

RESEARCH ARTICLE

tsiR: An R package for time-series Susceptible-Infected-Recovered models of epidemics

Alexander D. Becker^{1*}, Bryan T. Grenfell^{1,2,3}

1 Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, United States of America, **2** Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ, United States of America, **3** Fogarty International Center, National Institutes of Health, Bethesda, MD, United States of America

* adbecker@princeton.edu



Abstract

tsiR is an open source software package implemented in the R programming language designed to analyze infectious disease time-series data. The software extends a well-studied and widely-applied algorithm, the time-series Susceptible-Infected-Recovered (TSIR) model, to infer parameters from incidence data, such as contact seasonality, and to forward simulate the underlying mechanistic model. The **tsiR** package aggregates a number of different fitting features previously described in the literature in a user-friendly way, providing support for their broader adoption in infectious disease research. Also included in **tsiR** are a number of diagnostic tools to assess the fit of the TSIR model. This package should be useful for researchers analyzing incidence data for fully-immunizing infectious diseases.

OPEN ACCESS

Citation: Becker AD, Grenfell BT (2017) tsiR: An R package for time-series Susceptible-Infected-Recovered models of epidemics. PLoS ONE 12(9): e0185528. <https://doi.org/10.1371/journal.pone.0185528>

Editor: Hiroshi Nishiura, Hokkaido University Graduate School of Medicine, JAPAN

Received: May 26, 2017

Accepted: September 14, 2017

Published: September 28, 2017

Copyright: © 2017 Becker, Grenfell. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: A.D.B. was supported by a National Science Foundation Graduate Research Fellowship and the Center for Health and Wellbeing at Princeton University. B.T.G. was supported by the Research and Policy for Infectious Disease Dynamics program of the Science and Technology Directorate, Department of Homeland Security, the Fogarty International Center, National Institutes of Health, the Bill and Melinda Gates Foundation, and

Introduction

Mathematical models coupled with statistical inference techniques allow us to compare infectious disease theory and data, shedding light on transmission estimates, vaccine control strategies, and predicting future trends [1, 2]. These models (and inference methods) cover a spectrum from very simple (based on well-mixed, population-level assumptions) to highly complex representations (in which individual variation is modeled explicitly) [1, 3]. Even the simplest of such non-linear models can display very rich, elaborate, and potentially chaotic dynamics [4–6].

One of the simplest and most powerful of epidemic models is the family of mass-action formulations based on the Susceptible-Infected-Recovered (SIR) equations [1, 7]. The SIR model assumes a well-mixed population, and in the most basic form balances demographic processes (e.g., births, B , and deaths, μ) with infectious disease specific properties such as contact rate, β , and infectious period, inverse γ , for a single pathogen. In general, and especially for the prototypical example, measles, the transmission coefficient, β , varies seasonally. The deterministic skeleton of the SIR model is shown in Eq 1; λ is the force of infection, typically defined as $\frac{\beta I}{N}$. In addition, this model requires that immunity post-infection is life-long, although this

the U.S. Centers for Disease Control and Prevention.

Competing interests: The authors have declared that no competing interests exist.

assumption can be relaxed via a Susceptible-Infected-Recovered-Susceptible (SIRS) model.

$$\begin{aligned} \frac{dS}{dt} &= B - \frac{\lambda S}{N} - \mu S \\ \frac{dI}{dt} &= \frac{\lambda S}{N} - \gamma I - \mu I \\ \frac{dR}{dt} &= \gamma I - \mu R \end{aligned} \tag{1}$$

Despite its simplicity, calibrating the seasonally-forced SIR model against time-series data is a difficult mathematical and statistical problem, as evidenced by the extensive literature on this subject [8–13]. The desired outcome of SIR model inference is to extract parameter estimates from a given epidemic time-series for key values such seasonal variation. In terms of the data, fitting SIR-type models is a non-trivial inferential challenge for a number of reasons. First, only one state variable—the number of cases over time—is observed. Second, there is generally substantial under-reporting of disease incidence. Adding additional complexity is the fact that the model is a continuous-time process, whereas the data are generally collected on a weekly or monthly basis. While statistically robust and powerful methods (e.g., filtering based algorithms) are readily available in packages such as **pomp** [14], they may be slow to converge and may not be suitable for analyzing many (100+) time-series.

$$\begin{aligned} S_{t+1} &= B_{t+1} - S_t - I_{t+1} \\ \mathbb{E}[I_{t+1}] &= \beta_{t+1} S_t I_t^\alpha \end{aligned} \tag{2}$$

An alternative, computationally inexpensive, and highly tractable approach to these problems is provided by the time-series SIR model (TSIR model), shown in Eq 2 [8]. The TSIR model relies on two main assumptions: first, that the infectious period is fixed at the sampling interval of the data (e.g., bi-weekly for measles) and that over a long enough time (e.g., 10-20 years), the sum of births and cases should be approximately equal due to the high infectivity of pathogens such as measles and other childhood infections, in the pre-vaccine era. A full description of the TSIR model and algorithm can be found in [8]. A brief qualitative description of the algorithm follows in order to provide context for the subsequent development of the **tsiR** package.

In the TSIR framework, a regression model is first fitted between cumulative cases and cumulative births. Assuming they are equal, the slope will be the reporting rate ρ_t and the residuals of the regression model, Z_t provide the shape of the susceptible dynamics, S_t . Next, using Eq 2 and setting the expectation to the mean, the log-linear equation shown in Eq 3 can be acquired. The mean number of susceptible individuals across the time-series, \bar{S} , can be inferred using profile likelihood and a seasonally repeating contact rate (52 divided by the infectious period time points, e.g., 26 for measles); β_t , can then be estimated along with a homogeneity parameter α using a generalized linear model (GLM). The parameter α describes the epidemic saturation as well as a correction factor for switching from continuous to discrete time [15, 16]

$$\log I_{t+1} = \log \beta_{t+1} + \log (Z_t + \bar{S}) + \alpha \log I_t \tag{3}$$

The TSIR method has been used very successfully to analyze a number of childhood infections such as measles, whooping cough, diphtheria, mumps, varicella, and scarlet fever, as well as multi-strain pathogens such as dengue. [6, 8, 12, 17–20]. Analysis of this model has shown both birth rate and school-term forcing to be central drivers of the pattern of epidemics and periodicity, as well as improving our ability to predict infectious disease dynamics further into the future in both small and large populations ranging from London to Iceland [8, 12].

Although the underlying model is simple, the TSIR approach provides a number of different fitting options. When reconstructing the susceptible dynamics, options range from a simple linear regression fit between cumulative cases and cumulative births, to more sophisticated approaches such as Gaussian regressions. These choices, and which variable is the dependent variable in each case, also impact both whether the reporting rate ρ_t is time-invariant as well as how the residuals of the regression are calculated. Once the susceptible dynamics are reconstructed, the modeler is again faced with many more choices about the log-linear model, the GLM family and link, and which parameters to estimate and which to fix (commonly, for the study of measles dynamics, α is fixed to be 0.97, and \bar{S} is occasionally fixed as well to be 0.035 based on analysis of pre-vaccination data from the United Kingdom [6, 12, 18]). This decision process can also be implemented in a Bayesian framework, although is a computationally more extensive task. Further choices must be made when describing the distribution of the expected value in Eq 2, as well as a deciding between a completely forward prediction versus a step-ahead prediction. These options, while relatively straightforward, are cumbersome to implement especially while working with a number of time-series. Thus, while the TSIR model has been used extensively, there is a need for an open source software package which implements these options in a user-friendly way. We have developed the **tsiR** package to address this methodological challenge and facilitate a more straight-forward and widely-accessible model-fitting process.

1 Methods

The **tsiR** package, available on CRAN, is written completely in the **R** programming language. Package functions and a short description of their use is included at the end of this section. Package dependencies are **kernlab** [21], **ggplot2** [22], and **reshape2** [23]. All code is publicly available on GitHub (www.github.com/adbecker/tsiR). In the following section, typewriter font refers to function arguments and quotes refer to argument inputs.

1.1 Data

tsiR requires data to be a named data frame with a ‘time’ column (that can be understood by **ggplot2**—see [22] for details), and columns for ‘cases’, ‘births’ and ‘pop’ (population size). If the number of births and population size are on a different time scale than the reported cases, these data (as well as case data) must be interpolated to the generation time. When incidence data are reported weekly and demographic data (i.e., births and population size) are reported yearly, as is often the case, these individual vectors can be interpolated to the proper infectious period timescale via the **tsiRdata** function. Note that at each time point, the births must be the number of births that occur within the IP weeks, where IP is the infectious period (in weeks). For measles, this is typically taken as two weeks, however **tsiR** does not require IP to be an integer. A dataset, **twentymeas**, is included in the **tsiR** package as a list. This dataset is a list which contains biweekly data ($IP=2$) for twenty locations in England from 1944-1964 and was analyzed in [11].

1.2 Model inference and forward simulation

The main function in the **tsiR** package is **runtsir**. This function reconstructs the susceptible dynamics, fits the log-linear relationship in Eq 3, and then resimulates the TSIR model in Eq 2 forward under the fitted parameters.

An essential argument (and assumption of the TSIR model) in the **tsiR** package is that the time step is equal to the infectious period (IP), i.e. that IP is a consistent argument throughout the data. Written based on measles, in **tsiR** all functions default to $IP = 2$. Thus, if a

different disease is being analyzed, it is key to change the `IP` argument throughout the function inputs. As with all **R** functions, a full description of the function can be acquired via `?runtsir`. Here we describe the main arguments. The `xreg` argument indicates whether cumulative cases or cumulative births are on the x-axis of the regression. Thus, the options are simply 'cumcases' or 'cumbirths' with a default of 'cumcases'. A more extensive discussion of this choice can be found in [6]. After describing the formula for the susceptible reconstruction, the regression type must be specified via the `regtype` argument. Here, the options are 'gaussian', where the kernel size is profiled to maintain a reporting rate between zero and one, and 'lm', 'lowess', 'loess' or 'spline' regression, for a linear, a lowess, a loess, and a spline regression with 2.5 degrees freedom, respectively. If the Gaussian process fails to produce a reporting rate between zero and one, **runtsir** defaults to a loess regression. A Gaussian regression, as implemented in [12], appears to produce the most robust results in both small and large populations. However, using a Gaussian regression may produce exaggerated reporting rates when there is a single outbreak that is substantially larger in size than across the rest of the time series. The aforementioned arguments will specify the shape, Z_b , of the susceptible dynamics, S_t . At this stage, model parameters left to estimate are β_b , α , and \bar{S} , although the last two can be fixed as previously mentioned. Note that the number of parameters in β_t is dependent on `IP` such that the length of β_t is 52 divided by `IP`. The only exception is where `IP` is equal to one. In this scenario β_t is truncated to 26 points, each repeated twice to preserve statistical integrity. Parameters α and \bar{S} can be fixed via `alpha` and `sbar` if desired. For the log-linear regression in Eq 3 virtually any GLM family and link can be used, although the options are essentially 'quasipoisson', 'poisson', and 'gaussian' where 'poisson'/'quassipoisson' take a 'log' link and a 'gaussian' family takes either a 'log' or 'identity' link.

At this stage, all parameters are estimated, confidence intervals have been constructed (when appropriate and the computation can be completed), and the model can be forward simulated. The **runtsir** function defaults to a full time-series ahead forward prediction, although step-ahead can be inputted as well under the `pred` argument. For large populations, the forward prediction can generally be simulated without fear of fade-outs (i.e., the simulation declining to zero). However, for smaller populations with frequent local 'fade-outs' of infection between epidemics, the forward prediction may only be able to be run a simple epidemic-ahead [12]. This constraint on the prediction model can be specified via `epidemics = break` as well as designating a `threshold` parameter that indicates when a new epidemic is started and thus where to reset the forward simulation. By default, the initial conditions (S_0, I_0) are $(Z_0 + \bar{S}, C_0/\rho_0)$ where C_t is cases. Thus, if C_0 is zero, the forward prediction will fail. In this scenario, the data must be truncated to the first non-trivial case, or initial conditions can be fit using simple least squares per [6], although this feature could be made more robust in the future (see the [Conclusions](#) section). This can be specified using the `inits.fit = T` argument. Finally, the simulation distribution can be specified via `method` where the options are 'deterministic', 'negbin', and 'pois' for deterministic, Negative Binomial, and Poisson distributions. The number of simulations to perform is specified via `nsim`. The output of **runtsir** is a named list and can be fully plotted via the **plotres** function. The output of this model, as generated by the following code, can be seen below and is plotted in Fig 1.

```
LondonMeas <- twentymeas[["London"]]
LondonRes <- runtsir(data=LondonMeas, IP=2,
                    xreg='cumcases', regtype='gaussian',
                    alpha=NULL, sbar=NULL,
                    family='gaussian', link='identity',
                    method='negbin', nsim=100)
plotres(LondonRes)
```

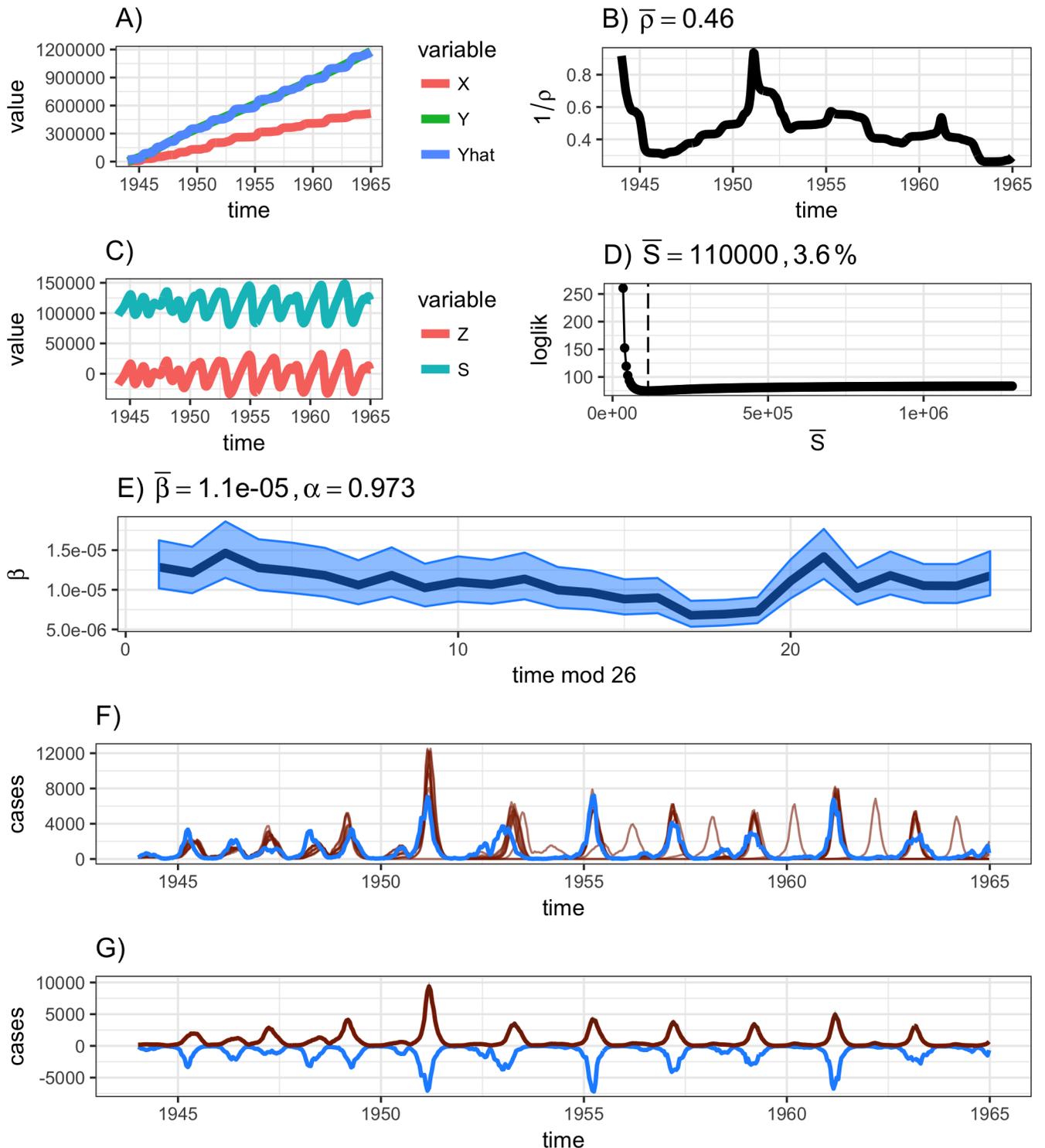


Fig 1. Output results from the runtsir function for London. Subplots A) and B) are the cumulative births against cumulative cases regression and estimated reporting rate, the C) and D) are the profiled S from Z_t and then reconstructed S, E) is 26-point β_t , with the α and mean β (indicated as $\bar{\beta}$) estimate, and F) and G) are the data (blue) against 10 randomly chosen stochastic simulations (red) and the (inverse) data against mean of the simulations with confidence intervals.

<https://doi.org/10.1371/journal.pone.0185528.g001>

The `runtsir` function is also decomposed via the `estpars` and `simulatetsir` functions. This may be more desirable if exploring a large number of simulations or analyzing sparse incidence data. In the following short example, we will examine such a data set from Northwich (Cheshire, England, population size in 1944 = 18,070). Here, we must define epidemic start and end times per [12]. Using a threshold of three, we can see where each epidemic is defined via the dashed line in Fig 2. Here, we are estimating parameters using different regression options available in `tsiR`, as well as simulating the model via a Poisson distribution and fixing α . Only the forward simulations are shown in Fig 3, where again the data is shown in blue and the simulation results are shown in red. Overall, the forward simulation captures the epidemic final size well for such a noisy time-series [12]. The code used to generate these fits and plots is shown below.

```
NorthwichMeas <- twenty meas[["Northwich"]]
NorthwichParms <- estpars (data = NorthwichMeas, IP = 2,
                           alpha = 0.97, sbar = NULL,
                           regtype = 'loess',
                           family = 'poisson', link = 'log')
plotbreaks (data = NorthwichMeas, threshold = 3)
NorthwichRes <- simulatetsir (data = NorthwichMeas, IP = 2,
                              parms = NorthwichParms,
                              epidemics = 'break', threshold = 3,
                              method = 'pois', nsim = 100)
plotcomp (NorthwichRes)
```

A brief summary of the main functions and their usage in the `tsiR` package follow below in Table 1. Please note for MCMC functionality, one must install `rjags` [24] independently. The MCMC functions, `mcmcestpars` and `mcmctsir`, follow generally the same arguments as their frequentist counterparts. Notable exceptions however are that a family and link no longer can be specified, and MCMC specific arguments (`n.chains` for the number of chains, `update.iter` for the number of MCMC iterations to use in the `rjags` update section, `n.iter` for number of MCMC iterations to perform, `n.adapt` for the adaptive number, and `burn.in` for the burn in number) can be specified. For more information on these arguments, we direct the reader to `rjags` documentation [24]. Additionally, annotated code for the London and Northwich estimations and forward simulations is included as a .R files in the Supporting Information.

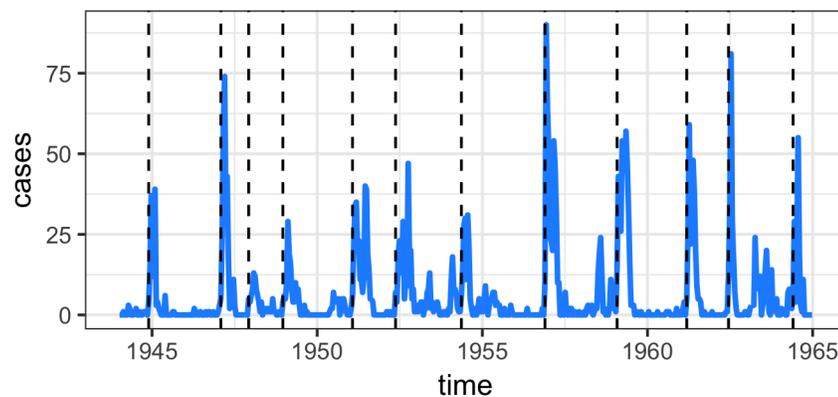


Fig 2. Northwich time-series data. Dashed lines are the time points in which the forward simulation resets in the `epidemics = 'break'` argument for a threshold for three.

<https://doi.org/10.1371/journal.pone.0185528.g002>

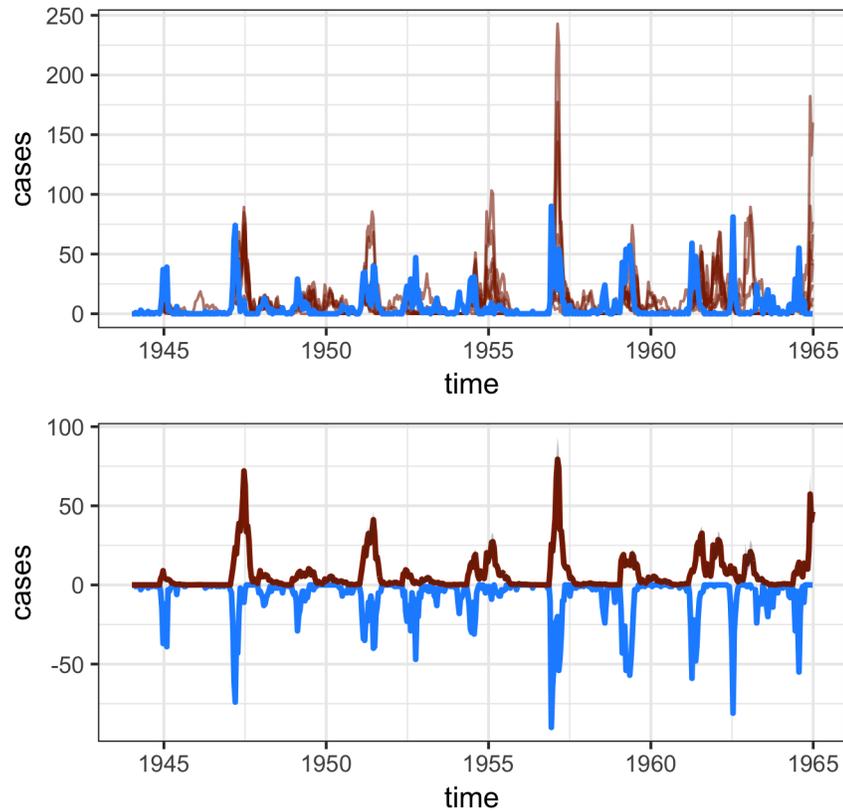


Fig 3. The forward simulations for the Northwich time-series data under an epidemic-ahead fit using a threshold of three. The color coding in the panels shown here are the same as in Fig 1.

<https://doi.org/10.1371/journal.pone.0185528.g003>

2 Conclusions, limitations, and future work

The **tsiR** package allows researchers to fit the time-series Susceptible-Infected-Recovered model using a number of different fitting options that are easy to change and compare. Per the model formulation, the frequentist fitting techniques are computationally tractable and, to a first approximation, work very well for childhood diseases in a number of settings.

However, the model does make a number of assumptions that are undesirable for certain data and pathogens. For example, fixing the infectious period may not be realistic for more chronic infections, and ignoring deaths may lead to biased conclusions when examined over a long enough time scale. Additionally, under certain settings and time scales, the assumption that cumulative cases approximates cumulative births may be flawed. Furthermore, the TSIR model only includes the observation process as a single reporting rate and not as a probability distribution, and stochasticity cannot be explicitly estimated. To include these complexities, methods such as Sequential Monte Carlo and Iterated Filtering can be used to perturb parameters in order to maximize the likelihood [10, 25]. These algorithms are included in the **R** package **pomp** [14]. Such flexibility does come at a cost, however, as Maximum Likelihood methods can be computationally expensive and optimization algorithms are often complex [11, 14, 26].

Regardless, for fully-immunizing childhood infections such as measles, the TSIR model is able to accurately capture the parameters of interest across a range of different scenarios and remains the most tractable approach (in particular for large numbers of time-series). Improvements and areas of future work to the **tsiR** package include incorporating spatial disease

Table 1. Summary and description of the main functions in the tsiR package.

Function	Description
tsiRdata	Interpolates (weekly) cases and (yearly) births and population vectors to the generation time of the disease
twenty meas	Named list of twenty biweekly (IP = 2) time-series from 1944-1964 England
plotdata	Plots the cases, births, and population time series
runttsir	Reconstructs susceptible dynamics, estimates parameters, and runs the forward simulation
plotres	Plots the fitted regressions, reporting, susceptible reconstruction, estimated parameters, and the fit diagnostics
plotcomp	Plots the data versus the simulations only
estpars	Reconstructs the susceptible dynamics and estimates parameters
maxthreshold	Optimizes the threshold parameter for sparse data using R^2
mcmcstpars	Using MCMC, reconstructs susceptible dynamics and estimates parameters
mcmcsttsir	Using MCMC, reconstructs susceptible dynamics, estimates parameters, and runs the simulation
simulatetsir	Runs the forward simulation taking in the output from estpars or mcmcstpars
plotregression	Plots the cumulative cases and cumulative births regression
plotrho	Plots the inferred reporting rate ρ
plotsbar	Plots the inferred \bar{S} and the profile log likelihood
plotbeta	Plots the inferred β
plotforward	Plots the data against the mean simulation with confidence intervals

<https://doi.org/10.1371/journal.pone.0185528.t001>

spread (e.g., the gravity model per [18, 27]), the addition of an immigration parameter to the force of infection, a more statistically robust method to estimate initial conditions, and a more streamlined Bayesian approach to the TSIR model. In the spirit of open science, other researchers are welcome to send suggestions, bug reports, as well as contributions to the software.

Supporting information

S1 File. R script. Annotated code to run the analysis in this paper (with additional plots). (R)

Acknowledgments

We thank Ottar Bjørnstad, Quentin Caudron, Benjamin Dalziel, Ayesha S. Mahmud, C. Jessica Metcalf, Sinead E. Morris, Saki Takahashi, Amy Wesolowski, and Colin Worby for their valuable suggestions, comments, and work using the TSIR model.

Author Contributions

Conceptualization: Alexander D. Becker, Bryan T. Grenfell.

Data curation: Bryan T. Grenfell.

Formal analysis: Alexander D. Becker.

Investigation: Alexander D. Becker.

Methodology: Alexander D. Becker.

Software: Alexander D. Becker, Bryan T. Grenfell.

Visualization: Alexander D. Becker, Bryan T. Grenfell.

Writing – original draft: Alexander D. Becker, Bryan T. Grenfell.

Writing – review & editing: Alexander D. Becker, Bryan T. Grenfell.

References

1. Anderson RM, May R M. Infectious diseases of humans: Dynamics and control. vol. 15. Oxford: Oxford University Press; 1992. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/hep.1840150131/abstract>.
2. Freckleton RP. The Ecological Detective: Confronting Models with Data. Hilborn R. & Mangel M. (1997). *Journal of Applied Ecology*. 1999; 36(5):842–846. <https://doi.org/10.1046/j.1365-2664.1999.04462.x>
3. Holland PW, Leinhardt S. An Exponential Family of Probability Distributions for Directed Graphs. *Journal of the American Statistical Association*. 1981; 76(373):33–50. <https://doi.org/10.2307/2287042>
4. May RM. Simple mathematical models with very complicated dynamics. *Nature*. 1976; 261(5560):459–467. <https://doi.org/10.1038/261459a0> PMID: 934280
5. Ferrari MJ, Grais RF, Bharti N, Conlan AJK, Bjørnstad ON, Wolfson LJ, et al. The dynamics of measles in sub-Saharan Africa. *Nature*. 2008; 451(7179):679–684. <https://doi.org/10.1038/nature06509> PMID: 18256664
6. Dalziel BD, Bjørnstad ON, van Panhuis WG, Burke DS, Metcalf CJE, Grenfell BT. Persistent Chaos of Measles Epidemics in the Pre-vaccination United States Caused by a Small Change in Seasonal Transmission Patterns. *PLOS Computational Biology*. 2016; 12(2):e1004655. <https://doi.org/10.1371/journal.pcbi.1004655> PMID: 26845437
7. Kermack WO, McKendrick AG. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 1927; 115(772):700–721. <https://doi.org/10.1098/rspa.1927.0118>
8. Finkenstädt BF, Grenfell BT. Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 2000; 49(2):187–205. <https://doi.org/10.1111/1467-9876.00187>
9. Wallinga J, Lévy-Bruhl D, Gay NJ, Wachmann CH. Estimation of measles reproduction ratios and prospects for elimination of measles by vaccination in some Western European countries. *Epidemiology and Infection*. 2001; 127(2):281–295. <https://doi.org/10.1017/S095026880100601X> PMID: 11693506
10. Ionides EL, Bretó C, King AA. Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(49):18438–18443. <https://doi.org/10.1073/pnas.0603181103> PMID: 17121996
11. He D, Ionides EL, King AA. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of The Royal Society Interface*. 2010; 7(43):271–283. <https://doi.org/10.1098/rsif.2009.0151>
12. Caudron Q, Mahmud AS, Metcalf CJE, Gottfresson M, Viboud C, Cliff AD, et al. Predictability in a highly stochastic system: final size of measles epidemics in small populations. *Journal of The Royal Society Interface*. 2015; 12(102):20141125. <https://doi.org/10.1098/rsif.2014.1125>
13. Takahashi S, Liao Q, Van Boeckel TP, Xing W, Sun J, Hsiao VY, et al. Hand, Foot, and Mouth Disease in China: Modeling Epidemic Dynamics of Enterovirus Serotypes and Implications for Vaccination. *PLOS Medicine*. 2016; 13(2):e1001958. <https://doi.org/10.1371/journal.pmed.1001958> PMID: 26882540
14. King AA, Ionides EL, Breto C, Ellner SP, Ferrari MJ, Kendall BE, et al. pomp: Statistical Inference for Partially Observed Markov Processes; 2015. Available from: <https://cran.r-project.org/web/packages/pomp/index.html>.
15. Glass K, Xia Y, Grenfell BT. Interpreting time-series analyses for continuous-time biological models—measles as a case study. *Journal of Theoretical Biology*. 2003; 223(1):19–25. [https://doi.org/10.1016/S0022-5193\(03\)00031-6](https://doi.org/10.1016/S0022-5193(03)00031-6) PMID: 12782113
16. He D, Dushoff J, Day T, Ma J, Earn DJD. Mechanistic modelling of the three waves of the 1918 influenza pandemic. *Theoretical Ecology*. 2011; 4(2):283–288. <https://doi.org/10.1007/s12080-011-0123-3>
17. Grenfell BT, Bjørnstad ON, Finkenstädt BF. Dynamics of measles epidemics: scaling noise, determinism, and predictability with the tsiR model. *Ecological Monographs*. 2002; 72(2):185–202. [https://doi.org/10.1890/0012-9615\(2002\)072%5B0185:DOMESN%5D2.0.CO;2](https://doi.org/10.1890/0012-9615(2002)072%5B0185:DOMESN%5D2.0.CO;2)

18. Bjørnstad ON, Finkenstädt BF. Dynamics of measles epidemics: Estimating scaling of transmission rates using a Time series SIR model. *Ecological Monographs*. 2002; 72(2):169–184. [https://doi.org/10.1890/0012-9615\(2002\)072%5B0169:DOMEES%5D2.0.CO;2](https://doi.org/10.1890/0012-9615(2002)072%5B0169:DOMEES%5D2.0.CO;2)
19. Metcalf CJE, Bjørnstad ON, Grenfell BT, Andreasen V. Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen. *Proceedings of the Royal Society of London B: Biological Sciences*. 2009; 276 (1676). <https://doi.org/10.1098/rspb.2009.1058>
20. Reich NG, Shrestha S, King AA, Rohani P, Lessler J, Kalayanarooj S, et al. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. *Journal of The Royal Society Interface*. 2013; 10(86). <https://doi.org/10.1098/rsif.2013.0414>
21. Karatzoglou A, Smola A, Hornik K, Zeileis A. Kernel-Based Machine Learning Lab [R package kernlab version 0.9-25]. *Journal of Statistical Software*. 2004; 11(9):1–20.
22. Wickham Hadley. *ggplot2: Elegant Graphics for Data Analysis*; 2009.
23. Wickham H. Reshaping Data with the reshape Package; 2007. Available from: <https://cran.r-project.org/web/packages/reshape2/>
24. Plummer M. Bayesian Graphical Models using MCMC [R package rjags version 4-6]. 2016;.
25. Bretó C, He D, Ionides EL, King AA. Time series analysis via mechanistic models. *The Annals of Applied Statistics*. 2009; 3(1):319–348. <https://doi.org/10.1214/08-AOAS201>
26. Groendyke C, Welch D, Hunter DR. A network-based analysis of the 1861 Hagelloch measles data. *Biometrics*. 2012; 68(3):755–765. <https://doi.org/10.1111/j.1541-0420.2012.01748.x> PMID: 22364540
27. Xia Y, Bjørnstad O, Grenfell B. Measles Metapopulation Dynamics: A Gravity Model for Epidemiological Coupling and Dynamics. *The American Naturalist*. 2004; 164(2):267–281. <https://doi.org/10.1086/422341> PMID: 15278849