A

CODECHECK certificate 2020-012

https://doi.org/10.5281/zenodo.3893617

B

Item	Value
Title	Report 23: State-level tracking of COVID-19 in the United States version 2 (28-05-2020)
Authors	H Juliette T Unwin, Swapnil Mistra, Valerie C Bradley, et al.
Reference	https://dx.doi.org/10.25561/79231
Codechecker	Stephen J. Eglen
Date of check	2020-06-14 14:00:00
Summary	R code for this paper shared with an earlier codecheck certifice (2020-011) from the same codebase.
Repository	https://github.com/sje30/covid19model-report23

Table 1: CODECHECK summary

C

Summary

The key findings in the "Report 23" from Imperial College were reproducible. I was able to re-run their code and generate qualitatively similar results to those shown in their manuscript. Differences in absolute values in results are due to the stochastic nature of the analysis. All code to reproduce the data worked as expected, and all key datasets were provided. I was able to regenrate the results in Figures 4–8 of the manuscript; code for Figures 1–3 was not available. (I did not attempt to go through all of the figures in the appendix, although Appendix D is an expanded version of Figure 6, showing summaries of each state.) The only significant complication in this reproduction was that some of the figures required the installation of system libraries. The final computations took about 17 hours on a multicore workstation.

In some cases, figures directly matched the layout in the manuscript; however, sometimes the figures have been post-processed as there are differences in layout. For example, in Figure 4 of the manuscript, the states have been re-ordered vertically in order of the value of R. Likewise, in Figure 8, the plots have been expanded out over three columns.

D

Output	Comment	Size (b)
usa/figures/rt_point__1006697.pdf	Manuscript Figure 4	10841
usa/figures/1006697_rt_map_chloropleth.pdf	Manuscript Figure 5	77748
usa/figures/WA_three_panel_1006697_.pdf	Manuscript Figure 6 (Washington)	15409
usa/figures/NY_three_panel_1006697_.pdf	Manuscript Figure 6 (New York)	14703
usa/figures/MA_three_panel_1006697_.pdf	Manuscript Figure 6 (Massachusetts)	14472
usa/figures/FL_three_panel_1006697_.pdf	Manuscript Figure 6 (Florida)	14642
usa/figures/CA_three_panel_1006697_.pdf	Manuscript Figure 6 (California)	14798
usa/figures/1006697_infectiousness_regions.pdf	Manuscript Figure 7	38005
usa/figures/WA_scenarios_56_0_20_40_1006697_deaths.pdf	Manuscript Figure 8 (Washington)	10032
usa/figures/NY_scenarios_56_0_20_40_1006697_deaths.pdf	Manuscript Figure 8 (New York)	10582
usa/figures/MA_scenarios_56_0_20_40_1006697_deaths.pdf	Manuscript Figure 8 (Massachusetts)	10432
usa/figures/FL_scenarios_56_0_20_40_1006697_deaths.pdf	Manuscript Figure 8 (Florida)	10039
usa/figures/CA_scenarios_56_0_20_40_1006697_deaths.pdf	Manuscript Figure 8 (California)	10052

Table 2: Summary of output files generated

E

CODECHECKER notes

The github repository https://github.com/ImperialCollegeLondon/covid19model was cloned, and renamed to "sje30/covid19model-report23". (I could not clone the project into the Github codecheckers group, as you cannot have two forks of the same project in the same organistion.)

This reproduction was performed after finishing the related certificate 2020-011; details of setting up the R environment are described in that certificate.

However, the R environment described was insufficient, as it didn't include *geofacet* and *rgdal* packages which needed system libraries to install. Once the sysadmin had installed extra libraries for unitdevs2 and gdal, I needed to run the following adhoc module provided locally:

```
module load ./gdal-2.1.2
```

```
install.packages("rgdal")
install.packages("geofacet")
install.packages("denstrip") #for plotting
```

An initial run of the FULL model didn't work because I had an older version of rstan package; this was upgraded to 2.19.3. The simulations were tested by running the simulation directly on a workstation:

```
time Rscript base-usa.r
```

Running the test mode took 41 minutes and generated outputs.

```
time Rscript base-usa.r -F
```

The final run time was 1020 minutes (17 hours). The code for reproducing figures 1,2 and 3 was not available in the repository, but all other key figures could be regenerated.

F

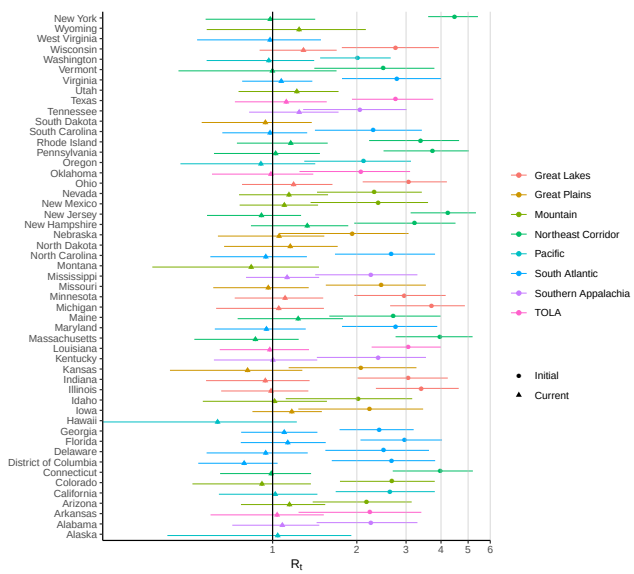

Figure C1: Manuscript Figure 4