are omitted. Individuals in the W compartment do not contribute to the force of infection.

$$R_0 = \frac{\beta}{\gamma_A} \left\{ 1 + \left(\frac{\gamma_A}{\gamma_I + \omega} \right) - \left[1 - \frac{\gamma_A}{\alpha} \right] \xi \right\} \quad (1)$$

To provide a scalable architecture for modeling, we built an automation pipeline that performs a variety of tasks ranging from data ingestion, smoothing, model invocation, and post-processing. The pipeline supports multi-processing, where each region is run by a separate instance of a docker image; allowing easy scaling of the clusters as required. This pipeline is discussed in the supplement.

Hospitalization and Intensive Care Unit Forecasting Model

In order to deliver actionable projections for decision making by healthcare
115 providers, as a post process we decomposed the incidence projections using rate
equations to forecast both total hospitalization and ICU demand at the county
level. The average rate (weighted by incidence) of hospitalization per case (in-
cidence) was 0.35 ± 0.04 and the average rate of ICU admission per hospitalized
patient as 0.10 ± 0.05 .

120 Tampa General Hospital in Hillsborough County, Florida, USA, implemented
this model to anticipate future resource demand from COVID-19 patients and
to advise decisions regarding possible need for extra capacity and reduction in
planned surgeries to free up bed space. They established a regional collabora-
tion for data and resource sharing facilitated by shared data collection and a
125 reporting dashboard. More information on this model is available in the sup-
plement. Projections based on both models were updated at least weekly since
April 2020.

Results

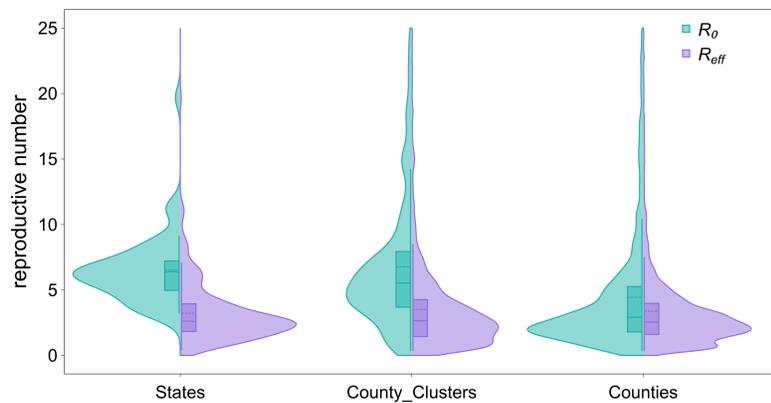
Derived Epidemiological Parameters

130 The model optimization process automatically adds discrete changes to the
transmission rate β , at discrete times, based on the dynamics observed in ground
truth incidence and deaths. In an attempt to distinguish between a "natural"
range of transmission rates, β , reflecting human contact rates in the absence of
interventions, and the larger range of β with interventions, the violin plots in
135 figure 2a are split. The left hand (cyan) distribution includes only the maxi-
mum β observed by region. The right (purple) distributions include all β values
(i.e. all changes in β from local policies and NPI, as well as the maximum
values). Similarly the distribution of R_0 derived from equation 1) is shown
in figure 2b. These distributions are also split with the left hand violin dis-
140 playing R_0 values computed from maximum β values observed by region. The
right hand violin plot (labeled R_{eff}) includes reproductive numbers computed
independently from *all* β values observed in all regions. The distribution of
transmission rates and reproduction numbers are multi-modal reflecting differ-
ences in disease dynamics between regions. Distributions for other parameters
145 and a table listing L-statistics for all optimized parameters and reproduction
numbers may be found in the supplementary Table 4 and 5 respectively. The
Table also shows the average model convergence error (nrmse) which was under
4% for simulation at the scale of U.S. States, systematically increasing to just
over 10% for simulation of U.S. counties.

150 The optimized parameter values and derived values for the basic and effective
reproduction numbers are within the range estimated from current clinical stud-
ies.^{20,21,22,23,24} Supplementary Table 5 provides statistics from the measured
distribution of reproduction number values obtained using the epidemiological
parameters in Table 4 at each spatial resolution.



(a) Transmission Rate



(b) Basic Reproductive Number

Figure 2: Regional distributions of transmission rates and reproduction numbers at different modeling resolutions. The violin plots are split to visualize both the distribution of maximal regional values and all regional values to show how non-pharmaceutical interventions may have lowered transmission rates and reproduction numbers (see text).

155 *Model Performance and Predictions*

Figure 3 shows example output of the epidemiological model for two states. The model was run at county, county cluster, and state resolution for all states (and DC). Figure 3 shows the actual and predicted (daily) incidence and (cumulative) deaths vs. time all county clusters within each state respectively. The data is presented as a stack chart so the envelop over aggregated clusters reflects total incidence and deaths at the state level. This is also indicated by the blue line above the stacked sub-regions for both the actual (reported) in-

160

It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

165 incidence and deaths, and the corresponding predictions. In each sub-figure the white background indicates the reference time period and data used for model optimization. The gray background shows actual and predicted incidence and deaths for subsequent days. This latter region was used to compute the MAPE.

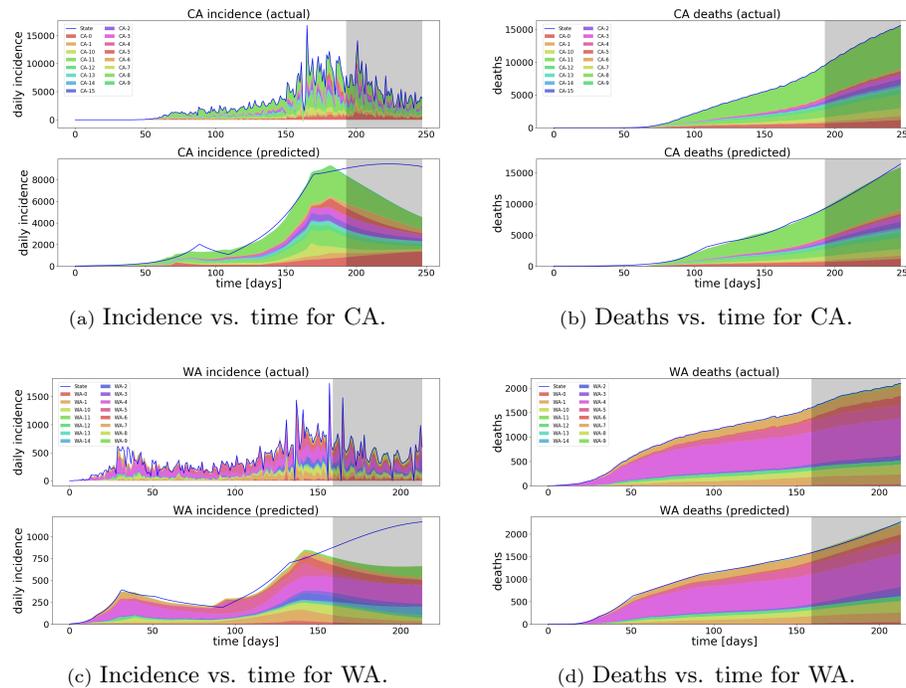


Figure 3: Actual and predicted (daily) incidence and (cumulative) deaths vs. time for county clusters within states. The white background shows the time period used for model optimization. The gray background shows incidence and deaths for subsequent days. The relative contribution to state level incidence or deaths from different sub-regions varies over time.

The data for California and Washington State (figures 3a and 3c) illustrate how the noise in daily incidence can adversely impact predictions made by an automated pipeline. This is evident as the model predictions at the state level (blue curve) deviate from the ground truth data after a particularly large noise spike. Predictions at individual and state aggregate county clusters do not show the same deviation. The stack charts also reveal that the relative contribution from different sub-regions to state level incidence or death varies over time indicating the states themselves are not accurately described as well mixed.

170 Prediction accuracy as a function of spatial scale is expected to be model dependent. Therefore, before evaluating the effect of modeling resolution on statistical error, we sought to test the predictive accuracy of the SACIR model relative to other models at the same resolution to establish its baseline performance.

175 For this purpose we selected models from the CDC challenge project⁷. The

majority of these models provide predictions at the state level. The pipeline predicts daily and cumulative incidence and cumulative deaths for all states at least four weeks into the future. Figure 4a, shows a self-consistent comparison of the accuracy of prediction, at state level resolution, for all models that predict both incidence and deaths at state level, for all states, and with data provided 1-4 weeks into the future. As discussed in the Methods section, some models were omitted from the comparison based on missing data.

Figures 4a and 4b compare the statistical accuracy of several models, averaged over all states, at one to four weeks beyond the time of optimization. The order of each bar is based on the MAPE (low to high) after the first week. For each model the average MAPE increases with time after the optimization window. The baseline predicted performance, when run a state level resolution, of the SACIR model compares favorably with the other CDC Challenge models run at the same resolution for the same ground truth data and time.

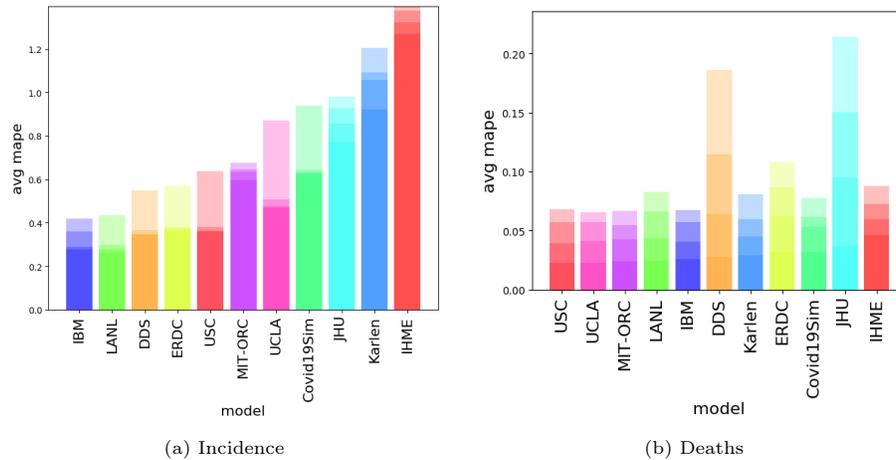


Figure 4: Mean absolute percentage error (MAPE) by model after 1-4 weeks. Models are sorted by prediction accuracy after one week

Effects of Spatial Scale

Given the state level baseline from figure 4a, we next compare the MAPE for predictions as a function of spatial scale. Figures 5a and 5c test the assumption that spatial regions defined by states represent well mixed populations. The figures show the statistical error in predicted incidence and predicted deaths, respectively, as a function of time. The red curves correspond to the average MAPE for models run at the state level (i.e. treating each state as single well mixed regions). The blue curves correspond to the average over states of the state level MAPE for models run at the county-cluster level. In this case the prediction for each state was obtained by summing all of the county cluster level predictions within the state, and the state level MAPE computed by comparing the sum to the ground truth data for the state. The green curves

were computed in the same way as the blue curves but running the model at county resolution. From figure 5a it is clear that modeling at either county or county-cluster resolution lowers the mean absolute percent predictive error by 30% for each of the four future weeks. County and county cluster resolutions yield comparable results. All three resolutions provide comparable predictive accuracy for cumulative deaths (figure 5c) at least in the first two weeks. State level prediction accuracy of cumulative deaths degrades slightly faster after two week.

Figures 5b and 5d test the complementary hypothesis. If larger regions are truly well mixed, then prediction at at higher resolution could be obtained by simply dividing or distributing predictions made for larger regions by the fraction of population in each sub-region. Figures 5b and 5d show the average MAPE for three methods to predict incidence or deaths at the county level. In each of the two figures, the red curve shows the average MAPE for county level predictions obtained by dividing the state level model output to obtain county level predictions. The blue curve represents county level predictions obtained by dividing county cluster level model output. The green curve is simply the average MAPE from modeling at the county level itself (not aggregated by state as in figures 5a and 5c).

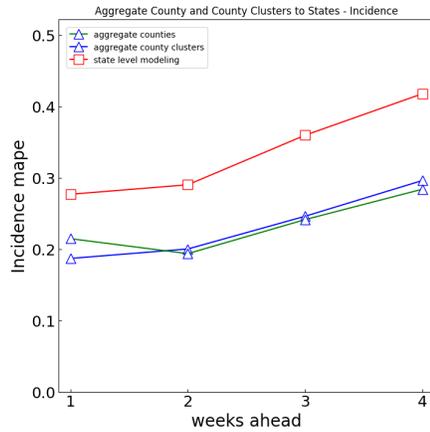
From the data in figures 5b and 5d, we list in Table 1 the ratio of MAPE in prediction of county incidence or deaths based on state or county-cluster regions relative to modeling at the county level. The error ratios decrease with time (prediction week) as the MAPE increases for all resolutions over time. Looking at predictions one week into the future, county level predictions based on state level modeling have $\sim 3.3x$ higher error for predicted incidence and over 20x higher error for predicted deaths. Modeling counties at the county-cluster level, compared to county level, results in 1.8x higher error for predicted incidence and over 8x higher error for predicted deaths.

Table 1: Table increasing MAPE (ratio) as spatial resolution is decreases from county to state (see Figures 5b and 5d).

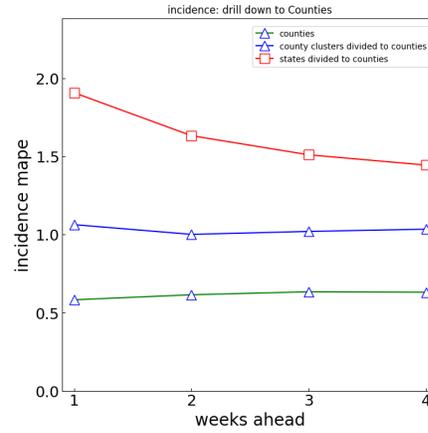
Region Resolution	Week 1	Week 2	Week 3	Week 4
Incidence				
state/county	3.3	2.7	2.4	2.3
county-clusters/county	1.8	1.6	1.6	1.6
deaths				
state/county	20.6	14.0	10.6	8.9
county-clusters/county	8.1	5.6	4.3	3.7

To demonstrate the practical implications of our analysis, we report the forecasts created to predict hospital admissions, and ICU demand for the Tampa General Hospital (TGH) network. Forecasts used predicted incidence data for Hillsborough County, and were performed at least weekly beginning in April 2020. Figure 6 shows actual and predicted COVID-19 ICU patients (magenta and blue solid curves respectively). The blue points indicate independent future projections. Total ICU patients, actual and predicted, are indicated by magenta and blue dashed curves respectively. Total ICU prediction in figure

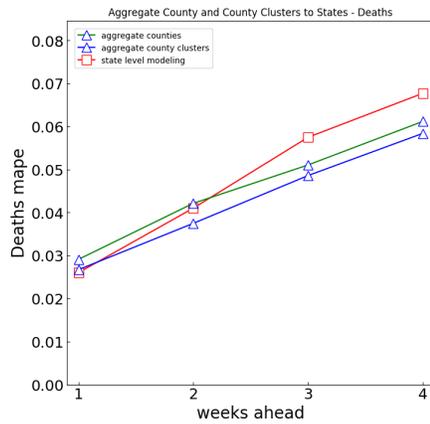
It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).



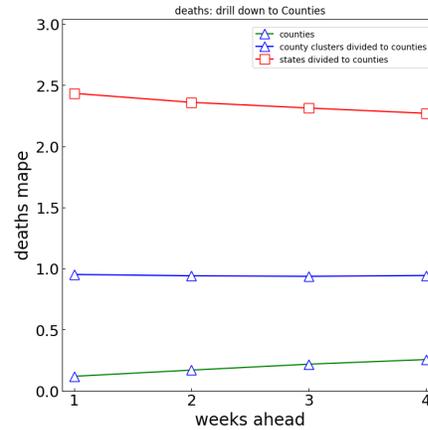
(a) Aggregating county cluster incidence predictions provides more accurate results at the state level.



(b) Dividing state level incidence predictions into smaller regions (by population) leads to much higher prediction error



(c) Aggregating county cluster death predictions provides more accurate results at the state level.



(d) Dividing state level death prediction into smaller regions (by population) leads to much higher prediction error

Figure 5: Comparing aggregation of smaller counties or county clusters provides more accurate state level prediction of deaths and incidence. Conversely, subdividing state level model results into counties or county clusters fails, demonstrating that most states are not approximated as well mixed regions.

6 was made by adding the predicted COVID-19 demand to a seven day moving average of non-COVID actual cases. The blue zone above and below the total ICU prediction represents the MAPE. The MAPE over the entire time range was 5.9%. The time-varying ICU bed capacity for the county is shown in red. Predicted and actual total ICU demand reached a maximum in June

245

(before the peak COVID-19 demand) and never exceeded the ICU bed capacity of ~ 350 beds at the time. Based on the total ICU demand projections, the hospital made the decision not to cancel other elective procedures and surgeries in the time period shown. Projections for county incidence, total hospital admissions, and statistical error (MAPE) for individual projections are available in the supplement.

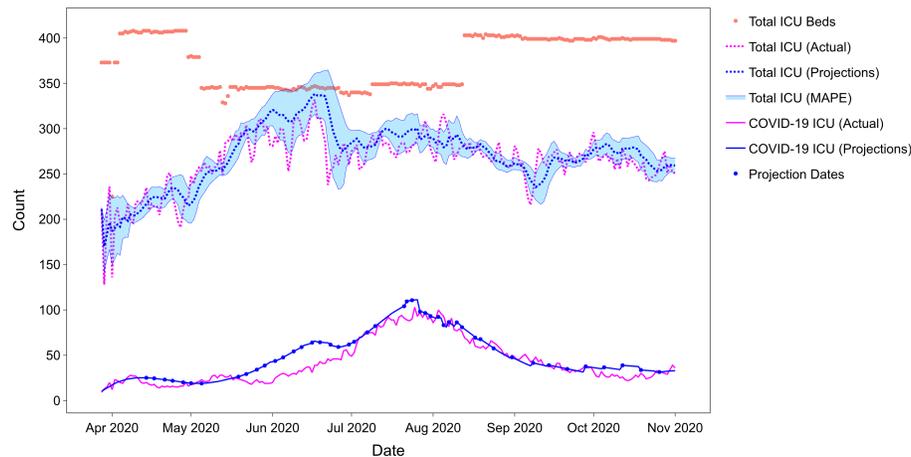


Figure 6: ICU Status for Hillsborough County, FL. The red points indicated capacity (which varied over time). The figure shows COVID-19 ICU Patients, all ICU Patients, and the full series of COVID-19 ICU projections made between April and November 2020.

Discussion

The data in figures 4a and 4b provides a quantitative comparison of the performance of several different models, including the current SACIR model, based on the mean absolute percent prediction error one to four weeks into the future. These models each use different approaches and make different assumptions. Some are compartmental epidemiological models, some are based on regression, and some adopt a hybrid approach. The goal of this paper is not to identify a best model or paradigm, but to explore the question of modeling resolution itself.

Figures 4a and 4b compare the predictive power of models at several spatial resolutions. Modeling at a spatial scale corresponding to administrative divisions of states may have advantages in terms of reduced computational complexity and cost. While public health reports by county, lowering spatial resolution allows aggregation of inherently noisy data which contributes to computational complexity and cost. However, if public health officials and policy makers require accurate insights at county level, modeling at the state level may not suffice.

Figure 3 and associated data in the supplement demonstrates - from raw data
270 alone - that different sub-regions within states make dominant contributions
to aggregate state level incidence and deaths at different times. Figure 5 and
Table 1 show that the statistical error (MAPE) in predicted incidence at the
state level is 30% higher for modeling states compared to modeling at the level
or counties or clusters of counties. Furthermore, testing the well mixed popu-
275 lation assumption, the same figures reveal that county level predictions made
by partitioning state level predictions (for a model performing well at the state
level) are 3x worse for predicted incidence and 20x worse for predicted deaths
relative to the same model at the county level. Figures 5a and 5c do indicate
the possibility of diminishing returns as average MAPE in predicted incidence
280 and deaths has similar (or even slightly lower) error at county-cluster resolution
than at county resolution. Aggregating data to county clusters does provide
noise reduction. Some of this noise is associated with artifacts of public health
case reporting where incidence may be reported based on clinical location as
opposed to patient residence. In other cases reporting location may be skewed
285 by local public health practices. This is evidenced by public health data made
available early in the pandemic where reported case counts and deaths were
zero in Bronx, NY, while the epidemic grew exponentially in all adjacent bor-
oughs²⁵. Local reporting bias suggests that the optimal trade-off between noise
reduction and spatial resolution may vary locally and even change over time
290 as public health systems become more efficient over the course of an extended
pandemic. Whatever decision one makes in optimizing modeling performance,
the measure of success must be based on the accuracy achieved at the spatial
resolution required by the consumers of the data. In the case of TGH Hills-
borough, modeling at the county level the statistical error was < 6% (figure 6).
295 Had the county prediction been based on state level modeling, the uncertainty
would have increased by a factor of ~3.3x to almost 20% (see Table 1), greatly
reducing confidence in the forecast.

With relative accuracy from the model performance, TGH was able to consume
the COVID forecasting model in two primary ways. In mid-June, as COVID-
300 19 cases began a rapid rise in Hillsborough County, TGH leadership faced a
difficult decision on whether or not to begin canceling operations to make ca-
pacity for the surging COVID-19 cases. As demonstrated in figure 6, the model
accuracy improved in June between actual vs forecast, and predicted a peak
incidence on 7/20/2020 and peak ICU on 7/26/2020. TGH then used the Hills-
305 borough COVID model projections to define a TGH specific hospital capacity
forecast and to inform leadership that TGH would have enough capacity to
handle the surge. As other health systems were canceling operations, based
on information the models presented, TGH leadership decided to continue sur-
gical procedures allowing for continuity in patient care and financial stability
310 of the health system. With the data monitored 2-3x a week, TGH was able
to strategically allocate resources, supplies, and staff to accommodate for the
surge. The model was directionally reliable enough to generate and execute
on this strategic COVID capacity plan. Note, the forecasts made in late July
should not be taken as predictive of the likelihood future epidemic waves will

315 or will not exceeding ICU capacity. The outcome of this work demonstrates
that epidemiological forecasting, when performed at the county level can have a
significant impact on the strategic planning of a COVID response for a county
and a hospital.

Acknowledgements

320 The authors would like to acknowledge valued discussions with Kun Hu and
Stefan Edlund.

Author contribution

JHK and AD conceived the work, JHK, VG, and SP created the model, VG
and SP created the automation pipeline, VG, SP, RS, JHK, SB, PH and AD,
325 analyzed the USAFacts COVID data, JHK, SB, VG, SP and SK performed the
COVID model analysis, EH, MD, PH and SK analyzed the ICU data, EH, MD,
and PH performed the ICU forecast, all authors wrote the manuscript.

References

- 330 [1] Haffajee RL, Mello MM. Thinking Globally, Acting Locally —
The U.S. Response to Covid-19. *New England Journal of Medicine*
2020;382(22):e75. URL: <http://www.nejm.org/doi/10.1056/NEJMp2006740>. doi:10.1056/NEJMp2006740.
- [2] Chin V, Samia NI, Marchant R, Rosen O, Ioannidis JP, Tanner MA,
et al. A case study in model failure? COVID-19 daily deaths and
335 ICU bed utilisation predictions in New York state. *European Journal of Epidemiology* 2020;35(8):733–42. URL: <https://doi.org/10.1007/s10654-020-00669-6>. doi:10.1007/s10654-020-00669-6. arXiv:2006.15997.
- [3] Voss J. Encoding changing country codes for the semantic web with iso 3166
340 and skos. In: *Metadata and Semantics*. Springer; 2009, p. 211–21.
- [4] Woolf SH, Chapman DA, Sabo RT, Weinberger DM, Hill L, Taylor DDH.
Excess Deaths From COVID-19 and Other Causes, March-July 2020. *JAMA*
2020;324(15):1562–4. URL: <https://doi.org/10.1001/jama.2020.19545>.
doi:10.1001/jama.2020.19545.
- 345 [5] Bergman A, Sella Y, Agre P, Casadevall A. Oscillations in U.S. COVID-
19 Incidence and Mortality Data Reflect Diagnostic and Reporting Factors.
mSystems 2020;5(4):1–6. doi:10.1128/msystems.00544-20.
- [6] Tolles J, Luong T. Modeling Epidemics with Compartmental Models. 2020.
URL: <https://jamanetwork.com/journals/jama/fullarticle/2766672>.
350 doi:10.1001/jama.2020.8420.

- [7] Cdc activities and initiatives supporting the covid-19 response and the president's plan for opening america up again. <https://www.cdc.gov/coronavirus/2019-ncov/downloads/php/CDC-Activities-Initiatives-for-COVID-19-Response.pdf>; 2020. Accessed: 2020-10-04. 355
- [8] Ray EL, Wattanachit N, Niemi J, Kanji AH, House K, Cramer EY, et al. Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S. medRxiv 2020;:2020.08.19.20177493URL: <https://doi.org/10.1101/2020.08.19.20177493>.
- [9] Us coronavirus cases and deaths. <https://usafacts.org/visualizations/coronavirus-covid-19-spread-map>; 2020. Accessed: 2020-10-04. 360
- [10] Adhikari S, Pantaleo NP, Feldman JM, Ogedegbe O, Thorpe L, Troxel AB. Assessment of community-level disparities in coronavirus disease 2019 (covid-19) infections and deaths in large us metropolitan areas. JAMA network open 2020;3(7):e2016938-. 365
- [11] Lasry A, Kidder D, Hast M, Poovey J, Sunshine G, Zviedrite N, et al. Timing of community mitigation and changes in reported covid-19 and community mobility—four us metropolitan areas, february 26–april 1, 2020. CDC Stacks 2020;. 370
- [12] Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the covid-19 pandemic. Nature 2020;584(7820):262–7.
- [13] Douglas JV, Bianco S, Edlund S, Engelhardt T, Filter M, Günther T, et al. Stem: An open source tool for disease modeling. Health security 2019;17(4):291–306. 375
- [14] Edlund SB, Davis MA, Kaufman JH. The spatiotemporal epidemiological modeler. In: Proceedings of the 1st ACM International Health Informatics Symposium. 2010, p. 817–20.
- [15] The eclipse spatiotemporal epidemiological modeler (stem) website. <http://www.eclipse.org/stem/>; 2006-2020. 380
- [16] Alamoudi E, Mehmood R, Aljudaibi W, Albeshri A, Hasan SH. Open source and open data licenses in the smart infrastructure era: Review and license selection frameworks. In: Smart Infrastructure and Applications. Springer; 2020, p. 537–59. 385
- [17] Schwartz IB, Kaufman JH, Hu K, Bianco S. Predicting the impact of asymptomatic transmission, non-pharmaceutical intervention and testing on the spread of covid19 covid19. medRxiv 2020;.

- 390 [18] Gurbaxani B.M. Schwartz I.B. ea. Sars-cov-2 open source models. <https://wiki.eclipse.org/STEM>; 2020.
- [19] Eclipse stem wiki (full documentation). <https://wiki.eclipse.org/STEM>; 2006-2020.
- 395 [20] Organization WH, et al. Modes of transmission of virus causing covid-19: implications for ipc precaution recommendations: scientific brief, 27 march 2020. Tech. Rep.; World Health Organization; 2020.
- [21] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (CoVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine* 2020;172(9):577–82. doi:10.7326/M20-0504.
- 400 [22] Liu Y, Yan LM, Wan L, Xiang TX, Le A, Liu JM, et al. Viral dynamics in mild and severe cases of covid-19. *The Lancet Infectious Diseases* 2020;.
- [23] Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of sars-cov-2 in different types of clinical specimens. *Jama* 2020;323(18):1843–4.
- 405 [24] Wolfel R, Corman V, Guggemos W, et al. Virological assessment of hospitalized cases of coronavirus disease. *Nature* 2020;581:465–9.
- [25] Dong E, Du H, Gardner L. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases* 2020;20(5):533–4.

Supplement

Supplemental Methods

410 *Supplemental Data Description*

USAFacts COVID-19 datasets⁹ were used for model optimization. The USAFacts dataset is widely used for modeling and analyses by many researchers^{10,11,12} and is also used in aforementioned CDC challenge⁷. In addition to the daily cumulative cases of incidences and death at county level, it also records the
415 number of cases that are currently "unassigned" to any county in a state, thus, enabling one to roll up the numbers to state level for modeling.

Hospitalization and Intensive Care Unit Forecasting Model

To support regional collaboration, COVID-19 hospital and ICU admissions were computed from daily predicted incidence as a post process using rate equations
420 determined from local hospital data. Hospital data and models do not represent any patient specific information (no PHI). All data sets used are publicly available through Florida Public Health via an open API that provides aggregated data at the Hillsborough county level¹. Rate parameters were calculated and fitted using the local hospital data and observations related to hospitalized
425 patients with COVID-19. Rate equations were updated and re-fitted approximately weekly, and more frequently during the second incidence peak in July. Over each independent projection of incidence, the average rate (weighted by incidence) of hospitalization per case (incidence) was 0.35 ± 0.04 and the weighted average rate of ICU admission per hospitalized patient as 0.10 ± 0.05 . The rates
430 for each projection period are available in the supplemental data.

Forming County Clusters

A major concern for modeling at fine grained resolution is possible sparsity of data. For example, cases from one county may be aggregated with and reported in a neighboring county. One solution is to cluster neighboring regions
435 and to aggregate data at the fine grained granularity. However, care should be taken that the process of amalgamating these finer resolution areas does not dilute the dynamics of the epidemic and/or violate the assumption of a well mixed population.

Towards this end, one could use statistical areas like Metropolitan Statistical
440 Area (MSA) as defined by U.S. Census statistics or like Combined Statistical Area (CSA) and Core-Based Statistical Area (CBSA) as defined by Office of Management and Budget. However, the challenge in using these county groupings is that such statistical area spans multiple states. Furthermore, these statistical areas do not cover all the counties in the U.S.

445 Consequently, we need to develop a custom mechanism to aggregate the counties in such a manner that it:

- Preserves the population mixing characteristics,
- Is sensitive to state boundaries,
- Dynamically identifies the number of correct clusters/groupings of the
450 counties in each state.

A deeper perusal of these above requirements highlights two important facets in formulating this problem - (i) A need to capture the people-people interaction between counties, and (ii) Whether there exist an inter-county affinity or whether the population mixing is primarily intra-county. With regard to the
455 second point, an example of the former (i.e., inter-county affinity) would be suburban regions and satellite township centred around large or economically major cities while for the latter (i.e., inter-county affinity) it would be large rural counties which usually tend to be self contained. It is worthwhile to reiterate here that although this definition is very similar to that of the CBSA, CBSA
460 does not respect state boundaries nor covers all the regions in the U.S. Hence, we use the census bureau's 2011-2015 5-Year ACS commuting flows between residence county to workplace county². This data is transformed into a weighted graph (with self-loops), where the weights are the commuter flows. The resulting asymmetric graph is then symmetrized. We use Louvain modularity
465 algorithm²⁴ that extracts communities using a greedy optimization approach. The algorithm first detects smaller communities and then progressively aggregates them if the resultant modularity value of the agglomeration is more than that of the individuals. This allows us to avoid specifying the number of communities/clusters to be extracted in advance. In other words, the groups of counties
470 are derived purely based on the work-site to residence mobility patterns. This approach results in a total of 685 groups of counties which we call county clusters.

Model Definition

The compartments shown in figure 1 are listed in Table 2. The flow diagram of
475 figure 1 translates into a system of ODE's reported in Eqns. 2.

Table 2: Table of epidemiological compartments for the model in Figure 1.

Compartment	Description	Initialization
S	Susceptible	grid search
A	Asymptomatic infectious	algorithmic
C	Pre-symptomatic infectious	algorithmic
I	Infectious (clinical)	grid search
R	Recovered	grid search
W	Worsening Symptoms (not infectious)	0
D	Disease Death	0

$$\begin{aligned}\frac{dS}{dt} &= -\beta \frac{S}{N} (I + A + C) + \rho R \\ \frac{dA}{dt} &= (1 - \xi) \left[\beta \frac{S}{N} (I + A + C) \right] - \gamma_A A \\ \frac{dC}{dt} &= \xi \left[\beta \frac{S}{N} (I + A + C) \right] - \alpha C \\ \frac{dI}{dt} &= \alpha C - (\omega + \gamma_I) I \\ \frac{dW}{dt} &= \omega I - (\mu_d + \gamma_W) W \\ \frac{dR}{dt} &= \gamma_I I + \gamma_A A + \gamma_W W - \rho R \\ \frac{dD}{dt} &= \mu_d W\end{aligned}\tag{2}$$

An important variable in the SACIR model is the case reporting rate, ξ . This parameter partitions new incidence across two paths; symptomatic infectious and asymptomatic infectious. The name suggests that asymptomatic cases are not reported whereas symptomatic cases are. However, the true fraction of reported cases is affected by a convolution of factors (including, but not only, the rate of asymptomatic infection). These includes cases that were truly asymptomatic, cases with symptoms so mild as to not result in an clinical encounter, and/or the efficiency of reporting process itself. For seasonal influenza the fraction of cases registered by public health has been reported to be as low as 3% in some jurisdictions²⁹. The asymptomatic fraction of SARS-CoV-2 has been reported to be in a very wide range of 15-90%^{52,55,27,43}. The reporting parameter is also affected by the underlying assumption that the population in a region being modeled is well mixed. If one constructs a model where a very large region, hence where the density of people is low, is treated as one well mixed node, this is a source of prediction error. Consider, for example, a patch model where a large country (Admin level 0) is treated as one well mixed region. The model puts the entire population of that country in one patch. But, if the actual epidemic is concentrated in one sub-region or city, that scenario has a built in error. Since the entire population is not well mixed, the denominator in the mass action terms, that is, the parameter N in Eqs. 2, should be the population of the affected province or city - not the whole country. In treating reference data from a single city as if it represents an entire country, simplex solver will compensate by decreasing the case reporting rate to an unrealistically low value. Therefore, derived parameter values can be influenced by artifacts related to spatial resolution. The A compartment (see figure 1) includes all non-reported incidence.

The symptom appearance rate (see Table 4) contributes a latency between infection and reporting (similar to incubation rate in an SEIR model) with pre-symptomatic individuals also contributing to the force of infection. Case report and mortality data indicate a delay between the infectious period and reported

disease mortality. The base SACIR model does not capture this additional latency. Therefore another compartment (W) was added between (I) and (D), as shown in figure 1. This compartment, along with the corresponding "worsening rate", ω , provides for that latency. Moreover, it also describes the fraction of patients whose symptoms become more severe after the clinical infectious period, and is a proxy for demand of clinical intensive care units (ICUs). Individuals in this compartment do not contribute the force of infection, as they are assumed to be clinically isolated in an ICU. Background birth and death rates are assumed to be zero. The differential equations for the SACIWR model are shown below. Model integration was performed with an order 8(5,3) Dormand-Prince solver from the Apache software foundation, and implemented in the STEM framework³³.

Model Parameters and Parameter Estimation

Epidemiological parameters were obtained by a combination of Nelder-Mead simplex, grid search, and grid search around literature estimates as shown in Table 4.

As R_0 is a ratio of transmission rate to recovery rate(s), (please refer to Equation 1), we need to vary only one of these two parameters to avoid over fitting. We chose to vary the transmission rate and set the recovery rate(s) based on literature estimates. Please also recall that in our approach, the transmission rate is time sensitive in the sense that we have multiple transmission rate to account for the different interventions. However, the recovery rate(s) are the same for the entire time period, as it is the characteristics of the disease and there has not been a significant change in the treatment regimen. A recent paper⁴⁰ has estimated that the recovery time of the asymptomatic individuals is roughly the same as that of those who exhibit symptoms, which is about 14-20 days. However, in our case, the infected individuals are split into two categories, compartment I and compartment W , where the latter represents patients with advanced stage of illness. Consequently, we split the recovery rates into three, asymptomatic recovery rate (γ_a), recovery rate for I (γ_i) and, recovery rate for W (γ_w). γ_a was set based to be 16 days, close to the midpoint region of the range provide in⁴⁰. γ_i was set to be on the lower extreme of the range - 14 days, and γ_w was set to the upper limit of the range - 20 days⁴⁵. It is important to highlight that since we vary the transmission rate over a range of continuous values, and as there are multiple transmission rates to fit the different portions of the data, the impact of the choice of recovery rates is mitigated so long as it is in the clinically valid range.

The transmission rates were allowed to vary from 0.02 to 0.85, which we determined through experiments covered all the ranges possible based on the data we had at the time of the experiment. Immunity loss rate was fixed based on an estimated 10 months period of infection³⁸.

As mentioned earlier, Nelder-Mead optimization algorithm was used to fit the variable parameters and the algorithm was run multiple times with reinitializations until a convergence in the parameter values was obtained across two runs. This was done to prevent trapping at local minima and for the reproducibility

of the parameter values; thus, ensuring that analysis of these parameters across regions are not confounded by numerical variations of the algorithm and reflect actual variation in epidemic dynamics. The average fitted nrmse at each resolution is also listed in Table 4. Figure 2a measures the distribution of each
555 of the optimized parameters across regions through a combination violin-plot, box-plot. The distribution for all other optimized parameters is shown in the supplement figure S1. L-statistics obtained from the box plot are listed in Table 4 for each parameter for comparing the distributions across spatial scales. The majority of the epidemiological parameters shown in figure 1 are variable
560 parameters with values determined by the AI optimization pipeline. The parameters were fitted using the data starting from the first day of the reported infection up to August 3rd, 2020. The evaluation MAPE was measured with respect to reference data reported for the following 4 weeks from August 3rd, 2020. Parameter values with uncertainties or range are provided in the Results.
565 All parameter values were obtained independently for all regions at three spatial resolutions (county, county cluster, and state). Resulting R_0 values are then compared to recent estimates from the literature (see Table 5). Over the course of the epidemic, changes in social distancing policies and other non-pharmaceutical interventions (npi) are manifest by a reduction in transmission
570 rate or changes in other epidemiological parameters. These modified values reflect the time varying effective reproductive number, R_{eff} (or R_t). R_{eff} changes with time due to depletion of the susceptible population as well as widespread policies (and policy removals). The maximum R_0 measured early in the epidemic reflects the "natural" infection dynamics give "normal" social interactions, absent npi.
575

Hyper-local Pipeline

To provide a scale-able architecture around the STEM model described in the sub-section Base Model, we built an automation pipeline that performs a variety of task ranging from data ingestion, smoothing to invoking the model and post-
580 processing (as required). The pipeline begins with a data set interface that allows the users to connect to multiple data set providers and convert them into a unified representation for the various downstream processes. For this paper, we used the USAFacts data set interface. The downloaded data is then pre-processed to fix any data quality
585 issues like negative daily cases and data smoothing. The pipeline provides access to different smoothing algorithms like Moving Average, Savitzky Golay and also Adaptive Savitzky Golay filters. For the purpose of this paper, we use Adaptive Savitzky Golay filter as it allows us to determine the right polynomial for smoothing with minimal data loss.
590 It is to be noted that as the pandemic evolves, there are number of responses by governments and individuals at various levels. In the case of COVID-19, such responses have mainly been through Non-Pharmaceutical Interventions (NPIs), which mainly aim at promoting social distancing and reducing the transmission risks. Further, these NPIs have different level of enforcement and consequently
595 its effects are manifested as various discontinuities on the case incidence and

death curves. Another source of discontinuities are due changes in mobility. These discontinuities in the curve can be thought to be different stages of the pandemic and warrants the model parameters to be sensitive to it. We use these discontinuities to define "segments"⁴⁸ and use the corresponding start and end
600 times as a part of Nelder-Mead optimization of transmission rate versus time. The pipeline also supports multi-processing, where each region is run by a separate instance of a docker image; thus, allowing easy scaling of the clusters based on the requirement.

Models evaluated for comparison

605 Our selection criteria for comparing to the models participated in the CDC COVID-19 forecasting challenge⁴ was based on the availability of the models' forecast submission on August 3rd, 2020. Only the models that had submission for both deaths and cases were included in our comparison regardless of the type of the model. The other model selection criterion was providing the forecasts
610 (both deaths and cases) for up to four weeks ahead. Additionally, for the state-level comparison, only the models that provided forecast for all the 50 states (excluding the District of Columbia) were included. List of the models with forecasting submission to the CDC challenge on August 3rd, 2020, that either met or did not meet the aforementioned criteria, can be found in Table 3.

Table 3: CDC Challenge Models (submitted 8/3/20). Data types C=cases, D = Deaths. \subseteq indicates subset. Models used for comparison provided predictions for both data types and all regions.

Model Name	Methods	data types	included ?	missing data
Auquan ³	Fitted SEIR	D	No	C
CDDEP ⁵	Bayesian SEIR	C	No	D
CMU ⁶	Autoreg. time-series	D	No	C
Columbia ⁷	Metapop SEIR	C,D	No	\subseteq C
Columbia-UNC ⁵⁷	Stat. survival-conv.	D	No	C
Covid19Sim ⁸	SEIR	C,D	Yes	n/a
DDS ⁹	Bayesian hierarchical	C,D	Yes	n/a
ERDC ¹⁰	SEIR	C,D	Yes	n/a
ESG ¹¹	Fitted gaussian dist.	D	No	C
Geneva ¹²	Statistical model	C,D	No	\subseteq weeks
GT-CHHS ³⁷	Agent-based	D	No	C
GT-DeepCOVID ⁴⁷	Deep learning	D	No	C
IHME ¹³	regression model	C,D	Yes	n/a
ISU ¹⁴	Spatiotemporal model	C,D	No	\subseteq regions
JCB ²⁵	Statistical model	D	No	C
JHU ¹⁵	Metapop. SEIR	C,D	Yes	n/a
Karlen ³⁶	Discrete time	C,D	Yes	n/a
LANL ¹⁶	Statistical model	C,D	Yes	n/a
LNQ ⁵⁸	Machine learning	C	No	D \subseteq weeks
LSHTM ²¹	Ensemble (3 models)	D	No	C
MIT-CovAlliance ²²	SIR	C,D	No	\subseteq regions
MIT-ORC ⁴¹	SEIR	C,D	Yes	n/a
MOBS ⁵⁶	Metapop. SLIR	D	No	C
NotreDame-FRED ³⁰	Agent-based	D	No	C
NotreDame-Mob ³¹	SEIR	D	No	C
Oliver Wyman ³⁹	SIR	C,D	No	\subseteq weeks
PSI ⁵³	Stoch SEIRX	D	No	C
QJHong ³⁴	ML	D	No	C
RPI-UW ⁵⁰	SIR	D	No	C
STH ³⁵	Stat growth	D	No	C
UA ²³	SIR	D	No	C
UCLA ⁶¹	mod SEIR	C,D	Yes	n/a
UCM ¹⁷	SEIR	D	No	C
UM ¹⁸	Ridge regression	C,D	No	\subseteq regions
UMass-MB ¹⁹	Mechanistic Bayesian	C,D	No	\subseteq regions
USC ⁵¹	SIR	C,D	Yes	n/a
UT ⁵⁹	Bayesian	D	No	C
UVA ²⁰	ensemble (3 models)	C	No	D
YYG ³²	SEIR and ML	D	No	C
Ensemble	ensemble	C,D	No	included above

615 *Supplemental Results*

Derived Epidemiological Parameters

As discussed in the main manuscript, derived recovery rates (Table 4) are within the range of current clinical studies.^{20,21,22,23,24} These studies document pre-symptomatic infections 5-14 days from the time of exposure.^{20,21,22,23,24} After symptom onset, recovery of replication-competent virus has been documented 10 – 20 days after symptom onset corresponding to a recovery rate between 0.05 – 0.1 [days⁻¹]⁵⁴. Another study documented shedding of virus up to 8 days following the symptomatic period²⁶.

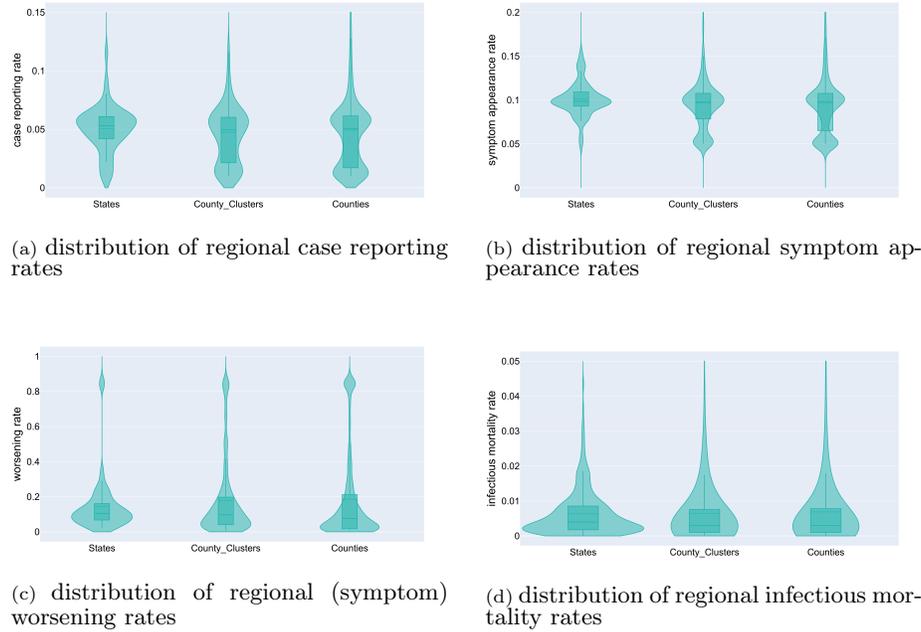


Figure S1: Distribution of regional values of parameters obtained via Nelder-Mead Simplex model optimization. The distribution of transmission rates and derived reproduction numbers are shown in manuscript figure 2a and 2b. Parameters vary by region and time as the disease dynamics change locally in response to non-pharmaceutical interventions and other local factors.

625 Table 5 provides statistics from the measured distribution of reproduction number values obtained using the epidemiological parameters in Table 4 at each spatial resolution. From the distributions shown in figure 2b the Table lists four L-statistics; first quartile, mean value, median value, and third quartile of the measured distributions for R_0 and R_{eff} as defined above. The Table also shows other estimates of the reproduction number from the literature. The examples
630 from the literature are simply ordered based on the upper limit for each reported range (and not related to the L-statistic categories). The median value for R_0 systematically decreases as the spatial resolution of the model increases from state to county. The distributions labeled R_{eff} (which reflect the diverse range of non-pharmaceutical interventions and range of compliance across all regions and time) do not show a systematic shift.
635

Table 4: Table of epidemiological parameters. NM = Nelder-Mead. GS=Grid Search.

Parameter	L-stat.	States	County Clusters	Counties	method √source
trans rate (max β)	q3 median mean q1	0.341 0.274 0.293 0.226	0.348 0.236 0.298 0.171	0.231 0.111 0.191 0.086	NM
trans rate (all β)	q3 median mean q1	0.177 0.120 0.145 0.082	0.192 0.120 0.155 0.065	0.173 0.111 0.145 0.072	NM
case reporting rate ξ	q3 median mean q1	0.061 0.053 0.050 0.042	0.060 0.050 0.048 0.021	0.061 0.050 0.049 0.017	NM
symptom appear. rate α	q3 median mean q1	0.109 0.098 0.101 0.093	0.108 0.098 0.097 0.079	0.108 0.098 0.097 0.065	NM
worsening rate ω	q3 median mean q1	0.161 0.104 0.145 0.067	0.197 0.097 0.179 0.041	0.212 0.075 0.185 0.017	NM
infectious mortality rate $\ddagger \mu_d$	q3 median mean q1	0.009 0.004 0.006 0.002	0.008 0.003 0.006 0.001	0.008 0.003 0.007 0.001	NM
recov γ_A	<i>fixed</i>	0.06	0.06	0.06	<i>lit.</i> ⁴⁵
recov γ_I	<i>fixed</i>	0.071	0.071	0.071	GS
recov γ_W	<i>fixed</i>	0.05	0.05	0.05	0.048-0.071
imm. loss ρ	<i>fixed</i>	0.1 <i>mos</i> ⁻¹	0.1 <i>mos</i> ⁻¹	0.1 <i>mos</i> ⁻¹	(8-12) ⁻¹ <i>mos</i> ⁻¹ <i>lit</i> ^{38,46}
<nrmse>	mean	0.038 ±0.008	0.065 ±0.040	0.093 ±0.052	NM

[†] rate units are per day unless otherwise indicated.

[‡] disease death rate (per day) not to be confused with case fatality rate (cfr)

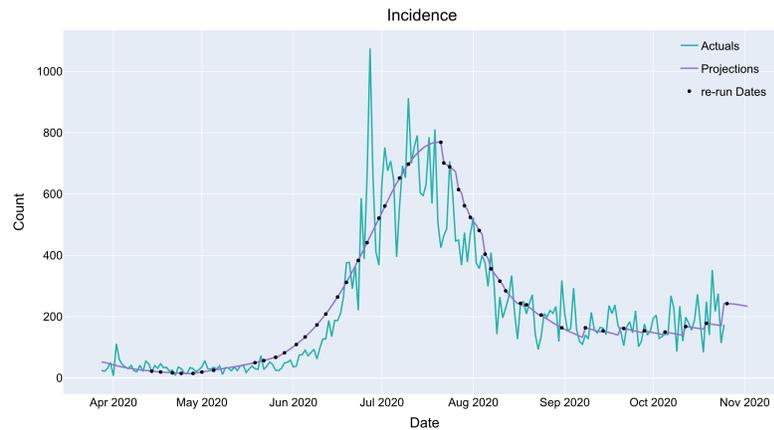
Table 5: This table shows the derived values of the maximal basic reproductive number of Sars-CoV-2, by spatial resolution, computed from the epidemiological parameter values in table 4. The final column shows estimates from current literature.

Param	L-stat.	States	County Clusters	Counties	<i>Lit</i> †
R_0	q_3	7.216	7.929	5.256	<i>up to 9.24</i> ⁶⁰
from	median	6.388	5.530	2.909	<i>1.3 - 6.5</i> ⁴⁹
max	mean	6.528	6.770	4.442	<i>2.4 - 3.1</i> ²⁸
β	q_1	4.966	3.688	1.790	<i>1.6 - 3.0</i> ⁴⁴
					<i>2.03 - 2.77</i> ⁴²
R_{eff}	q_3	3.942	4.260	3.975	
from	median	2.610	2.642	2.550	
all	mean	3.222	3.511	3.371	
β	q_1	1.833	1.436	1.602	

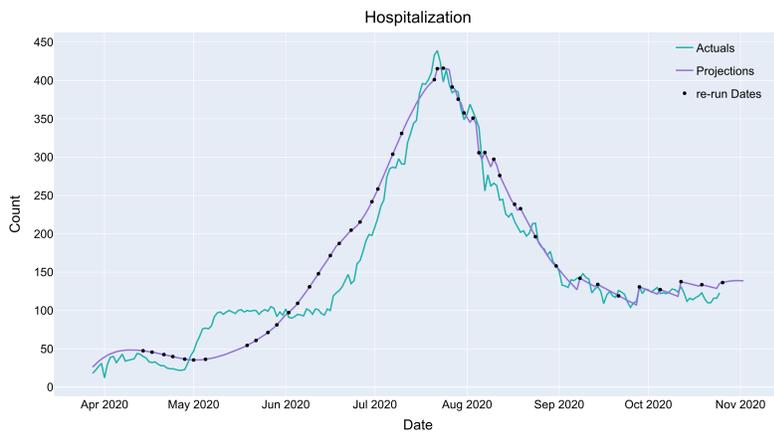
† Examples from literature are ordered by reported upper limit

Hospitalization and Intensive Care Unit Forecasting Results

Manuscript figure 6 shows predicted ICU demand for Tampa General Hospital (TGH), Hillsborough County, FL. Figure S2a shows the regular incidence predictions for the county (from the SACIR model), and figure S2b shows the predicted COVID-19 hospitalizations. The green curve is actual data and the magenta curve (and solid points) show independent projections made at least weakly beginning in April 2020.



(a) Predicted Incidence



(b) Predicted Hospital Admissions

Figure S2: Projected incidence and projected hospitalizations for Tampa General Hospital in Hillsborough County, Florida, USA.

Additional references

- 645 [1] Florida department of health in hillsborough county covid-19 dashboard. <https://www.hillsboroughcounty.org/en/residents/public-safety/emergency-management/stay-safe/covid-19-dashboard>.
- [2] 2011-2015 5-year acs commuting flows. <https://www.census.gov/data/tables/2015/demo/metro-micro/commuting-flows-2015.html>, 2020.
650 Accessed: 2020-10-04.
- [3] Auquan. <https://covid19-infection-model.auquan.com/>, 2020.
- [4] Cdc covid-19 forecasting. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/mathematical-modeling.html/>, 2020.
- [5] Cddep. <https://cddep.org/>, 2020.
- 655 [6] Cmu. <https://delphi.cmu.edu/>, 2020.
- [7] Columbia. <https://columbia.maps.arcgis.com/apps/webappviewer/index.html?id=ade6ba85450c4325a12a5b9c09ba796c>, 2020.
- [8] Covid19sim. <https://covid19sim.org/>, 2020.
- [9] Dds. <https://dds-covid19.github.io/index.html>, 2020.
- 660 [10] Erdc. <https://github.com/erdc-cv19/covid19-forecast-hub>, 2020.
- [11] Esg. <http://rwalraven.com/COVID19/>, 2020.
- [12] Geneva. <https://renkulab.shinyapps.io/COVID-19-Epidemic-Forecasting/>, 2020.
- 665 [13] Ihme. <https://covid19.healthdata.org/united-states-of-america>, 2020.
- [14] Isu. <https://covid19.stat.iastate.edu/>, 2020.
- [15] Jhu. <https://github.com/HopkinsIDD/COVIDScenarioPipeline/>, 2020.
- [16] Lanl. <https://covid-19.bsvgateway.org/>, 2020.
- [17] Ucm. <http://mechatronics.ucmerced.edu/covid19>, 2020.
- 670 [18] Um. <https://gitlab.com/sabcorse/covid-19-collaboration>, 2020.
- [19] Umass-mb. <https://github.com/dsheldon/covid>, 2020.
- [20] Uva. <https://biocomplexity.virginia.edu/>, 2020.

- 675 [21] Sam Abbott, Joel Hellewell, Robin N Thompson, Katharine Sherratt, Hamish P Gibbs, Nikos I Bosse, James D Munday, Sophie Meakin, Emma L Doughty, June Young Chun, et al. Estimating the time-varying reproduction number of sars-cov-2 using national and subnational case counts. Wellcome Open Research, 5(112):112, 2020.
- 680 [22] Jackie Baek, Vivek F Farias, Andreea Georgescu, Retsef Levi, Tianyi Peng, Deeksha Sinha, Joshua Wilde, and Andrew Zheng. The limits to learning an sir process: Granular forecasting for covid-19. arXiv preprint arXiv:2006.06373, 2020.
- [23] H. Biegle and J. Lega. Ua. <https://jocelinelega.github.io/EpiGro/>, 2020.
- 685 [24] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment, 2008(10):P10008, 2008.
- [25] J. Burant. Jcb. <https://github.com/JohnBurant/COVID19-PRM/>, 2020.
- 690 [26] De Chang, Guoxin Mo, Xin Yuan, Yi Tao, Xiaohua Peng, Fu-Sheng Wang, Lixin Xie, Lokesh Sharma, Charles S Dela Cruz, and Enqiang Qin. Time kinetics of viral clearance and resolution of symptoms in novel coronavirus infection. American journal of respiratory and critical care medicine, 201(9):1150–1152, 2020.
- 695 [27] NVV Chau. Thanh dung n, yen lm, minh nnq, hung lm, et al. The natural history and transmission potential of asymptomatic SARS-CoV-2 infection. Clin Infect Dis, 10, 2020.
- [28] Marco D’Arienzo and Angela Coniglio. Assessment of the sars-cov-2 basic reproduction number, r_0 , based on the early phase of covid-19 outbreak in italy. Biosafety and Health, 2020.
- 700 [29] Stefan Edlund, James Kaufman, Justin Lessler, Judith Douglas, Michal Bromberg, Zalman Kaufman, Ravit Bassal, Gabriel Chodick, Rachel Marom, Varda Shalev, et al. Comparing three basic models for seasonal influenza. Epidemics, 3(3-4):135–142, 2011.
- [30] G. et al. Espana. Notredame-fred. https://github.com/confunguido/covid19_ND_forecasting, 2020.
- 705 [31] G. et al. Espana. Notredame-mobility. https://github.com/TALexPerkins/covid19_NDmobility_forecasting, 2020.
- [32] Y. Gu. Yyg. <https://covid19-projections.com/about/>, 2020.
- 710 [33] Ernst Hairer, SP Norsett, and Gerhard Wanner. Solving ordinary, differential equations i, nonstiff problems/e. hairer, sp norsett, g. wanner, with 135 figures, vol.: 1. Technical report, 2Ed. Springer-Verlag, 2000, 2000.

- [34] Q. Hong. Qjhong. <https://github.com/qjhong/covid19>, 2020.
- [35] S. Horstman. Sth. <https://public.tableau.com/profile/covid19model/>, 2020.
- [36] D Karlen. Karlen. <https://pypm.github.io/home/>, 2020.
- 715 [37] Pinar Keskinocak, Buse Eylul Oruc Aglar, Arden Baxter, John Asplund, and Nicoleta Serban. The impact of social distancing on covid19 spread: State of georgia case study. medRxiv, 2020.
- [38] Stephen M Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H Grad, and Marc Lipsitch. Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. Science, 368(6493):860–868, 2020.
- 720 [39] U. et al. Koyluoglu. Oliver wyman. <https://pandemicnavigator.oliverwyman.com/>, 2020.
- [40] Seungjae Lee, Tark Kim, Eunjung Lee, Cheolgu Lee, Hojung Kim, Heejeong Rhee, Se Yoon Park, Hyo-Ju Son, Shinae Yu, Jung Wan Park, et al. Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with sars-cov-2 infection in a community treatment center in the republic of korea. JAMA internal medicine, 2020.
- 725 [41] M.L. et al. LI. Overview of delphi model v3 - covidanalytics. https://www.covidanalytics.io/DELPHI_documentation_pdf, 2020.
- 730 [42] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov-2). Science, 368(6490):489–493, 2020.
- [43] Hiroshi Nishiura, Tetsuro Kobayashi, Takeshi Miyama, Ayako Suzuki, Sung-mok Jung, Katsuma Hayashi, Ryo Kinoshita, Yichi Yang, Baoyin Yuan, Andrei R Akhmetzhanov, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (covid-19). International journal of infectious diseases, 94:154, 2020.
- 735 [44] Kiesha Prem, Yang Liu, Timothy W Russell, Adam J Kucharski, Rosalind M Eggo, Nicholas Davies, Stefan Flasche, Samuel Clifford, Carl AB Pearson, James D Munday, et al. The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: a modelling study. The Lancet Public Health, 2020.
- 740 [45] Eleanor M. Rees, Emily S. Nightingale, Yalda Jafari, Naomi R. Waterlow, Samuel Clifford, Carl A. Carl, Cmmid Working Group, Thibaut Jombart, Simon R. Procter, and Gwenan M. Knight. COVID-19 length of hospital stay: A systematic review and data synthesis. BMC Medicine, 18(1), 2020.
- 745

- 750 [46] Tyler J Ripperger, Jennifer L Uhrlaub, Makiko Watanabe, Rachel Wong, Yvonne Castaneda, Hannah A Pizzato, Mallory R Thompson, Christine Bradshaw, Craig C Weinkauff, Christian Bime, et al. Orthogonal sars-cov-2 serological assays enable surveillance of low prevalence communities and reveal durable humoral immunity. *Immunity*, 2020.
- 755 [47] Alexander Rodriguez, Anika Tabassum, Jiaming Cui, Jiajia Xie, Javen Ho, Pulak Agarwal, Bijaya Adhikari, and B Aditya Prakash. Deepcovid: An operational deep learning-driven framework for explainable real-time covid-19 forecasting. *medRxiv*, 2020.
- [48] Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*, pages 166–171. IEEE, 2011.
- 760 [49] Brendon Sen-Crowe, Mark McKenney, and Adel Elkbuli. Social distancing during the covid-19 pandemic: Staying home save lives. *The American journal of emergency medicine*, 2020.
- [50] Yunfeng Shi and Xuegang Ban. Capping mobility to control covid-19: A collision-based infectious disease transmission model. *medRxiv*, 2020.
- 765 [51] Ajitesh Srivastava, Tianjian Xu, and Viktor K Prasanna. Fast and accurate forecasting of covid-19 deaths using the sikj α model. *arXiv preprint arXiv:2007.05180*, 2020.
- [52] Desmond Sutton, Karin Fuchs, Mary D’alton, and Dena Goffman. Universal screening for sars-cov-2 in women admitted for delivery. *New England Journal of Medicine*, 382(22):2163–2164, 2020.
- 770 [53] J. et al. Turtle. Psi. <https://github.com/reichlab/covid19-forecast-hub/tree/master/data-processed/PSI-DRAFT>, 2020.
- 775 [54] JJA Van Kampen, DAMC van de Vijver, PLA Fraaij, BL Haagmans, MM Lamers, N Okba, JPC van den Akker, H Endeman, Cornelissen JJ Gommers DAMPJ, RAS Hoek, et al. Shedding of infectious virus in hospitalized patients with coronavirus disease-2019 (covid-19): duration and key determinants. *medrxiv*. Preprint posted online, 9, 2020.
- 780 [55] Sten H Vermund and Virginia E Pitzer. Asymptomatic transmission and the infection fatality risk for covid-19: Implications for school reopening. *Clinical Infectious Diseases*, 2020.
- [56] A. et al. Vespignani. Modeling of covid-19 epidemic in the united states. https://uploads-ssl.webflow.com/58e6558acc00ee8e4536c1f5/5e8bab44f5baae4c1c2a75d2_GLEAM_web.pdf, 2020.
- 785

- [57] Qinxia Wang, Shanghong Xie, Yuanjia Wang, and Donglin Zeng. Survival-convolution models for predicting covid-19 cases and assessing effects of mitigation strategies. medRxiv, 2020.
- [58] R. Wolfinger and D. Lander. Lnq. <https://www.kaggle.com/sasrdw/locknquay>, 2020. 790
- [59] S. et al. Woody. Ut. <https://covid-19.tacc.utexas.edu/projections/>, 2020.
- [60] Hyun Mo Yang, Luis Pedro Lombardi, Fabio Fernandes Morato Castro, and Ariana Campos Yang. Mathematical modeling of the transmission of sars-cov-2. evaluating the impact of isolation in *são paulo* state (brazil) and lockdown in spain associated with protective measures on the epidemic of covid-19. medRxiv, 2020. 795
- [61] Difan Zou, Lingxiao Wang, Pan Xu, Jinghui Chen, Weitong Zhang, and Quanquan Gu. Epidemic model guided machine learning for covid-19 forecasts in the united states. medRxiv, 2020. 800