

# Reproducing country-wide COVID-19 dynamics can require the usage of a set of SIR systems (#49502)

1

First revision

## Guidance from your Editor

Please submit by **26 Oct 2020** for the benefit of the authors (and your \$200 publishing discount) .



### Structure and Criteria

Please read the 'Structure and Criteria' page for general guidance.



### Author notes

Have you read the author notes on the [guidance page](#)?



### Raw data check

Review the raw data.



### Image check

Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

## Files

Download and review all files from the [materials page](#).

1 Tracked changes manuscript(s)  
1 Rebuttal letter(s)  
2 Latex file(s)  
6 Table file(s)



# Structure and Criteria

## Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. BASIC REPORTING
2. EXPERIMENTAL DESIGN
3. VALIDITY OF THE FINDINGS
4. General comments
5. Confidential notes to the editor

 You can also annotate this PDF and upload it as part of your review

When ready [submit online](#).

## Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your [guidance page](#).

### BASIC REPORTING

-  Clear, unambiguous, professional English language used throughout.
-  Intro & background to show context. Literature well referenced & relevant.
-  Structure conforms to [Peerj standards](#), discipline norm, or improved for clarity.
-  Figures are relevant, high quality, well labelled & described.
-  Raw data supplied (see [Peerj policy](#)).

### EXPERIMENTAL DESIGN

-  Original primary research within [Scope of the journal](#).
-  Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
-  Rigorous investigation performed to a high technical & ethical standard.
-  Methods described with sufficient detail & information to replicate.

### VALIDITY OF THE FINDINGS

-  Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
-  All underlying data have been provided; they are robust, statistically sound, & controlled.
-  Speculation is welcome, but should be identified as such.
-  Conclusions are well stated, linked to original research question & limited to supporting results.

# Standout reviewing tips

3



The best reviewers use these techniques

## Tip

**Support criticisms with evidence from the text or from other sources**

## Example

*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*

**Give specific suggestions on how to improve the manuscript**

*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*

**Comment on language and grammar issues**

*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*

**Organize by importance of the issues, and number your points**

- 1. Your most important issue*
- 2. The next most important item*
- 3. ...*
- 4. The least important points*

**Please provide constructive criticism, and avoid personal opinions**

*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*

**Comment on strengths (as well as weaknesses) of the manuscript**

*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

# Reproducing country-wide COVID-19 dynamics can require the usage of a set of SIR systems

**Eugene B Postnikov** Corresp. 1

<sup>1</sup> Kursk State University, Kursk, Russia

Corresponding Author: Eugene B Postnikov  
Email address: postnikov@kursksu.ru

This work shows that simple compartmental epidemiological models may not reproduce actually reported country-wide statistics since the latter reflects the cumulative amount of infected persons, which in fact is a sum of outbreaks within different patches. At the same time, the multilogistic decomposition of such epidemiological curves reveals components, which are quite close to the solutions of the SIR model in logistic approximations characterised by different sets of parameters including time shifts. This line of reasoning is confirmed and illustrated with the processing data for Spain and Russia.

# Reproducing country-wide COVID-19 dynamics can require the usage of a set of SIR systems

Eugene B. Postnikov

Department of Theoretical Physics, Kursk State University, Radishcheva st., 33, 305000 Kursk, Russia

Corresponding author:

E.B. Postnikov

Email address: postnikov@kursksu.ru

## ABSTRACT

This work shows that simple compartmental epidemiological models may not reproduce actually reported country-wide statistics since the latter reflects the cumulative amount of infected persons, which in fact is a sum of outbreaks within different patches. At the same time, the multilogistic decomposition of such epidemiological curves reveals components, which are quite close to the solutions of the SIR model in logistic approximations characterised by different sets of parameters including time shifts. This line of reasoning is confirmed by processing data for Spain and Russia in details and, additionally, is illustrated for several other countries.

## INTRODUCTION

The COVID-19 pandemic is recently one of the most urgent problems and challenges of public health, which affects almost all countries of the world, see Sohrabi et al. (2020); Wang et al. (2020); He et al. (2020); Gates (2020); Bedford et al. (2020) for a general review. Respectively, this situation states the problem of mathematical modelling of the local and global outbreaks aimed at understanding the epidemic process, revealing factors affecting the course of the outbreak, the effectiveness of applied measures and provided recommendations, which should prevent possible new waves of this disease Cobey (2020).

One of the natural approaches to such modelling is the usage of different variants of compartmental models Brauer et al. (2008), which subdivide the full populations into groups of susceptible  $S$ , infected  $I$  and recovered/removed  $R$  persons (SIR model), which can also be supplied with the group of exposed persons ( $E$ ), and even with more detailed subdivision. A number of recent studies Maier and Brockmann (2020); Fanelli and Piazza (2020); Postnikov (2020); Linka et al. (2020); Carcione et al. (2020) demonstrated a reasonable reproduction of the dynamics of the actually reported cases in a variety of countries on the time intervals from the growing stage of the outbreak till a vicinity of its peak as well as the work Dehning et al. (2020) analysed the well-developed outbreak with switching parameters of SIR model. It has been shown even for the case of the simplest SIR model chosen for two reasons: at the first, the lack of detailed data on intra-country distributions, movements and contacts in the country-wide reports, which give a strong mean-field picture, makes desirable minimisation of the hidden parameters, and in the second and important, the SIR model provides an opportunity for the direct data-driven approach.

It has been shown Postnikov (2020) such model defined by the system of ordinary differential equations

$$\frac{dS}{dt} = -kSI, \quad (1)$$

$$\frac{dI}{dt} = kSI - \tau^{-1}I, \quad (2)$$

$$\frac{dR}{dt} = \tau^{-1}I, \quad (3)$$

where  $k$  and  $\tau$  are the transmission and recovery rates, can be sequentially reduced to one Verhulst's

logistic equation with respect to the cumulative number of infected persons

$$\frac{d\Sigma I}{dt} = r\Sigma I \left(1 - \frac{\Sigma I}{K}\right), \quad (4)$$

where  $r$  is the growth rate, which is directly connected with the basic reproduction number  $\mathfrak{R}_0$  via parameters of the differential equations (1)–(3), see the explicit formulae below, in the section “Methods”. The parameter  $K$  is the saturation value. It depends on both  $\mathfrak{R}_0$  and the total number of potentially susceptible people, see “Methods” again; the problem of its proper determination, which can be connected to either spatially or socially separated subpopulations, i.e. there is a difficulty in its estimation *a priori*. Moreover, the cumulative mean-field epidemic curve can be affected by mixing such subpopulations that is the main focus of the present study.

At the same time, it should be kept in mind that the logistic function, which is the analytic solution to Eq. (4), is also an approximate solution to the SIR-system of ordinary differential equations (1)–(3) derived on a basis of existing invariants of this ODE system and assumption that  $\max I(t) \ll K$  that fulfils in practice, see Postnikov (2020) for the step-by-step derivation. The analysis carried out for different countries till the first decade of April confirmed quantitative adequacy of this approach for different countries among which were Italy, France, Germany, Spain.

However, new data registered to the decaying stage of the outbreak exhibit drastic deviations from such logistic curve. The same behaviour is noted for other compartmental models that induced different attempts to their more complicated modification, e.g. Oliveira (2020); Mamon (2020); Schaback (2020).

Note that the preliminary analysis evaluated for the case of Italy Vannucci and Vannucci (2020) demonstrated that reproducibility could be restored if to consider a combination of two logistic curves corresponding to two outbreak waves. The similar combination was applied earlier to the long-time modelling of tuberculosis Lavrova et al. (2017) based on a superposition of SIR models with different (switched with time) control parameters. Note also that a similar behaviour of data curves is already known in the theory of markets’ evolution and technological changes known as the logistic substitution model Meyer et al. (1999) and the respective software Loglet Lab 4 is developed to carry out such analysis. It is available online at <https://logletlab.com> and supplied with a variety of examples of such decomposition and details of the numerical algorithms.

Thus, the main goal of this work is to explore if it is possible to decompose the data reported for COVID-19 outbreaks at the country-wide scale into a set of curves, each of which satisfies the SIR-Verhulst model. In other words, if it is possible to consider the dynamics of such data as a superposition of outbreaks satisfying the simple compartmental model but operating with different subpopulations and parameters and, respectively, to get a tool for revealing a number of such individual characteristic modes from unique general data.

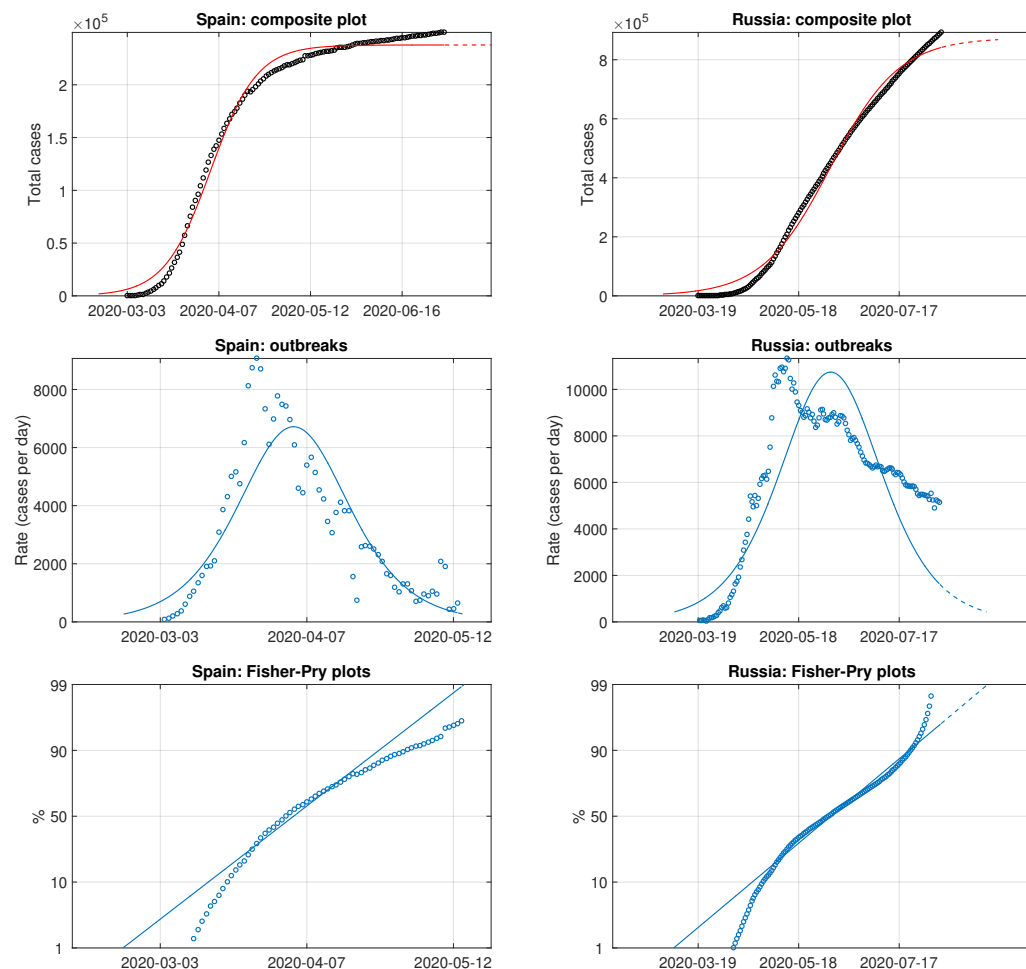
## METHODS

The project “Our world in Data” aggregating the data related to COVID-19 pandemics from different sources Roser et al. (2020) was used as the data source. These data were downloaded in an automatic regime with a home-written function, which belongs to the developed tool for MATLAB/GNU Octave made available in the GitHub repository

[https://github.com/postnicov/owid\\_loglet\\_interface](https://github.com/postnicov/owid_loglet_interface) and their subset containing the cumulative number of infected persons for a chosen country (total cases) was processed to form a text file suitable for uploading to the online software Loglet Lab 4 and carrying out the multilogistic decomposition. The results of this decomposition report the components of decomposition. Each such component is the logistic curve satisfying the SIR-Verhulst equation (4) and has the form

$$\Sigma I(t) = \frac{K}{1 + e^{-r(t-t_m)}}, \quad (5)$$

where  $t_m$  defines the time moment when the cumulative number of infected persons reaches half of the saturated values. It should be pointed out that the growth and saturation parameters  $r$  and  $K$  are directly connected with the parameters of the SIR model (1)–(3) as  $r = (\mathfrak{R}_0 - 1)\tau^{-1}$  and  $K = \mathfrak{R}_0^2 (2(\mathfrak{R}_0 - 1)N)^{-1}$ , where  $\mathfrak{R}_0 = k\tau$  is the basic reproductive number and  $N = S + I + R = \text{const}$ , see Postnikov (2020) for details.



**Figure 1.** The total cases of persons infected by COVID-19 according to “Our World in Data” for Spain (from March, 2 to July, 1) and Russia on August, 11 (circles) and their logistic fit (solid curves) as well as growth rates and Fisher-Pry plots for these data and their fitting.

An additional output is the so-called Fisher-Pry plot Fisher and Pry (1971) for each component transforming (5) into

$$\ln \frac{\Sigma I/K}{1 - \Sigma I/K} = -r(t - t_m), \quad (6)$$

i.e. one can easily check if the data already satisfy the logistic curve looking do they follow a straight line or not.

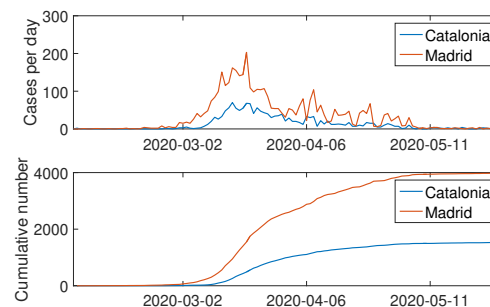
In addition, the average relative absolute deviation (AAD),

$$AAD = 100 \% \left\langle \frac{\Sigma I_{rec} - \Sigma I_{calc}}{\Sigma I_{rec}} \right\rangle$$

is used in this work as a quantitative criterion. Here  $\Sigma I_{rec}$  and  $\Sigma I_{calc}$  are the actually registered and the model-based calculated values of the cumulative number of infected persons, respectively. The average value is taken over the whole time interval covering values  $\Sigma I_{rec} > 0.05 \cdot \max(\Sigma I_{rec})$ .

## RESULTS

As two principal illustrative examples, the dynamics of COVID-19 outbreaks in two countries, Spain and Russia, will be considered. The work Postnikov (2020) showed that it satisfied the Verhulst-SIR model (4)



**Figure 2.** The plots demonstrating daily registered incidences of COVID-19 in two valuable regions of Spain spatially separated.

with high accuracy for the state-of-the-art up to the first decade of April. However, the actual longevity and cumulative effects of outbreaks exceed the predictions made using the data available for the early stages of the process.

Fig. 1 demonstrates the data downloaded from “Our World in Data”, uploaded for processing into Loglet Lab and post-processed again with the Octave script. Here the unique logistic curve was applied for the cumulative number of infected persons (total cases), i.e. the “composite plot” is automatically composed of one component.

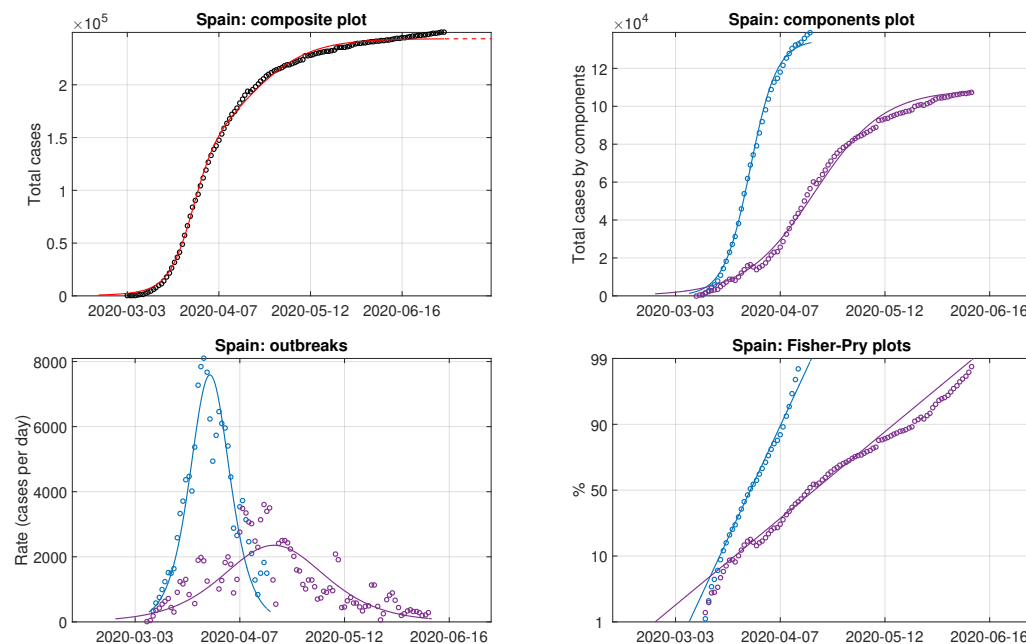
One can see that the single logistic curve fits the total cases only qualitatively on average ( $AAD = 6.66\%$  and  $AAD = 8.19\%$  for Spain and Russia, respectively). The quantitative details of approximation do not accurately reproduce the dynamics, black circles go under and over the red curve within relatively long intervals as well as there are significant displacements of recorded and model curves on the bell plots. In addition, this inconsistency is most obvious looking at the Fisher-Pry plots: they do not follow straight lines as it should be from (6) if the recorded data with a single defined  $K$  satisfy the logistic growth law. An attempt to improve the correspondence superposing maxima manually do not really improve the situation since there will emerge “long tails” of recorded data drastically overcoming fast-decaying model rate curve and underestimating the curves of total cases.

These facts argue in favour of the idea that the data accumulated over the whole country may in fact consist of different maxima and different growth rates for different regions, which start to be significant after the middle of April (before this date there was a quite accurate correspondence between the data and the SIR-Verhulst model, see Postnikov (2020)). As the first example, let us consider the case of Spain since the registered and the model curves in Fig. 1 visually differ mainly by the location of their maxima. Fig. 2 represents the examples of regional statistics for two separated regions of Spain, Madrid and Catalonia. The data for plotting were obtained from the web-site of the National Epidemiological Center of Spain CNE (2020). One can see that before the first decade of April both curves had the similar shape, similar location of their maxima and differ in the magnitude only. Therefore, such a synchronization leads to a possibility to represent the summed data as a single curve with some unique set of effective parameters of the SIR-Verhulst model. The same is true for the rest of the regions too and, as a result, one system (1)–(3) reduced to one equation (4) with one logistic functions as the solution is completely enough as it has been already demonstrated earlier.

On the contrary, the data for Madrid demonstrates the secondary outbreak between the first decades of April and May, which can not be predicted directly from the data recorded earlier. Moreover, since the SIR system (1)–(3) has an unimodal solution for  $I(t)$ , it can be hypothesized that the disease spread over a group of susceptible persons, which differ from the initial one.

Therefore, it is natural idea to carry out more detailed analysis, which includes consideration of an additional independent SIR system. As follows from the reducibility of the SIR ODE system to one Verhulst’s equation, this means that it is required to consider two logistic curves whose sum should fit the data not fitted by one curve only. This decomposition in practice can be easily carried out by means of Loglet Lab software. The respective result is shown in Figs. 3. Fig. 3 indicates that, in the case of Spain, the assumption of the existence of two logistic components gives sufficiently better reproducing the observed dynamics,  $AAD = 1.73\%$ . One component corresponds to the more fast and strong outbreak





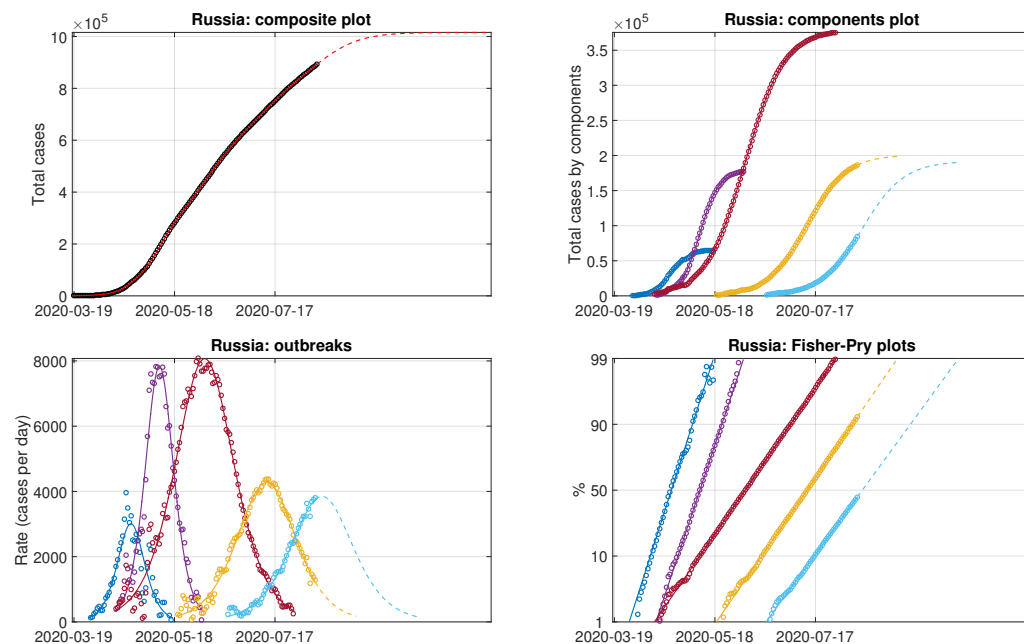
**Figure 3.** The plots generated after the downloading data from “Our World in Data” for Spain on July, 1, their processing with Loglet Lab and postprocessing. Different components revealed as comprising the total epidemic curve (cumulative number of revealed infected persons) are marked with different colours. The circles are recorded data and the lines are their reconstruction by the solutions of the SIR-based logistic equation (4).

with the peak on March, 27 and the second is slower with the peak on April, 18 that is agreed with the discussion of Fig. 2 above. Both components follow the solution of the Verhulst equation as seen from the respective plot in Fig. 3 and reasonable linearity of partial Fisher-Pry plots there. The last fact confirms the validity of logistic approximation for the both components considered as independent. The full set of parameters for both logistic functions are listed in Table. 1.

**Table 1.** The parameters of the separated logistic components (5) for the case of Spain, where  $r$  (days<sup>-1</sup>) is the components growth rate,  $t_m$  (days) is the time location of the inflection point for the logistic function, and  $K$  (persons) is the logistic curve’s maximum. The first day of the time series is March, 2.  $R^2$  denotes the coefficient of determination.

Phase	$K$	$t_m$	$r$	$R^2$
1	134912	26.0	0.225	0.973
2	108805	47.4	0.087	0.973

The next example, the case of Russia, which is sufficiently more spatially extended country than Spain, requires a larger number of logistic components. Fig. 4 demonstrates this five-modes decomposition valid to the current state of still progressing epidemics in Russia, which reproduces the actual data with a pretty good accuracy,  $AAD = 0.334\%$ . Each component curve corresponds to the logistic growth process with high accuracy too that is seen from all plot representations, for the components themselves, their growth rates, and the Fisher-Pry decompositions as well as from the high value of the coefficient of determination. Notably, that the growth coefficient values  $k$  given in Table 2 can be subdivided into two groups. At the first stage (from the middle of March till the middle of May), there are two subsequent component with large  $k$ , and then (since the middle of May up to date) three components with practically equal  $k$  again time-shifted. This is also demonstrably visible in the plot representing the Fisher-Pry transform. This division can be put in line with the governmental measures: the period March, 30 – May, 11 was non-working nationwide, the public places lockdown and strict travel restrictions were applied. Notably, this period parenthetically coincides with the period of existence of the major part of the first



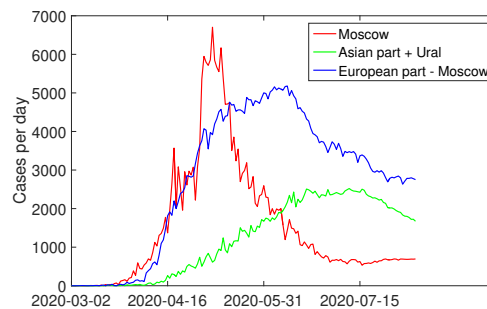
**Figure 4.** The plots generated after the downloading data from “Our World in Data” for Russian on August, 11, their processing with Loglet Lab and postprocessing. Different components revealed as comprising the total epidemic curve (cumulative number of revealed infected persons) are marked with different colours. The circles are recorded data and the lines are their reconstruction by the solutions of the SIR-based logistic equation (4). Dashed lines correspond to the model-based projections.

two components in Fig. 4. Therefore, it can be concluded that the measures applied localised the number of potentially susceptible persons by several spheres of close contacts of infected persons and led to the fast reaching the saturated values for these two first logistic components. Certainly, the contacts were not broken completely and the outbreak continued to spread countrywide but with sufficiently slower growth rate even over the background of more mild restrictions. The three subsequent logistic components can be associated with such a spatial spread.

**Table 2.** The parameters of the separated logistic components (5) for the case of Russia. The units of  $r$ ,  $t_m$ ,  $K$  are days,  $\text{days}^{-1}$ , persons, respectively. The first day of the time series is March, 18.  $R^2$  denotes the coefficient of determination.

Phase	$K$	$t_m$	$r$	$R^2$
1	65744	35.0	0.183	0.979
2	178111	52.0	0.176	0.996
3	379133	79.0	0.085	0.999
4	200709	116	0.085	0.997
5	192037	149	0.080	0.997

As above for Spain, the origin of such a behaviour for Russia can be explained exploring the geographically subdivision as it is shown in Fig. 5 using the regional data for Russia accumulated by the Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing accessed via Yandex’s web-service DataLens (2020). Moscow as the capital, the most populated city, and the largest transport hub is considered separately along with the European (without Moscow) and Asian (including the borderline Ural region) parts of country. From epidemiological point of view, one can see that Fig. 5 demonstrates that such a subdivision provides comparable amount of infected persons. Thus, it is possible to put in correspondence features in Figs. 4 and 5. First of all, it is visible that the outbreak in Moscow is the leading one during the first stage of the time course. In addition, the respective curve is bimodal. Therefore, the April’s peak can be associated with the primary introduced virus spreaders (Moscow is a



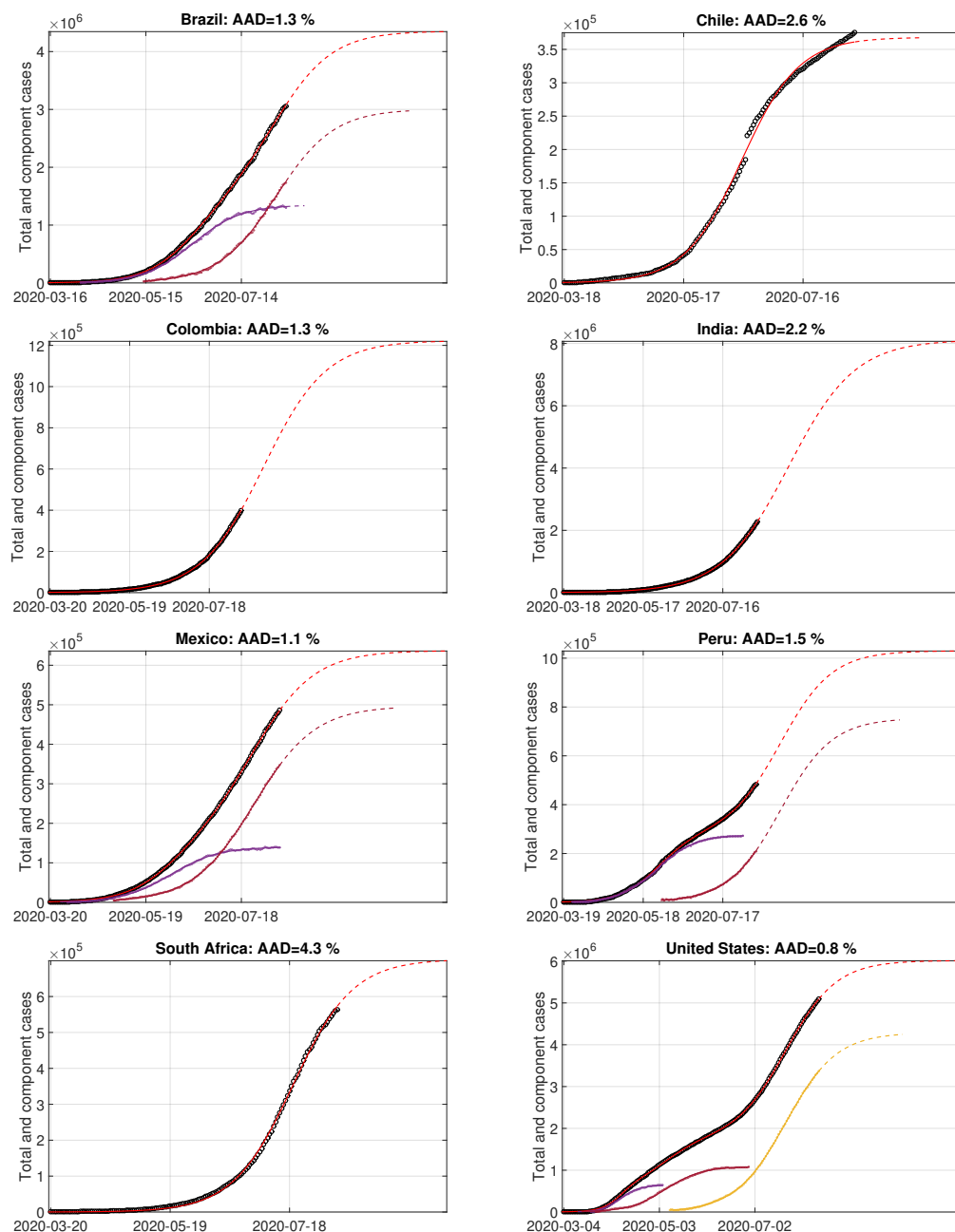
**Figure 5.** The plots demonstrating daily registered incidences of COVID-19 in Russia subdivided accordingly to principal socio-geographical division.

157 huge transport hub); this peak's behaviour also corresponds to the model results provided in Postnikov  
 158 (2020), where only the data up to the first decade of April were used. Further, the spreaders simultaneously  
 159 distributed the virus over Moscow (secondary peak of the red curve in Fig. 5) and over the country with a  
 160 prevalence to regions closer to Moscow (the European part, blue curve in in Fig. 5). Note, however, that  
 161 this region consists of multiple administrative units and, whence, the blue blue curve has not so accurate  
 162 bell-like shape as one city, Moscow. The third (green) curve corresponds to more distant from Moscow  
 163 and more weakly connected Ural, Siberian, and Far East part of Russia, and it demonstrates the larger  
 164 time delay in the outbreaks' development. At the same time, more-or-less uniform restrictive measures  
 165 "synchronize" the parameters of SIR model even for distinct weakly or non-connected locations whereby a  
 166 sum of almost independent data provides an opportunity to use the mean-field description resulting in the  
 167 accurate sequence of logistic representations for the countrywide data shown in Fig. 4. At the same time,  
 168 it should be kept in mind that the parameters  $K$  from Table 2 are not directly associated with particular  
 169 geographic regions within such a multiple-SIR but mean-field approach, they represent a characteristic  
 170 mixture of potentially susceptible groups distributed between different regions but simultaneously actual  
 171 for the particular stage of the epidemic spread.

## 172 DISCUSSION

173 The examples given above reveals that the simple logistic regression with a unique fitting function  
 174 corresponding to the solution of the SIR model can not give a fully adequate picture of the outbreak  
 175 in its full time course even if it was accurately applicable during almost complete initial growth phase,  
 176 see Postnikov (2020) for the same countries. The main reasons for this conclusion are the following: i)  
 177 the epidemics spreads over spatially distributed countries while the reported data curves gives summed  
 178 values, i.e. and the validity level of a mean-field model depends on a possible contrast in the course of  
 179 developed epidemics in different part of the country (more "compact" countries are more uniform in  
 180 the response that more "elongated" as seem from the comparison of Spain and Russia above and other  
 181 examples discussed below); ii) effects of counter-epidemic measures can change the outbreak size and  
 182 prolong its continuity that is reflected in a less value of the parameter  $k$  in Eqs. (1)–(2), or, equivalently  
 183 Eq. (4) that is also visible in Tables 1 and 2: the values of  $k$  for the second Spanish component and the  
 184 third-fifth Russian components are practically the same (the governmental restrictive measures were more  
 185 or less similar in the both countries).

186 However, these complications do not mean that the considered disease does not follow the classic  
 187 well-established SIR model of mathematical epidemiology (and its logistic-form) and the mean-field  
 188 consideration, which addresses the country-wide data. The analysis carried out in the present work  
 189 revealed that a more realistic data regression requires the decomposition over multiple logistic curves.  
 190 Each of such curve corresponds to its own SIR system with its own set of parameters. Among these  
 191 parameters, the possible maximal number of susceptible person within weakly connected communities  
 192 (on even approximated as uncoupled ones for simplicity of fitting) is one of the most important. Such  
 193 situation can have an interpretation in terms of metapopulation approach Colizza and Vespignani (2008);  
 194 Bichara et al. (2015); Ball et al. (2015); Chowell et al. (2016). In fact, the country-wide statistics unifies



**Figure 6.** The plots demonstrating the cumulative numbers of cases (circles), the respective solutions of logistic and multilogistic models (red solid line, the dashed part corresponds to the projection based on the simple continuation with the present state parameters) and their components (if any; dots denotes the sorted data and curves are the simulated continuous solutions) for eight countries among top-10 most affected ones. The parameters for all models/components are given in supplementary material on GitHub project (see “Methods”) as well as the data and code for their processing.

195 data from different regions and even from different social strata within each region. Due to differences  
 196 and different connectivity between such sub-populations, the local outbreaks can be shifted in time and  
 197 intensity. This problem especially arise under conditions of different measures of travel restrictions and  
 198 social distancing applied practically in all countries during COVID-19 pandemics.

199 Fig. 6 provides demonstrations additional to the considered above and covers the most affected

countries. One can see (not only visually but also basing on the values of the average relative absolute deviations) that the proposed decomposition reproduces the specific features of the cumulative number of cases with high accuracy and the multilogistic representation (i.e. superposition of several SIR systems associated with different sub-communities) is a widespread phenomenon. A single curve corresponds in the cases of Columbia and India to the situation, when the outbreak has not reach the peak yet(or, equivalently, the middle level with respect to the saturation estimated from the exiting data fitting). However, it should be pointed out that the forecast of the maximum total number of infected in such a case has a high level of uncertainty, see numerical tests in Alberti and Faranda (2020); Postnikov (2020) and an additional analysis of sensibility of sigmoid, including logistic, curves to perturbations in Faranda et al. (2020). This caveat is possibly true also for the second stage of the two-component representation of data for Peru.

Among the rest countries considered, where the epidemic curve looks overcoming the middle point, three of six demonstrate the multilogistic character: two components for Brazil and Mexico, and three for the USA. The latter is more spatially spread and, additionally, social turbulence of last months may influence the parameters of the third component there.

## CONCLUSION

The main message of this work is aimed at attracting attention to the importance of multimodal consideration of epidemic curves given at the macroscopic (country-wide) level and developing a method, which allows revealing time intervals of qualitative changes in dynamics as well as their quantitative estimations.

The complicated non-standard shape of epidemic curves not necessarily originates from some special features of the current pandemics providing a challenge in developing new very complicated multicomponent models, which contain a large number of parameters inaccessible directly. In fact, when one addresses the global mean-field data on the “macroscopic” scale of a country, a set of simple SIR model is plausible. In its turn, the number and temporal localisation of the revealed principal modes provide perspectives for building more detailed “microscopic” models, which should take into account spatial, social, etc. features, i.e. to reflect phenomena originated from the metapopulational dynamics. Note that the summation of individual growth curves attracted attention as an empiric approach from the classical problems of pure population dynamics of heterogeneous populations Reed and Pearl (1927) up to the mathematical problems of epidemic spread in metapopulations Chowell et al. (2016). In this work, it is shown that the effective logistic curve-based approach could play a role of an accurate tool for studying the recent COVID-19 pandemics in a tight connection with the classic SIR model approach. In contrast to the direct simulation of ODE’s including the case of parameter switching in *a priori* known time points (e.g. due to lockdown measures, see Dehning et al. (2020)), the proposed method solves the inverse problem revealing actual parameters and inflection points originated from the actual data.

Finally, this opens certain perspectives for future investigations in modelling by analogy to other field of science, which use dynamical systems as a principal tool. As some examples, the dimension reduction to several principal modes is used for simplified modelling of turbulence Sapsis and Majda (2013), distributed elastic mechanical systems Kerschen and Golinval (2002), chemical processes Qing et al. (2020), etc. Thus, the revealed possibility of describing the cumulative dynamics of the epidemics with a finite number of components calls for development of spatio-temporal models whose full-scale dynamics is reducible to the dynamics of several principal modes only reflected in the generalized data.

## ACKNOWLEDGEMENTS

I thank Dr. Irina L. Zhirnova for the motivational discussions of the epidemic growth curves dynamics and composition.

## REFERENCES

- Alberti, T. and Faranda, D. (2020). On the uncertainty of real-time predictions of epidemic growths: a COVID-19 case study for China and Italy. *Communications in Nonlinear Science and Numerical Simulation*, page 105372.
- Ball, F., Britton, T., House, T., Isham, V., Mollison, D., Pellis, L., and Tomba, G. S. (2015). Seven challenges for metapopulation models of epidemics, including households models. *Epidemics*, 10:63–67.

- 251 Bedford, J., Enria, D., Giesecke, J., Heymann, D. L., Ihekweazu, C., Kobinger, G., Lane, H. C., Memish,  
252 Z., Oh, M.-d., Schuchat, A., Ungchusak, K., and Wieler, L. H. (2020). COVID-19: towards controlling  
253 of a pandemic. *Lancet*, 395(10229):1015–1018.
- 254 Bichara, D., Kang, Y., Castillo-Chavez, C., Horan, R., and Perrings, C. (2015). SIS and SIR epidemic  
255 models under virtual dispersal. *Bulletin of Mathematical Biology*, 77:2004–2034.
- 256 Brauer, F., van den Driesche, P. V., and Wu, J., editors (2008). *Mathematical Epidemiology*. Springer,  
257 Berlin.
- 258 Carcione, J. M., Santos, J. E., Bagaini, C., and Ba, J. (2020). A simulation of a covid-19 epidemic based  
259 on a deterministic seir model. *arXiv:2004.03575*.
- 260 Chowell, G., Sattenspiel, L., Bansal, S., and Viboud, C. (2016). Mathematical models to characterize  
261 early epidemic growth: A review. *Physics of Life Reviews*, 18:66–97.
- 262 CNE (2020). <https://cnecovid.isciii.es/covid19/>.
- 263 Cobey, S. (2020). Modeling infectious disease dynamics. *Science*, 368(6492):713–714.
- 264 Colizza, V. and Vespignani, A. (2008). Epidemic modeling in metapopulation systems with heterogeneous  
265 coupling pattern: Theory and simulations. *Journal of Theoretical Biology*, 251:450–467.
- 266 DataLens, Y. (2020). <https://datalens.yandex/7o7is1q6ikh23>.
- 267 Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., and Priesemann, V. (2020).  
268 Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*,  
269 369:eabb9789.
- 270 Fanelli, D. and Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and  
271 France. *Chaos, Solitons & Fractals*, 134:109761.
- 272 Faranda, D., Castillo, I. P., Hulme, O., Jezequel, A., Lamb, J. S., Sato, Y., and Thompson, E. L. (2020).  
273 Asymptotic estimates of SARS-CoV-2 infection counts and their sensitivity to stochastic perturbation.  
274 *Chaos*, 30:051107.
- 275 Fisher, J. C. and Pry, R. H. (1971). A simple substitution model of technological change. *Technological  
276 forecasting and social change*, 3:75–88.
- 277 Gates, B. (2020). Responding to Covid-19 – a once-in-a-century pandemic? *New England Journal of  
278 Medicine*, 382(18):1677–1679.
- 279 He, F., Deng, Y., and Li, W. (2020). Coronavirus Disease 2019 (COVID-19): What we know? *Journal of  
280 Medical Virology*, 76(71-76).
- 281 Kerschen, G. and Golinval, J.-C. (2002). Non-linear generalization of principal component analysis: from  
282 a global to a local approach. *Journal of Sound and Vibration*, 254:867–876.
- 283 Lavrova, A. I., Postnikov, E. B., Manicheva, O. A., and Vishnevsky, B. I. (2017). Bi-logistic model  
284 for disease dynamics caused by Mycobacterium tuberculosis in Russia. *Royal Society open science*,  
285 4:171033.
- 286 Linka, K., Peirlinck, M., Sahli Costabal, F., and Kuhl, E. (2020). Outbreak dynamics of covid-19 in europe  
287 and the effect of travel restrictions. *Computer Methods in Biomechanics and Biomedical Engineering*,  
288 pages 1–8.
- 289 Maier, B. F. and Brockmann, D. (2020). Effective containment explains subexponential growth in recent  
290 confirmed COVID-19 cases in China. *Science*.
- 291 Mamon, G. A. (2020). Fit of french covid-19 hospital data with different evolutionary models: regional  
292 measures of  $r_0$  before and during lockdown. *arXiv:2005.06552*.
- 293 Meyer, P. S., Yung, J. W., and Ausubel, J. H. (1999). A primer on logistic growth and substitution: the  
294 mathematics of the Loglet Lab software. *Technological forecasting and social change*, 61:247–271.
- 295 Oliveira, G. (2020). Refined compartmental models, asymptomatic carriers and covid-19.  
296 *arXiv:2004.14780*.
- 297 Postnikov, E. B. (2020). Estimation of COVID-19 dynamics “on a back-of-envelope”: Does the simplest  
298 SIR model provide quantitative parameters and predictions? *Chaos, Solitons & Fractals*, 135:109841.
- 299 Qing, X., Jin, J., Niu, Y., and Zhao, S. (2020). Time–space coupled learning method for model reduction  
300 of distributed parameter systems with encoder-decoder and RNN. *AIChE Journal*, page e16251.
- 301 Reed, L. J. and Pearl, R. (1927). On the summation of logistic curves. *Journal of the Royal Statistical  
302 Society*, 90:729–746.
- 303 Roser, M., Ritchie, H., Ortiz-Ospina, E., and Hasell, J. (2020). Coronavirus Pandemic (COVID-19). *Our  
304 World in Data*. <https://ourworldindata.org/coronavirus>.
- 305 Sapsis, T. P. and Majda, A. J. (2013). Statistically accurate low-order models for uncertainty quantification

- 306 in turbulent dynamical systems. *Proceedings of the National Academy of Sciences USA*, 110:13705–  
307 13710.
- 308 Schaback, R. (2020). On covid-19 modelling. *arXiv:2005.07004*.
- 309 Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Iosifidis, C., and Agha, R.  
310 (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus  
311 (COVID-19). *International Journal of Surgery*, 76:71–76.
- 312 Vannucci, V. and Vannucci, L. (2020). Estimation of the ending time of Covid-19 emergence in Italy: a  
313 2-waves model as sum of logistics. (doi:10.13140/RG.2.2.26675.58403).
- 314 Wang, L.-s., Wang, Y.-r., Ye, D.-w., and Liu, Q.-q. (2020). A review of the 2019 Novel Coronavirus  
315 (COVID-19) based on current evidence. *International Journal of Antimicrobial Agents*, page 105948.