Brazilian Journal of Physics (2022) 52: 15
https://doi.org/10.1007/s13538-021-01010-6

GENERAL AND APPLIED PHYSICS

Uncertainty Reduction in Logistic Growth Regression Using Surrogate Systems Carrying Capacities: a COVID-19 Case Study

Bruno Hebling Vieira[1] · **Nathalia Hanna Hiar**[2] · **George C. Cardoso**[1]

Received: 4 June 2021 / Accepted: 12 October 2021 / Published online: 30 November 2021

Abstract
Logistic growth regressions present high uncertainties when data are not past their inflection points. In such conditions, the uncertainty in the estimated carrying capacity K, for example, can be of the order of K. Here, we present a method for uncertainty reduction in logistic growth regression using data from a surrogate logistic growth process. We illustrate the method using Richards' growth function to predict the inflection points of COVID-19 first-wave accumulated causalities in Brazilian cities. First waves of epidemics are known to be reasonably well modeled a posteriori by Richard's growth function. Yet, we make predictions using early data that end before or around the inflection point. For that goal, we estimate K by logistic growth regression using data from surrogate international cities where the epidemics are clearly past their inflection points. The constraint stabilizes the logistic growth regression for the Brazilian cities, reducing the uncertainty in the prediction parameters even when the surrogate K is a rough estimate. The predictions for COVID-19 first-wave peaks in Brazilian cities agree with official data. The method may be used for other logistic models and logistic processes, in areas such as economics and biology, when surrogate populations or systems are identified.

Keywords Decision-making · Richard's curve · Logistic functions · Epidemic peaks · Uncertainties

1 Introduction

Logistic models describe monotonic growth processes that present an accelerating phase, an inflection point, a decelerating phase, and asymptotically saturate to a maximum value K, called the carrying capacity. Examples of logistic models include the Verhulst function and the Richards' curve [1, 2] that present exponential growth at the beginning of the accelerating phase. Logistic functions can adequately model processes such as population growth, tumor growth, and disease spreading [3–5].

Classical epidemic descriptions use compartmental models that assume a homogeneous contact network, thus being inadequate to deal with an emerging epidemic in an unknown complex network [6]. Alternative approaches model disease propagation using agent networks but require knowledge of pathogen transmission probability, and of the social network [7–9]. Moreover, in actual epidemics, there is often a scarcity of data and reduced knowledge about the disease. Logistic models have successfully described the cumulative curve of cases $C(t)$ of the first wave of immunizable diseases [2, 5, 10, 11]. Indeed, logistic models are adequate to make short and mid-term predictions of $C(t)$ in the decelerating phase, using only historical data and a few fitting parameters [4, 5]. However, a logistic model cannot predict the inflection point of a logistic growth that is still in the accelerating phase [2, 12]. Uncertainties in predictions using data regression would be unacceptably large, and the uncertainty in the carrying capacity K would be of the order of K [4]. This shortcoming is an unfortunate limitation. Because estimating the time and magnitude of the maximum growth rate (inflection point) is an important goal in logistic growth modeling. In an epidemic example, such

✉ George C. Cardoso
gcc@usp.br

Bruno Hebling Vieira
bruno.hebling.vieira@usp.br

Nathalia Hanna Hiar
nathaliahiar@usp.br

[1] Department of Physics, FFCLRP, Universidade de Sao Paulo, Ribeirao Preto, SP 14040-901, Brazil

[2] Department of Biology, FFCLRP, Universidade de Sao Paulo, Ribeirao Preto, SP 14040-901, Brazil

information would give the approximate date and the peak number of active cases per week.

Here we use independent estimates of the carrying capacity K of a logistic model, using a surrogate population to stabilize a logistic growth regression on a different data series that is still in the acceleration phase. "The aim of this paper is to suggest a method to work around these intrinsic limitations logistic functions present. This is not a paper on epidemics, but for convenience we use COVID-19 epidemic data as real-world cases that can be approximated by logistic functions. The logistic function chosen for illustration and evaluation of the methodology is the Richard's curve function, because it is one of the simplest functions that have been successful in making short-term predictions in epidemics [2, 4]. As an illustration of the process, we apply the Richard's curve function to the COVID-19 first-wave epidemic data of Brazilian cities still in the acceleration phase, to make short-term predictions. As surrogate populations, we select Western countries and cities where the inflection point of the first COVID-19 wave has been identified, to estimate plausible values for K. We chose the Richard's curve for convenience, because it has been shown in the literature to work well for COVID-19 data regression [4, 13, 14]. The carrying capacity K for a city is normalized by the city's total population. For reliability, we use official accumulated fatalities count data instead of the accumulated number of cases, and correct fatalities rates using demographic information. Predictions of the inflection point using a range of surrogate carrying capacities agree with actual observed data. Our key contribution to epidemics is offering an accessible early prediction tool to gauge ranges of plausibility for epidemic peak timing and magnitude in late-onset, low-resources cities. Other applications for the proposed method include forecasts for processes known to obey logistic functions, but still in their accelerating phases. Identification of a proper surrogate process is beyond the scope of this paper and depends on the knowledge about the processes under investigation.

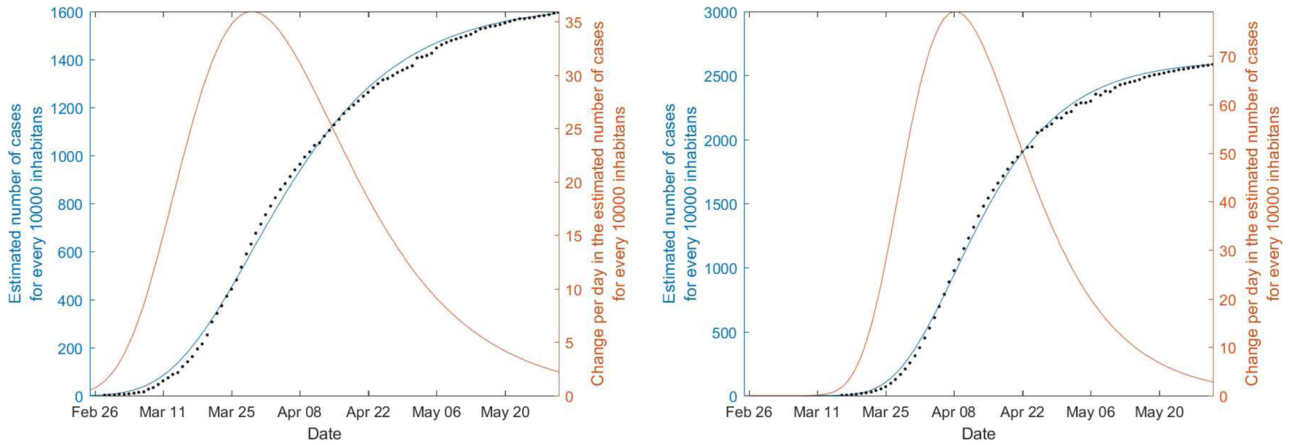
2 Methods

2.1 Logistic Function for Data Regression

We use the Richard's curve [1], which is a type of logistic curve compatible with first-wave data for airborne transmissible diseases [2, 5]:

$$Y(t) = \frac{K}{\left(1 + \alpha e^{-\alpha r (t - t_c)}\right)^{\left(\frac{1}{\alpha}\right)}}, \tag{1}$$

where K is an asymptotic value that represents the carrying capacity, r is the initial growth rate, α depends on the epidemic dynamics in the network, and t_c is the abscissa for the inflection point. Two examples of Richards' curves and their time derivatives are shown in Figure 1. Notice the adequate quality of the regressions, which we assume to have been because interventions were soft or, once started, remained constant. To stabilize the logistic growth regressions of Equation 1 to data of cities of interest still in the accelerating phase of an epidemic, we use K values derived from logistic growth regressions to data from surrogate cities, where the epidemic has reached the deceleration phase. A key assumption, which will be shown to be reasonable a posteriori, is that K written as a fraction of the population is



(a) Lombardy, Italy **(b)** New York, United States of America

Fig. 1 (Color online) Examples of Richard's curve regressions to COVID-19 data of select regions, and respective time derivative. The proportion of fatalities divided by the infection fatality rate is shown in black dots. Notice the adequacy of the regressions to the data-series

approximately the same both for the surrogate city and for the city of interest.

2.2 Data

We have analyzed fatality data from select regions from the USA, Italy, Spain, the UK, and Belgium, which had reached the decelerating phase of fatalities per day. Data for the Queens borough, NY, USA, were collected from the US census (https://www.census.gov/), and corresponding COVID-19 data we collected from the NYC government website (https://www1.nyc.gov/site/doh/covid/).

For the remaining locations above, the data were collected from the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (https://coronavirus.jhu.edu/). Populations and fatality data for Brazilian cities were collected from official census data https://www.ibge.gov.br/cidades-e-estados, from Oswaldo Cruz Foundation (COVID-19), Rio de Janeiro, https://bigdata-api.fiocruz.br/relatorios, and official death registrations https://transparencia.registrocivil.org.br. The period analyzed comprises the period between the 15th registered COVID-19 death in each city and June/21/2020, for all cities included in this study.

2.3 Modeling

A major issue with modeling of early stages of pandemic outbreaks is that the carrying capacity K (Figure 1) must be determined by data regression alone. This is a major source of uncertainty in the model. An alternative is the use of exponential growth modeling ($K \to \infty$), which forsakes the identification of the inflection point. We proposed instead to model different scenarios for a city of interest based on plausible values for K derived by fitting a logistic function on data from surrogate cities where the epidemic is clearly past the inflection point.

For the regression analyses, we use nonlinear least-squares optimization to estimate the free-parameters (K, α, r, t_c) in Equation 1 for the surrogate cities, and to estimate (α, r, t_c) for the cities of interest, where in the latter case K was kept fixed. Optimization was performed with a trust-region-reflective algorithm, implemented in the function "lsqcurvefit", from the Optimization Toolbox in MATLAB (2015). This optimizer, albeit more complex than the traditional Levenberg–Marquardt-based ones, can better cope with the negative curvature of the estimate of the Hessian of the objective (for a quick primer, see [15]). Variables α and r are both constrained to the [0, 1] interval in the model, for stability [5]. Because t_c is not constrained, it may be orders of magnitude larger than α and r; we divide its gradient by 10^3. Otherwise, its gradient magnitude would dominate over other terms.

Since our objective function is non-convex on the parameters, one hundred random initializations were tried. Local minima initializations were not counted, but their resulting models remained in the pool of candidate models. Model selection was done by minimizing the Akaike information criterion. In our particular case, this approach resulted in capturing the model with the set of parameters giving the smallest mean-square error, given the Gaussian error assumption of nonlinear least-squares.

The cumulative number of reported cases ($C(t)$) is notoriously susceptible to testing and sampling biases. Thus, instead of using $C(t)$, to estimate the inflection points, we used cumulative fatalities curves $D(t)$ to determine both the surrogate K's and the regressions using data of Brazilian cities of interest. The number of fatalities $D(t)$, while not completely free from biases, is more robust to them. Infection Fatality Rates (IFR) for Brazilian locations were estimated to be $\approx 1\%$, based on IFR and demographics of Chinese and Italian locations, corrected to match the demographics of Brazilian locations (See supporting information and data). We used $C(t - t_D) = D(t)/\text{IFR}$ to estimate the cumulative number of reported cases in the plots. The delay t_D between diagnosis and death is a further uncertainty in the determination of $C(t)$.

2.4 Estimating the Inflection Point: the Epidemic Peak

In Equation 1, t_c is the inflection point: the time when the derivative $Y'(t) = dY(t)/dt$ is maximal, as shown in the curve in Figure 1. However, derivatives amplify noise—especially given that the fitting parameters contain uncertainties, and COVID-19 data are certainly noisy. Thus, instead of directly determining dY/dt, we chose to determine an average number of fatalities per day, ΔY_{peak}, in an interval approximately equal to the duration of the disease [16]: $T = 15$ days centered around t_c, as shown in Equation 2.

$$\Delta Y_{\text{peak}} = \frac{Y(t_c + T/2) - Y(t_c - T/2)}{T}, \tag{2}$$

where t_c is determined from data regression using Equation 1, using (or not) a surrogate K. Regression to real-world data with Equation 1 is extremely likely to have uncertainties in the fit parameters, even for fixed K assumed without uncertainty. The uncertainty of ΔY_{peak} can be approximated by the variance formula:

$$\sigma^2_{\Delta Y_{\text{peak}}} = \frac{\sigma^2_{Y(t_c+T/2)} + \sigma^2_{Y(t_c-T/2)}}{T^2} \tag{3}$$

where the uncertainties $\sigma^2_{Y(t)}$ can be obtained from equation (1), using the fitting parameters and their respective

uncertainties, and the conventional Delta method for uncertainty propagation [17].

3 Results

First, we used the Richard's curve to fit the logistic growth data of epidemic evolution in surrogate cities where the epidemic is beyond its inflection point. Results for K values for surrogate cities are shown in Table 1 for data, corresponding to the first wave of COVID-19. Using surrogate values for K, we performed Richard's curve regression to Brazilian cities' fatalities curves starting from the onset of COVID-19 until June, 21/2020. Uncertainties in surrogate K values were not explicitly taken into account in our analysis. In practice, the ranges of plausibility, such as $10\% < K < 25\%$ that we discuss, act as uncertainty in the surrogate Ks.

Figure 2 shows Richard's curve regressions for different fixed values of K, and also for K as a free parameter (automatic), for select Brazilian cities. Figure 3 shows the prediction of the day of the epidemic peak, given by t_c in Equation 1, and the estimate of the number of fatalities averaged in an interval of $T = 15$ days centered on t_c, as described in Equation 2. Figures for other cities studied can be found in the supporting information and data. Our predictions for Brazilian cities agree with the actual observed curves by October/2020, with uncertainties for the inflection points that are up to tens of times narrower than when K is left as a free parameter.

4 Discussion

Logistic functions model phenomena with initial exponential growth and eventual saturation (carrying capacity) [2], presenting an inflection point. However, predictions using logistic regressions to data that end before the inflection point result in overfitting and lead to uncertainties of the order of 100% of the mean values [5]. The duration of the exponential growth, and therefore the carrying capacity K, is overestimated. Here, we discuss the use of plausible ad hoc K values extracted from surrogate physical systems expected to be representative of the system of interest. Using plausible surrogate K values for logistic regressions of the data of interest creates plausible prediction scenarios for the evolution of the data. In this discussion, we show that the use of surrogate K values improves at least short-term prediction, and can improve prediction of the inflection point. For the sake of definiteness, we chose examples using data from the COVID-19 pandemic, and Richards' curve as the logistic function—which has been shown to fit well the cumulative cases curve $C(t)$ for epidemics [5]. Despite the challenge presented by an epidemic context, predictions made with surrogate K values are compatible with later observed results.

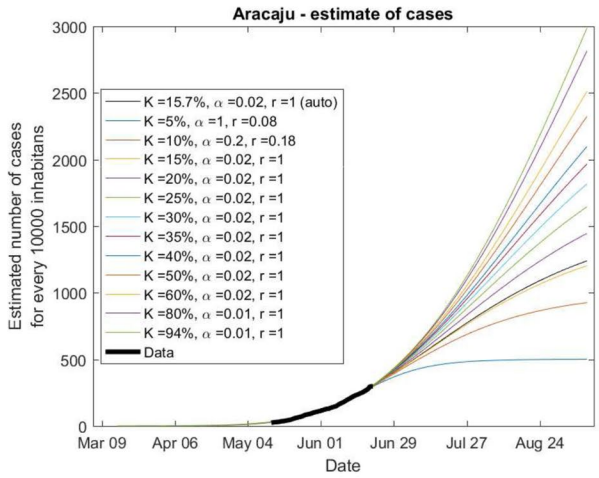
The Richards' curve data regression with all parameters free over-predicts K when the accumulated cases $C(t)$ are still in the exponential stage (Fig. 2(a) and (b)—K = 100%). Our goal is not to predict K, but to decrease the overestimate of K to avoid overestimates of C(t) predictions even in the short term (Fig. 3(a) and (b)); also, an uncertainty in K would be carried to any prediction made from the data. Surrogate carrying capacities K from select international cities, which have epidemic data well past their inflection points (Table 1, $5\% < K < 26\%$), can be hypothesized to present plausible evolution scenarios for the cities of interest, at least for the short term (see curves for $5\% < K25\%$ in Fig. 2). Again, surrogate cities' carrying capacities K do not necessarily predict those cities' epidemics' final sizes but represent the best logistic regressions' K to be used for short and midterm predictions in the cities of interest. Notice that some low surrogate K can be eliminated from the possibilities. For example, in Fig. 2(c) and (d) $C(t) > 10\%$, therefore surrogates $K <= 10\%$ are implausible.

Surrogate international cities that were well advanced in the epidemic (such as in Fig. 1) have $5\% < K < 26\%$ (Table 1), which encompass the K of Brazilian cities that just passed the inflection point (Belém, Fortaleza, and Rio de Janeiro, Fig. 3(c)

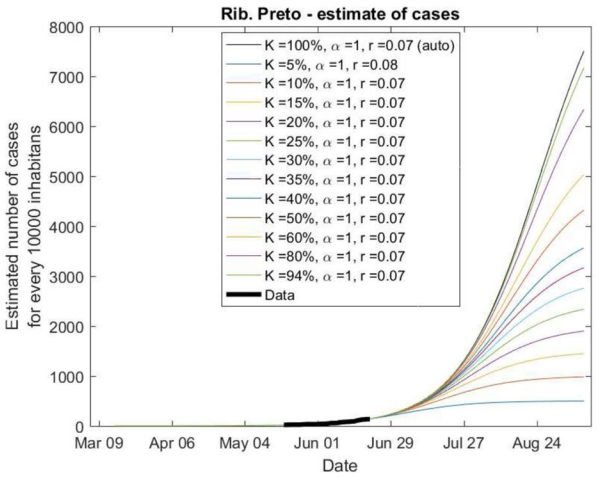
Table 1 Richard's curve regression on data of international localities. The estimated carrying capacity K, presented as a fraction of the total population of the location, the 95% confidence interval is show in parentheses

Country	City	Carrying capacity (K)
USA	Los Angeles	0.0378 (0, 0.2180)
	New York	0.2617 (0.1326, 0.3907)
	Rockland	0.2011 (0, 0.4554)
	Philadelphia	0.1073 (0, 0.2853)
	Fairfield	0.1456 (0.0578, 0.2334)
	Essex (MA)	0.1434 (0.0377, 0.2491)
Italy	Lombardia	0.1627 (0.0504, 0.2750)
Spain	Madrid	0.2302 (0, 1)
UK	London	0.0672 (0.0006, 0.1338)
Belgium	-	0.0832 (0, 0.1683)

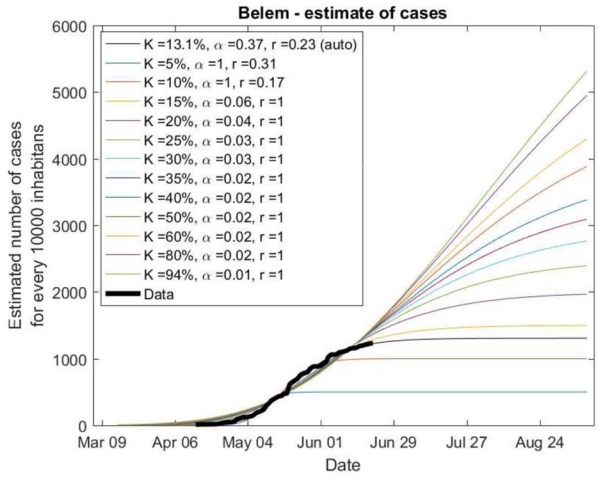
Fig. 2 (Color online) Regressions of Richard's Growth function to COVID-19 time series. Except for the curves represented by "auto", K (the "carrying capacity") is not used as a regression variable. Observe that each regression made with fixed K represents a possible growth outcome. Surrogate cites can be used to establish which K range is more plausible. Notice in (**b**) that for a time-series still in the beginning of the accelerating phase, K =100% when left as variable (auto), which is not expected in epidemic problems. For fixed K, the inflection point changes position with K, but this change cannot be determined analytically because it depends on the regression to the data. The uncertainties in the regressions are not shown



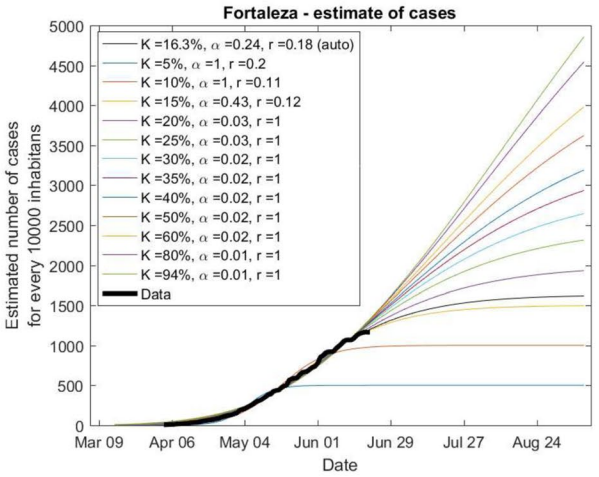
(a) Aracaju



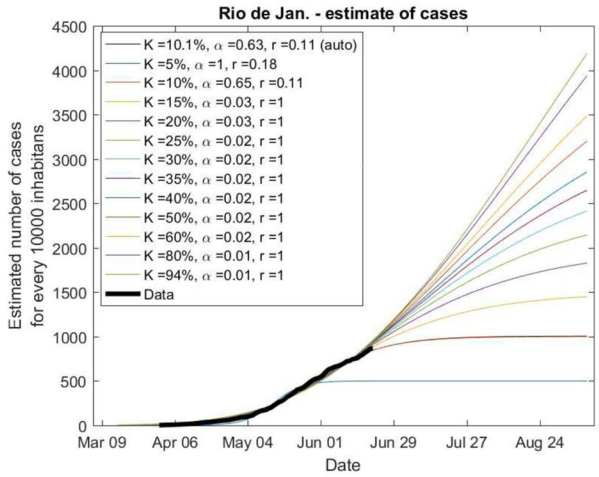
(b) Ribeirão Preto



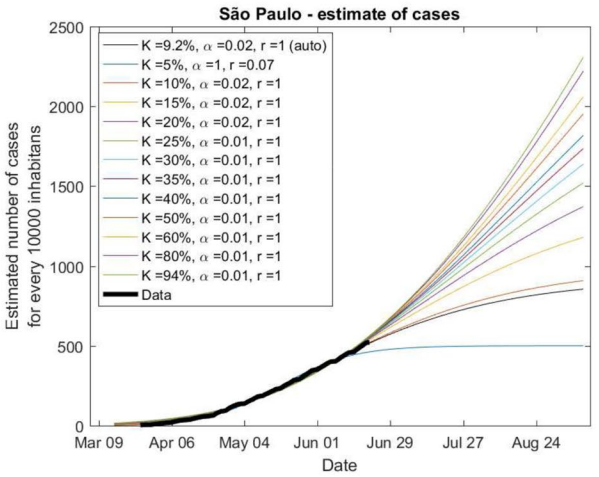
(c) Belém



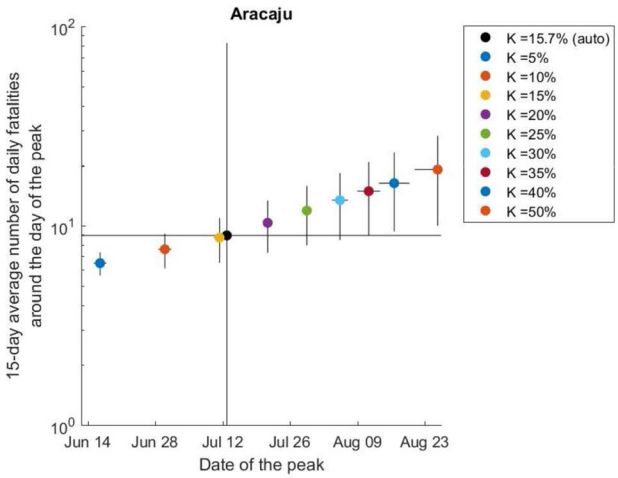
(d) Fortaleza



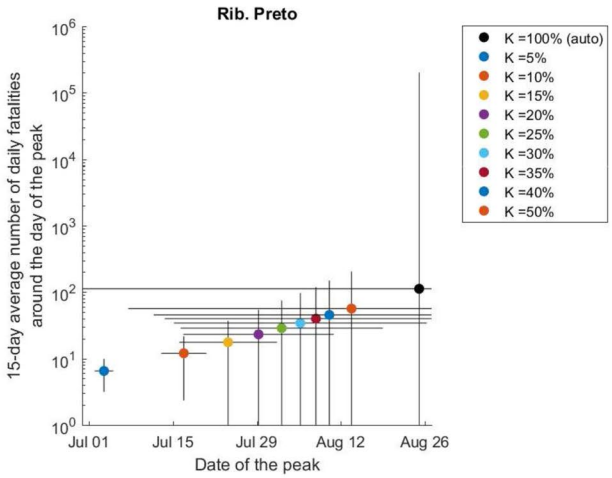
(e) Rio de Janeiro



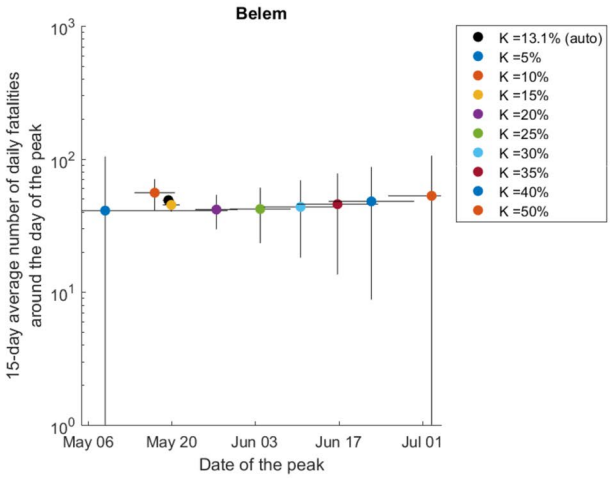
(f) São Paulo



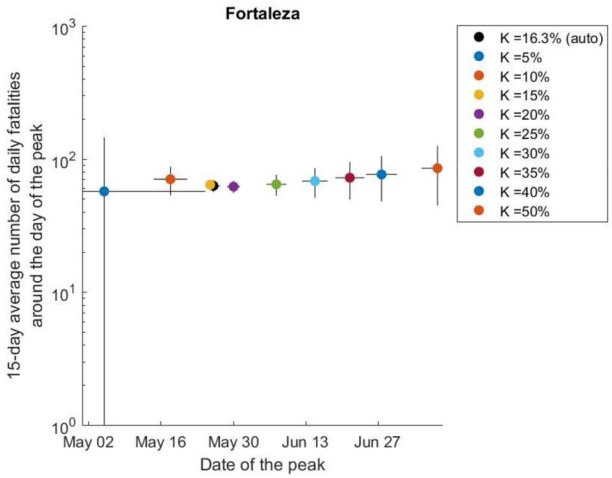
(a) Aracaju



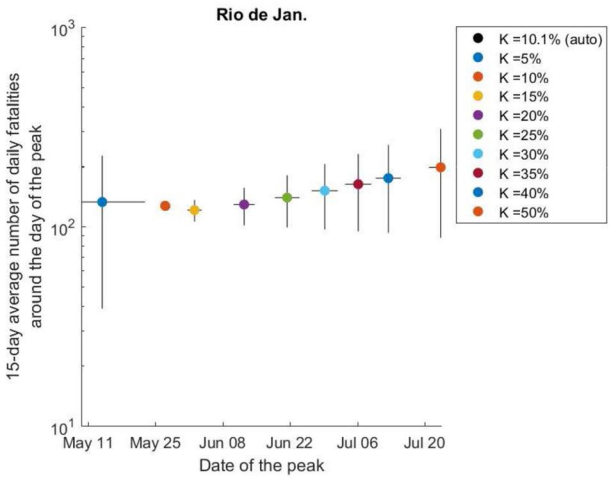
(b) Ribeirão Preto



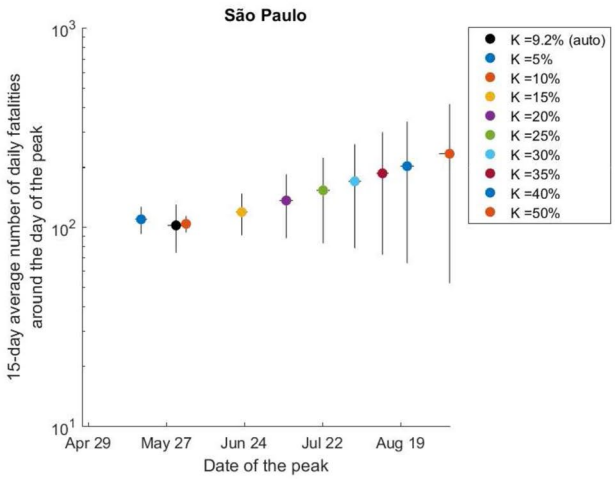
(c) Belém



(d) Fortaleza



(e) Rio de Janeiro



(f) São Paulo

◂**Fig. 3** (Color online) Scenarios for epidemic peak for Brazilian cities. Each point represents a short-term prediction for the peak of the derivative of a fitted logistic function vs. time. The black dot (auto) represents a model where the carrying capacity K was automatically optimized as a parameter. Results are averaged over 15 days centered on the peak day. Notice that in (**a**) and (**b**) the automatic K (without use of surrogate K) have large uncertainties compare to the same cities when a surrogate K is used. This behavior indicates (**a**) and (**b**) are still in the accelerating phase of growth. Cities (**c**) - (**f**) are near the inflection point. We can also observe possible scenarios for larger K values. Such short-term predictions are stable with respect to small increases in the surrogate K

to (e), with free parameter fit resulting in $10\% < K < 16\%$). We interpret the congruence between Brazilian and international cities as a sign of plausibility of the surrogate method. Now, because they are socio-economically closer, we can use these Brazilian cities as surrogates to narrow down plausible (short-term) predictions for the inflection points of Aracaju and Ribeirao Preto (Fig. 3(a) and (b)). Notice the substantial decrease in the range of predicted peak (inflection point) weeks, even for $K = 25\%$, and the relative stability in the prediction of the inflection point for small changes in K. Later observed data for Aracaju and Ribeirao Preto—peak on July, 2nd with 14 fatalities/day, and a broad peak in July with ~ 10 fatalities at the peak [19], respectively—are compatible with $K \simeq 12\%$. Finally, auto-fit for the city of São Paulo showed a larger than typical uncertainty in the inflection point's position for $K = 9.2\%$ prediction, suggesting that the estimate may be improved (because the fit is good in the region with data). Actually, for Sao Paulo city, a broad was reported in June [19], with a 7-day moving average of 123 fatalities/day, which is also compatible with $K = 12\%$. We can argue that surrogate "carrying capacities" enable not only plausible scenarios but also good short-term predictions for systems describable by logistic growth, which are still at or behind their inflection points.

COVID-19 has revealed that sophisticated epidemic models, in the absence of granular agent-level knowledge of the disease, might not be enough to make good epidemic predictions [18, 23]. For example, a Bayesian hierarchical model based on Richards' growth curve, and integrating multiple countries' data for infection trajectory prediction of the spread of COVID-19 [20], underestimated the number of cases in main COVID-19 countries by a factor of at least five, even for countries of interest that were past their inflection points. The fact that the surrogate K method produced good results in such a challenging context suggests that it would be adequate for other logistic growth problems.

We chose to illustrate the surrogate process using Richards' curve for its simplicity. The Richards' curve is one of the simplest logistic functions that adequately represent epidemic data [2]. Limitations to the Richards' curve includes the narrow range for the height of its inflection point that is limited between $K/2$ (for $\alpha = 1$) and K/e (for $\alpha \to 0$) for α values compatible with epidemic regressions [2]. In our study, α vs. r remained unresolved, but we observed that the product αr is proportional to the slope of the logistic function [20]. In addition, different from [12], here K is pre-determined, allowing for continuous transitions between curves for different K's. Because of the difficulty in determining both α and r independently, in some references $\alpha r = r_o$, where r_o is not represent the initial growth of the process [2]. Such limitations did not seem to affect short-term predictions using surrogate Ks. If needed, the surrogate system method may use more sophisticated logistic models, such as using generalized exponential [21, 22].

To build Fig. 2, we used infection fatality rates (IFR) based on Italian and Chinese cities' IFR, and corrected for the demographics of Brazilian cities, reaching an of 1% IFR for Brazilian cities of interest (supporting information and data). Saturation in the health system would change IFR, making it important to determine IFRs before the health system saturates. For some poorer cities, saturation in the health system may occur before the inflection point. Thus, the surrogate K technique is especially important, especially if a non-saturated surrogate city is identified. For Fig. 3 we used the actual fatality rate assuming no saturation, to avoid the use of uncertain IFR. One limitation in our study is not having considered possible saturation of the health systems. However, such issues might be irrelevant for other logistic growth problems—other than epidemics.

We cannot tell a priori which cities (or systems) are adequate "surrogates". In an epidemic case, surrogate city selection is based on sociocultural aspects and is out of the scope of this discussion. Our choice of the international cities listed in Table 1 was heuristic. In the absence of a surrogate city (or system), the surrogate carrying capacity K could, in principle, be estimated from an appropriate theoretical model. As discussed, even a range of plausible Ks improves prediction. At the exponential phase of the growth, a logistic function with surrogate K is expected to provide better predictions than a free K logistic regression, which in the exponential phase of the growth is no better than a simple exponential regression. The "surrogate K" method is not intended to predict the carrying capacity of the system whose data is being adjusted: it is more adequate for predictions in short and intermediary timescales. In a real scenario, the logistic regressions must be rerun every time a data point is added, and sometime after the inflection point has been reached, there will be no need for a surrogate system. Adequacy of the regressions can be evaluated from the uncertainties in the fitting parameters. Short-term predictions are more reliable.

The proposed surrogate carrying capacity approach seeks to mitigate the logistic functions' over-fitting problem in the accelerating phase of the data, which causes uncertainty and overestimation of forecasted values. This surrogate prescription enables more accurate short-term

extrapolations and permits the visualization of plausible unfolding scenarios for mid-term predictions. We have discussed an illustration of the surrogate system method using COVID-19 data, and the Richards' logistic growth model. However, the method may be applied to any system obeying any logistic curve, including systems in physical, economical, and biological sciences. The lower uncertainty allowed with the surrogate system carrying capacity allows also for optimistic and pessimistic mid- and longer-term predictions without the need for sophisticated models or big data resources, and gives the experimentalist or data scientist anchor points for decision-making.

Acknowledgements We thank Professor Alexandre S. Martinez for critical reading of the manuscript. This work is partially supported by FAPESP (São Paulo Research Foundation), 18/11881-1 (BHV).

Data Availability Statement Code and data are available on: https://github.com/bhvieira/CovidRichards/

Declarations

Conflict of Interest The authors declare no potential conflict of interests.

References

1. F. Richards, J Exp Botany **10**(2), 290 (1959)
2. X.S. Wang, J. Wu, Y. Yang, J Theo Bio **313**, 12 (2012)
3. G. Chowell, Infectious Disease Modelling **2**(3), 379 (2017)
4. E. Aviv-Sharon, A. Aharoni, Infectious Disease Modelling **5**, 502 (2020)
5. G.L. Vasconcelos, A.M. Macêdo, R. Ospina, F.A. Almeida, G.C. Duarte-Filho, A.A. Brum, I.C. Souza, PeerJ **8**, e9421 (2020)
6. I. Holmdahl, C. Buckee, New England Journal of Medicine (2020)
7. G.M. Nakamura, A.C.P. Monteiro, G.C. Cardoso, A.S. Martinez, Sci Rep **7**(1), 1 (2017)
8. G.M. Nakamura, A.C.P. Monteiro, G.C. Cardoso, A.S. Martinez, Math Comp Appl **24**(2), 44 (2019)
9. G.M. Nakamura, G.C. Cardoso, A.S. Martinez, Royal Society Open Science **7**(2) (2020). https://doi.org/10.1098/rsos.191504
10. G. Chowell, Infectious Disease Modelling **2**(3), 379 (2017). http://dx.doi.org/10.1016/j.idm.2017.08.001
11. C. Pongkitivanichkul, D. Samart, T. Tangphati, P. Koomhin, P. Pimton, P. Dam-O, A. Payaka, P. Channuie, Physica Scripta **95**(8) (2020). https://doi.org/10.1088/1402-4896/ab9bdf
12. F. Clark, B.W. Brook, S. Delean, H. Reşit Akçakaya, C.J. Bradshaw, Methods in Ecology and Evolution **1**(3), 253 (2010)
13. K. Wu, D. Darcet, Q. Wang, D. Sornette, Nonlinear Dynamics **101**(3), 1561 (2020)
14. E. Pelinovsky, A. Kurkin, O. Kurkina, M. Kokoulina, A. Epifanova, Chaos. Solitons & Fractals **140**, 110241 (2020)
15. F.V. Berghen. Levenberg-Marquardt algorithms vs Trust Region algorithms (2004). http://www.applied-mathematics.net/LMvsTR/LMvsTR.pdf
16. W. Dhouib, J. Maatoug, I. Ayouni, N. Zammit, R. Ghammem, S.B. Fredj, H. Ghannem, Syst Rev **10**(1), 1 (2021)
17. J.M. Ver Hoef, The American Statistician **66**(2), 124 (2012)
18. V. Chin, N.I. Samia, R. Marchant, O. Rosen, J.P. Ioannidis, M.A. Tanner, S. Cripps, Euro J Epidemiology **35**(8), 733 (2020)
19. Associação Nacional dos Registradores de Pessoas Naturais. Especial COVID-19. https://transparencia.registrocivil.org.br/especial-covid
20. S.Y. Lee, B. Lei, B. Mallick, PloS one **15**(7), e0236860 (2020)
21. A.S. Martinez, R.S. González, C.A.S. Terçariol, Physica A: Stat Mech Appl **387**(23), 5679 (2008)
22. C. Tsallis, U. Tirnakli, Front Phys **8**, 217 (2020)
23. L. Hébert-Dufresne, B.M. Althouse, S.V. Scarpino, A. Allard, J Royal Soc Int **17**(172), 20200393 (2020)