

Modelling COVID-19 in South Africa at a provincial level

By Louis Rossouw

Presented at the Actuarial Society of South Africa's 2020 Virtual Convention
6–8 October 2020

ABSTRACT

This paper documents a model of the COVID-19 epidemic in South Africa. Mobility data are used to model the reproduction number of the COVID-19 epidemic over time using a Bayesian hierarchical model. This is achieved by adapting the work by Imperial College London researchers for South Africa. In this case the model is calibrated to excess deaths and not case counts or reported COVID-19 deaths. The model uses mobility movement indexes by province produced by Google. Furthermore the model includes a fixed effect for interventions introduced at the start of level 4 lockdown. The model and report are automatically generated on a regular basis using R. This version contains data available on 16 September 2020; however the model and report is updated regularly.

KEYWORDS

COVID-19, SARS-CoV-2, model, mobility, reproduction number, Bayesian, automation

CONTACT DETAILS

Louis Rossouw, Head of Research & Analytics, Gen Re | Life/Health Canada, South Africa, Australia, NZ, UK & Ireland; Email: LRossouw@GenRe.com; Mobile: +27 (0)71 355 2550

1. INTRODUCTION

This paper documents a model of the COVID-19 epidemic in South Africa. Mobility data are used to model the reproduction number of the COVID-19 epidemic over time using a Bayesian hierarchical model. This is achieved by adapting the work by Imperial College London researchers [1] for South Africa. Authors from Imperial College London have built similar models for Brazil [2] and the United States [3]. In this case the model is calibrated to excess deaths [4] and not case counts or reported COVID-19 deaths.

The model uses mobility movement indexes by province produced by Google [5]. Furthermore the model includes a fixed effect for interventions introduced at the start of level 4 lockdown.

The model and report are automatically generated on a regular basis using R [6]. This version contains data available on 16 September 2020.

An online, regularly updated, version of this report is available [here](#).

2. UPDATES

As this paper is updated over time this section will summarise significant changes. The code producing this paper is tracked using git. The git commit hash for this project at the time of generating this paper was dc7cd1a548142bbaf1d43461ab60821974105638 and this change was made on 2020-09-19.

2020-05-31	Document is made available online (pending review).
2020-06-01	Added backtesting.
2020-06-04	Incorporated feedback from reviewer 1.
2020-06-08	- Incorporated feedback from reviewers 2 and 3. - Added a projection for South Africa (based on sum of the provinces). - Created a CSV with the daily projections. See the projection section for details. - Added an updates section to track changes to this model.
2020-06-09	- Incorporate feedback from reviewer 4. - Remove the under review designation. As at this date the model has feedback from four reviewers and also comments from others. The author will continue to update this model and seek review as needed. - Due to mobility increasing from level 3, additional projection scenarios have been introduced involving decreased mobility (back to level 4 levels). - Numerous updates on plots including some residual plots.
2020-06-10	Remove parks index from average mobility index in line with [7]. Will drop residential index soon too.
2020-06-20	- Update IFRs for HIV. Possibly need further increases based on research coming out in near future. - Remove residential mobility, such that there is only one mobility parameter left.
2020-06-22	Add parameter for fixed effects correlated with the start of level 4 lockdown aimed at capturing the effect of mask wearing and workplace screening
2020-07-04	- Updated standard deviation of prior for β to reflect [8]. This allows for a greater impact of mask wearing on $R_{t,m}$. β_m^s prior assumption updated to allow increases and decreases in effect of interventions implemented at start of level 4 lockdown

2020-07-07	- Major changes to allow for excess deaths exceeding reported COVID-19 deaths. - Expanded calibration graphs for more provinces.
2020-07-16	- Update excess reporting based on [9] data. - Increase uncertainty around completeness of reporting factors.
2020-07-26	- Remove 14-day backtesting to reduce total run time. - Further update to excess deaths calculation to give reporting data an additional week of run-off.
2020-08-19	- Incorporate autoregressive term in modelling of $R_{t,m}$ to improve fit using approach in [3]. - Switch to using excess deaths as the outcome to be modelled rather than reported deaths. - Remove backtesting to improve run time. Will add it back once model is stable.
2020-08-27	Update to using the latest excess deaths from [10]. This has resulted in a significant lowering of attack rates, and future projections. Added some comments to this effect.
2020-09-04	Update excess deaths to [11].
2020-09-05	Update percentage of excess deaths (ψ) distribution to narrow the variance.
2020-09-09	Update excess deaths to [12].
2020-09-16	Update excess deaths to [4].

3. METHODOLOGY

A detailed description of the methodology and assumptions is provided below. The key features of the approach employed are summarised here:

- It uses a Bayesian hierarchical model to calibrate model parameters based on observed excess death data and prior assumptions.
- Changes in the reproduction number are linked to mobility data [5], the introduction of other non-pharmaceutical interventions (NPIs), as well as an autoregressive term.
- The reproduction number over time is used to model the number of infections occurring over time.
- The model uses population-weighted infection fatality ratios (IFRs) by province with some noise added as well as assumptions about the time it takes to die to model deaths from infections.
- There is a single combined model for all provinces. This model shares information between provinces, but province-specific effects are also allowed.
- The model is fitted such that it uses the prior assumptions about various parameters as well as the excess death data in such a way to provide an updated parameter distribution and outcome distribution that reflect both the prior assumptions as well as the excess death data.

4. MOBILITY

4.1 National mobility data

The national mobility data from [5] is plotted below in Figure 1. The model uses the indexes at a provincial level but here the national indexes are plotted for convenience. Clear trends are observable:

- Mobility generally reduces before the lockdown on 27 March.

- There is an increase in mobility in the days just before the lockdown. In particular the Grocery & Pharmacy and Retail indexes. Perhaps an indication of pre-lockdown “panic buying”?
- Mobility is relatively stable during level 5 lockdown at low levels.
- Mobility increases when level 4 starts.

Figure 2 summarises the average mobility index used in the model. This is the average of mobility indicators excluding parks and residential. This follows [7].

There are some risks of the mobility data above as it may be biased to the Android operating system, people with smartphones and people with enough data on their smartphone to share their location.

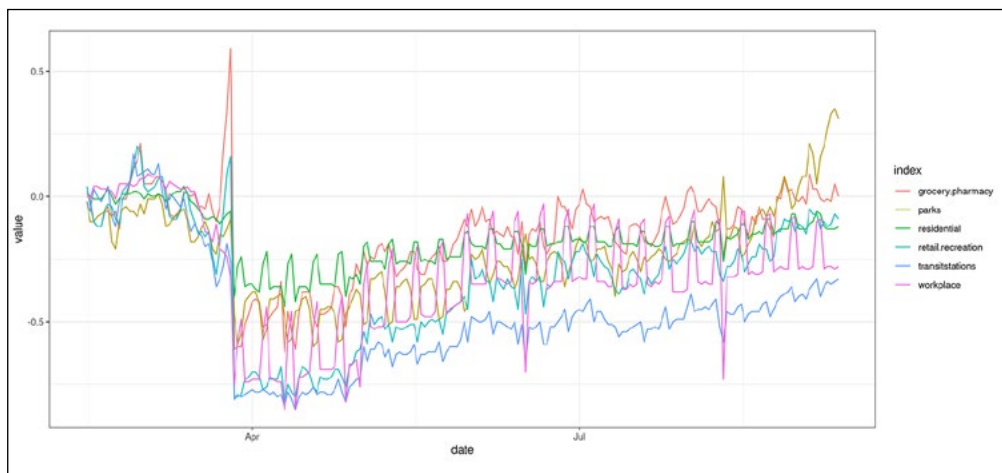


FIGURE 1 Google mobility indexes for South Africa

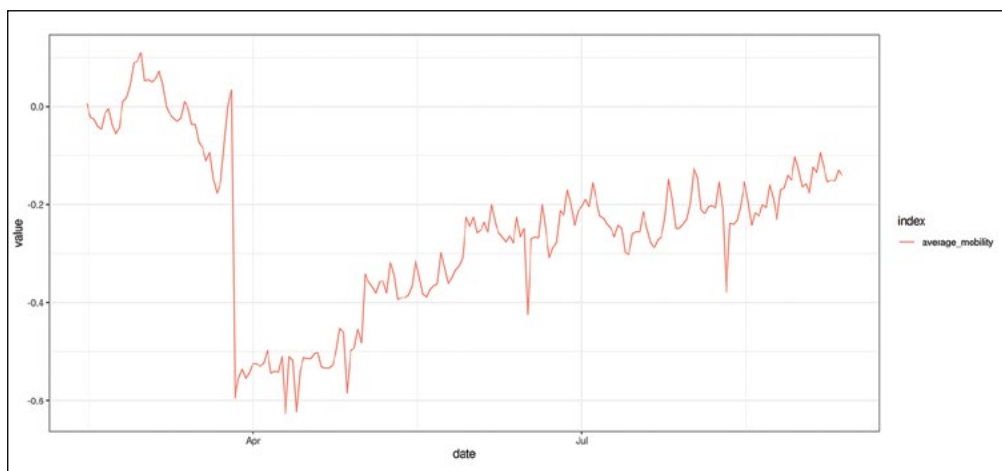


FIGURE 2 Average mobility

This bias does not really represent a problem though unless it changes over time because then the model would not be able to produce accurate projections. There is a risk that mobility changes disproportionately between users contributing data to these indexes and those not. It would not seem reasonable to expect a major change in the bias over time (unless the calculation is somehow compromised). Provincial biases may exist, but the model also allows the provinces to exhibit provincial-specific mobility effects for both parameters.

4.2 Provincial mobility data

The average mobility (excluding residential and park indexes) is plotted in Figure 3 for the various provinces. From this plot it seems apparent that Western Cape mobility has reduced the most during lockdown and is, perhaps, increasing slower.

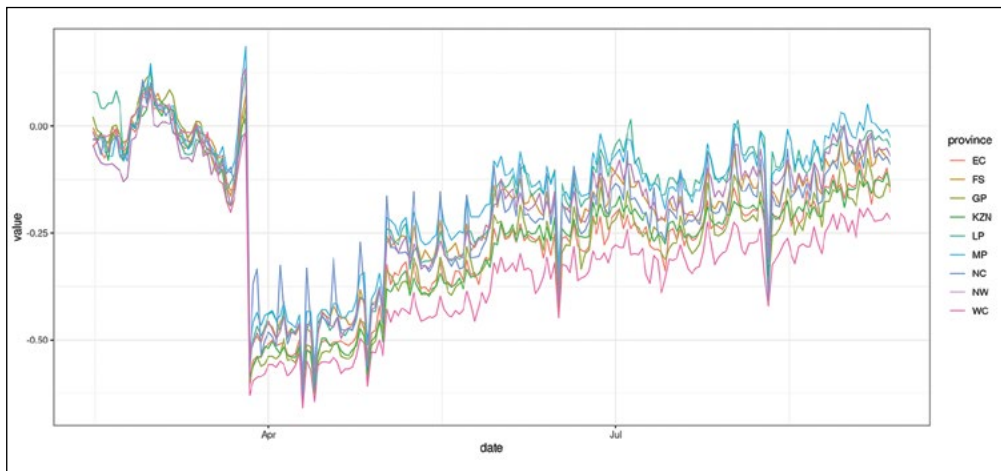


FIGURE 3 Average mobility by province

5. CALIBRATION

The sections below show how well the model is reproducing past excess death data. As the model is fitted to past death data we would want to see how well it is reproducing such data. Each section covers a province.

Three panels are plotted for each province:

1. The first panel shows modelled *weekly* deaths (blue) and 90% of excess deaths (brown). Bars are plotted at the start of the week (which coincides with [4]). A good model would produce death forecasts such that the actual reported past deaths would appear in the confidence intervals of the prediction.
2. The second shows the daily count of deaths reported in the province (brown) and the estimated deaths in blue.
3. The third panel shows the modelled daily number of infections (blue) compared to confirmed case counts (brown) as reported for the province. Note that this model does not calibrate to case data, but this data is shown for reference. We expect the model

- to produce infections exceeding confirmed cases as testing is likely only finding a fraction of the true infections.
4. The fourth panel shows the estimates for reproduction number ($R_{t,m}$) over time (t) for each province (m). These values are generating the infections in the model which result in deaths. Note that due to the delay of deaths following infections the last week or two of $R_{t,m}$ values already present an extrapolation from the data. Deaths as a result of these recent infections have generally not occurred as yet.
 5. The final panel shows the estimated value of $-\varepsilon_{w_m(t),m}$. This represents the unexplained component. A positive value indicates that the $R_{t,m}$ needed to be increased over and above the value implied by other parameters. A negative value indicates the reverse.

In all the charts the darker shaded area represents a confidence interval of 50% and the lighter shaded area represents a confidence interval of 95%.

5.1 Eastern Cape

Figure 4 shows 90% of excess deaths per [4] and the modelled deaths. For the Eastern Cape the fit appears reasonable. Also to be noted is the levelling off of excess deaths in the last week.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 5. It appears that reported deaths and modelled deaths are moving closer together, possibly because completeness of reporting is improving.

Modelled infections are plotted compared to confirmed cases in Figure 6. Over the last 14 days it would appear that 1.8% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 7. The $R_{t,m}$ started low but increased significantly during lockdown peaking at the start of June and declining afterwards.

The $-\varepsilon_{w_m(t),m}$ is plotted in Figure 8 on page 8. This parameter has needed to be negative to enable the $R_{t,m}$ to remain at low levels.

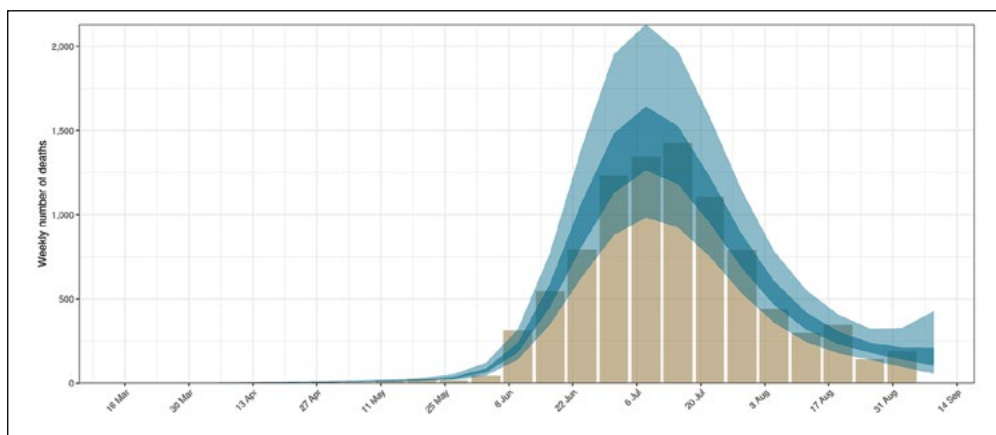


FIGURE 4 Projected deaths and 90% of excess deaths in Eastern Cape

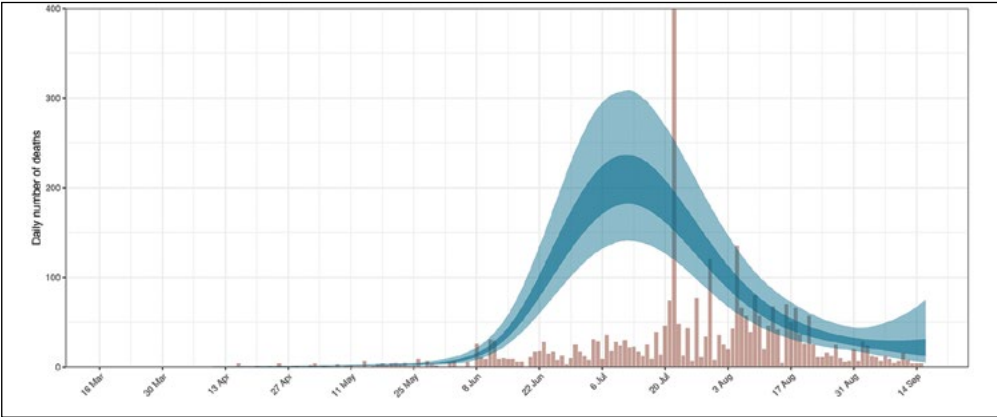


FIGURE 5 Projected deaths compared to reported COVID-19 deaths in Eastern Cape

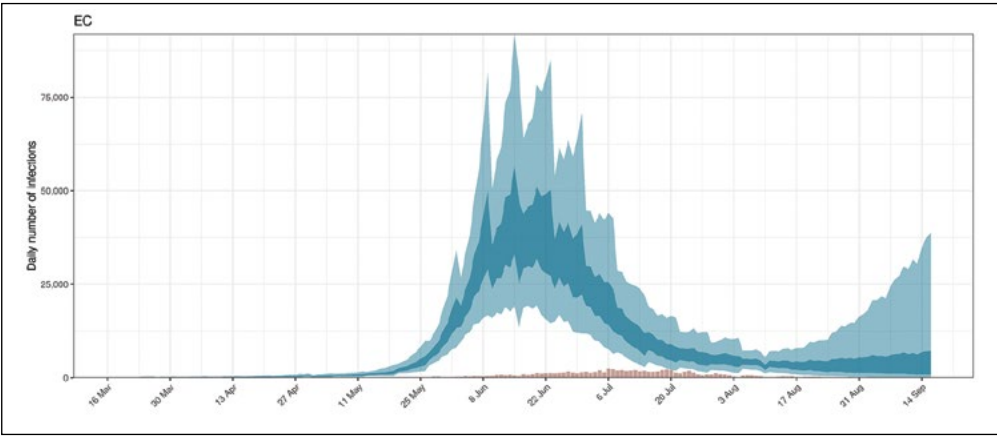


FIGURE 6 Projected infections compared to confirmed cases in Eastern Cape

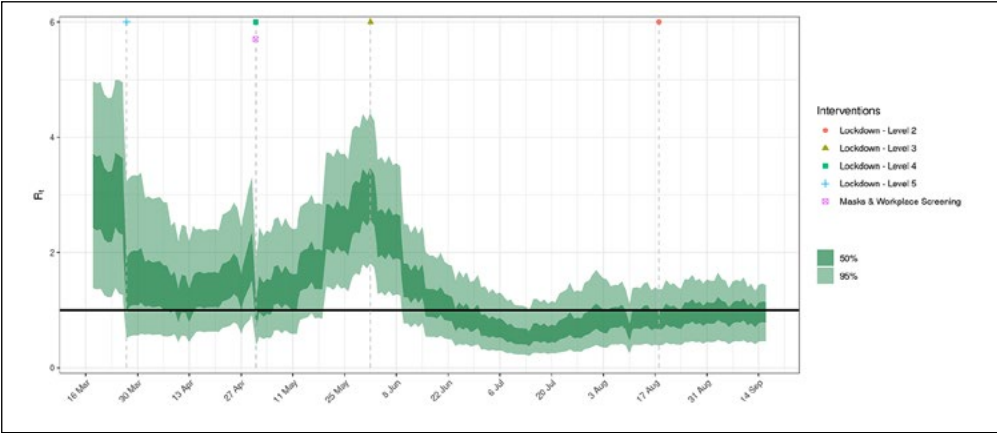


FIGURE 7 Effective reproduction number in Eastern Cape

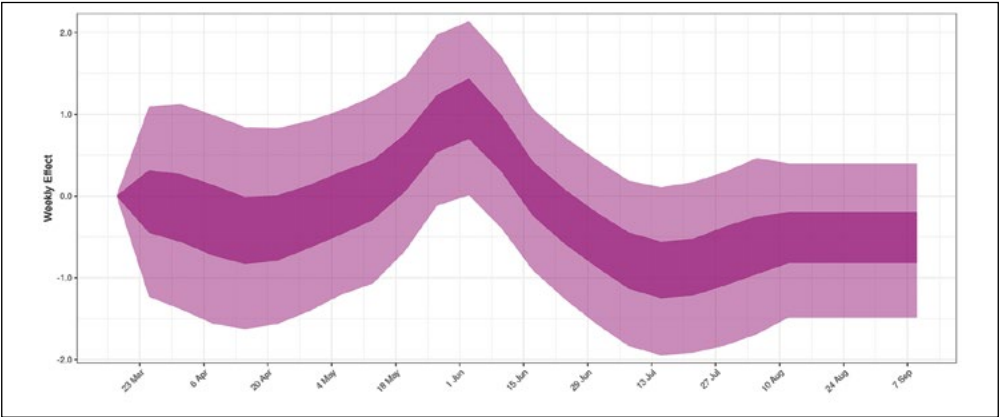


FIGURE 8 Weekly AR(2) effect plot for Eastern Cape

5.2 Free State

Figure 9 shows 90% of excess deaths per [4] and the modelled deaths. For the Free State the fit appears acceptable.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 10. Modelled infections are plotted compared to confirmed cases in Figure 11. Over the last 14 days it would appear that 9.6% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 12. The $R_{t,m}$ started low but increased significantly during lockdown peaking during June and declining afterwards.

The $-\varepsilon_{w_m(t),m}$ is plotted in Figure 13 on page 10.

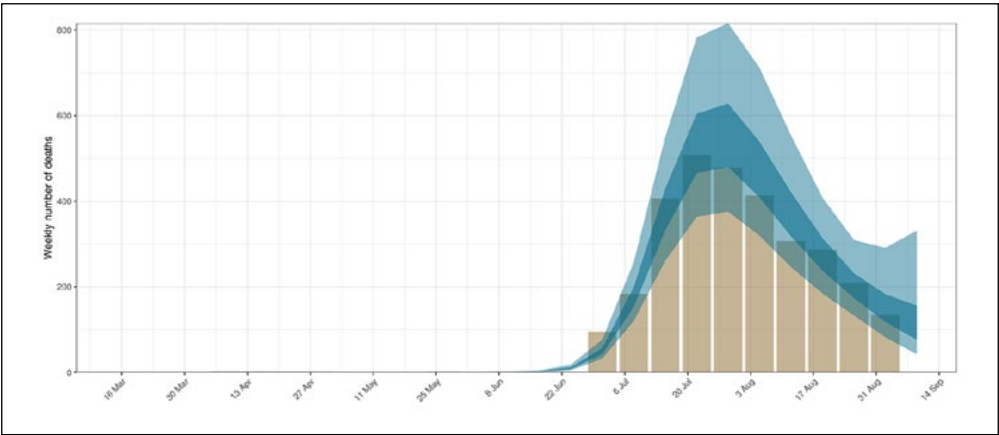


FIGURE 9 Projected deaths and 90% of excess deaths in Free State

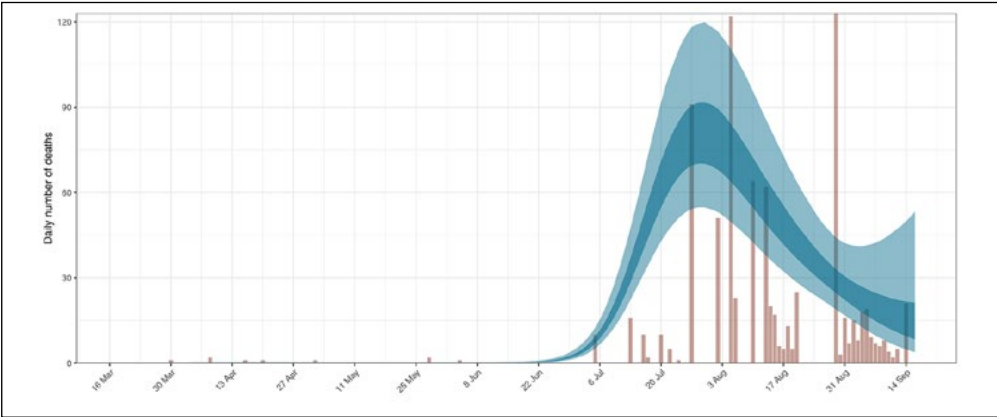


FIGURE 10 Projected deaths compared to reported COVID-19 deaths in Free State

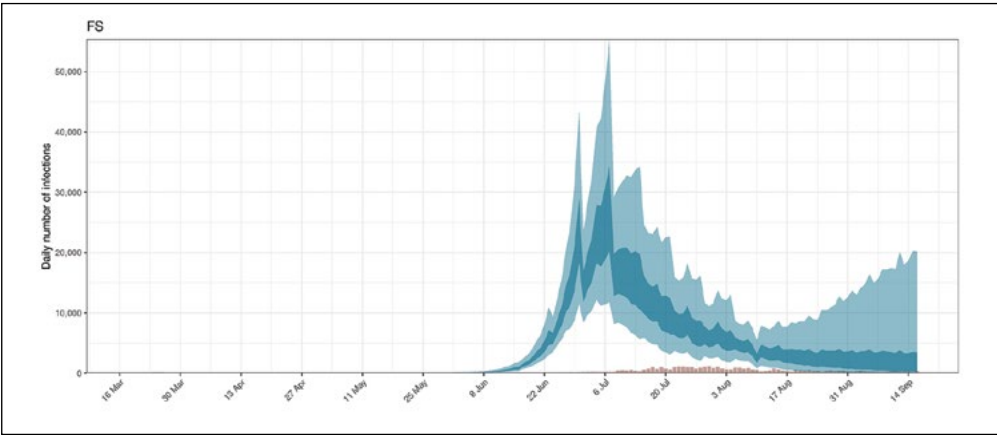


FIGURE 11 Projected infections compared to confirmed cases in Free State

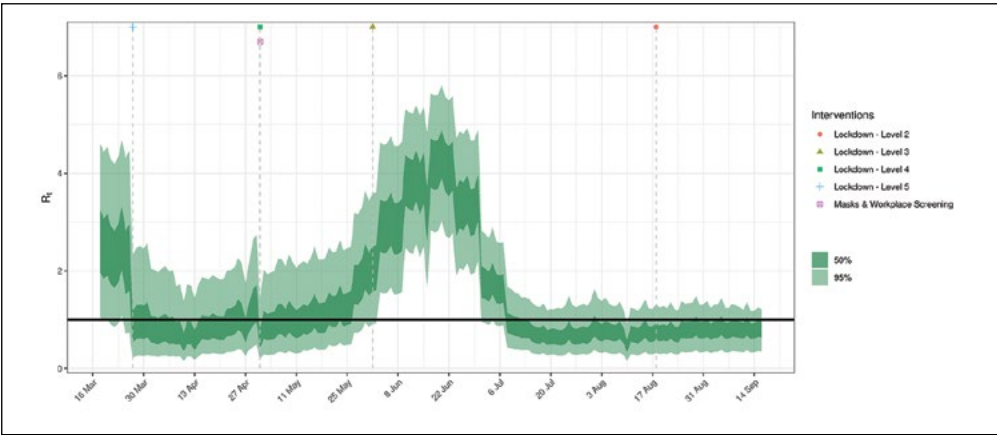


FIGURE 12 Effective reproduction number in Free State

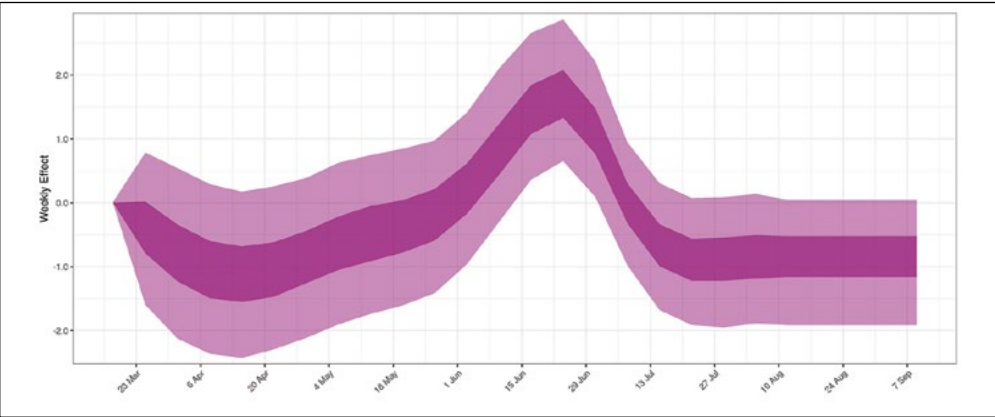


FIGURE 13 Weekly AR(2) effect plot for Free State

5.3 Gauteng

Figure 14 shows 90% of excess deaths per [11] and the modelled deaths. The fit for Gauteng appears reasonable. The levelling off of excess deaths in the last week or two should also be noted.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 15. Modelled deaths are now becoming similar to reported deaths.

Modelled infections are plotted compared to confirmed cases in Figure 16. Over the last 14 days it would appear that 7.6% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 17. The $R_{t,m}$ started low but increased significantly during lockdown peaking during June and declining afterwards.

The $-\epsilon_{w_m(t),m}$ is plotted in Figure 18 on page 12. This parameter has needed to be negative to enable the $R_{t,m}$ to remain at low levels to replicate the slowdown in excess deaths in Gauteng.

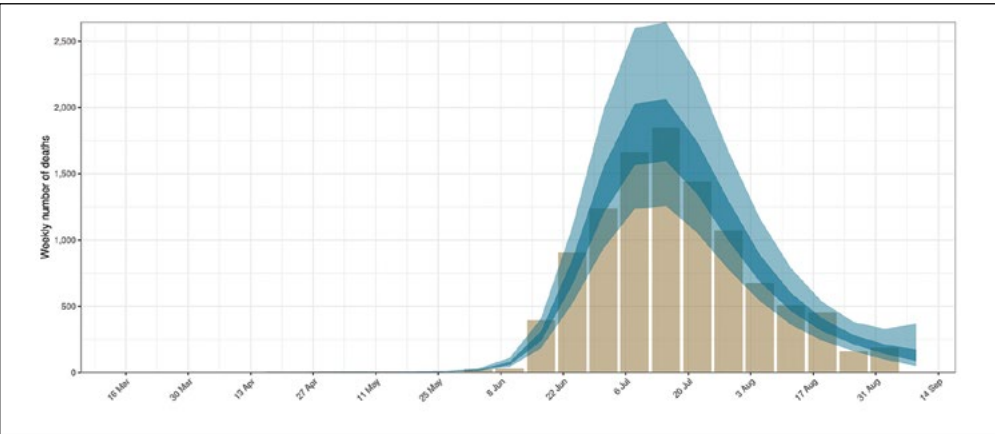


FIGURE 14 Projected deaths and 90% of excess deaths in Gauteng

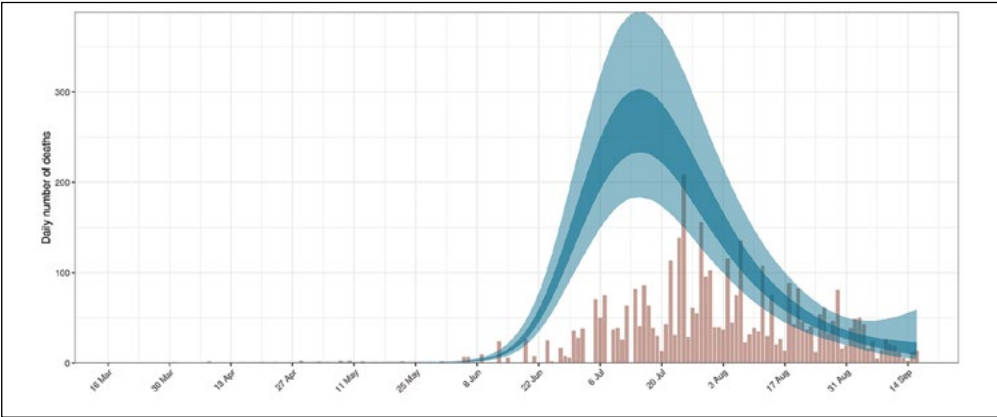


FIGURE 15 Projected deaths compared to reported COVID-19 deaths in Gauteng

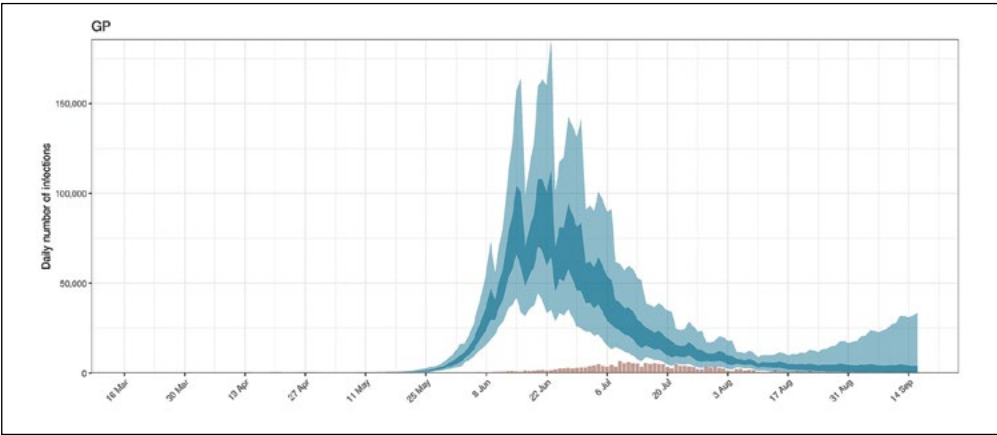


FIGURE 16 Projected infections compared to confirmed cases in Gauteng

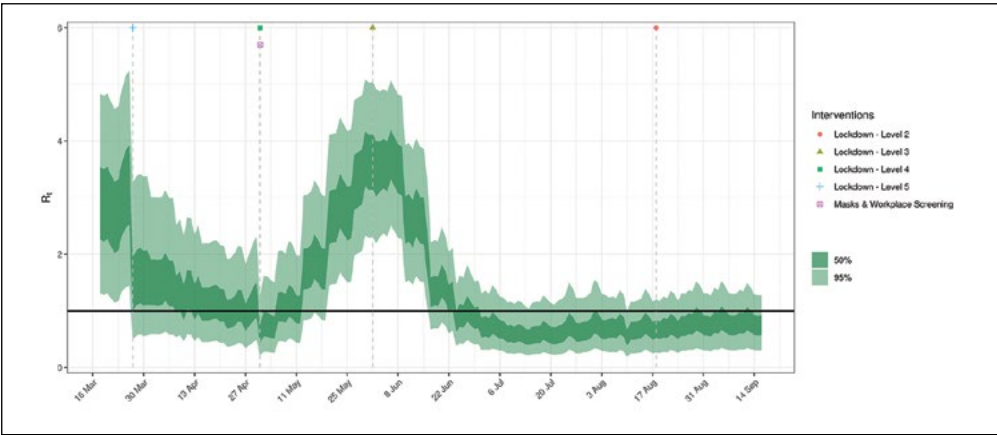


FIGURE 17 Effective reproduction number in Gauteng

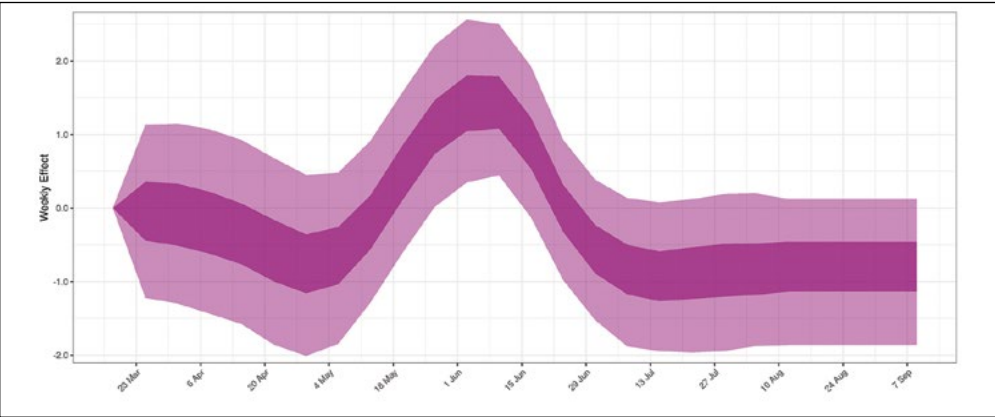


FIGURE 18 Weekly AR(2) effect plot for Gauteng

5.4 KwaZulu-Natal

Figure 19 shows 90% of excess deaths per [4] and the modelled deaths. The fit for KwaZulu-Natal appears reasonable.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 20. Modelled infections are plotted compared to confirmed cases in Figure 21. Over the last 14 days it would appear that 2% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 22. The $R_{t,m}$ started low but increased significantly during lockdown peaking during June and declining afterwards.

The $-\varepsilon_{w_m(t),m}$ is plotted in Figure 23 on page 14. This parameter has needed to be negative to enable the $R_{t,m}$ to remain at low levels to replicate the slowdown in excess deaths in this province.

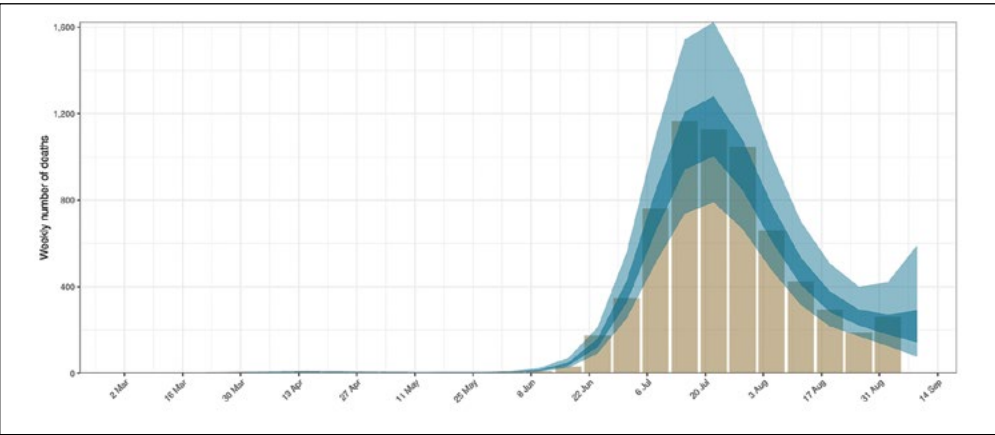


FIGURE 19 Projected deaths and 90% of excess deaths in KwaZulu-Natal

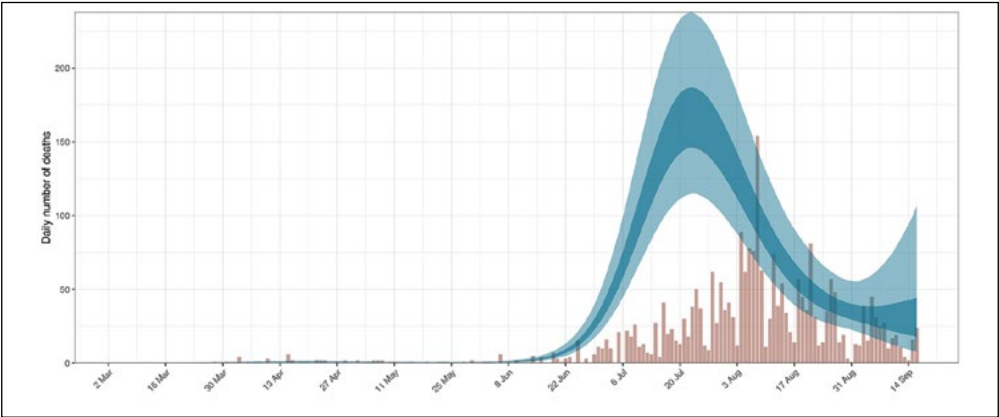


FIGURE 20 Projected deaths compared to reported COVID-19 deaths in KwaZulu-Natal

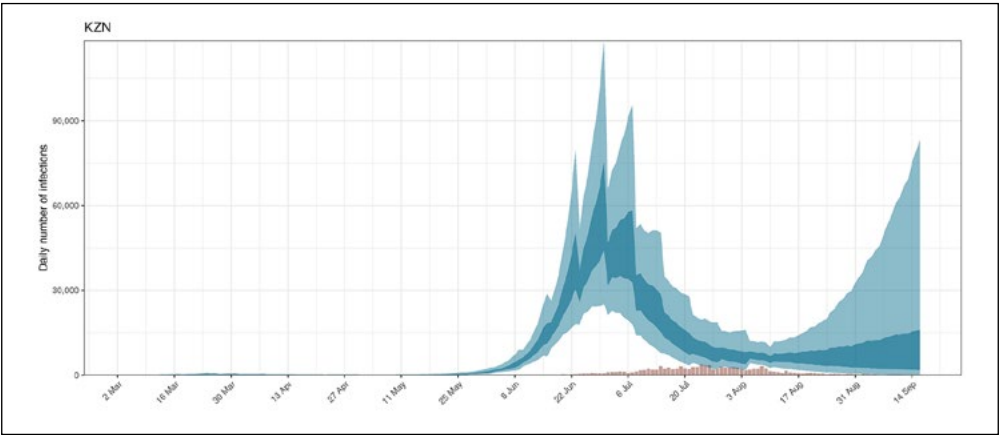


FIGURE 21 Projected infections compared to confirmed cases in KwaZulu-Natal

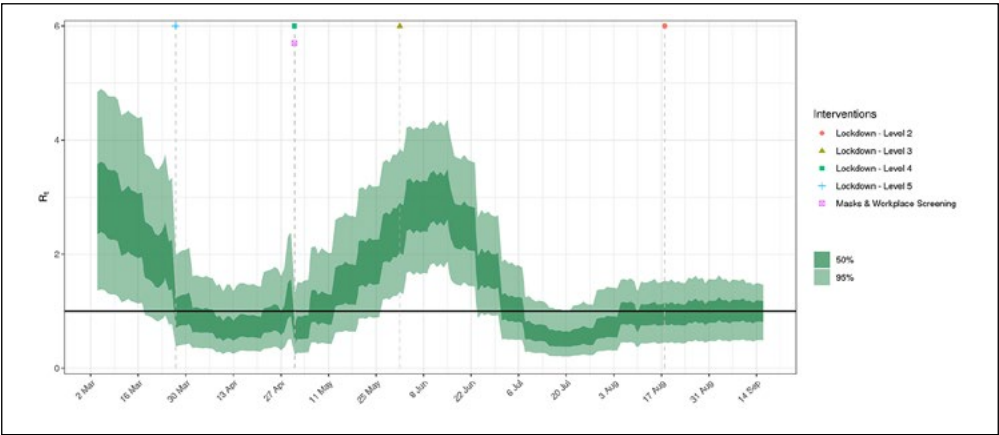


FIGURE 22 Effective reproduction number in KwaZulu-Natal

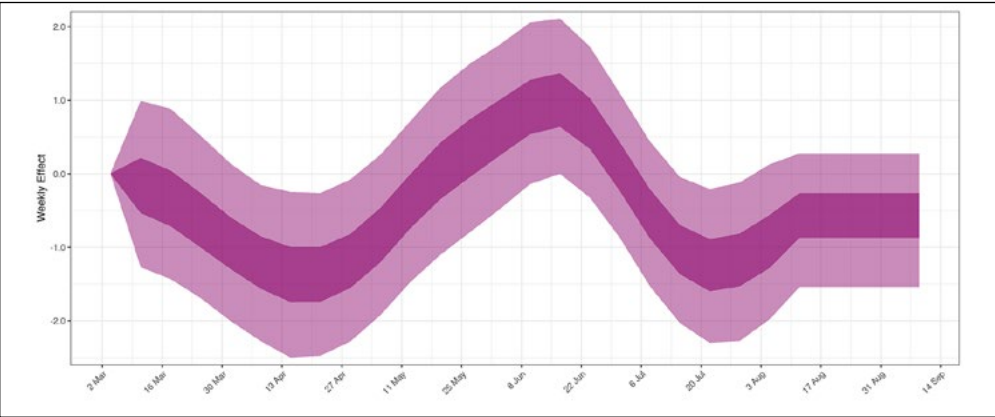


FIGURE 23 Weekly AR(2) effect plot for KwaZulu-Natal

5.5 Limpopo

Figure 24 shows 90% of excess deaths per [4] and the modelled deaths. The fit for Limpopo appears reasonable though possibly over-predicting for the last week or two.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 25. Modelled infections are plotted compared to confirmed cases in Figure 26. Over the last 14 days it would appear that 1% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 27. The $R_{t,m}$ peaks during June and declines afterwards. The $-\epsilon_{w_m(t),m}$ is plotted in Figure 28 on page 16.

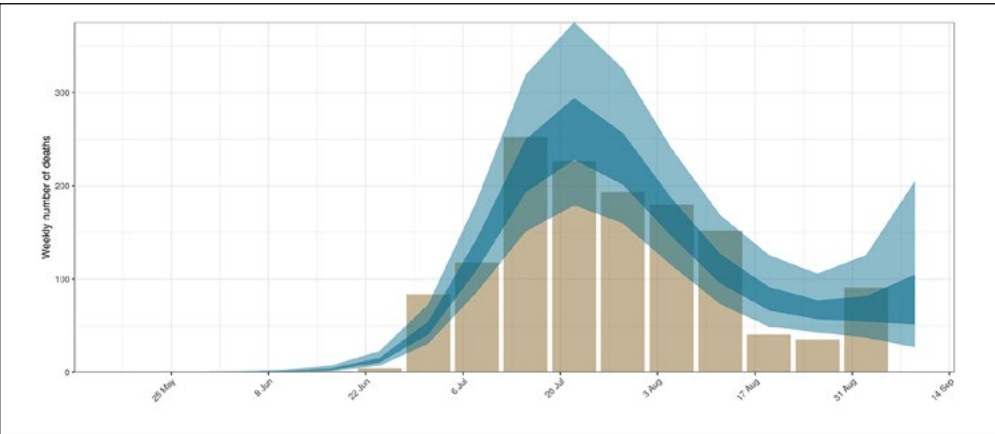


FIGURE 24 Projected deaths and 90% of excess deaths in Limpopo

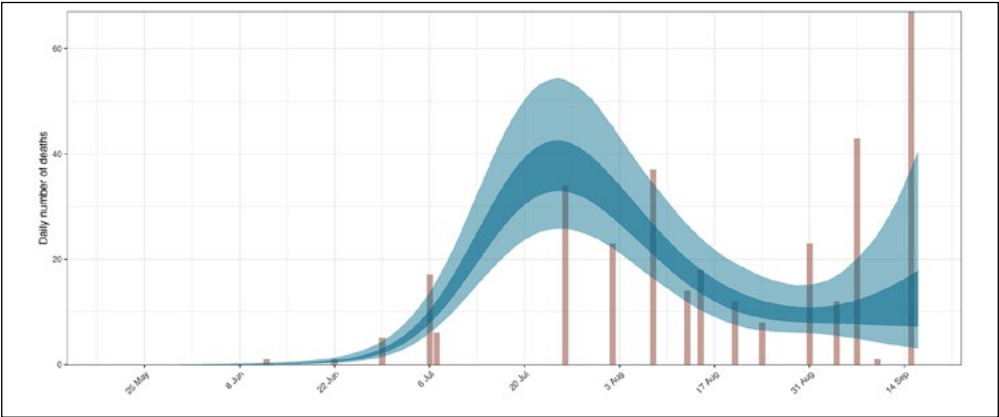


FIGURE 25 Projected deaths compared to reported COVID-19 deaths in Limpopo

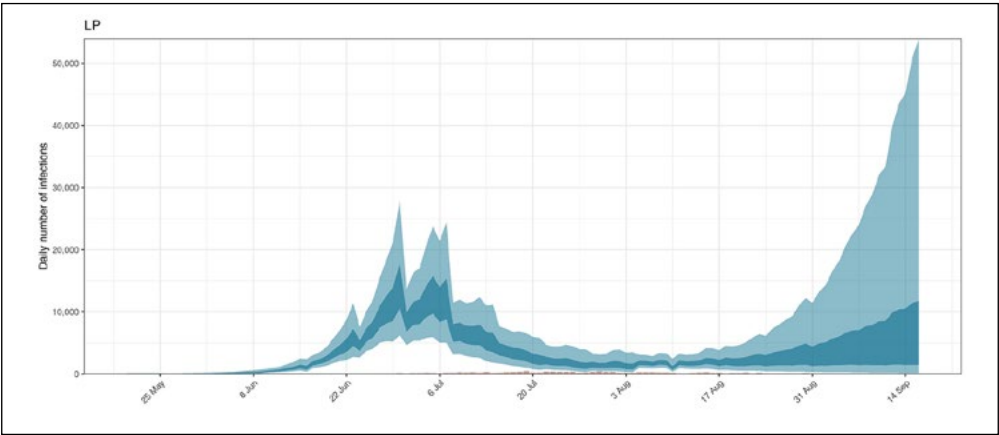


FIGURE 26 Projected infections compared to confirmed cases in Limpopo

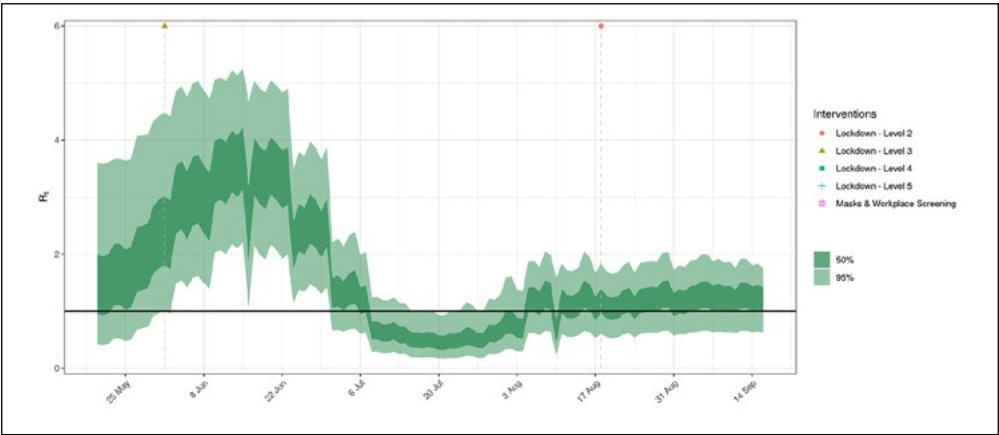


FIGURE 27 Effective reproduction number in Limpopo

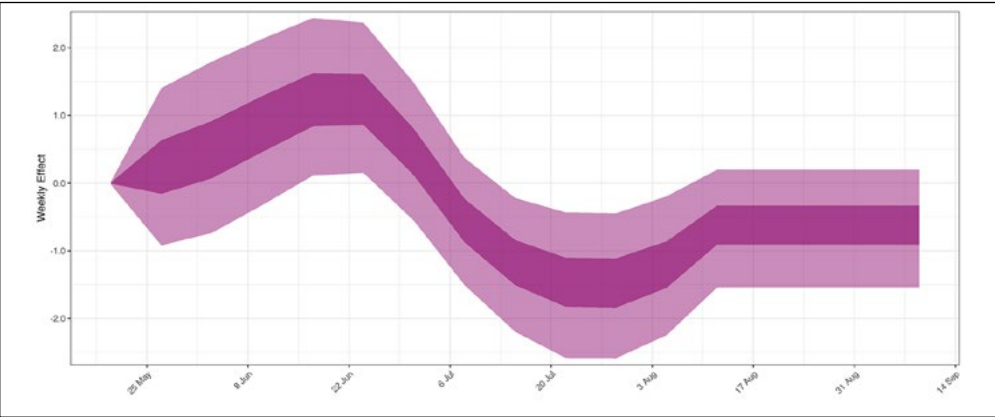


FIGURE 28 Weekly AR(2) effect plot for Limpopo

5.6 Mpumalanga

Figure 29 shows 90% of excess deaths per [11] and the modelled deaths. The fit for Mpumalanga appears reasonable.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 30. Modelled infections are plotted compared to confirmed cases in Figure 31. Over the last 14 days it would appear that 4.5% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 32. The $R_{t,m}$ peaks during June and declines afterwards. The $-\epsilon_{w_m(t),m}$ is plotted in Figure 33 on page 18.

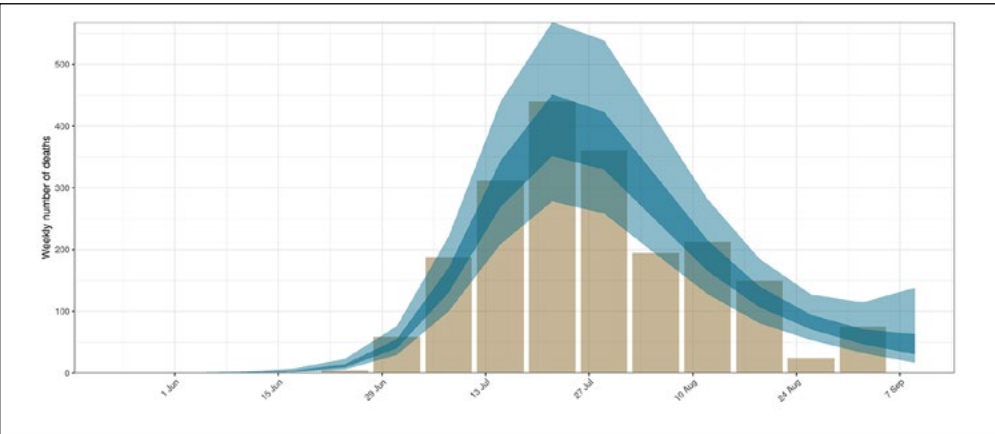


FIGURE 29 Projected deaths and 90% of excess deaths in Mpumalanga

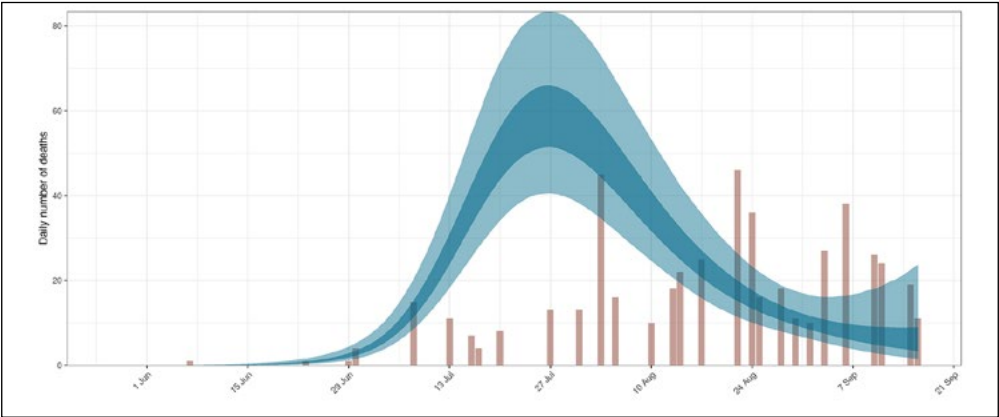


FIGURE 30 Projected deaths compared to reported COVID-19 deaths in Mpumalanga

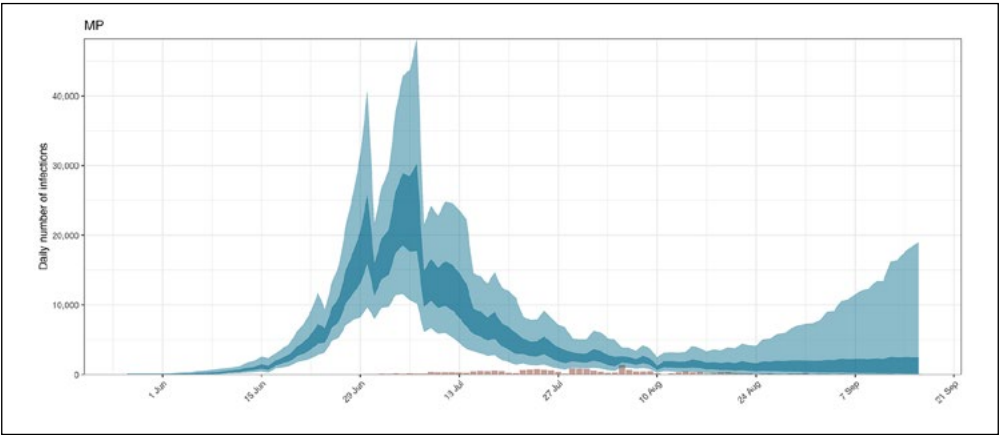


FIGURE 31 Projected infections compared to confirmed cases in Mpumalanga

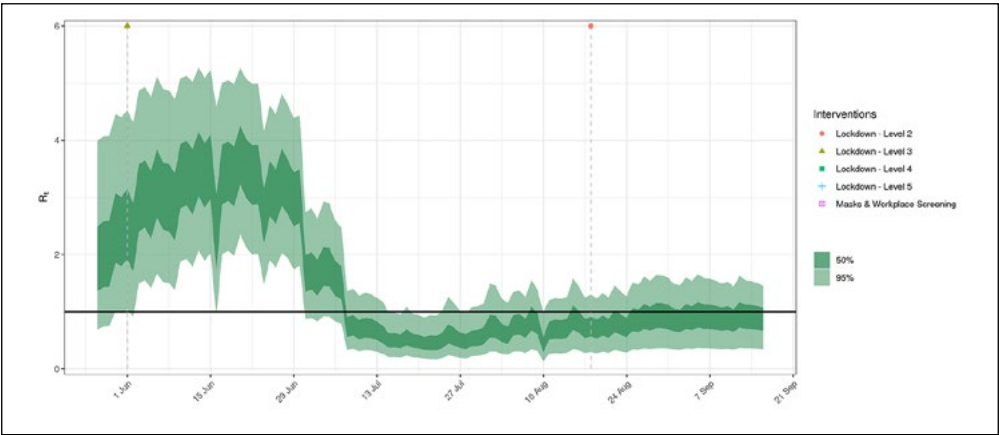


FIGURE 32 Effective reproduction number in Mpumalanga

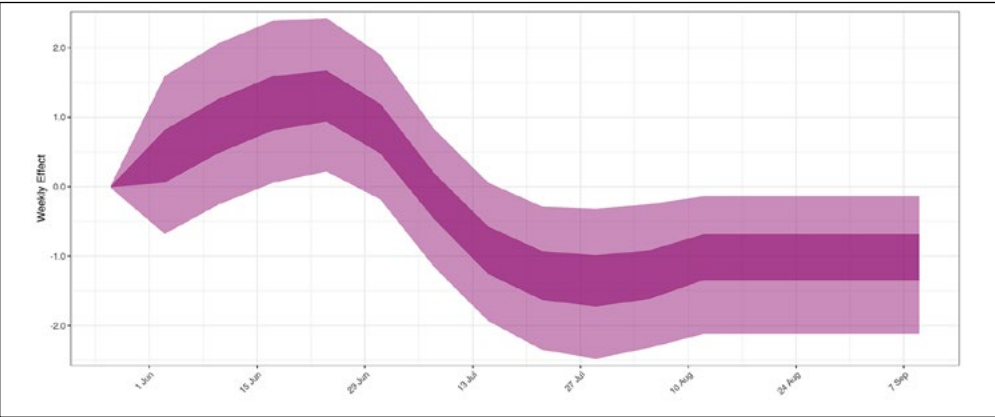


FIGURE 33 Weekly AR(2) effect plot for Limpopo

5.7 Northern Cape

Figure 34 shows 90% of excess deaths per [4] and the modelled deaths. The fit for Northern Cape appears reasonable though possibly over-predicting for the last week or two.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 35. Modelled infections are plotted compared to confirmed cases in Figure 36. Over the last 14 days it would appear that 5.4% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 37. The $R_{t,m}$ peaks during June and declines afterwards. The $-\epsilon_{w_m(t),m}$ is plotted in Figure 38 on page 20.

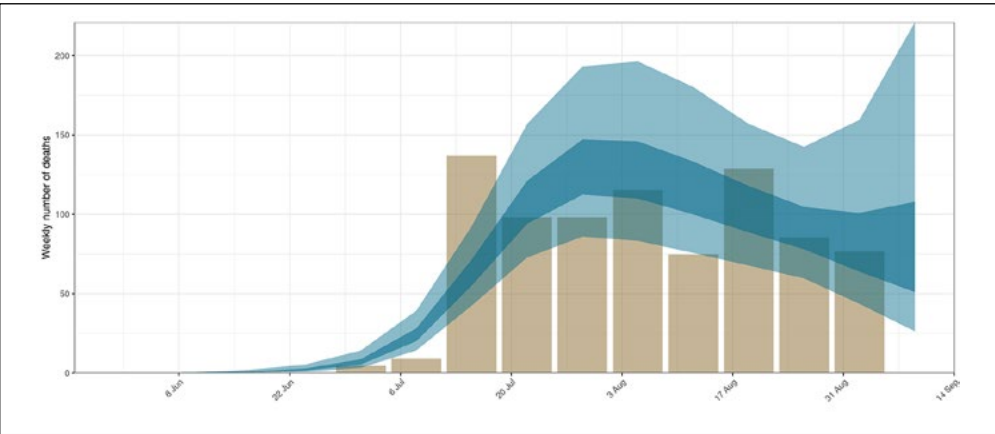


FIGURE 34 Projected deaths and 90% of excess deaths in Northern Cape

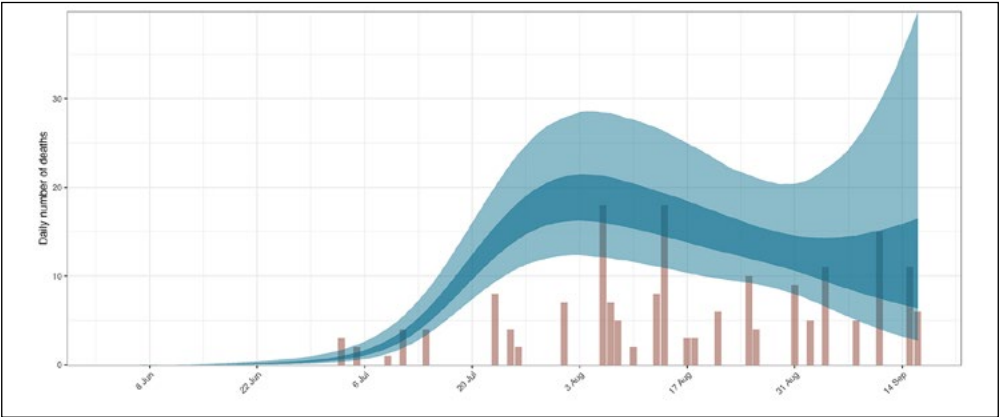


FIGURE 35 Projected deaths compared to reported COVID-19 deaths in Northern Cape

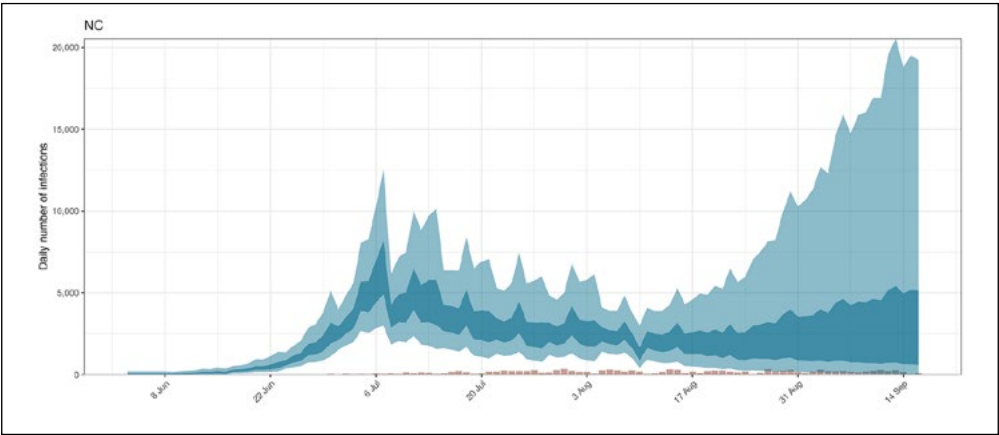


FIGURE 36 Projected infections compared to confirmed cases in Northern Cape

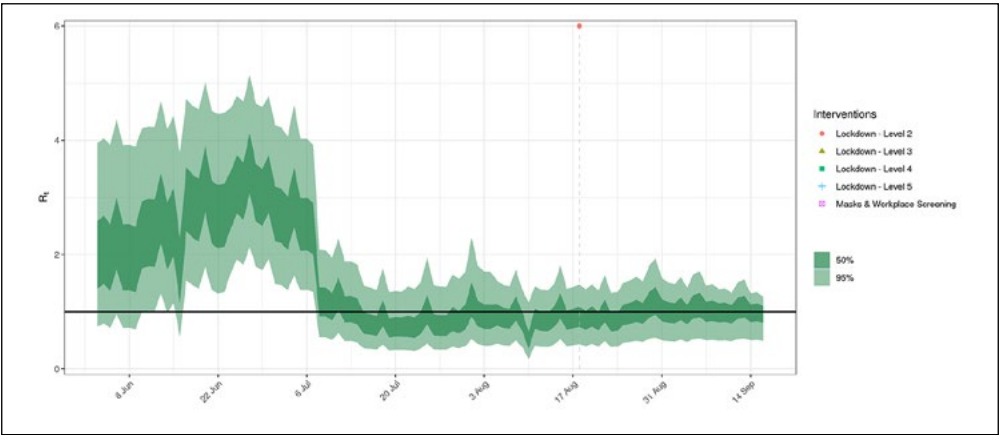


FIGURE 37 Effective reproduction number in Northern Cape

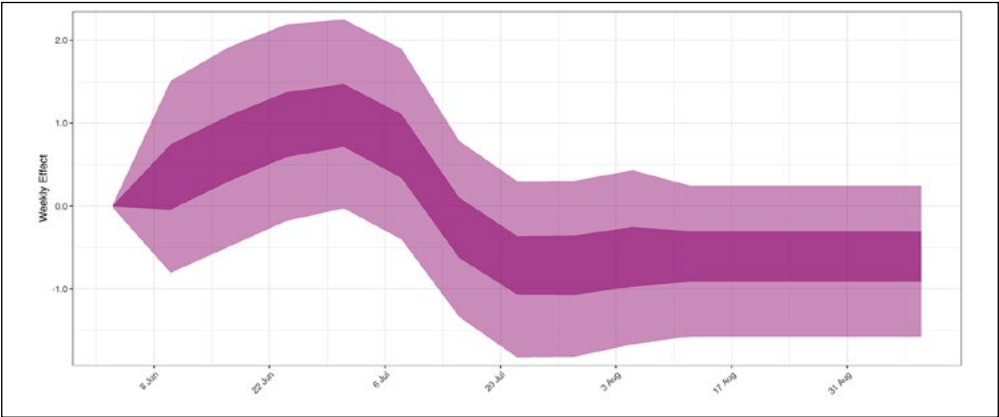


FIGURE 38 Weekly AR(2) effect plot for Northern Cape

5.8 North West

Figure 39 shows 90% of excess deaths per [4] and the modelled deaths. The fit for North West appears reasonable though the peak mid July appears not to be modelled accurately.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 40. Modelled infections are plotted compared to confirmed cases in Figure 41. Over the last 14 days it would appear that 7.1% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 42. The $R_{t,m}$ peaks during June and declines afterwards. The $-\epsilon_{w_m(t),m}$ is plotted in Figure 43 on page 22.

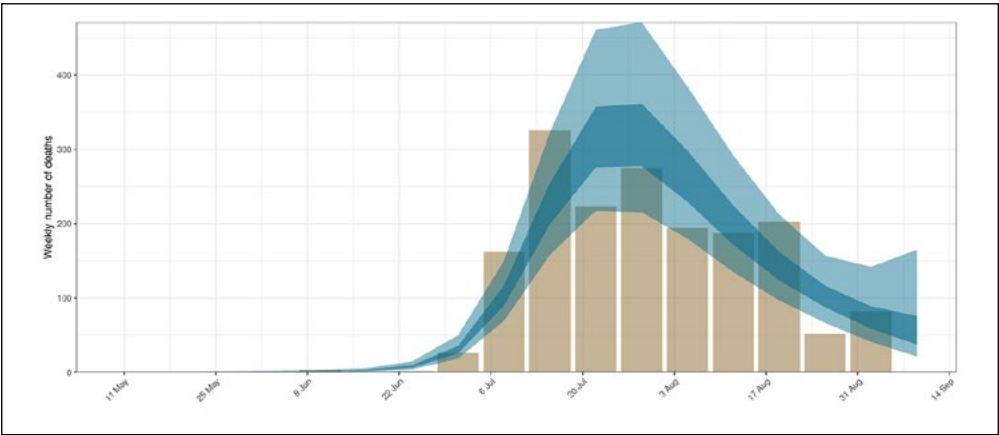


FIGURE 39 Projected deaths and 90% of excess deaths in North West

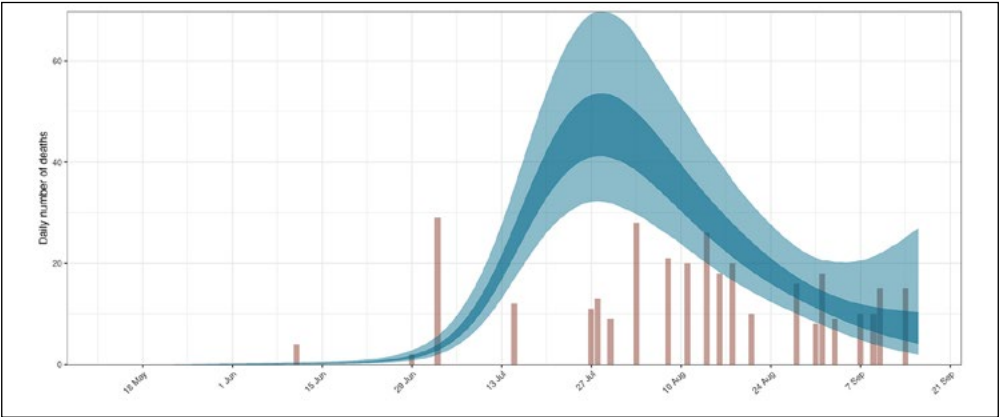


FIGURE 40 Projected deaths compared to reported COVID-19 deaths in North West

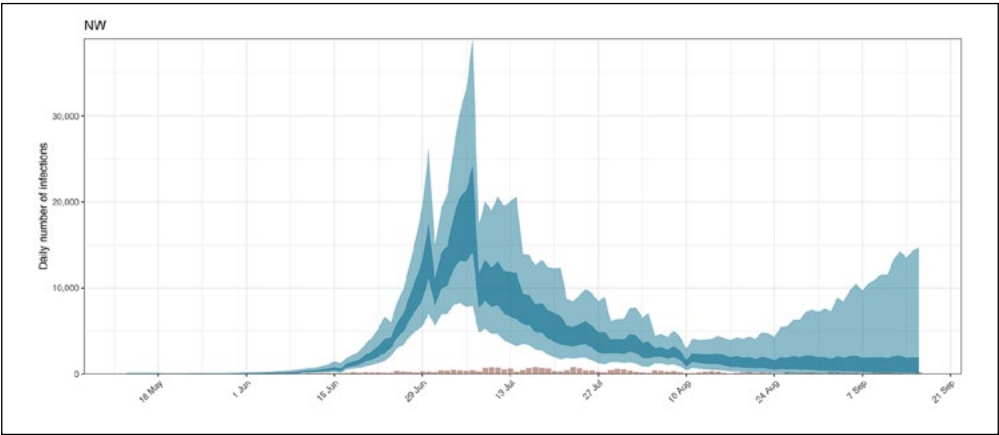


FIGURE 41 Projected infections compared to confirmed cases in North West

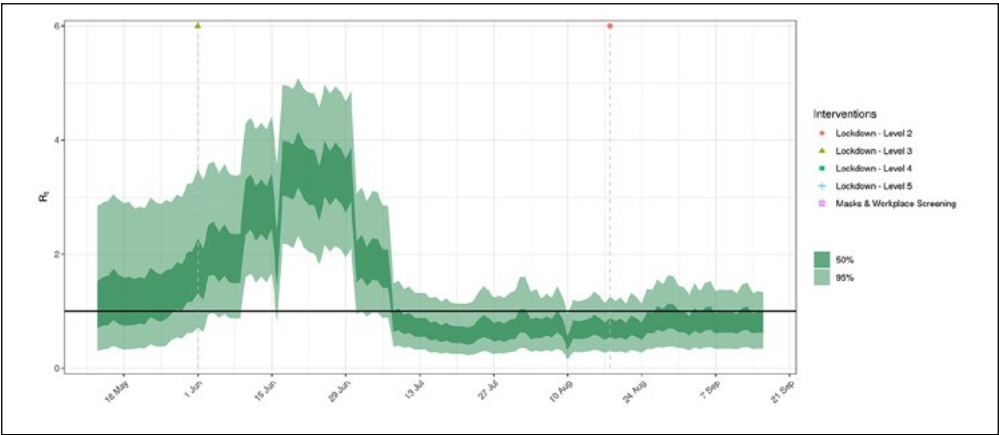


FIGURE 42 Effective reproduction number in North West

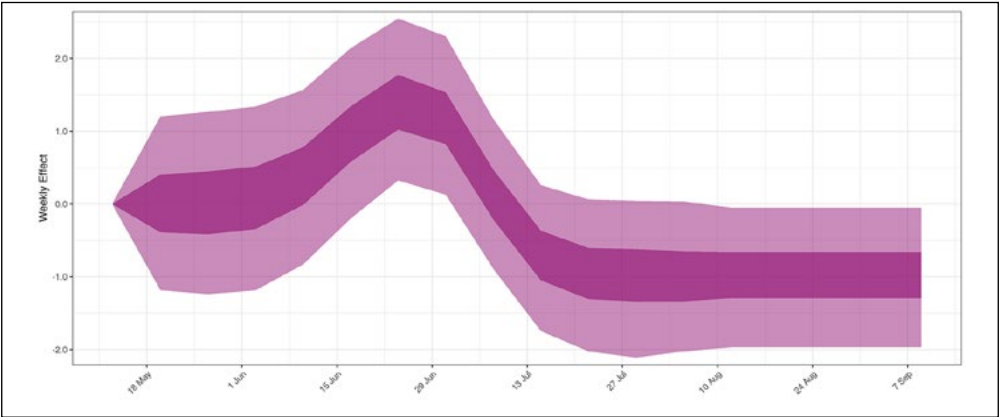


FIGURE 43 Weekly AR(2) effect plot for North West

5.9 Western Cape

Figure 44 shows 90% of excess deaths per [11] and the modelled deaths. The fit appears reasonable. The recent blip in excess deaths does appear to have affected the projections made by the model.

The reported deaths (in brown) and estimated deaths (in blue) are plotted in Figure 45. Modelled deaths appear to be lower than reported deaths, indicating either that reported deaths include prior period deaths or that the model is under-projecting the current situation.

Modelled infections are plotted compared to confirmed cases in Figure 46. Over the last 14 days it would appear that 6.1% of all new infections were tested.

The $R_{t,m}$ is plotted in Figure 47. The $R_{t,m}$ declines at the start of lockdown and continues to decline over time. The $-\varepsilon_{w_m(t),m}$ is plotted in Figure 48 on page 24. This parameter has needed to be negative to enable the $R_{t,m}$ to remain at low levels to replicate the slowdown in excess deaths in this province.s

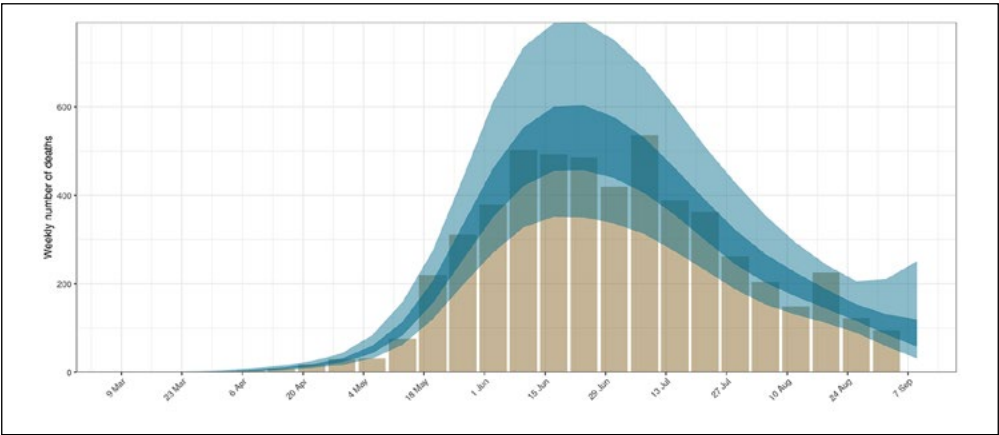


FIGURE 44 Projected deaths and 90% of excess deaths in Western Cape

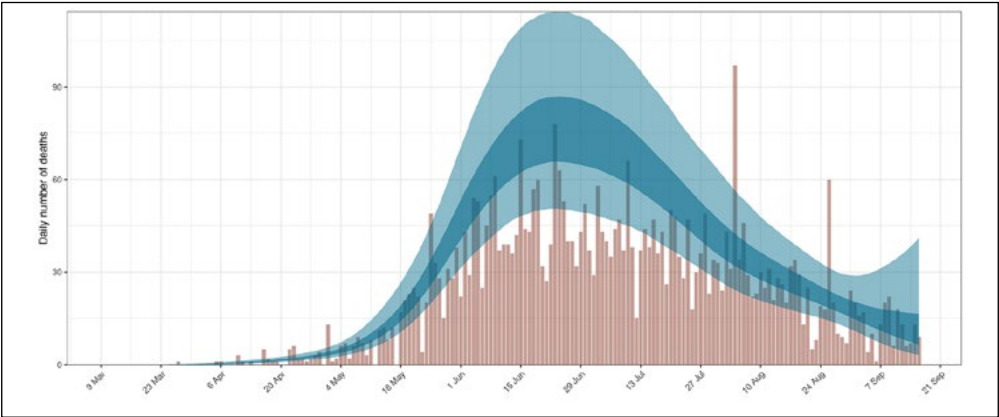


FIGURE 45 Projected deaths compared to reported COVID-19 deaths in Western Cape

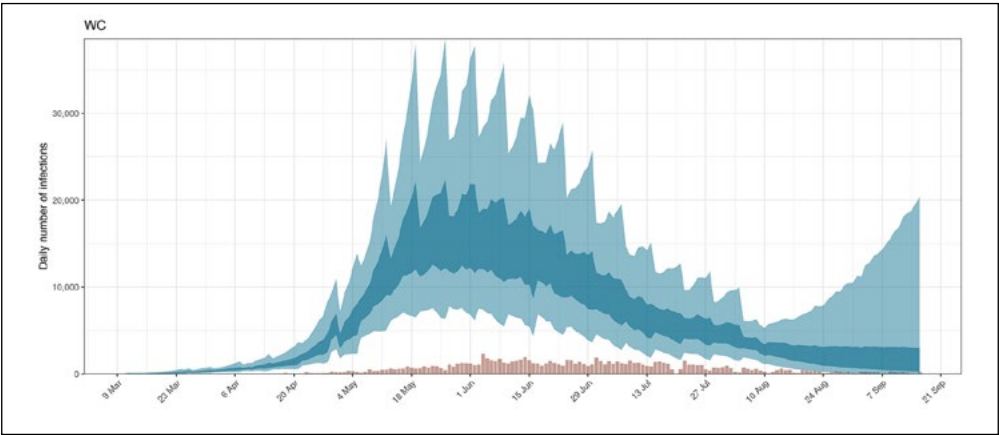


FIGURE 46 Projected infections compared to confirmed cases in Western Cape

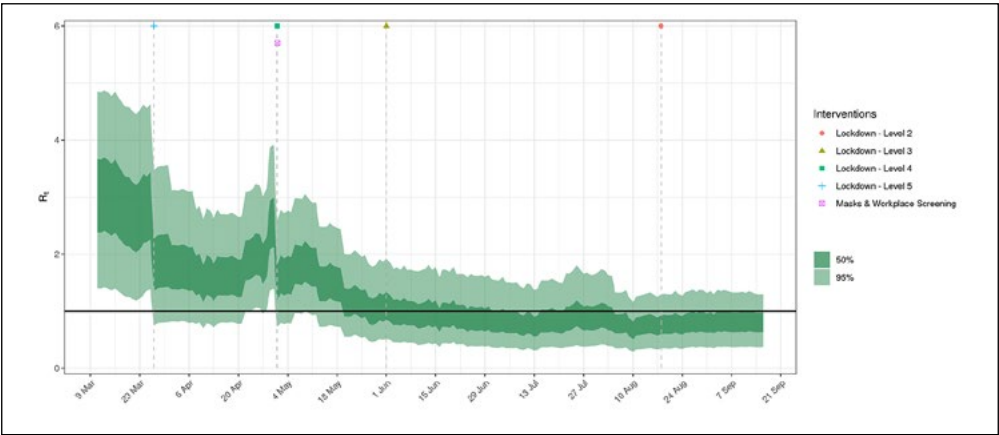


FIGURE 47 Effective reproduction number in Western Cape



FIGURE 48 Weekly AR(2) effect plot for Western Cape

6. PARAMETER ESTIMATES

To understand the net parameter estimates for α and α_m^s and their impact on $R_{t,m}$ we plot the percentage reduction in $R_{t,m}$ assuming the particular index is 1 (representing a 100% reduction in average mobility).

This is equivalent to plotting $1 - 2 \cdot \phi^{-1} \left(- \left(\alpha + \alpha_m^s \right) \right)$ for a particular province m . We also plot $1 - 2 \cdot \phi^{-1} \left(- \left(\beta + \beta_m^s \right) \right)$.

The confidence intervals are relatively wide with some provinces not showing much impact of mobility, including the Western Cape. These same provinces show higher impact of NPIs introduced with Level 4 (such as mask wearing and workplace screening).

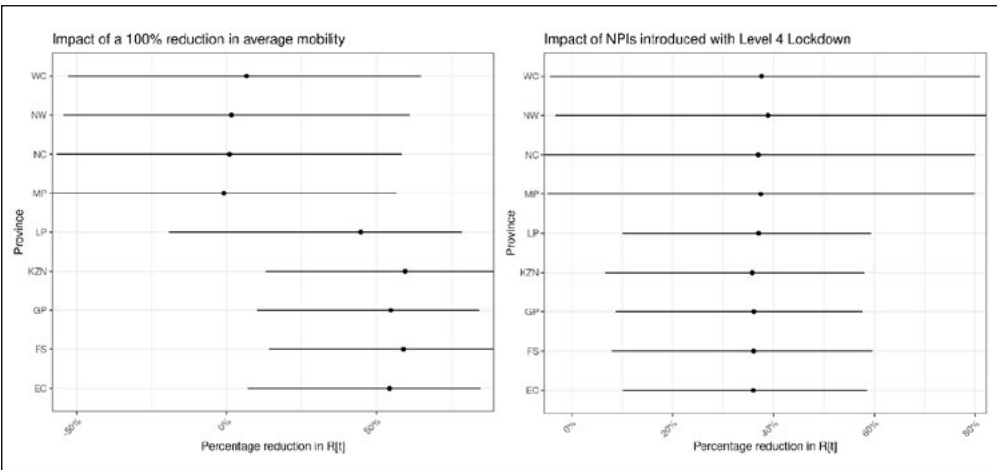


FIGURE 49 Model parameter estimates with 95% confidence intervals

7. REPRODUCTION NUMBER ESTIMATES

7.1 Initial reproduction number

Estimates for $R_{0,m}$ for each province are plotted in Figure 50. It should be noted that the provinces with low $R_{0,m}$ tend to be the ones with less data early in the epidemic. This indicates that $R_{0,m}$ may not be correctly estimated for these provinces as there is no data to differentiate it from post lockdown $R_{t,m}$. This may be the cause for the limited impact of mobility for these provinces.

In any case, for all provinces, it is probable that the posterior estimates for $R_{0,m}$ are not heavily influenced by the data. This is probably due to the relatively early lockdown implemented in South Africa. There were probably not enough deaths that resulted from infection prior to lockdown to develop an independent estimate for each province of $R_{0,m}$.

7.2 Current reproduction number as at 16 September 2020

Current estimates and 95% confidence intervals for $R_{t,m}$ (current effective reproduction number) are plotted for each province (Figure 51). A value below 1 would indicate an epidemic that is slowing while a value above 1 indicates an epidemic that is growing. It is clear that the spread of the epidemic is somewhat slowed compared to the initial $R_{0,m}$. For some provinces the confidence interval does not include 1 which would indicate this to a stronger degree.

8. ATTACK RATE AS AT 16 SEPTEMBER 2020

The estimated attack rate (with 95% confidence intervals) is tabulated in Table 1. This is the proportion of the population infected to date. This figure has to be estimated, because many that are infected experience no or mild symptoms, thus they may not seek medical advice and hence will never be tested.

The confidence intervals for these estimates remain very wide. The highest attack rates are occurring in Northern Cape, North West and Free State, but this may be reflecting limited data. WC attack rate appears to be relatively low compared to other provinces.

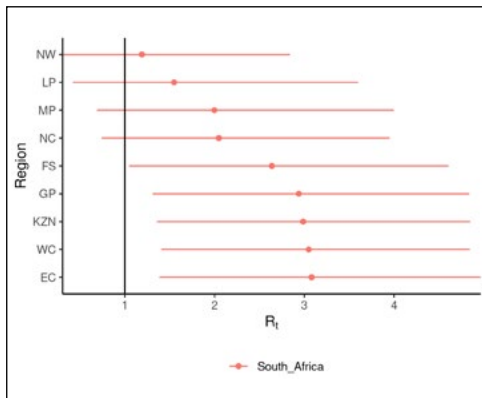


FIGURE 50 Initial reproduction number by province

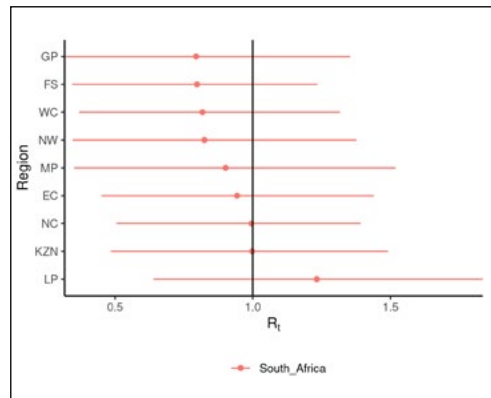


FIGURE 51 Current reproduction number by province

TABLE 1 Current attack rate by province

Province	Attack rate
EC	24.64% [17.09%-35.46%]
FS	24.49% [16.32%-38.42%]
GP	17.75% [12.37%-25.25%]
KZN	15.18% [10.12%-24.51%]
LP	7.15% [4.07%-14.80%]
MP	11.31% [7.62%-17.29%]
NC	20.32% [11.41%-41.78%]
NW	10.72% [7.21%-16.78%]
WC	17.81% [12.87%-24.88%]
South Africa	16.37% [13.97%-19.63%]

9. PROJECTIONS

All models are wrong but some are useful—George Box [13]

One of the reasons we build models is so that we can make sense of the future or indeed the past. We can project forward models to assess the impact of varying assumptions on future outcomes. This gives us a sense of how changes in actions may impact the future, i.e. it allows us to answer “what if” questions. Note however that in projecting the future we are extrapolating, and due care needs to be taken. There are numerous limitations to this model and these projections discussed below but the author is also of the opinion that the projections add value in that they indicate a significant range across the scenarios projected and in such a manner inform discussion.

An incorrect model can be useful because it enables a better understanding of the model and the phenomena being modelled. This section should be used with caution because, as we show in calibration, the model seems to be projecting higher deaths than observed for some provinces. Note detailed projection output can be found [here](#).

Projections are done using constant, increasing and decreasing mobility but other interventions are left constant. It should also be noted that the $\varepsilon_{w_m(t),m}$ are projected forward at constant levels and are not adjusted for projection purposes.

The $\varepsilon_{w_m(t),m}$ is held at “current” levels when projecting forwards. It should be noted that these levels are set 28 days prior as recent deaths only affect infections that occurred prior to that. Based on the observed calibration for Western Cape there is a possibility that this may be resulting in too low recent infections.

9.1 No change in mobility

In this projection, mobility is held constant at current levels. Level 4 interventions are left intact. The result of this scenario as at 31 December 2020 is shown in Table 2. Note based on these it would appear that some provinces have already reached the peak in daily deaths.

TABLE 2 Projection to 31 December 2020 with no change in mobility

Province	Attack rate	Deaths	Peak daily deaths	Peak date
EC	33.84% [17.88%-72.54%]	13 959 [7 941- 30 951]	211	2020-07-11
FS	30.12% [16.80%-65.82%]	4 589 [2 758- 10 191]	82	2020-07-29
GP	21.82% [12.56%-58.94%]	13 975 [8 851- 36 456]	271	2020-07-15
KZN	28.22% [10.63%-71.32%]	14 159 [5 872- 37 777]	168	2020-07-23
LP	34.55% [4.46%-80.65%]	9 457 [1 387- 25 002]	116	2020-11-10
MP	19.96% [7.87%-67.95%]	3 868 [1 714- 13 805]	59	2020-07-27
NC	35.54% [12.10%-76.57%]	2 348 [838- 5 387]	21	2020-10-15
NW	16.54% [7.41%-59.49%]	3 099 [1 534- 11 414]	48	2020-07-28
WC	22.47% [13.16%-59.09%]	7 376 [4 721- 18 961]	77	2020-06-24
South Africa	25.99% [16.43%-39.79%]	72 830 [47 534-110 420]	864	2020-07-21

9.2 Increase in mobility

This scenario assumes future mobility half-way between current mobility levels and normal levels. Other interventions are left intact. The attack rate and deaths after increasing mobility are tabulated below (as at 31 December 2020).

TABLE 3 Projection to 31 December 2020 with increase in mobility

Province	Attack rate	Deaths	Peak daily deaths	Peak date
EC	36.25% [17.96%-76.11%]	14 849 [8 015- 32 838]	211	2020-07-11
FS	31.04% [16.87%-68.52%]	4 715 [2 762- 10 630]	82	2020-07-29
GP	24.36% [12.61%-68.67%]	15 321 [8 879- 43 770]	271	2020-07-15
KZN	31.80% [10.69%-75.51%]	15 817 [5 903- 40 785]	168	2020-07-23
LP	36.05% [4.49%-81.63%]	9 868 [1 390- 25 493]	123	2020-11-10
MP	20.19% [7.87%-68.52%]	3 909 [1 714- 13 928]	59	2020-07-27
NC	38.99% [12.22%-80.35%]	2 565 [845- 5 651]	23	2020-10-20
NW	18.18% [7.46%-65.05%]	3 368 [1 539- 12 620]	48	2020-07-28
WC	25.19% [13.27%-68.01%]	8 128 [4 748- 22 336]	77	2020-06-24
South Africa	28.33% [16.93%-44.43%]	78 539 [48 509-122 780]	864	2020-07-21

Based on the above an increase mobility could mean roughly 6 000 more deaths by the end of the year.

9.3 Decreased mobility

This scenario assumes future mobility is decreased by 50% more than currently. The attack rate and deaths after decreased mobility are shown in Table 4 (as at 31 December 2020).

TABLE 4 Projection to 31 December 2020 with decrease in mobility

Province	Attack rate	Deaths	Peak daily deaths	Peak date
EC	31.87% [17.76%-68.74%]	13 234 [7 885- 29 099]	211	2020-07-11
FS	29.34% [16.73%-63.51%]	4 481 [2 751- 9 733]	82	2020-07-29
GP	20.26% [12.50%-48.32%]	13 137 [8 823- 29 697]	271	2020-07-15
KZN	25.28% [10.55%-67.15%]	12 813 [5 848- 35 208]	168	2020-07-23
LP	33.10% [4.44%-79.47%]	9 061 [1 383- 24 565]	109	2020-11-10
MP	19.74% [7.86%-67.26%]	3 830 [1 713- 13 647]	59	2020-07-27
NC	32.61% [12.01%-73.37%]	2 165 [833- 5 078]	19	2020-08-03
NW	15.25% [7.37%-53.93%]	2 886 [1 531- 10 113]	48	2020-07-28
WC	20.78% [13.05%-50.04%]	6 903 [4 706- 15 880]	77	2020-06-24
South Africa	24.24% [16.02%-36.35%]	68 511 [46 584-102 233]	864	2020-07-21

Based on the above a decrease in mobility could mean roughly 4 000 fewer deaths by the end of the year.

9.4 South Africa

We plot the results for South Africa as a whole in Figure 52. This is the sum of the provincial projections. The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 53.

9.5 Eastern Cape

We plot the projections for Eastern Cape in Figure 54. The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 55.

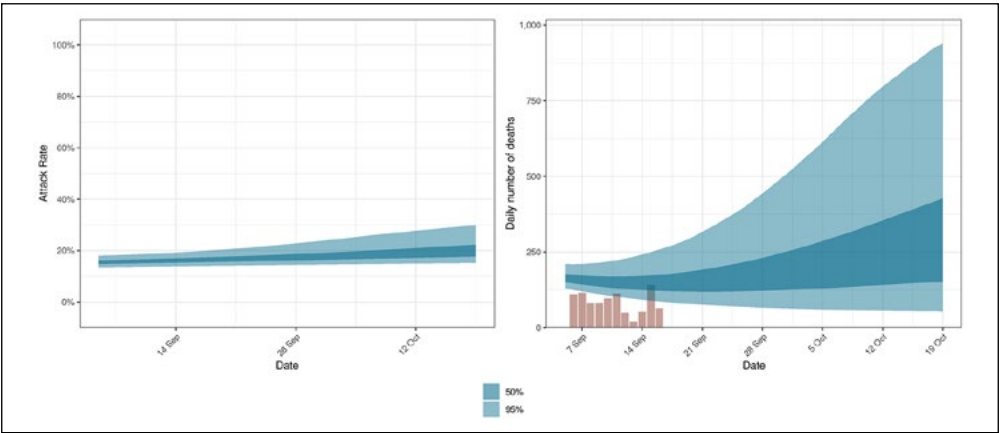


FIGURE 52 Projected attack rate and daily deaths in South Africa

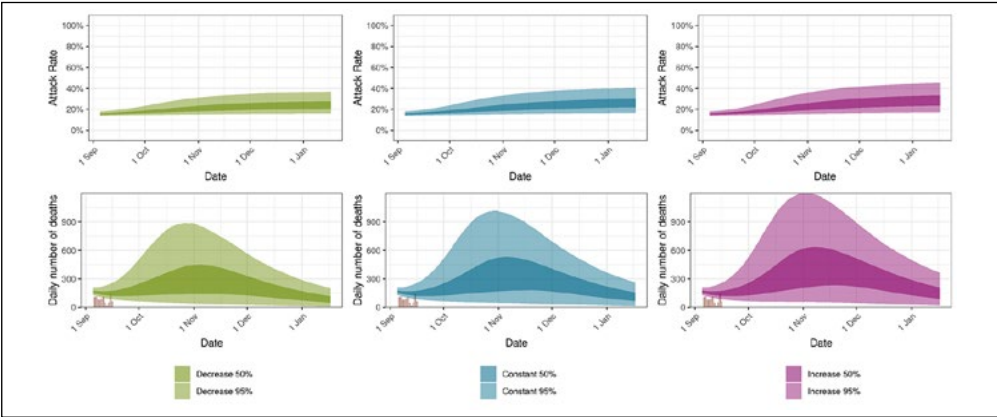


FIGURE 53 Projected attack rate and daily deaths in South Africa by scenario

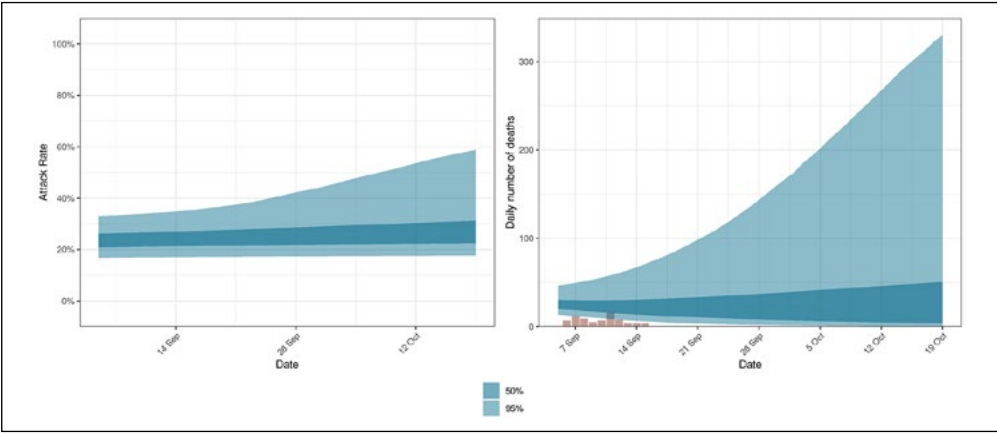


FIGURE 54 Projected attack rate and daily deaths in Eastern Cape

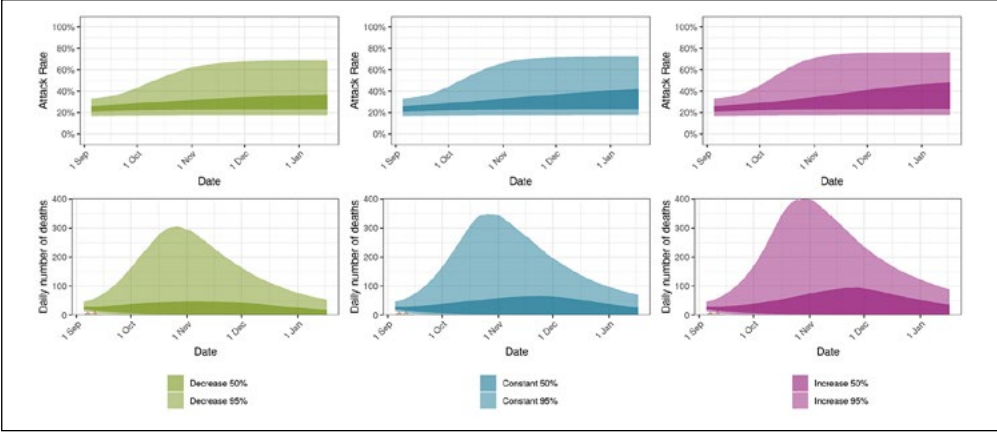


FIGURE 55 Projected attack rate and daily deaths in Eastern Cape by scenario

9.6 Free State

We plot the projections for the Free State in Figure 56. The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 57.

9.7 Gauteng

We plot the projections for Gauteng in Figure 58. The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 59.

9.8 KwaZulu-Natal

We plot the projections for KwaZulu-Natal in Figure 60.

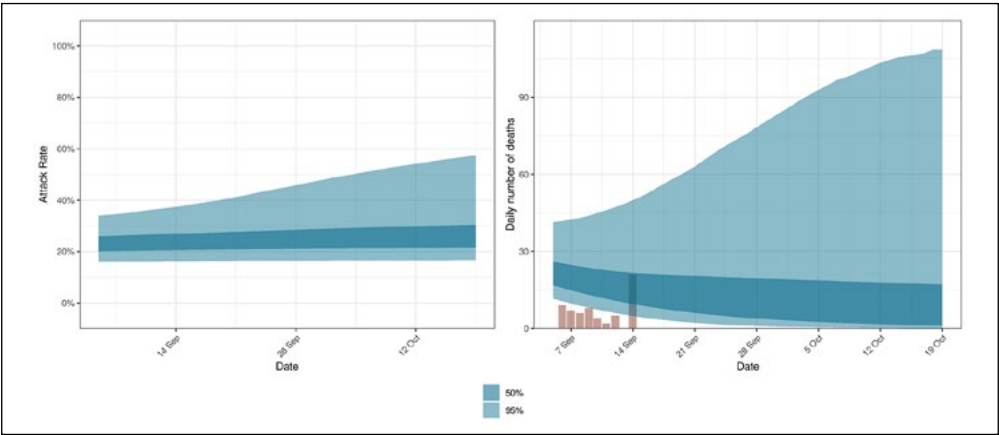


FIGURE 56 Projected attack rate and daily deaths in Free State

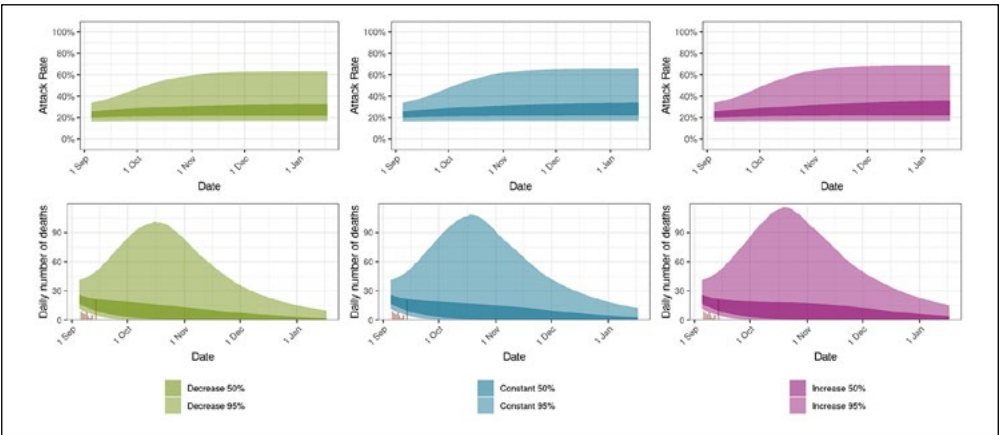


FIGURE 57 Projected attack rate and daily deaths in Free State by scenario

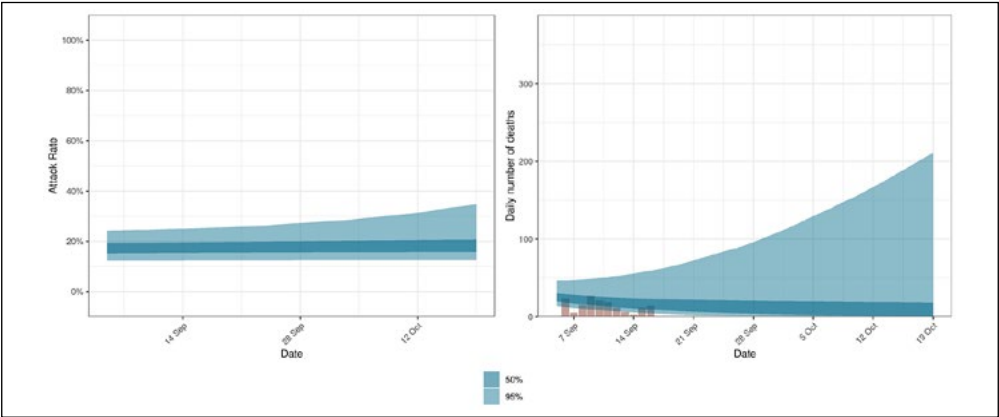


FIGURE 58 Projected attack rate and daily deaths in Gauteng

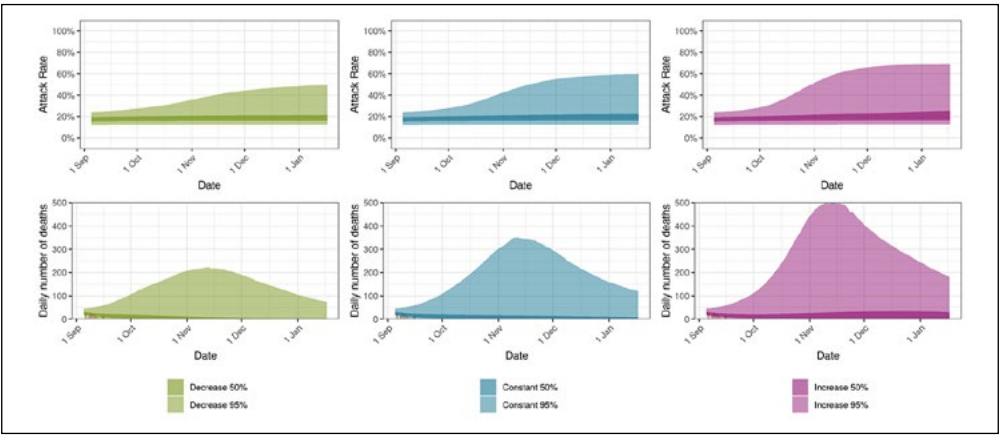


FIGURE 59 Projected attack rate and daily deaths in Gauteng by scenario

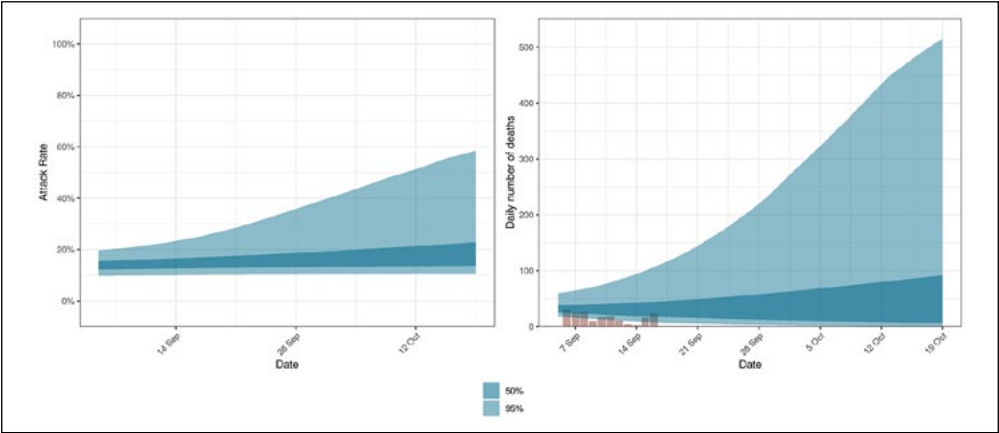


FIGURE 60 Projected attack rate and daily deaths in KwaZulu-Natal

The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 61.

9.9 Limpopo

We plot the projections for Limpopo in Figure 62. The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 63.

9.10 Mpumalanga

We plot the projections for Mpumalanga in Figure 64. The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 65.

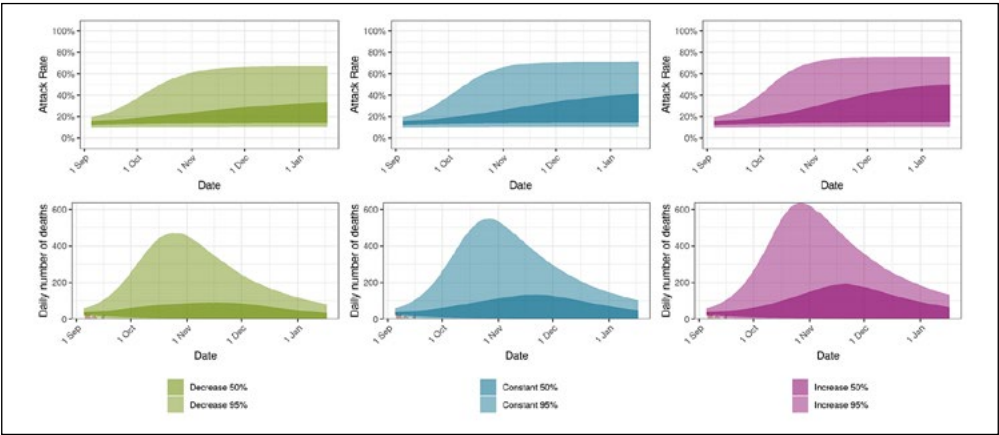


FIGURE 61 Projected attack rate and daily deaths in KwaZulu-Natal by scenario

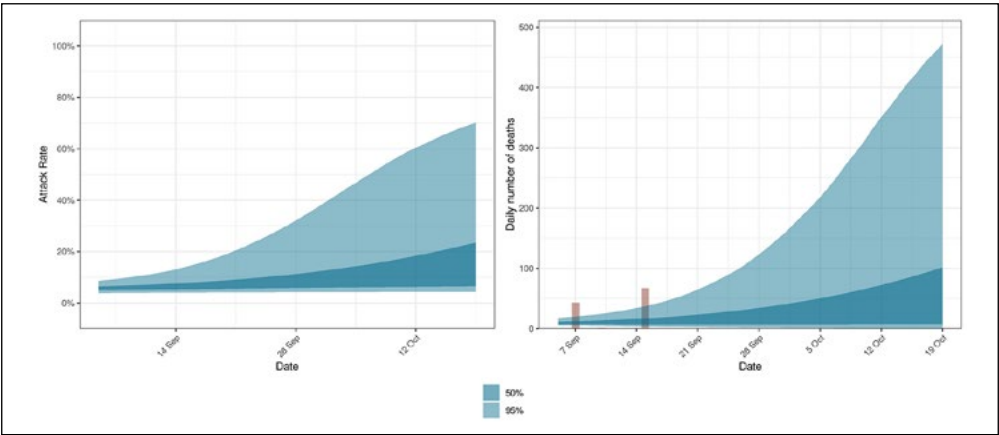


FIGURE 62 Projected attack rate and daily deaths in Limpopo

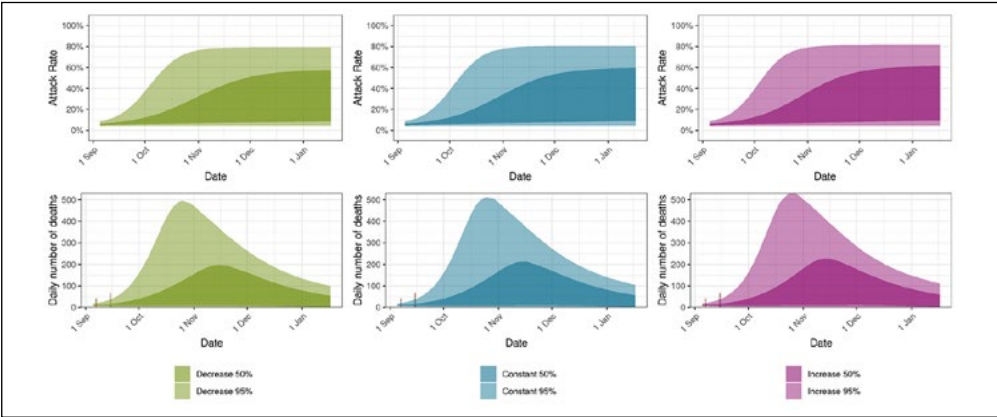


FIGURE 63 Projected attack rate and daily deaths in Limpopo by scenario

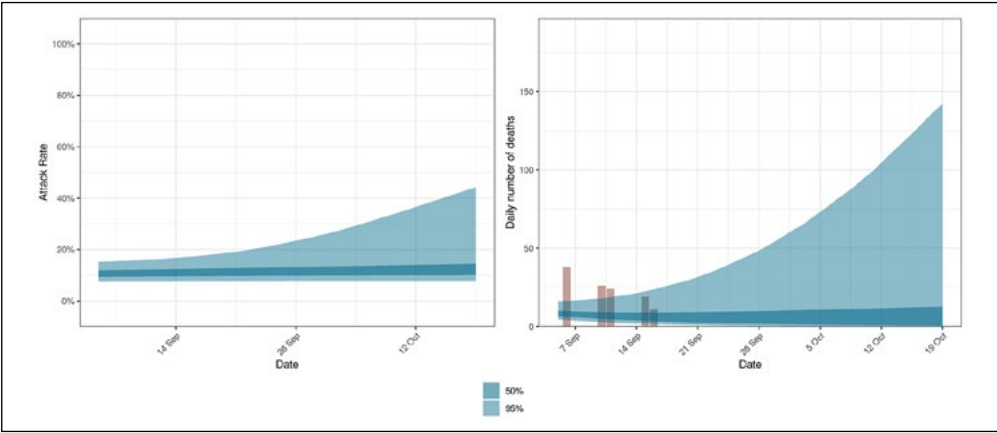


FIGURE 64 Projected attack rate and daily deaths in Mpumalanga

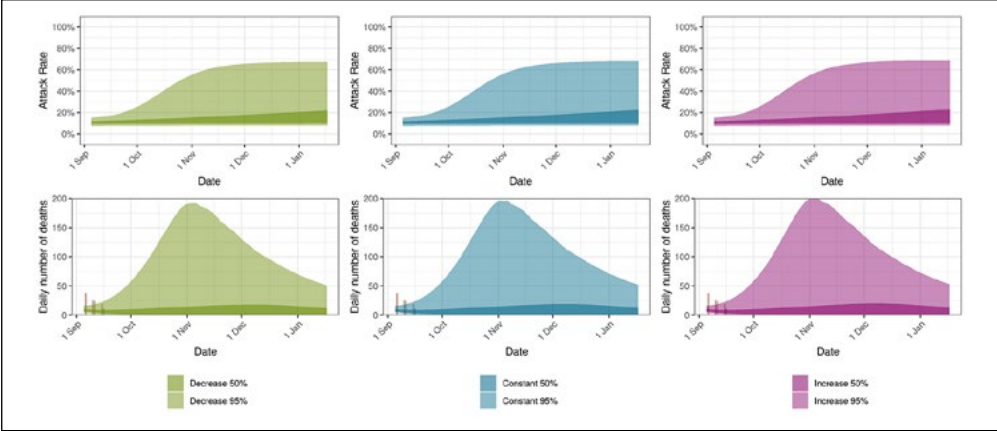


FIGURE 65 Projected attack rate and daily deaths in Mpumalanga by scenario

9.11 Northern Cape

We plot the projections for the Northern Cape in Figure 66.

The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 67.

9.12 North West

We plot the projections for the North West in Figure 68.

The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 69.

9.13 Western Cape

We plot projections for the Western Cape in Figure 70.

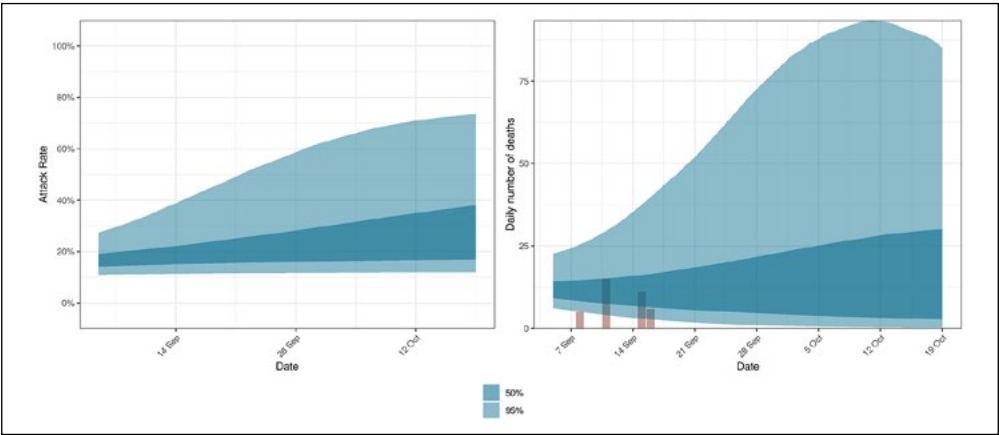


FIGURE 66 Projected attack rate and daily deaths in Northern Cape

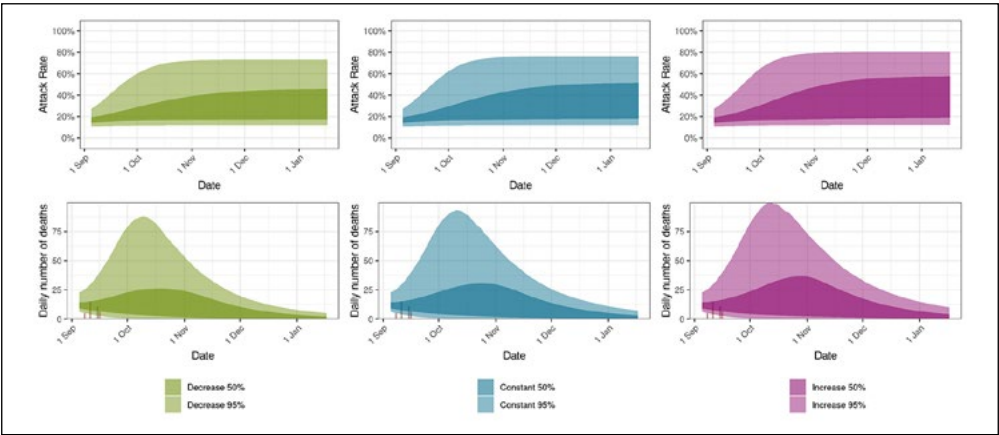


FIGURE 67 Projected attack rate and daily deaths in Northern Cape by scenario

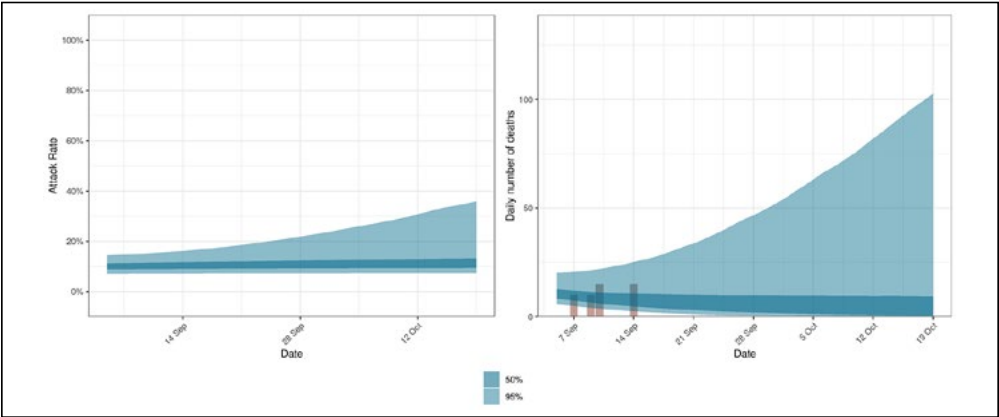


FIGURE 68 Projected attack rate and daily deaths in North West

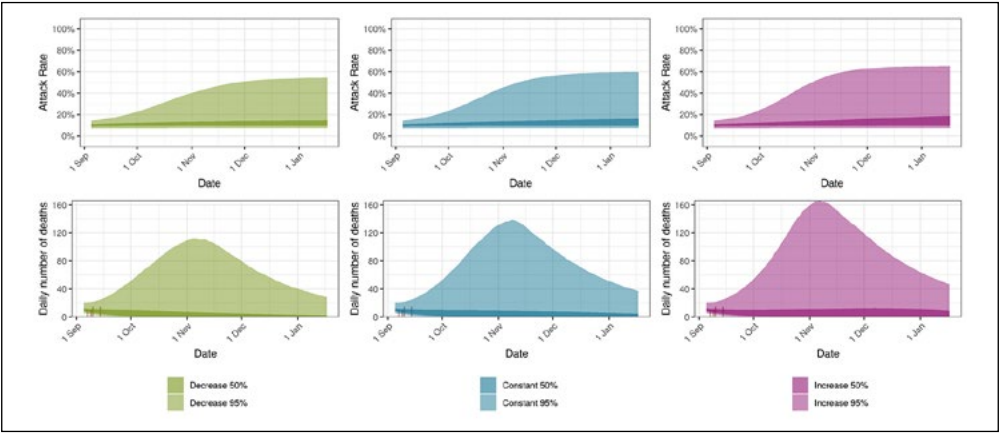


FIGURE 69 Projected attack rate and daily deaths in North West by scenario

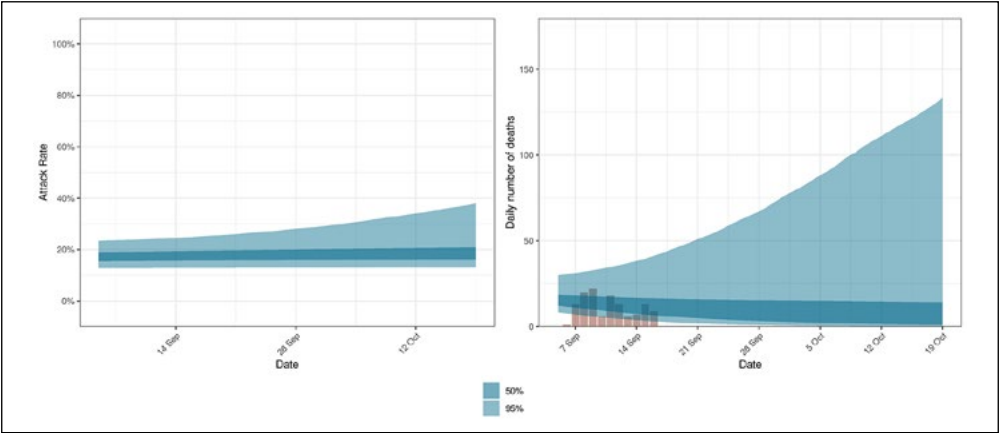


FIGURE 70 Projected attack rate and daily deaths in Western Cape

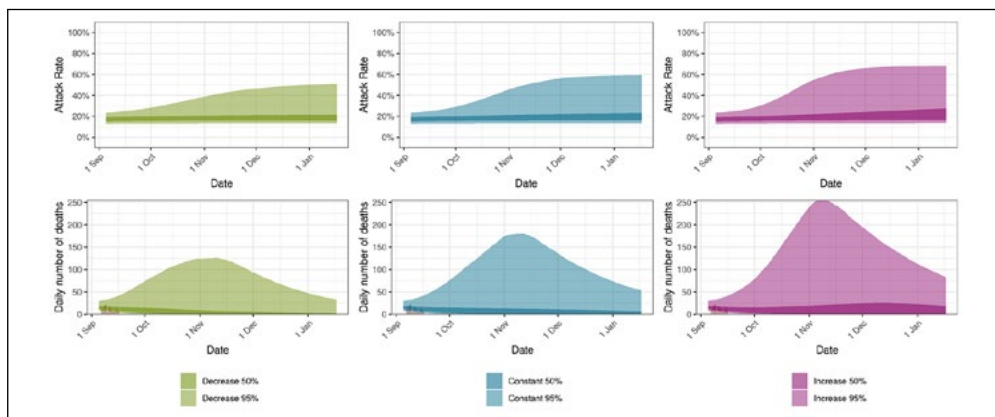


FIGURE 71 Projected attack rate and daily deaths in Western Cape by scenario

The attack rate and deaths over a longer period for both the constant mobility and increased mobility scenarios are plotted in Figure 71.

10. DISCUSSION

There remains significant uncertainty with all matters relating this epidemic. This includes uncertainty in data, assumptions and many epidemiological and medical matters relating to this disease. It is therefore likely that the model will not be accurate. It is also reflected in the wide confidence intervals of many of the results of the model.

Some researchers have gone as far as to say that it is not possible to accurately estimate the mean number of fatalities of pandemics based on data ([14] & [15]). This would indicate the goal of the paper may be futile. This paper has however attempted to allow for uncertainty using the Bayesian approach here which results in at least some distribution of possible results, some of which are definitely catastrophic and some of which are less severe.

10.1 Limitations

This analysis has various limitations. We highlight a few below but there are many—perhaps the most important being the lack of good quality data.

- The model is based on somewhat limited data for South Africa. In particular the effects of lockdown in some of the provinces may be simple “extrapolations” using the fixed effect (α) and shared prior of $R_{o,m}$ from other provinces. As these fits are updated with new data though the model should be more accurate for each province resulting in better estimates given.
- The model is a simplistic high-level single population model. It contains no differentiation by age groups or any further details. This is required to make the hierarchical models manageable. It still seems to provide useful information on potential movements in the epidemic. However, this homogeneity may result in overestimation of attack rates and hence deaths.

- Similar to the above point, various groups even of the same age may have varying susceptibility to the virus. Varying susceptibility may reduce the effective herd immunity threshold of the disease.
- The model does not take account of changes in many interventions that do not impact general mobility such as social distancing, handwashing and testing, tracing and isolation. The model does take account of the average impact of these interventions over time, but changes in these interventions will not be taken account of. It could be argued that the AR(2) autoregressive term is taking account of these factors.
- The paper does take account of mask wearing, workplace screening and other NPIs introduced with the start of level 4 lockdown. This is done approximately and only measures the implementation of the regulation and not actual compliance with the regulations. These parameters may also not be capturing only this effect and may include other effects that occurred at the same time. Similarly these variables may not be fully capturing increased compliance with these regulations over time.
- Google mobility data may not effectively cover the South African population. It is unclear how uniform the coverage is in, say, lower income populations who may have limited data and hence may not contribute significantly to the mobility indexes. Smartphone penetration in South Africa is estimated to exceed 90% [16] though. So the representative nature may be driven mainly by access to data as well as how many people have enabled the feature used to generate the data (location history). Changes in mobility may have a different impact on various segments of the population with unexpected results.
- IFR assumptions may be wrong:
 - It is based on Chinese data and may therefore be higher or lower in SA. The IFR employed though has been found consistent with various sources [17].
 - It does not consider health system capacity. In adverse scenarios the IFR is likely to increase if health system capacity is exceeded.
- The model assumes an average of 90% of excess deaths are attributable to COVID-19 directly. This may be wrong or in fact change over time.
- The autoregressive term may change in future. In current projections the term is kept constant going forward. If the term changes in future it may result in increased infections and deaths that cannot be projected by the model as is.
 - In times of accelerating spread this would probably result in over-prediction of the epidemic.
 - In times of slowing spread this may result in under-estimation of future infections.
- Given the calibration to deaths, this model is unlikely to be able to predict a “second wave” until deaths start increasing.

10.2 Pattern of spread

The epidemic starts in most provinces with a relatively high $R_{t,m}$ (based on assumptions). This then reduces as mobility reduces due to lockdown. In some provinces

an additional reduction can be seen at the time mask wearing and workplace screening is introduced. However in most provinces the epidemic starts spreading during lockdown due to increased mobility and also increases in the weekly autoregressive term. This term is adjusting to fit the increased numbers of deaths that are observed.

During the last number of weeks however the spread slows (without a slowdown in mobility) and the weekly adjustment has to be decreased in most provinces as this decrease is not explained by herd immunity or other interventions.

- The reason for this slowdown in spread is not clear. Some possible interpretations are:
- Changes in behaviour of the population in response to the epidemic, for example mask-wearing compliance or distancing.
 - Variation in contact rates of individuals may result in the virus infecting those with higher contact rates first, and possibly slowing down as it reaches those with lower contact rates. This may be due to some subsets of the population being more willing or better able to avoid getting infected.
 - Inherent variation in susceptibility between individuals resulting in a slowdown in infection rates as the pool of lives with high susceptibility is exhausted.

10.3 Impact of interventions

From the results it is clear that $R_{t,m}$ in some provinces has reduced from the starting values of $R_{0,m}$ and this has slowed the spread of the epidemic in South Africa, saving lives. However, in Europe the lockdowns have been able to reduce R_t below 1 for various countries [18]. As shown here, South Africa's lockdown and other interventions have not been as successful as in most the European countries shown in [18].

10.4 Further relaxation of lockdown

Projections on current mobility levels already result in significant peaks in deaths. Further relaxation of lockdowns will result in increases in mobility, which would worsen the reproduction numbers.

Based on the modelling it is expected to result in roughly 6 000 more deaths by the end of the year. This ignores other impacts such as ICU availability and the impact on deaths from other causes which would be expected to increase these figures, but also the impact of the homogeneity assumption which might be exacerbating this number.

10.5 Reducing mobility

Reducing mobility could result in roughly 4 000 fewer deaths by the end of the year.

10.6 Relevance of IFR

The IFR is not treated as a parameter but as a constant with random noise. Changes to the IFR will change the modelled infections that correlate with the observed deaths. Sensitivity to the IFR could be modelled.

If the IFR is higher than what is assumed (which would seem likely), we would expect

lower attack rates at present and thus the epidemic could result in a higher ultimate death toll as well. A lower than assumed IFR would imply more people have been infected already and we would be further along in the epidemic.

10.7 Using mobility data

Using mobility data is useful not only to measure the impact of government interventions but also to include the societal response to those interventions (as they affect mobility). This means that changes in the reaction to new regulations can be modelled. It may also be useful going forward as many new regulations are introduced possibly at a provincial level to summarise the impact of interventions numerically. From this research it appears that mobility cannot explain the pattern of spread of the epidemic fully.

10.8 Untangling mobility and other interventions

This paper projects mobility changes forward. We can see that as the model assigns some value to initiatives introduced at the start of level 4 around mask wearing and workplace screening. These reduce the impact of increasing mobility such that in the Western Cape the $R_{t,m}$ actually reduced in level 4 compared to level 5.

10.9 Health system capacity

When health system capacity is under pressure in various provinces it would seem likely that the mortality rate of infected individuals in those provinces would increase. This model needs to take account of that for two reasons:

1. Given the relationship between deaths and infections in this model a sudden increase in deaths due to capacity constraints would incorrectly be treated as a rise in infections.
2. To project the deaths forward accurately, this model needs to allow for future capacity constraints as well.

The model does not currently do this.

10.10 Poor fit in some provinces

Some provinces show a slight over-projection of the excess deaths in the last week or two where data are available. It would seem likely that the AR(2) term is not reducing fast enough for the drop in excess deaths.

10.11 Autoregressive term

Previous versions of the model over-predicted deaths to a large degree. The introduction of the autoregressive process as per [3] to model further residuals has improved this aspect of the model. This term may be reflecting other NPIs, general population awareness and behaviour adjustment or other unmodelled disease effects. However these should be treated as an error term in that they do not aid understanding of the disease and show deficiencies in this model.

10.12 Further work

Along with the above, the author intends to investigate the following:

- Search and include further indexes that may track other interventions to model their effect.
- Analyse the sensitivity of results to changes in IFR.
- Consider incorporating health system constraints and their impact on IFRs.
- Allowances for improvements in IFRs due to improved treatment protocols and higher health system capacity in some provinces.
- Consider adjusting completeness of reporting over time.

11. DETAILED ASSUMPTIONS AND METHODOLOGY

The model assumes that the current reproduction number, $R_{t,m}$, is a function of the initial reproduction number, $R_{0,m}$, and changes in mobility over time. It then calculates infections as a function of $R_{t,m}$ over time and then, using these infections, calculates deaths from the infections based on a distribution of time to death.

Various prior distributions are assumed. The model structure is similar to that in [1] and is briefly documented below. The parameters are estimated jointly using a single hierarchical model covering all provinces. This means that data in different provinces are combined to inform all parameters in the model. As per [1], fitting was done in R using Stan with an adaptive Hamiltonian Monte Carlo sampler.

11.1 The Model

The model is documented here but is similar to [1]. The model assumes a base reproduction number ($R_{0,m}$) for each province (m) and then it models future values of the reproduction number using mobility indexes as follows:

$$R_{t,m} = R_{0,m} \cdot 2 \cdot \phi^{-1} \left(-(\alpha + \alpha_m^s) I_{t,m}^\alpha - (\beta + \beta_m^s) I_{t,m}^\beta - \varepsilon_{w_m(t),m} \right)$$

Here:

- t is the day.
- ϕ^{-1} is the inverse logit function.
- $I_{t,m}^\alpha$ is the average mobility indexes from Google mobility [5] data for each province over time. Typical movements are represented by increases or decreases to the above.
- α is the effect of $I_{t,m}^\alpha$ on $R_{t,m}$ independent of province.
- α_m^s is the province specific effect of $I_{t,m}^\alpha$ on $R_{t,m}$.
- $I_{t,m}^\beta$ is the effect of other interventions introduced with level 4 lockdown. $I_{t,m}^\beta$ is 0 before the start of level 4 and remains at 1 after the start of level 4 lockdown.

- β is the effect of $I_{t,m}^\beta$ on $R_{t,m}$ independent of province.
- β_m^s is the province specific effect of $I_{t,m}^\beta$ on $R_{t,m}$.
- $\varepsilon_{w_m(t),m}$ is a AR(2) process modelled in weeks. It is centred on 0. This term represents unexplained variation in the model over time.
- $w_m(t)$ is the week index associated with t for province m .

Infections are modelled as:

$$c_{t,m} = S_{t,m} R_{t,m} \sum_{\tau=0}^{t-1} c_{\tau,m} g_{t-\tau} \text{ where } S_{t,m} = 1 - \frac{\sum_{i=0}^{t-1} c_{i,m}}{N_m}.$$

Infections, $c_{t,m}$ at time t are a function of proportion of population not yet infected ($S_{t,m}$), the reproduction number ($R_{t,m}$) and infections prior to that $c_{\tau,m}$ as well as an infectiousness curve $g_{t-\tau}$. N_m is the population in province m .

Deaths, $d_{t,m}$ are modelled as:

$$d_{t,m} = ifr_m^* \sum_{\tau=0}^{t-1} c_{\tau,m} \pi_{t-\tau}$$

Here:

- ifr_m^* is the average IFR in the province (with random noise added).
- $\pi_{t-\tau}$ is a distribution of time to death since infection.

It is assumed that COVID-19 deaths are a proportion of excess deaths (ψ_m). We calculate weekly excess deaths as:

$$d'_{t,m} = \frac{\sum_{week} d_{t,m}}{\psi_m}$$

Here t' represents time in weeks rather than days.

Weekly excess deaths $D'_{t',m}$ are assumed to have the following distribution:

$$D'_{t',m} \sim \text{Negative Binomial} \left(d'_{t',m}, d'_{t',m} + \frac{d_{t',m}^2}{\phi} \right)$$

If $Y \sim N(\mu, \sigma)$ then we define N^+ to mean the distribution of $|Y| \sim N^+(\mu, \sigma)$. We assume that:

$$\phi \sim N^+(0, 5)$$

Then: $d'_{t',m} = E(D'_{t',m})$.

11.2 Prior assumptions and random noise

The following assumptions are taken as is from [1] except where indicated. We add random noise to the IFR as follows:

$$ifr_m^* = ifr_m \cdot N(1, 0.1)$$

α and β are distributed as follows:

$$\alpha \sim N(0, 0.5)$$

$$\beta \sim N^+(0, 1)$$

The above distribution for β has been increased to reflect [8] which indicated an effect of mask wearing. In this paper a relative risk of 0.56 is indicated for mask wearing in non-health-care settings. The above prior assumption would indicate a 0.62 expected relative risk. β is not a parameter in the model used in [1].

Then for the provincial-specific index effects we use:

$$\alpha_m^s \sim N(0, \gamma^\alpha) \text{ with } \gamma^\alpha \sim N^+(0, 0.5)$$

$$\beta_m^s \sim N(0, \gamma^\beta) \text{ with } \gamma^\beta \sim N^+(0, 0.5)$$

β_m^s is a parameter in the model used in [1].

The $R_{0,m}$ are defined to be distributed normally as follows:

$$R_{0,m} \sim N(3.28, \kappa) \text{ with } \kappa \sim N^+(0, 0.5)$$

$\varepsilon_{w_m(t),m}$, the weekly province specific effect above is as described in [3]:

- It is modelled as an AR(2) process.
- It has a stationary standard deviation σ_w .
- $\varepsilon_{1,m} \sim N(0, \sigma_w^*)$
- $\varepsilon_{w,m} \sim N(\rho_1 \varepsilon_{w-1,m} + \rho_2 \varepsilon_{w-2,m}, \sigma_w^*)$
- ρ_{01} and ρ_{02} are independent with $\rho_{01} \sim N(0.8, 0.05)$ and $\rho_{02} \sim N(0.1, 0.05)$ both conditioned to be in $[0, 1]$.
- And $\sigma_w \sim N^+(0, 0.2)$
- Then standard deviation of the weekly updates is $\sigma^* = \sigma_w \sqrt{1 - \rho_1^2 - \rho_2^2 - 2\rho_1^2 \rho_2 / (1 - \rho_2)}$

Infectiousness follows this distribution:

$$g \sim \text{Gamma}(6.5, 0.62)$$

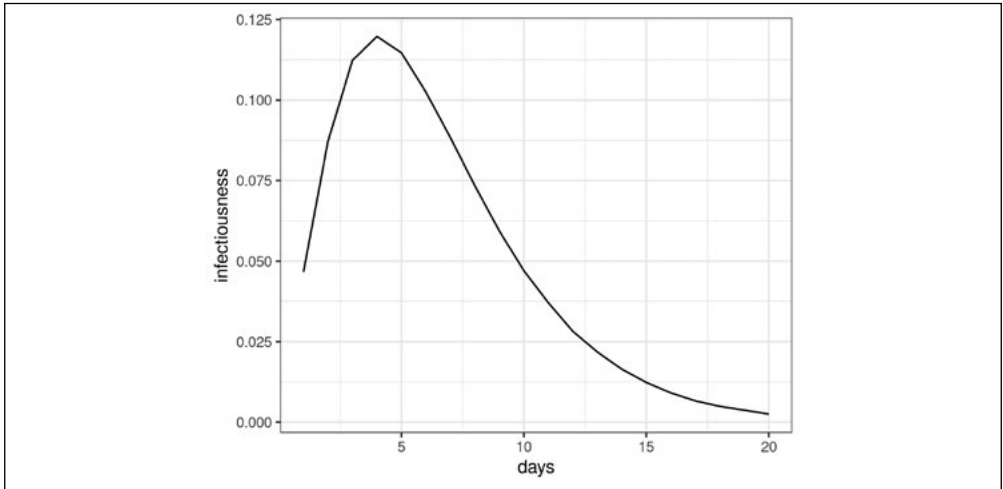


FIGURE 72 Infectiousness / serial interval

Time to death follows this distribution:

$$\pi \sim \text{Gamma}(5.1, 0.86) + \text{Gamma}(17.8, 0.45)$$

The distribution of time to death is plotted in Figure 73.

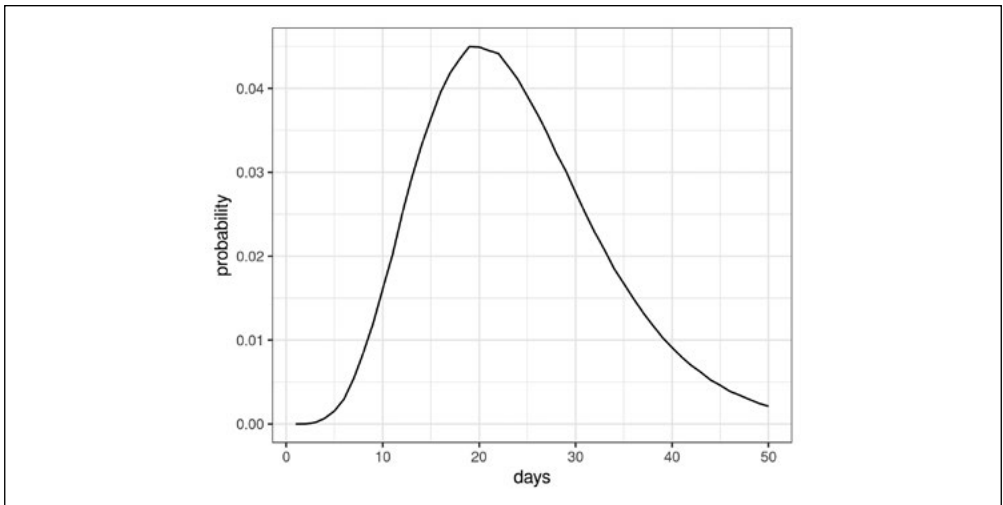


FIGURE 73 Distribution of time to death since infection

The above implies an average time to death of 23 days.

We assume a beta distribution for ψ_m :

$$\psi_m \sim B(27, 3)$$

The above distribution has a mean of 0.9, which is an assumption based on early indications that the majority of excess deaths are due to COVID-19. The distribution above implies that the probability ψ_m is below 0.8 is roughly 0.05. It is estimated that reported deaths in Western Cape is 70–80% of excess deaths, indicating a possible lower bound on the proportion of excess deaths that are due to COVID-19.

11.3 Death and confirmed case data

Death and case data were used from [19]. This data set contains provincial case and death data as released by the National Institute of Communicable Diseases (NICD). These data are shown for comparison purposes only as the model calibrates against excess deaths.

11.4 Excess deaths

Excess deaths are taken from [4]. These excess deaths only contain deaths from 6 May 2020 and later (depending on when excess deaths were observed to start in each province). Reported COVID-19 deaths prior to these dates were included and treated as excess deaths.

We plot weekly excess deaths for each province in Figure 74.

11.5 Infection fatality rates (IFR)

The IFR (ifr_m) for each province was calculated using the output of the *squire* R package [20]. It produces age-specific infection attack rates (IAR), infections and deaths. The per age band IFR were used together with the per age band IAR and these were applied to provincial populations [21]. The IFRs from *squire* package are based on [22], [23] and [20].

The projection was done using the default parameters for South Africa and the resultant attack rate (a_x) and IFR (ifr_x) for each 5-year age band was obtained (x).

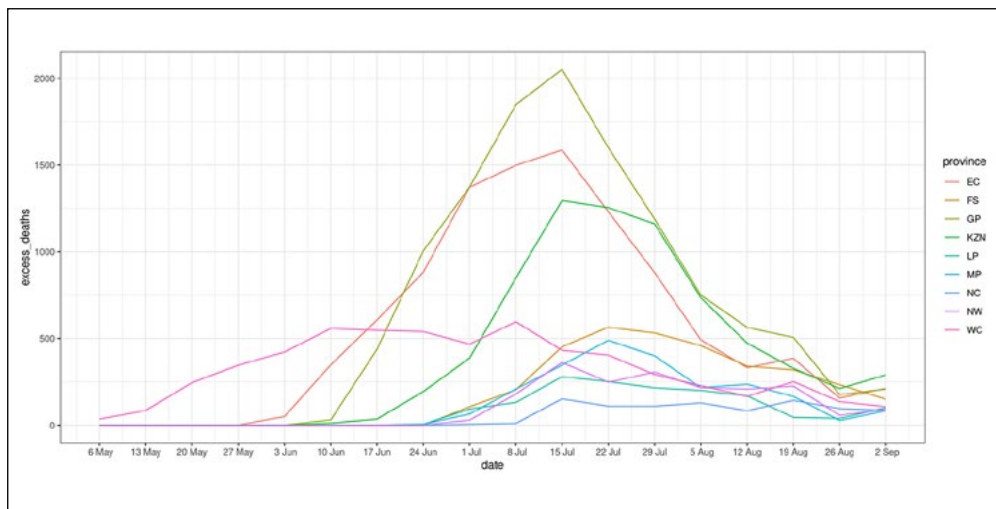


FIGURE 74 Excess deaths by province and week

Additionally HIV prevalence by age band and province, $i_{x,m}^{HIV}$ was taken into account from [24] as well as results from [25]. In [25] it is shown that lives with HIV have higher COVID-19 mortality once infected. Based on these results we assume three times COVID-19 mortality below 55 and double mortality from age 55 upwards (m_x^{HIV}).

These a_x and ifr_x and HIV adjustments were then applied to the South African population per province and age band as per [21] and weighted IFRs calculated as follows.

$$ifr_m = \frac{\sum_x N_{x,m} \cdot a_x \cdot ifr_x \left((1 - i_{x,m}^{HIV}) + i_{x,m}^{HIV} \cdot m_x^{HIV} \right)}{\sum_x N_{x,m}}$$

Here $N_{x,m}$ is the population in a particular province (m) and age band (x).

Below we tabulate the resultant ifr_m :

TABLE 5 Infection fatality rate by province

Province	IFR
EC	0.64%
FS	0.55%
GP	0.44%
KZN	0.47%
LP	0.5%
MP	0.45%
NC	0.54%
NW	0.49%
WC	0.5%
South Africa	0.49%

The differences between provinces reflect the different age profiles in those provinces as per [21]. This seems low compared to the IFRs in [17], but may be reasonable given the younger profile of the South African population.

11.6 Mobility indexes

Google has released aggregated mobility data for various countries and for sub-regions in those countries. These data are generated from devices that have enabled the Location History in Google Maps. This feature is available both on Android and iOS devices but is off by default.

For South Africa, these data contain the mobility indexes for each province. These are described in [5] as follows:

- **Grocery & pharmacy** Mobility trends for places like grocery markets, food warehouses, farmers' markets, specialty food shops, drug stores, and pharmacies.
- **Parks** Mobility trends for places like local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens.
- **Transit stations** Mobility trends for places like public transport hubs such as subway, bus, and train stations.
- **Retail and recreation** Mobility trends for places like restaurants, cafés, shopping centres, theme parks, museums, libraries, and movie theatres.
- **Residential** Mobility trends for places of residence.
- **Workplaces** Mobility trends for places of work.

The above categories measure changes in mobility relative to a baseline established before the epidemic. For example, -30% implies a 30% reduction in mobility from pre-COVID-19 mobility.

As per [7] these data were combined into an average mobility index for each province which was an average of all mobility indexes excluding:

- Residential
- Parks

In [1] three indexes were used (residential, transit and the rest). This was reduced for this paper due to limited data.

12. CONCLUSION

This paper has set out to show how a Bayesian model can be fitted to the South African data. From the process it becomes clear that there remains significant uncertainty with all matters relating this epidemic. It is therefore likely that the model will not be accurate.

This paper has however attempted to allow for uncertainty using the Bayesian approach, which results in at least some distribution of possible results, some of which are definitely catastrophic and some of which are less severe.

It is hoped that this is a useful contribution to the ongoing research on the epidemic.

DISCLAIMER

The views in this document represent those of the author and may not represent those of Gen Re. Also note that given the significant uncertainty involved with the parameters, data and methodology, care should be taken with these numbers and their use.

REFERENCES

- [1] M Vollmer et al. (2020). "Report 20: A sub-national analysis of the rate of transmission of COVID-19 in Italy," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-20-italy/>
- [2] T Mellan et al. (2020). "Report 21: Estimating COVID-19 cases and reproduction number in Brazil," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-21-brazil/>
- [3] H Unwin et al. (2020). "Report 23: State-level tracking of COVID-19 in the United States," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-23-united-states/>
- [4] D Bradshaw, R Laubscher, R Dorrington, P Groenewald & T Moultrie (2020). "Report on weekly deaths in South Africa 1 January–8 September 2020 (Week 36)," Burden of Disease Research Unit, South African Medical Research Council [Online]. Available: <https://www.samrc.ac.za/sites/default/files/files/2020-09-09/weekly1September2020.pdf>
- [5] Google LLC (2020). "Google COVID-19 community mobility reports." [Online]. Available: <https://www.google.com/covid19/mobility/>
- [6] R Core Team (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. [Online]. Available: <https://www.R-project.org/>
- [7] P Nouvellet et al. (2020). "Report 26: Reduction in mobility and COVID-19 transmission," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-26-mobility-transmission/>
- [8] DK Chu et al. (2020). "Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis," *The Lancet*, vol. 395, no. 10242, pp. 1973–1987, Jun. 2020, doi: 10.1016/S0140-6736(20)31142-9. [Online]. Available: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)31142-9/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31142-9/fulltext)
- [9] D Bradshaw, R Laubscher, R Dorrington, P Groenewald & T Moultrie (2020). "Report on weekly deaths in South Africa 1 January–7 July 2020 (Week 27)," Burden of Disease Research Unit, South African Medical Research Council. [Online]. Available: <https://www.samrc.ac.za/sites/default/files/files/2020-07-19/WeeklyDeaths7July2020.pdf>
- [10] D Bradshaw, R Laubscher, R Dorrington, P Groenewald & T Moultrie (2020). "Report on weekly deaths in South Africa 1 January–18 August 2020 (Week 33)," Burden of Disease Research Unit, South African Medical Research Council. [Online]. Available: <https://www.samrc.ac.za/sites/default/files/files/2020-08-26/weekly18August2020.pdf>
- [11] D Bradshaw, R Laubscher, R Dorrington, P Groenewald & T Moultrie (2020). "Report on weekly deaths in South Africa 1 January–25 August 2020 (Week 34)," Burden of Disease Research Unit, South African Medical Research Council. [Online]. Available: <https://www.samrc.ac.za/sites/default/files/files/2020-09-03/weekly25August2020.pdf>
- [12] D Bradshaw, R Laubscher, R Dorrington, P Groenewald & T Moultrie (2020). "Report on weekly deaths in South Africa 1 January–1 September 2020 (Week 35)," Burden of Disease

- Research Unit, South African Medical Research Council. [Online]. Available: <https://www.samrc.ac.za/sites/default/files/files/2020-09-09/weekly1September2020.pdf>
- [13] GEP Box (1979). "Robustness in the strategy of scientific model building," in *Robustness in statistics*, Elsevier, pp. 201–236 [Online]. Available: <https://doi.org/10.1016/b978-0-12-438150-6.50018-2>
- [14] NN Taleb, Y Bar-Yam & P Cirillo (2020). "On single point forecasts for fat-tailed variables." [Online]. Available: <https://arxiv.org/abs/2007.16096>
- [15] P Cirillo & NN Taleb (2020). "Tail risk of contagious diseases," *Nature Physics*, vol. 16, no. 6, pp. 606–613, May 2020, doi: 10.1038/s41567-020-0921-x. [Online]. Available: <https://www.nature.com/articles/s41567-020-0921-x>
- [16] Independent Communications Authority of South Africa (2020). "The state of the ICT sector report in South Africa" [Online]. Available: <https://www.icasa.org.za/legislation-and-regulations/state-of-the-ict-sector-in-south-africa-2020-report>
- [17] G Meyerowitz-Katz & L Merone, "A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates" [Online]. Available: <https://www.medrxiv.org/content/10.1101/2020.05.03.20089854v3>
- [18] S Flaxman et al. (2020). "Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/>
- [19] V Marivate et al. (2020). "Coronavirus disease (COVID-19) case data – South Africa." Zenodo. [Online]. Available: <https://zenodo.org/record/3888499>
- [20] P Walker et al. (2020). "Report 12: The global impact of COVID-19 and strategies for mitigation and suppression," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-12-global-impact-covid-19/>
- [21] Statistics South Africa (2019). "Mid-year population estimates 2019," Republic of South Africa. [Online]. Available: <https://www.statssa.gov.za/publications/P0302/P03022019.pdf>
- [22] R Verity et al. (2020). "Estimates of the severity of coronavirus disease 2019: A model-based analysis," *The Lancet Infectious Diseases*, vol. 20, no. 6, pp. 669–677, Jun. 2020, doi: 10.1016/s1473-3099(20)30243-7. [Online]. Available: [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(20\)30243-7/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext)
- [23] N Ferguson et al. (2020). "Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand," Imperial College London. [Online]. Available: <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/>
- [24] L Johnson, R Dorrington & H Moolla (2017). "Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa," *Southern African Journal of HIV Medicine*, vol. 18, no. 1, p. 8. doi: 10.4102/sajhivmed.v18i1.694. [Online]. Available: <https://sajhivmed.org.za/index.php/hivmed/article/view/694>
- [25] M-A Davies (2020). "Brief update on surveillance of COVID-19 infections in the Western Cape." 03-Jun-2020. [Online]. Available: <http://www.medicine.uct.ac.za/covid19-echo-clinic>