Title
Modeling the Transmission Dynamics of Pertussis Using Recursive Point Process and SEIR model

Author
Yang, Ah Sung

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Modeling the Transmission Dynamics of Pertussis Using Recursive Point Process and SEIR model

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Applied Statistics

by

Ah Sung Yang

2019

ABSTRACT OF THE THESIS

Modeling the Transmission Dynamics of Pertussis
Using Recursive Point Process and SEIR model

by

Ah Sung Yang
Master of Applied Statistics
University of California, Los Angeles, 2019
Professor Frederic R. Paik Schoenberg, Chair

Here we used the recursive point process and SEIR model to describe the transmission dynamics of pertussis. The recursive point process is a refined version of the Hawkes point process which defines productivity as a function of conditional intensity. The SEIR model is a variant of the basic compartmental model, SIR, which additionally takes into account a period when an individual is infected but not infectious yet. We attempted to fit the models to the reported cases of pertussis in Nevada since 1940. The evaluation of the model fit and prediction errors suggests that the recursive point process performs better in modeling the spread of pertussis and predicting the weekly incidence. However, we still recommend using it along with the SEIR model considering that the compartmental model has been more commonly used for an epidemic study.

The thesis of Ah Sung Yang is approved.

Nicolas Christou

Ying Nian Wu

Frederic R. Paik Schoenberg, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

1 Introduction **1**

2 Recursive Process Model **4**

3 Compartmental Model SEIR **7**

4 Description of Data and Pertussis **10**

5 Estimation **12**

6 Prediction **19**

7 Conclusion and Remarks **23**

References **36**

LIST OF FIGURES

1 Probability of points being attributable to μ according to the fitted recursive point process. 27

2 Histogram of points on the left and a plot of λ estimated by the recursive point process on the right. 27

3 Density of points and fitted intensity estimated by the recursive point process. . 28

4 Histogram of the standardized inter-event times u_i of super-thinned residuals produced by the recursive point process. 28

5 Lag plot of the standardized inter-event times u_i of super-thinned residuals produced by the recursive point process. 29

6 Super-thinned residuals and their corresponding standardized inter-event times u_i of the recursive point process. 29

7 Histogram of the standardized inter-event times u_i of super-thinned residuals produced by the SEIR model. 30

8 Lag plot of the standardized inter-event times u_i of super-thinned residuals produced by the SEIR model. 30

9 Super-thinned residuals and their corresponding standardized inter-event times u_i of the SEIR model. 31

10 Histogram of points on the left and a plot of λ estimated by the SEIR model on the right. 31

11 Density of points and fitted intensity estimated by the SEIR model. 32

12 Cumulative weekly incidence of pertussis. Predicted values are the sum of daily intensities over the forecast week estimated by the fitted recursive point process. 33

13 Cumulative weekly incidence of pertussis. Predicted values are the total number of simulated points over the forecast period. The simulation was conducted by running the thinning algorithm on the fitted recursive point process. 34

14 Cumulative weekly incidence of pertussis. Predicted values are the total number of simulated points over the forecast period. The simulation was conducted by running the tau-leaping algorithm on the SEIR model. 35

LIST OF TABLES

6.1 Sum of squared errors for different models, on the training and test data sets, respectively 22

6.2 Mean of squared errors for different models, on the training and test data sets, respectively 22

CHAPTER 1

Introduction

The accurate modeling and understanding of the spread of infectious diseases is crucial for prevention and control of epidemics. Due to its significant contribution to public health, many statistical models have been applied to a variety of infectious diseases to explain their transmission dynamics and further attempt to predict future outbreaks.

One of the frequently used models for an epidemic study is the compartmental model explored by Kermack and McKendrick (1927), which became the foundation of the SIR (Susceptible-Infected-Recovered) model and its variants (Chaffee et al., 2018). SIR is the simplest form of the compartmental model, which consists of a system of three ordinary differential equations characterizing a change in the number of susceptible (S), infected (I), and recovered (R) individuals in a given population (Goufo et al., 2014). More complicated models have been developed since the appearance of the SIR. For example, the SEIR model takes into account individuals that are exposed to the disease but not yet infectious, considering an incubation period of the disease. The SIRS model assumes the recovered individuals return to the susceptible state as their immunity wanes. Ganna Rozhnova and Ana Nunes used the several variations of the SIR model to model the dynamics of pre-vaccination pertussis (2012). Similarly, Patricia T Campbell et al. also used the standard SIRS and its modified version called boosting SIRWS, which takes into account an enhanced immune response to re-exposure to the disease, to describe the spread of pertussis (Campbell et al., 2015). In addition, Ottar N. Bjørnstad et al. and Amira Rachah et al. applied the SEIR model to measles and Ebola virus data, respectively (Bjørnstad et al., 2002; Rachah & Torres, 2017).

A spatial-temporal point process is also considered as an alternative approach to model the spread of diseases despite its more common use in seismology. One of the point processes

that have been commonly employed to model clustered phenomena is the Hawkes point process (Hawkes, 1971). One of the key characteristics of the Hawkes process is that it self-excites in the sense that each arrival of an event increases the rate of future arrivals for some period of time (Laub et al., 2015). The Hawkes process and its modified versions were used to model the spread of smallpox in Brazil by Becker (1977), to describe the effect of vaccinations on measles by Farrington et al. (2003), and to explain the incidence of invasive meningococcal disease (IMD) in humans by Balderama et al. (2012). The Hawkes process is also considered as the de facto modeling choice for a social media process, where information is treated as a viral contagion that diffuses in the similar fashion to infectious diseases (Rizoiu et al., 2018). Marian-Andrei Rizoiu et al. created a novel approach by establishing a connection between the Hawkes process and SEIR model to explain the pattern of information diffusion (2018). Ryota Kobayashi and Renaud Lambiotte attempted to predict the Retweet activity by applying Time-Dependent Hawkes process which takes into account the circadian nature of users and aging of information (2016). Similarly, P. K. Srijith et al. developed several Hawkes process models with different properties related to general features of users and their connections and applied them to the Twitter data (2017). The Hawkes model was also employed to model dynamics of YouTube video views and popularity of online content (Rizoiu et al., 2017).

One of the properties of the Hawkes process is that its productivity is static, which assumes the rate of transmission by one infectious individual to other remains constant. However, the fixed productivity imposes limitation on modeling the spread of infectious diseases since the transmission rate of a disease is likely to vary depending on the prevailing conditions (Schoenberg et al., 2017). For example, in early stage of an epidemic, when prevalence of the disease is low, the rate of transmission is expected to be high. However, at a later stage when the disease has become prevalent, the transmission rate is likely to be low due to human intervention to control the epidemic, a lower number of potential hosts, and immunity gained by the population (Schoenberg et al., 2017). In order to resolve this issue, Frederic P. Schoenberg et al. developed the recursive point process, an extension of the Hawkes process that defines productivity as a function of conditional intensity (2017).

Here we attempted to fit the recursive point process to the reported cases of pertussis in Nevada and compared its model fit and performance to the SEIR model, one of the classic compartmental models. The structure of the paper is as follows. We will begin the paper explaining the Hawkes and recursive point processes in chapter 2, followed by a review of compartmental models in chapter 3. After the introduction of the models, a description of pertussis and data will be given in chapter 4. Finally, this will be followed by estimation of the model parameters in chapter 5 and predictions in chapter 6 as well as concluding remarks in chapter 7.

CHAPTER 2

Recursive Process Model

The recursive point process is a refined version of the Hawkes process that allows productivity to vary depending on conditional intensity. Having productivity as a function of intensity enables the model to have flexibility to describe transmission dynamics better.

The Hawkes process is a counting or point process that models a sequence of arrivals of a certain event over time such as earthquakes, gang violence, trade orders, or bank defaults (Laub et al., 2015). Its defining characteristic is that it is self-exciting, which means each arrival of the event increases the probability of a subsequent arrival (Laub et al., 2015).

A point process is typically characterized by its conditional intensity function as opposed to conditional arrival distribution (Laub et al., 2015). The conditional intensity represents the infinitesimal rate at which points are accumulating at a specific location of space-time, given all information up to the time t (Schoenberg et al., 2017). For a simple temporal Hawkes process, its conditional intensity function is defined as follows:

$$\lambda(t) = \mu + K \int_0^t g(t - t')dN(t') = \mu + K \sum_{i:t_i<t} g(t - t_i)$$

where $\mu > 0$ is the background intensity, $g(u)$ is the excitation or triggering function relating the effect of previous events to the current intensity, and K is the productivity. The background rate describes the arrival of events triggered by external sources independently from the previous events (Rizoiu et al., 2017). The triggering function defines how much each occurrence at time t_i contributes to the secondary occurrence. The self-exciting feature of the process arises from the summation or integral of the triggering function (Rizoiu et al., 2017).

A common choice for the triggering function is a monotonically decreasing one such as the exponential decay function $\beta \exp(-\beta(t-t'))$ or the power law function $g(t-s) = \frac{1}{(c+(t-s))^p}$ (Laub et al., 2015). These functions allow more recent events to have higher influence on the current intensity than the ones having occurred further away in time (Rizoiu et al., 2017). The remaining discussion in this paper will use the exponential decay function as the triggering function of Hawkes and recursive processes. The parameter in the exponential decay function can be interpreted as follows: each new arrival of the event will instantaneously increase the conditional intensity by βK; however, its impact decays at a rate of β over time (Laub et al., 2015).

The recursive model was originated from the idea that the static productivity assumed in the Hawkes process is not optimal for describing transmission dynamics of epidemics (Schoenberg et al., 2017). When prevalence of a disease is very low, the current intensity is also expected to be low. However, its productivity, the transmission rate by one infectious individual to other, is likely to be high. This can be attributed to lack of immunity carried by potential hosts in the population, thereby increasing likelihood of transmission of the disease (Schoenberg et al., 2017). On the other hand, when the outbreak reaches its peak, the conditional intensity is expected to be high, whereas the productivity is low. This is due to human intervention and efforts carried out to contain the epidemic as well as immunity gained by individuals who contracted the disease (Schoenberg et al., 2017). This interplay between the prevalence of a disease and its transmission rate suggests an inverse relationship between conditional intensity and productivity at a specific time t.

Hence, the recursive model was modified to take into account this relationship by defining productivity as a decreasing function of conditional intensity. For the remaining discussion, we will use $H(x) = kx^{-\alpha}$ as the productivity function. Hence, the temporal recursive model can be written as:

$$\lambda(t) = \mu + k \int_0^t \lambda_{t'}^{-\alpha} \beta \exp(-\beta(t-t'))dN(t')$$

One of the key differences between the recursive and Hawkes processes is the mean of

the process $\frac{1}{T}E(N(S))$. The mean of the recursive model is $\mu + \frac{k}{T}E\int_0^T \lambda_{t'}^{1-\alpha} d\mu(t')$, which is simplified to $\mu + k$ when α is set to 1 (Schoenberg et al., 2017). This highlights the difference in an impact of the background rate on the expected number of events. For the Hawkes process, when the background rate is doubled, it doubles the expected number of events (Schoenberg et al., 2017). However, for the recursive model, a change in the background rate has a significantly less effect on the expected number of events (Schoenberg et al., 2017).

Another major difference is the productivity. The productivity of Hawkes process is assumed to be constant at K. On the other hand, the productivity of the recursive process at a point τ_i is dependent on the corresponding λ at the point, which is given by $H\{\lambda(\tau_i)\}$ (Schoenberg et al., 2017). When α is set to 1, the expected value of the total productivity $\sum_{i=1}^{n} H\{\lambda(\tau_i)\}$ is given by (Schoenberg et al., 2017):

$$kE\int_0^T \frac{1}{\lambda_t} dN_t = kE\int_0^T \frac{1}{\lambda_t}\lambda_t d\mu = kT$$

The average productivity is thus $\frac{k}{\mu+k}$ since $\frac{N(S)}{T}$ converges to $\mu + k$ when α is set to 1. The productivity of the recursive model starts at $\frac{1}{\mu}$ and gradually averages to $\frac{k}{\mu+k}$, which differs from the constant productivity at K assumed in the Hawkes process (Schoenberg et al., 2017).

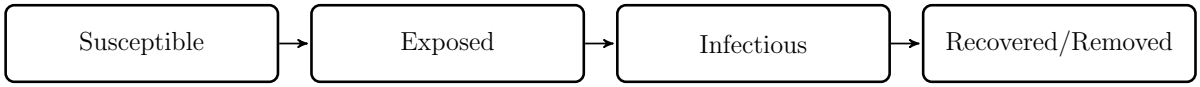
CHAPTER 3

Compartmental Model SEIR

The compartmental models are based on dividing the host population into a small number of compartments based on their stage of infection (Brauer et al., 1945). An individual in the population can only be assigned to one compartment at a given time and be transferred to another compartment.

Many variants of the basic SIR model have been developed to take into account different dynamics of disease transmission. For example, the SIRS model assumes recovery from disease does not confer permanent immunity, and rather it wanes at a constant rate (Rozhnova & Nunes, 2012). The SIRWS model takes into account boosting of immunity when recovered individuals are re-exposed to the disease (Rozhnova & Nunes, 2012). As another example, the SIRIS model is designed for the case when recovered individuals have immunity to a severe infection but still are susceptible to a mild infection (Rozhnova & Nunes, 2012). Lastly, the SIRSI model considers the case when recovered individuals become susceptible to the disease but with less susceptibility (Rozhnova & Nunes, 2012).

For our study, we will use the Susceptible-Exposed-Infectious-Recovered (SEIR) model, which is an extended version of the basic SIR model taking into account an exposed period when an individual is infected but not infectious yet. The SEIR model consists of four compartments as follows (Brauer et al., 1945):



- **Susceptible (S):** individuals who have no immunity against the disease and potentially become infected if exposed

- **Exposed (E):** individuals who are infected with the disease but currently unable to transmit the disease since their symptoms are not visible. This stage reflects an incubation period of disease
- **Infectious (I):** individuals who are infected and can transmit the disease
- **Recovered/Removed (R):** individuals who obtained life-long immunity to the infection, thus no longer being involved in transmission dynamics of disease

The total host population, denoted as N, is equal to the sum of individuals in all states, S + E + I + R = N. A set of differential equations are then defined to specify how the size of each compartment changes over time (Brauer et al., 1945). The rates of transfer between the compartments are specified as follows:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta(t)\frac{SI}{N} \\
\frac{dE}{dt} &= \beta(t)\frac{SI}{N} - \sigma E \\
\frac{dI}{dt} &= \sigma E - \gamma I \\
\frac{dR}{dt} &= \gamma I
\end{aligned}$$

Here $\beta(t)$ represents the transmission rate of infection at a given time t. For our analysis, we will assume that it exponentially decays over time at a rate k considering availability of whole-cell vaccination against pertussis in the 1940s (Kuchar et al., 2016).

$$\beta(t) = \beta_0 \exp(-kt)$$

βN represents the number of people with whom one individual makes a contact on average assuming the contact was sufficient to transmit disease (Brauer et al., 1945). $\frac{S}{N}$ represents a probability that a random contact by an infectious is with a susceptible. Thus, the number of new infections per unit time is equal to $\beta N \frac{S}{N} I$ (Brauer et al., 1945). In our equation,

we divide it by N so that it represents the per capita rate at which susceptible individuals acquire infection (Blackwood & Childs, 2018).

The infected individuals are then transferred to the exposed state during which they are not yet infectious to transmit the disease. σ is defined as 1/incubation period, representing the percentage of exposed individuals becoming infectious per unit time.

The infectious state therefore has an influx of σE number of individuals per unit time while simultaneously losing γI number of individuals who recovered from the disease and acquired life-long immunity. Here γ represents the rate of recovery, given by an inverse of the infectious period.

The above SEIR model does not take into account demographic effects such as birth and death rates of the population. This implies that there is no entry into or departure from the population except for deaths from the disease (Brauer et al., 1945). When the time scale of disease is much faster than the time scale of births and deaths, vital dynamics on the population may be ignored (Brauer et al., 1945). In case of pertussis, the span of disease is fairly short; symptoms of pertussis last on average 12 weeks after 1 to 2 weeks of an incubation period. Since the number of births and deaths in unit time, one day in our case, is negligible, it seems reasonable to exclude birth and death rates from the equations.

The key quantity of interest in the SEIR model is the basic reproductive number, R_0, which determines whether the disease is considered epidemic or not. R_0 is computed by multiplying $\beta(t)$ with the average duration of infectiousness, $\frac{1}{\gamma}$. It defines the number of secondary infections occurred by a single infectious individual over the course of infection (Chaffee et al., 2018). If R_0 is below 1, it implies the disease will fade out in the early stage, whereas if it is above 1, it will become a sustainable epidemic that can spread to a large proportion of the population (Brauer et al., 1945; Chaffee et al., 2018)

CHAPTER 4

Description of Data and Pertussis

Pertussis, commonly known as whopping cough, is an infectious disease caused by the bacterium called Bordetella pertussis (Centers for Disease Control and Prevention [CDC], 2015). The bacteria attach to the cilia of the respiratory cells, release toxins, and cause inflammation of the respiratory tract (CDC, 2015). The bacteria spread from person to person through the air when an infected person sneezes or coughs. Pertussis primarily affects infants younger than 6 months old, causing severe symptoms and possibly fatal complications (CDC, 2015). Children, adolescents and adults are less likely to become seriously ill with pertussis than infants and young children (CDC, 2015).

An incubation period of pertussis is usually 7-10 days, which is followed by three stages of illness (CDC, 2015). The first stage, the catarrhal stage, lasts 1-2 weeks with onset of runny nose, sneezing, fever, and a mild, occasional cough, which appears similar to a common cold (CDC, 2015). The second stage, the paroxysmal stage, is when the diagnosis of pertussis is usually suspected (CDC, 2015). During this stage, a patient has paroxysms of numerous, rapid coughs, and the symptoms usually last 1 to 6 weeks (CDC, 2015). In the last stage called convalescent stage, a patient gradually recovers from the illness and has less paroxysmal coughs that disappear in 2-3 weeks (CDC, 2015). An individual infected with pertussis is most infectious during the catarrhal period and the first 2 weeks after onset of cough, which is approximately 21 days (CDC, 2015). For the SEIR model, we will use 8.5 days and 21 days for an incubation period and duration of infection in order to determine σ and γ.

In 20th century, pertussis was one of the most common childhood diseases and a major cause of childhood mortality in the United States (CDC, 2015). Since introduction of whole-

cell pertussis vaccine in the 1940s, incidence of pertussis had significantly declined compared to the pre-vaccine era (CDC, 2015). However, since the early 1980s, the reported cases of pertussis have been gradually increasing (CDC, 2015). Particularly, the proportion of adolescent (11 to 18 years) and adult (19 years and older) cases has been increasing in recent years (CDC, 2015).

We obtained our data from Project Tycho, www.tycho.pitt.edu (Van Panhuis et al., 2013). The data consists of weekly number of pertussis occurrences reported in Nevada from 1888 to 2017. For our study, we decided to remove the cases having occurred prior to 1940 as the pertussis vaccine began to be in widespread use in mid-1940 (CDC, 2015). From 1940 to 2017, there were in total 1,122 reported cases of pertussis in Nevada. For the recursive and Hawkes processes, we converted the weekly aggregate number of events into a series of times drawn uniformly from the corresponding 7-day interval. For example, if 5 cases were recorded during the first week of 1940, we uniformly chose 5 different times within the week and considered them as when each transmission of pertussis happened. All times were then converted into the number of days elapsed since 1 January 1940. We split the data into training and test data sets so that we can evaluate predictive power of the model using out-of-sample data. We chose data points in the first 90% of the time range as the training data, which is from September 29, 1940 to January 24, 2010. The remaining data points were considered as the test data, which is from January 25, 2010 to October 7, 2017. In total, the training and test data sets consist of 946 and 176 cases, respectively.

CHAPTER 5

Estimation

The parameters of a point process such as the Hawkes and recursive processes can be estimated by maximizing its log-likelihood (Schoenberg et al., 2017). The log-likelihood of the Hawkes and recursive models can be written as (Ozaki, 1979):

$$\log L(t_1, t_2, \ldots, t_n|\theta) = \int_0^T \log \lambda(t|\theta) dN(t) - \int_0^T \lambda(t|\theta) dt$$

The gradient and Hessian of the log-likelihood are given by (Ozaki, 1979):

$$\frac{\delta \log L}{\delta \theta_i} = -\int_0^T \frac{\delta \lambda(t|\theta)}{\delta \theta_i} dt + \int_0^T \Big(\frac{\delta \lambda(t|\theta)}{\delta \theta_i} \Big/ \lambda(t|\theta) \Big) dN(t)$$

$$\frac{\delta^2 \log L}{\delta \theta_j \delta \theta_i} = -\int_0^T \frac{\delta^2 \lambda(t|\theta)}{\delta \theta_j \delta \theta_i} dt + \int_0^T \Big[\frac{\frac{\delta^2 \lambda(t|\theta)}{\delta \theta_j \delta \theta_i} \lambda(t|\theta) - \frac{\delta \lambda(t|\theta)}{\delta \theta_i} \frac{\delta \lambda(t|\theta)}{\delta \theta_j}}{\lambda(t|\theta)^2} \Big] dN(t)$$

The parameter vectors θ that maximize the log-likelihood can be estimated by applying the Newton-Raphson method at each updating stage of the function maximization (Ozaki, 1979). The standard errors can be also approximated by taking the diagonal elements of the inverse of the Hessian of the log-likelihood (Schoenberg et al., 2017).

One of the ways to evaluate the model fit is to compute the Stoyan-Grabarnik statistic, which is given by (Baddeley et al., 2005):

$$\frac{\sum_{i=1}^n \frac{1}{\lambda(t_i)}}{T}$$

In addition, the following ratio is also used as a diagnostic tool to assess the model fit:

$$\frac{\int_0^T \lambda(t)dt}{N}$$

Since the expected values of $\int \frac{1}{\lambda(t_i)}$ and $\int \lambda(t)$ are T and N as shown below, as the maximum likelihood estimators converge to their optimal states, the above ratios become extremely close to 1 (Schoenberg et al., 2017).

$$\begin{aligned} E\big[\sum_{i=1}^{n} \frac{1}{\lambda(t_i)}\big] &= E\big[\int_0^T \frac{1}{\lambda(t)} dN\big] \\ &= E\big[\int_0^T \frac{1}{\lambda(t)} \lambda(t)dt\big] \\ &= E\int_0^T 1dt \\ &= T \end{aligned}$$

$$E\big[\int_0^T \lambda(t)dt\big] = E\big[\int dN\big] = E(n) = N$$

For the recursive model, the sum of λ can be approximated as follows (Schoenberg et al., 2017):

$$\begin{aligned} \int_S \lambda(s)d\mu &= \int_0^T \{\mu + \int_0^t H(\lambda(s))g(t-s)dN(s)\}dt \\ &= \mu T + \int_0^T H(\lambda(s)) \int_0^{T-s} g(u)dN(s)du \\ &\approx \mu T + \int_0^T H(\lambda(s))dN(s) \\ &= \mu T + \sum_i H(\lambda(\tau_i)) \end{aligned}$$

For the Hawkes model, the approximation of $\int \lambda(t)$ can be written as $\mu T + nK$, where $\sum H(\lambda(\tau_i))$ is replaced with the sum of n constant Ks.

As an alternative, the model fit can be assessed by examining super-thinned residuals. Super-thinning is an integration of thinning and superposition techniques that are commonly

used for a residual analysis of spatial-temporal point processes (Clements et al., 2013). Since a test based on residuals formed by either method tends to have lower power when the conditional intensity of an original process is volatile, a more powerful approach was created by combining the two techniques (Clements et al., 2013). Super-thinning, a combination of thinning observed points and superposing simulated points, is supposed to create a homogeneous residual point process if and only if the estimate of conditional intensity λ of the original point process is correct (Clements et al., 2013).

First, a tuning parameter, denoted by k, has to be defined. For our analysis, we will use the mean of lambdas at time t in the training data as k. A thinned residual process is obtained by keeping each observed point independently with probability of $\frac{k}{\hat{\lambda}(t)}$ only when $\hat{\lambda}(t)$ is greater than k. This is followed by simulating a homogeneous Poisson process with a rate k and independently keeping each simulated point with probability of $\max\{\frac{k-\hat{\lambda}(t)}{k}, 0\}$. This step is equivalent to superposing points from a Poisson process with a rate equal to $\max\{k - \hat{\lambda}(t), 0\}$ (Clements et al., 2013). Then, the combined process is expected to form a homogeneous Poisson process with rate k if and only if $\hat{\lambda}$ is equal to the true λ almost everywhere. Hence, goodness-of-fit of $\hat{\lambda}$ can be assessed by inspecting uniformity of points in the super-thinned process (Clements et al., 2013). Given that the super-thinned process is a homogeneous Poisson process, its inter-event times are expected to follow the exponential distribution with mean of $\frac{1}{k}$ (Schoenberg et al., 2017). In addition to super-thinned residuals, uniformity of standardized inter-event times can be inspected to assess fit of the model.

The estimated parameters of the recursive model are (μ, k, β, α) = (0.006/day, 0.816, 0.495/day, 0.014) with corresponding standard error estimates (0.0005/day, 0.0341, 0.0332/day, 0.0194). The small value of α suggests pertussis is likely to have constant productivity, which indicates the fitted model is almost reduced to the Hawkes process. The estimated background rate is 2.3 cases per year, which implies a new outbreak occurs once every 6 months. The fitted recursive model can be written as follows:

$$\lambda(t) = 0.006 + 0.816 \int_0^t \lambda_{t'}^{-0.014} 0.495 \exp(-0.495(t - t'))dN(t')$$

Figure 1 displays probability of observed points being attributable to a new outbreak as opposed to contagion, which was computed by dividing μ by $\lambda(t)$. The points having probability smaller than 0.5 account 84% of the total number of cases in the training data set; however, the vast majority of them are concentrated around the early 1940s. Since 1970, we can notice the points attributable to the background rate account a higher proportion of the observed cases. The background rate μ can be interpreted as the rate of outbreaks that were not triggered by contagion within Nevada. In other words, it describes the infections transmitted by any other sources than infectious individuals in Nevada. Since our data is spatially limited to the state, these cases are considered as new outbreaks. Hence, we can infer from the plot that the proportion of infections that were caused by external sources has increased since 1970.

The β of 0.5 suggests that any new arrival of event will increase the conditional intensity roughly by 0.4 (βK) if we ignore the $\lambda_{t'}^{-\alpha}$ part given the α value close to zero, while this impact decays at a rate of 0.5 per unit time. The inverse of β can be interpreted as the mean triggering time for each transmission. According to our model, β of 0.5 corresponds to a triggering time of 2 days. In other words, an occurrence of infection is expected to be followed by a subsequent one in 2 days.

The mean triggering time is significantly shorter than the current understanding of an incubation period. One of the possible reasons for the difference is the assumption of constant μ in our model. As mentioned above, the background rate μ represents the rate of outbreaks caused by external sources other than contagion. It is possible that the background rate has been increasing due to various reasons such as frequent traveling, mutation of bacteria or waning immunity. However, since μ is assumed to be constant in our model, β has to be adjusted to reflect any change in the background rate. Hence, the higher number of new outbreaks would have resulted in raising β, which in turn shortens the mean triggering time.

Based on the fitted parameters, $\int_0^T \hat{\lambda}(t)dt \big/ N(0,T)$ and $\sum_{i=1}^n \frac{1}{\hat{\lambda}(t_i)} \big/ T$ correspond to 0.9984 and 0.9994, which indicates the maximum likelihood estimation successfully converged to its optimal point. Figure 2 displays a histogram of observed cases and a plot of estimated lambdas at time t calculated based on the fitted parameters. Figure 3 illustrates density of

points with estimated lambdas in a red dashed line. Figure 4 and 5 are a histogram and a lag plot of standardized inter-event times u_i of the super-thinned residuals.

From Figure 2 and 3, we can infer that the fitted lambdas represent the rate of events fairly well. According to Figure 5, the inter-event times seem evenly scattered without displaying any empty spaces or clusters. Figure 6 illustrates standardized inter-event times u_i of the super-thinned residuals at time t as well as cumulative inter-event times and their 95% confidence intervals based on 1,000 simulations. Similar to Figure 5, u_i appears uniformly scattered, and their cumulative values form a straight line lying in the 95% confidence interval. Both Figure 5 and 6 confirm goodness-of-fit of the recursive model.

The estimated parameters of the Hawkes process are (μ, K, β) = (0.006/day, 0.827, 0.4997/day), and these values are fairly close to the recursive model. This indicates for pertussis additional flexibility in productivity does not significantly contribute to improving the model fit. This is also supported by a marginal improvement in log-likelihood from -1899.799 to -1899.538, when the recursive model is used. From the estimated K in Hawkes and recursive processes, we can infer that roughly 82-83% of observed cases are attributable to the current rate of contagion.

In addition to the point processes, we also attempted to fit the SEIR model to data by using maximum likelihood estimation. The model considered σ and γ as fixed variables and the remaining ones, β_0 and k, as floating variables. As mentioned in chapter 4, we assumed an incubation period and duration of infection as 8.5 days and 21 days, respectively. The inverses of these values were then used as σ and γ. The transmission rate was assumed to exponentially decay since December 1, 1940, which is the first reported case in our data. We believe it is a reasonable assumption considering that the whole-cell vaccination against pertussis was developed in the 1930s and became widely available in the 1940s (CDC, 2015; Kuchar et al., 2016). The initial value of the total population was set to 1 million, given by the average historical population in Nevada from 1940 to 2010. Starting from a single infectious individual, the cumulative number of infected cases, denoted by C, was simulated by deriving solutions to $\frac{\delta C}{\delta t} = \sigma E$. The ode function from deSolve package in R was used to numerically solve differential equations. The simulated infected cases C were then used to

obtain maximum likelihood estimates based on the assumption that occurrences of pertussis follow the Poisson process (Chaffee et al., 2018). The MLE was solved by using the Nelder-Mead optimization algorithm implemented in the optim function in R (Althaus, 2014).

The fitted parameters of the SEIR model were (β_0, k) = (0.0693, 0.0007). The β_0 of 0.0693 implies that a contact with an infectious individual initially had 0.0693 probability of transmitting disease, and this rate has exponentially decayed at a rate of 0.0007 per unit time since December 1, 1940. In other words, approximately one person out of 14 contacts was likely to be infected with the disease back in 1940s; however, this impact has been declined since then at a rate of k. The primary quantity of interest in the SEIR model is $R_0(t)$, given by the transmission rate $\beta(t)$ multiplied by the average duration of infectiousness, $\frac{1}{\gamma}$ (Chaffee et al., 2018). $R_0(t)$ represents the number of transmissions of disease by one infectious individual until the infectious individual recovers or dies (Chaffee et al., 2018). In our case, $R_0(t)$ turns out to be 1.4555, which defines pertussis as a sustainable epidemic.

For the SEIR model, goodness-of-fit was also assessed by applying the super-thinning method. The conditional intensity rate, $\lambda(t)$, was estimated by multiplying $\beta(t)$ with the prevailing number of infectious individuals at time t (Chaffee et al., 2018). Given that an infectious period of pertussis is 21 days, the aggregate number of cases reported during 21 days prior to the corresponding time t was used as the prevailing number of infectious individuals. Similar to the previous case, the average of lambdas calculated based on the above method was used as a tuning parameter k.

Figure 7, 8 and 9 display plots of standardized inter-event times of super-thinned residuals simulated based on the SEIR model. In general, sparsity of points indicates areas where the model over-predicts the number of cases, while clustering indicates areas where it under-predicts (Chaffee et al., 2018). This is because over-prediction causes more points to be thinned out and fewer points to be superposed (Clements et al., 2013). From Figure 9, we can observe points are particularly sparse between 1940 and 1950, which suggests the SEIR model is likely to over-predict this period. However, from Figure 10, we can notice the SEIR model has lower estimates for lambda than the recursive model during this period. This can be explained by different choices for the tuning parameter k since the relative amount of

thinning and superposition is controlled by it. The tuning parameter k used for the SEIR and recursive models were 0.64 and 1.26 respectively determined by the mean of previous lambdas. Since the super-thinning technique considers a relative difference between lambdas and tuning parameter to determine the point to be thinned and superposed, it is possible to display sparsity despite lower estimates of lambdas. From the analysis of super-thinned residuals, we can infer that the recursive model fits the data better than the SEIR model.

CHAPTER 6

Prediction

The weekly incidence of pertussis was projected using the fitted models and compared to the test data to evaluate predictive performance. Forecasting started from the second week of observing more than zero cases so that the model has at least one week of data to incorporate in its projection.

For the recursive model, two different methods were used for prediction. The first method is simply summing conditional intensities on each day in the forecast week given that the conditional intensity describes a daily rate of contagion. The conditional intensities, $\lambda(t)$, were computed based on observed points up to the first date of each forecast period.

The second method is to simulate a recursive point process for each forecast week using the thinning technique of Lewis and Shedler (1979) and consider the total number of simulated points over 7 days as the estimated weekly incidence. The thinning technique is based on the idea that an inhomogeneous Poisson process can be simulated by thinning points from a homogeneous process (Chen, 2016). If we generate points from a homogeneous Poisson process with rate b, where b is set to some large value, and delete the sorted points independently with probability of $1 - \frac{\lambda(t)}{b}$, the remaining points will form a point process with the conditional intensity function of $\lambda(t)$ (Schoenberg et al., 2018). $\lambda(t)$ at each candidate point was computed running the recursive model on all observed points as well as kept simulation points up to the time t. For our analysis, the maximum of lambdas at all points in our data was used as the tuning parameter b, which corresponds to 8.9.

For the SEIR model, the adaptive tau-leaping algorithm was applied to simulate a real-world outbreak. The algorithm simulates transitions from S to E, E to I, and I to R states based on a continuous-time Markov process. The transitions occur according to the transi-

tion rates calculated based on the model parameters, R_0, k, σ and γ as well as prevailing state populations. After each iteration of simulation, state populations are adjusted according to the simulation result. The updated populations are then used to recalculate the transition rates for the next iteration. The simulation process continues until it reaches a user-determined stopping time t. The tau-leaping mechanism accelerates the process by identifying time periods with length τ when all transition rates remain approximately constant, and all state variables remain greater than zero with probability close to 1 (Johnson, 2016). The mechanism then leaps over the identified periods and for each period, adds the net effect of the Poison-distributed number of transitions that would have happened (Johnson, 2016). The net effect can be written as follows:

$$X(t+\tau) \approx X(t) + \sum_j y_j \Delta_j$$

where $y_j \sim \text{Poisson}(\tau\lambda_j)$

In order to use the tau-leaping algorithm, initial values of state populations have to be determined upfront. We used the Nevada population in 2010, which was 2.7 million, as the initial value of the total population. The initial number of infectious individuals was estimated by summing the number of cases observed over 21 days prior to the start date of each simulation. The initial number of exposed and recovered individuals was assumed to be zero. The susceptible population was then computed by subtracting the prevailing number of infectious individuals from the total population.

We ran each simulation over 7 days and considered the total number of people transferred from susceptible to exposed states as prediction of the weekly incidence of pertussis. After each simulation, the total population, exposed and recovered individuals were reset to the initial values, and the infectious population were re-calculated based on the start date of the following simulation period. For our analysis, the tau-leaping method was applied using the adaptivetau package in R.

Figure 12, 13 and 14 illustrate the cumulative weekly number of observed and predicted cases. The prediction values in Figure 12 are the sum of lambdas calculated based on the

recursive model. For Figure 13 and 14, the prediction values were calculated by running simulations on the recursive process and the SEIR model, respectively. The line between the two points was drawn to visualize the size of prediction error.

From Figure 12, we can infer that the sum of lambdas tends to under-predict incidence of pertussis. In particular, the model was unable to forecast a sudden leap in the number of cases, suggesting its limitation on predicting unprecedentedly high occurrence rate. Most of the steep increases in the number of cases are accompanied by overestimation on the following week and a sharp decline afterwards.

As shown in the Figure 13, the simulation method also failed to predict unusually high number of cases. Interestingly, it appears to overestimate the following week more excessively than the sum of lambdas. Also, the simulation method appears to produce more irregular predictions; for example, 16 infections were predicted to happen during the 2nd week of 2014 when there were zero cases observed for the previous 7 weeks. Since the thinning algorithm relies on randomness to determine simulated points, it is not surprising to observe more extreme errors than the sum of λ.

From Figure 14, we can observe the SEIR model seems to have a tendency to overestimate which differentiates it from the recursive model. Similar to the thinning algorithm, the tau-leaping algorithm also produced more extremely deviated predictions than the sum of lambdas.

Besides the plots, the sum and mean of squared errors were used to assess goodness-of-fit and predictive performance of the models.

Model, Method	Test Data	Train Data
Recursive, Sum of $\lambda(t)$	569	12,231
Recursive, Thinning Algorithm	1,137	15,504
SEIR, Tau Leaping Algorithm	796	18,615

Table 6.1: Sum of squared errors for different models, on the training and test data sets, respectively

Model, Method	Test Data	Train Data
Recursive, Sum of $\lambda(t)$	1.48	3.40
Recursive, Thinning Algorithm	2.96	4.31
SEIR, Tau Leaping Algorithm	2.07	5.18

Table 6.2: Mean of squared errors for different models, on the training and test data sets, respectively

Note. For the training period, the average of historical populations in Nevada from 1940 to 2010 was used as the initial total population for SEIR model predictions.

From the sum of squared errors, we can infer that the recursive model describes the spread of pertussis better than the SEIR model. In terms of prediction methods, the sum of lambdas seems to be a better approach than simulation since simulation tends to generate highly deviated predictions. Hence, the sum of lambdas estimated by the recursive model appears to be the best method for forecasting weekly incidence of pertussis despite its tendency to underestimate.

CHAPTER 7

Conclusion and Remarks

In this study, we have presented the recursive process, a refined version of the Hawkes process, as well as the SEIR model to describe the spread of pertussis. Based on our analysis, the recursive process appears to perform well in explaining the spread of pertussis despite its rare application to epidemics. This suggests that the recursive point process can be considered as a useful method for modeling and forecasting the spread of infectious diseases.

According to the recursive point process, 0.006 infections of pertussis per day occur as a new outbreak independently from the previous cases. Prior to 1970, the vast majority of reported cases were attributable to secondary contagion from the existing outbreaks. During the last few decades, however, new outbreaks seem to account a higher proportion of cases. The estimated triggering time for each transmission, given by inverse of β, was approximately 2 days. Based on the fitted parameters, we can infer that pertussis is more likely have the constant transmission rate independent from conditional intensity. According to the SEIR model, the transmission rate of pertussis was 0.07 in the 1940s and has been exponentially decaying since then at a rate of 0.0007 per day. The R_0 turns out to be higher than 1, defining pertussis as an epidemic.

From the evaluation of super-thinned residuals and sum of squared errors, the recursive point process appears to perform better than the SEIR model in explaining transmission dynamics of pertussis. The supper-thinned residuals of the recursive model were more evenly scattered without displaying any noticeable clusters or empty spaces. The sum and mean of squared errors were reduced by 28.5% when the recursive process was used. Since the recursive process and SEIR models are based on a fundamentally different mathematical approach, they provide two different descriptions of the spread of pertussis. Thus, even if

our analysis suggests the recursive model performs better, it is still recommended to use it along with the compartmental model so that we can understand how infectious diseases spread from the various angles.

One of the areas for further study is exploring different options for a triggering kernel of the recursive point process. Although an exponential decay function is a common choice, other parametric kernels such as the power-law, Tsallis Q-Exponential or Rayleigh can be considered as an alternative (Lima & Choi, 2018).

- **Power-law:**

$$PWL(K, c, p) = \frac{K}{(t + c)^p}$$

  It is appropriate for modeling a slower form of decay than the exponential (Lima & Choi, 2018).

- **Tsallis Q-Exponential:**

$$QEXP(a, q) = \begin{cases} ae^{-t}, & \text{if } q = 1 \\ a[1 + (q - 1)t]^{\frac{1}{1-q}}, & \text{if } q \neq 0 \text{ and } 1 + (1 - q)t > 0 \\ 0, & \text{if } q \neq 0 \text{ and } 1 + (1 - q)t \leq 0 \end{cases} \tag{7.1}$$

  It models decay in a hybrid way between exponential and power-law, being widely used in quantum optics and atomic physics as well as in statistics for variance stabilization and other purposes (Lima & Choi, 2018).

- **Rayleigh:**

$$RAY(\gamma, \eta) = \gamma t \exp(-\eta t^2)$$

  It is useful in modeling a distinct and non-monotonical decay (Lima & Choi, 2018).

Alternatively, kernel functions can be non-parametrically estimated when it is difficult to define triggering structure by one of the classic decaying functions. For example, Ke Zhou et al. studied application of the MMEL algorithm to non-parametrical learning of the

triggering function (2013). Adam Chaffee et al. non-parametrically estimated a triggering function and other parameters of the Hawkes process by applying a method assuming it is a piece-wise constant step function (2018).

Besides the triggering kernel, the productivity function can be further explored given that the current assumption of an inverse relationship with conditional intensity does not significantly improve the model fit from the original Hawkes process. Given that the rate of transmission depends on external factors such as hygiene, population density, a transmission path or availability of vaccination, the productivity might be explained better if these variables are included in the function. Alternatively, the transmission rate can be modeled to have a lagged relationship with conditional intensity considering that human efforts in containment of disease might take time to become fully effective.

Other area to explore is the assumption of fixed background rate μ. The background rate might vary depending on factors such as mutation of virus, waning immunity, a release of new vaccination or seasonality of the disease. Takahiro Omi et al. explored a Hawkes process model with a time-varying background rate for analyzing high-frequency financial data (2017). The log of time-dependent $\mu(t)$ was defined as a linear regression model of m basis functions and parameters a_j as follows: $\log \mu(t) = \sum_{j=1}^{m} a_j f^j(t)$ (Omi et al., 2017). The assumption of non-stationary background rate might improve the model fit especially when the data spans a long period of time.

For the SEIR model, we can consider exploring different compartmental models that describe transmission dynamics of pertussis better. For example, considering that vaccination against pertussis does not confer lifelong immunity, the SIRS model, which allows recovered individuals to become susceptible again once they lose immunity, can be explored as an alternative structure.

Also, a function for the transmission rate, β, can be further explored to improve the model fit. For example, Hanh T.H. Nguyen and Pejman Rohani defined β as a function of time representing the aggregation of children in school in their study of epidemics of pertussis (2007). The function returns a higher transmission rate during the school term

and a low transmission rate during the holidays. Additionally, Ottar N. Bjørnstad et al. incorporated seasonality and community size into estimation of the transmission rate for analyzing measles epidemics (2002). Currently, in our model, β_0 is assumed to exponentially decay since December 1940; however, this assumption results in a very small value of $\beta(t)$ for the period after 2000. This might not be an appropriate assumption given that incidence of pertussis has been increasing since 1980 (CDC, 2015). Hence, an equation reflecting this trend is likely to model the transmission rate better than a continuously decreasing one.

A further study of these considerations will allow us to refine the models and achieve better understanding of pertussis. We believe the concurrent use of point processes such as the recursive or Hawkes process with other epidemic models will describe the transmission dynamics of pertussis better and further help us prevent and mitigate the disease.

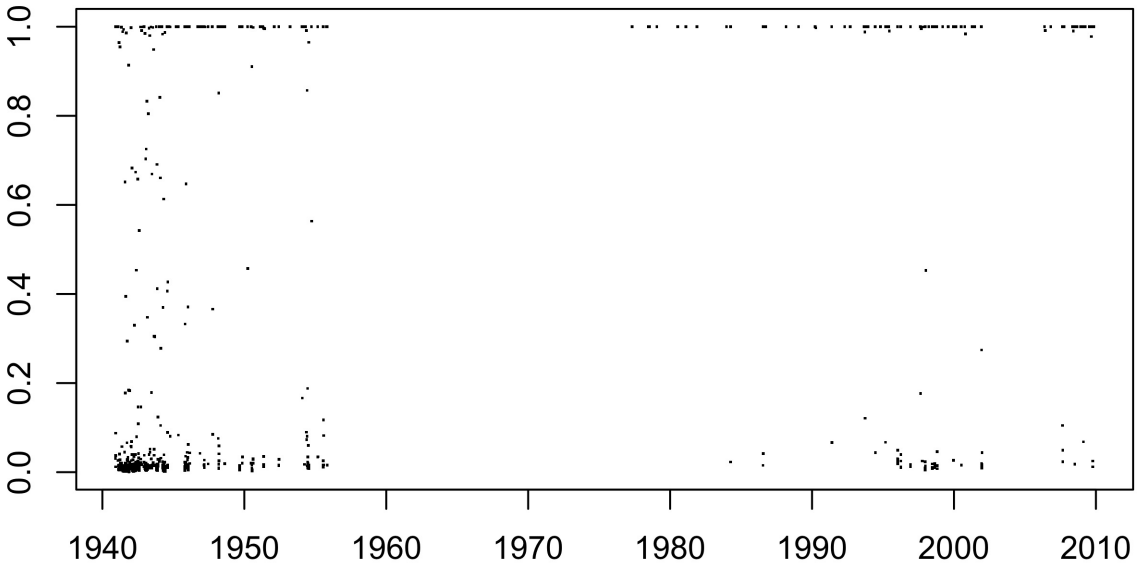


Figure 1: Probability of points being attributable to μ according to the fitted recursive point process.

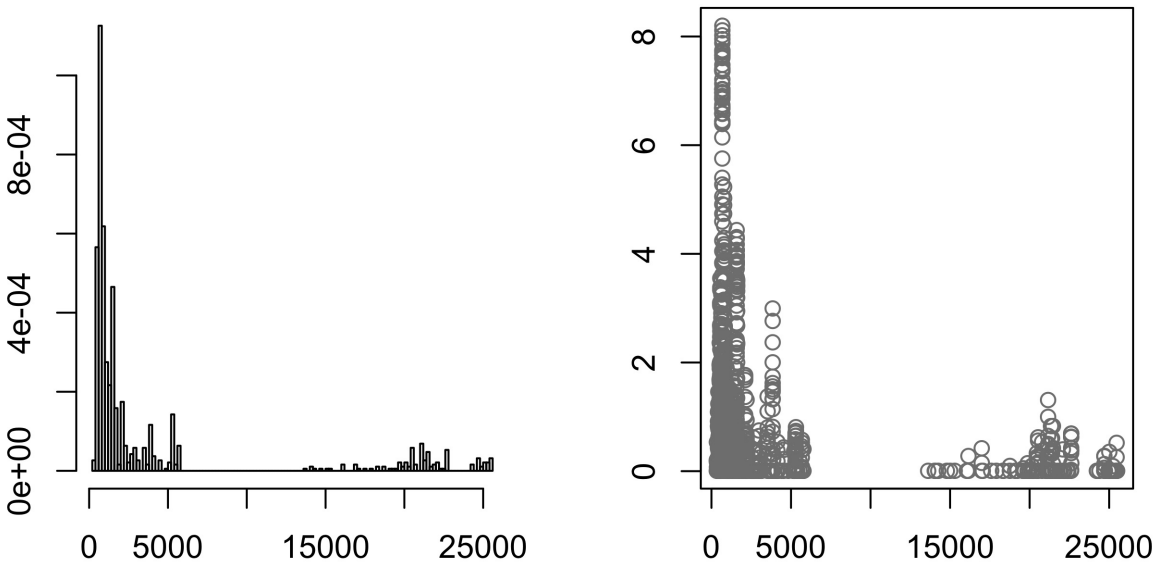


Figure 2: Histogram of points on the left and a plot of λ estimated by the recursive point process on the right.

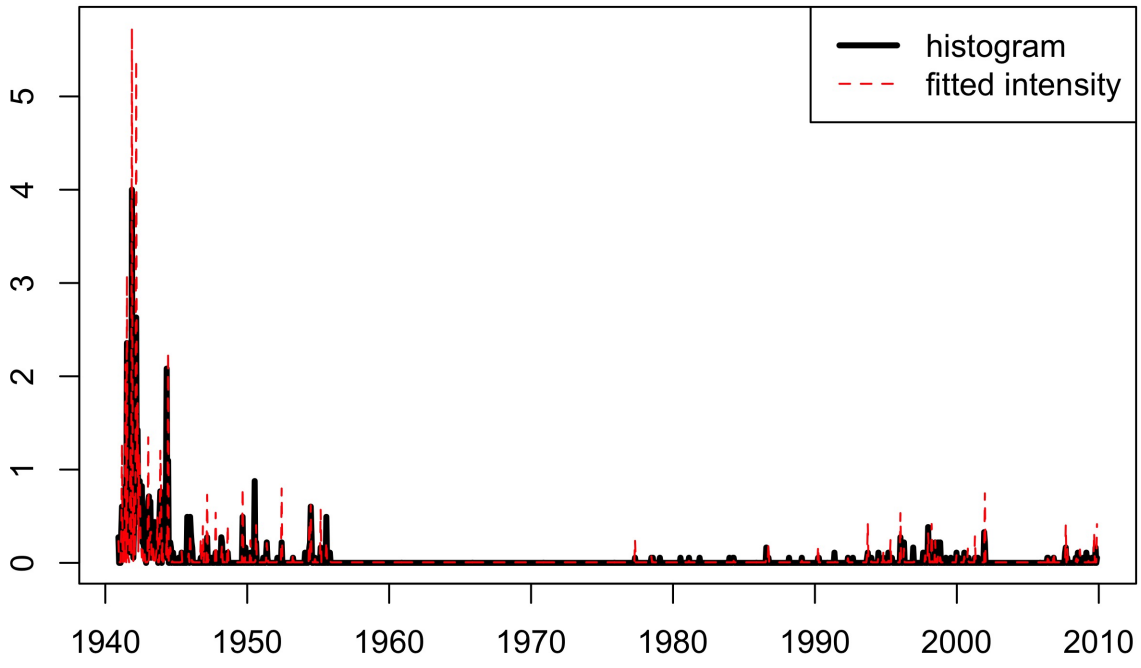


Figure 3: Density of points and fitted intensity estimated by the recursive point process.

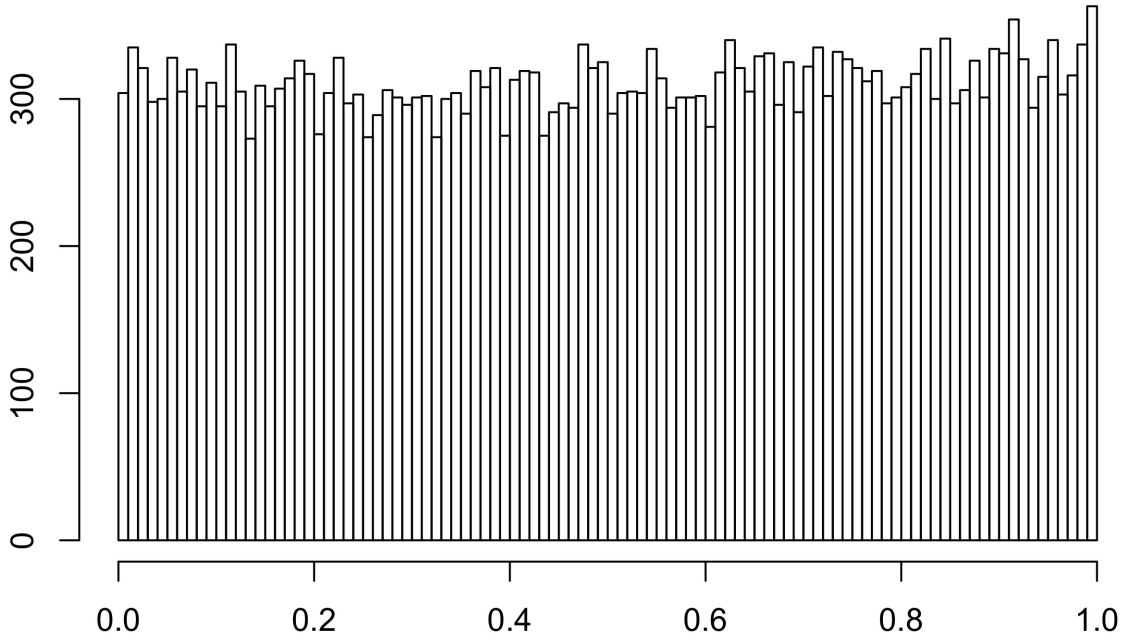


Figure 4: Histogram of the standardized inter-event times u_i of super-thinned residuals produced by the recursive point process.

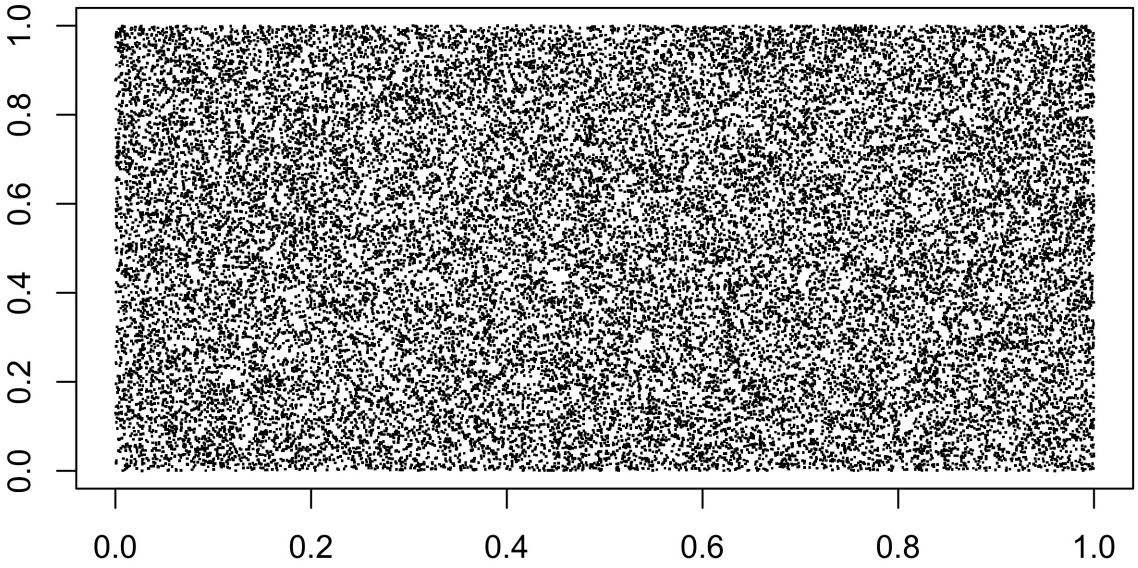


Figure 5: Lag plot of the standardized inter-event times u_i of super-thinned residuals produced by the recursive point process.

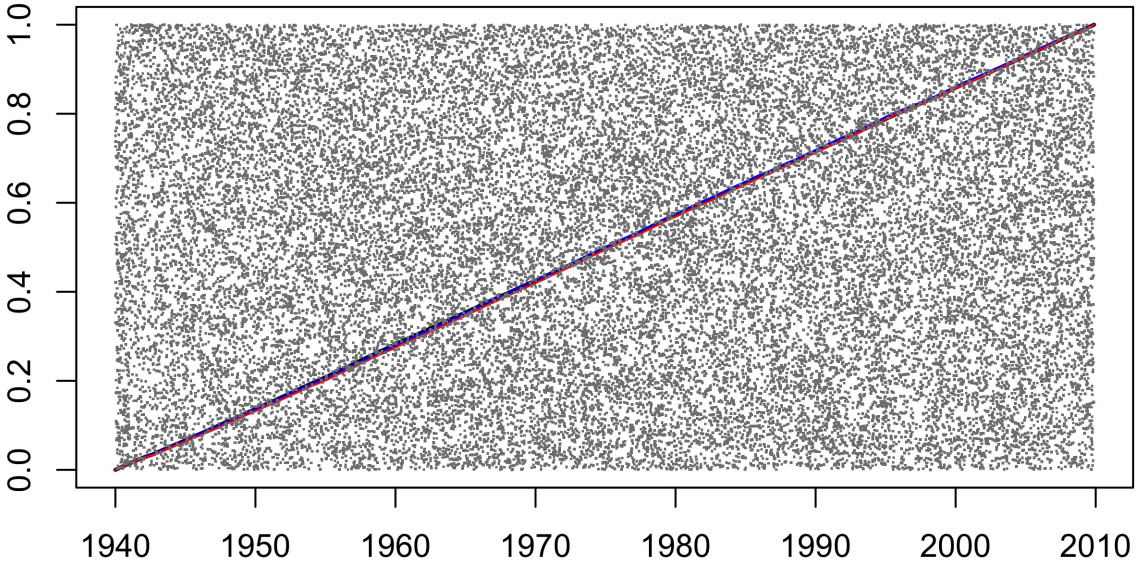


Figure 6: Super-thinned residuals and their corresponding standardized inter-event times u_i of the recursive point process.

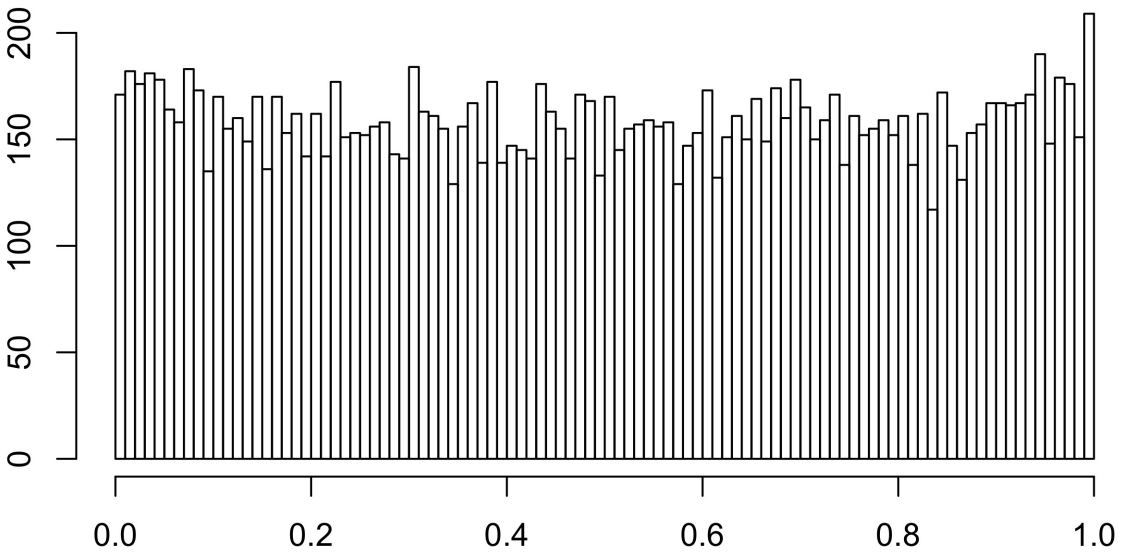


Figure 7: Histogram of the standardized inter-event times u_i of super-thinned residuals produced by the SEIR model.

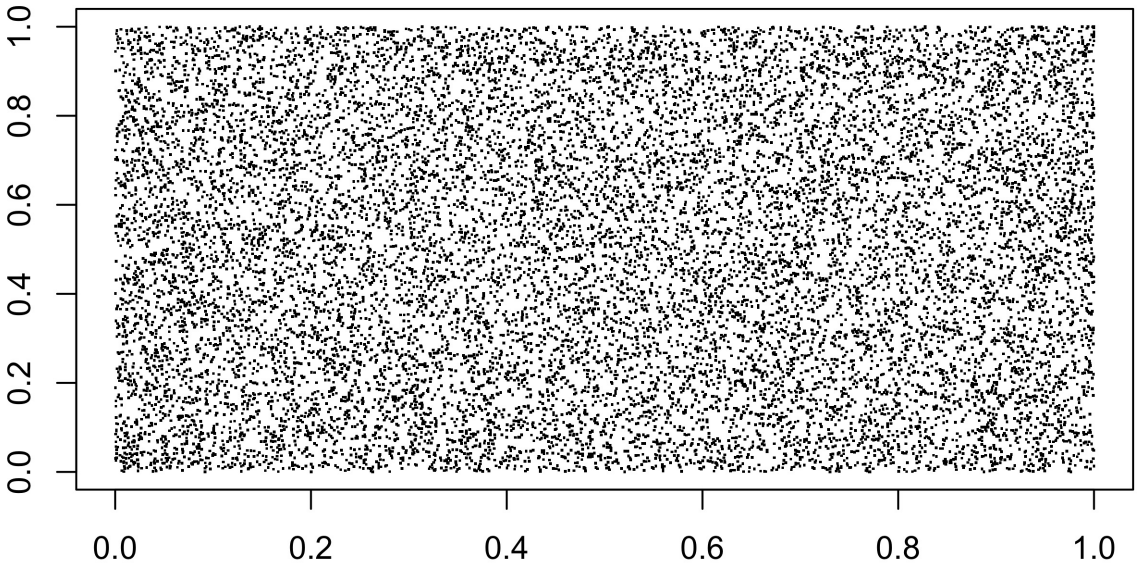


Figure 8: Lag plot of the standardized inter-event times u_i of super-thinned residuals produced by the SEIR model.

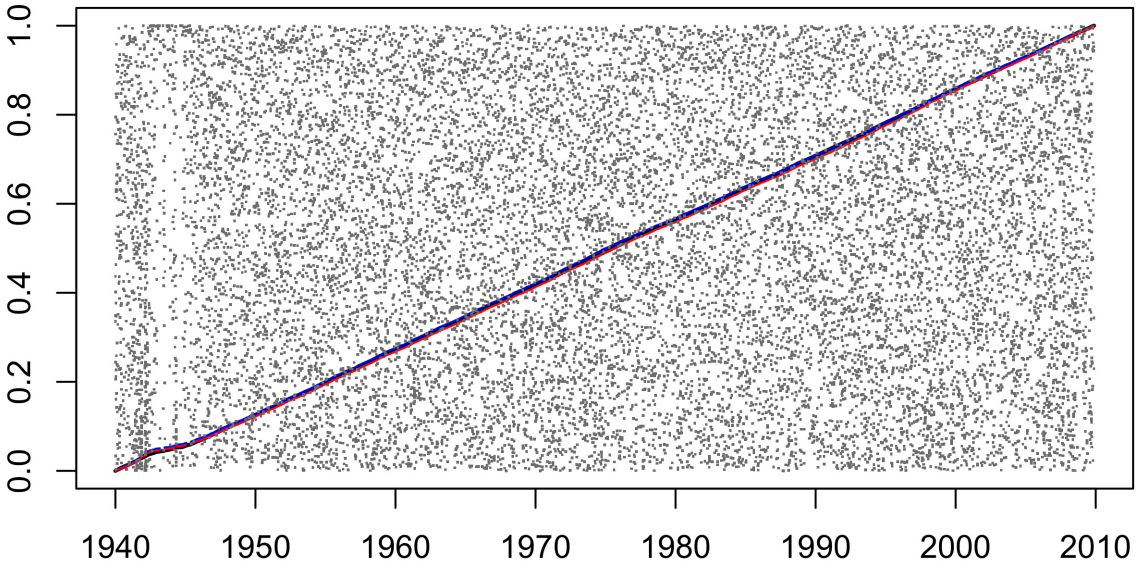


Figure 9: Super-thinned residuals and their corresponding standardized inter-event times u_i of the SEIR model.

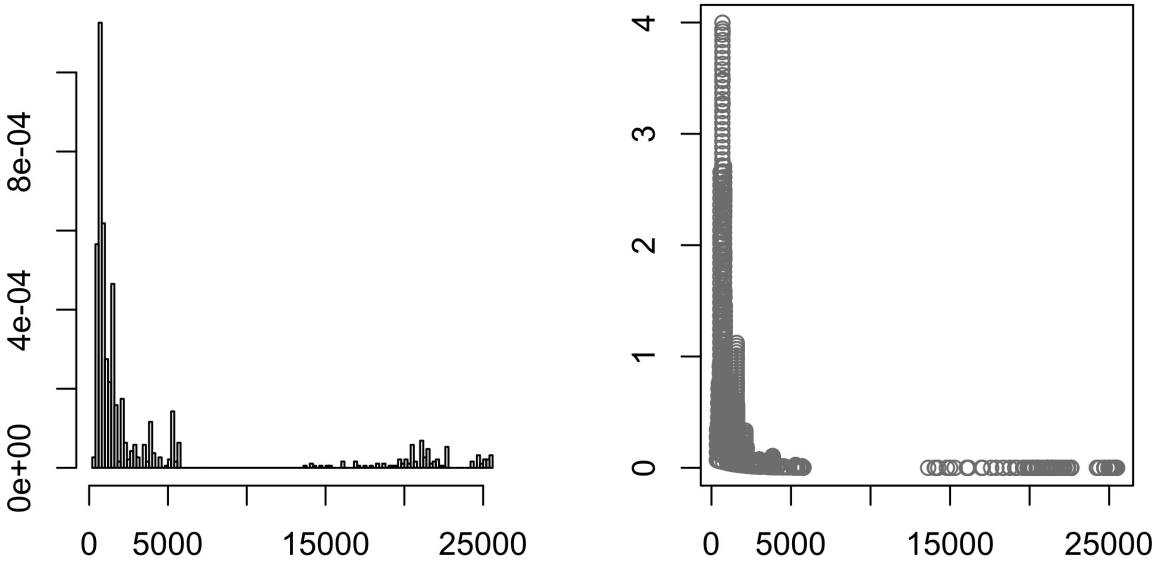


Figure 10: Histogram of points on the left and a plot of λ estimated by the SEIR model on the right.

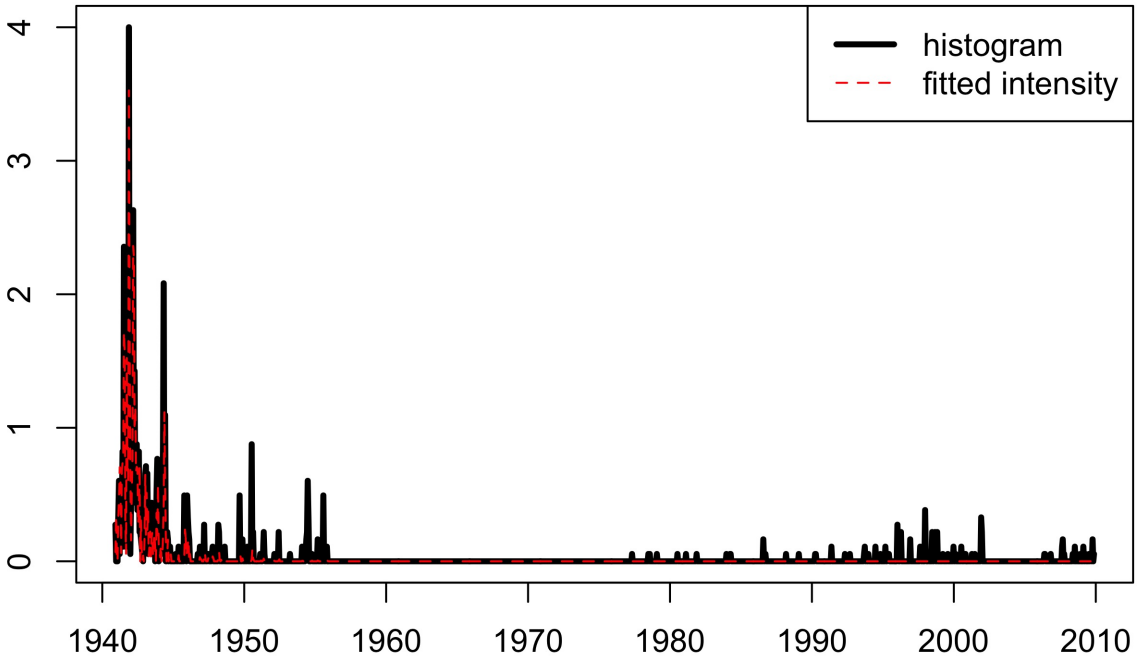


Figure 11: Density of points and fitted intensity estimated by the SEIR model.

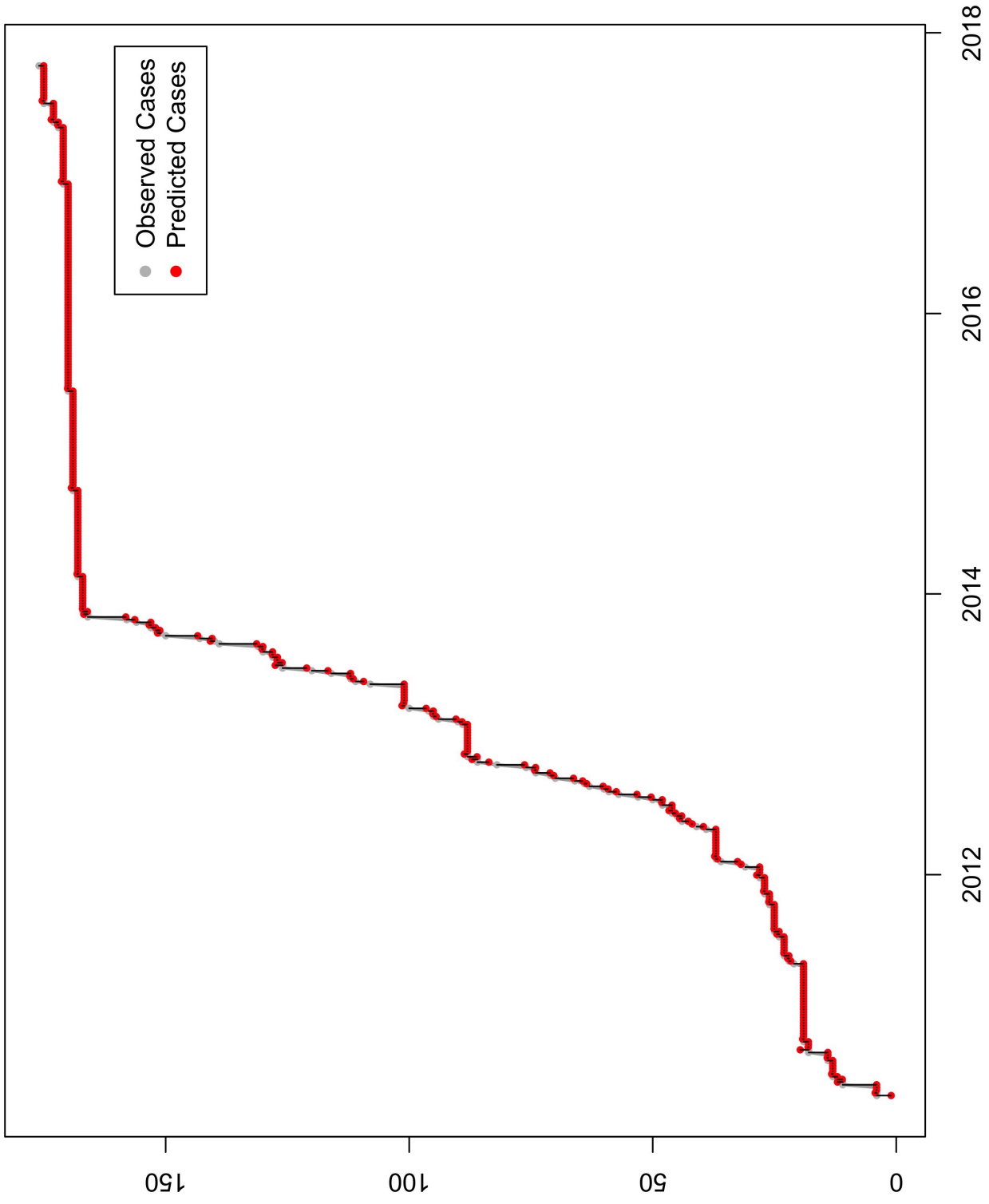


Figure 12: Cumulative weekly incidence of pertussis. Predicted values are the sum of daily intensities over the forecast week estimated by the fitted recursive point process.

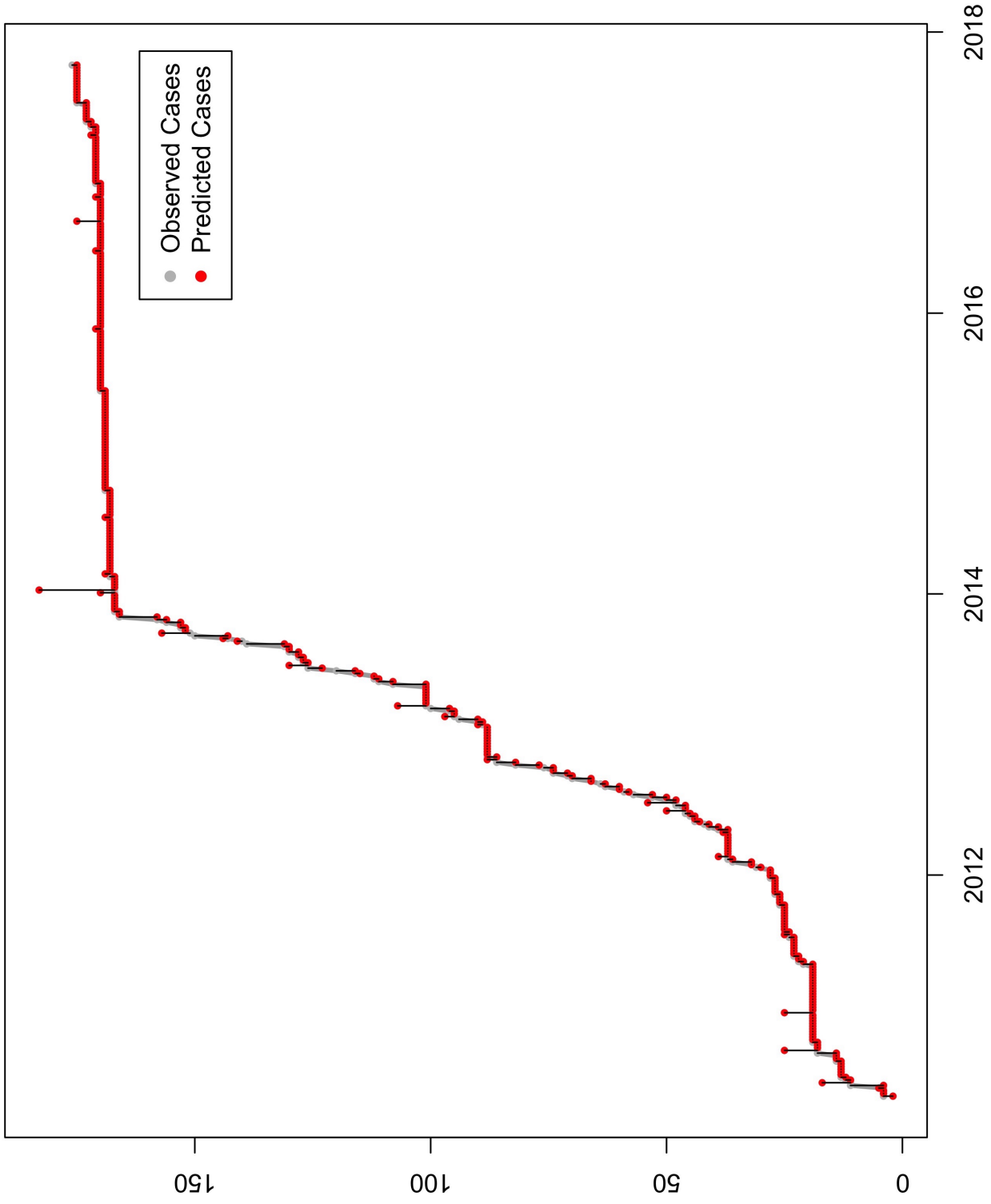


Figure 13: Cumulative weekly incidence of pertussis. Predicted values are the total number of simulated points over the forecast period. The simulation was conducted by running the thinning algorithm on the fitted recursive point process.

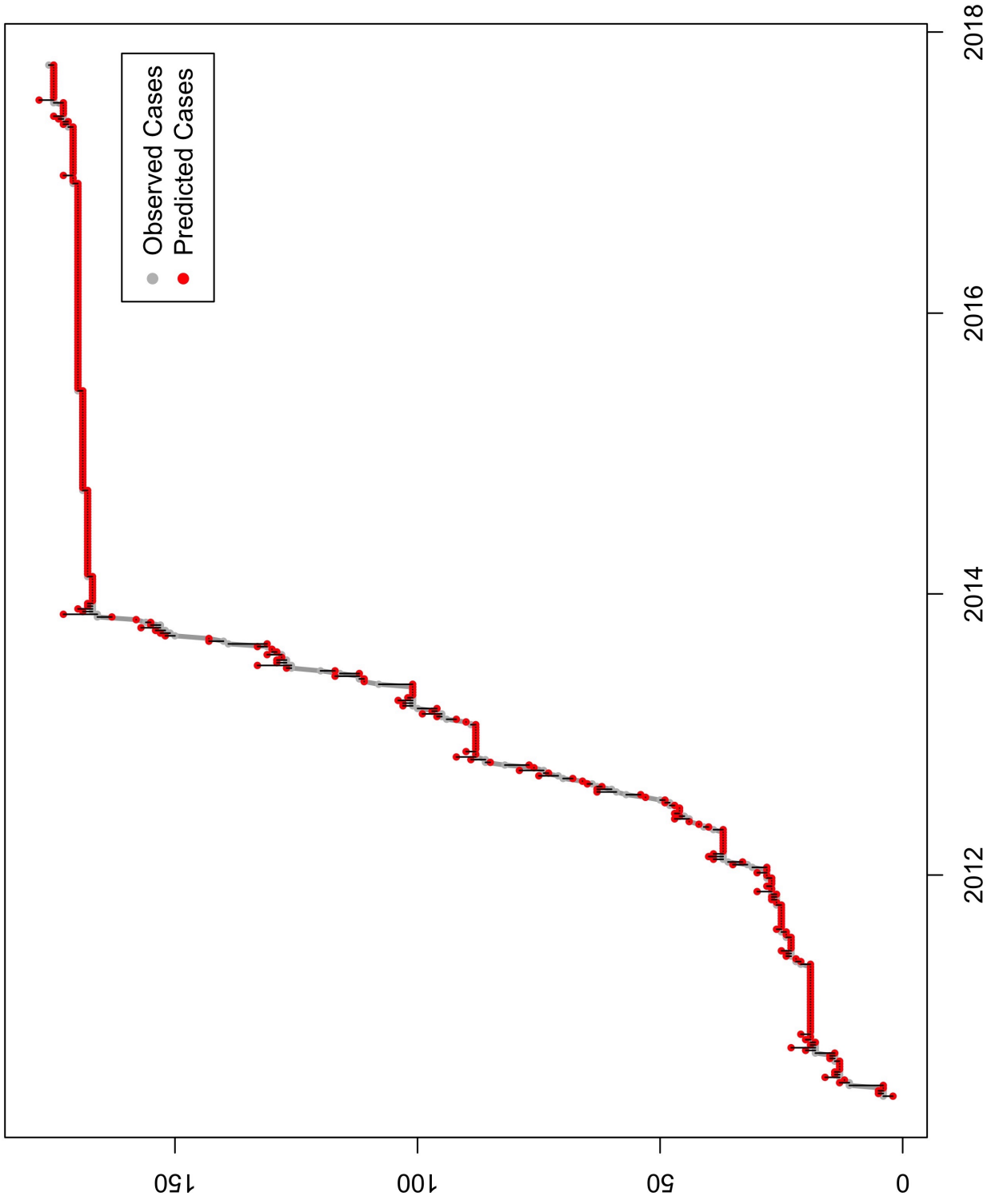


Figure 14: Cumulative weekly incidence of pertussis. Predicted values are the total number of simulated points over the forecast period. The simulation was conducted by running the tau-leaping algorithm on the SEIR model.

REFERENCES

[1] Althaus, C.L. (2014). Estimating the Reproduction Number of Ebola Virus (EBOV) During the 2014 Outbreak in West Africa. *PLOS Current Outbreaks*.

[2] Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005). Residual analysis for spatial point processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67 617-666.

[3] Balderama, E. et al. (2012). Application of branching point process models to the study of invasive red banana plants in Costa Rica. *JASA*, 107.498, pp. 467-476.

[4] Becker, N. (1977). Estimation for discrete time branching processes with application to epidemics. *Biometrics*, 33.3, pp. 515-522.

[5] Bjørnstad, O.N., Finkenstädt, B., Grenfell, B.T. (2002). Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates Using a Time Series SIR Model. *Ecological Monographs*, 72(2), 2002, pp. 169-184

[6] Blackwood, J.C., & Childs, L.M. (2018). An introduction to compartmental modeling for the budding infectious disease modeler. *Letters in Biomathematics*, 5:1, 195-221, DOI: 10.1080/23737867.2018.1509026

[7] Brauer, F., Driessche, P., Wu, J. (1945). *Mathematical Epidemiology*. Springer

[8] Campbell, P. T., McCaw, J. M., McVernon, J. (2015). Pertussis models to inform vaccine policy. *Human Vaccines & Immunotherapeutics*, 11:3, 669-678, DOI: 10.1080/21645515.2015.1011575

[9] Centers for Disease Control and Prevention. (2015). Chapter 16: Pertussis In *Epidemiology & Prevention of Vaccine-Preventable Diseases textbook*, 261-278.

[10] Chaffee, A. W., Park, J., Harrigan, R. J., Schoenberg, F. P. (2018). A non-parametric Hawkes model of the spread of Ebola in West Africa. *JASA*

[11] Chen, Y. (2016) Thinning Algorithms for Simulating Point Processes, Florida State University, Tallahassee, FL, 2016; available at https://www.math.fsu.edu/∼ychen/research/.

[12] Clements, R.A., Schoenberg, F.P., Veen, A. (2013). Evaluation of space-time point process models using super-thinning. *Environmetrics*, 23.7, 606-616

[13] Farrington, C.P., Kanaan, M.N., and Gay, N.J. (2003). Branching process models for surveillance of infectious diseases controlled by mass vaccination. *Biostatistics*, 4.2, pp. 279-295.

[14] Goufo, E.F., Maritz, R., Munganga, J.M. (2014). Some properties of the Kermack-McKendrick epidemic model with fractional derivative and nonlinear incidence. *Advances in Difference Equations*, 2014, 1-9.

[15] Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes, *J. Roy. Statist. Soc.*, **B33**, 438-443.

[16] Johnson P. (2016). adaptivetau: Efficient Stochastic Simulations in R.*R Foundation for Statistical Computing*; available at https://cran.r-project.org/web/packages/adaptivetau/vignettes/adaptivetau.pdf. Accessed February 1, 2016.

[17] Kermack, W.O. & McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society*, A 115.771, pp. 700-721.

[18] Kobayashi, R., & Lambiotte, R. (2016). TiDeH: Time-Dependent Hawkes Process for Predicting Retweet Dynamics. *ArXiv*, abs/1603.09449.

[19] Kuchar, E., Karlikowska-Skwarnik, M., Han, S., Nitsch-Osuch, A. (2016). Pertussis: History of the Disease and Current Prevention Failure. In: Pokorski M. (eds) Pulmonary Dysfunction and Disease. Advances in Experimental Medicine and Biology, vol 934. Springer, Cham

[20] Laub, P. J., Taimre, T., Pollett, P. K. (2015). Hawkes Processes *ArXiv*, preprint arXiv:1507.02822

[21] Lewis, P.A.W., and Shedler, G.S. (1979). Simulation of non-homogeneous Poisson processes by thinning. *Naval Res. Logistics Quart.* 26(3), 403-413.

[22] Lima, R., & Choi, J. (2018). Hawkes Process Kernel Structure Parametric Search with Renormalization Factors. arXiv:1805.09570v3

[23] Nguyen, H.T., & Rohani, P. (2007). Noise, nonlinearity and seasonality: the epidemics of whooping cough revisited. *Journal of the Royal Society, Interface*, 5 21, 403-13.

[24] Omi, T., Hirata, Y., Aihara, K. (2017). Hawkes process model with a time-dependent background rate and its application to high-frequency financial data. *Physical review*, E, 96 1-1, 012303.

[25] Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31, 145-155.

[26] Rachah, A. & Torres, D. F. M. (2017). Analysis, simulation and optimal control of a SEIR model for Ebola virus with demographic effects. *Commun. Fac. Sci. Univ. Ank. Ser. A1 Math. Stat.*, 67 (2018), no. 1, 179-197

[27] Rizoiu, M., Lee, Y.S., Mishra, S., Xie, L. (2017). A Tutorial on Hawkes Processes for Events in Social Media. *ArXiv*, abs/1708.06401.

[28] Rizoiu, M., Mishra, S., Kong, Q., Carman, M.J., Xie, L. (2018). SIR-Hawkes: Linking Epidemic Models and Hawkes Processes to Model Diffusions in Finite Populations. *WWW*

[29] Rozhnova, G., & Nunes, A. (2012). Modeling the long-term dynamics of pre-vaccination pertussis. *Journal of the Royal Society, Interface*, 9 76, 2959-70.

[30] Schoenberg, F.P., Hoffmann, M., Harrigan, R.J. (2017). A recursive point process model for infectious diseases. *Annals of the Institute of Statistical Mathematics*, 1-17.

[31] Srijith, P. K., Lukasik, M., Bontcheva, K., Cohn, T. (2017). Longitudinal Modeling of Social Media with Hawkes Process Based on Users and Networks. *ASONAM*

[32] van Panhuis, W.G., Grefenstette, J., Jung, S.Y., Chok, N.S., Cross, A., Eng, H., Lee, B.Y., Zadorozhny, V., Brown, S., Cummings, D., and Burke, D.S. (2013). Contagious diseases in the United States from 1888 to the present. *NEJM* , 369(22), 2152-2158.

[33] Zhou, K., Zha, H., Song, L. (2013). Learning Triggering Kernels for Multi-dimensional Hawkes Processes. *ICML*