Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Forecasting COVID-19 infections with the semi-unrestricted Generalized Growth Model

Pablo Pincheira-Brown a,⁎, Andrea Bentancor b

a Business School, Universidad Adolfo Ibáñez, Chile
b Facultad de Economía y Negocios, Universidad de Talca, Santiago, Chile

A R T I C L E   I N F O

Keywords:
SARS-CoV-2
Coronavirus disease
COVID-19
Forecasting
Out-of-sample comparison
Growth model
Phenomenological models
Generalized Growth Model

A B S T R A C T

Recently, the Generalized Growth Model (GGM) has played a prominent role as an effective tool to predict the spread of pandemics exhibiting subexponential growth. A key feature of this model is a damping parameter p that is bounded to the [0, 1] interval. By allowing this parameter to take negative values, we show that the GGM can also be useful to predict the spread of COVID-19 in countries that are at middle stages of the pandemic. Using both in-sample and out-of-sample evaluations, we show that a semi-unrestricted version of the model outperforms the traditional GGM in a number of countries when predicting the number of infected people at short horizons. Reductions in Root Mean Squared Prediction Errors (RMSPE) are shown to be substantial. Our results indicate that our semi-unrestricted version of the GGM should be added to the traditional set of phenomenological models used to generate forecasts during early to middle stages of epidemic outbreaks.

1. Introduction

In May 17th, 2020, the emergence of a novel coronavirus (COVID-19) in the city of Wuhan in China (Wu et al., 2020a,b) had caused approximately 4.8 million cases of confirmed infections and almost 320,000 deaths around the world. Since vaccines and anti-viral drugs were still not available, governments were addressing the COVID-19 outbreak through the implementation of a wide set of public health measures in an attempt to flatten the curve, to delay the peak of the epidemic and to reduce its magnitude (ECDC, 2020; Koo et al., 2020; Pung et al., 2020; Kucharski et al.; Prem et al., 2020). In the span of a few weeks both international and regional authorities had been forced to deal with a new disease that is highly contagious and extremely demanding of national and regional health system resources. Moreover, the lack of a formal evaluation of the accuracy of different available forecasts for the number of infected people has undermined authorities’ capacity to implement adequate policies in order to avoid the collapse of health systems.

In this paper we contribute to the improvement of modeling and forecasting epidemics’ spreads through an in-sample and out-of-sample evaluation of the Generalized Growth Model (GGM) using the current COVID-19 outbreak data coming out from the top ten most affected countries with the disease. We do so by comparing the predictive performance of the traditional GGM with that of an extension of this model that allows for the key damping parameter p to take negative values instead of restricting it to lie within the [0, 1] interval. Our analysis and results for p ∈ (−∞, 1] are consistent with Tolle’s closed-form solution (Tolle, 2003) and with other authors’ posterior research (Chowell and Viboud, 2016). For convenience and clarity, whenever we compare estimates of our proposed extension of the GGM with the GGM itself, we add the expression “Semi-unrestricted”. Then, through this paper, our main analysis focuses on the predictive ability of both the GGM and the Semi-unrestricted GGM (SUGGM). As an extension, however, we also compare the SUGGM with other phenomenological models frequently used in the literature. Our findings indicate that the SUGGM fares very well relative to the Logistic Growth Model (LGM). When compared to the Generalized Logistic Growth Model (GLGM), however, mixed results emerge: for a few countries the SUGGM is able to outperform the GLGM, but in other countries the GLGM is clearly superior to the SUGGM.

It is worth noting that the GGM is widely used in the literature (Chowell and Viboud, 2016; Viboud et al., 2016; Chowell, 2017, etc.) not only because its parsimony and relevance in the public policy debate, but also because it has shown a good fit in early stages of previous pandemics.

Other authors also examine the trajectory of emerging infectious disease outbreaks using phenomenological models (Chowell et al., 2019, 2015; Wu et al., 2020c). Chowell et al. (2019) analyze the GLGM, which also restrict a similar damping parameter to the [0, 1] interval.

https://doi.org/10.1016/j.epidem.2021.100486
Received 18 June 2020; Received in revised form 11 June 2021; Accepted 23 July 2021
Available online 11 August 2021

⁎ Any errors or omissions are responsibility of the authors.
* Correspondence to: Avenida Diagonal Las Torres 2640, Edificio C, Oficina 521-C, Peñalolén, Santiago, Chile.
E-mail addresses: pablo.pincheira@uai.cl (P. Pincheira-Brown), andrea.bentancor@usalca.cl (A. Bentancor).
Chowell et al. (2015) examine the exponential and the polynomial growth models, taking the cumulative number of persons infected as the target variable. Recently, Wu et al. (2020a) calibrated the LGM, the GLGM, the GGM and the Generalized Richards Model (GRM) to the reported number of infected people in the early COVID-19 outbreak in 29 Chinese provinces.

Our in-sample and out-of-sample results show that the SUGGM can outperform the GGM when forecasting the total number of infected people at short horizons, especially in countries that are at middle stages of the spread of the disease. Although most of the studies in the literature of phenomenological models focus mainly on initial stages of pandemics, we think that it is also beneficial to use these models to provide accurate forecasts across the different stages of the disease. For instance, these models can be used as informed benchmarks for more complex mechanistic models, and also they could be used to construct an ensemble forecast. According to Chowell et al. (2020) ensemble forecasts can at times be more accurate than any of the individuals models contributing to the ensemble. In this sense we show that the SUGGM is a fairly useful forecasting device in concave segments of the epidemic trajectories, where a set of negative estimates of the damping parameter are usually obtained.

It is worthwhile to discuss in a few lines the reasons for conducting both in-sample and out-of-sample evaluations. First, while in-sample analyses are widely used and popular, it is also well known in the forecasting literature that they are prone to selection biases and overfitting, even more so than their out-of-sample counterparts, see for instance Ashley et al. (1980), Fair and Shiller (1990) and White (2000). In fact, using an extensive set of Monte-Carlo simulations Clark (2004) provides evidence supporting this view. Second, some other authors prefer out-of-sample relative to in-sample analyses because they better capture the potential instabilities that are likely to be present in the data and because they are more adequate to mimic the actual uncertainty faced by practitioners and policy makers when building forecasts in real time (see, Ashley et al., 1980; Fair and Shiller, 1990; De Gooijer and Klein, 1992; Giacomini and Rossi, 2010; Rossi, 2013). Despite these arguments, it is also true that out-of-sample analyses have some disadvantages relative to their in-sample counterparts. In particular they are typically more computationally involved and criticized for providing tests with low power (see, for example, Inoue and Kilian, 2005). Our reading of this literature is that both types of analyses have their own advantages and disadvantages. As a consequence, in this paper we follow both approaches and check for the robustness of our results across both methodologies.

2. Materials and methods

2.1. Data

We consider daily data on cumulative cases of confirmed COVID-19 infections in the ten countries with the highest numbers of confirmed cases as of May 17th 2020 according to Worldometers. These countries are: Brazil, Germany, Italy, India, Iran, Russia, Spain, UK, the US and Turkey. We exclude France due to an irregularity in its observations (there is a negative figure of new cases in May 16th). Our data starts on the day that all these countries reached at least one hundred of accumulated cases (Turkey is the last one to reach this threshold of confirmed infections in March 18th 2020). Accordingly, our sample of observations covers the period March 18th to May 17th 2020. This means that for each country we have a time series of 61 observations in the cumulative number of infected people, and 60 observations on the first difference of this variable.

In Fig. 1, we present our data. This figure shows five countries with clear concave trajectories at the end of the sample period (Spain, Italy, Germany, United Kingdom, and Turkey). Three other countries show either convex or linear patterns only (Russia, Brazil, India), while the US seems to follow mostly a linear trajectory with a minor concave behavior in the last days of our sample period. Finally, there is one country displaying a mixed pattern: convex at the beginning, then concave and afterwards convex again (Iran).

2.2. Models

The Generalized Growth Model (GGM) is determined by the following simple differential equation

\[
\frac{dC(t)}{dt} = rC(t)^p \quad p \in [0, 1], \quad r > 0
\]  

where \(C(t)\) represents the total cumulative number of people that has been reported as infected during the time interval \([t_0, t]\). We assume that the outbreak of the disease occurs at time \(t_0\) when a number of people equal to \(C(t_0) \equiv C_0\) has been detected as infected. This model has been used by Chowell (2017) and Viboud et al. (2016), mainly to characterize the early incremental phase of contagious in epidemics outbreaks. The GGM is a subtle modification of the exponential growth model. The only difference being the inclusion of the damping \(p\) parameter, which is introduced so to model the spread of a disease at a subexponential growth rate.

The SUGGM that we use in this paper is defined by the same differential equation in (1) but relaxing one restriction on the damping parameter \(p\), which is now allowed to take negative values in the real line:

\[
\frac{dC(t)}{dt} = rC(t)^p \quad p \in (-\infty, 1], \quad r > 0
\]  

Fig. 1. Cumulative cases of confirmed Covid-19 infections (cases/country population in thousands). Source: Cumulative cases of confirmed infections from worldometers.info and population from data.worldbank.org. The “zero” in the x axis represents March 18th 2020.
The solution of the totally unrestricted differential equation in (2) is given by Tolle (2003):

\[
C(t) = \begin{cases} 
\frac{1}{C_0^{p-1}} \left( \frac{t}{C_0^{p-1}} \right)^{\frac{1}{p}} & \text{if } p > 1 \\
\left( \frac{t + pC_0}{p} \right) & \text{if } p = 1 \\
\left( \frac{t - pC_0}{p} \right) & \text{if } p < 0 
\end{cases}
\]

The richness of the proposed SUGGM relies on the fact that different values of the parameter imply quite different speed rates in the spread of the disease: when \( p = 1 \) the model is just exponential, when \( p \) lies in the \((0, 1)\) interval the speed of the disease is subexponential. For \( p = 0 \) we have linear growth, whereas for negative values of the parameter we still have positive growth but at slow rate, slower than linear.

Three additional observations are worth mentioning. First, the solution given by Tolle (2003) for \( p > 1 \) satisfies the positivity constraint for \( C(t) \) only for

\[
t < \frac{1}{(p-1)rC_0^{p-1}}
\]

This means that the totally unrestricted version of the GGM allows faster than exponential growth in “short” periods of time. Given that some authors rule out the case of superexponential growth, in the SUGGM we only relax the positivity constraint in the damping parameter, while keeping the upper bound to \( 1 \).

Second, if one considers the integer \( m \), so that

\[
p = 1 - \frac{1}{m}
\]

then the solution when \( p \in (0, 1) \) is nicely written as

\[
C(t) = \left( \frac{t^m}{m} + C_0 \right) : m > 1
\]  \hspace{1cm} (3)

Expression (3) is fairly eloquent at indicating that for \( 0 < p < 1 \), the growth rate is polynomial and therefore subexponential, which has been useful to characterize early expansion stages in epidemics outbreaks, see Chowell (2017) and Viboud et al. (2016). Third, and following Tolle (2003), whenever \( p < 0 \), we can consider an integer \( n \geq 2 \) such that

\[
p = 1 - n
\]

which allows us to express the solution to (2) as follows:

\[
C(t) = \sqrt[p]{nt + C_0^n} ; n \geq 2
\]

This expression indicates that whenever \( p \) is negative, the disease spreads with the slow speed of a root \( n \) function.

Taking the second derivative in (2) it is clear that

\[
\frac{d^2C(t)}{dt^2} = prC(t)^{p-1} \frac{dC(t)}{dt} = prC(t)^{p-1} C(t)^p = pr^2C(t)^{2p-1}
\]

So the sign of this expression, and hence, the convexity or concavity of the curve \( C(t) \) hinges on the sign of the parameter. In (1) this sign is positive, so the GGM is always characterizing a convex expansion of the disease, which is consistent with its early stages. As the disease enters at middle stages, however, the shape of the curve becomes concave. By allowing the parameter in the SUGGM to take negative values, this model is able to capture exponential growth \((p = 1)\), subexponential growth \((p \in (0, 1))\), linear growth \((p = 0)\) and sub-linear growth \((p < 0)\) which means concavity in the epidemic curve.

2.3. The semi-unrestricted generalized growth model with breaks

Fig. 2 illustrates the prototypical different stages in the spread of a disease. While a curve like that can be generated by a logistic-type of model with two or three constant parameters, it can also be generated in a different way. In particular, Fig. 2 is generated by the SUGGM but allowing for 4 different regimes that are indicated with legends and different colors in the curve. While the usual GGM and SUGGM specifications depend only on two parameters \( p \) and \( r \), it is also possible to generate a curve with a single model but allowing for different regimes in different subsample periods. Each regime is characterized by a different pair of parameters \( (p, r) \). In Fig. 2 we use four different values for the parameters \( p \) and \( r \) to generate the whole curve. Notice that even allowing for breaks it would have been impossible for the GGM to generate a concave part of the curve (the red section) under the traditional restriction of a positive \( p \) parameter. In sharp contrast, when allowing for breaks in its parameters, the SUGGM is able to reproduce a whole epidemiological curve thanks to the semi-unrestricted nature of the \( p \) parameter.

Let us illustrate the use of the SUGGM with breaks using an empirical example. Table 1 next shows very interesting results for the cases of Italy and Germany. This table reports estimates of the \( p \) parameter for the aforementioned countries when the SUGGM is allowed to have

---

1 See Vasconcelos et al. (2021).
Table 1
In-Sample Analysis of the SUGGM allowing for breaks using the UD-Max test of Bai-Perron.

| Panel A: Germany | (1) | (2) | (3) | (4) |
|------------------|-----|-----|-----|-----|
|                  | "p" parameter estimate | Sample period | Number of observations |
| Regime 1         | 0.942 | Feb. 25 - Mar. 11 | 16 |
| Regime 2         | 0.661 | Mar. 12 - Mar. 25 | 14 |
| Regime 3         | -0.413 | Mar. 26 - Apr. 7 | 13 |
| Regime 4         | -3.697 | Apr. 8 - Apr. 21 | 14 |
| Regime 5         | -10.717 | Apr. 22 - May 5 | 14 |
| Regime 6         | -15.292 | May 6 - May 17 | 12 |
| **Total Observations** | **88** |               |               |
| **R-squared**    | **0.961** |               |               |

| Panel B: Italy | (1) | (2) | (3) | (4) |
|----------------|-----|-----|-----|-----|
|                  | "p" parameter estimate | Sample period | Number of observations |
| Regime 1         | 0.787 | Feb. 20 - Mar. 19 | 29 |
| Regime 2         | -0.249 | Mar. 20 - Apr. 2 | 14 |
| Regime 3         | -1.052 | Apr. 3 - Apr. 15 | 13 |
| Regime 4         | -3.065 | Apr. 16 - May 2 | 17 |
| Regime 5         | -8.204 | May 3 - May 17 | 15 |
| **Total Observations** | **88** |               |               |
| **R-squared**    | **0.978** |               |               |

Authors’ elaboration.
Method: Least Squares with Breaks.
HAC standard errors according to Newey and West (1987, 1994).

The maximum of 5 breaks in both parameters \( r \) and \( p \) simultaneously. We use the UD-Max test of Bai and Perron (1998) to endogenously detect breaks in our series. If both parameters \( r \) and \( p \) are stable along time, the UD-Max test should detect no breaks, but if the parameters are not stable, the test will detect these breaks and will re-estimate these parameters in different subsample periods.

Let us consider the case of Italy in Table 1 Panel B. The test detects 4 breaks and therefore 5 different regimes for the SUGGM. Regime 1 goes from 02/20/2020 to 03/19/2020 and indicates an estimate of the \( r \) parameter equal to \( r = 0.787 \), so in this first regime we have sub-exponential growth as mentioned by Viboud et al. (2016). The second regime goes from 03/20/2020 to 04/02/2020 with a \( p \) estimate given by \( p = -0.249 \), that represents sub-linear growth, which is the emphasis of our paper. A similar situation happens in the following three regimes. Table 1 also shows a coefficient of determination (R-squared statistic) equal to 0.978 for Italy. This excellent fit is achieved by allowing for breaks in the SUGGM model. For this result it is critical to allow the \( p \) parameter to be negative. Otherwise it would not be possible to characterize the concave part of the epidemic curve. In this sense our model is more flexible than the traditional GGM. Even allowing for breaks, the restrictions imposed on the \( p \) parameter would have made impossible for the GGM to reproduce a whole prototypical epidemic curve. Notice that the declining phase of the disease is captured by allowing both breaks and negative values of the parameter \( p \).

Fig. 3 next shows the fitted curve for Italy for the period 02/20/2020–05/17/2020 using the results shown in Table 1 Panel B. Fig. 4 shows the actual curve of cumulative cases for Italy. The most impressive feature in Fig. 3 is that it looks just like Fig. 4. They are extremely similar. As mentioned before, the R-squared statistic in Table 1 Panel B is 0.978, which indicates an excellent fit of the SUGGM with breaks to the Italian data. It is important to mention that without breaks the R-squared statistic is lower: 0.707.

2.4. Short-term forecasts

We are interested in short term forecasts. Specifically we consider \( h \)-days-ahead forecasts, where \( h \) goes from 1 to 10. We think that for policy makers it would be useful to have accurate forecasts of infections at these forecasting horizons. Within 10 days, there is plenty of time to reallocate resources from one region to another and, in particular, from one hospital to another. So accurate forecasts at these horizons allow a more efficient performance of health systems.

Let us introduce some notation. We will assume that we are at time \( t \) interested in the construction of forecasts for the cumulative number of infected people at time \( t+h \), where \( h \) represents the forecast horizon. Therefore our target variable is \( C(t+h) \). The forecast for this variable, built at time \( t \) with model \( A \) is denoted by \( C_A(t+h) \). We use a similar notation for the forecast error \( e_A(t+h) \) that is defined as follows:

\[
e_A(t+h) = C(t+h) - C_A(t+h)
\]

Consider now the SUGGM in expression (2) again. Dividing both sides of this equation by \( C(t) \) we obtain

\[
\frac{d\ln(C(t))}{dt} = rC(t)^{p-1}
\]

But

\[
\frac{d\ln(C(t))}{dt} = rC(t)^{p-1}
\]

so the SUGGM is also characterized by the expression

\[
\frac{d\ln(C(t))}{dt} = rC(t)^{p-1}
\]

For the construction of our forecasts we consider a discrete version of this expression as follows:

\[
\frac{\Delta \ln(C_{t+1})}{\Delta(t+1)} = rC_{t+1}^{p-1}
\]

Here \( \Delta \) represents the traditional “difference” operator in time-series, so that

\[
\Delta \ln(C_{t+1}) = \ln(C_{t+1}) - \ln(C_t)
\]

\[
\Delta(t + 1) = (t + 1) - (t) = 1
\]

Therefore our discrete version of the SUGGM reduces to

\[
\frac{\Delta \ln(C_{t+1})}{\Delta(t+1)} = rC_{t+1}^{p-1}
\]

To estimate the unknown parameters \( p \) and \( r \) we use a regression-based approach. By taking logarithms to both sides of expression (4) we end up with the following expression

\[
\ln(\Delta \ln(C_{t+1})) = \ln(r) + (p - 1) \ln(C_t)
\]
By augmenting this last expression by an error term \( \epsilon_{t+1} \) with standard orthogonality assumptions \( (E(\epsilon_{t+1}) = E(\epsilon_{t+1}\ln(C_t)) = 0) \) we finally have the following linear regression:

\[
\ln(\Delta \ln(C_{t+1})) = \ln(\hat{r}) + (p - 1) \ln(C_t) + \epsilon_{t+1}
\]

Let \( \hat{r} \) and \( \hat{p}_1 \) be the least squares estimates of \( \ln(\hat{r}) \) and \( (p - 1) \). The one-step-ahead forecast of \( \ln(\Delta \ln(C_{t+1})) \) is constructed as \( \hat{r} + \hat{p}_1 \ln(C_t) \).

Consequently, the one-step-ahead forecast for \( \ln(C_{t+1}) \) is simply given by

\[
\hat{C}_t^{(1)} = e^{\hat{r} + \hat{p}_1 \ln(C_t)}
\]

We construct multistep ahead forecasts using the iterated approach, explained for instance, in Pincheira and West (2016). According to this approach, at time \( t \) the \( h \)-step ahead forecast for \( \ln(C_t) \) is built in terms of the \( (h - 1) \)-step ahead forecast as follows

\[
\hat{C}_t^{(h)} = e^{\hat{r} + \hat{p}_1 \hat{C}_t^{(h-1)}}
\]

Accordingly, the \( h \)-step ahead forecast of \( C_t \) is simply built as:

\[
\hat{C}_t^{(h)} = e^{\hat{C}_t^{(h)}}
\]

In the case of the GGM, we proceed in the same fashion. The only difference is that the estimates of the \( p \) parameter are restricted to remain in the \([0, 1]\) interval whereas in the SUGGM these estimates are only restricted to remain lower or equal than 1.

2.5. In-sample and out-of-sample evaluations

At each forecasting horizon \( h \) we evaluate the accuracy of the predictions coming from both models (GGM and SUGGM) by means of the Root Mean Squared Prediction Error (RMSPE) defined as

\[
RMSPE_h = \sqrt{\frac{1}{P_h} \sum_{i=1}^{P_h} (\epsilon_i(h))^2}
\]

In practice we approximate this performance metric with sample analogs like this:

\[
\hat{RMSPE}_h = \sqrt{\frac{1}{P_h} \sum_{i=1}^{P_h} (\hat{\epsilon}_i(h))^2}
\]

where \( P_h \) represents the number of \( h \)-days ahead prediction errors. This is a standard performance metric, which is basically the same as the Root Mean Squared Error (RMSE) used by Chowell et al. (2019) and Chowell (2017), for instance. Our emphasis on the “Predictive” word is to point out that our metric depends on the forecasting horizon, and
that can be applied to both residuals coming from an in-sample analysis and to out-of-sample forecast errors.

In our out-of-sample evaluations we split the available time series sample in two halves of 30 observations each. The first half goes from March 19th to April 17th 2020. In the forecasting literature this sample is usually called the “initial estimation window” because it is only used to obtain an initial estimate of the parameters of the models. The second window, that goes from April 18th to May 17th 2020 is usually called “the evaluation window”, although this name is a little misleading because this window has a double duty: it is used to update the parameter estimates of the models and also to build out-of-sample forecasts. The out-of-sample evaluation is summarized in the following steps:

1. First, we set our estimation window as the “initial estimation window” and \( t \) as the date of the last day available in that estimation window.
2. Second, we estimate the parameters of both models (GGM and SUGGM) with the 30 observations available in the estimation window.
3. Third, we build \( h \)-days ahead forecasts of \( C(t+h) \) with both models, for \( h = 1, 2, 3, \ldots, 10 \).
4. Fourth, we construct the out-of-sample forecast errors for each model and forecast horizon as follows:

   \[
   e_{t}^{GGM}(h) = C_t^i - C_{t+h}^{GGM} \\
   e_{t}^{SUGGM}(h) = C_t^i - C_{t+h}^{SUGGM} 
   \]

5. Fifth, we reset our estimation window discarding the first observation of the previous estimation window and adding a new observation at the end, so that the new estimation window still has 30 observations.

If there are no more observations to forecast the procedure stops. If it is possible to forecast an extra observation, we reset \( t \) as the date of the last day available in the new estimation window and we go back to point 2.

The idea in out-of-sample evaluations is to separate the estimation and evaluation stages so to avoid overfitting problems, and to mimic as close as possible a real time forecasting problem, in which a decision maker cannot estimate his/her preferred models with observations that are not available at the moment of the construction of the forecasts. In real time applications it is impossible to build and to evaluate forecasts at the same time, because the evaluation requires observations that will be available in the future. Out-of-sample evaluations try to mimic this process by splitting the available samples in the estimation and evaluation windows described above. More details of out-of-sample evaluations can be found in West (1996, 2006).

The more traditional in-sample evaluations do not split the data in two different windows, so the evaluation process is carried out with the same sample that is used to estimate the parameters. With this approach, in principle one could build forecasts for the entire sample period, but for the sake of a fair comparison with the out-of-sample approach we will restrict to compute in-sample RMSPE considering the period, but for the sake of a fair comparison with the out-of-sample approach, in principle one could build forecasts for the entire sample and evaluation stages so to avoid overfitting problems, and to mimic as close as possible a real time forecasting problem, in which a decision maker cannot estimate his/her preferred models with observations that are not available at the moment of the construction of the forecasts. In real time applications it is impossible to build and to evaluate forecasts at the same time, because the evaluation requires observations that will be available in the future. Out-of-sample evaluations try to mimic this process by splitting the available samples in the estimation and evaluation windows described above. More details of out-of-sample evaluations can be found in West (1996, 2006).

The more traditional in-sample evaluations do not split the data in two different windows, so the evaluation process is carried out with the same sample that is used to estimate the parameters. With this approach, in principle one could build forecasts for the entire sample period, but for the sake of a fair comparison with the out-of-sample approach we will restrict to compute in-sample RMSPE considering the forecast errors corresponding to the last 30 observations of our sample.

### 3. Results

In this section we show two exercises providing evidence in favor of the SUGGM. First, in Fig. 5 we show the least squares estimates of the damping parameter \( p \) estimated in 30 rolling windows of 30 observations each. Estimations are carried out allowing for negative values but restricting the estimates to be lower or equal than 1.

The first window goes from March 19th to April 17th 2020, while the last one begins in April 17th and finishes in May 16th. We show the curves
patterns and their semi-unrestricted summary, there is a clear correspondence between countries’ trajectory are the countries that show either convex or linear patterns only. So, in no negative estimates were found are Russia, Brazil and India. These of cumulative cases we find negative values of the damping parameter that for the seven countries displaying some concavity in the trajectory for all 10 countries in our sample. Fig. 5 is consistent with Fig. 1 in Authors’ elaboration.

Note: Countries are sorted by the number of infected people as of May 18th according to Worldometer.info.

Figures lower than 1 favors the Semi-Unrestricted Model (SUGGM). RMSPE ratios between the SUGGM and the GLGM at different forecasting horizons Table 4

Results in Table 2 show that the cost of using the SUGGM is fairly low when the GGM does a better job. In the worst case scenario, the loss in forecast accuracy is only 4%. Nevertheless, the benefits of using the SUGGM are sometimes very high when the SUGGM outperforms the GGM. Table 2 also reports average results, indicating that reductions in RMSPE are higher than 10% when using the SUGGM.

Notice also that in 69% of the entries in Table 2 the out-of-sample relative performance of the SUGGM is better than its in-sample relative performance. This is probably due to the use of rolling windows. Rolling estimates are very useful when parameter estimates are not stable in different time periods. According to Fig. 5, this is what is happening in the majority of our countries.

Finally, in the Appendix we report Tables A.1 and A.2 in which we show the RMSPE of both models (SUGGM and GGM) in both exercises: in- and out-of-sample. It is interesting to notice that there is an important difference in RMSPE between these two analyses, especially at long horizons, where the out-of-sample RMSPE is almost always higher than its in-sample counterpart. If the out-of-sample analysis offers a more realistic measure of uncertainty as claimed by certain literature, the use of in-sample analyses would be underestimating this uncertainty by quite a large margin.

Though most of the studies in the literature that analyze phenomenological models focus mainly on initial stages of pandemics, we find that the SUGGM improves the forecasting ability of the GGM for many countries that are going through middle stages in the propagation of the disease. This is important, because these simple phenomenological models can be used as informed benchmarks for more complex mechanistic models, and also because they could be used to construct an ensemble forecast. See Chowell et al. (2020).

As a final remark and for robustness, we also carried out the same forecasting exercise described in this section based on a different estimation strategy for our models: Least Absolute Deviations (LAD). for all 10 countries in our sample. Fig. 5 is consistent with Fig. 1 in that for the seven countries displaying some concavity in the trajectory of cumulative cases we find negative values of the damping parameter $p$, at least in some rolling windows. The only three countries for which no negative estimates were found are Russia, Brazil and India. These are the countries that show either convex or linear patterns only. So, in summary, there is a clear correspondence between countries’ trajectory patterns and their semi-unrestricted $p$ estimates in Fig. 5.

The results of a second exercise are presented in Table 2. This table shows the ratio of RMSPE between the SUGGM and the GGM (SUGGM RMSPE/GGM RMSPE) for each of the one ten the countries and each one of the forecast horizons analyzed in this paper (one-day to ten-day ahead forecasts). Figures lower than 1 indicate that a semi-unrestricted estimation of the damping parameter $p$ provides more accurate forecasts. Figures greater than 1 mean the opposite: the nonnegativity restriction in the estimation of $p$ produces more accurate forecasts. Results are reported for both the “in-sample” analysis (top panel in Table 2) and the “out-of-sample” analysis (bottom panel in Table 2).

Results in Table 2 are consistent with those of Fig. 5 in that the SUGGM shows no substantial gains in predictive accuracy relative to the GGM for the three countries for which no negative estimates of the damping parameter were found: Russia, Brazil and India. For the US, UK and Turkey we obtain mixed results: the in-sample analysis shows no gains for the SUGGM, whereas the out-of-sample analysis shows consistent non negligible gains in forecast accuracy in favor of the SUGGM. For the other four countries in our sample: Spain, Italy, Germany and Iran, gains in predictive accuracy are substantial when using the SUGGM. This is consistent across forecasting horizons and both type of analyses, in- and out-of-sample. For Spain, Italy and Germany results are striking in favor of the SUGGM where reductions in RMSPE are sizable: around 50% or even more.
Instead of minimizing the sum of squared errors, LAD minimizes the sum of the errors’ absolute value. This estimator is considered more robust relative to traditional least squares because it is less affected by outliers. Extreme observations leading to extreme errors are highly influential for the least squares estimator that tries to minimize the sum of squared errors. LAD instead considers only the absolute value of the errors. Even in the presence of big errors, this estimator puts relatively less weight on these extreme observations. In this sense LAD is more robust. Table A.3 in the appendix is akin to Table 2. It shows the RMSPE ratios between the SUGGM and the GGM at different forecasting horizons. The only difference now is that parameter estimates of the models are based on the LAD estimator and not on least squares. Interestingly, results with both estimation strategies, LAD and least squares, lead to fairly similar outcomes from the qualitative point of view, with minor quantitative differences. In particular, both approaches indicate that important gains in forecast accuracy are obtained in the cases of Spain, Italy, Germany and Iran when using the SUGGM. With both estimation approaches we obtain mixed results for the US, Turkey and the UK. Finally, both estimation strategies generate a similar forecasting performance between the SUGGM and GGM for Russia, Brazil and India. According to these results, we conclude that our main message is robust to both estimation methods: least squares and LAD.

4. Extensions

In this section we compare the forecasting performance of the SUGGM with other phenomenological models that are frequently used to forecast the cumulative number of cases in epidemic outbreaks. We also provide a simulation study with synthetic data comparing the performance of our regression approach when estimating the key parameters of the SUGGM and GGM in two different scenarios: one in which the true damping parameter $p$ belongs to the $[0, 1]$ interval, and a second scenario in which this damping parameter $p$ is set to a negative value.

4.1. Forecasting performance relative to other common benchmarks

Here we present a formal forecasting evaluation comparing our SUGGM model with the LGM and the GLGM models used, for instance, by Roosa et al. (2020). These models are characterized by the following differential equations:

$$\frac{dC(t)}{dt} = rC(t) \left(1 - \frac{C(t)}{K}\right), \quad r > 0 \quad (LGM)$$

$$\frac{dC(t)}{dt} = rC(t)^q \left(1 - \frac{C(t)}{K}\right), \quad q \in [0, 1], \quad r > 0 \quad (GLGM)$$

Using a discrete version of these models, just like we do with the GGM and the SUGGM, we estimate the parameters of the LGM and GLGM using least squares. Then, we compute multi-step-ahead forecasts for $h = 1$, through $h = 10$ days, just like we explain in Section 2.

We evaluate the predictive performance of these models by means of the RMSPE. We present a summary of our evaluation results in the following Tables 3–4. These tables contain RMSPE ratios for each one of the ten countries and each one of the forecast horizons analyzed in our paper (one-day to ten-day ahead forecasts). In Table 3 we compare the SUGGM with the LGM, whereas in Table 4 we compare the SUGGM with the GLGM. Figures lower than 1 favors the SUGGM, while figures greater than 1 indicate that the SUGGM is outperformed by the corresponding competing model. We organize the information in two panels: the first one provides in-sample results, while the bottom one provides out-of-sample results.

Let us focus first on Table 3. This table shows a remarkable forecasting performance for the SUGGM relative to the traditional LGM. As it can be seen through both methodologies, in-sample and out-of-sample, we observe robust and substantial gains in accuracy for almost all countries and horizons when forecasting with the SUGGM. On average the RMSPE of the SUGGM is about a half of the RMSPE of the LGM.

Table 4 shows a different picture. In this table results are mixed, sometimes favoring the SUGGM, some other times favoring the GLGM. Results are relatively stable across horizons, though, in the sense that if one model outperforms the other for a given country, it does so at every forecasting horizon (the only exception is the US in the in-sample comparison). If we focus on the out-of-sample analysis, The SUGGM is clearly outperforming the GLGM for India, Iran, Brazil, Russia, Spain and the US. For the other countries, the GLGM is preferred, and sometimes for a wide margin. In other words, with the only exceptions of Spain and Iran, the GLGM is doing a better job out-of-sample in the countries in which the disease is at middle or late stages.
In conclusion, we see that the SUGGM is not only able to outperform its natural GGM competitor, but also the traditional LGM and, at least for some countries, the GLM. We think that these results indicate that the SUGGM is a competitive predictive tool that should be taken into consideration when forecasting COVID-19 infections with phenomenological models.

4.2. Simulation study with synthetic data

Here we provide a simulation study with synthetic data comparing the performance of our regression approach when forecasting and estimating the key parameters of the SUGGM and GGM models. In the next Table 5 we present in-sample estimates of the parameter \( p \) for the SUGGM. We also show estimates of the parameter \( \ln(r) \) and standard errors for all the estimates computed for the sample period March 18th 2020 through May 17th 2020. In Table 5 we observe outcomes of the estimation of the SUGGM for the 10 countries in our sample. In 4 cases we see negative estimates of the damping parameter \( p \) (Spain, Italy, Germany and Iran). In other 6 cases we observe positive estimates lying within the [0, 1] interval. Consequently we analyze two prototypical cases: one with a negative parameter \( p \), and other with a positive parameter lower than 1. To that end, we pick pairs of parameter values \((p, \ln(r))\) inspired in the estimates of Italy (negative \( p \)) and Brazil (positive \( p \)). Therefore we generate the following two prototypical epidemics curves:

1. Data Generating Process 1 inspired in Brazil: we pick \( p = 0.7 \), \( \ln(r) = 0.5 \) and \( C(0) = 100 \).
2. Data Generating Process 2 inspired in Italy: we pick \( p = -1 \), \( \ln(r) = 20 \) and \( C(0) = 100,000 \).

We generate 1000 replications of these curves using the bootstrap method outlined in Chowell (2017) based on Poisson innovations. Basically, once we have the epidemic curve generated using the estimated parameters from Brazil and Italy, we generate 1000 bootstrap replications of 61 observations each, assuming that daily new cases come from a Poisson distribution with parameter \( \lambda \) equal to the daily new cases implied by the originally estimated cumulative curves. Each bootstrap replication contains 61 observations covering the same sample period in our data from March 18th 2020 to May 17th 2020. We pick an initial number of cases of 100 for Data Generating Process 1 (DGP1) to mimic the case of the early stages in an outbreak. For DGP2 we consider a case in which we are at middle stages, so we generate 61 observations starting from \( C(0) = 100,000 \), which is inspired in the average accumulated number of cases for Italy in our sample period.

Our first exercise is to estimate the parameters of the GGM and SUGGM models with least squares using our regression approach for the 1000 bootstrap replications of each DGP. Summary results of this exercise are displayed in Table 6.

For DGP 1 both models (GGM and SUGGM) provide excellent estimates for the parameter \( p \) with an average across the 1000 replications of 0.71 for the SUGGM and an average of 0.69 for the GGM. Estimates of \( \ln(r) \) are slightly biased in both models. While the true value is 0.50, the average across the 1000 replications yields a figure of 0.42 with the SUGGM and of 0.66 for the GGM. Despite the bias detected when estimating \( \ln(r) \), in general terms both models provide estimates that are close to the true values with which we generated our replications.

For DGP2 we have a different picture because the GGM is unable to detect the negative value of the true \( p \) parameter. In the GGM the parameter \( p \) is restricted to the \([0, 1]\) interval. Consequently, our estimate of that parameter imposes that restriction in the estimation process. As a result, for the GGM we obtain always an estimate equal to zero for this damping parameter. This is intuitive: least squares estimates try to minimize the objective function and go as low as they can get given the restriction imposed on them. As a collateral effect the estimate of \( \ln(r) \) is extremely biased as well. While the true value is 20, the average across 1000 bootstrap replications is 7.85. In sharp contrast, the SUGGM that allows for negative values in the \( p \) parameter, provides fairly adequate estimates: -0.97 on average for the \( p \) parameter that is set to -1, and 19.58 for the \( \ln(r) \) parameter which has a true value of 20.

In summary, for DGP2 the SUGGM clearly outperforms the GGM in estimating the parameters of the models. This is the key mechanism that we are trying to raise in our paper. The nonnegativity restriction on the \( p \) parameter can be detrimental when trying to fit and forecast the cumulative number of cases in different stages of an outbreak.

As a final exercise, with all the 1000 bootstrap replications for each DGP, we repeat the exact same forecasting evaluation described in Section 2. We compute 1 through 10 day ahead forecasts for each series of our synthetic data and compute RMSPE. We do this for both models SUGGM and GGM. Then, we compute RMSPE ratios and take the average of these ratios across our 1000 replications. Results are displayed in Table 7, in which figures lower than 1 favor the SUGGM.

Table 7 shows average RMSPE ratios very close to 1 for DGP 1. Actually, they all range between 0.95 and 1.04. This is consistent with both models (SUGGM and GGM) being able to correctly estimate their parameters \( p \) and \( \ln(r) \). In sharp contrast, for DGP 2 Table 7 shows very low average RMSPE ratios: between 0.02 and 0.11, indicating an outstanding forecasting superiority of the SUGGM over the GGM. This is consistent again with results in Table 6, which clearly shows that the GGM is unable to correctly estimate the negative damping parameter \( p \) in DGP2. As this model estimates \( p \) with a zero, it is forecasting assuming linear growth instead of the correct sub-linear growth pattern associated to the true value of \( p = -1 \).

In summary, our exercises with synthetic data shows that whenever the damping parameter \( p \) is negative, the GGM is outperformed by the SUGGM when estimating the key parameters of the models and when generating short term forecasts for the cumulative cases of COVID-19.

| Table 5 | Point estimates and Standard Errors of the parameters of the SUGGM | In sample analysis. |
|---|---|---|
| Semi-unrestricted generalized growth model | Parameter | \( \text{F} \) | Standard Error (F) | \( \ln(r) \) | Standard Error (\( \ln(r) \)) |
| United States | 0.30 | 0.08 | 6.12 | 1.02 |
| Russia | 0.78 | 0.03 | -0.03 | 0.31 |
| Spain | -0.32 | 0.19 | 11.97 | 2.25 |
| Brazil | 0.70 | 0.04 | 0.50 | 0.36 |
| United Kingdom | 0.37 | 0.08 | 4.02 | 0.95 |
| Germany | -0.68 | 0.27 | 15.44 | 3.08 |
| Turkey | 0.36 | 0.10 | 3.74 | 1.02 |
| Iran | -0.16 | 0.21 | 9.12 | 2.36 |
| India | 0.77 | 0.03 | -0.33 | 0.29 |

Source: Authors’ elaboration.
5. Discussion

Accurate forecasts for the number of total infected people during disease outbreaks are critical for the adequate implementation of control measures and an effective allocation of medical resources. This is particularly important in the case of the current COVID-19 outbreak, given that anti-viral drugs are still not available and an important share of the population will require medical attention, stressing healthcare systems at a national level. The implementation of social distancing, stay-at-home recommendations, disease surveillance, mass gathering cancellations, among others public health measures, could be more effective if they are implemented with an accurate appraisal of the most likely upcoming scenarios.

The COVID-19 pandemic has had a drastic impact around the globe. Furthermore, the outbreak has been sudden and fast. In many countries the available time series on infections, deaths and recoveries had less than a hundred observations at the time we carried out our main analyses. This is extremely challenging when dealing with heavily parameterized models that require a large number of observations to obtain reliable estimates of their key parameters. In these circumstances, simple and parsimonious models displaying a low number of unknown parameters may be extremely useful. Accordingly, phenomenological models emerge as an important tool to build real-time forecasts in this type of epidemic outbreaks.

While having models to build forecasts is important, their blind application without a proper evaluation of their accuracy can be misleading. It is also necessary to carry out a formal assessment of their precision, especially in a context in which many options are available and it is not clear which model will outperform the others. In addition, as suggested by Chowell et al. (2020), the danger of overfitting may be very pervasive, and traditional in-sample measures of accuracy suffering from this problem may underestimate the actual uncertainty faced by health system authorities. Consequently, formal forecast evaluations that can mitigate the potential problems of overfitting are required.

With the focus on the GGM, a model that has been shown to generate adequate forecasts in the early stages of pandemic outbreaks, in this paper we make a twofold contribution. First we propose an extension of the GGM that is useful to generate forecast of infections in a number of countries. Second, we evaluate the accuracy of forecasts coming from the GGM and our extension (the SUGGM) using both in-sample and out-of-sample approaches. The latter has been reported by Clark (2004) as an effective method to mitigate overfitting problems.

According to our results, the SUGGM can substantially improve the accuracy of the forecasts relative to its natural competitor (the GGM) and other traditional phenomenological models. Although the most striking gains in forecast accuracy relative to the GGM are shown for countries that are at middle stages of the COVID-19 disease (Italy, Spain and Germany), we also show important gains in other countries like Iran, that at the time in which we collected our data seemed to be experiencing a second outbreak, and Turkey and the US, that were at earlier stages of the disease. Notice that good forecasts are also needed at middle stages, especially when a high number of infected people continue to put a lot of pressure on national health systems in the more affected countries.

Our reported in-sample and out-of-sample RMSPE are quite different. In particular, at the longest horizon of ten days, out-of-sample RMSPE are much higher than their in-sample counterparts for most of the countries in our sample. This is intriguing and deserves attention. On the one hand it is possible to think that in-sample analyses might provide a more realistic measure of the true uncertainty faced by forecasters. This is so because if the models are correctly specified then the use of the total number of available observations for estimation would imply more accurate estimates, forecasts and hence a better assessment of forecast uncertainty. On the other hand, if the relatively high accuracy reported by our in-sample analysis is the result of data-induced overfitting, then in-sample RMSPE estimates should be considered with caution and more attention should be placed on out-of-sample results.

A final point is worthy of mention: some other models within the phenomenological family make use of a nonnegativity constraint for their damping parameters. It would be interesting to explore if the approach we propose here could be beneficial for these models as well.
### Table A.1
Out-of-sample RMSPE of the GGM and SUGGM at different forecasting horizons
Figures represent number of infected people.

| Forecast horizons in days (1 through 10) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|----------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| **Out-of-sample evaluation SUGGM**     |     |     |     |     |     |     |     |     |     |     |
| United States                          | 6,153 | 11,774 | 17,048 | 22,095 | 27,009 | 32,462 | 39,247 | 47,953 | 58,024 | 68,719 |
| Russia                                 | 1,591 | 3,392 | 5,448 | 7,917 | 10,754 | 14,012 | 17,847 | 22,799 | 27,801 | 34,117 |
| Spain                                  | 1,230 | 2,130 | 3,022 | 4,214 | 5,584 | 6,941 | 8,372 | 9,704 | 10,285 | 11,894 |
| Brazil                                 | 3,230 | 6,448 | 9,551 | 12,687 | 16,891 | 21,195 | 26,519 | 32,943 | 40,423 | 48,971 |
| United Kingdom                         | 1,154 | 2,375 | 3,684 | 5,077 | 6,586 | 8,257 | 10,109 | 12,197 | 14,514 | 16,941 |
| Italy                                  | 502 | 1,000 | 1,553 | 2,141 | 2,800 | 3,561 | 4,424 | 5,388 | 6,240 | 7,525 |
| Germany                                | 789 | 1,596 | 2,438 | 3,265 | 4,103 | 5,053 | 6,169 | 7,407 | 8,732 | 10,117 |
| Turkey                                 | 1,426 | 3,055 | 4,929 | 7,124 | 12,518 | 15,681 | 19,188 | 23,059 | 27,295 |
| Iran                                   | 438 | 909 | 1,432 | 1,985 | 2,568 | 3,158 | 3,768 | 4,385 | 5,023 | 5,676 |
| India                                  | 529 | 979 | 1,496 | 2,102 | 2,753 | 3,448 | 4,205 | 5,031 | 5,915 | 6,870 |

Note: Countries are sorted by the number of infected people as of May 18th according to Worldometer.info.
Source: Authors’ elaboration.

### Table A.2
In-sample RMSPE of the GGM and SUGGM at different forecasting horizons
Figures represent number of infected people.

| Forecast horizons in days (1 through 10) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|----------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| **In-sample evaluation SUGGM**         |     |     |     |     |     |     |     |     |     |     |
| United States                          | 6,544 | 12,397 | 17,927 | 23,283 | 28,525 | 34,107 | 40,770 | 49,130 | 58,697 | 68,742 |
| Russia                                 | 1,570 | 3,343 | 5,365 | 7,794 | 10,584 | 13,788 | 17,561 | 22,038 | 27,354 | 33,588 |
| Spain                                  | 1,372 | 2,920 | 4,947 | 7,531 | 8,149 | 9,207 | 10,788 | 13,016 | 16,938 | 18,956 |
| Brazil                                 | 3,233 | 6,456 | 9,561 | 12,687 | 16,891 | 21,195 | 26,519 | 32,943 | 40,423 | 48,971 |
| United Kingdom                         | 1,235 | 2,512 | 3,862 | 5,282 | 7,021 | 7,755 | 9,548 | 11,614 | 13,516 | 14,504 |
| Italy                                  | 1,307 | 2,701 | 4,199 | 5,775 | 7,458 | 9,275 | 11,223 | 13,288 | 15,442 | 17,693 |
| Germany                                | 1,454 | 2,957 | 4,562 | 6,225 | 7,963 | 9,836 | 11,874 | 14,047 | 16,340 | 18,733 |
| Turkey                                 | 1,602 | 3,383 | 5,382 | 7,666 | 10,267 | 13,130 | 16,271 | 19,688 | 23,407 | 27,427 |
| Iran                                   | 583 | 1,188 | 1,829 | 2,479 | 3,141 | 3,801 | 4,474 | 5,169 | 5,869 | 6,952 |
| India                                  | 531 | 984 | 1,506 | 2,119 | 2,778 | 3,482 | 4,250 | 5,088 | 5,984 | 6,950 |

Note: Countries are sorted by the number of infected people as of May 18th according to Worldometer.info.
Source: Authors’ elaboration.

### CRediT authorship contribution statement

Pablo Pincheira-Brown: Conceptualization, Design, Formal analysis, Writing – original draft, Writing – review & editing. Andrea Bentancor: Conceptualization, Design, Formal analysis, Writing – original draft, Writing – review & editing.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Appendix

See Tables A.1–A.3.
Giacomini, R., Rossi, B., 2010. Forecasts comparisons in unstable environments. J. Appl.

Fair, R.C., Shiller, R., 1990. Comparing information in forecasts from econometric models. Amer. Econ. Lit. 51 (4), 1063–1169. http://dx.doi.org/10.1257/aer.51.4.1063.

Tolle, J., 2003. Can growth be faster than exponential, and just how slow is the logarithm? The Mathematical Gazette 87, 522–525, https://www.jstor.org/stable/36180361.

Chowell, G., Viboud, C., 2016. Is it growing exponentially fast? Impact of assuming polynomial growth rates. PLoS Curr. http://dx.doi.org/10.1371/currents.outbreaks.58b55f4bad99ac5c5db3663e916803261.

Chowell, G., Viboud, C., 2016. A comparison of some out-of-sample tests of predictability in iterated multi-step-ahead forecasts. Econ. Res. 70 (2016), 304–319. http://dx.doi.org/10.1016/j.eire.2016.03.002.

Prem, K., Liu, Y., Russell, T.W., Kucharski, A.J., Eggo, R.M., Davies, N., 2020. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Publ. Health 5, e261–e270. http://dx.doi.org/10.1016/S2468-2667(20)30144-4.

Epидемические 37 (2021) 100486

| Forecast horizon in days (1 through 10) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------------------------------------|---|---|---|---|---|---|---|---|---|---|
| United States                          | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| Russia                                 | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| Spain                                  | 0.76|0.71|0.68|0.65|0.62|0.60|0.58|0.56|0.54|0.53|
| Brazil                                 | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| United Kingdom                         | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| Italy                                  | 0.48|0.46|0.44|0.43|0.43|0.41|0.41|0.40|0.38|0.37|
| Germany                                | 0.47|0.45|0.42|0.40|0.39|0.38|0.37|0.36|0.35|0.35|
| Turkey                                 | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| Iran                                   | 1.00|1.00|0.98|0.97|0.96|0.94|0.93|0.92|0.91|0.90|
| India                                  | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| **Average**                            | 0.87|0.86|0.85|0.85|0.85|0.84|0.83|0.83|0.82|0.82|
| **Minimum**                            | 0.47|0.45|0.42|0.40|0.39|0.38|0.37|0.36|0.35|0.35|
| **Maximum**                            | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|
| **Out-of-sample evaluation**           | 0.81|0.81|0.81|0.81|0.81|0.81|0.81|0.81|0.81|0.81|
| **Minimum**                            | 0.41|0.40|0.40|0.40|0.40|0.40|0.40|0.40|0.40|0.40|
| **Maximum**                            | 1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|1.00|

Note: Countries are sorted by the number of infected people as of May 18th according to Worldometer.info.

Source: Authors’ elaboration.

References

Anon, 2020. Considerations Relating To Social Distancing Measures in Response To COVID-19 – Second Update. ECDC, European Centre for Disease Prevention and Control, https://www.ecdc.europa.eu/en/publications-data/considerations-relating-social-distancing-measures-response-covid-19-second-update.

Ashley, R., Granger, C.W.J., Schmalensee, R., 1980. Advertising and aggregate consumption: an analysis of causality. Econometrica 48, 1149–1167. http://dx.doi.org/10.2307/20191217.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66 (1), 47–78. http://dx.doi.org/10.2307/2998546, https://www.jstor.org/stable/2998546.

Chowell, G., 2017. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. Infect. Dis. Modelling 2, 379–398. http://dx.doi.org/10.1016/j.idm.2017.08.001.

Chowell, G., Luo, R., Sun, K., Roosa, K., Tariq, A., Viboud, C., 2020. Real-time forecasting of epidemic trajectories using computational dynamic ensembles. Epidemics (30), http://dx.doi.org/10.1016/j.epidem.2019.100379, 2020.

Chowell, G., Tariq, A., Hyman, J.M., 2019. A novel sub-epidemic modelling framework for short-term forecasting epidemic waves. BMC Med. 17, 164. http://dx.doi.org/10.1186/s12916-019-1406-6.

Chowell, G., Viboud, C., 2016. Is it growing exponentially fast? Impact of assuming exponential growth for characterizing and forecasting epidemics with initial near-exponential growth dynamics. Infect. Dis. Modelling 1, 71–78. http://dx.doi.org/10.1016/j.idm.2017.08.001.

Chowell, G., Viboud, C., 2016. Power law behaviour in the saturation regime of fatality curves of the COVID-19 pandemic. Sci. Rep. 11, 4619. http://dx.doi.org/10.1038/s41598-021-84165-1.
Viboud, C., Simonsen, L., Chowell, G., 2016. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. Epidemics 15, 27–37. http://dx.doi.org/10.1016/j.epidem.2016.01.002.

West, K.D., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067-1084, https://www.ssc.wisc.edu/~kwest/publications/1990/Asymptotic%20Inference%20about%20Predictive%20Ability.pdf.

West, K.D., 2006. Forecast evaluation. In: Graham Elliott, Clive W.J. Granger, Timmermann, Allan (Eds.), Handbook of Economic Forecasting, Vol. 1. Elsevier, http://dx.doi.org/10.1016/S1574-0706(05)01003-7 (Chapter 3).

White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126. http://dx.doi.org/10.1111/1468-0262.00152.

Wu, K., Darce, D., Wang, Q., Sornette, D., 2020a. Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. https://www.medrxiv.org/content/10.1101/2020.03.11.20034363v1.full.pdf.

Wu, J.T., Leung, K., Bushman, M., Kishore, N., Niehuis, R., de Salazar, P.M., Cowling, B.J., Lipsitch, M., Leung, G.M., 2020b. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nat. Med. Lett. 26, 506–510. http://dx.doi.org/10.1038/s41591-020-0822-7.

Wu, J.T., Leung, K., Leung, G.M., 2020c. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 395, 689–697. http://dx.doi.org/10.1016/S0140-6736(20)30260-9.