Modelling the COVID-19 Fatality Rate in England and its Regions

Nigel Saunders

Thermotech Ltd., Surrey Technology Centre, The Surrey Research Park,
Guildford, Surrey GU2 7YG, U.K.

Abstract

A model to account for the fatality rate in England and its regions is proposed. It follows the clear observation that, rather than two connected waves, there have been many waves of infections and fatalities in the regions of England of various magnitudes, usually overlapping. The waves are self-limiting, in that clear peaks are seen, particularly in reported positive test rates. The present model considers fatalities as the data reported are more reliable than positive test rates, particularly so during the first wave when so little testing was done.

The model considers the observed waves are essentially similar in form and can be modelled using a single wave form whose final state is only dependent on its peak height and the start date of the particular wave. The basic wave form was modelled using the observed fatality rates for London, which unlike the other regions, exhibited almost completely as a single wave in the "first wave". Its form matches rather well with the "Do Nothing" model reported by Imperial College on 16th March 2020, but reduced substantially from its expected peak.

There are, essentially, only two adjustable parameters used in the model, the start date of the relevant wave and its height. The modelled fatalities for each wave are summated per day and a cumulative curve is matched to that reported. The minimal number of adjustable parameters, alongside the fact that the waves invariably overlap, provides highly stringent conditions on the fitting process.

Results are presented for each region for both the "first" and "second' waves. High levels of accuracy are obtained with R^2 values approaching 100% against the ideal fit for both waves. It can also be seen there are fundamental differences between the underlying behaviour of the "first" and "second" waves.

1. Introduction

While there are daily reports of the progress of COVID-19 infections in the UK, it seems little work has been done in trying to understand how the virus has actually spread in the UK. While regional R rates are regularly presented, infection or fatality rates over the whole period for the regions are rarely, if ever presented. However, by doing so, a far more detailed view of how the infection has spread is revealed. Fig.1 below shows reported population adjusted infection rates for the regions of England from the beginning of the pandemic, up to 1st January 2021 [(1)]. A 7-day average of cases by specimen date (*averaged at midweek*) is used and the last 5 days not used due to likely incompleteness as advised by the GOV.uk site.

1. UK Government website for data and insights on Coronavirus (COVID-19) https://coronavirus.data.gov.uk/

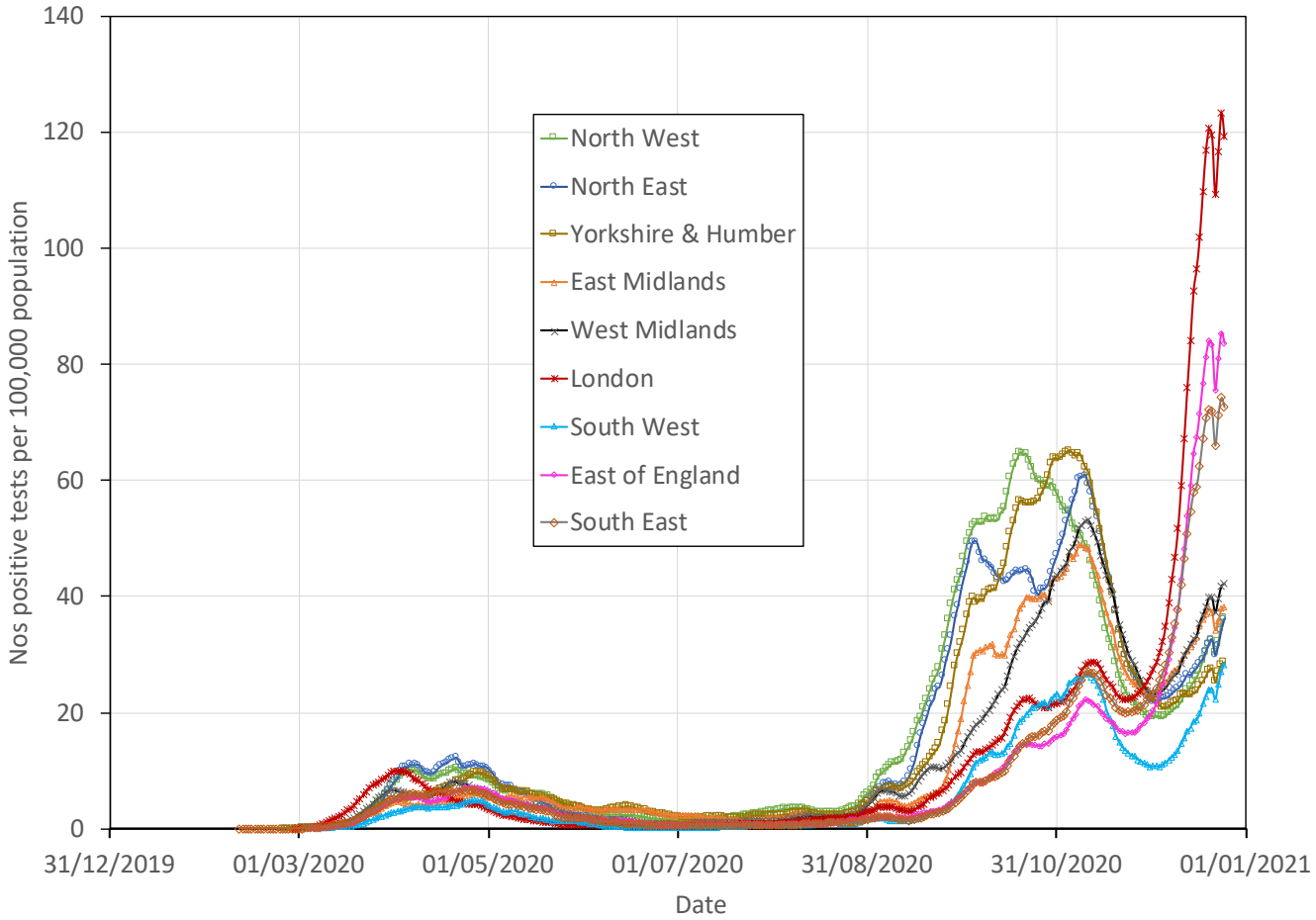

Fig.1 Reported population adjusted infection rates for the regions of England from the beginning of the pandemic, up to 1st January 2021 (1) using a 7-day average of cases by specimen date

Examination of Fig.1 provides clear light on how the virus has spread and grown across England. (Note, the variability in the curves at the end of December appears due to under recording around Christmas day.)

- There have been many waves of infection in the regions of various magnitudes, many major, but with some small but distinctly observable.

- In the "first wave", with the exception of London, which exhibited as an essentially singular wave, multiple overlapping waves can be seen in the other regions, exhibiting as multiple peaks within a broad "first" regional wave.

- After the first lockdown, infections continued to fall in a quite continuous fashion. There were some minor waves but, effectively, rates plateaued, showing only a slight rise until week 35, after which rates increased in all regions nationwide.

- Since week 35 there appears to be multiple overlapping, but distinct waves in all of the regions. While the lower rate regions exhibit these waves quite clearly, it is also clear they exist in the high-rate regions and can be seen as inflections in the slope of the test data, which occur as an increasingly dominant wave adds to the infection curve.

- During the second week of October, with the exception of the North West which peaked in the second half of September, there was a fall in infection rates for all regions, with the Northern and Midland regions exhibiting particularly sharp falls. Subsequently, around the end of November/beginning of December, infection rates rose in all regions, but most sharply in London, the South East and the East of England.

2. ***Development of the model***

From Fig.1 London is the only region that exhibited as mainly a single wave during the "first wave", which is rather symmetrical in form, and can be used a basis for developing a model consisting of multiple waves. Rather than using positive test data, the model has been developed taking as its basis fatality rates. The data is far more reliably reported over the whole period and the reported test data also does not accurately reflect the true total number of infections due to asymptomatic cases.

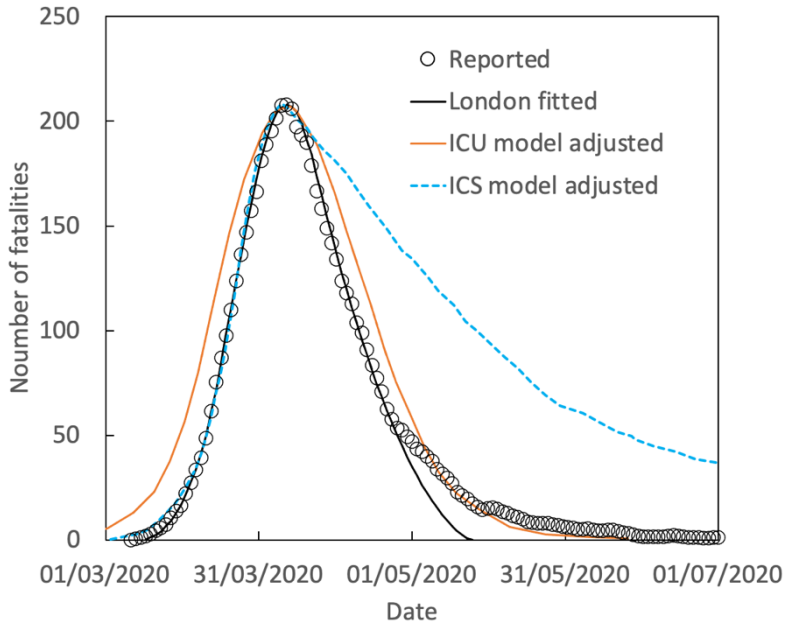

Fig.2 Fatality rates reported for London including those from the present model and using rate adjusted ICU and ICS models.

Fig.2 shows reported fatalities by date of death [(1)], as a 7-day average for London. The wave is quite symmetrical in nature until the 29th April, at which point a further wave becomes apparent, adding to the first, before diminishing to almost zero. The model used here fits the curve up until then, before itself diminishing to zero (Appendix 1). Fig.2 also includes two further curves taken after the Imperial College report of 16th March 2020 [(2)].

(i) A curve based on the "Do Nothing" scenario (hereafter called the Imperial College Unconstrained (ICU) model), scaled to match the peak height of the reported curve for London and date adjusted.

(ii) A curve based on the scenario of Schools and Universities Closure, Case Isolation and Population-Wide Social Distancing (hereafter called the Imperial College Suppression (ICS) model). The reported curve is provided in terms of critical care beds occupied and, based on the assumption it is directly proportional to subsequent fatalities, scaled to match the peak height for London and date adjusted.

The scaled ICU curve matches the reported data for the first wave rather well, though is a little broader, while the ICS curve matches the reported data almost exactly to the peak height before markedly diverging.

The fitted curve can now be used in a multiple wave model which requires three sets of input. The first is the number of waves, which is usually clear to see or becomes required during the fitting process. Otherwise, the fitting process requires only two parameters (i) the start date of the relevant wave and (ii) its height. The calculated fatalities for each wave are then summated and the cumulative curve matched to the reported data. The model was implemented here using spreadsheets but should suit a computer programme utilising a least squares methodology to minimise the error.

2. N.M. Ferguson et al., Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, (Imperial College COVID-19 Response Team, 16 March 2020)

3. ***Results***

3.1 "First Wave"

Figs. 3-11(a&b) show results for the fitted curves to reported data alongside plots for the divergence from the "ideal fit" in terms of the value of R^2. The data is given as fatalities, reported by date of death [1], as a 7-day average.

Considering the simplicity of the model, it matches the general behaviour of fatality rates very accurately, but there are some spikes in fatalities which cannot be matched. It may be, straightforwardly, that the model is unable to account for minor variations in behaviour or is due to clusters of fatalities in settings where fatality rates would be higher than for the overall population average.

Fig. 12 shows the sum of all the waves of the regions compared to fatality data reported for England [1] as a 7-day average, while Fig.13 compares the raw data as reported by date of death for completeness.

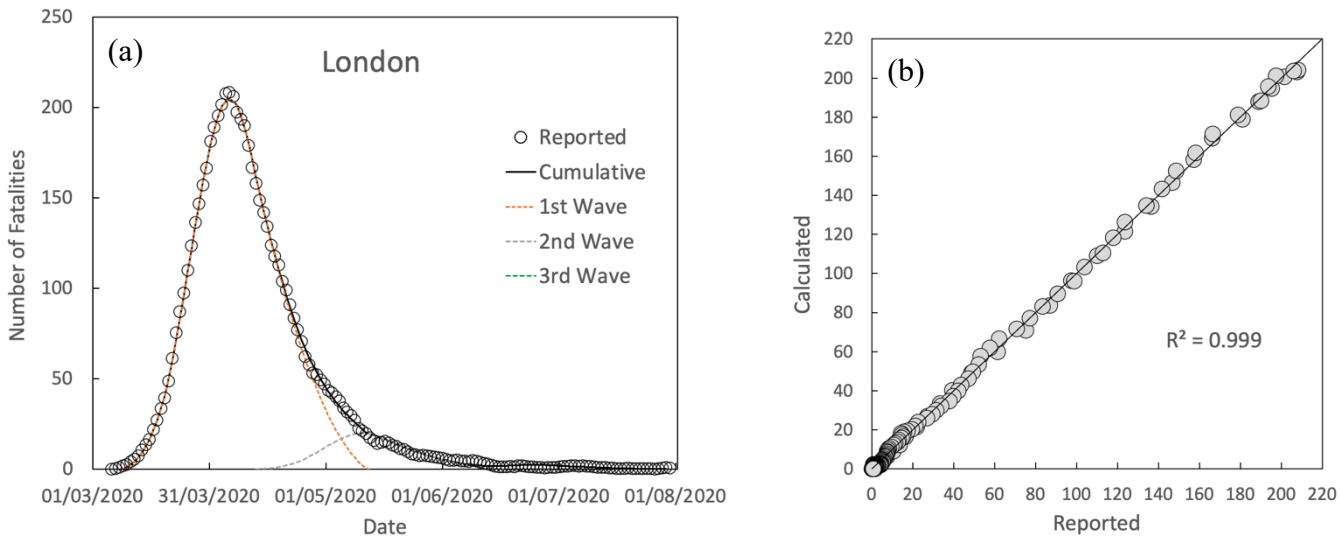

Fig.3 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for London in the "first wave"

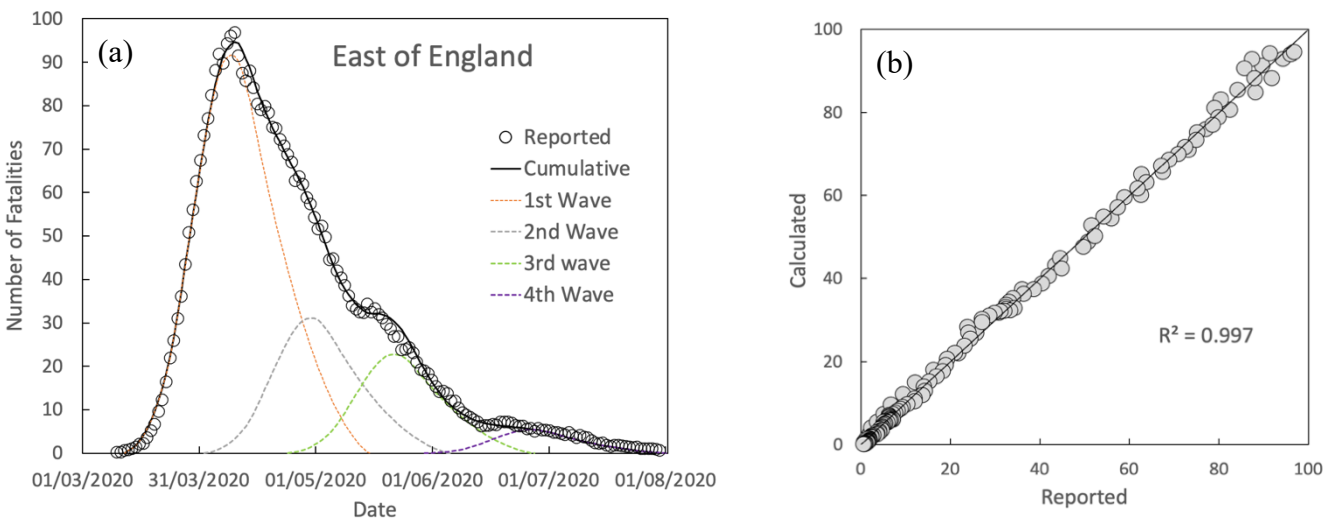

Fig.4 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the East of England in the "first wave"

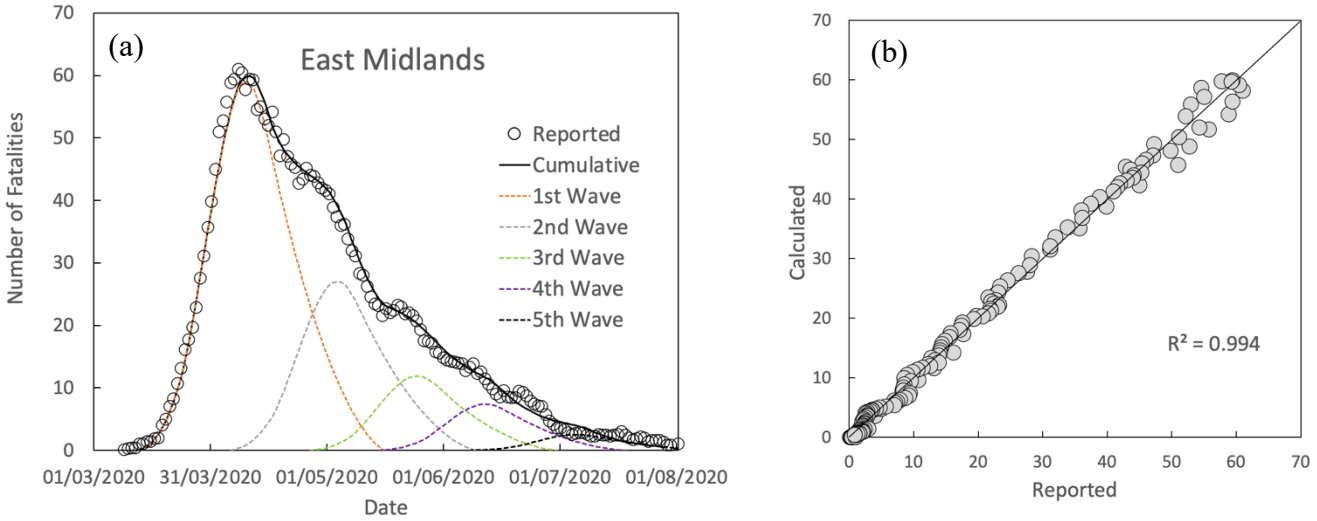

Fig.5 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the East Midlands in the "first wave

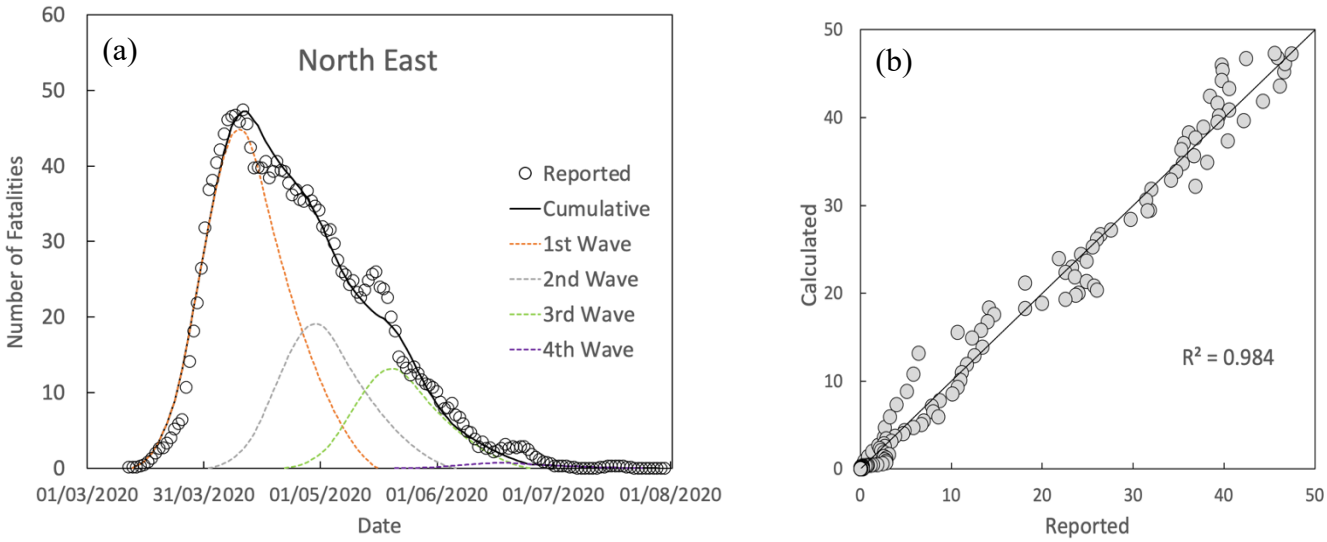

Fig.6 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the North East in the "first wave"

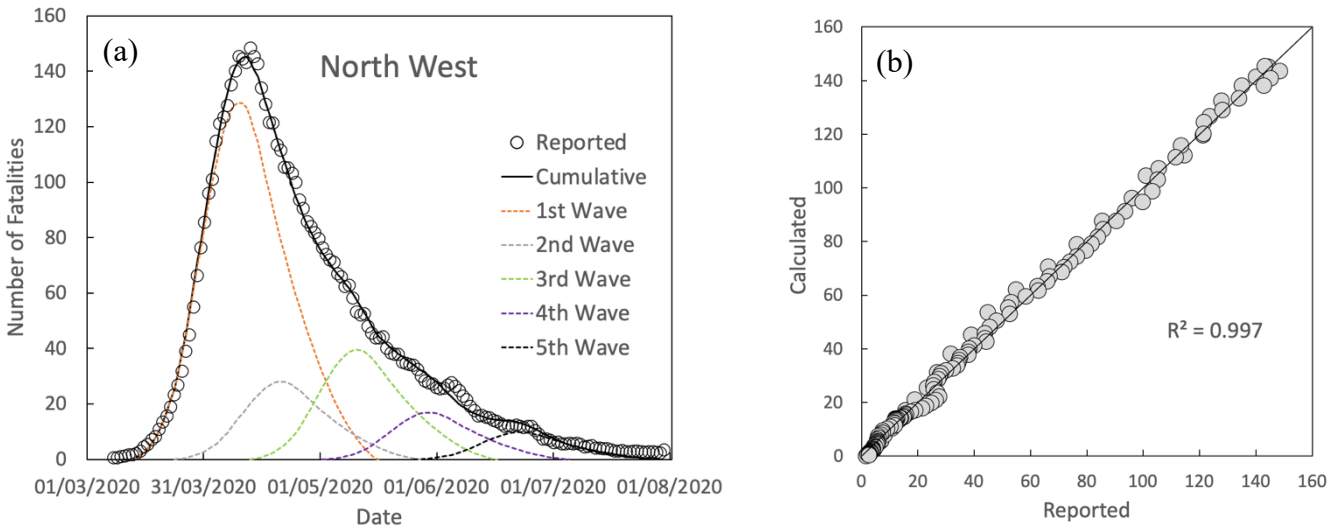

Fig.7 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the North West in the "first wave"

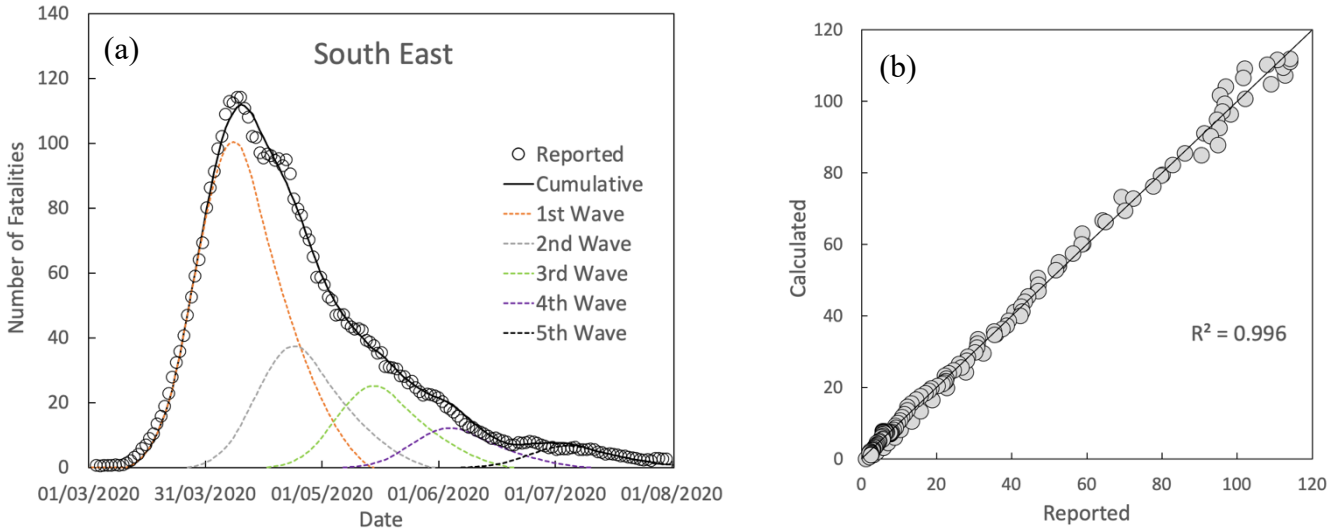

Fig.8 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the South East in the "first wave"

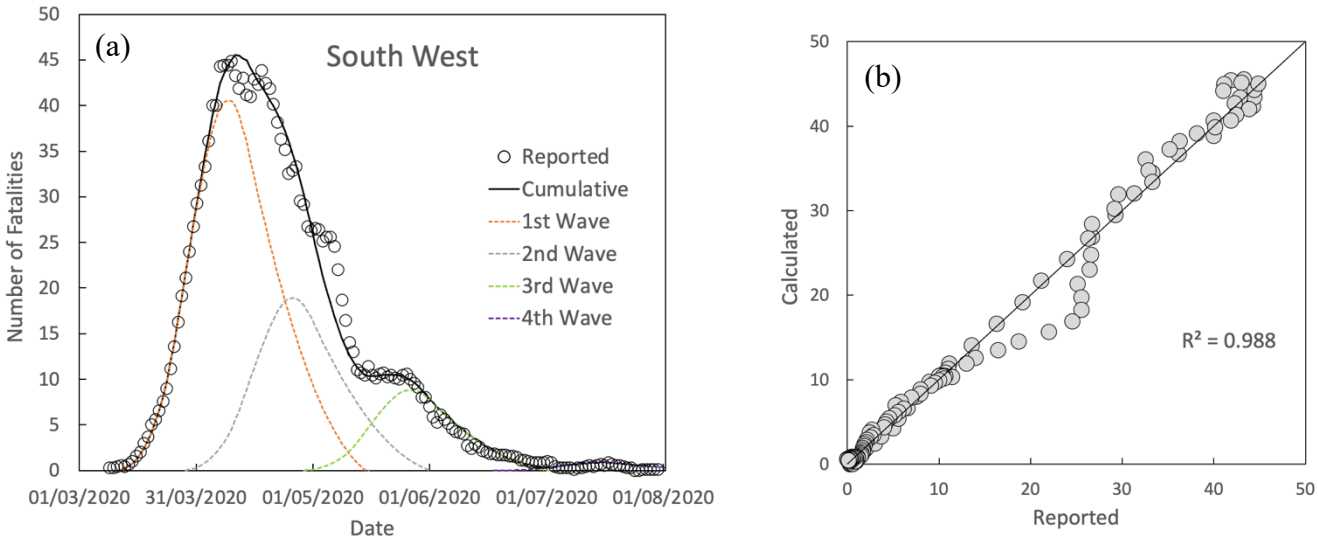

Fig.9 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the South West in the "first wave"

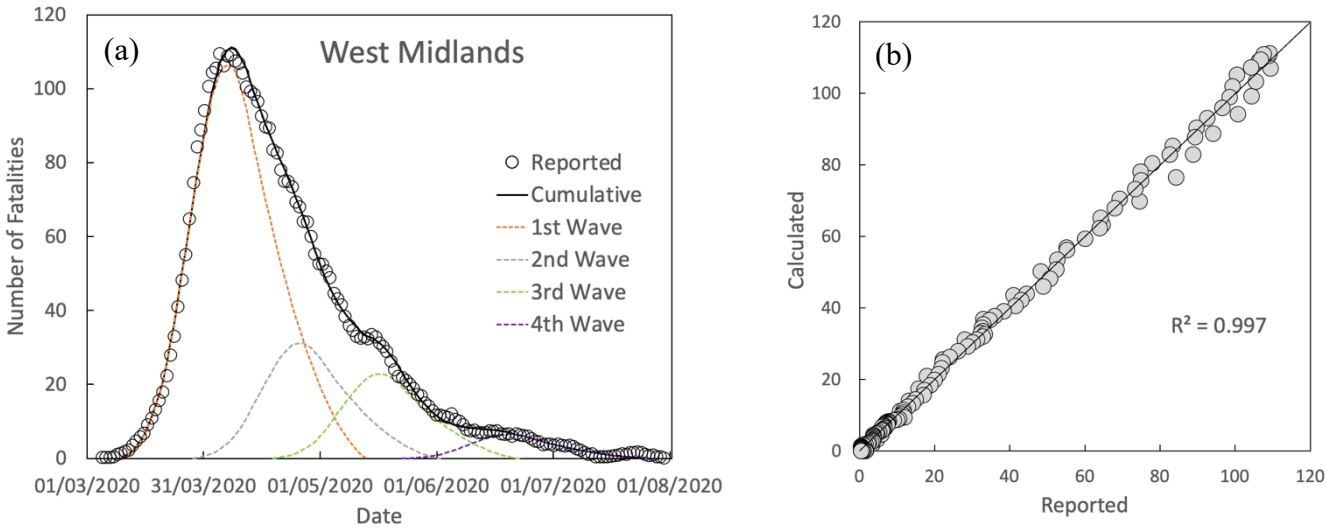

Fig.10 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the West Midlands in the "first wave"

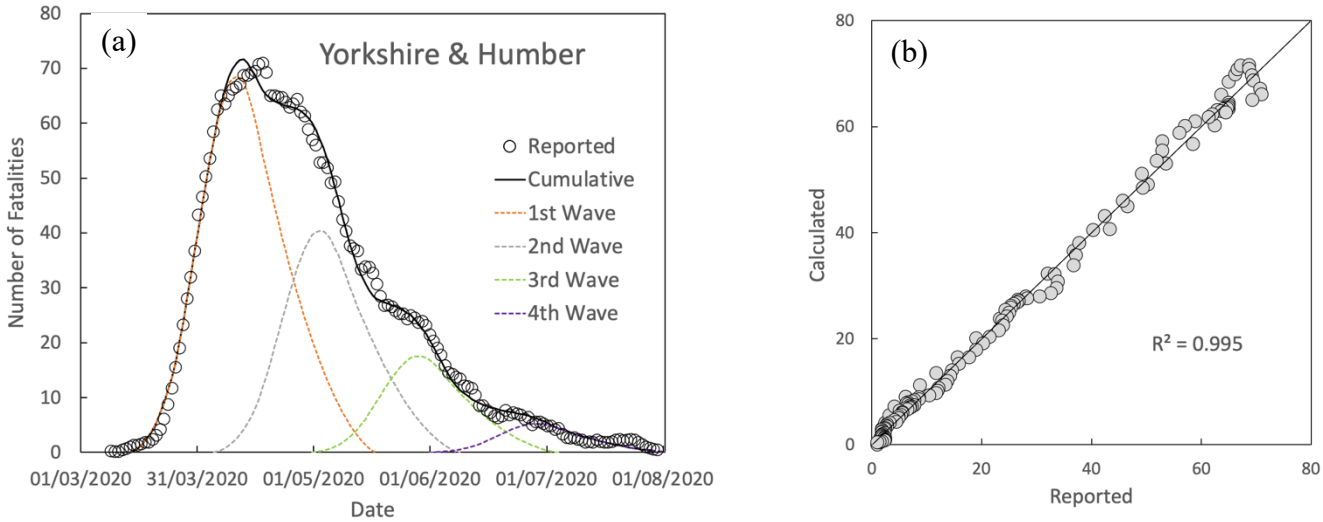

Fig.11 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for Yorkshire & Humber in the "first wave"

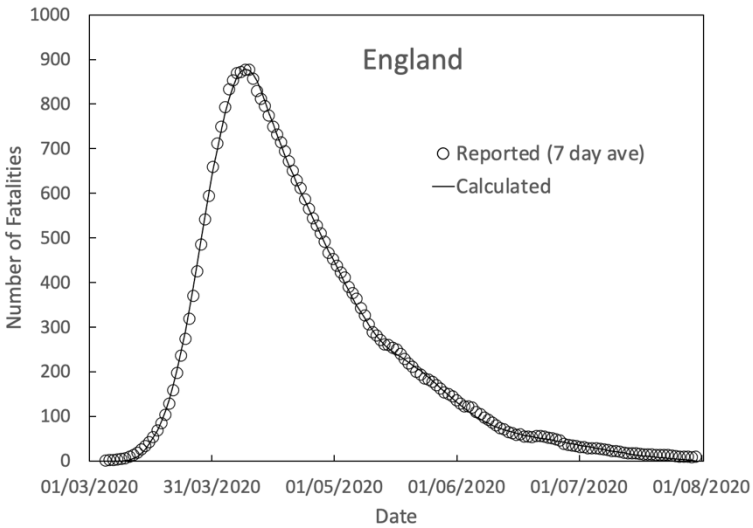

Fig. 12 Comparison between the sum of all calculated waves of the regions compared to reported fatalities by date of death (7-day average) for England during the "first wave"

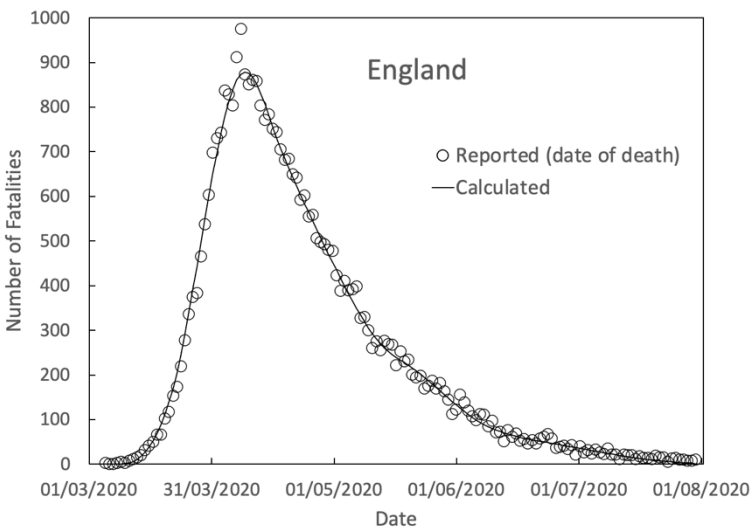

Fig. 13 Comparison between the sum of all calculated waves of the regions compared to reported fatalities by date of death for England during the "first wave"

3.2 "Second Wave"

Figs. 14-22 (a&b) show, similarly to the "first wave", results for the fitted curves to reported data [1] alongside plots for the divergence from the "ideal fit" in terms of the value of R^2. As for Fig.1 the last 5 reported days are not considered in the reported data. However, it is noted that there is significant delay in reporting, with fatality results still being reported well after the 5-day delay accounted for here. The reported and subsequently fitted rates for the later dates, therefore, have significant uncertainty. Fig.23 show the sum of all the waves of the regions compared to fatality data reported for England [1] as a 7-day average, while Fig.24 compares the raw data as reported per day for completeness.

All regions show at least 5 waves and, with the exception of the Midlands regions, a further 6th wave is becoming apparent in the other regions. It should be considered highly likely that a further wave will manifest itself in the Midlands regions during early 2021.

A common feature for all regions is the periodic nature of the waves. The model cannot explain why that should be the case but, a periodic nature is also seen quite clearly in the infection curves. It therefore cannot be considered as an artifact of the model and needs to be fundamentally understood.

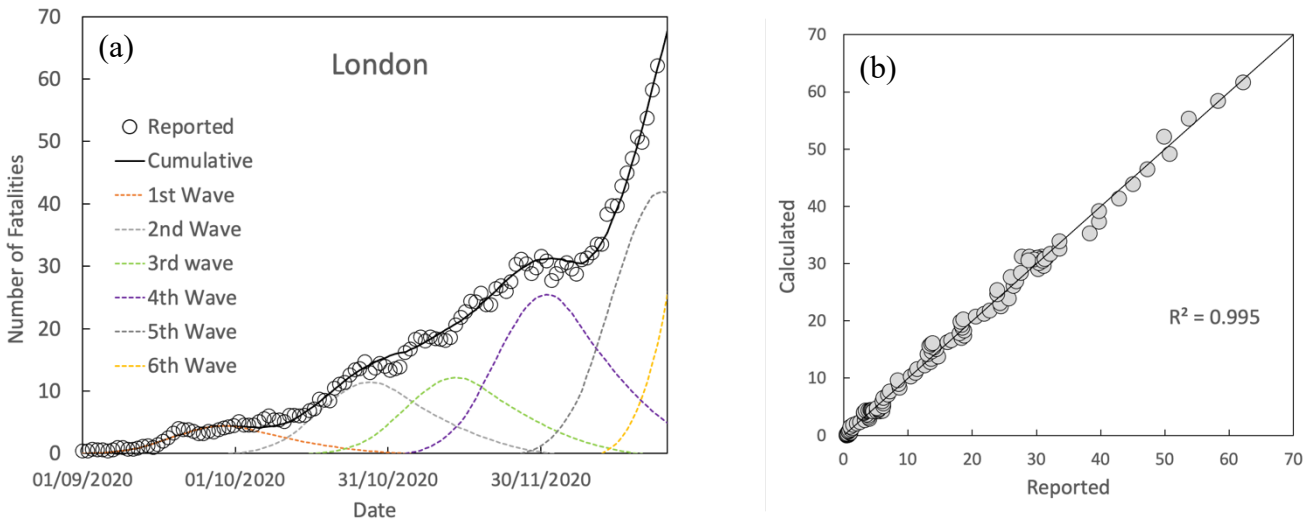

Fig.14 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for London in the "second wave"

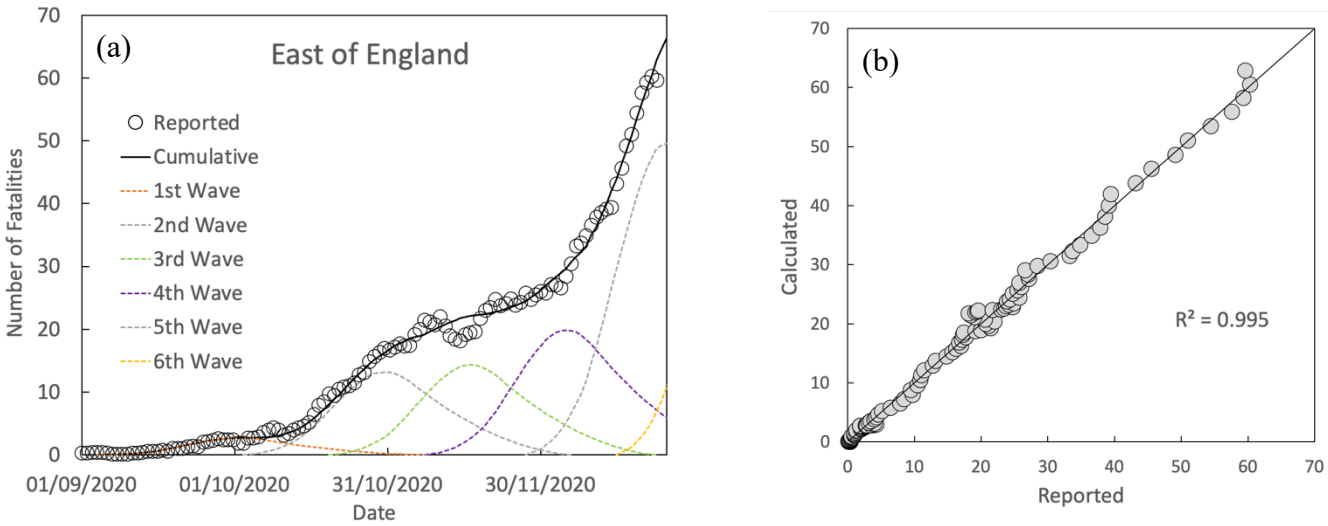

Fig.15 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the East of England in the "second wave"

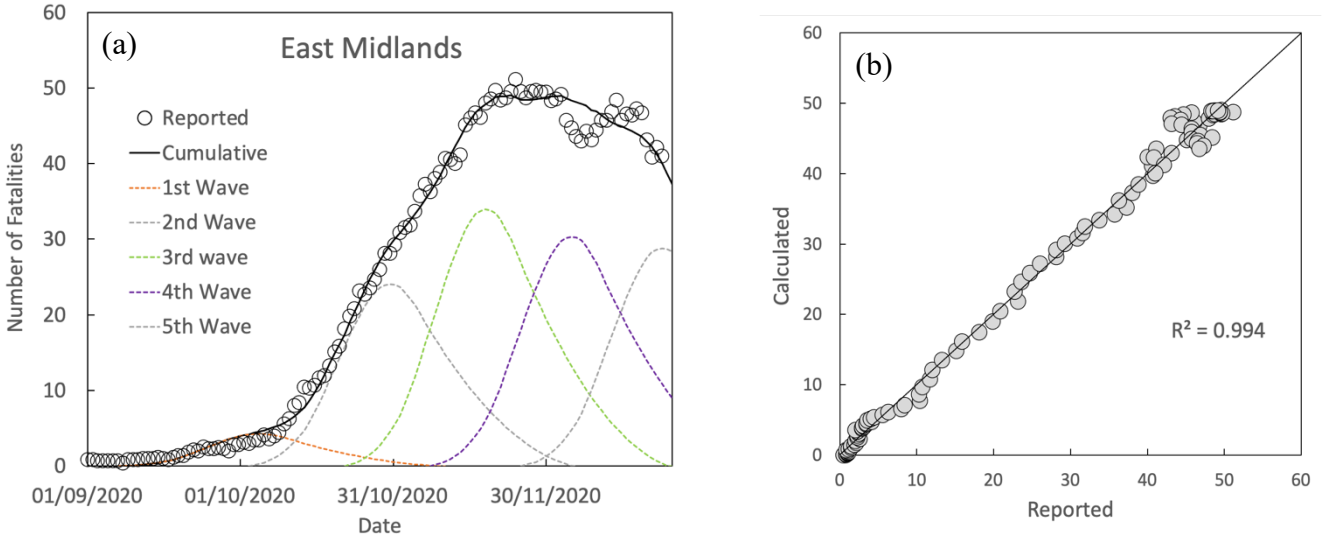

Fig.16 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the East Midlands in the "second wave"

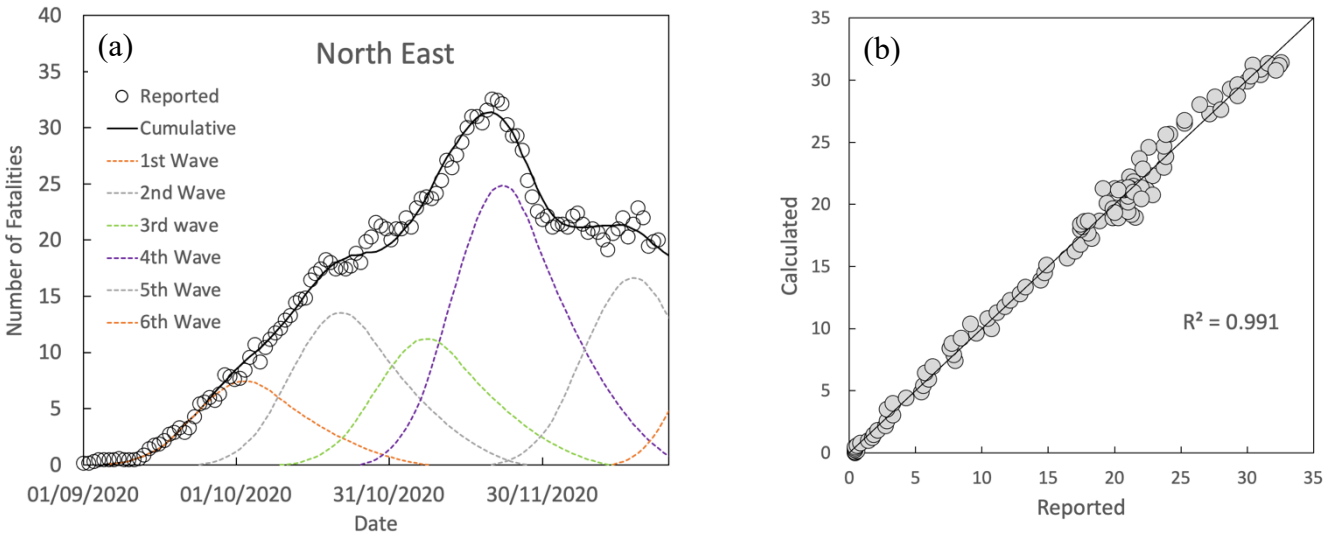

Fig.17 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the North East in the "second wave"

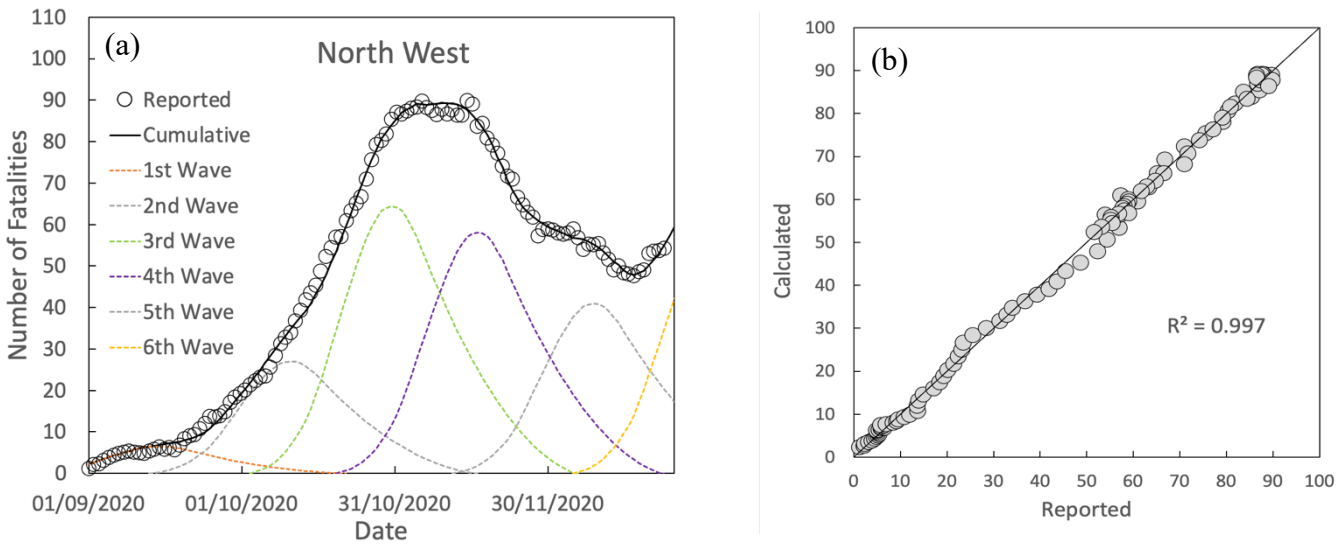

Fig.18 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the North West in the "second wave"

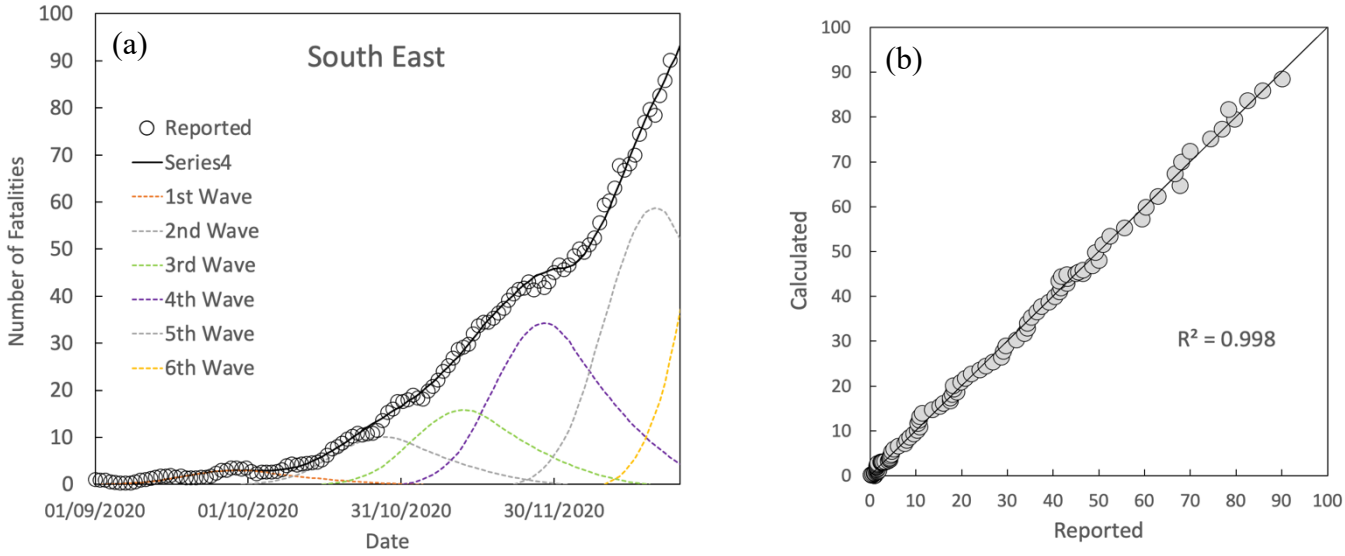

Fig.19 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the South East in the "second wave"

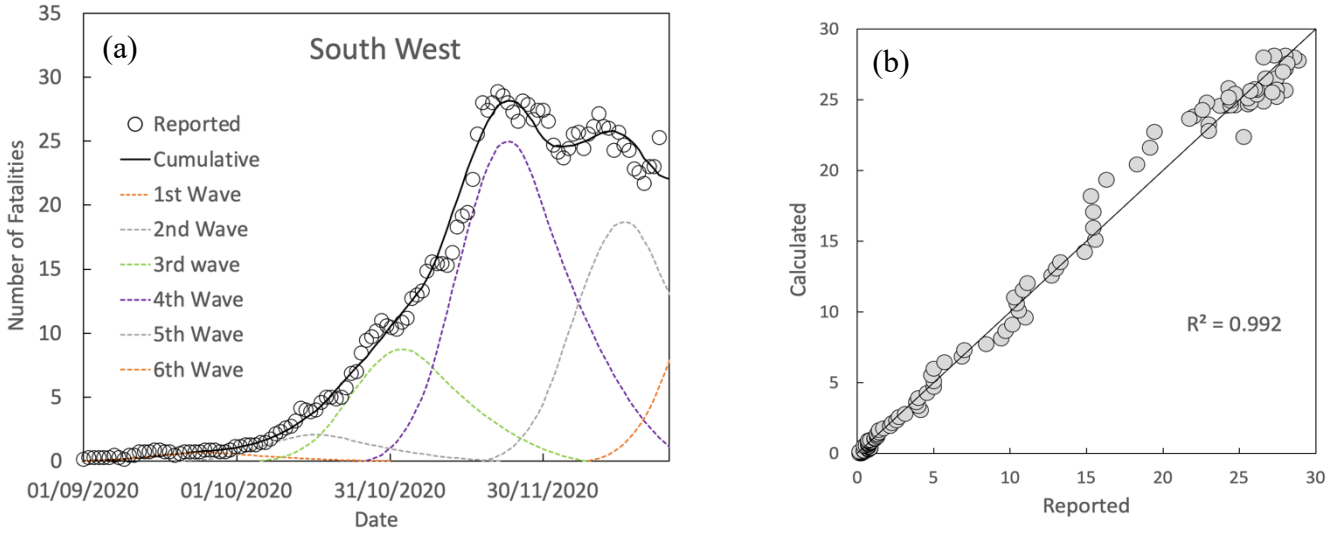

Fig.20 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the South West in the "second wave"

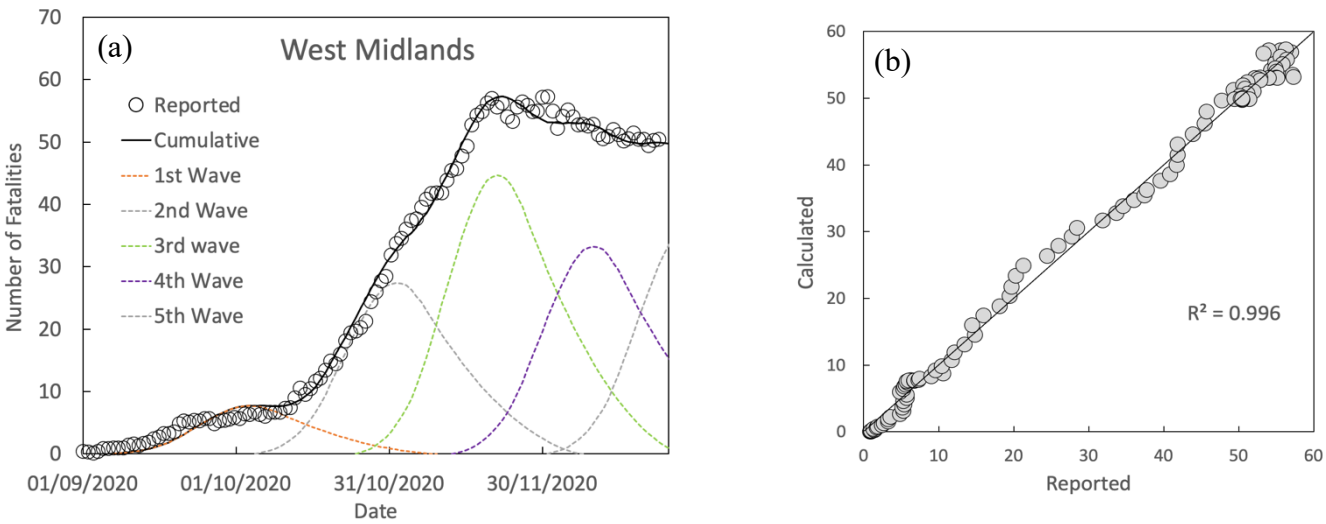

Fig.21 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the West Midlands in the "second wave"

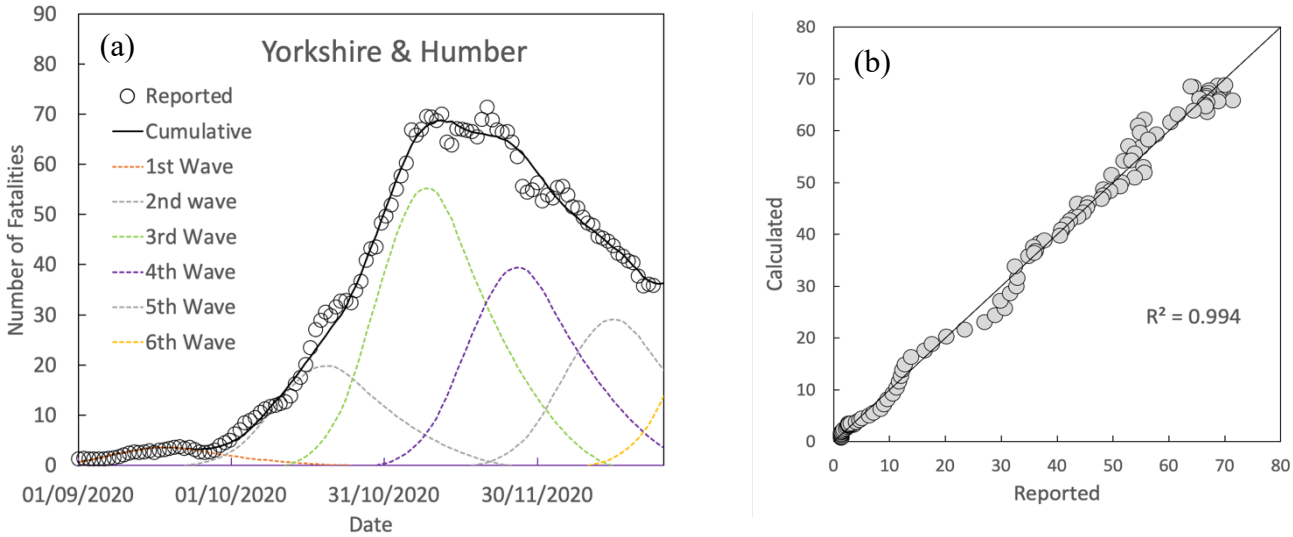

Fig.22 (a) Comparison between calculated and reported fatalities by date of death (7-day average) and (b) divergence from the "ideal fit" for the West Midlands in the "second wave"

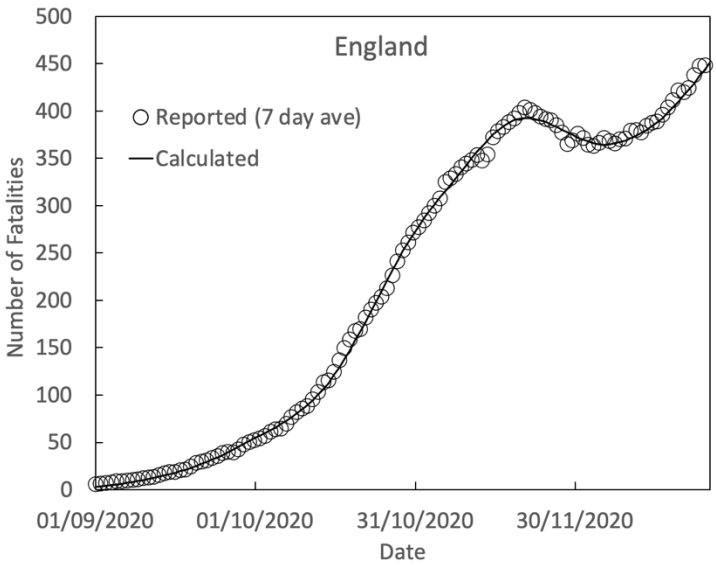

Fig. 23 Comparison between the sum of all calculated waves of the regions compared to reported fatalities by date of death (7-day average) for England during the "second wave"

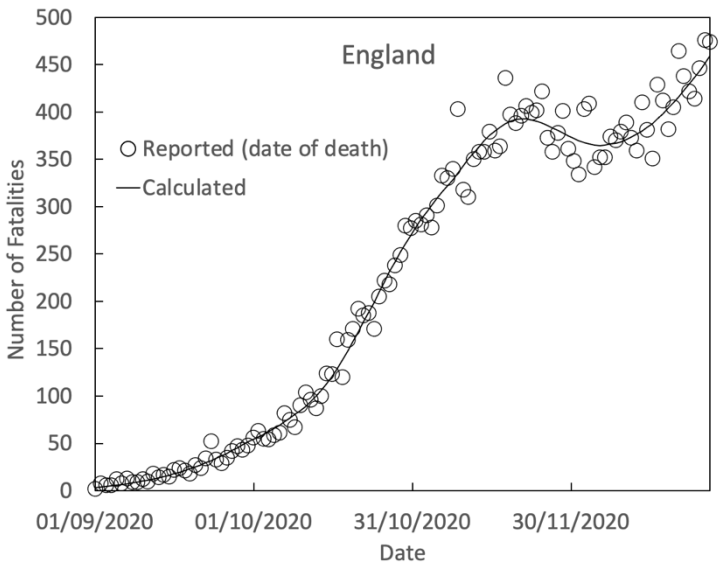

Fig. 24 Comparison between the sum of all calculated waves of the regions compared to reported fatalities by date of death for England during the "second wave"

4. Discussion

The accuracy of the fitted curves to reported fatality rates, with R^2 values approaching 100% against the ideal, strongly supports the validity of the current model and it reveals detail that would otherwise remain hidden or considered purely speculative. For example, in the "first wave" some regions show smoothly falling fatality rates with time, while others clearly show a multiple wave format. The strength of the model is it accounts for both types of behaviour in a unified approach.

It also accounts for the observed "exponential" rises in infection/fatality rates often quoted during the "second wave" which, when extrapolated into the future, substantially overestimated subsequent reported rates. Such apparent rises can happen when multiple waves, each increasing in magnitude with time, are added to each other and appear to be one continuous wave, e.g., Yorkshire & Humber.

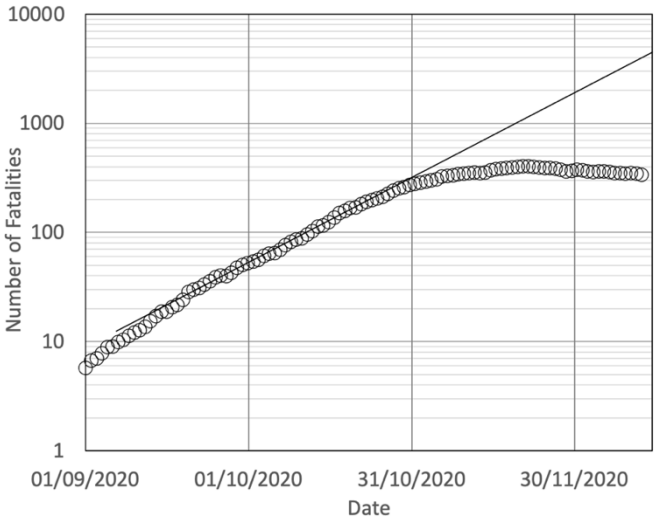

Fig. 25 Extrapolation of the fatality curve for England in the "second wave" showing overestimation of future fatality rates by assuming exponential growth

Fig.25 shows reported fatality rates as a 7-day average for England [(1)] plotted on an exponential y-axis. Extrapolation, using data for the first 2-3 weeks of October, leads to a prediction of 2,000 fatalities per day by the end of November, while the real value had peaked around 400 before then and subsequently fell further.

By its very nature, the method used here is an evidence-based approach and quite unlike the more theoretical approaches adopted by groups such as in Imperial College. It has the advantage that it uses pre-existing evidence to develop an approach that can quantitatively match reported behaviour. As such, it is not within the scope of the present report to undertake any detailed discussion on the differences between the two approaches other than to note and reference that the observed behaviour of London in the "first wave", which exhibited almost wholly as a single wave, is well matched by a much-reduced ICU model, rather than the ICS model.

The behaviour of the "first" and "second" waves is very different. The "first wave" consisted of a large initial spike in fatalities in all regions and was followed by a series of significantly sized secondary waves which diminished with time. London, which is the most highly interconnected of the regions, exhibited almost completely as a single wave unlike the other regions. It therefore seems reasonable to assume that the secondary waves in the other regions was due to the virus spreading into previously unaffected areas.

The "second wave" began after a long period of very low fatality rates and in all regions simultaneously. Unlike the "first wave", the "second wave" is characterised by a series of periodic waves which, to begin with, generally increased in size. Notably, London now exhibits as a series of periodic waves rather than a predominantly single wave. The magnitude of the waves peaked in the Northern, Midland and South West regions during November. However, fatalities in the South East and the East of England continued to rise, and quite sharply, while London slowed briefly before also sharply rising.

4.1 Regional Variation in Fatality Rates

While it is not immediately obvious from the fatality curves, there has been a clear regional divide in infection/fatality rates during the pandemic, which has been accentuated in the "second wave". Table 1 shows the cumulative death toll for each region at the start of the "second wave" and up until the 31st December, given both as a total and as per head of population. It also ranks the regions with respect to the lowest rate region which, at both times, was the South West.

At both times the North East and North West had the highest fatality rates and its ratio increased during the "second wave". While Yorkshire & Humberside sat in the middle of the regions at the start of the "second wave", it has climbed to just below the other Northern regions and the East Midlands has now closed with the West Midlands. London's position has dropped sharply during the "second wave" falling below the East of England. At both times the South East and South West had the lowest fatality rates per head of population.

Table.1 The regional variation in fatalities [1].

Region	***Number of Fatalities***	***per million of population***	***Ratio between highest and lowest region***
31 August 2020			
North West	6,173	841	*2.50*
North East	2,123	795	*2.36*
West Midlands	4,563	769	*2.28*
London	6,180	690	*2.05*
East of England	4,126	662	*1.96*
Yorkshire & Humber	3,617	658	*1.95*
East Midlands	2,953	610	*1.81*
South East	4,979	543	*1.61*
South West	1,892	337	*1.00*
31 December 2020			
North West	11,931	1,625	*2.89*
North East	4,095	1,534	*2.72*
Yorkshire & Humber	7,663	1,393	*2.47*
West Midlands	8,069	1,361	*2.42*
East Midlands	5,986	1,237	*2.20*
East of England	6,377	1,023	*1.82*
London	8,621	962	*1.71*
South East	8,373	913	*1.62*
South West	3,162	563	*1.00*

Conclusions

The current report makes no claims to provide a theoretical explanation for how COVID-19 has transmitted in England and its regions. Its purpose is to see if it is possible to quantitatively model the clear observation from measured infection/fatality rates that there have been many waves of infection and subsequent fatalities in the regions of England. The fact that the model can provide such exact correlation with observed fatality rates suggests it has a fundamentally sound basis and there would be merit in developing a theoretical approach to describe how such a multiple wave form has arisen, particularly so for the "second wave".

It further provides a clear and quantitative explanation as to why exponential extrapolations which have been made during the "second wave" have invariably failed by such large margins and accounts for fatality curves which quite clearly exhibit multiple waves as well as those that appear quite continuous.

It is also shown that there is a clear North/South divide in fatality rates during the pandemic, with the Northern regions exhibiting substantially higher rates per head of population than the more southern regions, with the Midlands regions sitting in-between. It is important to fundamentally understand why there has been such a large geographical disparity and, once understood, include it in any future theoretical modelling.

About the author

The author is a long-standing researcher and scientific modeller in the physical sciences and has written this article in a personal capacity as a contribution to the further understanding of the spread of COVID-19.

Appendix 1. Fitted curve for London Fatalities per day based on a 7-day average of reported data [1]

The curve below is used as a "master wave". It is based on the reported 7-day average for London and, for the first part, manually adjusted to provide a smoothly changing curve particularly around the peak. Subsequently, a polynomial function was used, partly fitted to the reported 7-day average and subsequently extrapolated to maintain consistency with the offset from the reduced ICU curve and match the apparent wave form itself.

While the fitted curve matches the reported fatalities for London very well, it would clearly be desirable for a more meaningful wave function to be developed to match the behaviour of London. It would also be of interest to examine other highly interconnected regions/cities, particularly if reasonably isolated from other populated areas, to see if there are common features in their behaviour during the "first wave" of the pandemic.

09/03/2020	0.00
10/03/2020	1.30
11/03/2020	2.40
12/03/2020	4.00
13/03/2020	6.40
14/03/2020	9.20
15/03/2020	12.50
16/03/2020	16.80
17/03/2020	21.80
18/03/2020	27.50
19/03/2020	34.00
20/03/2020	41.00
21/03/2020	50.00
22/03/2020	61.00
23/03/2020	72.00
24/03/2020	85.00
25/03/2020	98.00
26/03/2020	111.00
27/03/2020	123.57
28/03/2020	136.43
29/03/2020	149.00
30/03/2020	161.00
31/03/2020	172.00
01/04/2020	182.00
02/04/2020	191.00
03/04/2020	198.00
04/04/2020	204.00
05/04/2020	206.50
06/04/2020	207.60
07/04/2020	207.00
08/04/2020	204.50
09/04/2020	199.00
10/04/2020	191.50

11/04/2020	184.33
12/04/2020	174.21
13/04/2020	164.37
14/04/2020	154.82
15/04/2020	145.54
16/04/2020	136.54
17/04/2020	127.81
18/04/2020	119.37
19/04/2020	111.21
20/04/2020	103.32
21/04/2020	95.72
22/04/2020	88.39
23/04/2020	81.35
24/04/2020	74.58
25/04/2020	68.09
26/04/2020	61.89
27/04/2020	55.96
28/04/2020	50.31
29/04/2020	44.93
30/04/2020	39.84
01/05/2020	35.03
02/05/2020	30.50
03/05/2020	26.24
04/05/2020	22.27
05/05/2020	18.57
06/05/2020	15.16
07/05/2020	12.02
08/05/2020	9.16
09/05/2020	6.58
10/05/2020	4.28
11/05/2020	2.26
12/05/2020	0.52
13/05/2020	0.00