

Forecasting subnational COVID-19 mortality using a day-of-the-week adjusted Bayesian hierarchical model

Justin J. Slater^{*1,2} | Patrick E. Brown^{1,2} | Jeffrey S. Rosenthal²

¹Centre for Global Health Research, St. Michael's Hospital

²Department of Statistical Sciences, University of Toronto

Correspondence

Email: justin.slater@mail.utoronto.ca

Abstract

The death toll from the COVID-19 pandemic has risen over 900,000 worldwide. Reliable estimates of mortality due to COVID-19 are important to guide intervention strategies such as lockdowns and social distancing measures. In this paper, we develop a data-driven model that accurately and consistently estimates COVID-19 mortality at the regional level early in the epidemic, using only daily mortality counts as the input. We use a Bayesian hierarchical skew-normal model with day-of-the-week parameters to provide accurate projections of COVID-19 mortality. We estimate day-of-the-week effects and cumulative COVID-19 Mortality for Canada, U.S, Spain, and Brazil. We validate our projections by comparing our model to the projections made by the Intitute for Health Metrics and Evaluation, and highlight the importance of hierarchicalization and day-of-the-week effect estimation.

KEYWORDS:

COVID-19, Forecasting, Bayesian statistics, Time Series, Hierarchical Model

1 | INTRODUCTION

The death toll from the COVID-19 pandemic has to risen over 900,000 deaths worldwide, with deaths in many regions now rising again following decreases in late spring and early summer. Although this number is likely an under estimate of the true number of deaths, it is more reliable than the reported number of cases, which is largely a function of the number of patients tested. Reliable estimates of mortality due to COVID-19 are useful for guiding intervention strategies such as lockdowns and social distancing measures. These estimates should be at the regional (e.g provincial or state) level, as the spread of the disease can vary greatly within a particular country.

There have been many attempts at forecasting COVID-19 cases and mortality. Extensions of Susceptible, Infectious, or Recovered (SIR) models have been considered (Anastassopoulou, Russo, Tsakris, and Siettos (2020); Sarkar, Khajanchi, and Nieto

(2020)). Various time series models have also been considered (Chakraborty and Ghosh (2020); Perc, Gorišek Miksić, Slavinec, and Stožer (2020); Petropoulos and Makridakis (2020)). Perhaps most notably, the Institute for Health Metrics and Evaluation (IHME) has made their predictions available since March 25th 2020 (Friedman, Liu, and Gakidou (2020)), and have been cited as the gold standard regional level projections. However, in all of these forecasting methods, there has been little attempt at accounting for differences in deaths by day-of-the-week, which if left unaccounted for, can drastically bias long-term forecasts depending on which day of the week the observed data ends. Additionally, making projections for regions can be difficult where there are a relatively small number of deaths. The methodology needs to be able to handle low daily mortality counts, and still provide reasonable cumulative mortality estimates for that region.

Figure 1 shows the COVID-19 daily death counts in the United States from March 2nd to June 25th. There are several key features of these daily deaths that seem to be prevalent in every country or region's Coronavirus mortality counts. Firstly, note the rapid rise in daily deaths relative to the decline. Capturing this skewness in the daily death counts is essential for accurately forecasting COVID-19 mortality, and is not captured using the Normal density initially used by the IHME. Additionally, notice the weekly periodicity in daily death counts. It appears that certain days of the week tend to have higher death counts than others, which is an important feature to capture, so that our cumulative death forecasts don't depend on what day of the week they are made. This weekly periodicity has been confirmed using spectral analysis (Bukhari, Jameel, Massaro, D'Agostino, and Khan (2020)), but the relative risks of mortality between certain days-of-the-week are still indeterminate. Studying day-of-the-week effects has proven useful in other fields such as actuarial science (Crevecoeur, Antonio, and Verbelen (2019)) and economics (Berument and Kiyamaz (2001)).

The goal of this paper is to develop a data-driven model that accurately and consistently estimates COVID-19 mortality at the regional level early in the epidemic, using daily mortality counts as the input. We do so by developing a hierarchical Bayesian model where the daily death counts are assumed to follow a skew-normal distribution, and vary by day-of-the-week. In doing so, we estimate the number of daily deaths at the peak of the epidemic, the date of the peak of the epidemic, and other interesting features. The hierarchical nature of our model will allow for accurate estimation of cumulative death counts in regions where the epidemic is still in the early stages by borrowing information from regions where the epidemic has matured. We compare the forecasting performance of our model to the projections made by the Institute for Health Metrics and Evaluation (IHME), as well as the observed mortality values and highlight the importance of hierarchicalization and accounting for day-of-the-week effects when projecting mortality counts during an epidemic.

This work is an extension of the model presented in Brown, Jha, et al. (2020), which has shown promise as a national-level forecasting model. The two main contributions of our work are to allow for subnational level data (hierarchicalization), and estimation/modelling of day-of-the-week effects.

2 | METHODS

2.1 | The Skew-Normal Model

We saw in figure 1 that a key feature of daily coronavirus mortality counts is the rapid rise relative to the fall in daily death counts. A natural choice to model these counts would be to assume that the mean deaths per day in a region follows a skew normal density. The skew-normal provides a good base for our model, but is far too simple to capture the full range of shapes of epidemics by itself. Firstly, although a majority of the deaths in a region occur during the main epidemic, a small number of deaths can occur outside of this epidemic. Additionally, in order to be able to compare various regions' mortality counts, we need to "standardize" our estimates based on how many deaths we would expect to see in that region from all causes. Lastly, we need to add a multiplicative term to our skew-normal that accounts for differences in daily mortality counts by day-of-the-week. By including all of these considerations, we arrive at the full model.

The statistical model (referred to as the "DOW model") used to estimate daily mortality, Y , in region i of country j is given by:

$$\begin{aligned}
 Y_{ij} &\sim \text{NegBinom}[\lambda_{ij}(t), \tau_j] \\
 \lambda_{ij}(t) &= R_{j,m[t]} E_{ij} [C_{ij} f(t; A_{ij}, B_{ij}, K_{ij}) + D_{ij}] \\
 C_{ij} &\sim \text{Gamma}(\alpha_j / \theta_C, \theta_C) \\
 B_{ij} &\sim \text{Gamma}(\eta_j / \theta_B, \theta_B) \\
 K_{ij} &\sim \text{Gamma}(\zeta_j / \theta_K, \theta_K)
 \end{aligned} \tag{1}$$

The priors for our model can be found in table 1. A_{ij} represents the "location" parameter, indicating the date at which the daily deaths reaches its peak and is analogous to the mean of a normal distribution. B_{ij} represents the duration of the epidemic, and is analogous to the standard deviation in the Normal scenario. K_{ij} represents the "skewness" parameter, which describes the ratio of the initial incline relative to the decline, and is a key parameter for capturing the shape of daily mortality counts. E_{ij} represents the age standardized death counts in each region of the included countries. This was computed by obtaining age distribution information from census data of each country and comparing it to the deaths-by-age breakdown in Italy on March 29 2020. Note that E_{ij} is assumed constant, as it is calculated a priori for each region. Calculation of this E_{ij} allows for comparable heights of peaks between regions, which is captured in the parameter C_{ij} . A high C_{ij} means that there were a large number of coronavirus related deaths, relative to the expected number of deaths in that region. The parameter D_{ij} , known as the "spark" term, captures the few deaths that were outside of the main epidemic. $R_{j,m[t]}$ captures the day-of-the-week effect for country j on day-of-the-week $m[t]$, relative to Sunday. For example, if coronavirus related deaths in Spain occurred on Tuesday twice as often as Sunday, then $R_{\text{Spain}, \text{Tuesday}} = 2$. The overdispersion parameter, τ_j , allows the variance of the daily mean deaths to vary by country by a multiplicative factor.

The number of parameters in this model can grow very quickly depending on the number of countries/regions included in analysis, and can be difficult to implement without carefully chosen priors. One of the advantages to Bayesian analysis is the incorporation of prior knowledge to guide parameter estimation. That is, we can use information about duration and severity from epidemics that are already over (e.g Spain) to set priors for regions where the epidemic has yet to peak (e.g Brazil).

Note that C_{ij} , B_{ij} and K_{ij} are all modelled hierarchically. The advantage of this is that for regions that are in the early stages of the epidemic can "borrow" information from other regions in the same country, to estimate the severity, duration, and skewness of the epidemic. This is because when we model, for example C_{ij} , hierarchically, C_{ij} will be a weighted average between the country's mean and the region's mean, but will tend more toward the country's mean when the number of events is small (Gelman and Hill (2006)). Modelling A_{ij} hierarchically was considered, however it was deemed inadvisable because for regions with small death counts, the location parameter would tend toward the country average. This is problematic because the reason that the death counts are low in that region is likely because the epidemic has yet to run its course. For this reason, we decided to estimate the location parameter separately for each region. Modelling the day-of-the-week effects, $R_{j.m[t]}$, hierarchically was also considered because it would provide a day-of-the-week effect estimate for each region. However, this was deemed computationally too cumbersome, as this would add over 300 parameters to our model.

2.2 | Daily mortality projections for four countries

We applied our model to daily COVID-19 death counts from 95 regions from four countries: U.S states; Canadian provinces; Spanish Autonomous Communities; and Brazilian states. Data are from www.coronavirus.app, where any region with 50 or more deaths as of June 25th was included in the analysis. Parameters for our models were estimated using No-U-turn sampling (Hoffman and Gelman (2014)) within the Stan software (Carpenter et al. (2017)). Four chains were used with 3000 iterations of warm-up and 1000 iterations of sampling, which were then thinned by a factor of 10 (leaving 400 posterior samples for each parameter). Convergence of Markov Chains was assessed using trace plots alongside the Gelman-Rubin Statistic ($\text{Rhat} < 1.05$) (Gelman, Rubin, et al. (1992)).

Forecasts were created from the posterior samples of $\lambda_{ij}(t)$ up until October 1st 2020. Given that λ_{ij} has a day-of-the-week effect, the posterior samples of $\lambda_{ij}(t)$ will be oscillatory. Forecasts were also made at the country level by computing

$$\bar{\lambda}_{ij}(t) = \sum_j \lambda_{ij}(t)$$

for each posterior sample.

2.3 | Day-of-the-week effect estimates

From this model, we obtain 24 day of the week effect parameters: one for each day of the week (except for Sunday which is fixed at 1) for each of the four countries. In order to gain some insight as to whether the estimated day-of-the-week effects are simply due to differences in reporting, we pulled proportions of deaths by day-of-the-week in Canada from various non-COVID related causes: circulatory, pulmonary, circulatory and pulmonary, and non-circulatory pulmonary. We will compare the estimated mortality rates from these causes to the estimated day-of-the-week effects to see if there is a similar trend. If not, then this suggests that coronavirus may be more likely to cause death on certain days of the week, or simply that coronavirus mortality has its own unique reporting trends.

2.4 | Validating Forecasts

The IHME has made projections for the United States available since March 25th 2020, and has since expanded the number of regions they include in their model. In order to validate our model projections, we ran our model using the same daily mortality data as the IHME, and compared both model's cumulative death counts to the observed death count on June 30th 2020 for 14 different time points ranging from April 1st to June 25th. Additionally, we ran our model without a day-of-the-week effect to see whether the forecasting performance of our model relies on the day-of-the-week, and is not only outperforming the IHME model for other reasons. This model will be referred to as the "non-DOW model".

Models were compared by assessing consistency and accuracy of projections throughout time. Consistency was assessed by examining the amount of overlap between successive intervals, with subsequent intervals hopefully being narrower and mostly contained in previous intervals. Overlap was measured at 13 time points, since the first time point does not have a previous interval. At each of these 13 time points, the proportion of the interval that is contained in the previous interval is calculated for each region individually, resulting in 13 proportions per region. These proportions are then averaged across all regions to determine which model, on average, had the most consistent predictions.

Accuracy of model forecasts was assessed by computing the proportion of time points that the prediction intervals contain the observed cumulative death counts on June 30th 2020. To avoid the issue of excessively wide intervals appearing the best, we also plot the mean log-length of the intervals for each model, at each time point. The model which contains the true value the most often, relative to the mean log-length of the intervals, was considered the favourable model by this metric.

3 | RESULTS

3.1 | Forecasts of Daily Mortality in Four Countries

Forecasts for all regions can be found at: https://github.com/cghr-toronto/public/blob/master/covid/DOW/all_daily.pdf. The forecasts for all of Brazil are shown in figure 3a, with the red points representing the data that was used to fit the model, and the purple representing the observed values from June 26th to August 31st. Up until July, our model fits the data quite well, indicated by the red points clustered around the posterior samples, and the day-of-the-week effect is well captured. However, starting in July, our model slightly underestimates the mean daily death counts. Figures 3b-c show the daily mortality forecasts in Sao Paulo, the region in Brazil with the most deaths, and Acre, a region with few deaths. In Acre, our model captures the trend of daily deaths reasonably well, suggesting that the hierarchical nature of our model is helping provide good forecasts for small regions. Starting in July, we are underestimating daily deaths in Sao Paulo. Although we estimate the mean daily deaths in Sao Paulo before July, it appears that we are slightly underestimating the day-of-the-week effect in Sao Paulo, indicated by the dispersion of the red and purple points having higher variance than the day-of-the-week effect allows for. This is due to the fact that we only allowed for one day-of-the-week effect for each country.

The forecasts for all of the United States are shown in Figure 4a. Notice that our model under predicts daily mortality counts throughout July and August. This is likely due to the fact that many states in the U.S have loosened their lockdown restrictions, which our model is unable to account for. Figures 4b-c show plots of the projections for Illinois and California. Our projections for Illinois seem to be quite accurate, as this is one of the states that is yet to experience any repercussions from reopening. However, our model is underestimating deaths in California likely because of loosening COVID-19 related restrictions.

Results for Spain and Canada are presented in figures 5a and 5b, and are less interesting due to the fact that the epidemics are largely over in these countries. When looking at the daily death plot for Spain, we see a small second wave. But the fact that this plot is on the log scale makes small second waves appear larger than they are. Our model seems to project Canada's COVID-19 mortality reasonable well, likely due to the relatively firm COVID-19 restrictions in Canada. Ultimately, our model seems to predict daily COVID-19 mortality well in regions with firm COVID-19 restrictions. In 3.3, we will validate our projections under the assumption that COVID-19 related restrictions are held constant.

3.2 | Day-of-the-week effects

The 2.5th, 50th and 97.5th percentiles of the posterior distributions of the day-of-the-week effects are presented in figure 2. With the exception of Spain, Sunday appears to report the lowest death counts, followed by Monday. In all countries, death counts seem to rise on Tuesday, and remain high until Friday or Saturday, and are still elevated relative to Sunday. The day-of-the-week effect is most pronounced in the United States, where Tuesday - Friday are all very similar, but are vastly different than the other

days of the week. Brazil also shows a strong day-of-the-week effect, indicated by Monday's credible interval having almost no overlap to any other day-of-the-week. As expected, Canada's credible intervals are the widest, due to the fewest overall death counts.

The proportion of deaths by day-of-the-week relative to Sunday for Canada are shown in Table 2. Note that there is very little, if any, similarities between our model's day of the week estimates and this data. These numbers only fluctuate by a few percentage points, and do not generally show a large spike on tuesday as we saw in our COVID-19 day-of-the-week estimates. This could indicate that people are more likely to die due to COVID-19 on particular days, but is more likely explainable by differences in reporting between data sources.

3.3 | Model Validation

Projections made at each of the 14 dates in all regions for the 3 models are shown at https://github.com/cghr-toronto/public/blob/master/covid/DOW/validation_all.png. Figures 6-8 show projections for the four regions with the most deaths in each nation. In any plot where the IHME's results are missing, it is because they were not yet available for that region at that time. In the plots where the DOW or non-DOW model were missing, it is because they had not yet achieved the minimum 50 deaths required to be included in our analysis. For regions with a large number of deaths (such as New York), the day-of-the-week model was very consistent, where 95% credible intervals have a large amount of overlap from date-to-date. The IHME projections are somewhat inconsistent for this region, indicated by non-overlapping intervals. However, in regions like Florida, the IHME model seems to outperform the DOW and non-DOW models, requiring a more formal investigation into model projection assessments. The mean proportion of overlap between successive intervals for all regions is shown in figure 9a. The DOW model tends to show the highest mean overlap at 8/13 time points, the non-DOW model had the most overlap at 3/13 time points, follow by the IHME which had the most overlap at 2/13 time points. This suggests that the DOW model produces the most consistent projections of the 3 models, and indicates that in the long-run our projections are likely to be better suited than the IHME's.

The accuracy of the models was assessed visually in 9b. This figure shows the proportion of regions that each model's interval contained the observed June 30th cumulative death count. Somewhat surprisingly, we see a downward trend as we get closer to June 30th, as shorter term predictions should become easier. The DOW and non-DOW models seem to do very well early on in the epidemic, capturing over 90% of true values on April 17th. The IHME model tends to do better in the 1-month projection range. All models tend to perform poorly for very short-term projections (i.e <2 weeks).

Although the IHME model appears to be better suited in 1-month projections based on our accuracy metric, figure 9c shows that this is likely due to the increased interval width. The IHME's intervals are approximately two times as wide when making

1-month projections. Despite the IHME's wider intervals at almost every time point, the DOW model outperforms the IHME projections in terms of consistency (i.e interval overlap), and is more accurate for longer-term projections.

4 | DISCUSSION

Based on our model results, it is not totally unreasonable to suggest that COVID-19 patients are more likely to pass away on certain days of the week. It may be the case that people are more likely to contract the disease on given days of the week (e.g weekdays), which may cause them to pass away at higher rates in the following days. It is also possible that on certain days of the week, hospitals are more crowded and have fewer available resources, which could increase mortality on those days. It could also be the case that deaths are equally likely to occur on any day-of-the-week, but are simply more likely to be reported on certain days due to hospital administration procedures. Either way, our results show that regardless of whether or not these are true day-of-the-week effects or are simply an artifact created by inconsistent reporting, accounting for day-of-the-week effects is important when projecting mortality during epidemics. Our non-DOW model also outperformed the IHME projections, which shows the benefits of using a skew-Normal and/or hierarchicalizing the model parameters.

Since June 30th 2020, many regions such as Florida have seen a second increase in COVID-19 deaths. Further extensions to our model need to be explored to account for a second increase in daily deaths. Prediction of these second increases could be performed using region-level covariates such as lock-down severity or mask usage in the region. The second peak itself could be potentially be modelled by adding a second skew-Normal, where the location and height of the peak daily deaths are related to the first skew-Normal's parameters. Additionally, some regions that appear to be undergoing a second increase but may just be having multiple epidemics in different subregions (e.g counties in the U.S). Having smaller-area level data could improve model performance because of this.

Another extension to this model could be to have a different day-of-the-week parameter per region, but this is likely only possible for regions that have at least several weeks of data. Hierarchicalizing the day of the week parameters is also an option. For example, the monday effect for regions with less mature data could be estimated by "borrowing" information from regions where the epidemic has matured. However, in the midst of an epidemic, computational efficiency is of the utmost importance, as these models can easily take over a week to run. Obtaining robust estimates quickly can help aid policy decisions which can ultimately save lives, so having projections within a day or two is largely beneficial.

Despite these limitations in the DOW model, it seems to forecast mortality related to COVID-19 in the first "wave" quite well when compared to the gold standard projections. Our projections made in Section 3.1 can be seen as accurate projections assuming that COVID-19 restrictions were never loosened in each region. The day-of-the-week effect was shown to be important when forecasting COVID-19 mortality, as the IHME and non-DOW model seemed to be inconsistent with their projections over

time, which could be because the projections were made on different days of the week. Our skew-Normal Bayesian hierarchical model with day-of-the-week effects could be used for future COVID-19 mortality predictions where the epidemic is less mature, or perhaps for future outbreaks where accurate and consistent mortality projections are needed, where only daily mortality data is available as the input.

5 | DATA AND CODE AVAILABILITY

Data, code, and results are all available at the Centre for Global Health Research's public github:

<https://github.com/cghr-toronto/public/tree/master/covid/DOW>

6 | ACKNOWLEDGEMENTS

This research was supported by Canadian Statistical Sciences Institute's rapid response program on COVID-19, and the Natural Sciences and Engineering Research Council.

References

- Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PloS one*, 15(3), e0230405.
- Berument, H., & Kiyamaz, H. (2001). The day of the week effect on stock market volatility. *Journal of economics and finance*, 25(2), 181–193.
- Brown, P., Jha, P., et al. (2020). Mortality from COVID-19 in 12 countries and 6 states of the united states. *medRxiv*.
- Bukhari, Q., Jameel, Y., Massaro, J. M., D'Agostino, R. B., & Khan, S. (2020). Periodic oscillations in daily reported infections and deaths for coronavirus disease 2019. *JAMA Network Open*, 3(8), e2017521–e2017521.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1).
- Chakraborty, T., & Ghosh, I. (2020). Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. *Chaos, Solitons & Fractals*, 109850.
- Crevecoeur, J., Antonio, K., & Verbelen, R. (2019). Modeling the number of hidden events subject to observation delay. *European Journal of Operational Research*, 277(3), 930–944.
- Friedman, J., Liu, P., & Gakidou, E. (2020). Predictive performance of international COVID-19 mortality forecasting models. *medRxiv*.

-
- Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
- Gelman, A., Rubin, D. B., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4), 457–472.
- Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1), 1593–1623.
- Perc, M., Gorišek Miksić, N., Slavinec, M., & Stožer, A. (2020). Forecasting COVID-19. *Frontiers in Physics*, 8, 127.
- Petropoulos, F., & Makridakis, S. (2020). Forecasting the novel coronavirus COVID-19. *PloS one*, 15(3), e0231236.
- Sarkar, K., Khajanchi, S., & Nieto, J. J. (2020). Modeling and forecasting the COVID-19 pandemic in india. *Chaos, Solitons & Fractals*, 110049.

Parameter	Prior	Description
$A_{i,j}$	$N(\text{Mar } 29, 100^2)$	Location Parameter
$A_{i,Brazil}$	$N(\text{Jun } 17, 45^2)$	Location Parameter
$\frac{1}{\sqrt{\tau_j}}$	$\text{Exp}(1/10)$	Overdispersion Parameter
D_{ij}	$\text{Exp}(1/10000)$	Spark Term
C_{ij}	α_j	Mean of C_{ij} , the severity parameters
	θ_C	Scale of the severity parameters
B_{ij}	η_j	Mean of B_{ij} , the duration parameter
	θ_B	Scale of the duration parameters
K_{ij}	ζ_j	Mean of the K_{ij} , the skewness parameters
	θ_K	Scale of the skewness parameters
$R_{j,m}$	$N_+(1, 2^2)$	Day-of-the-week parameters

TABLE 1 Prior distributions for the DOW model in (1)

Cause	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Circulatory	1.016	1.003	1.005	0.995	1.003	0.997
Pulmonary	0.994	0.996	0.975	0.985	0.996	0.988
Circulatory and Pulmonary	1.011	1.002	0.999	0.993	1.002	0.995
Non-Circulatory Pulmonary	0.988	0.998	1.002	1.007	0.998	1.005

TABLE 2 Relative risks of days-of-the-week (relative to Sunday) for non COVID-19 related causes in Canada

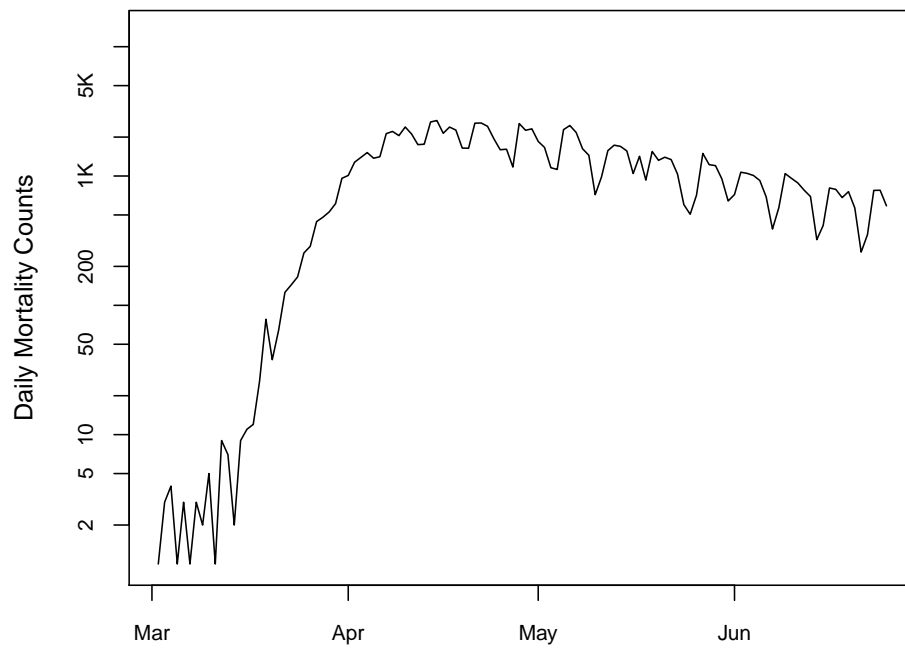


FIGURE 1 United States daily COVID-19 Mortality from www.coronavirus.app from March 1, 2020 to June 25th 2020



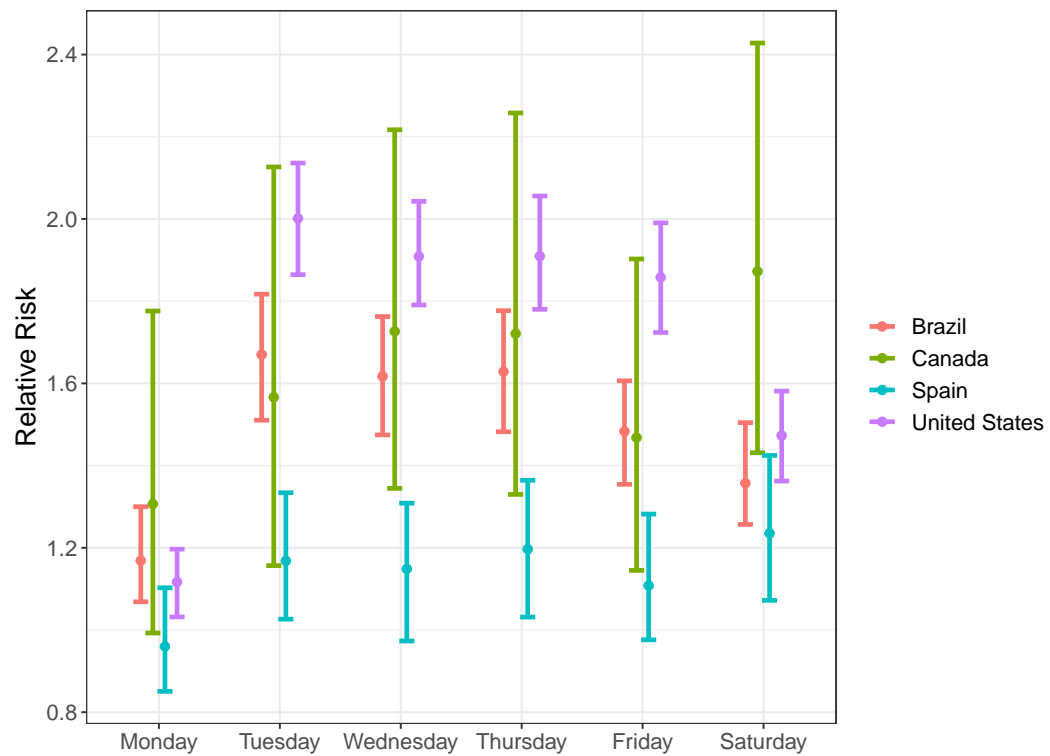
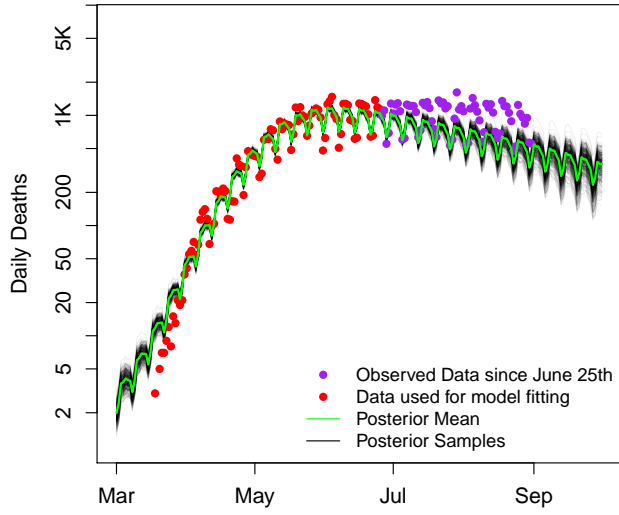
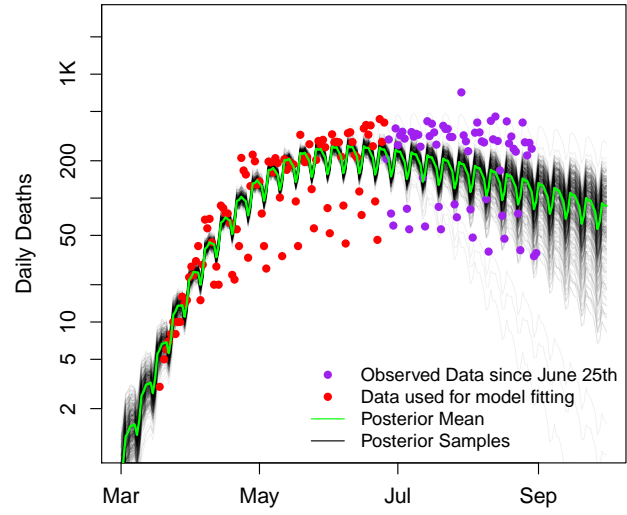


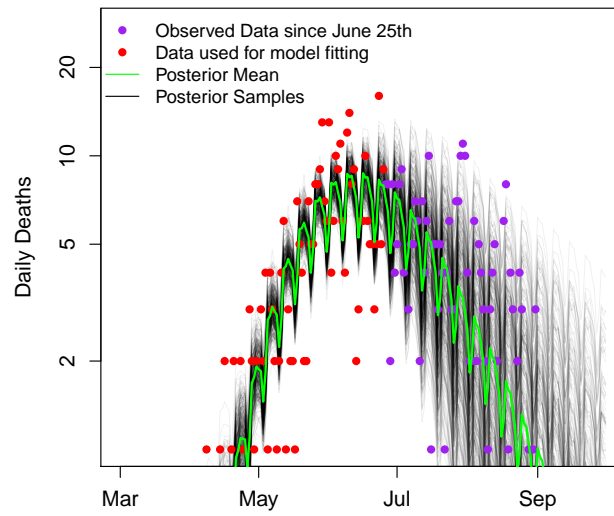
FIGURE 2 2.5th, 50th, and 97.5th percentiles of posterior distributions for day of the week parameters relative to Sunday (which is fixed at 1.)



(a) Brazil

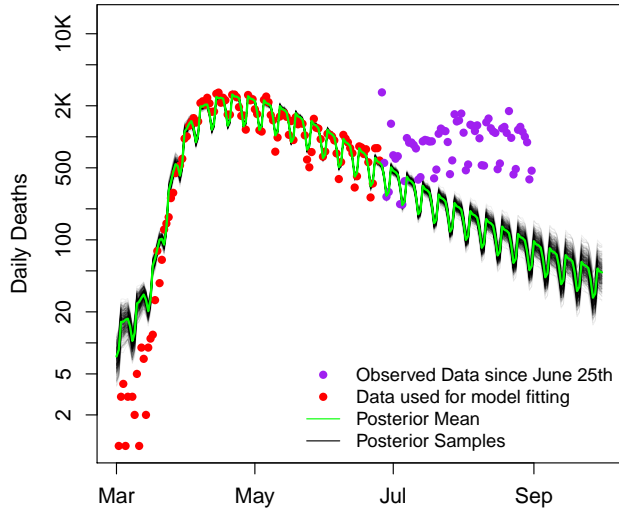


(b) Sao Paulo

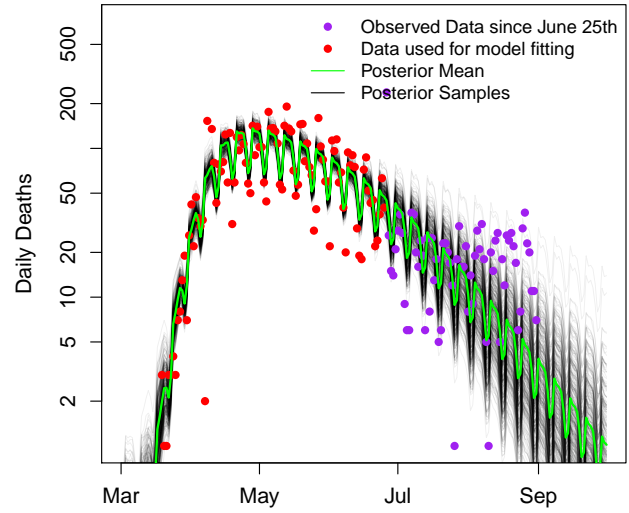


(c) Acre

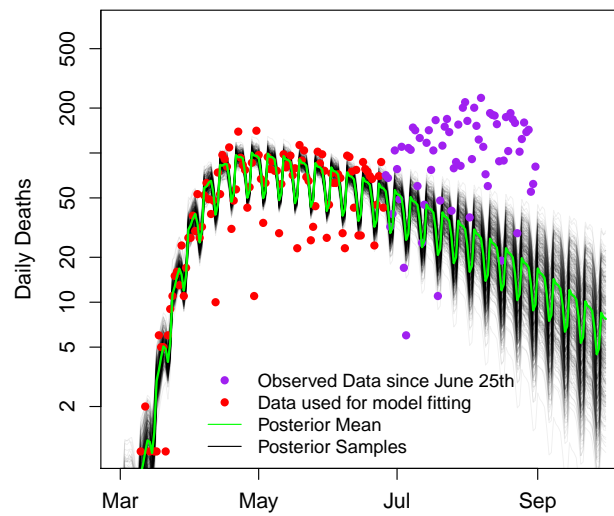
FIGURE 3 Forecasting daily and cumulative deaths in Brazil as a whole, and the states of Sao Paulo and Acre.



(a) United States

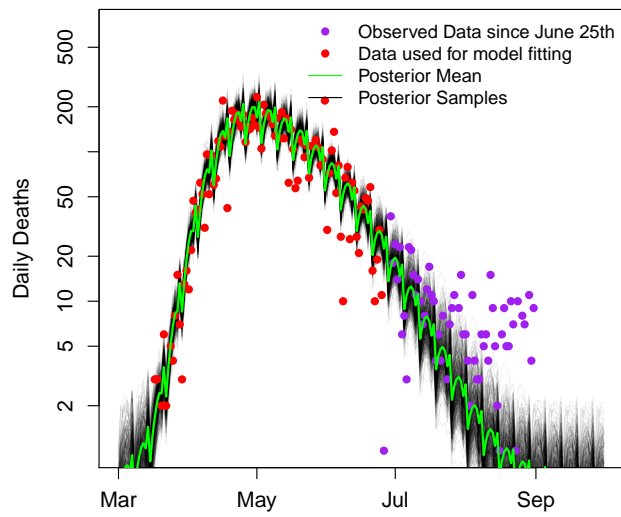


(b) Illinois

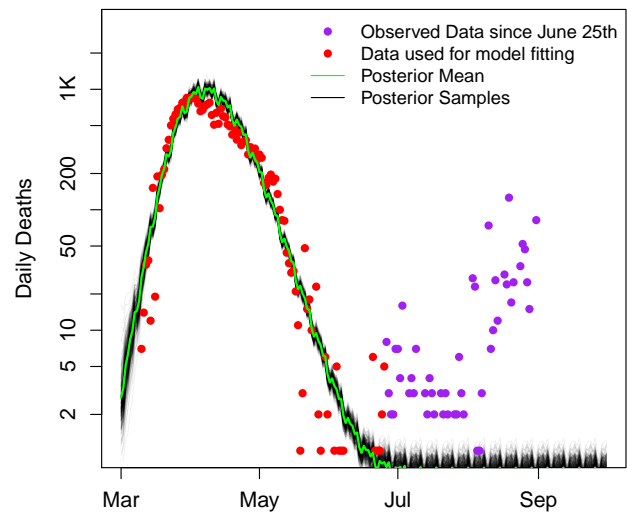


(c) California

FIGURE 4 Forecasting daily and cumulative deaths in the United States



(a) Canada



(b) Spain

FIGURE 5 Forecasting daily deaths in the Canada and Spain

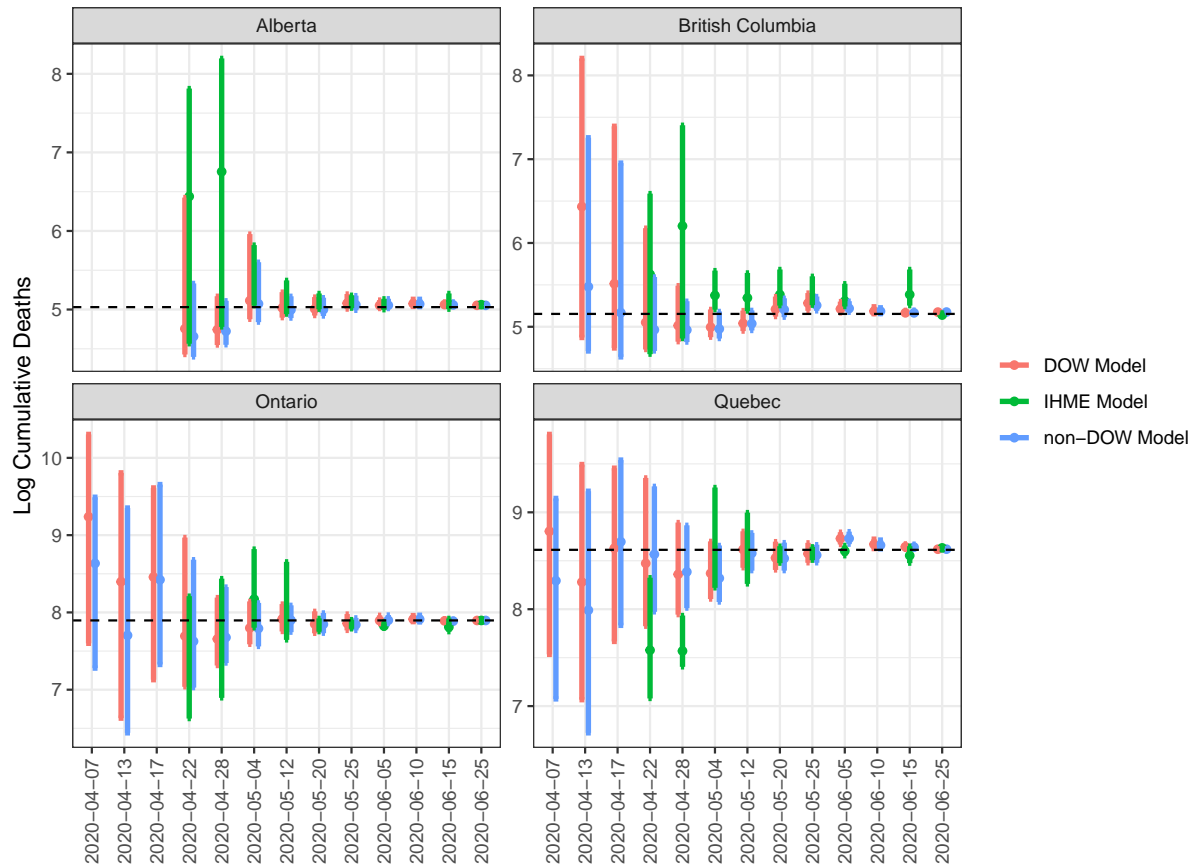


FIGURE 6 Cumulative mortality projections for June 30th made at 14 different dates starting April 1st 2020. Results are shown for the four Canadian provinces with the most COVID-19 deaths.

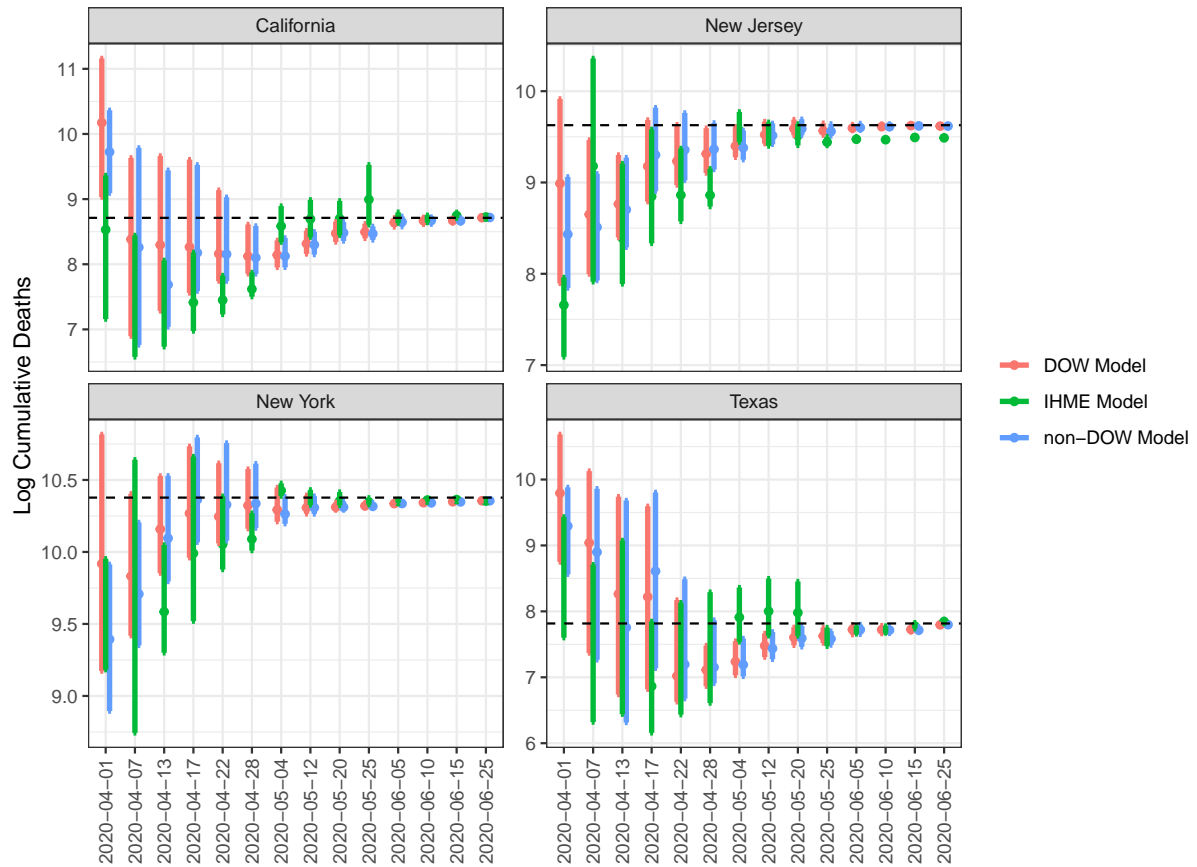


FIGURE 7 Cumulative mortality projections for June 30th made at 14 different dates starting April 1st 2020. Results are shown for the four U.S states with the most COVID-19 deaths.

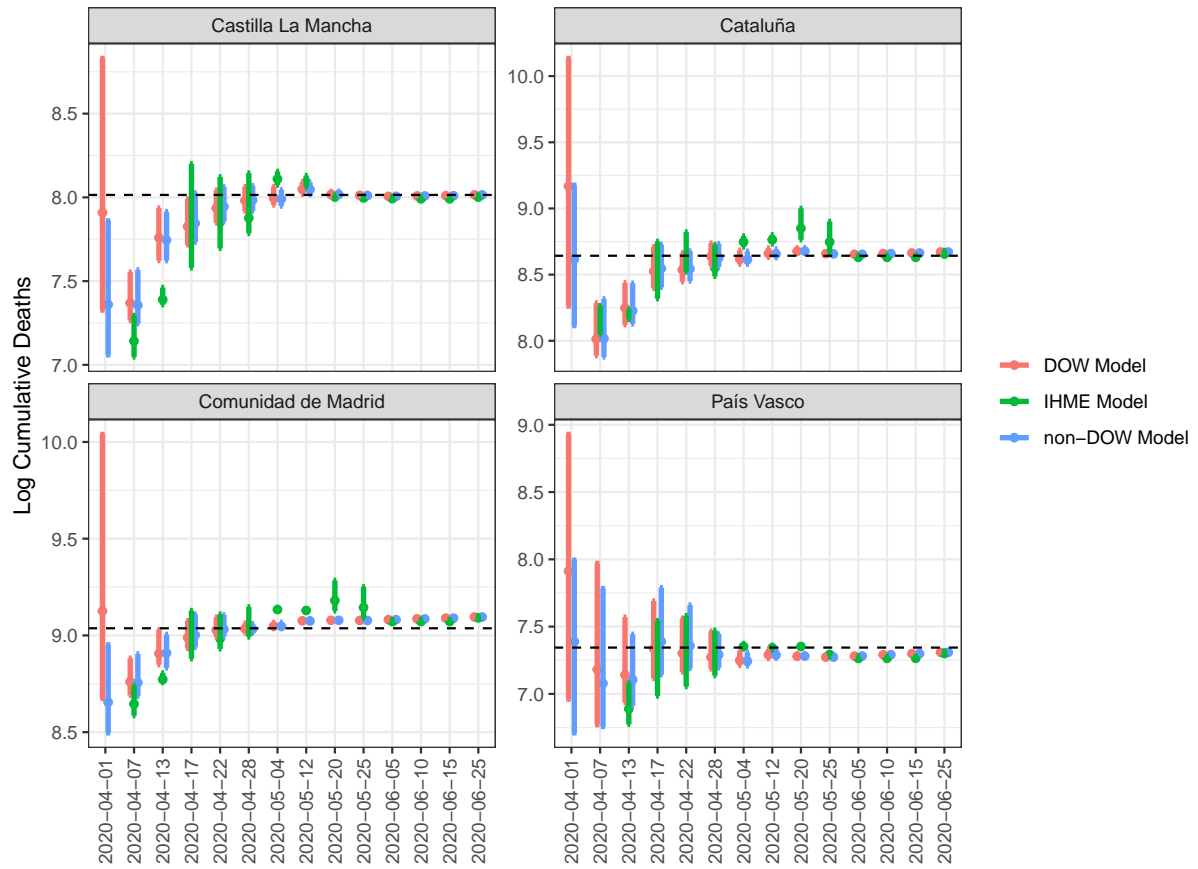
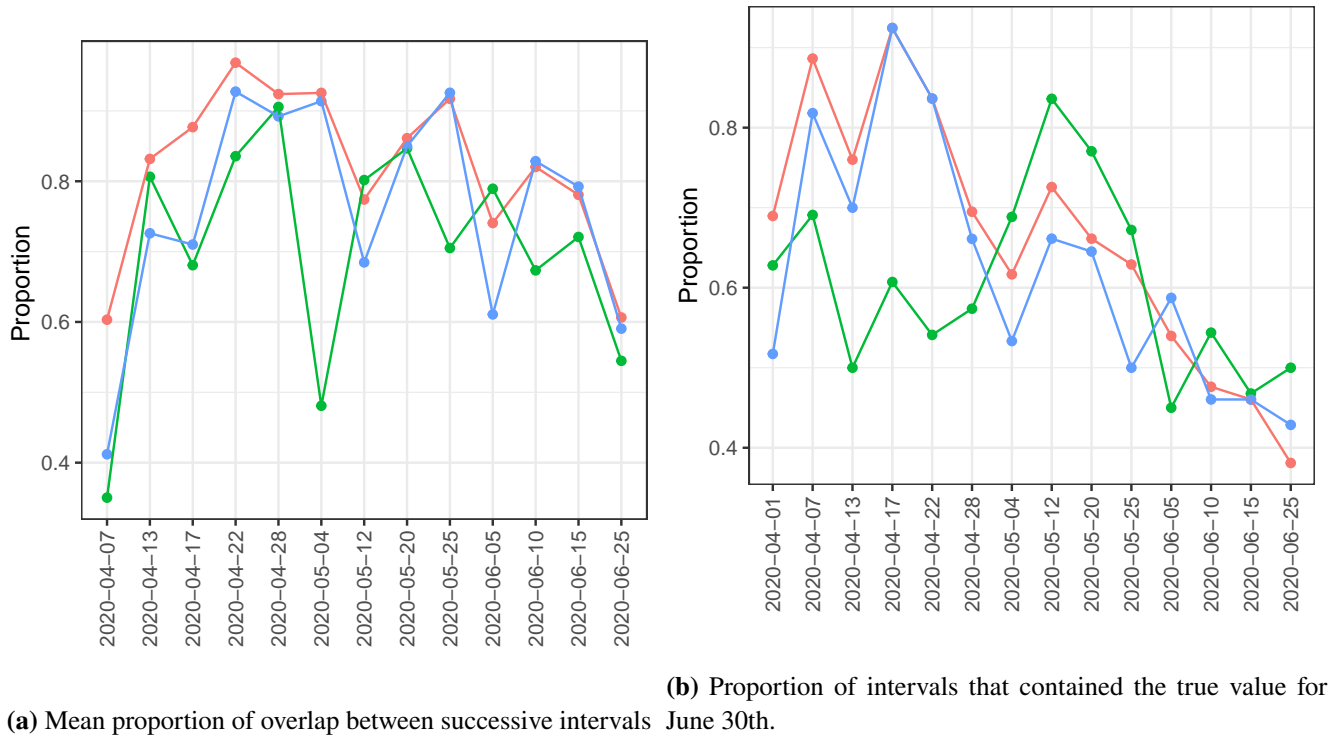
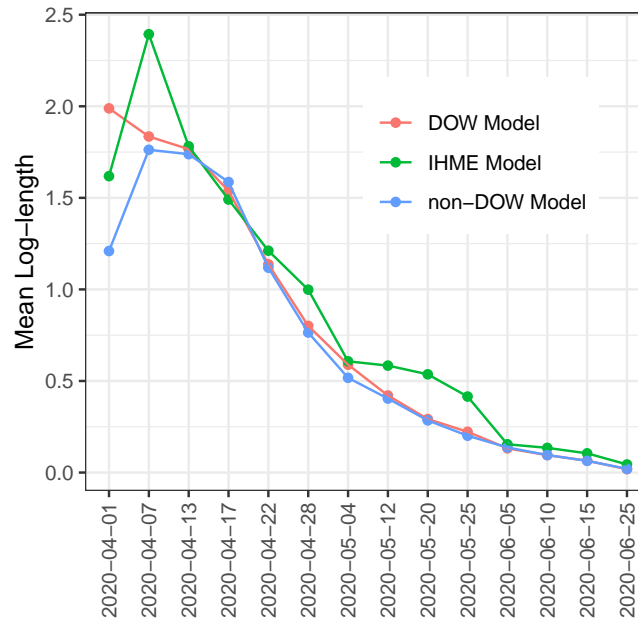


FIGURE 8 Cumulative mortality projections for June 30th made at 14 different dates starting April 1st 2020. Results are shown for the four Spanish Autonomous Communities with the most COVID-19 deaths.



(a) Mean proportion of overlap between successive intervals June 30th. (b) Proportion of intervals that contained the true value for



(c) Mean log-length of intervals

FIGURE 9 Comparing the DOW, non-DOW, and IHME models based on how much overlap there was between successive intervals, how often their intervals contained the true value, and their interval lengths