Short-term forecasts of COVID-19 - preliminary version 2020-03-19

Jennifer L. Castle, Jurgen A. Doornik and David F. Hendry*
Department of Economics, Nuffield College, Magdalen College, and Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, UK

January 2020

Abstract

JEL classifications: C51, C22.
KEYWORDS: COVID-19; Model Selection; Robustness; Outliers; Location Shifts; Trend Indicator Saturation; *Autometrics*.

1 Introduction

Our aim is to provide short-term forecasts of the number of confirmed cases and deaths attributed to COVID-19. These forecasts may be a useful guide as to what happens in the next few days. For example, the reports on 2020-03-17 that Italian deaths increased by 16% was largely in line with our forecast of 18%, and should not have been a surprise.

The methodology to construct the forecasts involves several steps. First, the observed daily time series is decomposed in a trend and a remainder term. The trend is estimated by taking moving windows of the data and saturating these by linear trends. Selection from these trends is made with an econometric machine learning algorithm, and the selected linear trends are then averaged to give the overall flexible trend. Next, the trend and remainder terms are forecast separately using the Cardt method and recombined in a final forecast. Cardt is a somewhat improved version of Doornik, Castle, and Hendry (2020).

The target variables are cumulative daily counts, which grow exponentially in the initial endemic phase. For forecasting to be effective, we need to step down, not just to the daily increments, but the change in the daily increments. At some stage in the spread of a virus the counts will settle down: at most the entire population can be affected, but usually this is well before that happens.

Estimation of the trend is subject to several data challenges. First of all, policy interventions aim to suppress the transmission of the SARS-CoV-2 virus that causes COVID-19. Futhermore, countries have different testing strategies and technologies, and these are occasionally revised. So the counts will be subject to structural breaks, underreporting, and errors, but nonetheless they are the only readily available information.

*This is a preliminary draft: please do not quote without permission from the authors. Financial support from the Robertson Foundation (award 9907422) and Institute for New Economic Thinking (grant 20029822) is gratefully acknowledged, as are helpful comments from Andrew B. Martinez. email: jennifer.castle@magd.ox.ac.uk, jurgen.doornik@nuffield.ox.ac.uk and david.hendry@nuffield.ox.ac.uk

The models used here are not epidemiological, in that they do not model the spread of the virus, or the sigmoid path that transmission takes. Instead, they are no more than extrapolations of the data, and so only useful for a few days or a week ahead. But, because there is no need to make assumptions about what will happen next, they may briefly provide more accurate forecast paths. In that sense they are more akin to weather forecasts, and in contrast to the epidemiological models which are more like climate models.

We report here on our first tentative round of forecasts. While forecasts can be made for all countries, US states, or Chinese provinces, we do not use spatial information: countries seem to be on quite different trajectories, and by now travelling between countries is restricted. The exception is in a few cases where countries seem to move closely together. The forecasts presented here have been made after two days of modelling, and further improvements are likely to be possible.

2 Data source

We use the data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering, This is currently updated daily and located at `github.com/CSSEGISandData/COVID-19`.

A dataset for modelling is created from this with minor adjustments. First, observations are put in columns with ISO date labels, and a few countries are renamed to be closer to their ISO name. Next, for France, Denmark, United Kingdom, and Netherlands we only include the mainland tallies. Aggregates for China, US, and EU-27 are constructed.

3 First results for confirmed cases

Forecasts of confirmed cases of COVID-19 and deaths from COVID-19 are obtained for a few countries only in the first instance. Data is available from 2020-01-22 to 2020-03-16. In many countries, the first confirmed case is later in the sample, e.g. 2020-01-31 in the UK. After some experimentation it was found that one specification of the model was most effective, combined with a choice of two forecasting methods. Details are given in the sections below.

In each graph we report the observed value in a grey line marked with dots. Forecasts are made from 9 March onwards from estimates up to 8 March. These are the red crosses in the graphs. Next, we make two weeks of daily forecasts from 17 March onwards: the red circles. The thin lines are 60% forecast intervals – these are very large and may not be reliable.

Figure 1 shows the forecasts of confirmed cases for the UK, EU-27, US and China. The UK and EU seem to be on similar trajectories. The UK had 1543 confirmed cases on 2020-03-16, and we forecast $10\,000$ around 25/26 March. The adopted mitigation policies may reduce this, which then will be reflected in updated forecasts.

The EU had almost $60\,000$ cases at the end of the sample, and is predicted to reach $250\,000$ around 25 March. But again, measures are in place now in an effort to reduce this.

The US here is on a more rapid growth path. This could be a reflection of the sparsity of testing for COVID-19. In that case, a few more days of data could improve (and possibly reduce) the forecasts.

In each case, the older forecasts are remarkably effective. Where they overlap, the

Figure 2 shows the forecasts of confirmed cases for Denmark on the left and the Netherlands on the right (mainland only for both countries). The surface area of Denmark and the Netherlands is almost identical, but the population of the Netherlands is three times higher. Correspondingly, we set the vertical scale of the graph for the Netherlands to three times that of Denmark. The trajectories are quite different, and, in Denmark's case, show a trend break. The Netherlands has one outlying observation.

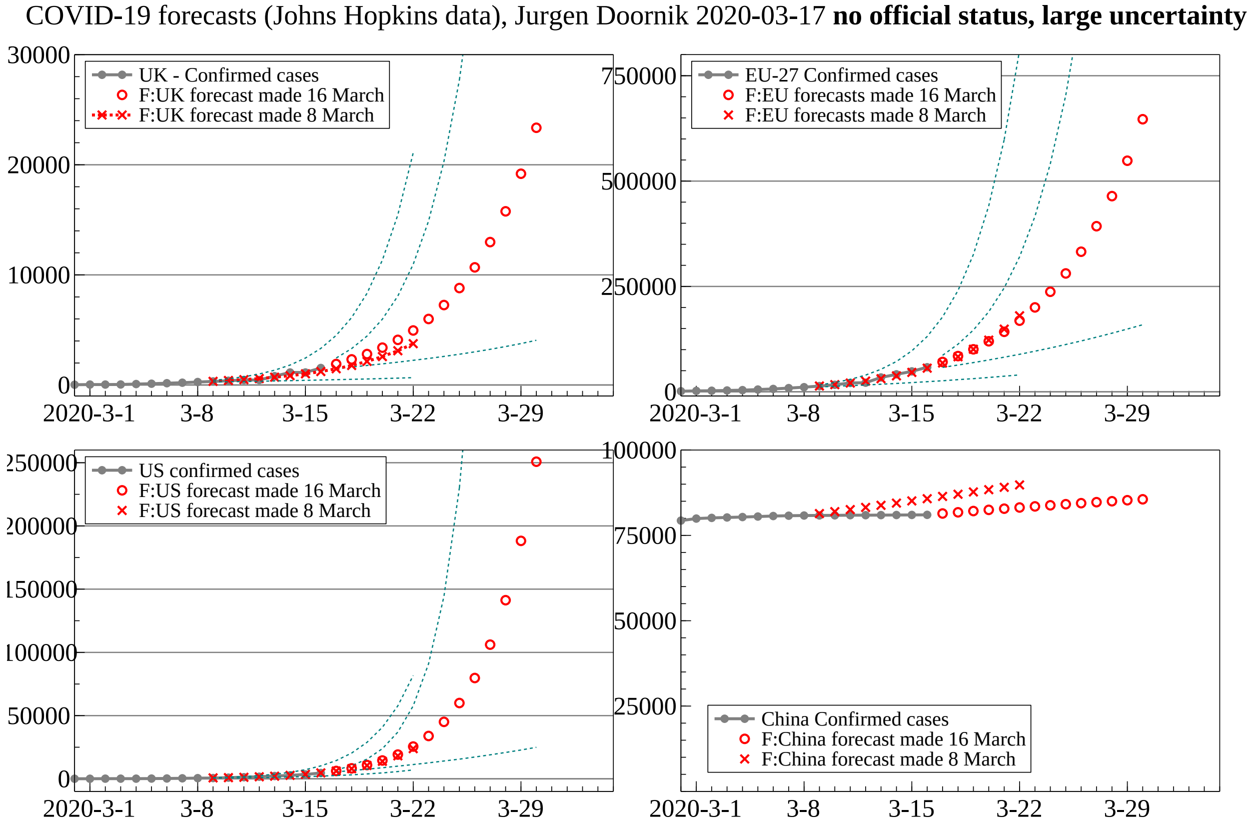

Figure 1: Forecasts of confirmed cases of COVID-19 for UK, EU-27, US, and China. Data from 2020-03-17.

4 Results for death counts

Figure 3 reports preliminary short-term forecasts for the death count. The observations on mortality are still limited, making forecasting more difficult. Italy is further ahead, and forecasts are more reliable in this case.

5 Methodology

We use local averaged time trend estimation (LATTE) to decompose the dependent variable $y_t, t = 1, ..., T$ in a *trend* term μ_t, and *residual* or irregular ε_t; seasonality is assumed to be absent. For the logarithmic model:

$$\log y_t = \widehat{\mu}_t + \widehat{\varepsilon}_t,$$

from which we obtain

$$y_t = \exp(\widehat{\mu}_t)\exp(\widehat{\varepsilon}_t).$$

To allow for the counts of zero before the virus took hold, we replace the specification with

$$\log(y_t + 1) = \widehat{\mu}_t + \widehat{\varepsilon}_t, \tag{1}$$

$$y_t \approx [\exp(\widehat{\mu}_t) - 1]\exp(\widehat{\varepsilon}_t), \tag{2}$$

with (2) not exactly following from exponentiation of (1).

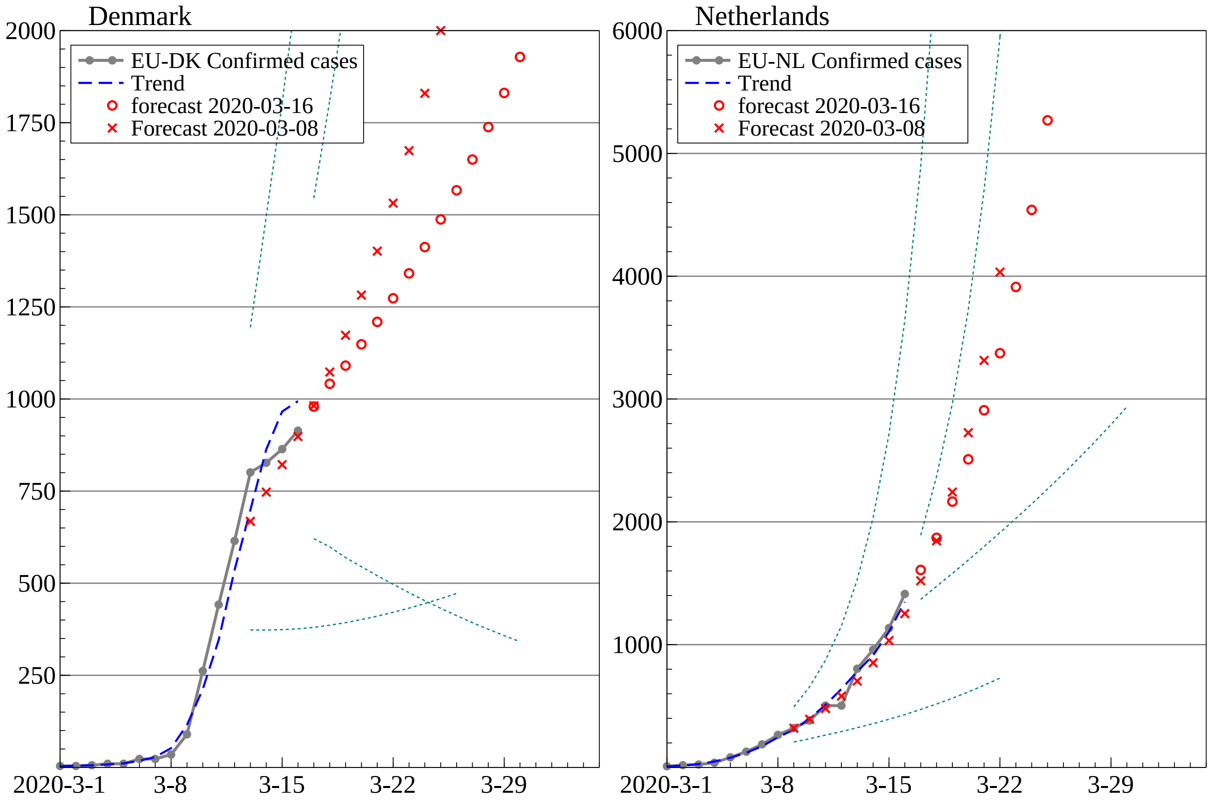

Figure 2: Forecasts of confirmed cases of COVID-19 for mainland Denmark and the Netherlands. Data from 2020-03-17.

The sample is split in overlapping windows, and for each window a trend indicator saturation (TIS) model is estimated: the model is saturated with linear trends and selection with *Autometrics* (Doornik, 2009) then obtains a sparse linear regression model. For a typical window w:

$$\widehat{x}_{w,t} = \widehat{\alpha}^F_w + \widehat{\beta}^F_w t + \sum_{s \in \mathcal{T}_w} \widehat{\theta}_{w,s}(t-s-1)I(t \le s), \quad t = T_w, ..., T_{w+1}-1. \tag{3}$$

The linear trend $(t-s-1)I(t \le s)$ starts at a negative value, then increases by unity until it hits zero, after which it stays at zero. The superscript F indicates that those terms are always included. In contrast to many other trend-cycle decomposition methods, this approach can handle abrupt breaks as well as smooth changes (although that seems less important in the current setting).

The LATTE estimates of the trend and residual is from W windows are:

$$\widehat{\mu}_t = W^{-1} \textstyle\sum_{w=1}^{W} \widehat{x}_{w,t},$$
$$\widehat{\varepsilon}_t = y_t - \widetilde{\mu}_t.$$

Model (3) is called L-TIS. The following two variants allow for a quadratic and cubic trend respectively:

$$\text{DL-TIS} \quad \Delta\widehat{x}_{w,t} = \widehat{\alpha}^F_w + \widehat{\beta}^F_w t + \sum_{s \in \mathcal{T}_w} \widehat{\theta}_{w,s}(t-s-1)I(t \le s), \tag{4}$$

$$\text{DDL-TIS} \quad \Delta\Delta\widehat{x}_{w,t} = \widehat{\alpha}^F_w + \widehat{\beta}^F_w t + \sum_{s \in \mathcal{T}_w} \widehat{\theta}_{w,s}(t-s-1)I(t \le s). \tag{5}$$

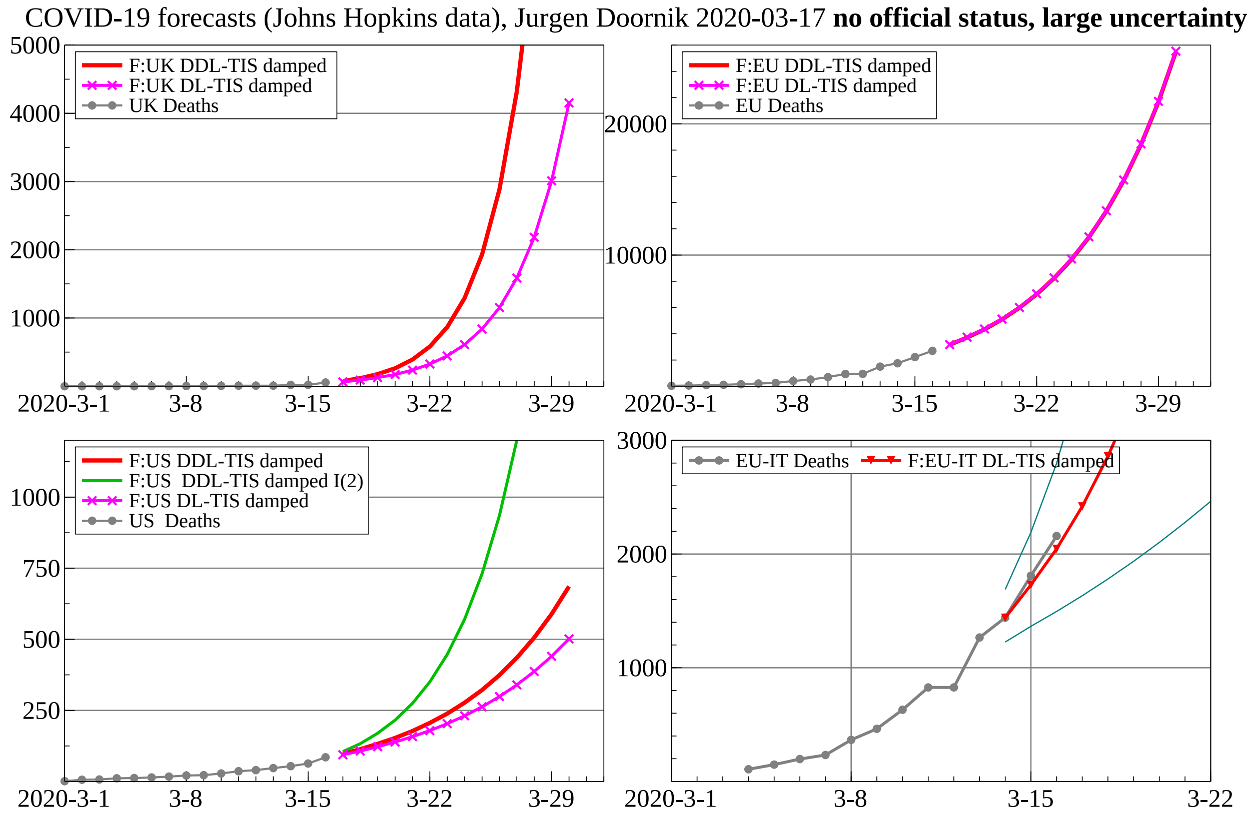

Figure 3: Forecasts of deaths from COVID-19 for UK, EU-27, US, and Italy. Data from 2020-03-17.

DL-TIS introduces cumulated differences in the trend, making $\widehat{\mu}_t$ an I(1) variable[1] with quadratic trend, while DDL-TIS makes $\widehat{\mu}_t$ I(2) with up to a cubic trend. For economic data it is common to restrict the model to a linear trend. More information on (LATTE) can be found in **?**).

The generality provided by the I(2) model with up to cubic trend can become a burden for forecasting: too much flexibility can lead to wild forecasts. The adopted forecasting device is Cardt (Castle, Doornik, and Hendry, 2019), which allows for up to I(1) with linear trend, making automatic decisions about the specific form. Cardt takes the average of three forecasting models, followed by calibration where the forecast are treated as pseudo-observed values. This performs very well on the data from the M4 and M3 forecast competitions (Makridakis, Spiliotis, and Assimakopoulos, 2020), but is too limited here, leading to three possible forecasts of the trend:

$$\widehat{\mu}^{(0)}_{T+h} = \text{Cardt}(h \mid \widehat{\mu}_1, ..., \widehat{\mu}_T) \qquad \text{[standard]} \tag{6}$$

$$\widehat{\mu}^{(1)}_{T+h} = \widehat{\mu}_T + \sum_{s=1}^{h} \text{Cardt}(s \mid \Delta\widehat{\mu}_2, ..., \Delta\widehat{\mu}_T) \qquad \text{[up to I(2)]} \tag{7}$$

$$\widehat{\mu}^{(2)}_{T+h} = (\widehat{\mu}^{(0)}_{T+h} + \widehat{\mu}^{(1)}_{T+h})/2 \qquad \text{[damped I(2)].} \tag{8}$$

Forecasts (6) are targeted at economic applications. Forecasts (7) apply Cardt to the differenced trend,

[1] An I(1) variable is stationary when differenced, an I(2) variable needs differencing twice for stationarity, see Johansen (1995), Doornik and Juselius (2018) *inter alia*.

Data	*Countries*	LATTE	*Forecasts starting*	
			2020-03-09	2020-03-17
Confirmed	UK,EU	DDL-TIS	damped I(2)	damped I(2)
Confirmed	US	DDL-TIS	damped I(2)	I(2)
Confirmed	China	DDL-TIS	damped I(2)	damped I(2)
Confirmed	DK	DDLX-TIS	standard	damped I(2)
Confirmed	NL	DDL-TIS	damped I(2)	damped I(2)
Deaths	UK,EU,US	DL-TIS, DDL-TIS		damped I(2)/I(2)
Deaths	IT	DL-TIS		damped I(2)

Table 1: Model and forecast specifications used in the reported results.

which is then reintegrated. The differences can then have a damped trend. Finally, (8) is the simple average of the previous two. This may seem ad hoc, but leads to an effective forecasting device.

Cardt forecasts are also made for the residual term. This is combined into the forecast, which for the damped I(2) version yields:

$$\widehat{y}_{T+h} = \left[\exp(\widehat{\mu}^{(2)}_{T+h}) - 1\right] \exp(\widehat{\varepsilon}_{T+h}).$$

6 Application

Several different specifications of the model and forecasting approach were tried. The adopted versions are in Table 1. When countries are listed together, a multivariate model DDLTIS model was estimated, but each forecast is made separately.

In most cases the DDL-TIS model with damped I(2) forecasts was adopted.

7 Extensions

Seeveral extensions could be considered to improve the quality of the forecast:

1. The data are cumulative count, so more congruent representation would enforce that the estimated trend has integer values and cannot decrease. This will have no noticeable impact on the forecasts.
2. The forecasts can be moderated with observed outcomes, or longer term forecasts, from countries or regions that are more advanced in the reduction of COVID-19.
3. If deaths are delayed from infection, so, if they are close to a fixed percentage of the number of cases, this can be used in forecasting.
4. Model and forecast specification is not yet fully automated. This would facilitate realtime forecasting (and forecast evaluation).

8 Conclusion

9 About the authors

References

Castle, J. L., J. A. Doornik, and D. F. Hendry (2019). Some forecasting principles from the M4 competition. Economics discussion paper 2019-w01, Nuffield College, University of Oxford.

Doornik, J. A. (2009). Autometrics. In J. L. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics: Festschrift in Honour of David F. Hendry*. Oxford: Oxford University Press.

Doornik, J. A., J. L. Castle, and D. F. Hendry (2020). Card forecasts for M4. *International Journal of Forecasting 36*, 129–134. doi.org/10.1016/j.ijforecast.2019.03.012.

Doornik, J. A. and K. Juselius (2018). *CATS 3: Cointegration Analysis of Time Series in OxMetrics*. London: Timberlake Consultants Press.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Makridakis, S., E. Spiliotis, and V. Assimakopoulos (2020). The m4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting 36*, 54–74. doi.org/10.1016/j.ijforecast.2019.04.014.