Time series estimation and forecasting of Covid in Norway

Gunnar Bårdsen   Ragnar Nymoen

Network meeting, FHI
3 November 2022

Introduction

- Covid time series are typical examples of changing data-generating processes, because of
  - mutations and policy responses
  - over time and across economies
- Likely that constant-parameter models will fail in forecasting, even if they are correct in a constant parameter world.
- First, a framework for forecasting new cases (incidence), hospital admissions and hospital beds is presented. This project, named CovidMod, produced 21 days ahead forecasts each working day from 17 March 2021 to 1 April 2022 (real time forecasts).

Introduction (cont.)

- Second, the model is extended to allow for endogenous effects of policy responses, like lockdowns and vaccinations, to counter the events of exogenous effects, like new mutations. The threshold of the target variable for policy reactions is estimated. The forecasting performance is compared to the forecasts of the Norwegian Institute of Public Health (NIPH) as well as to the real time forecasts of CovidMod.
- The approach and chosen method, based on forecasting methodology of macroeconomic time series, is to be considered as a complement to other forecasting approaches used.

A model based on autoregressive processes

- Builds on Nymoen (2022).
- Let Y_t denote new cases on day t,

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \epsilon_t \tag{1}$$

where $\alpha_i, i = 0, 1., ...p$ are parameters and ϵ_t is an error term which we assume linearly independent of $Y_{t-1}, ..., Y_{t-1-p}$. Equation (1) can be re-written as:

$$Y_t = \alpha_0 + \beta(\sum_{i=1}^{p-1} Y_{t-i}) + \sum_{i=1}^{p-1} \alpha_i^{\ddagger} \Delta Y_{t-i} + \epsilon_t \tag{2}$$

where

$$\beta = \frac{\sum_{i=1}^{p} \alpha_i}{p-1} \tag{3}$$

and $\alpha_i^{\ddagger}$ $(i = 1, 2, ..., p-1)$ are combinations of the original autoregressive coefficients .

A model based on autoregressive processes (cont.)

- The accumulated number of confirmed cases for day t is denoted S_{Yt}. It is given by the definition:

$$S_{Yt} = Y_t + S_{Yt-1}, \tag{4}$$

Equation (2) can therefore be expressed with the change $S_{Yt-1} - S_{Yt-p}$ on the right hand side:

$$Y_t = \alpha_0 + \beta(S_{Yt-1} - S_{Yt-p}) + \sum_{i=0}^{p-1} \alpha_i^{\ddagger} \Delta Y_{t-1-i} + \epsilon_t. \tag{5}$$

A working hypothesis could be that the amount of virus in the population, call it S_t^* (which is not observable) is increasing in the accumulated number of confirmed cases over a period of p days:

$$S_t^* = f(S_{Yt-1} - S_{Yt-p}),\ f' > 0, \tag{6}$$

Expression (5) and (4) highlights the positive feedback between between new cases and how much corona virus there is in the population.

The model contains four endogenous variables:

- Y_t, number of new infected with Covid-19, day t.
- Y_{St}, accumulated number of new infected with Covid-19, day t.
- HA_t, number of new hospitalisations wih Covid-19, day t.[1]
- HB_t, number of hospitalisations with Covid-19, day t.[2]

[1] https://www.fhi.no/sv/smittsomme-sykdommer/corona/dags--og-ukerapporter

[2] https://www.helsedirektoratet.no/statistikk/antall-innlagte-med-pavist-covid-19-for-nedlasting

A flow chart of the model

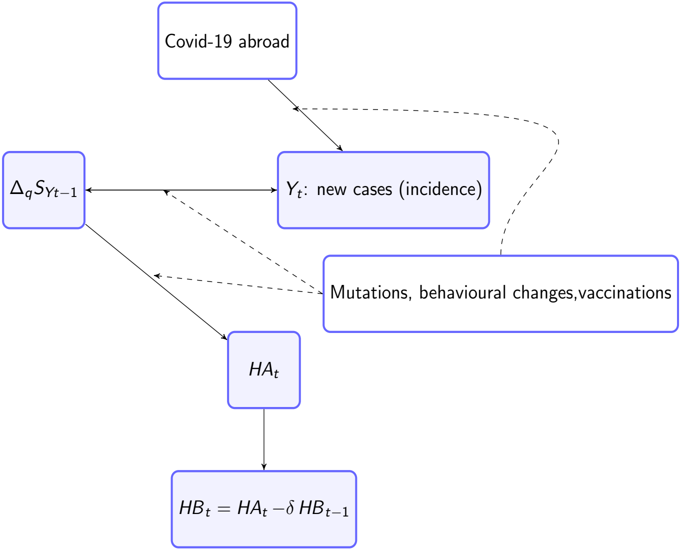


Figure: Relationships in CovidMod

Estimation of the model

Since March 2021 and until early in May 2022, forecasts were produced five times weekly (ie., work-days). During that period, abrupt and more gradual changes in the dynamics of the pandemic had been represented in the model by step-indicator variables (step dummies) and an by variables that measures the percentage of the population that have received vaccination (one of more doses). The step-indicator variables can be divided in two categories:

- Abrupt exogenous events, like mutations (D).
- Non-pharmaceutical policy responses (G).

Estimation of the model (cont.)

The estimated equation for new cases Y_t, using equation (5), was:

$$
\begin{aligned}
\widehat{Y}_t = & \underset{(61.9)}{0.072} (S_{Yt-1} - S_{Yt-14}) \\
& + \underset{(6.49)}{0.029} (S_{Yt-1} - S_{Yt-14}) D_{\alpha t} \\
& + \underset{(2.58)}{0.0014} (S_{Yt-1} - S_{Yt-14}) D_{\delta t} \\
& - \underset{(-5.23)}{0.013} (S_{Yt-1} - S_{Yt-14}) G_{Mar21t} \\
& + \underset{(9.94)}{0.017} (S_{Yt-1} - S_{Yt-14}) G_{Oct21t} \\
& - \underset{(-4.81)}{0.05} (S_{Yt-1} - S_{Yt-14}) VAC_t \\
& + \sum_{i=0}^{13} \hat{\alpha}_i^{\ddagger} \Delta Y_{t-1-i} + \text{other short term facors}
\end{aligned}
\tag{7}
$$

Estimation period: $2.21.2020 - 11.22.2021$, 641 obs.

Estimation of the model (cont.)

- $D_{\alpha t}$ and $D_{\delta t}$ are two step-dummies that are zero before the Alpha and Delta variants became dominant, and 1 after.
- G_{Mar21t} represents the renewed-lockdown in March 2021.
- G_{Oct21t} is a dummy for a partial opening-up of the society in October 2021.
- VAC_t represents higher vaccination rate: increasing in a continuous manner towards 1.

Estimation of the model (cont.)

The second estimated equation in the model is a conditional model of new admissions to hospital with Covid-19 (NH):

$$
\begin{aligned}
\widehat{HA}_t = {} & \underset{(4.50)}{0.276}\,(HA_{t-1}) + \underset{(5.64)}{0.248}\,Y_t \\
& + \underset{(3.77)}{0.00039}\,(S_{Yt-3} - S_{Yt-19}) \\
& + \underset{(6.15)}{0.00007}\,(S_{Yt-3} - S_{Yt-19})\,D_{\alpha t} \\
& + \underset{(6.41)}{0.000479}\,(S_{Yt-3} - S_{Yt-19})\,G_{Oct21t} \\
& - \underset{(-5.64)}{0.0009}\,(S_{Yt-1} - S_{Yt-14})\,VAC_{t-7} \\
& + \sum_{i=0}^{2} \hat{\gamma}_i \Delta NH_{t-1-i} + \text{other short term factors}
\end{aligned}
\tag{8}
$$

Estimation period: $7.14.2020 - 11.22.2021$, 497 obs.

Estimation of the model (cont.)

The third empirical equation in the model estimated on 22 November 2021 was the simple law of motion for the number of hospital beds with Covid-19:

$$\widehat{HB}_t = HA_t + \underset{(346.1)}{0.898} HB_{t-1} \tag{9}$$

$$\text{Estimation period:} 7.14.2020 - 11.22.2021, \; 497 \text{ obs.}$$

Simulation of the model, 1.7-20.11.21

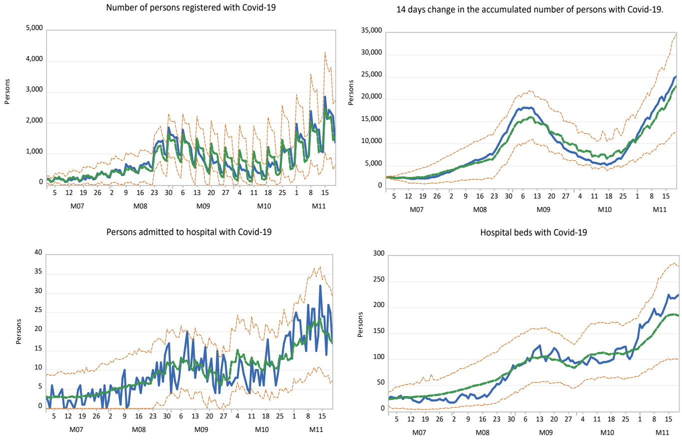


Figure: Line graphs represent actual values, dashed lines simulated values and dotted and lines represent upper and lower bounds of 90 % uncertainty intervals.

Forecast performance: New cases

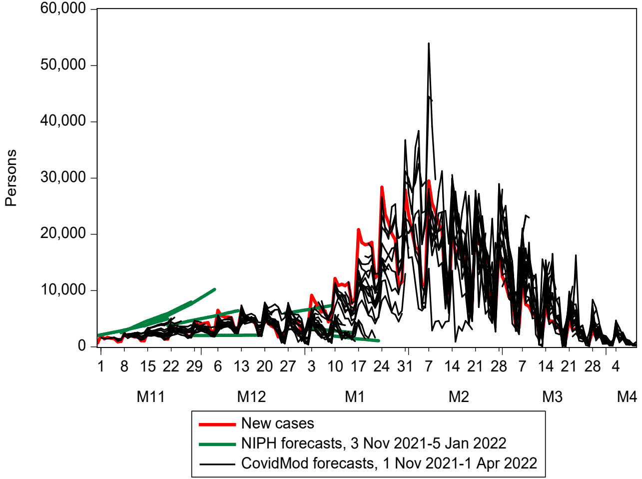


Figure: Forecasted number of new cases (daily data) and forecasts from CovidMod and NIPH (National regional model).

Forecast performance: Hospital beds

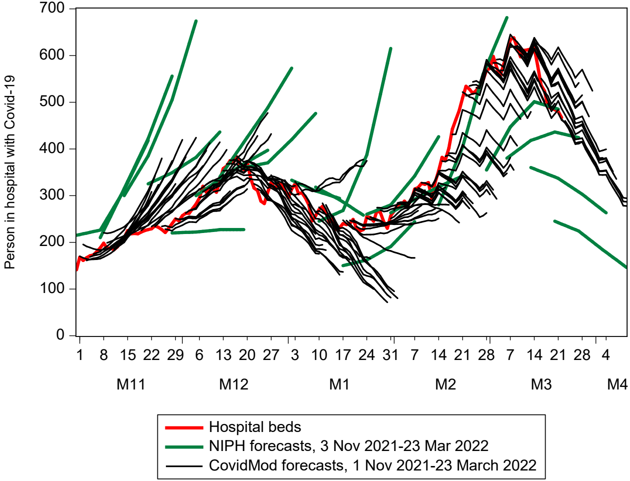


Figure: Forecasted number of hospital beds and forecasts from CovidMod and NIPH (National regional model).

A model with exogenous breaks and endogenous policy response regime shifts.

The general framework is to divide the breaks or the switching part of the model into exogenous regime changes D_t, like mutations, and effects of endogenous policy regime changes G_t, like lockdowns.
To illustrate, consider the following stylized model:

$$Y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t D_t + \beta_3 x_t G_t + u_t \quad (10)$$

An exogenous event, or shock, makes the binary variable D_t change the effect of x_t on Y_t from β_1 to $\beta_1 + \beta_2$.
To model the endogenous policy effects G_t, we use the standard Smooth Transition Model. See f. ex. van Dijk, Teräsvirta, and Franses (2002) for an overview. A transition function changes smoothly from 0 to 1 with an increasing policy target variable, here taken to be hospital beds HB_{t-k}

A model with exogenous breaks and endogenous policy response regime shifts. (cont.)

with a treshhold value of HB^*. As is common, the transition function is the logistic specification

$$G\left(HB_{t-k};\gamma,HB^*\right)=\frac{1}{1+\exp\left[-\gamma\left(HB_{t-k}-HB^*\right)\right]},$$

implying that:
If $HB_{t-k}=HB^*$, then

$$G\left(HB_{t-k}-HB^*\right)=0.5.$$

If $\left(HB_{t-k}-HB^*\right)\rightarrow-\infty$

$$G\left(HB_{t-k}-HB^*\right)\rightarrow 0,$$

A model with exogenous breaks and endogenous policy response regime shifts. (cont.)

and with $(HB_{t-k} - HB^*) \to \infty$

$$G\left(HB_{t-k} - HB^*\right) \to 1.$$

The steepness parameter of the transition function γ and the threshold value HB^* are both estimated.
With a very high γ, the effects of switching between zero or full policy effects can approximate a binary variable.

Modelling new infections $\boldsymbol{Y_t}$

The equation for Y_t is estimated as:

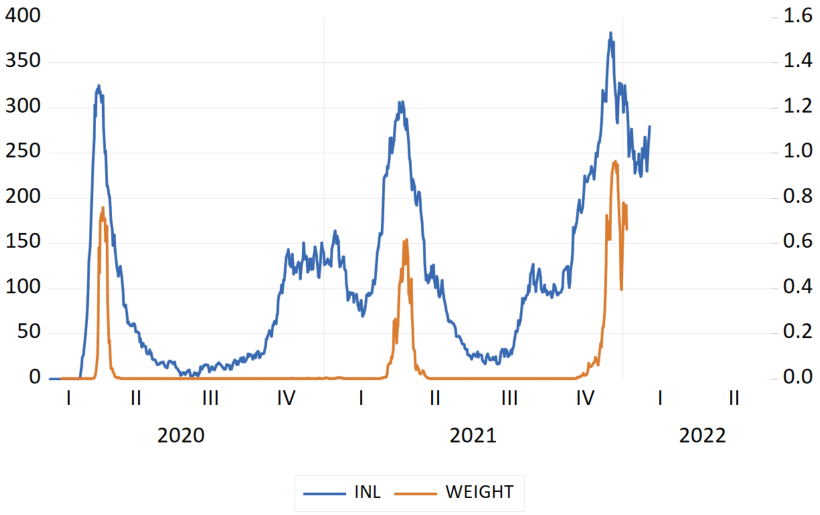
$$\begin{aligned}
\hat{Y}_t &= \underset{(0.003)}{0.060}\,(S_{Yt-1} - S_{Yt-14}) \\
&+ \underset{(0.0009)}{0.007}\,(S_{Yt-1} - S_{Yt-14})\,D_t \\
&- \underset{(0.003)}{0.007}\,(S_{Yt-1} - S_{Yt-14})\,G_t \\
&+ \text{lagged}(\Delta Y_{t-j}) + \text{ residual} \\
T &= 15.2.2020 - 5.1.2022,\ 691 \text{ obs.}
\end{aligned} \tag{11}$$

where

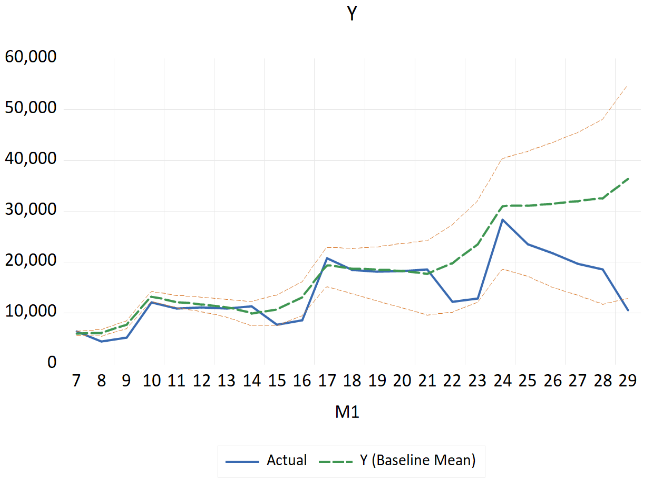
$$D_t = f(Tyrol, Alpha, Delta, Omicron)$$

$$G_t = \frac{1}{1 + \exp\left[-\underset{(0.028)}{0.038} \left(HB_{t-5} - \underset{(22.58)}{294} \right) \right]}.$$

Hospitalisations *HB* and smooth transition function *G*



The forecasts and realizations of Y

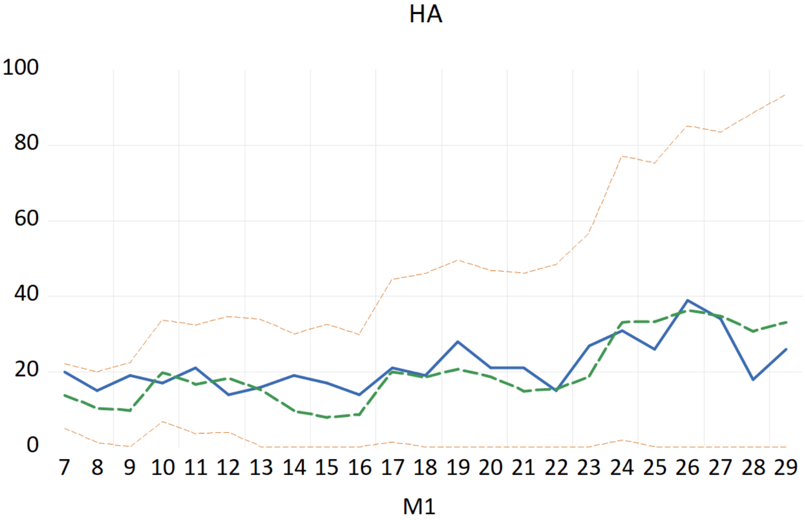


The HA and HB equations

$$
\begin{aligned}
\widehat{HA}_t &= -\underset{(0.0001)}{0.0006}\,(S_{Yt-3} - S_{Yt-9})\,Omicron_{t-6} \\
&\quad + \underset{(0.001)}{0.012}\,Y_t \\
&\quad - \underset{(0.0009)}{0.007}\,Y_t \times NINLDN_t \qquad (12) \\
&\quad - \underset{(0.0004)}{0.002}\,Y_t \times NINLDNnov21_t \\
&\quad + \text{lagged } HA_{t-j}) + \text{ residual} \\
\widehat{HB}_t &= HA_t + \underset{(0.002}{0.901}\,HB_{t-1} \qquad (13) \\
T &= 14.7.2020 - 5.1.2022,\ \text{541 obs.}
\end{aligned}
$$

The forecasts and realizations of *HA*

Forecast intervals including parameter uncertainty



The forecasts and realizations with parameter uncertainty

The forecasts and realizations without parameter uncertainty

Comparisons of forecasts

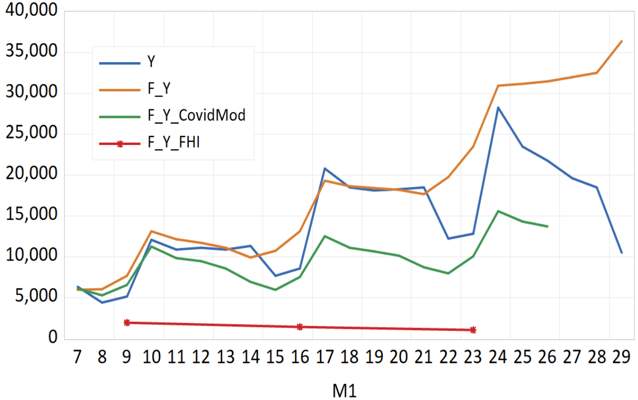


Figure: Incidents. Source: "Situational awareness and forecasting for Norway".FHI COVID-19 modelling team Week 1, 5 January 2022. Table 2. Linear interpolation between 7, 14 and 21 days ahead forecasts from forecast origin date.

Comparisons of forecasts (cont.)

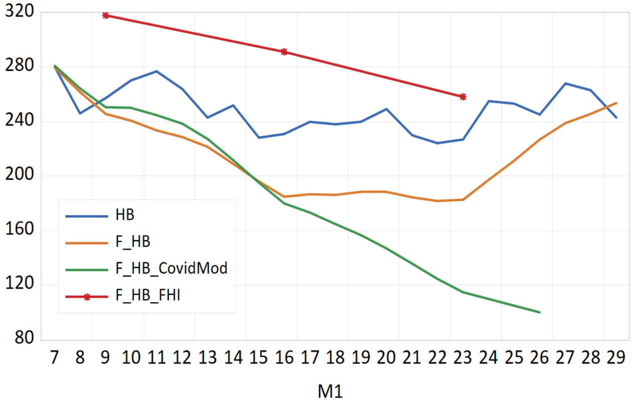


Figure: Hospitalisations. Source: "Situational awareness and forecasting for Norway".FHI COVID-19 modelling team Week 1, 5 January 2022. Table 2 and Figure 5. Linear interpolation between 7, 14 and 21 days ahead forecasts from forecast origin date.

Conclusions

- Likely that constant-parameter models will fail in forecasting, even if they are correct in a constant parameter world.
- A framework for forecasting new cases (incidence), hospital admissions and hospital beds is presented.
- The model is extended to allow for endogenous effects of policy responses.
- The approach and chosen method, based on forecasting methodology of macroeconomic time series, is to be considered as a complement to other forecasting approaches used.

References

Nymoen, R. (2022).
Dynamisk modellering og framskrivning av nye smittede og innlagte med Covid-19 i Norge.
Samfunnsøkonomen (1), 5–13.

van Dijk, D., T. Teräsvirta, and P. H. Franses (2002).
Smooth transition autoregressive models — a survey of recent developments.
Econometric Reviews 21(1), 1–47.