Infectious Disease Modelling 2 (2017) 268–275

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

A primer on stable parameter estimation and forecasting in epidemiology by a problem-oriented regularized least squares algorithm

Alexandra Smirnova [a, *], Gerardo Chowell [b, c]

[a] Department of Mathematics and Statistics, Georgia State University, Atlanta, USA
[b] School of Public Health, Georgia State University, Atlanta, USA
[c] Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

ARTICLE INFO

Article history:
Received 17 March 2017
Received in revised form 15 May 2017
Accepted 20 May 2017
Available online 25 May 2017

Keywords:
Generalized Richards model
Parameter estimation
Regularization methods
Epidemic forecasting
Ebola

ABSTRACT

Public health officials are increasingly recognizing the need to develop disease-forecasting systems to respond to epidemic and pandemic outbreaks. For instance, simple epidemic models relying on a small number of parameters can play an important role in characterizing epidemic growth and generating short-term epidemic forecasts. In the absence of reliable information about transmission mechanisms of emerging infectious diseases, phenomenological models are useful to characterize epidemic growth patterns without the need to explicitly model transmission mechanisms and the natural history of the disease. In this article, our goal is to discuss and illustrate the role of regularization methods for estimating parameters and generating disease forecasts using the generalized Richards model in the context of the 2014–15 Ebola epidemic in West Africa.

1. Introduction

Developing tools for stable parameter estimation and reliable forecasting of emerging and re-emerging infectious disease epidemics represents a key priority for public health officials and government agencies in their work to prevent and mitigate disease threats. At the early stage of an outbreak, when incidence data are limited and subject to reporting delays, it is often premature to characterize the transition rates of individuals between various disease epidemiological compartments using, for instance, SEIR-type systems of differential equations, which may involve a substantial number of unknown parameters. For the early transmission period, phenomenological models of a logistic type, describing the progression of the epidemic in terms of the cumulative number of reported cases, C, provide a simple alternative. In what follows, we employ the generalized Richards model (Chowell et al., 2016; Smirnova et al., 2017; Turner, Bradley, Kirk, & Pruitt, 1976)

* Corresponding author.
E-mail addresses: asmirnova@gsu.edu (A. Smirnova), gchowell@gsu.edu (G. Chowell).
Peer review under responsibility of KeAi Communications Co., Ltd.

http://dx.doi.org/10.1016/j.idm.2017.05.004

$$\frac{dC}{dt} = rC^p\left[1-\left(\frac{C}{K}\right)^a\right], \quad \frac{C}{K}(\tau) = \left(\frac{p}{a+p}\right)^{\frac{1}{a}}, \quad \tau\in(0,T), \tag{1.1}$$

to estimate crucial disease parameters, such as the intrinsic growth rate (r), the deceleration of growth (p), the final size of the epidemic (K), the disease turning point (τ), and the extent of deviation from the S-shaped dynamics of the classical logistic-growth curve (a), see Fig. 1. In (1.1), T is the duration of the outbreak.

For the original Richards model ($p = 1$), the closed-form solution is available

$$C(t) = \frac{K}{\left[1+ae^{-ar(t-\tau)}\right]^{1/a}}, \tag{1.2}$$

which simplifies its theoretical and numerical study. However, the statistical analysis of the fitting accuracy for the original and generalized Richards models indicates that the generalized Richards model outperforms its original version for epidemics characterized by an early sub-exponential growth phase (Chowell et al., 2016). For such outbreaks, the value of parameter p in (1.1) is less than 1 ($0<p<1$), and the reduction in the residual sum of squares appears to be statistically significant. Several mechanisms could give rise to initial sub-exponential growth in case incidence including (Chowell, Viboud, Hyman, & Simonsen, 2015; Viboud, Simonsen, & Chowell, 2015): (i) spatially clustered contact structures (e.g., high clustering levels) (Chowell et al., 2015; Szendroi & Csanyi, 2004); (ii) early onset of population behavioral changes and control interventions (Chowell et al., 2015; Szendroi & Csanyi, 2004); and (iii) substantial heterogeneity in susceptibility and infectivity of the host population that introduces high local variability in the local reproduction number that fluctuates around the epidemic threshold at 1.0. The two limiting cases of the generalized Richards model are those in which $p = 0$ (constant incidence growth) or $p = 1$ (exponential growth) (Viboud et al., 2015).

Fig. 2 illustrates representative epidemic profiles, $\frac{dC}{dt}$, that the generalized Richards model supports, as the deceleration of growth parameter, p, is varied. Overall, as parameter p increases, the epidemic profile shows faster growth that converges to exponential rate as $p\to 1$ while keeping the epidemic size, K, fixed. On the other hand, parameter a modulates the epidemic turning point (Fig. 2), which occurs earlier as this parameter increases.

Thus, in our investigation we utilize the generalized four-parametric model (1.1), and solve the ODE-constrained least squares problem to estimate r, p, K, τ, and a from early data of the 2014–15 Ebola epidemics in Guinea, Sierra Leone, and Liberia. We then use the reconstructed parameter values to forecast future incidence cases by propagating uncertainty in the system forward in time.

2. The regularized least squares problem

To present a regularized numerical algorithm for stable parameter estimation, we introduce

$$b := \frac{r}{K^{1-p}}, \quad H(t) := \frac{C(t)}{K}, \tag{2.1}$$

and arrive at the normalized model (Cavallini, 1993)

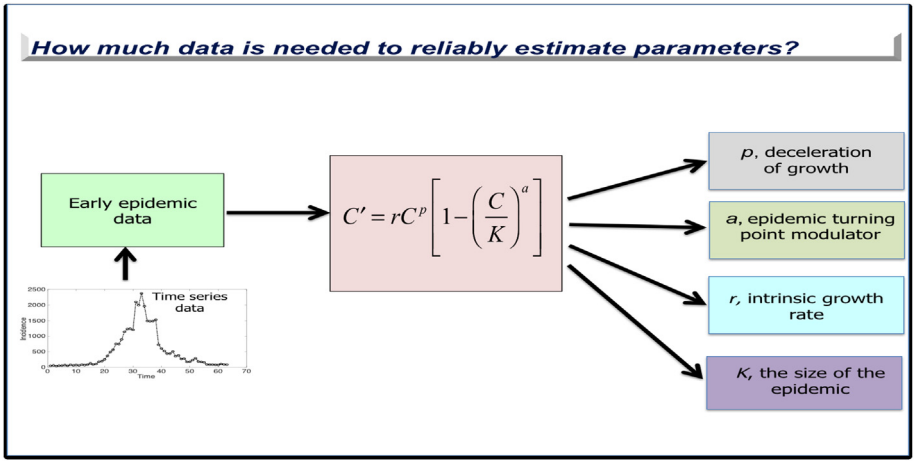

Fig. 1. Parameter identification process from early incidence data using the generalized Richards model.