Application of the Logistic Model to the COVID-19 Pandemic in South Africa and the United States: Correlations and Predictions

David H. Roberts[1]

Department of Physics MS-057
Brandeis University
Waltham, MA 02454-0911 USA
roberts@brandeis.edu

Revised 15 January 2022

ABSTRACT

We apply the simple logistic model to the four waves of COVID-19 taking place in South Africa over the period 2020 January 1 through 2022 January 11. We show that this model provides an excellent fit to the time history of three of the four waves. We then derive a theoretical correlation between the growth rate of each wave and its duration, and demonstrate that it is well obeyed by the South Africa data.

We then turn to the data for the United States. As shown by Roberts (2020a, 2020b), the basic logistic model provides only a marginal fit to the early data. Here we break the data into six "waves," and treat each one separately. For four of the six the logistic model is useful, and we present full results. We then ask if these data provide a way to predict the length of the ongoing Omicron wave in the US (commonly called "wave 4," but the sixth wave as we have broken the data up). Comparison of these data to those from South Arica, and internal comparison of the US data, suggest that this last wave will die out by about 15 days from the peak, or about 30 January 2022.

1. Introduction

Since the spring of 2020 a pandemic of infection of a novel coronavirus (SARS-CoV-2) has overspread the world. Epidemiologists have struggled to describe adequately this

[1]ORCID 0000-0001-5226-4540.

pandemic and to predict its future course. This is a complex undertaking, involving the mathematics of epidemiology and a huge amount of uncertain data that serve as input to the models. In this paper we examine the ability of a simple epidemiological model, the *logistic model,* to describe the course of the pandemic in the Republic of South Africa and the United States. After showing that it is adequate to the task for South Africa, we derive a theoretical correlation between the growth rate of each wave of the pandemic and its duration, and show that it is well obeyed for the South Africa data. Finally, we discuss the use of this correlation to predict the course of the fourth (Omicron driven) wave in the United States, and show that it should die out very soon.

The data are from the World Health Organization (WHO 2022).

2. The Logistic Model

2.1. Derivation

The exposition in this section is from Roberts (2020a).

A simple model for the evolution of an pandemic is based on the logistic differential equation. This describes an pandemic that begins with a small number f_0 of infected individuals, and subsequently spreads through a population. The motivation for this model is as follows. If the population of infected individuals as a function of time is $f(t)$, simple exponential growth with growth rate r is determined by the differential equation

$$\frac{df}{dt} = rf,$$

with the growing exponential solution

$$f(t) = f_0 e^{rt},$$

where $f_0 = f(0)$. This is what happens with an infinite pool of subjects. However, for a finite pool of subjects, as the population of infected individuals grows the number of subjects available to be infected gets smaller. This is taken into account by modifying the exponential differential equation to become

$$\frac{df}{dt} = r(1 - \frac{f}{K})f, \tag{1}$$

where K is the total available pool of individuals. The solution of this equation[2] is

$$f(t) = \frac{K f_0 e^{rt}}{K + f_0(e^{rt} - 1)}, \tag{2}$$

which satisfies the required limits $f(0) = f_0$ and $f(\infty) = K$. The time course described by Equation 2 is the familiar "S curve" used to describe bacterial growth and other phenomena (see Figure 3).

An analysis of the total number of cases as a function of time $f(t)$ is just one way to compare model and data. Instead we can examine the number of new cases per day as a function of time; this is the time derivative of $f(t)$.[3] This is easily found by substituting Equation 2 into Equation 1,

$$\frac{df}{dt} = f_0 K r e^{rt} \left[\frac{K - f_0}{((e^{rt} - 1)f_0 + K)^2} \right]. \tag{3}$$

In either case, for the logistic model the three parameters to be adjusted are f_0, K, and r.

2.2. COVID data for the United States and the Republic of South Africa

The plots in this section show the total and daily distributions of the number of cases of COVID-19 in the United States and in South Africa (WHO 2022). Day 1 is 03 January 2020.

[2]This is a version of the Bernoulli differential equation $f' = f(1 - f)$.

[3]Fitting the daily numbers is statistically the preferred procedure as the data points are independent, unlike those for daily totals.

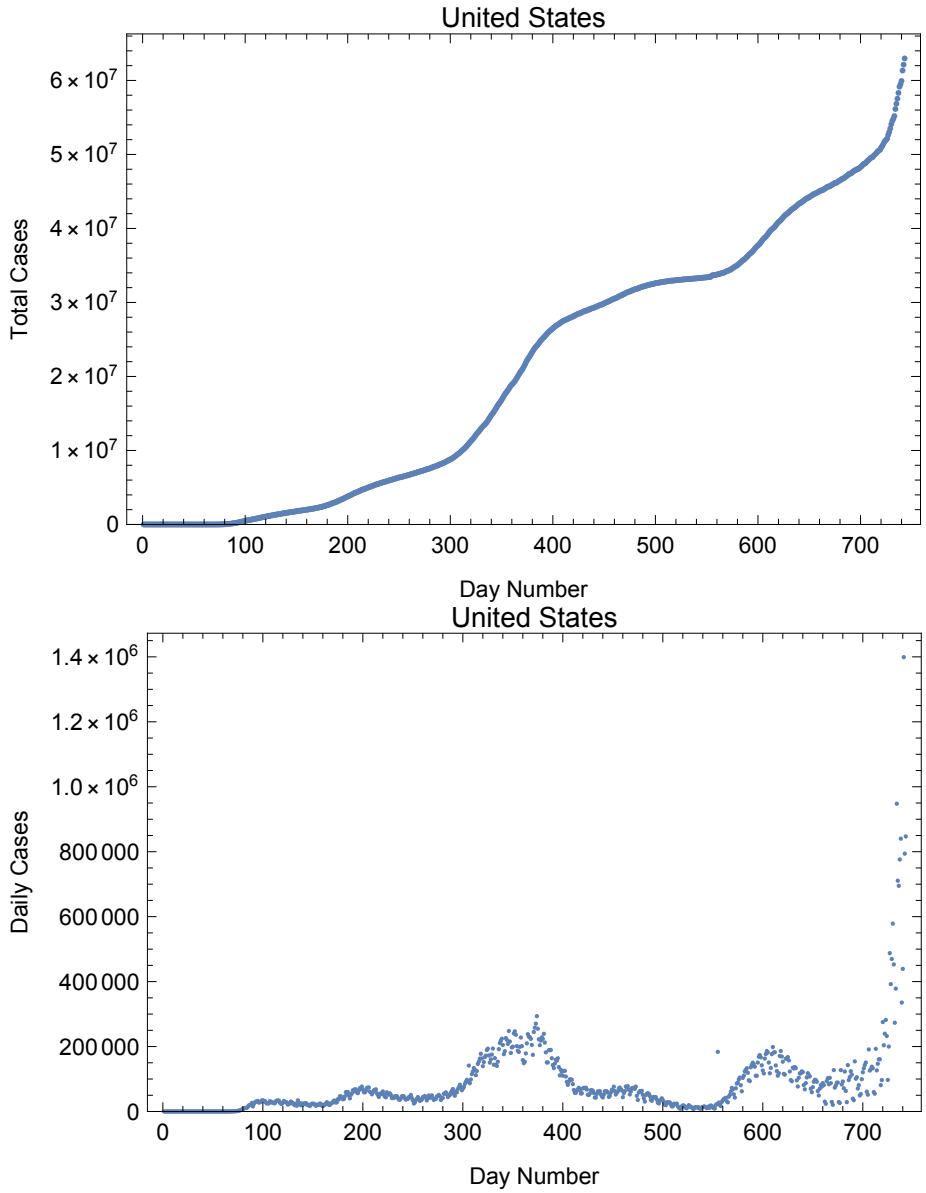

Fig. 1.— Cases data for the entire pandemic for the United States.

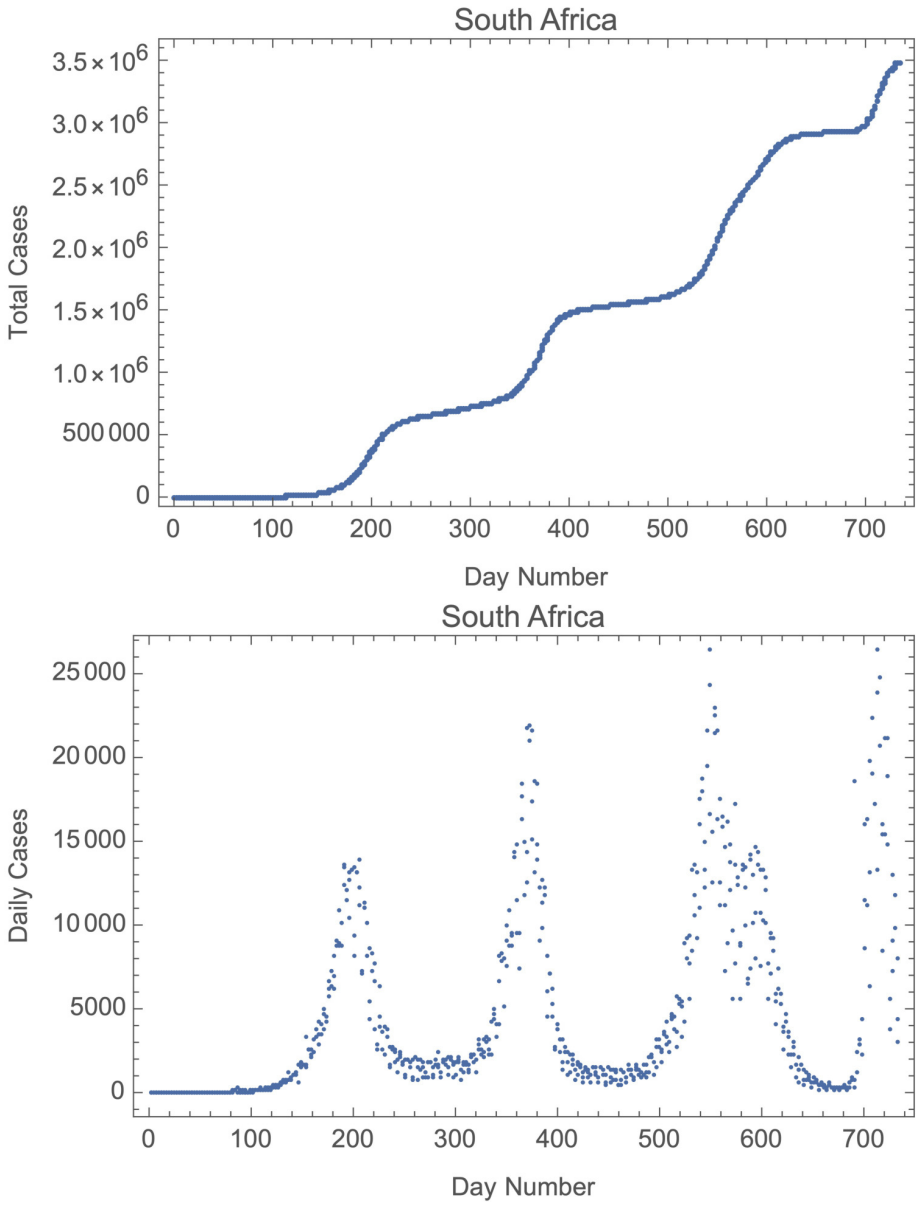

Fig. 2.— Cases data for the entire pandemic for the Republic of South Africa..

2.3. Fit to the COVID-19 Pandemic in the Republic of South Africa

When applied to the ongoing COVID-19 pandemic in the United States the logistic model does only a fair job of accounting for the actual history of the total number of infected individuals as a function of time in the first wave of COVID-19, which took place in the first half of 2020 (see Roberts 2020a). It fails even more spectacularly for later waves (Roberts 2020b).

In this section we apply the logistic model to three of the four waves in South Africa. Solutions were found by numerical minimization of either the sum of the squares of the differences between model and data (least squares fit, or "LSQ"), or the sum of the absolute values of the differences ("L1").

2.4. RSA Wave 1

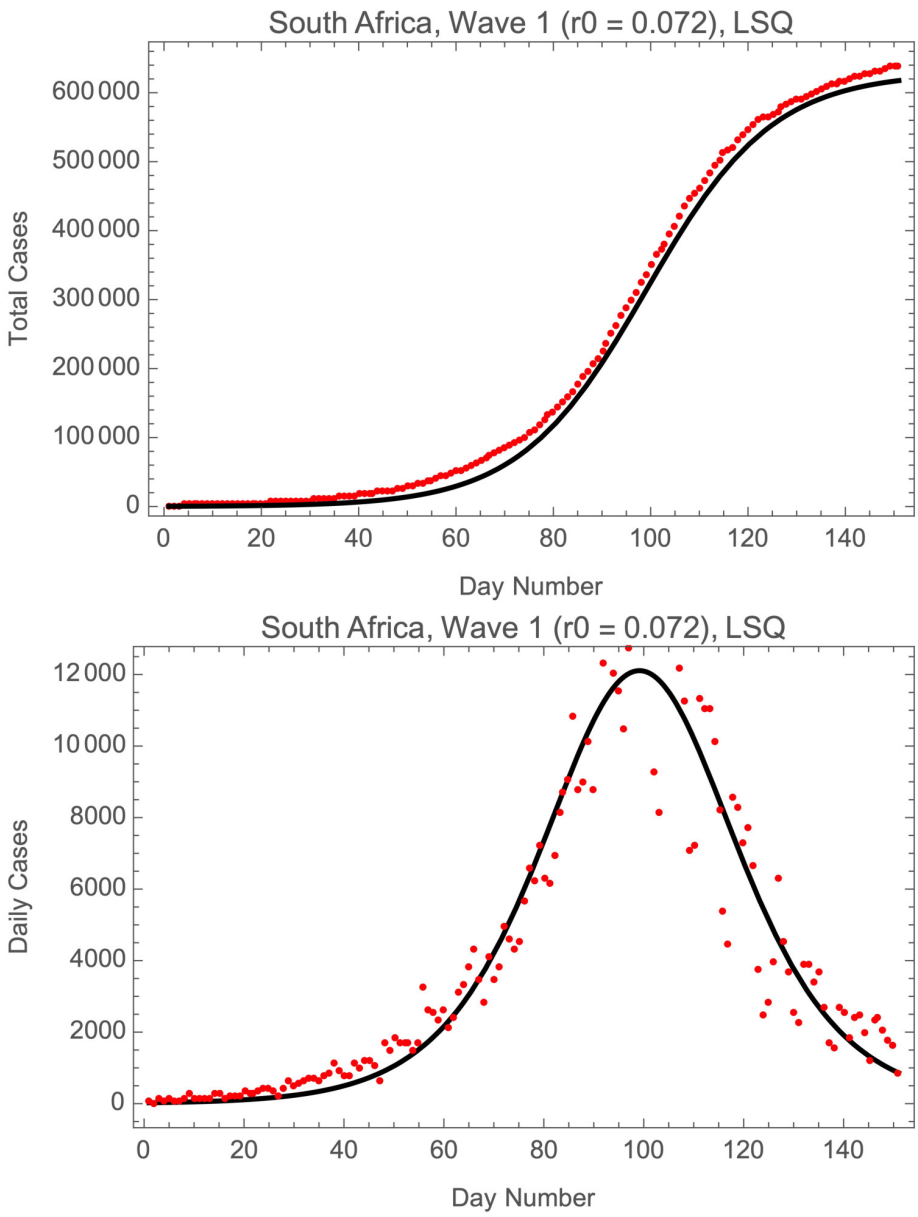

Fig. 3.— Fits of the logistic model to RSA cases of COVID-19 in wave 1. Day 1 is 11 April 2020.

2.5. RSA Wave 2

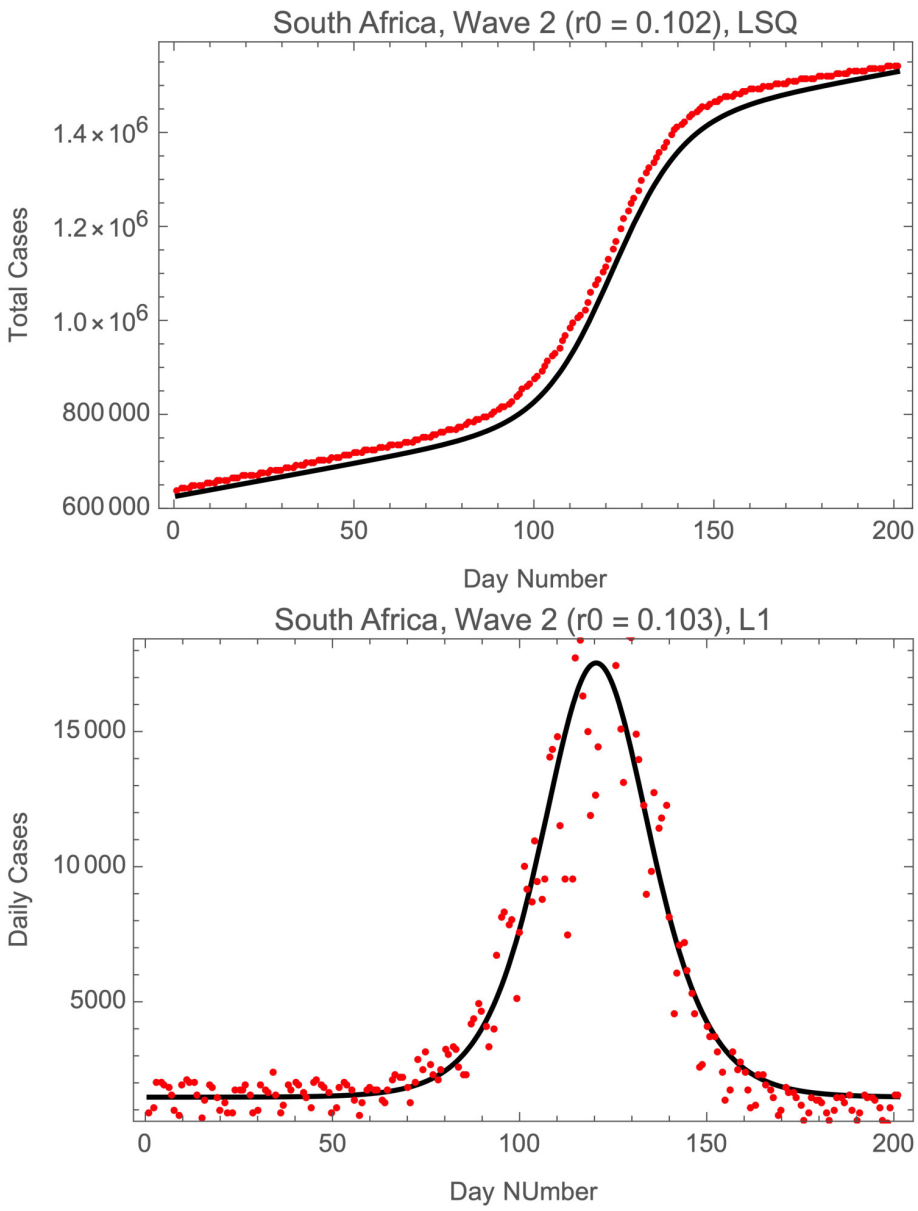

Fig. 4.— Fits of the logistic model to RSA cases of COVID-19 in wave 2. Day 1 is 8 September 2020.

2.6. RSA Wave 4

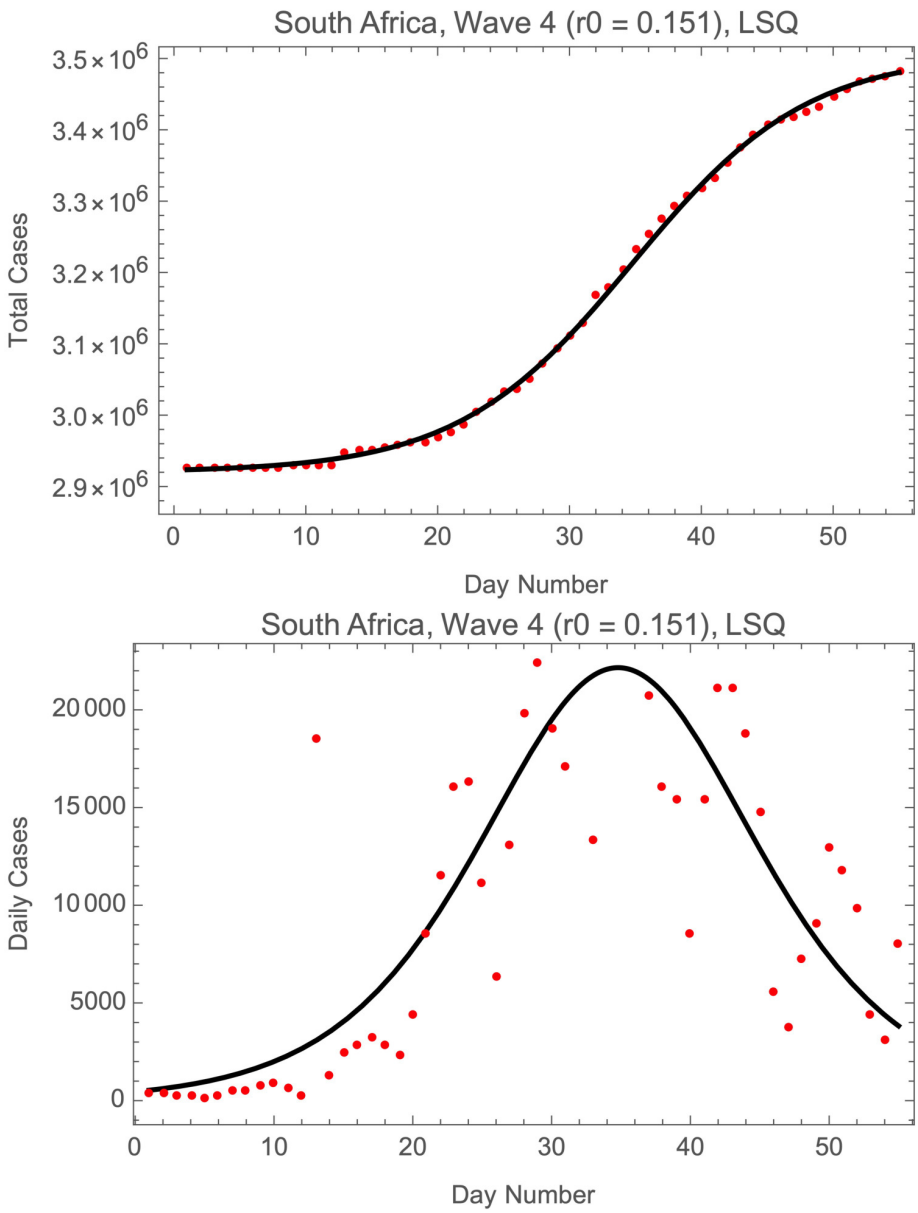

Fig. 5.— Fits of the logistic model to RSA cases of COVID-19 in wave 4. Day 1 is 12 November 2021.

3. Correlation of the Growth Rate and Duration for the COVID-19 Waves in the Republic of South Africa

It is natural that the growth rate and the duration of the waves of a pandemic should be correlated, higher growth rates leading to shorter durations. It can be shown (see the Appendix) that for a logistic model the daily cases distribution has a full-width-half-maximum given by

$$\mathrm{FWHM} = \frac{3.526}{r}. \tag{4}$$

Note that this result is independent of f_0 and K. The data and the prediction of Eq. 4 are compared in Fig. 6, where we see that the agreement is excellent.

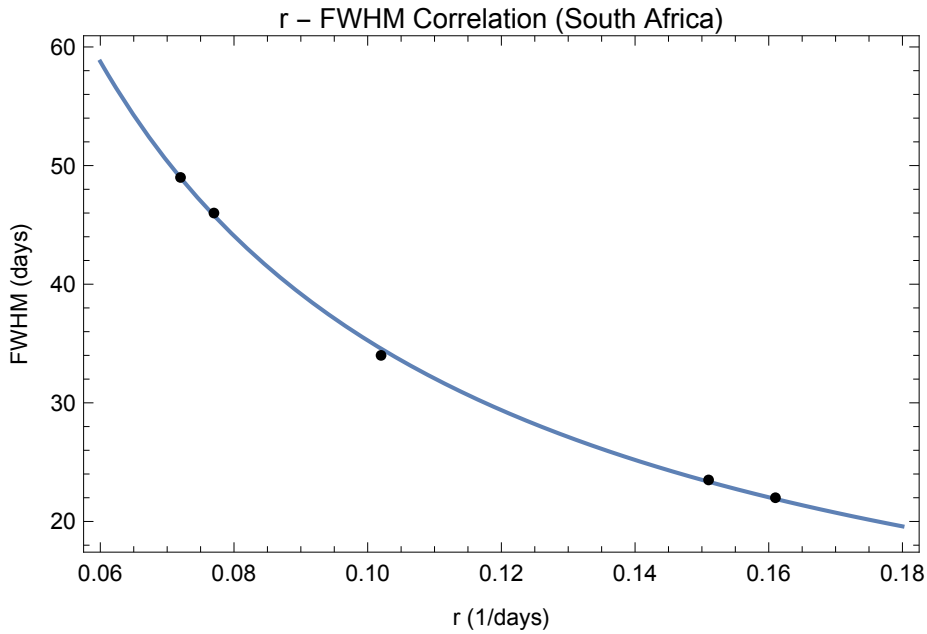

Fig. 6.— Correlation of the growth rate r and the FWHM of each wave for the COVID-19 pandemic in the Republic of South Africa. The curve is Eq. 4. The two pairs of points at the ends are from the LSQ and L1 fits to waves 1 and 4.

4. Six Waves in the United States

In this section we apply the logistic model to four of the waves of COVID-19 in the United States, using the techniques described above.

4.1. USA Wave 1

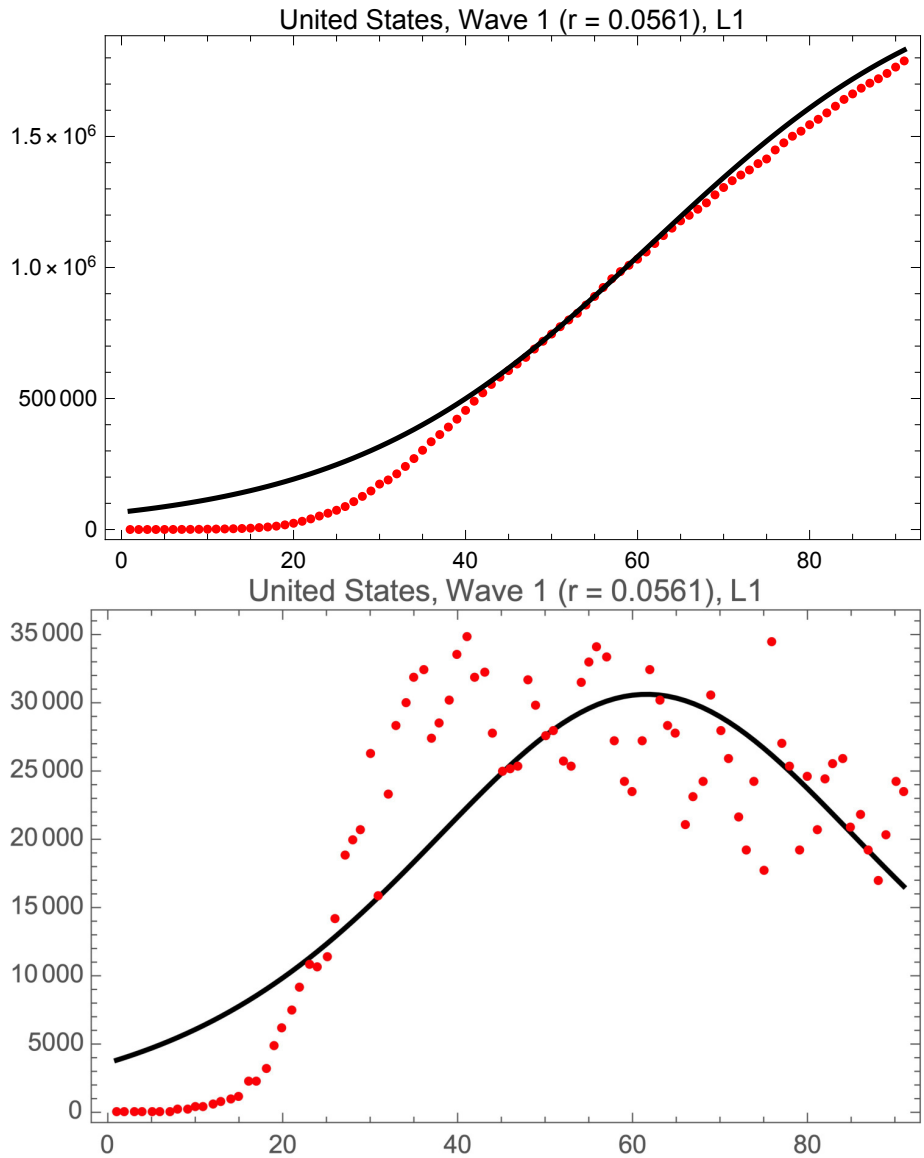

Fig. 7.— L1 fits of the logistic model to USA cases of COVID-19 in wave 1. Day 1 is 20 March 2020.

4.2. USA Wave 3

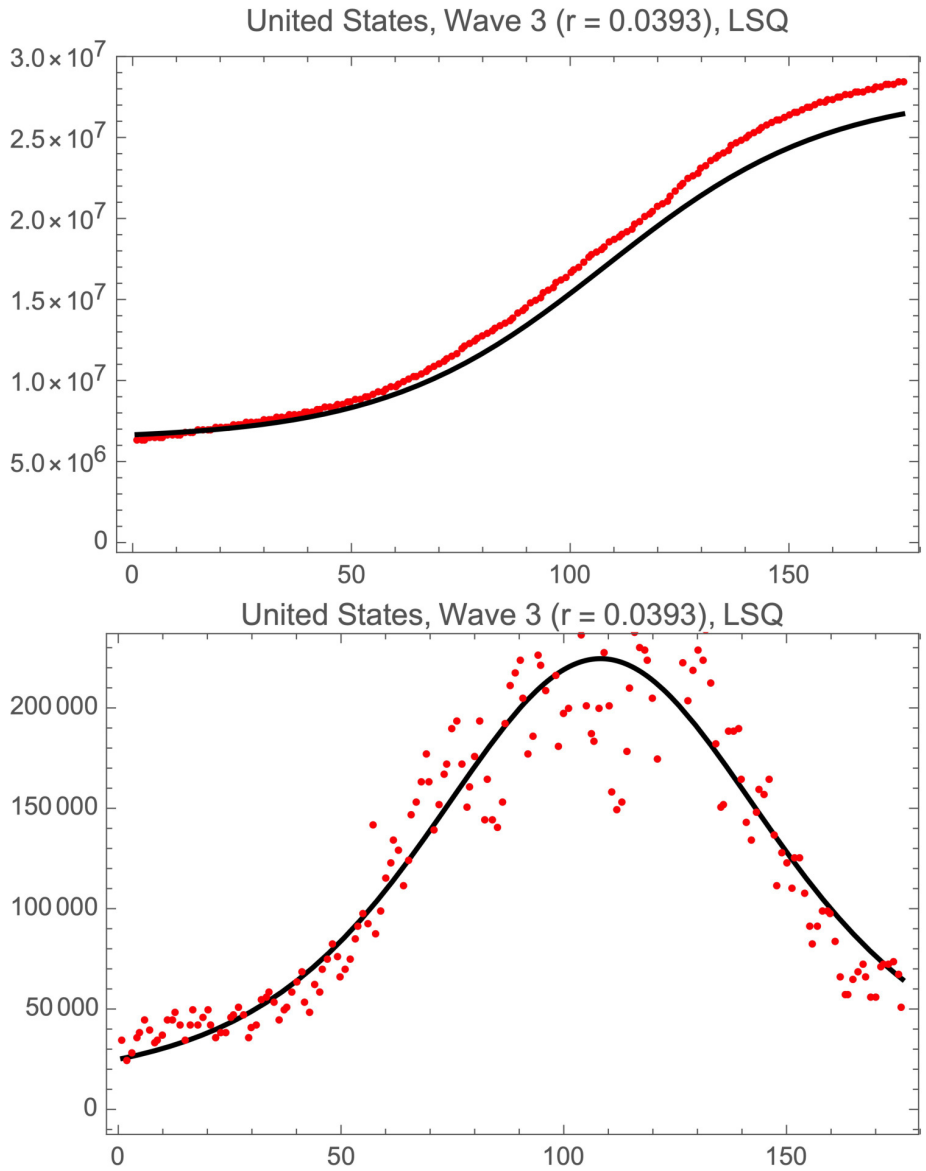

Fig. 8.— LSQ fits of the logistic model to total USA cases of COVID-19 in wave 3. Day 1 is 8 September 2020. is

4.3. USA Wave 5

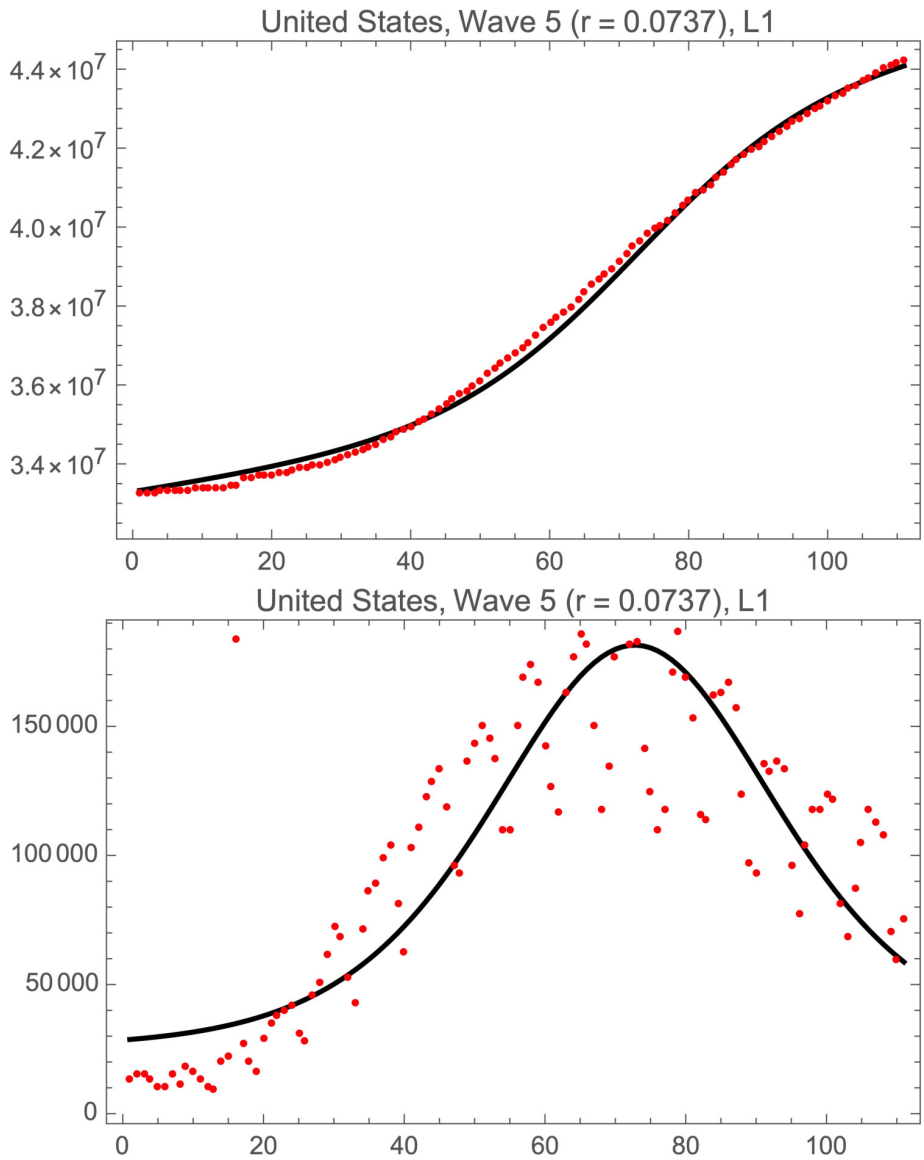

Fig. 9.— L1 fits of the logistic model to USA cases of COVID-19 in wave 5. Day 1 is 25 June 2021.

4.4. USA Wave 6

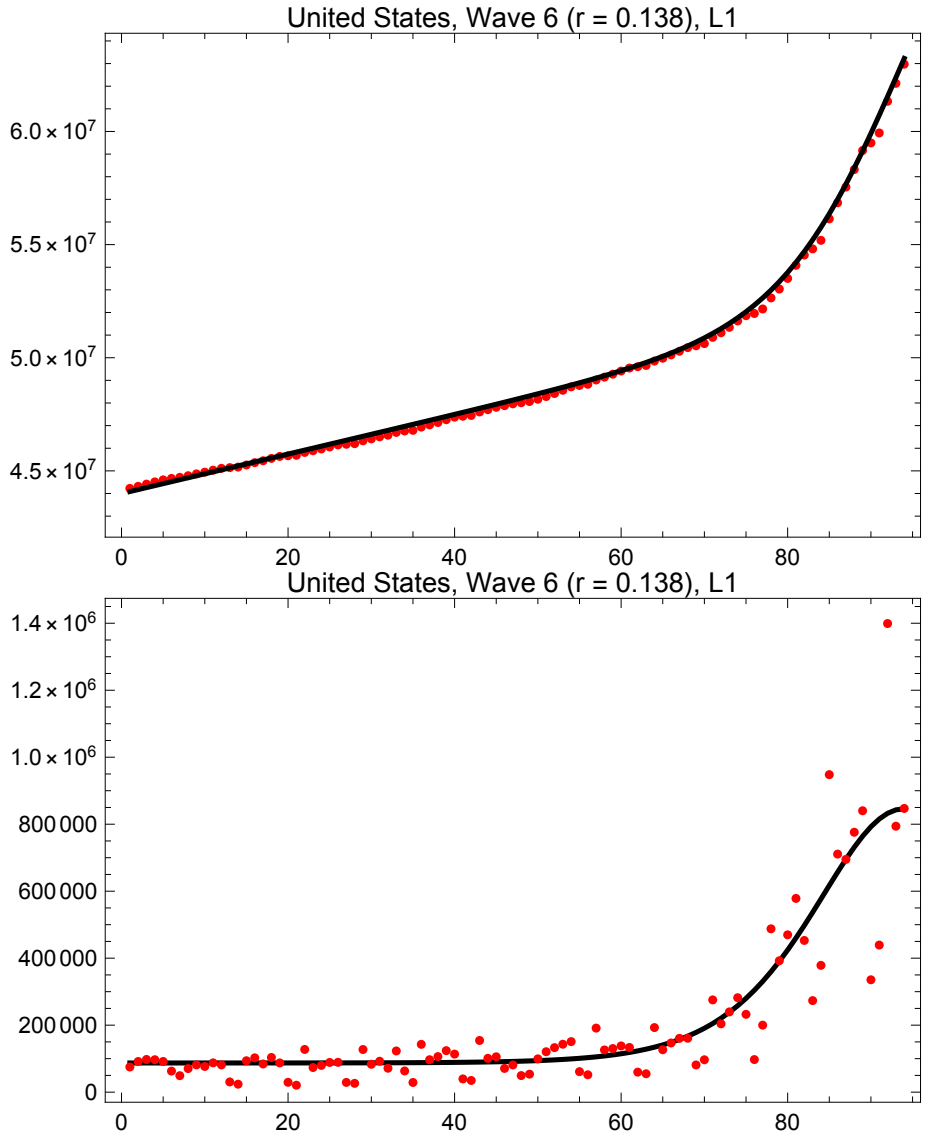

Fig. 10.— L1 fits of the logistic model to total USA cases of COVID-19 in wave 6. Day 1 is 13 October 2021.

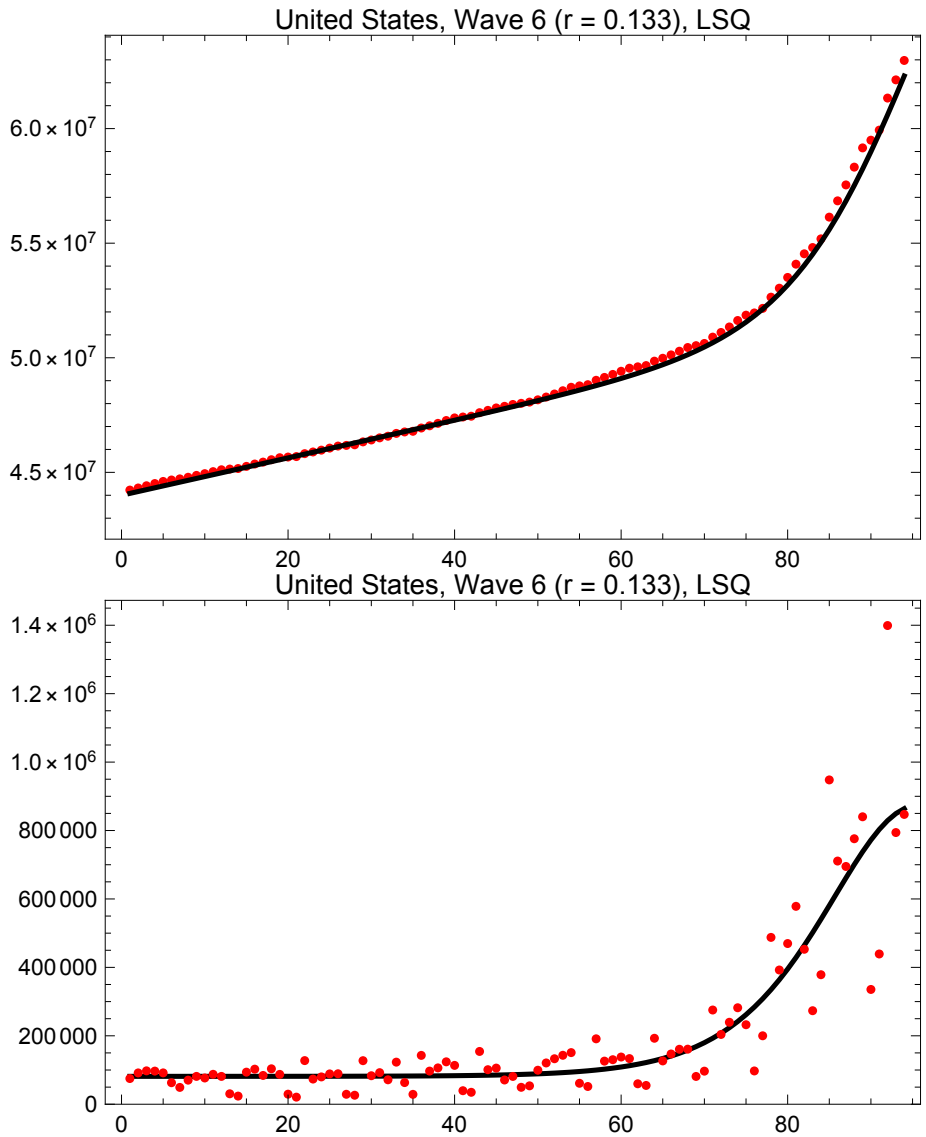

Fig. 11.— LSQ fits of the logistic model to USA cases of COVID-19 in wave 6. Day 1 is 13 October 2021.

5. Predictions for United States Wave 6

Fig. 12 shows the correlation between growth and duration of the various waves in the United States. The curve is Eq. 4, identical to the one in Fig. 6.

In Figs. 13 & 14 we show the predictions for wave 6 that follow from the L1 and LSQ solutions. The two provide equivalent fits to the data (see Figs. 10 & 11), so we average the predictions made from the two. Thus we expect the Omicron wave in the US to be dramatically reduced within about 15 days from now, roughly 30 January 2022. Needless to say, this would be very welcome.

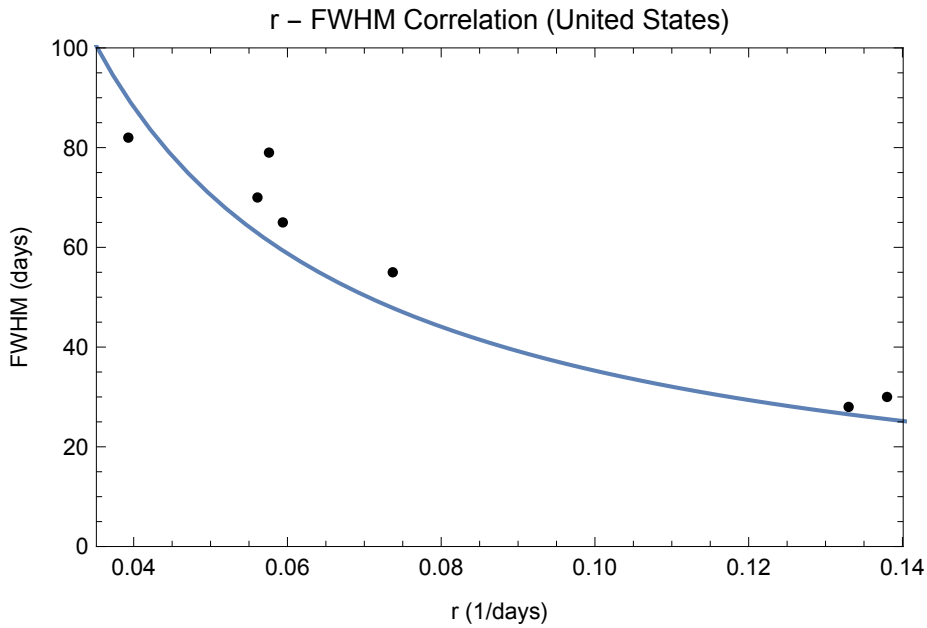

Fig. 12.— Correlation of the growth rate r and the FWHM of four waves of the COVID-19 pandemic in the United states. The curve is Eq. 4.

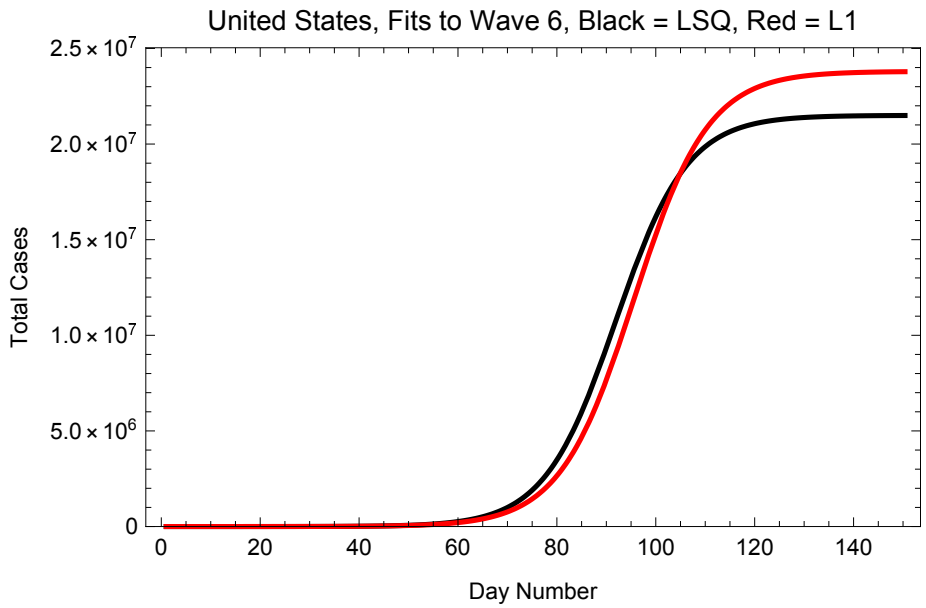

Fig. 13.— Predictions of the logistic model for the total USA cases of COVID-19 in wave 6.

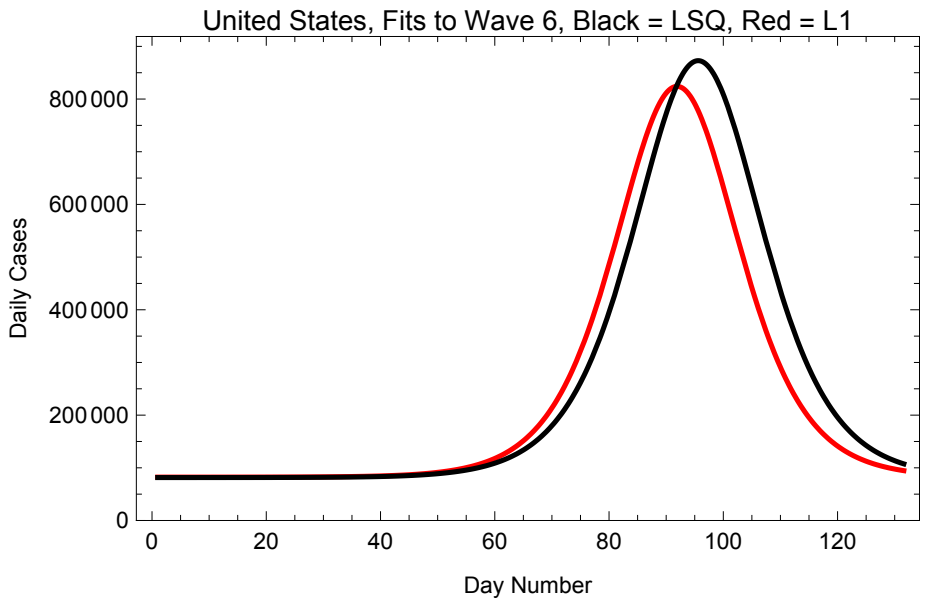

Fig. 14.— Predictions of the logistic model for the daily USA cases of COVID-19 in wave 6.

6. Conclusions

In this paper we have used the logistic model for epidemics to describe the COVID-19 outbreaks in the Republic of South Africa and in the United States. We find a universal analytic relationship between the growth rates and the durations of each wave, and this is closely followed by the (very clean) data from South Africa. When applied to the United States, the data are messier, but a tentative prediction is possible for the expected duration of the current Omicron wave in the US – it should be mostly over by about 15 days from now (roughly 30-January-2022).

7. Acknowledgements

We thank Brian Boyle, Mary Roberts, and Bob Sauer for their helpful comments and suggestions.

8. References

Roberts, D. H. (2020a), "Two New Models for Epidemics with Application to the COVID-19 Pandemic in the United States, Italy, and the United Kingdom," MedRxiv:2020.07.13.20152686v1

Roberts, D. H. (2020b), "A New Adaptive Logistic Model for Epidemics and the Resurgence of COVID-19 in the United States," MedRxiv:2020.07.17.20156109v1

World Health Organization (2022), world-wide COVID histories, version of 14 January 2022.

9. Appendix: Derivation of Equation 4

The differential distribution for the logistic function is[4]

$$\frac{df}{dt} = \frac{e^{rt} f_0 K(-f_0 + K) r}{\left((-1 + e^{rt}) f_0 + K\right)^2}, \tag{5}$$

and is derivative is

$$\frac{d^2 f}{dt^2} = -\frac{2e^{2rt} f_0^2 K(-f_0 + K) r^2}{\left((-1 + e^{rt}) f_0 + K\right)^3} + \frac{e^{rt} f_0 K(-f_0 + K) r^2}{\left((-1 + e^{rt}) f_0 + K\right)^2}. \tag{6}$$

[4]All the calculations were done with Mathematica.

This is zero at

$$t_{max} = \frac{\ln \frac{K-f_0}{f_0}}{r}, \tag{7}$$

and df/dt at this maximum is

$$\left(\frac{df}{dt}\right)_{max} = \frac{K(-f_0+K)^2 r}{\left(K + f_0\left(-1 + \frac{-f_0+K}{f_0}\right)\right)^2}. \tag{8}$$

Setting df/dt equal to half of this value and solving for the locations yields a complicated expression that reduces to

$$\text{FWHM} = \frac{3.526}{r}. \tag{9}$$

Note that neither f_0 nor K enters this expression.