COVIDPostiveTestingSA

Belay Birlie Yimer, . . . , Ziv Shkedy

02/03/2021

Introduction

- Individuals rely on test results to guide their medical treatment and decisions on whether to self-isolate.
- Public health officials rely on the results to track the state of the pandemic, and policymakers use this information to guide decisions on reopening schools and businesses.
- One number—the "percent positive"—is often cited in these decisions.
- The percent positive is the percentage of all coronavirus tests performed that are actually positive, or: (positive tests)/(total tests) x 100%.
- The percent positive helps public health officials answer questions such as:
  - What is the current level of SARS-CoV-2 (coronavirus) transmission in the community?
  - Are we doing enough testing for the amount of people who are getting infected?

Methods

Data

We downloaded data from Coronavirus COVID-19 (2019-nCoV) Data Repository for South Africa maintained by Data Science for Social Impact research group at the University of Pretoria [ref]. The data repository captures the daily number of new cases, number of tests, number of deaths and recoveries. Our primary outcome of interest was the daily number of newly diagnosed COVID-19 cases and the unit of time used in modelling was a day. We used the daily case reports from March 12, 2020, until February 27, 2021, in our analysis.

Statistical analysis

We assume the number of postive cases to have a binomial distribution. Let π_t denote the daily postive testing rate per test, Y_t be the daily number of confirmed COVID-19 cases out of the n_t indiviudals tested at day t. Our aim is to model the probability π_t over time and produce a model-based estimate for it's first-order derivative. The model for the daily poitive testing rate is then given by

$$\begin{aligned} Y_t \sim & Binomial(n_t, \pi_t), \quad t = 1, ..., T, \\ logit(\pi_t) = & f(t). \end{aligned}$$

The smooth function $f(t)$ is represtend by basis expansion of rank K. That is,

$$f(t) = \sum_{i=1}^{K} \beta_i b_i(t)$$

where $b_i(t)$s are set of thin plate spline basis functions and β_is are unknown coefficients of the basis functions. The spline model parameters can be estimated by considering the goodness-of-fit and the degree of smoothness. Writing all the basis coefficents in one parameter vector $\boldsymbol{\beta}$,

$$l(\beta) - \frac{\lambda}{2}\sum_{j=1}^{M}\beta^T S\beta$$

where S is metrices of known coefficents and λ is the smoothing parameter to be estimated.

Results

Figure 1 presents the daily number of reported COVID-19 cases from 12 March 2020 to 27 February 2021. Similar to elsewhere in the world, South Africa pass through a two-wave pandemic. The pandemic's first peak was on 07 July 2020, where up to 13944 new COVID-19 cases reported, followed by a second peak in January 2021, where more than 21,000 daily cases reported. Figure 2 presents the cumulative number of new reported COVID-19 cases and tests performed. To date, 8,838,937 tests have been conducted, and a total of 1,500,677 cases reported.

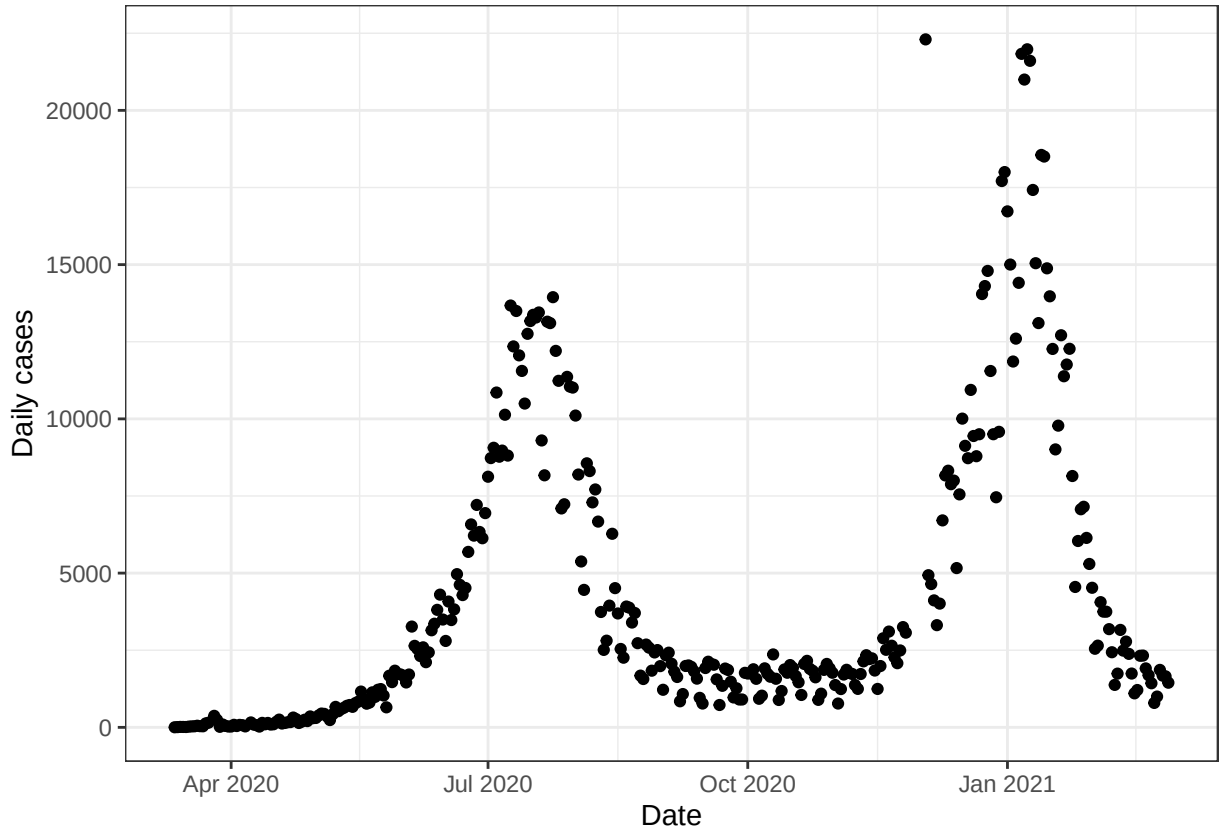

Figure 1: Daily number of COVID-19 cases in South Africa from 12/03/2020-27/02/2021.

```
## [1] 345   6
## Warning: Ignoring unknown parameters: ize
## Saving 6.5 x 8 in image
```

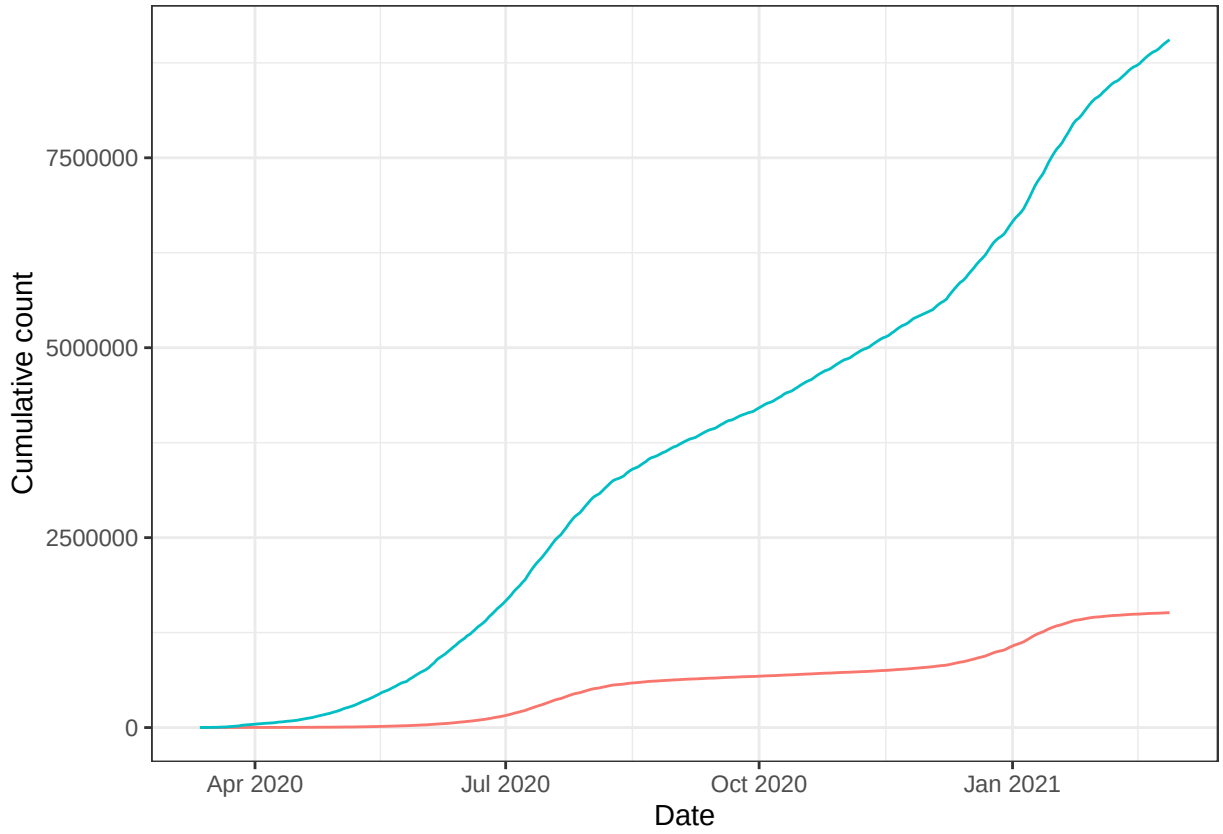

Figure 2: The cummulative number of COVID-19 cases and Cummulative number of tests in South Africa from 12/03/2020-27/02/2021. Red-line denote the number of cases and blue-line denotes the number of tests.

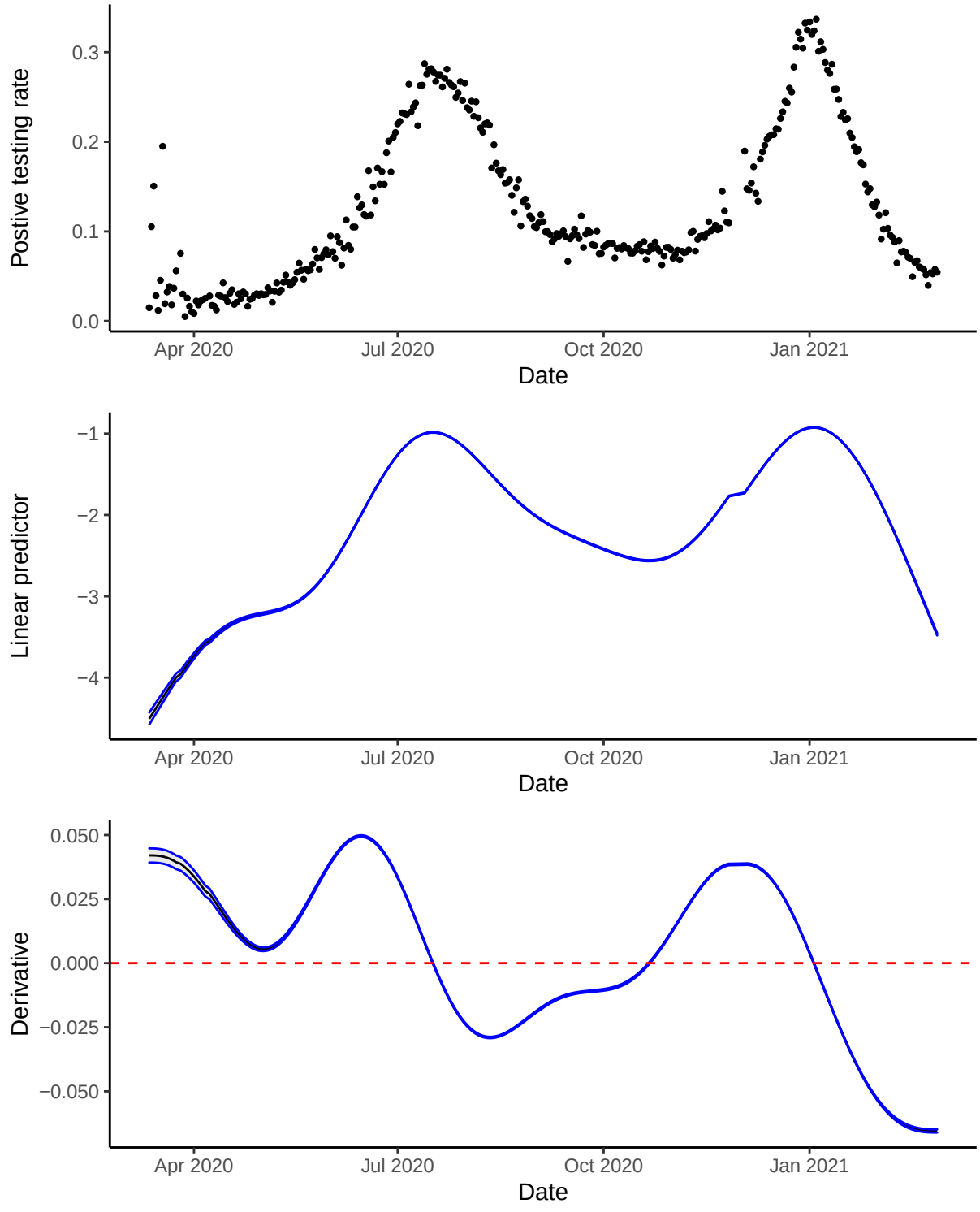

Figure 3: Upper Panel: Observed postive testing rate. Middle panel: The linear predictor of the smoother with 95% creadible interval. Lower panel: The derivative of the estimated probability of postive testing with the 95% creadible interval.