A scenario-based forecasting framework for COVID-19:
Case study application for Canada

Colin Daniel[1], Leonardo Frid, Bronwyn Rayfield & Alex Embrey
Apex Resource Management Solutions
www.apexrms.com
May 17, 2020

Introduction

Decisions are being made every day, at the level of individuals to countries, in response to the outbreak of COVID-19. Ideally these decisions are informed by the best possible forecasts. The better our forecasts, the better we can prepare for the future – whether it be mobilizing health care resources to deal with the anticipated number of future infections, or understanding, in advance, the potential consequences of changes in public health measures. While the need to forecast disease progression is nothing new, the speed and extent of the current COVID-19 pandemic has challenged our abilities, as forecasters, like never before. Early data on the disease's epidemiology is still very limited; records of cases and infections are far from complete. And in most parts of the world the dynamics of the disease are still changing daily. Yet despite these challenges, decision makers still need to push forward and make new decisions every day.

There are several different approaches to forecasting present and future levels of COVID-19 infections and deaths. Forecasts can be as simple as estimating the future number of infections and deaths based on published infection and fatality rates. More complicated mathematical models, however, can also be used. An important limitation of most existing models, however, is their prescriptive nature: forecasts are typically "published" by a team of modelers, with the modeling scenarios decided upon a priori by the modeling team. For example the results of the U.S. IHME COVID-19 model (Murray 2020) are published every few days for all U.S. states (and some countries) using pre-defined scenario assumptions; as a result there is no opportunity for policy makers in these jurisdictions to decide for themselves what scenarios to consider. Similarly, forecasts from the U.K.'s Imperial College London model (Flaxman et al. 2020) periodically publish forecasts based on specific sets of historical public health interventions for a predetermined set of jurisdictions.

In contrast we have developed a general software *framework* for providing real-time forecasts of COVID-19 infections and deaths that can be rapidly deployed for use anywhere in the world. The framework provides access to any open-source (i.e. GitHub-based) forecasting model, along with real-time daily data, in a standardized and user-friendly format. Using the latest source control techniques, the framework allows scientists to continually adapt and improve the underlying models as understanding of COVID-19 evolves over time. The framework also defines consistent data formats and definitions for shared model inputs and outputs, ensuring that forecasters are able to generate projections using various combinations of local and international data; in this way the framework allows decision makers to assess and compare alternative model projections for local accuracy and relevance, thus building confidence over time in their forecasts. Finally, the framework provides the flexibility for policy makers to introduce their own local "what-if" scenarios regarding the effects of possible future changes to public health measures.

[1] Primary author email: colin.daniel@apexrms.com

Approach

Our modelling framework is built on our existing SyncroSim software platform (www.syncrosim.com), a product we originally developed to support forecasting of animal populations and vegetation change, two areas in which, like the early stages of the COVID-19 outbreak, there is typically very high uncertainty and very little data.

As shown in Figure 1, our framework is designed to work with a range of existing scientific models: scientists can quickly add their own existing models or choose from a library of existing online open source models (stored in GitHub), and continually adapt and improve their models as understanding of COVID-19 evolves over time. Individual models can be chained together (e.g. a first model projecting infections, followed by a second model projecting hospitalizations), or run side-by-side (e.g. two alternative models both projecting infections and deaths). Because we are using a framework, rather than a specific model per se, most of the assumptions within the models can be modified locally by users as they see fit. The modelling framework is stochastic, in that it allows uncertainties to be specified for model inputs; this, in turn, allows the consequences of these uncertainties to be reflected in its forecasts. In addition, because the framework is able to run alternative models side-by-side, it is able to account for forecasting uncertainties due to projection differences between models. Finally, and we believe most importantly, the framework is scenario-based, allowing users to easily explore alternative "what-if" scenarios regarding model structure, model assumptions, and the timing and extent of future public health measures.

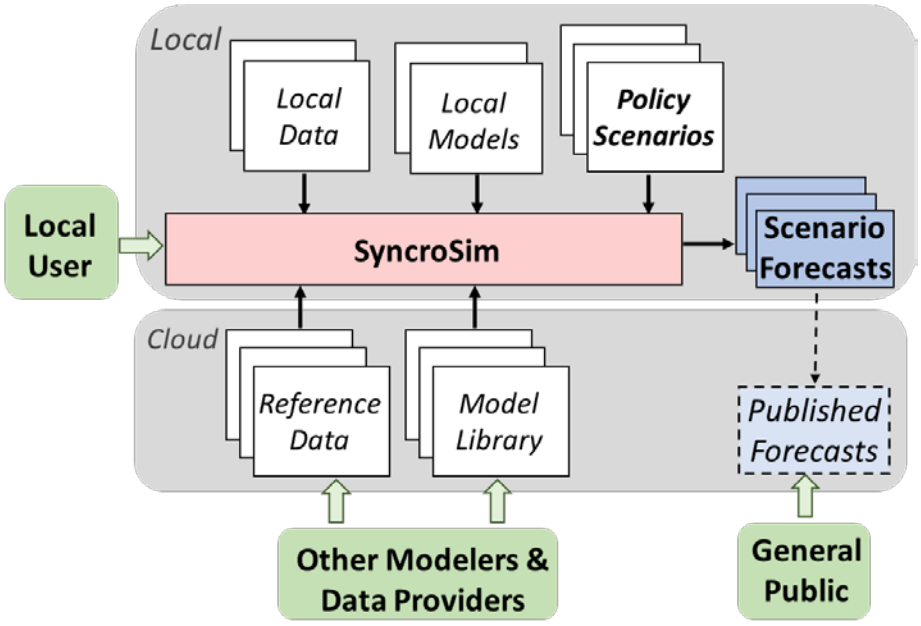

Figure 1: Overview of the scenario-based forecasting framework. At the core of the framework is the SyncroSim software, which mediates interactions between users and their forecasting models, data, and policy scenarios. Local users have the choice of selecting from one or more cloud (i.e. GitHub) based data sources and models, as developed and supported by experts in other jurisdictions, and/or adding their own local data and models. SyncroSim then allows users to generate a range of forecasts based on one or more locally-derived "what-if" future policy scenarios, and to compare scenario forecasts across models and data sources. Forecasts are, by default, generated locally, with users having the option to publish portions of these forecasts online.

Because our proposed solution is based on a mature product, with hundreds of existing users, the solution is stable and deployment ready. Over the past month we have configured our SyncroSim product to support the specific requirements of real-time COVID-19 forecasting; a demonstration version of our software system, providing daily forecasts for Canada and six of its provinces, is available at www.apexrms.com/covid19. In the remainder of this paper we provide additional details on how we have tailored the framework to forecast COVID-19 infections and deaths in Canada.

Case Study: COVID-19 Forecasts for Canada

The best way to understand the framework is to see it in action – that is, to see how it might be applied to make real-time COVID-19 forecasts. For this case study we chose our home country of Canada, demonstrating how the framework can be used to simultaneously generate daily forecasts of COVID-19 infections and deaths for three different jurisdictions: the entire country (*Canada*), a province (*Ontario*), and a city (*Ottawa*).

The following section outlines how the framework was configured for this case study, according to each of the steps shown previously in Figure 1.

Reference Data. As shown previously in Figure 1, the first step in using the framework is to identity relevant *reference data*. Any reference data required by the model must be made available online to SyncroSim. All of the required reference data is downloaded automatically by the framework each day (see https://github.com/ApexRMS/covid19sim). For this case study this includes both Canadian and international daily death data (Figure 3).

Model Library. The next step is to identify one or more available forecasting models from the online *model library*. For this case study we use a single open-source model, which we developed ourselves to forecast future COVID-19 infections and deaths in Canada (see Appendix for a full description of the model). As with the reference data, all of the source code for this model, including the code required to connect the model to SyncroSim, is made available online on GitHub (see https://github.com/ApexRMS/epidemic). Thanks to SyncroSim's tight integration with GitHub-based source control, the framework, in turn, is able to update its projections automatically whenever updates are made to the model.

Local Data. For the *Canada* and *Ontario* jurisdictions, forecasts use only reference data available online; for *Ottawa*, however, we provide an additional spreadsheet of local daily death data for the city, as this data is not readily available online.

Local Models. No additional local models were used in this case study.

***Policy Scenarios*.** For all three jurisdictions (i.e. *Canada*, *Ontario*, and *Ottawa*), we consider the policy option of continuing the public health policy measures indefinitely into the future. In addition, the *Ottawa* forecasts considered an additional scenario that included the gradual lifting of public health measures, where we increased the daily growth rate of infections by 0.5% each week for three weeks, from the current provincial level of approximately 2.5%, starting on May 18, 2020.

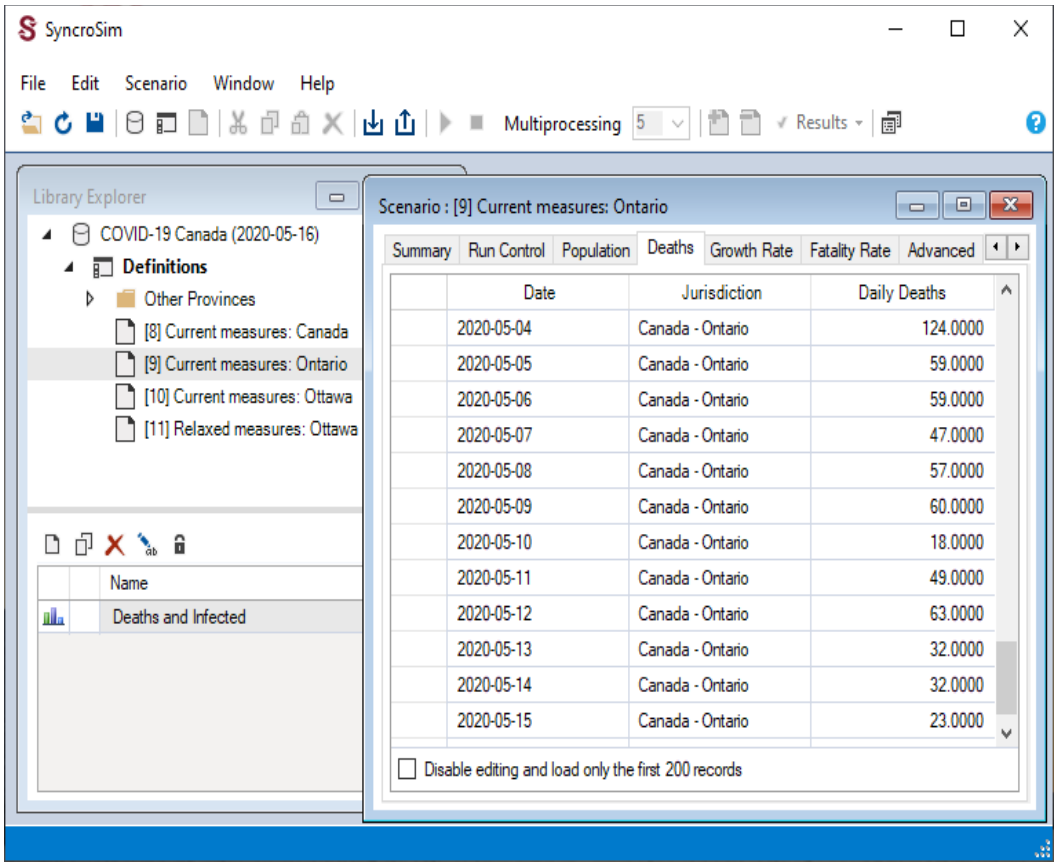

Figure 3: Daily reference death data for the province of Ontario, as downloaded automatically and then displayed in SyncroSim Studio on May 16, 2020.

Scenario Forecasts. Figure 4 shows an example of daily forecasts generated using the modeling framework for the *Canada* and *Ontario* jurisdictions, under a scenario where current public health measures are forecast to continue into the future. Note that the data and projections presented here represent a one-time snapshot forecast, based on the data available at the time of writing this document (i.e. COVID-19 death data publicly available up to and including May 15, 2020); real-time daily forecasts for Canada and its provinces, as generated by this same model, can be found online at www.apexrms.com/covid19. Figure 5 shows the forecasts generated using the framework for *Ottawa* by modifying the scenario for *Ontario* using locally collected data on COVID-19 deaths. Shown here are the results for two alternative public health measure scenarios: a first scenario in which current public health measures continued indefinitely, and a second scenario in which these measures were gradually relaxed over a three-week period.

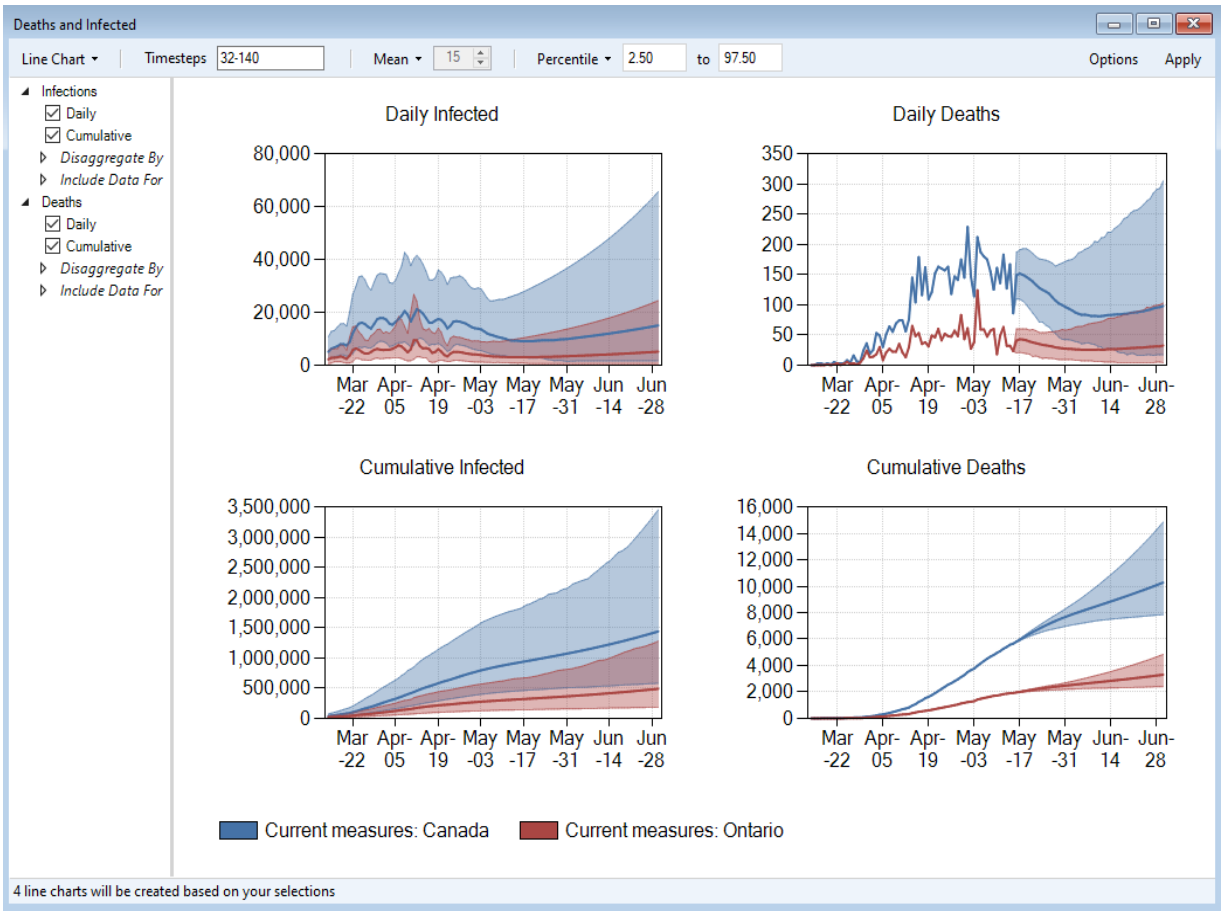

Figure 4: Example of a daily forecast of infections and deaths for Canada and the province of Ontario generated on May 16, 2020, as displayed in SyncroSim Studio under a scenario in which current public health measures remain in place. Solid line indicates the mean number of infections (daily and cumulative) and deaths (daily and cumulative) each day, while zones indicate the 95% Monte Carlo confidence intervals. Deaths up to and including May 15th are actual reported deaths.

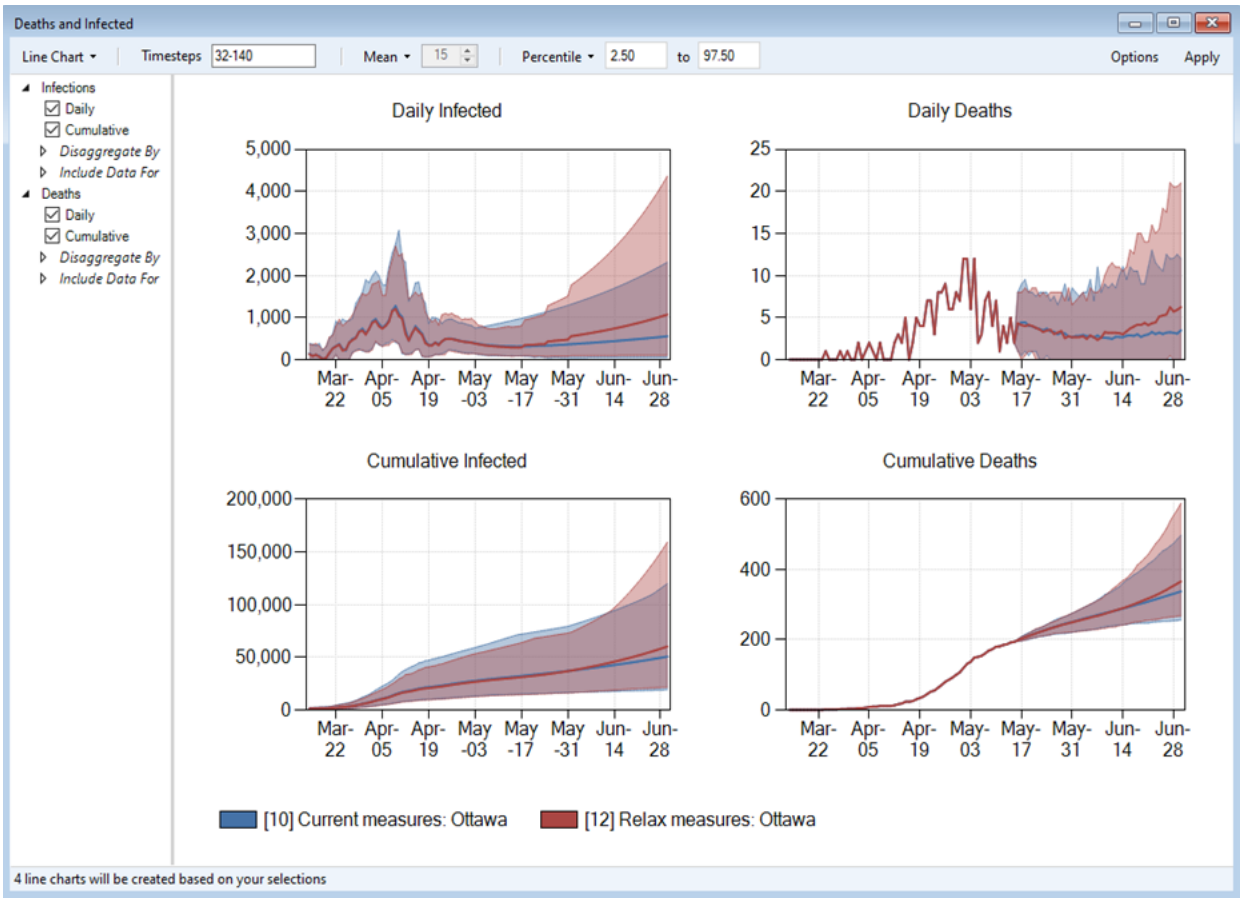

Figure 5: Example of a daily forecast of infections and deaths for the City of Ottawa generated on May 16, 2020, as displayed in SyncroSim Studio under two alternative scenarios: one in which current public health measures remain in place, and a second in which public health measures are gradually relaxed. Solid line indicates the mean number of infections (daily and cumulative) and deaths (daily and cumulative) each day, while zones indicate the 95% Monte Carlo confidence intervals. Deaths up to and including May 15th are actual reported deaths.

Published Forecasts. The final (and optional) step in the forecasting framework is to publish a subset of the locally generated forecasts online in order to share them with others. For this example we automatically push the base (i.e. "Current measures") scenario forecasts, for *Canada* and *Ontario*, back to a GitHub site each day (https://github.com/ApexRMS/covid19sim), and then display these forecasts using a simple web application (Figure 6).

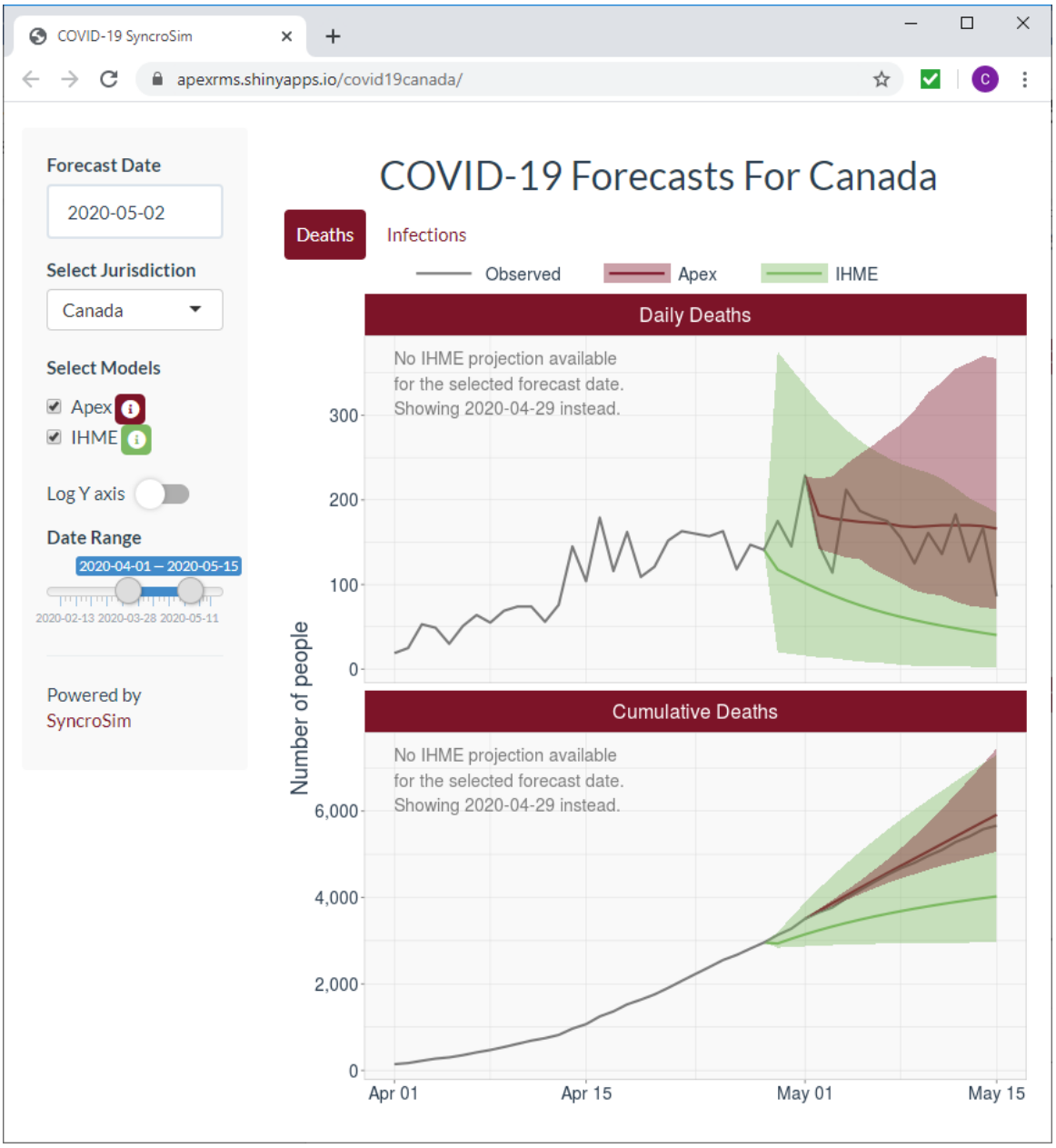

Figure 6: Example of a daily forecast of infections and deaths, as published online for the Canada on May 2, 2020. Through this website, forecasts generated daily since April 18, 2020 for Canada and provinces can be compared to both actual deaths and forecasts from other models. See www.apexrms.com/covid19 for additional forecasts.

Concluding Remarks

The modeling framework presented here provides a novel approach to projecting COVID-19 infections and deaths, allowing users to generate real-time forecasts specific to their jurisdiction and questions. Built upon our existing SyncroSim software platform, the framework brings together the best of the world's open-source forecasting models and real-time data in a consistent format. The framework allows users to rapidly generate their own jurisdiction-specific forecasts for alternative "what-if" scenarios regarding future public health care measures, and to compare and contrast projections made by different models for each of these scenarios. Finally, the framework simplifies the process of sharing locally generated forecasts online. The result is a tool that generates responsive, meaningful, and ultimately actionable forecasts.

Acknowledgments

The case study presented here, along with the underlying infection and death model, was self-funded by Apex Resource Management Solutions. Core development of the SyncroSim software platform upon which this framework is built has been supported by The Nature Conservancy, U.S. Geological Survey, U.S. Forest Service, and Apex Resource Management Solutions. Special thanks to the many visionary scientists from south of the border that helped make SyncroSim what it is today, namely Mr. Ben Sleeter, Dr. James L. Smith, Dr. Louis Provencher, Dr. Zhi-Liang Zhu, Dr. Brian Miller, Dr. Catherine Jarnevich, Dr. Jennifer Costanza, Dr. Jeffrey Morisette , Dr. Wendel Hann and Dr. James Merzenich.

References

Berry, I., J.-P. R. Soucy, A. Tuite, D. Fisman, and C.-C. O. D. W. Group. 2020. Open access epidemiological data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. CMAJ. 2020 Apr 14;192(15):E420. DOI: https://doi.org/10.1503/cmaj.75262

Dong, E., H. Du, and L. Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases. DOI: https://doi.org/10.1016/S1473-3099(20)30120-1

Flaxman, S., S. Mishra, A. Gandy, H. Unwin, H. Coupland, T. Mellan, H. Zhu, T. Berah, J. Eaton, and P. Perez Guzman. 2020. Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. Imperial College London. DOI: https://doi.org/10.25561/77731

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Annals of internal medicine. DOI: https://doi.org/10.7326/M20-0504

Murray, C. J. 2020. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. medRxiv:2020.2003.2027.20043752. DOI: https://doi.org/10.1101/2020.03.27.20043752

Verity, R., L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, and H. Fu. 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases. DOI: https://doi.org/10.1016/S1473-3099(20)30243-7

Appendix: Case Study Model Description

April 20, 2020

The model we use here to demonstrate our framework is similar to the approach used by the IHME model (Murray 2020), in that we fit growth rate curves to time series of COVID-19 death data. In addition to actual death data, our model requires users to make assumptions regarding future growth rates of COVID-19 deaths, including changes in growth rates under alternative public health interventions. As we will demonstrate, such growth rate "scenarios" are typically developed through an analysis of existing death data and/or data and projections from other jurisdictions. Other model inputs include the infection fatality rates, attack rate, and infection period, for which initial estimates are available in the published literature.

Calculating infections and deaths

Given an incubation period (C_t) and symptom-to-death period (S_t) for each day t in our simulation, the model can determine the total infection period each day as $i = C_t + S_t$. The model then uses a time series of actual daily death data, $\{D_t : i + 1 \leq t \leq i + n\}$, for days $i + 1$ to $i + n$, to back-calculate the prior cumulative number of infections, $\{I_t : 1 \leq t \leq \text{n}\}$, for days 1 to n, based on the relationship between daily infections and daily deaths, namely $\Delta I_t = I_t - I_{t-1} = D_{t+i} / F_t$, where F_t is the infection fatality rate. This is repeated for all days t, $1 \leq t \leq n$, for which death data exists. For example, assuming an infection fatality rate 1%, along with an incubation period 5 days and a symptom-to-death period 15 days, the model would calculate that a single death occurring 20 days from now would correspond to 1 / 0.01 = 100 daily infections today. Note that the variables C_t , S_t and F_t are all represented as discrete-time stochastic processes – in other words, they are all random variables that can vary over time; as a result the projected cumulative daily infections, I_t , is also a stochastic process.

The model then uses a simple time series growth model to project the daily cumulative infections forward beyond the last day, n, for which infections can be directly estimated from deaths (i.e. to project I_t for $t > n$). Currently two forms of growth models are supported: exponential and logistic. If the exponential growth model is selected, the future number of daily cumulative infections is calculated as $I_{t+1} = I_t (1 + R_t)$, where R_t is growth rate for day t. A limit to the total number of cumulative infections can also be set as A_tP, where A_t is the user-defined attack rate on day t and P is the total population size. By contrast if the logistic growth model is selected, the future daily cumulative infections is calculated using the discrete form of the logistic equation: $I_{t+1} = I_t R_t ((1 - (I_t - 1 / A_tP)) + 1)$, where R_t is the user-supplied maximum growth rate for day t, and A_tP represents the logistic model's traditional carrying capacity. Note that these two model forms were selected as they both can be parameterized using at most two parameters: a growth rate and an optional attack rate; additional model forms could be easily added in the future (Figure 1). Note also that, as with the other model variables outlined above, the growth rate (R_t) and attack rate (A_t) are also both discrete-time stochastic processes, and thus can also be specified as distributions that vary over time. Finally, the future number of daily deaths, D_t, for those days $t > n$ for which deaths are not provided as a model input, are calculated from the projected number of daily cumulative infections as $D_t = \Delta I_{t-i} F_{t-i}$.

Figure A-1 contrasts the two different forms of growth models that are currently available in the model. With the logistic model the growth rate is exponential at low population levels, but then declines over time as the cumulative number of infections approaches the product of the population and attack rate. By contrast, with the exponential model the cumulative number of infected increases exponentially at a fixed daily growth rate until

the maximum possible infections are reached. However because the growth rate can optionally vary over time, it is possible to create a third type of model, which we refer to as "time-varying" (or non-stationary) exponential growth model, whereby the growth rate of the exponential model changes over time – e.g due to a gradual reduction in growth rate as a result of the introduction over time of public health measures. It is this final option that provides the greatest flexibility for specifying future patterns of infection growth rates in the model, as ultimately any empirical pattern of infection growth can be captured in this way.

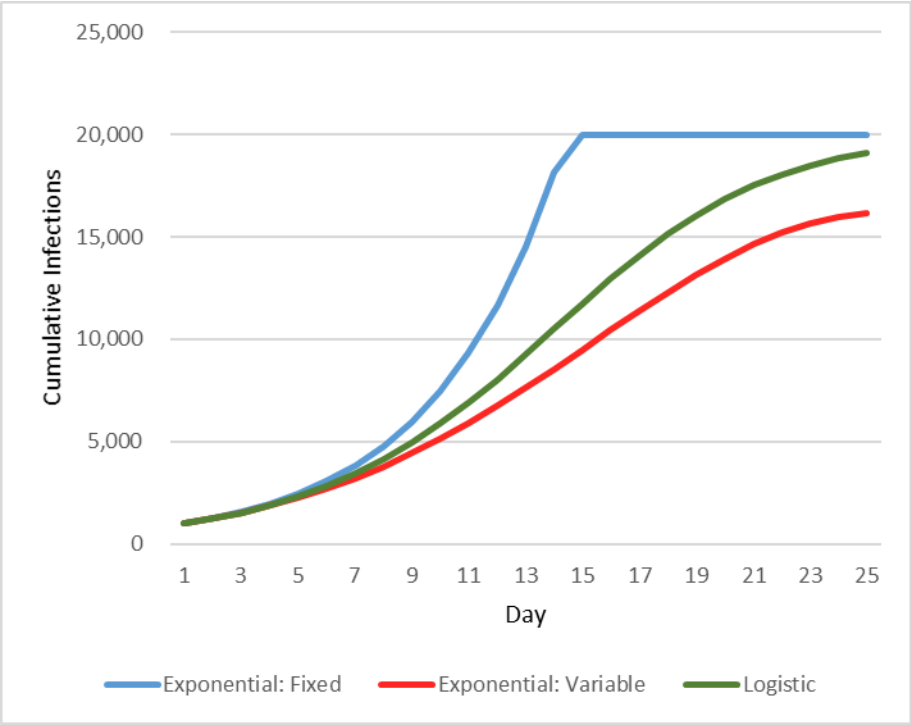

Figure A-1: Cumulative infections projected by the exponential and logistic forms of the infection growth model. Example shows deterministic projections over time starting with 1000 infections, growth rate of 25%, population of 40,000, and attack rate of 50%. The variable exponential curve lowers the initial growth rate of 25% by 1% every day, demonstrating how time-varying growth rates can be used to capture the effects of public health measures on cumulative infections.

Parameterizing and running the model

The equations in the previous section detail the calculations required for a single time series projection of daily infected and deaths for one jurisdiction. The model framework is designed to allow users to run multiple jurisdictions simultaneously (e.g. several counties within a state, or several countries within a region), whereby some inputs are shared by all jurisdictions (e.g. incubation and symptom-to-death periods), while others can vary between jurisdictions (e.g. deaths, population and fatality rates). The framework is also designed to allow some variables (e.g. growth rates) to vary over time; the choice of which variables to vary by jurisdiction and/or date is fully customizable by the user. Finally, a key element of the SyncroSim modeling framework is its ability

to capture uncertainty in projections through the use of random variables and Monte Carlo simulations. As in all Monte Carlo simulations the model can be run for multiple Monte Carlo realizations, with the results of each realization differing from each other due to the random sampling of input values for those model inputs that are uncertain. As shown in Table A-1, most of the model parameters can be specified as distributions, rather than single values; these distributions can also vary spatially (i.e. by jurisdiction) and temporally (i.e. by day). In this way the model can capture a wide range of uncertainty and variability in model parameters.

With the support of the SyncroSim's modeling framework, users are able to chose from a wide range of distributions in order to characterize the uncertainty in model inputs, including empirically generated frequency distributions. Users are also able to vary these distributions over time, and by jurisdiction (e.g. when running a multi-jurisdictional model that might divide a country into several states or provinces). This is one of the key benefits of developing a simulation model using SyncroSim: the framework manages all the Monte Carlo simulations, feeding the required sampled inputs to the core model for each realization and timestep.

The final output of the model is a set of model realizations for the projected time series of daily infections and deaths, which can either be post-processed externally (e.g. using scripts written in R with the help of the rsyncrosim R package), or summarized directly within the SyncroSim platform into time series of mean and Monte Carlo confidence intervals for daily and cumulative infections and deaths.

Table A-1: List of the model inputs and outputs, including which variables can be specified by jurisdiction, date and/or Monte Carlo realization. Input variables that vary by realization can be characterized using either parametric distributions or user-specified empirical frequency distributions. Parametric distributions currently supported by SyncroSim include uniform, normal, beta and gamma.

Type	Variable Name	Equation Symbol	X = required o = optional		
			Jurisdiction	Date	Realization
Input	Population	P	X		
	Deaths	D_t	X	X	
	Growth rate	R_t	o	o	o
	Attack rate	A_t	o	o	o
	Fatality rate	F_t	o	o	o
	Incubation period	C_t	o	o	o
	Symptom-to-death period	S_t	o	o	o
	Model type		o	o	
Output	Daily Infections	ΔI_t	X	X	X
	Daily Deaths	D_t	X	X	X

The source code for the current model is available in open-source format on GitHub at https://github.com/ApexRMS/epidemic. To streamline development, the first version of the core model is written in C# – doing so allowed us to finish and distribute a fully functional version of the model in under two weeks. As SyncroSim supports models written in any language the model could easily be translated to other programming languages (e.g. R or Python) in the future.

Case study model inputs

There are several inputs required in order to run our demonstration model for Canada and its provinces; a summary of the values assigned to each of these inputs for the case study is provided in Table 2.

Table 2: Model inputs for the Canadian case study.

Model Input	Scenario	Values	Approach	Data Source
Population	All	Reported population by jurisdiction	As reported	Statistics Canada (2020)
Deaths	All	Reported daily deaths by jurisdiction	As reported	Berry et al. (2020)
Growth rate	Current Measures	Sampled rates by jurisdiction, date & realization	7-day moving average of future death growth rates sampled from United States, Italy, France, Spain and South Korea	Dong et al. (2020)
Fatality rate	Base Fatality	Distribution by jurisdiction: CA: 1.02% (0.55-1.96) AB: 0.80% (0.43-1.55) BC: 1.07% (0.58-2.07) ON: 1.01% (0.55-1.94) QC: 1.11% (0.60-2.13)	Age standardized fatality rate for each jurisdiction, based on published infection fatality ratio (with 95% credible intervals) and age-structured population for each jurisdiction. Fitted to gamma distribution.	Verity et al. (2020) Statistics Canada (2020).
	High Fatality	As above X 1.5	Base rates increased by 50%	Verity et al. (2020)
Incubation period	All	Single distribution: 4.5-5.8 days	Uniform distribution over published 95% credible interval	Lauer et al. (2020)
Symptom-to-death period	All	Single distribution: 16.9-19.2 days	Uniform distribution over published 95% credible interval	Verity et al. (2020)

The first class of inputs are those that characterize the behaviour of the disease in general, namely the incubation period (i.e. time from infection to first symptoms), and the duration from onset of symptoms to death. Estimates for the distribution of incubation periods for COVID-19 were taken from Lauer et al. (2020), while estimates for the distribution of onset-of-symptoms-to-death periods were taken from Verity et al. (2020). For this demonstration we assumed these values are the same across all jurisdictions in Canada, and that these values do not vary over time. We captured uncertainty in these values by sampling once for each Monte Carlo realization from the distribution of reported values for each of these two periods (Table A-2).

The infection fatality rate was calculated for the *Base Fatality* sub-scenarios using age-specific infection fatality ratios (Verity et al., 2020). These rates were then age-standardized based on the age structure of each jurisdiction's population. Credible intervals (95%) were also estimated for each jurisdiction using age-weighted averaging of the original published age-specific credible intervals. The mean and credible intervals for each jurisdiction were then fit to a gamma distribution and sampled once for each realization. For the *High Fatality* sub-scenarios we increased the mean of the base gamma distributions by 50% (yet left the standard deviation unchanged).

For the *Continue Current Measures* scenario we used an exponential growth model with a time-varying growth rate in order to project the future infections each day. Daily reported deaths in Canada, by province and date, were taken from Berry et al. (2020). Growth rates for the *Continue Current Measures* scenario were estimated for the period beyond which infections cannot be estimated directly from deaths. With death data available at the time of writing through until April 18, this corresponds to the growth in infections beginning on March 27 (i.e. 23 days prior). While there are many different methods that could be used to estimate a distribution for future growth rates for each of our modeled jurisdictions, for our demonstration we chose to base our future growth rates on trends observed for other "reference" jurisdictions – i.e. jurisdictions that had enacted public health measures similar to those used in Canada, yet were ahead of Canadian jurisdictions with respect to the date upon which these measures were taken. Figure 3 shows the 7-day moving average of cumulative death growth rates for several jurisdictions, both in Canada and internationally, adjusted to show the progression in

death growth rates relative to when outbreaks began in each jurisdiction. With the exception of the United States, there is a consistent trend across both the Canadian and international jurisdictions with respect to the decline in growth rates over time. Italy and South Korea are the furthest advanced in terms of time since the start of outbreak, with both countries approximately 3 weeks ahead of Canada in this regard, while France and Spain are about 2 weeks ahead.

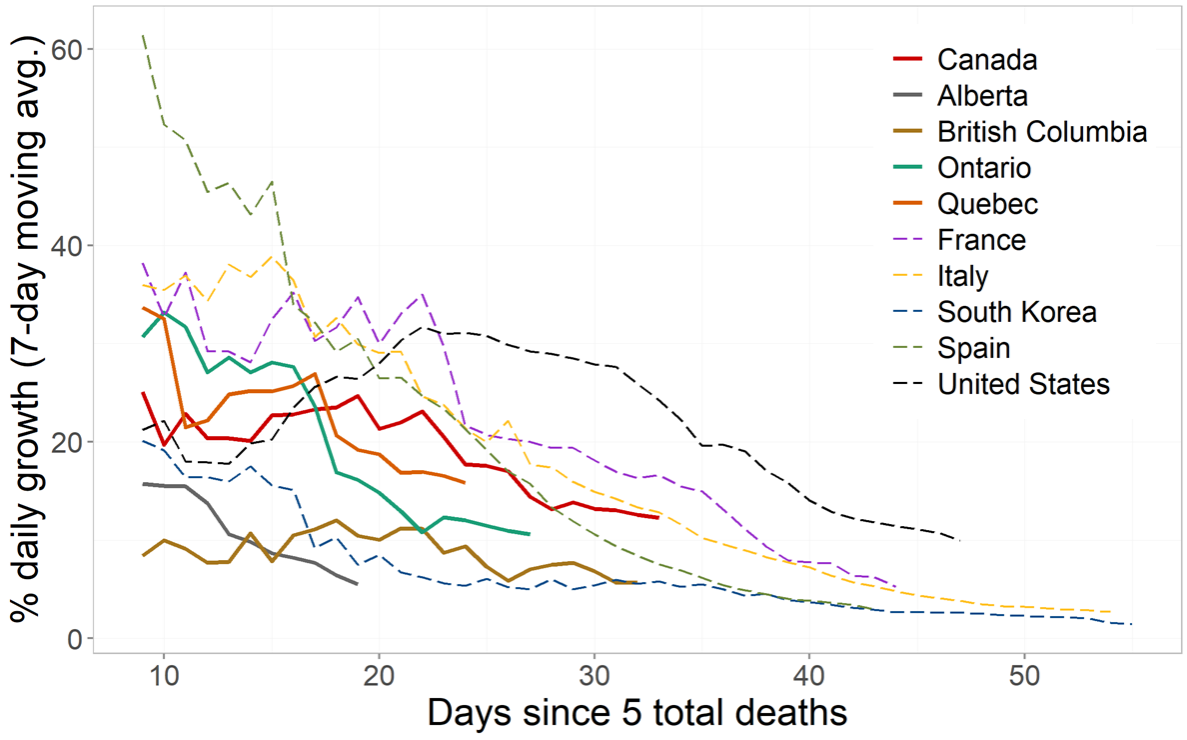

Figure 3: Seven-day moving average of the daily growth rates in cumulative COVID-19 deaths for Canada and four provinces (solid lines), as compared to 5 other countries (dashed lines). Growth rates for each country are displayed relative to the first day in each jurisdiction for which 5 total deaths were reported. Data shown is current to April 18, 2020. Sources: Berry et al. (2020) & Dong et al. (2020).

For our initial demonstration we selected four countries – Italy, Spain, France and South Korea – to act as reference jurisdictions for our projections. For each of our modeled jurisdictions we found the first point in time on each reference country's 7-day moving average growth curve that was equal to the last recorded 7-day moving average growth rate for our modeled jurisdiction. The reference jurisdiction's growth rate beyond this point in time was then used to extrapolate the future growth rate each day for the modeled jurisdiction, with the last growth rate value for each reference country held constant for all days beyond the end of the reference growth rate time series. Figure 4 illustrates the use of this extrapolation technique for the *Canada* jurisdiction, resulting in four possible future growth rate trajectories for this modeled jurisdiction. We repeated this same approach for all five of our modeled jurisdictions.

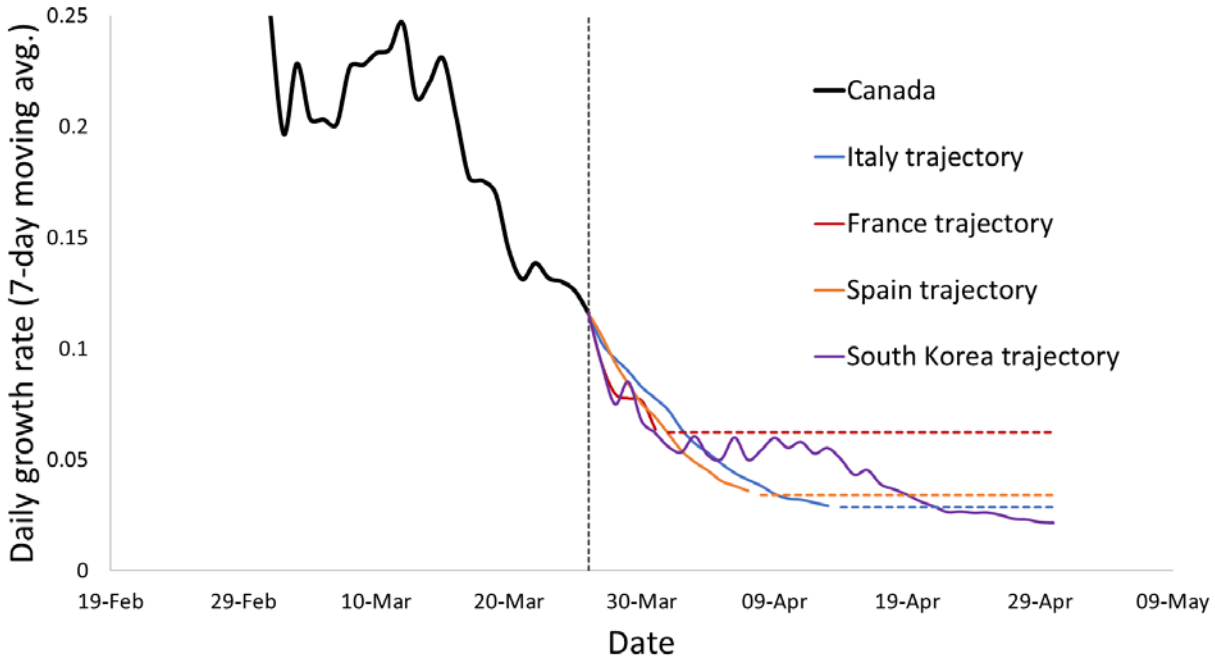

Figure 4: Growth rates in cumulative COVID-19 infections, as supplied as inputs to the model for the *Current Measures* scenario for the modeled *Canada* jurisdiction. Black line represents actual death growth rates calculated from a 7-day moving average of the recorded cumulative deaths, shifted back 23 days from when deaths were recorded. Dotted line indicates the last day for which infections in the model were projected directly from death data; as a result the data represented by the black line are not directly used in the model. Solid coloured lines indicate the trajectory of growth rates corresponding to each reference jurisdiction, based on 7-day moving average of actual growth rates and shifted in time so as to continue the trajectory of the modeled jurisdiction's actual growth rate trend. Dotted lines indicate those reference trajectories that were held constant over time beyond their last actual growth rate value. Data shown are current to April 18, 2020. Sources: Berry et al. (2020) & Dong et al. (2020).

Given the uncertainty that exists as to which of these reference trajectories each of our modeled jurisdictions will follow most closely in the future, we set up our simulations to sample one of these four curves for each of our realizations, each with equal probability. This provides the model with an empirical distribution of future growth curves that spans the full range of possible reference jurisdiction growth rates. Note that this algorithm for estimating future growth rates is particularly powerful because it self-updates the distribution of projected growth rates every day based on the latest death data from each modeled jurisdiction and the four reference jurisdictions; the algorithm also allows reference jurisdictions to be easily added and removed as conditions change in the future.

The start date for all of our case study simulations was set to February 12, 2020, 25 days before the first reported death in Canada; we chose this date in order to capture all of the projected infections that might result through the stochastic back-calculation of infections from deaths. Each of our five modeled jurisdictions were initialized with their actual death data up to and including April 18, 2020. Simulations were run to May 2, 2020 (i.e. for a total of 81 days), two weeks beyond the last day for which deaths were reported. All simulations were repeated for 1000 Monte Carlo realizations.