

Analysis of Covid Cases in North Dakota and the US

Kevin Thompson

Southern Methodist University Data Science Program

Abstract

The SARS-coronavirus-2 pandemic ravages the United States both economically, socially, and physically. Rapid, tailored responses to regional spikes in cases depend on our ability to forecast the future state of the pandemic. The inherently explosive and non-stationary behavior of a pandemic presents its own modelling challenges along with the many sources of measurement error that come from both individual and structural faults.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: covid-19, SARS-coronavirus-2, time series analysis

Related Work

Predicting the behavior of viruses from the outbreak stage to the pandemic stage has been an active field of study since the SIR model was introduced in Kermack and McKendrick (1927). The standard SIR begins with the modeling assumptions that viruses behave as lifeforms do (i.e. birth, life, reproduction, death) and that all members of a fixed population fall into the trichotomy expressed in the name of the model. At every point in time, each individual in the model is either susceptible, infectious, or recovered (SIR). Moreover, there is a fixed transmission and recovery rate. The SIR also assumes that the

This project was completed in fulfillment of an assignment for a time series course at Southern Methodist University, though I plan to expand the scope of this analysis beyond the project requirements in the near future.

compartments in the trichotomy are ordered and that members of the population can only move in one direction. If an individual goes from infectious to recovered, they cannot be infected again. Along with the above assumptions, the standard SIR model is a system of three non-linear ordinary differential equations:

$$\frac{dS}{dt} = -\beta \frac{IS}{N} \quad (1)$$

$$\frac{dI}{dt} = \beta \frac{IS}{N} - \gamma I \quad (2)$$

$$\frac{dR}{dt} = \gamma I \quad (3)$$

where S denotes the susceptible subpopulation size, I the infected subpopulation size, and R the recovered subpopulation size (Nadler, Wang, Yang, & Guo, 2020). γ and β are the unknown recover rate and transmission rate parameters, respectively.

Methods

Data

The models in this paper are fit using data from the Johns Hopkins University Coronavirus Research Center.

Univariate Models

Multivariate Models

SIR Model

No Modelling of Residuals.

With Modelling of Residuals. Given that the conventional time series methods like ARUMA models assume that time is discrete, we will have to discretize equations (1)-(3). Thankfully, this just leads to first differences because the differential equations are only first order:

$$\frac{dS}{dt} \approx S_{t+1} - S_t \quad (4)$$

$$\frac{dI}{dt} \approx I_{t+1} - I_t \quad (5)$$

$$\frac{dR}{dt} \approx R_{t+1} - I_t \quad (6)$$

Following Nadler et al. (2020) I will label S as the total population and R will denote the number of confirmed cases.

Participants

Material

Procedure

Data analysis

We use Nadler et al. (2020).

Results

Discussion

References

- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A*, 115(772), 700–721. <https://doi.org/10.1098/rspa.1927.00118>
- Nadler, P., Wang, S., Yang, X., & Guo, Y. (2020). An epidemiological modelling approach for covid-19 via data assimilation. *European Journal of Epidemiology*, 35(8), 749–761. <https://doi.org/10.1007/s10654-020-00676-7>