

Article

# A Simultaneous Estimation of the Baseline Intensity and Parameters for Modulated Renewal Processes

Jiancang Zhuang <sup>1,\*</sup> and Hai-Yen Siew <sup>2</sup>

<sup>1</sup> The Institute of Statistical Mathematics, 10-3 Midori-Cho, Tachikawa 190-8562, Tokyo, Japan

<sup>2</sup> Independent Researcher, 4-27-115 Sakuragaoka, Higashi-Yamato 207-0022, Tokyo, Japan; haiyen79@gmail.com

\* Correspondence: zhuangjc@ism.ac.jp

**Abstract:** This paper proposes a semiparametric solution to estimate the intensity (hazard) function of modulated renewal processes: a nonparametric estimate for the baseline intensity function together with a parametric estimate of the model parameters of the covariate processes. Based on the martingale property associated with the conditional intensity, we construct a statistic from a residual analysis to estimate the baseline renewal intensity function, when the model parameters of the covariate processes are known. In addition, when the baseline intensity is obtained, the model parameters can be estimated using the usual maximum likelihood estimation. In practice, both the baseline intensity and model parameters are suggested to be estimated simultaneously via an expectation–maximization (E–M)-type iterative algorithm. A more important feature of the newly proposed algorithm is that, given  $n$  events in the observation dataset, its computation time is of order  $O(n^2)$ , while the Nelson–Aalen–Breslow estimator takes a computation time of order  $O(n^3)$ . For illustration, we apply the proposed estimation procedure to a set of data simulated from a modulated gamma renewal process and the aftershock sequence following the  $M_s$  8 Wenchuan earthquake, which occurred in Sichuan Province, China on 12 May 2008.

**Keywords:** martingale; residual analysis; Nelson–Aalen–Breslow estimator; survival analysis; point process

**MSC:** 62G05; 62M20; 60G55



**Citation:** Zhuang, J.; Siew, H.-Y. A Simultaneous Estimation of the Baseline Intensity and Parameters for Modulated Renewal Processes. *Axioms* **2022**, *11*, 303. <https://doi.org/10.3390/axioms11070303>

Academic Editors: Stefano Isola and Kevin H. Knuth

Received: 13 January 2022

Accepted: 20 June 2022

Published: 23 June 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The renewal process and its variants have been widely used in the studies of recurrent events [1,2] including episodes of illness, failures of a computer network system, wildfires in a particular forest, and major earthquakes on an active fault. Usually, a renewal process is taken as a starting point to model a sequence of such recurrent events, in which the lengths of the time intervals between each pair of neighboring events are identically and independently distributed. When the interevent times are also influenced by some external factors, the process then becomes a modulated renewal process (see e.g., [3]), which is formulated mathematically in the following way:

Consider a point process observed over time  $t \in [0, T]$ , and suppose that events occur at time  $0 < T_1 < T_2 < \dots < T_n$ . Let  $N([s, t])$  ( $s < t$ ) and  $\mathcal{H}_t$  denote the number of events occurring in the time interval  $[s, t]$ , and the history of the process at time  $t$ , respectively. We also write  $N(t)$  for  $N([0, t])$  and  $N(t_-)$  for  $N([0, t))$ . In general, a model for such a process is specified by its hazard function, also known as conditional intensity, defined as

$$\lambda(t) = \lim_{\Delta \rightarrow 0} \frac{\mathbf{E}\{N([t, t + \Delta]) \mid \mathcal{H}_t\}}{\Delta}, \quad (1)$$

i.e., the expectation of the number of events occurring in a unit time interval at  $t$  given the past history. For a modulated renewal process, the conditional intensity is often expressed in a multiplicative form:

$$\lambda(t) = \mu(t) Z_{\beta}(t), \tag{2}$$

where  $\mu(t) = \mu_0(t - T_{N(t-)})$ ,  $\mu_0(x)$  being the baseline hazard function and  $t - T_{N(t-)}$  being the left-continuous backward recurrence time at  $t$ , and  $Z_{\beta}(t)$  is a function of a time-varying covariate process with parameter vector  $\beta$ . In particular, when  $Z_{\beta}(t) = 1$ ,  $N(t)$  reduces to an ordinary renewal process, where the time differences between two consecutive events,  $X_i = T_i - T_{i-1}$ , for  $i = 1, \dots, n$ , are independently and identically distributed. In practice,  $\mu_0(t)$  can be interpreted as the internal property of the system, and  $Z_{\beta}(t)$  explains the external influences to the intensity, or the responses of the system to the external environments. For example, when modeling the recurrence of a characteristic earthquake in a particular fault or region,  $\mu_0$  describes the intrinsic recurrence time, and  $Z_{\beta}(t)$  represents the changes in tectonic environments, which may accelerate or decelerate the renewal of earthquake occurrences (e.g., [4]).

In past studies, the rich literature on survival analysis based on clinical data from a sample of patients or independent cluster samples (see, e.g., [3,5,6]) has deeply influenced studies on renewal processes and their variants. In survival analysis, the failure time  $X_i$  for the  $i$ th individual ( $i = 1, \dots$ ) are mutually independent with a modulated hazard function

$$h_i(x) = h_0(x) Z_{i,\beta}(x) \tag{3}$$

where  $h_0$  is the baseline hazard function common to all individual, and  $Z_i(x; \beta)$  is the covariate function specific to the  $i$ th individual. Refs. [3,7,8] proposed a partial likelihood estimate for  $\beta$  in the absence of knowledge of the baseline intensity, i.e., the partial likelihood estimator  $\hat{\beta}$  is obtained by maximizing in  $\beta$  the partial likelihood

$$\text{lik}(\beta) = \prod_{i=1}^n \frac{Z_{i,\beta}(X_i)}{\sum_{j=1}^n \mathbf{1}(X_j \geq X_i) Z_{j,\beta}(X_i)}, \tag{4}$$

where  $\mathbf{1}(A)$  denotes the indicator function of the event  $A$ . Cox also suggested that the asymptotic inference about  $\beta$  proceeds as if the partial likelihood were an ordinary likelihood function.

For the estimation of the cumulative baseline hazard function, [9,10] considered a nonparametric estimate for the cumulative intensity from censored survival data to assess the fit of parametric models graphically for engineering applications. Under the framework of counting processes, [11,12] illustrated the use of this estimate in the studies of Markov chains and other event history models. The Nelson–Aalen estimator takes the form of

$$\hat{H}_0(x) = \sum_{i: X_i < x}^n \frac{1}{\sum_{j=1}^n \mathbf{1}(X_j \geq X_i) Z_{j,\beta}(X_i)}, \tag{5}$$

for the integrated baseline hazard function

$$H_0(x) = \int_0^x h_0(u) du. \tag{6}$$

Ref. [12] also discussed the small and large sample properties of the estimate using the martingale methods. Stimulated by [9,13] replaced  $\beta$  by the Cox’s estimate  $\hat{\beta}$ , i.e.,

$$\hat{H}_0(x) = \sum_{i: X_i < x}^n \frac{1}{\sum_{j=1}^n \mathbf{1}(X_j \geq X_i) Z_{j,\hat{\beta}}(X_i)}, \tag{7}$$

The asymptotic properties were discussed by [5,14].

The connection between the renewal process and survival analysis is that instead of a series of absolute occurrence times,  $\{T_i : i = 1, 2, \dots\}$ , if we express the renewal process as a series of interevent intervals,

$$\{X_i : X_1 = T_1, X_i = T_i - T_{i-1}, i = 2, 3, \dots\},$$

the modulated renewal process is turned into survival data. This treatment enables us to apply the analyzing techniques from survival analysis to the modulated renewal processes. For the intensity estimation, for instance, [15] introduced a pseudo-martingale estimating equation based on such a time rearranging scheme. The word “pseudo-martingale” was used because the authors developed an estimating equation from the filtration with respect to the time intervals, but not the original occurrence times. For the modified renewal process, they in fact gave the same estimators as Cox’s partial likelihood estimator for  $\beta$  and the Nelson–Aalen–Breslow estimator for the cumulative baseline hazard function.

Unlike the tradition of avoiding a simultaneous estimation of the baseline intensity and model parameters in survival studies, we propose a pure martingale-based estimating procedure in the manner of an expectation–maximization (E–M)-type iterative algorithm by making use of the ideas of residual analysis of [16]. The ensuing sections are organized as follows. In Section 2, we first obtain a new estimating equation based on the martingale property of the conditional intensity to estimate the cumulative baseline intensity. Then, making use of the technique of partial integration, we also establish the maximum likelihood estimates for the model parameters. Section 3 briefly compares our estimate for the cumulative baseline intensity with the Nelson–Aalen and Breslow estimates. Section 4 illustrates the estimating performance of the proposed method by applying it to a simulated dataset from a modulated gamma renewal process and the aftershock data of the M8.0 Wenchuan earthquake, which occurred on 12 May 2008, in Sichuan Province, China. Finally, a residual analysis based on the transformed time sequence is conducted to test the goodness-of-fit in Section 5.

## 2. Theory and Methodology

### 2.1. Estimation of Baseline Intensity

To estimate the baseline intensity  $\mu_0(x)$ , we make use of the residual analysis for the temporal or spatiotemporal point process developed by [16], which was stimulated from the residual analysis for spatial point processes (see [17]). Let  $N$  be a simple temporal point process in the interval  $[0, T]$ , admitting a conditional intensity  $\lambda(t)$ . According to the martingale property of the conditional intensity, for any predictable process (measurable with respect to the  $\sigma$ -algebra that is generated by sets of the form  $(s, t] \times E$ , where  $E \in \mathcal{F}_s$ , and  $\mathcal{F}_s$  is the time filtration,  $\zeta(t) \geq 0$ ,

$$\mathbf{E} \left[ \int_D \zeta(s) dN(s) \right] = \mathbf{E} \left[ \int_D \zeta(s) \lambda(s) ds \right] \tag{8}$$

where  $D$  is any regular (Lebesgue measurable) subset of  $[0, T]$ . Equation (8) holds since  $dM(t) = dN(t) - \lambda(t)dt$  is a zero-mean martingale.

The first-order innovation with respect to a predictable function  $\zeta(t)$  and a measurable set  $D \subset [0, T]$  is defined by

$$V(D, \zeta, \lambda) = \int_D \zeta(s) dN(s) - \int_D \zeta(s) \lambda(s) ds \tag{9}$$

Because of (8),  $E[V(D, \zeta, \lambda)] = 0$ . The first-order residual is defined by replacing the true model in Equation (9) by the fitted model, i.e.,

$$R(D, \hat{\zeta}, \hat{\lambda}) = \int_D \hat{\zeta}(s) dN(s) - \int_D \hat{\zeta}(s) \hat{\lambda}(s) ds, \tag{10}$$

where we use  $\hat{\zeta}$  instead of  $\zeta$  for the reason that  $\zeta$  may include the estimated model parameters. If the fitted model is sufficiently close to the true model, then  $R(D, \hat{\zeta}, \hat{\lambda}) \approx 0$ .

Ref. [16] also illustrated how to use the above techniques to reconstruct different components in the formation of a branching-type clustering model for describing the spatiotem-

poral clustering features of earthquake occurrences, which was called the epidemic-type aftershocks sequence (ETAS) model ([18,19]). In the following, we apply similar techniques to the multiplicative modulated processes in order to reconstruct the baseline intensity function.

Let  $N$  be a modulated renewal process with occurrence times  $T_1, \dots, T_n$  and intensity function (2). Furthermore, let  $T_0 = 0, T_{n+1} = T$ , and  $X_i = T_i - T_{i-1}, i = 1, 2, \dots, n + 1$ . For  $x > 0$ , choose  $\zeta_x(t)$  as

$$\zeta_x(t) := \frac{\mathbf{1}(t < T_{N(t-)} + x) \mu(t)}{\lambda(t)} = \frac{\mathbf{1}(t < T_{N(t-)} + x)}{Z_\beta(t)}, \quad 0 \leq t \leq T, \quad x \geq 0,$$

and then (8) gives

$$E \left[ \int_0^T \zeta(s) dN(s) \right] = E \left[ \int_0^T \mathbf{1}(s < T_{N(t-)} + x) \mu(s) ds \right]. \tag{11}$$

As the integral of the left hand side of (11) is actually equivalent to summing the integrand over the jump times  $T_1, \dots, T_n$  of the process, we rewrite it as

$$\int_0^T \frac{\mathbf{1}(s < T_{N(t-)} + x)}{Z_\beta(s)} dN(s) = \sum_{i=1}^n \frac{\mathbf{1}(X_i < x)}{Z_\beta(T_i)}. \tag{12}$$

Additionally, we partition the integral of the right hand side of (11) into  $n + 1$  parts:

$$\int_0^T \mathbf{1}(s < T_{N(t-)} + x) \mu(s) ds = \sum_{i=1}^{n+1} \int_{T_{i-1}}^{T_i} \mathbf{1}(s < T_{i-1} + x) \mu(s) ds. \tag{13}$$

Now, dropping the expectation sign and inserting (12) and (13) into (11), we get an approximation

$$\sum_{i=1}^n \frac{\mathbf{1}(T_i - T_{i-1} < x)}{Z_\beta(T_i)} \approx \sum_{i=1}^{n+1} [\mathbf{1}(X_i < x) U_0(X_i) + \mathbf{1}(X_i \geq x) U_0(x)] \tag{14}$$

where

$$U_0(x) = \int_0^x \mu_0(s) ds$$

is the cumulative baseline intensity function. Rearranging (14), we obtain a recursive estimate of  $U_0(x)$ :

$$\hat{U}_0(x) = \frac{\sum_{i=1}^n \mathbf{1}(X_i < x) / Z_\beta(T_i) - \sum_{i=1}^{n+1} \mathbf{1}(X_i < x) \hat{U}_0(X_i)}{\sum_{i=1}^{n+1} \mathbf{1}(X_i \geq x)}, \tag{15}$$

by setting  $U_0(0) = 0$ . It is easy to see that  $\hat{U}_0(x)$  is a right-continuous function, if we change  $<$  and  $\geq$  to  $\leq$  and  $>$  in (14), respectively, we can obtain another estimate for  $U_0(x)$ , which has the same values as  $\hat{U}_0(x)$  at almost all the locations except that it is a left-continuous function. Thus, to eliminate the ambiguity of  $\hat{U}_0(x)$  at each observed interevent time, we consider the following function,

$$\tilde{U}_0(x) = \begin{cases} 0, & x \in [0, X_{(1)}/2] \\ \omega_i(x) \hat{U}(X_{(i-1)}) + [1 - \omega_i(x)] \hat{U}(X_{(i)}), & x \in [(X_{(i-1)} + X_{(i)})/2, (X_{(i)} + X_{(i+1)})/2], \\ & i = 1, 2, \dots, n \\ \omega_n(x) \hat{U}(X_{(n-1)}) + [1 - \omega_n(x)] \hat{U}(X_{(n)}), & x \in [(X_{(n)} + X_{(n+1)})/2, \infty), \end{cases} \tag{16}$$

where  $\{X_{(i)} : i = 1, 2, \dots, n + 1\}$  is the reordering of  $\{X_i : i = 1, 2, \dots, n + 1\}$  in an ascending order,  $X_{(0)} = 0$ , and  $\omega_i(x) = [2x - (X_{(i-1)} + X_{(i)})] / (X_{(i+1)} - X_{(i-1)})$  represents the proportion of the segment cut by  $x$  from the left end of  $[(X_{(i-1)} + X_{(i)})/2, (X_{(i)} + X_{(i+1)})/2]$ .

2.2. Estimation of Parameters in the External Process

The log-likelihood of a point process equipped with conditional intensity  $\lambda(t)$  is

$$\log L = \sum_{i=1}^n \log \lambda(T_i) - \int_0^T \lambda(s) ds$$

(see [20]). When the baseline intensity is known, the maximum likelihood estimates of parameter  $\beta$  of model (2) is the maximizer of

$$\log L = \sum_{i=1}^n [\log \mu_0(X_i) + \log Z_\beta(T_i)] - \sum_{i=1}^{n+1} \int_0^{X_i} \mu_0(s) Z_\beta(s + T_{i-1}) ds. \tag{17}$$

The integral term of (17) can be rewritten by using partial integration

$$\begin{aligned} & \int_0^{X_i} \mu_0(s) Z_\beta(s + T_{i-1}) ds \\ &= \int_0^{X_i} Z_\beta(s + T_{i-1}) dU_0(s) \\ &= Z_\beta(X_i + T_{i-1}) U_0(X_i) - \int_0^{X_i} U_0(s) Z'_\beta(s + T_{i-1}) ds. \end{aligned} \tag{18}$$

Recall that we can estimate  $U_0$  by using (15) or (16) and that  $\hat{U}_0$  is a stepwise function; (18) can be approximated by

$$\int_0^{X_i} \mu_0(s) Z_\beta(s + T_{i-1}) ds \approx Z_\beta(X_i + T_{i-1}) \hat{U}_0(X_i) - \int_0^{X_i} \hat{U}_0(s) Z'_\beta(s + T_{i-1}) ds.$$

Denote  $\{X_{(k)} : k = 1, 2, \dots, n + 1\}$  in the ascending order of  $\{X_j : j = 1, 2, \dots, n + 1\}$ . The integral term in the above equation can be rewritten as

$$\int_0^{X_i} \hat{U}_0(s) Z'_\beta(s + T_{i-1}) ds = \sum_{k: X_{(k)} < X_i} \hat{U}(X_{(k)}) [Z_\beta(X_{(k+1)} + T_{i-1}) - Z_\beta(X_{(k)} + T_{i-1})]. \tag{19}$$

In summary, the likelihood can be approximated by

$$\begin{aligned} \log L \approx & \sum_{i=1}^n \left\{ \log \mu_0(X_i) + \log Z_\beta(T_i) - Z_\beta(T_i) \hat{U}_0(X_i) \right. \\ & \left. + \sum_{k: X_{(k)} < X_i} \hat{U}(X_{(k)}) [Z_\beta(X_{(k+1)} + T_{i-1}) - Z_\beta(X_{(k)} + T_{i-1})] \right\}. \end{aligned}$$

As  $\sum_{i=1}^n \log \mu_0(X_i)$  is independent of the model parameters, it can be omitted in the calculation of the maximum of the likelihood function. Similarly, the derivative of the likelihood function with respect to the model parameter can be approximated by

$$\begin{aligned} \frac{\partial \log L}{\partial \beta} \approx & \sum_{i=1}^n \left\{ \frac{1}{Z_\beta(T_i)} \frac{\partial Z_\beta(T_i)}{\partial \beta} - \frac{\partial Z_\beta(T_i)}{\partial \beta} \hat{U}_0(X_i) \right. \\ & \left. + \sum_{k: X_{(k)} < X_i} \hat{U}(X_{(k)}) \left[ \frac{\partial Z_\beta(X_{(k+1)} + T_{i-1})}{\partial \beta} - \frac{\partial Z_\beta(X_{(k)} + T_{i-1})}{\partial \beta} \right] \right\}. \end{aligned}$$

### 2.3. Estimation Algorithm

As shown in Section 2.2, when  $\mu_0(x)$  is known, we can estimate the model parameter in  $Z_\beta$  through a maximum likelihood estimation (MLE). Furthermore, when  $Z_\beta$  is given,  $\mu_0(x)$  can be estimated by taking the derivative of a smoothed version of  $\hat{U}_0(x)$  in Equation (15). However, in practice, both  $\mu_0(x)$  and  $Z_\beta$  are unknown to us. What we have is a series of occurring times,  $T_1, \dots, T_n$ , in a given observation period, say  $[0, T]$ , with or without other related information such as an external process or marks (such as sizes and location) of each event. To estimate  $\mu_0(x)$  and  $\beta$  simultaneously, we use the following iterative algorithm:

Step 1: Set  $\ell \leftarrow 0$ ,  $\hat{\mu}_0^{(\ell)}(s) = n/T$ , or  $\hat{U}_0^{(\ell)}(s) = ns/T$ , where  $n$  is the total number of event and  $[0, T]$  is the observing time window.

Step 2 (Maximization): Estimate  $\beta^{(\ell)}$  through an MLE, with  $\hat{U}_0^{(\ell)}$ , i.e.,

$$\begin{aligned} \hat{\beta}^{(\ell)} &= \arg \max_{\beta} \log L \\ &= \arg \max_{\beta} \sum_{i=1}^n \left\{ \log Z_{\beta}(T_i) - \int_0^{X_i} \mu_0^{(\ell)}(s) Z_{\beta}(s + T_{i+1}) ds \right\} \\ &\approx \arg \max_{\beta} \sum_{i=1}^n \left\{ \log Z_{\beta}(T_i) - Z_{\beta}(T_i) \tilde{U}_0^{(\ell)}(X_i) \right. \\ &\quad \left. + \sum_{k: X_{(k)} < X_i} \hat{U}^{(\ell)}(X_{(k)}) \left[ Z_{\beta}(X_{(k+1)} + T_{i-1}) - Z_{\beta}(X_{(k)} + T_{i-1}) \right] \right\}, \end{aligned} \tag{20}$$

where

$$\tilde{U}_0^{(\ell)}(x) = \begin{cases} 0, & x \in \left[ 0, \frac{X_{(1)}}{2} \right] \\ \omega_i(x) \hat{U}^{(\ell)}(X_{(i-1)}) + [1 - \omega_i(x)] \hat{U}_0^{(\ell)}(X_{(i)}), & x \in \left[ \frac{X_{(i-1)} + X_{(i)}}{2}, \frac{X_{(i)} + X_{(i+1)}}{2} \right] \\ \omega_n(x) \hat{U}_0^{(\ell)}(X_{(n-1)}) + [1 - \omega_n(x)] \hat{U}_0^{(\ell)}(X_{(n)}), & x \in \left[ \frac{X_{(n)} + X_{(n+1)}}{2}, \infty \right), \end{cases}$$

and  $\omega_i(x) = [2x - (X_{(i-1)} + X_{(i)})] / (X_{(i+1)} - X_{(i-1)})$ .

Step 3 (Expectation): Update

$$\hat{U}_0^{(\ell+1)}(x) = \frac{\sum_{i=1}^n \mathbf{1}(X_i < x) / Z_{\hat{\beta}^{(\ell)}}(T_i) - \sum_{i=1}^{n+1} \mathbf{1}(X_i < x) \hat{U}_0^{(\ell+1)}(X_i)}{\sum_{i=1}^{n+1} \mathbf{1}(X_i \geq x)}.$$

Step 4 (Rescaling): Let  $\hat{U}_0^{(\ell+1)}(x) = \hat{U}_0^{(\ell+1)}(x) / \max \hat{U}_0^{(\ell+1)}(x)$ .

Step 5: If  $\max |\hat{U}_0^{(\ell+1)} - \hat{U}_0^{(\ell)}| < \epsilon$ , where  $\epsilon$  is a small number, stop; otherwise, let  $\ell \leftarrow \ell + 1$ , and go back to Step 2.

In Step 2, the term  $\sum_{i=1}^n \log \mu_0(X_i)$  is dropped out in (20) because it is independent from  $\beta$ . Step 4 is introduced here because there is a free linear factor between function  $Z_\beta$  and  $U_0(x)$ , that is, if we divide  $Z_\beta$  by a constant  $C$  and then multiply  $C$  by  $U_0(x)$ , such a solution  $\{Z_\beta / C, C U_0(x)\}$  is equivalent to  $\{Z_\beta, U_0(x)\}$ .

### 3. Comparison with Existing Estimates

As mentioned in the introduction, it has been a tradition to construct statistical inferences of renewal processes in the manner of a survival analysis. Given the occurrence times of a renewal process,  $\{T_1, T_2, \dots, T_n\}$ , in an observation time window  $[0, T]$ , let  $T_0 = 0$  and  $X_i = T_i - T_{i-1}$ ,  $i = 1, \dots, n$ , then the new sequence  $\{X_1, X_2, \dots, X_n\}$  forms up a survival process starting with  $n$  individuals. To study the differences between our estimate and those

in survival processes, we compare our estimate of the cumulative baseline intensity with the Nelson–Aalen and Breslow estimates that were constructed based on partial likelihood (see [3,21]).

As mentioned in Section 1, after rearranging each interevent time as a survival sample that starts at time 0 and dies off after a period of the same length as the interevent time, the modulated renewal process is converted into a survival model with conditional intensity

$$\lambda(x) = \mu_0(x) \sum_{i=1}^n \mathbf{1}(X_i \geq x) Z_{\beta}(T_{i-1} + x).$$

If  $\beta$  is known, the cumulative baseline intensity  $U_0(x)$  could be estimated by the Nelson–Aalen estimator

$$A(x) = \int_0^x \frac{dN(s)}{\sum_{i=1}^n \mathbf{1}(X_i \geq s) Z_{\beta}(T_{i-1} + s)} = \sum_{X_j \leq x} \frac{1}{\sum_{i=1}^n \mathbf{1}(X_i \geq X_j) Z_{\beta}(T_{i-1} + X_j)}. \tag{21}$$

If  $\beta$  is unknown and is replaced by its estimate  $\hat{\beta}$ ,  $U_0(x)$  could be estimated by the Breslow estimator

$$B(x) = \int_0^x \frac{dN(s)}{\sum_{i=1}^n \mathbf{1}(X_i \geq s) Z_{\hat{\beta}}(T_{i-1} + s)} = \sum_{X_j \leq x} \frac{1}{\sum_{i=1}^n \mathbf{1}(X_i \geq X_j) Z_{\hat{\beta}}(T_{i-1} + X_j)} \tag{22}$$

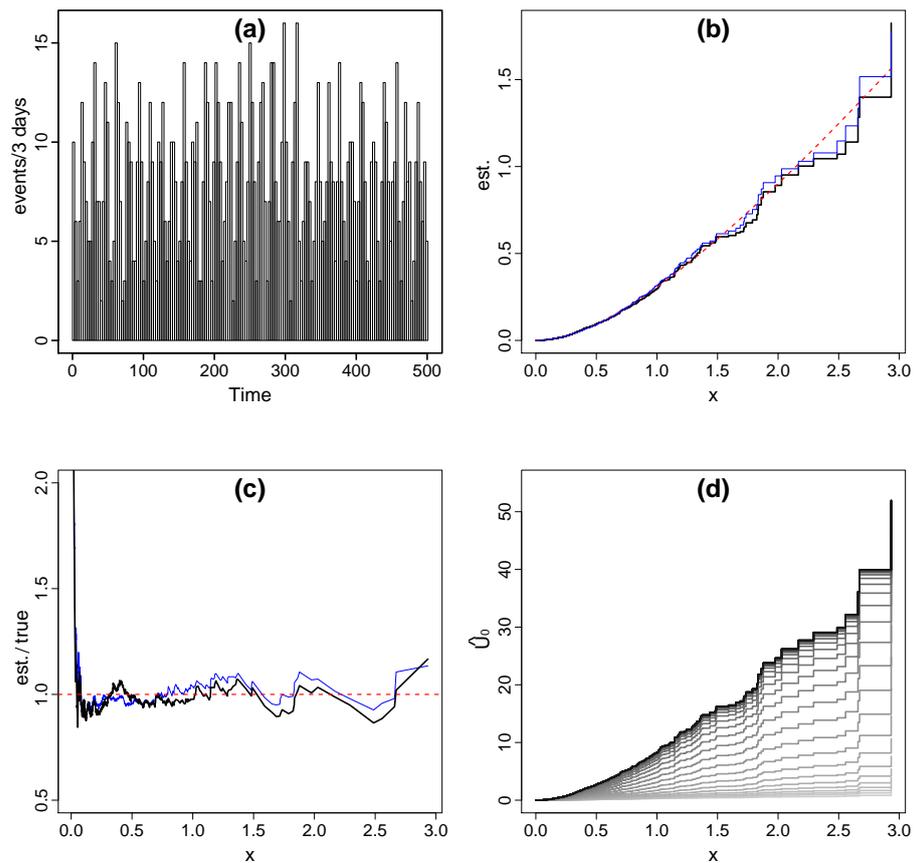
The main difference between these two estimates and ours is the inclusion of the last interval of the event times in the estimation. In our study, we assume that the last interval is censored at time  $T$ , that is, the recurrence time of the  $(n + 1)$ th observation is longer than  $X_{n+1}$ , whereas in the partial likelihood estimation, there are no censored data in the last interval. In addition, concerns regarding the validity of using the Nelson–Aalen and Breslow estimators in modulated renewal processes were raised by [22,23], as the timescale reordering, from real time  $t$  to time interval  $x$ , might invalidate the conditioning argument underlying partial likelihood. As our estimate is martingale-based, its validity naturally holds according to the asymptotic properties related to martingales.

### 4. Applications

#### 4.1. Simulation Study

In this example, we consider a modulated gamma renewal process with  $Z(t) = 2 + 30 \sin^2(0.2t + 1.5)$  and  $\mu_0(x) = x e^{-x} / \Gamma(2, x)$ , where  $\Gamma$  denotes the incomplete gamma function. We simulated this process using the thinning algorithm given in Appendix A. Figure 1a displays the number of events occurring in a 3-day interval. We first assumed that the true  $Z_{\beta}(t)$  is given, then applied (15) and (21) to the simulated data to get estimates of the cumulative baseline intensity,  $\hat{U}_0(x)$  and  $A(x)$ , as shown in Figure 1b. Figure 1c gives the ratio of the estimates to the true  $U_0(x)$ . From Figure 1b,c, we can see that the two estimates are almost identical to the true model, especially for small time intervals, and starts to diverge slightly when the size of the time interval increases.

To examine the convergence of our algorithm proposed in Section 2.3, we assumed that  $Z(t) = a + b \sin^2(\omega t + d)$ , where  $(a, b, \omega, d)$  is the parameter vector to be estimated. After some iterations, the final estimation converged to  $\hat{Z}(t) = 0.0799 + 1.0040 \sin^2(0.2000t + 1.4466)$ . The convergence steps of  $\hat{U}_0(t)$  are plotted in Figure 1d. To show the convergence steps more clearly in Figure 1d, we modified the rescaling step (Step 4) to  $\hat{U}_0^{(\ell+1)}(x) = \sqrt{\hat{b}^{(\ell+1)}} \hat{U}_0^{(\ell+1)}(x)$ . The ratio of the finally estimated  $\hat{U}_0(x)$  to the true  $U_0(x)$ ,  $\hat{U}_0(x) / U_0(x)$ , is similar to that in Figure 1c, but fluctuates around a different constant 28, which when multiplied with  $\hat{Z}(t)$ , yields a good estimate to the real values of parameters.



**Figure 1.** (a) Histogram of the simulated periodic gamma renewal process; (b) cumulative baseline intensity:  $\hat{U}_0$  (thick),  $A(x)$  (thin), and true (dashed); (c) ratio of  $\hat{U}_0(x)$  (thick) and  $A(x)$  (thin) to the true cumulative baseline intensity; (d) convergence steps of  $\hat{U}_0(x)$ . The iteration steps are indicated by different grayscales from the lightest gray for the initial step to the darkest for the final results.

#### 4.2. Real Data: Wenchuan Earthquake

In our second example, we applied our method to the aftershock data of the 2008 Wenchuan earthquake. This mainshock was measured at  $M_s$  8.0 and  $M_w$  7.9 and occurred at 14:28:01 CST (06:28:01 UTC) on Monday, 12 May 2008 in the Sichuan province of China, killing about 68,000 people. We selected the aftershocks of  $M_s \geq 4.0$  for the analysis. In addition, as aftershocks immediately after a large mainshock are usually missing due to detection and recording problems (see, e.g., [24]), we neglected the event occurring within 0.225 days (about 5.4 h) after the occurrence of the mainshocks (see also [25]. Figure 2a gives the daily occurrence numbers of earthquakes.

In seismology, the Omori–Utsu formula ([26]; see [27], for a review) is a generally accepted empirical law for describing how the frequency of aftershocks decays with time. It asserts that the occurrence rate of aftershocks can be written as an inverse power law of time, i.e.,

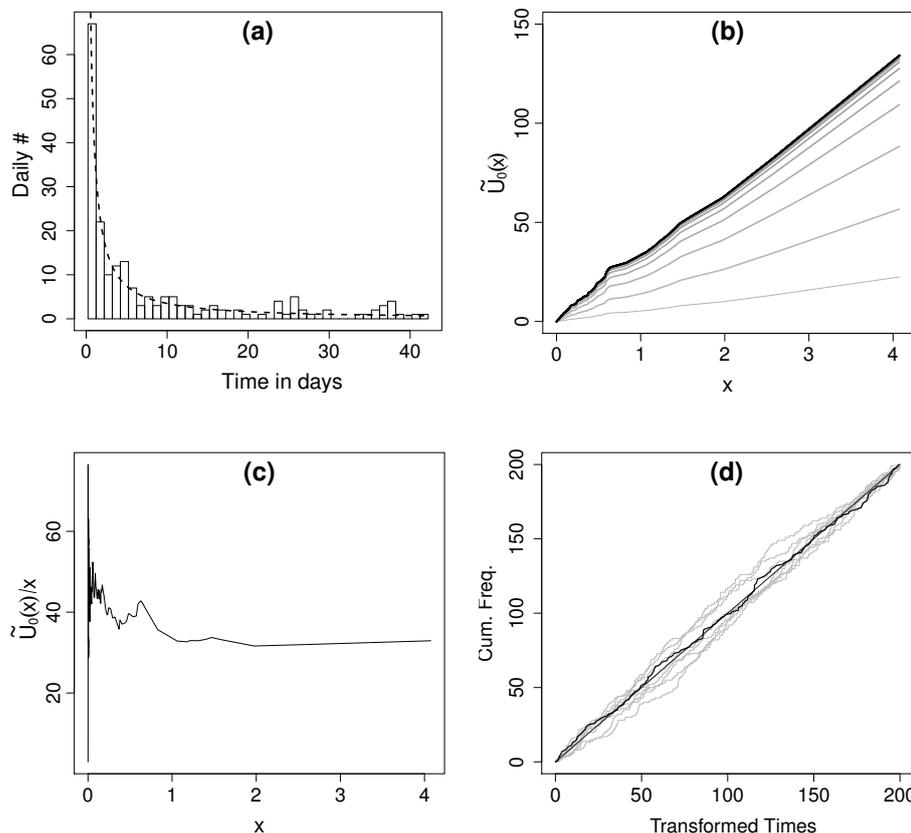
$$n(t) = \frac{\nu}{(t + c)^p}, \tag{23}$$

where  $t$  is the time since the occurrence of the mainshock,  $\nu$  is the amplitude,  $c$  is a “time-offset” parameter, and  $p$  modifies the decay rate of the aftershocks. All parameters are positive. In Figure 2a, the dashed line represents the results from fitting the Omori–Utsu formula, with  $\hat{\nu} = 46.21$ ,  $\hat{c} = 0.1449$  day, and  $\hat{p} = 1.1160$ .

It is interesting to know whether the Omori–Utsu formula can describe the aftershock activity sufficiently; that is, if we regard the inverse power decay as the external modulation from the mainshock, whether the baseline hazard function is temporally constant. If so, the interevent time between the consecutive aftershocks behaves like a Poisson process,

after the excitation effect from the mainshock is removed. We applied a modulated renewal model to the data, where the covariate process  $Z_{\beta}(t) = \nu_1 / (t + c_1)^{\beta}$  was chosen to follow the Omori–Utsu formula.

The final estimates of the parameters in the fitted modulated renewal model were  $\hat{\nu}_1 = 1.0010$ ,  $\hat{c}_1 = 0.0743$  day, and  $\hat{\beta}_1 = 0.8270$ . Figure 2b shows the convergence steps of  $\hat{U}_0(x)$ . in Figure 2c, which gives the ratio of  $\hat{U}_0(x)$  to the cumulative conditional intensity of a Poisson process of unit rate, i.e.,  $U_0^{\text{Poisson}}(x) = x$ , we can see that the estimated baseline intensity has higher values at small time intervals and decreases to a constant when the time interval is large. This implies that, even if the triggering effect of the mainshock is modeled by the Omori–Utsu formula as the external modulating process and is eliminated by using the proposed technique, the interevent times between consecutive events do not always follow a Poisson process. For longer time intervals, the constant rate of the baseline intensity indicates that the larger interevent times are exponentially distributed. However, for the shorter time intervals,  $U_0(x)/x$  is much larger than at longer time intervals, indicating that, beyond the clustering effect excited by the mainshock, there are further clusters occurring immediately after some of the aftershocks. In other words, the aftershock sequence from the Wenchuan earthquake shows a multistage triggering effect: not only the mainshock, but also the aftershocks can trigger their own aftershocks.



**Figure 2.** (a) Daily occurrence numbers of the  $M_{5.4.0+}$  Wenchuan aftershocks and the dashed line represents the intensity function from fitting the original Omori–Utsu formula to this dataset; (b) convergence steps of  $\hat{U}_0(x)$ : the iteration steps are indicated by different grayscale colors from the lightest gray for the initial step to the darkest for the final results; (c) Ratio of the final estimate  $\hat{U}_0(x)$  to  $U_0(x) = x$  in the Omori–Utsu formula modulated renewal model; (d) the cumulative numbers of events against the transformed times.

### 5. Transformed Time Analysis

When a model comparison is not presented, the technique of time-sequence transformation is a naive method to test goodness-of-fit. Notice that the following transformation of the original occurrence times of events,

$$T_i \rightarrow \tau_i = \int_0^{T_i} \lambda(s) ds \tag{24}$$

forms up a sequence  $\{\tau_i : i = 1, 2, \dots, n\}$ . If the model is true, then  $\{\tau_i : i = 1, 2, \dots, n\}$  should follow a standard Poisson process with unit rate (see, e.g., [18]). Thus, graphically, the plot of the cumulative frequencies of a fitted model against the transformed times would be close to a straight line if it fits the data well. For our study, the transformed time sequence for modulated renewal processes could be calculated based on Equation (19), i.e.,

$$\begin{aligned} \tau_0 &= 0; \\ \tau_{i+1} &\approx \tau_i + \tilde{U}_0(X_i) Z_{\hat{\beta}}(T_i) - \int_0^{X_i} \hat{U}_0(s) Z'_{\hat{\beta}}(s + T_{i-1}) ds \\ &= \tau_i + \tilde{U}_0(X_i) Z_{\hat{\beta}}(T_i) - \sum_{k: X_{(k)} < X_i} \hat{U}(X_k) [Z_{\hat{\beta}}(X_{(k+1)} + T_{i-1}) - Z_{\hat{\beta}}(X_{(k)} + T_{i-1})]. \end{aligned} \tag{25}$$

Applying the above analysis to the Wenchuan aftershocks, the cumulative number of events against the transformed time sequence for the Wenchuan aftershocks are plotted in Figure 2d. The closeness between the curve of cumulative frequencies and the straight line of the unit slope shows the goodness-of-fit of the fitted-Omori-Utsu-formula modulated renewal process for the aftershock data. To understand the variation of the transformed times caused by randomness, we also simulated 10 realizations of a Poisson process with the same number of events as the Wenchuan aftershocks and rescaled their occurrence times in the interval  $[0, n]$  (see [28], for justification of  $\int_0^T \hat{\lambda}(u) du = n$ ). The curve of the cumulative number of the transformed time sequence lies among the simulated curves, again indicating the goodness-of-fit of the fitted-Omori-Utsu-formula modulated renewal process.

### 6. Discussion

From (21) and (22), we can see that, for each  $x$ , the computational complexity of the Nelson-Aalen and Brewslow estimators are  $O(n^2)$ , since these estimators have two layers of summations over the events in the processes. Considering that  $A(x)$  or  $B(x)$  are both stepwise constant functions, which change their values at nodes of  $n$  interevent times, the total computational complexity for estimating the entire  $A(x)$  or  $B(x)$  is  $O(n^3)$ . From (15), the calculation of  $\hat{U}_0(x)$  takes a computation time of the order  $O(n)$  at one location, and  $O(n^2)$  at all locations. This is a great improvement from the Nelson-Aalen-Brewslow estimators and the method proposed by [15].

Another point is that we used the true likelihood function for estimating the model parameters in the modulating covariate processes instead of Cox’s maximizing partial likelihood estimator. In our procedure, once  $\hat{U}_0(x)$  in (20) is obtained, the computation time of  $\log L$  in (20) is also of the order of  $O(n^2)$ , the same as the partial likelihood in Cox’s estimator in (4). In other words, the proposed algorithm does not increase the computational complexity in the parametric part.

### 7. Conclusions

This study presented a semiparametric method for simultaneously estimating the baseline hazard function and the parameters of a modulated renewal process. The core of this approach was an E-M-type iterative algorithm. The E step was the estimation of the integrated baseline hazard intensity obtained through constructing a zero-mean martingale, by assuming that the parameters of the external modulating effects were known. The

M step was the MLE of the parameters in the external process, assuming the integrated baseline intensity was known. Compared to the Nelson–Aalen–Breslow estimator, the computation time of this algorithm was of the order of  $O(n^2)$  for a dataset consisting of  $n$  events, while the Nelson–Aalen–Breslow estimator took a time of order  $O(n^3)$ . The power of such a semiparametric method was illustrated through two examples: a simulated periodically modulated gamma renewal process and the aftershock sequences of the  $M_s$  8.0 Wenchuan earthquake, which occurred on 21 May 2008 in Sichuan Province, China. We also used the transformed time analysis to show the goodness-of-fit of our procedure to the Wenchuan earthquake.

**Author Contributions:** Conceptualization, J.Z.; methodology, H.-Y.S.; software and formal analysis, H.-Y.S. both authors have written, edited, and finalized the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** H.-Y. Siew was supported by the Global COE program of Meiji University for this study. Jiancang Zhuang is partially supported by JSPS Kakenhi Grant No. 19H04073. Helpful comments from two anonymous referees are gratefully acknowledged.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Algorithm for Simulating a Modulated Renewal Process

A modulated renewal process can be simulated based on the following thinning algorithm, which was designed according to Daley and Vere-Jones (2003, Algorithm 7.5.IV, Chapter 7):

Suppose that the process, which is equipped with a conditional intensity  $\lambda(t) = \mu_0(t - t_{N(t)}) Z(t)$ , has been simulated up to time  $t$ , with a record of  $\{T_1, T_2, \dots, T_k\}$ .

- Step 1. Find three positive numbers  $D$ ,  $M_1$ , and  $M_2$  such that  $M_1 \geq \mu_0(s - T_k)$  and  $M_2 \geq Z(s)$  for  $s \in (t, t + D)$ .
- Step 2. Starting from  $t$ , generate the next event according to a Poisson process with rate  $M_1 M_2$ . Suppose that its occurrence time is  $t'$ , that is to say,  $t' - t$  has an exponential distribution with a rate of  $M_1 M_2$ .
- Step 3. Stop the simulation if  $t'$  is greater than the termination time.
- Step 4. If  $t' > t + D$ , then let  $t = t + D$  and return to step 1.
- Step 5. Generate an r. v.  $U \sim U[0, 1]$ . If  $U < \frac{\mu_0(t' - T_k)}{M_1} \frac{Z(t')}{M_2}$ , accept  $t'$ , i.e., set  $T_{k+1} = t'$  and  $k = k + 1$ ; otherwise, let  $t = t'$  and return to step 1.

## References

1. Cox, D.R. *Renewal Theory*; Methuen: London, UK, 1962.
2. Cox, D.R.; Lewis, P.A.W. *The Statistical Analysis of Series of Events*; Methuen: London, UK, 1966.
3. Cox, D.R. The statistical analysis of dependencies in point processes. In *Stochastic Point Processes*; Lewis, P.A.W., Ed.; John Wiley: New York, NY, USA, 1972; pp. 55–66.
4. Nomura, S.; Ogata, Y.; Nadeau, R.M. Space-Time model for repeating earthquakes and analysis of recurrence intervals on the San Andreas Fault near Parkfield, California. *J. Geophys. Res.* **2014**, *119*, 7092–7122. [[CrossRef](#)]
5. Andersen, P.K.; Gill, R.D. Cox's regression model for counting processes: a large sample study. *Ann. Stat.* **1982**, *10*, 1100–1120. [[CrossRef](#)]
6. Lin, D.Y.; Wei, L.J.; Yang, I.; Ying, Z. Semiparametric regression for the mean and rate functions of recurrent events. *J. R. Stat. Soc. Ser.* **2000**, *62*, 711–730. [[CrossRef](#)]
7. Cox, D.R. Regression models and life-tables (with discussion). *J. R. Stat. Soc. Ser.* **1972**, *34*, 187–220.
8. Cox, D.R. Partial likelihood. *Biometrika* **1975**, *62*, 269–276. [[CrossRef](#)]
9. Nelson, W. Hazard plotting for incomplete failure data. *J. Qual. Technol.* **1969**, *1*, 27–52. [[CrossRef](#)]

10. Nelson, W. Theory and applications of hazard plotting for censored failure data. *Technometrics* **1972**, *14*, 945–965. [[CrossRef](#)]
11. Aalen, O.O. Statistical Inference for a Family of Counting Processes. Ph.D. Thesis, University of California, Berkeley, CA, USA, 1975.
12. Aalen, O.O. Nonparametric inference for a family of counting processes. *Ann. Stat.* **1978**, *6*, 701–726. [[CrossRef](#)]
13. Breslow, N.E. Discussion of the paper by D. R. Cox. *J. R. Stat. Soc. Ser.* **1972**, *34*, 216–217.
14. Tsiatis, A.A. A large sample study of Cox's regression model. *Ann. Stat.* **1981**, *9*, 93–108. [[CrossRef](#)]
15. Lin, F.; Fine, P. Pseudomartingale estimating equations for modulated renewal process models. *J. R. Stat. Soc. Ser.* **2009**, *71*, 3–23. [[CrossRef](#)]
16. Zhuang, J. Second-order residual analysis of spatiotemporal point processes and applications in model evaluation. *J. R. Stat. Soc. Ser.* **2006**, *68*, 635–653. [[CrossRef](#)]
17. Baddeley, A.; Turner, R.; Møller, J.; Hazelton, M. Residual analysis for spatial point processes (with discussion). *J. R. Stat. Soc. Ser.* **2005**, *67*, 617–666. [[CrossRef](#)]
18. Ogata, Y. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *J. Am. Stat. Assoc.* **1988**, *83*, 9–27. [[CrossRef](#)]
19. Ogata, Y. Space-Time point-process models for earthquake occurrences. *Ann. Inst. Stat. Math.* **1998**, *50*, 379–402. [[CrossRef](#)]
20. Daley, D.D.; Vere-Jones, D. *An Introduction to Theory of Point Processes*, 2nd ed.; Volume 1: Elementary Theory and Methods; Springer: New York, NY, USA, 2003.
21. Oakes, D.; Cui, L. On semiparametric inference for modulated renewal processes. *Biometrika* **1994**, *81*, 83–90. [[CrossRef](#)]
22. Berman, M. Inhomogeneous and modulated gamma processes. *Biometrika* **1981**, *68*, 143–152. [[CrossRef](#)]
23. Oakes, D. Survival times: Aspects of partial likelihood. *Int. Stat. Rev.* **1981**, *49*, 235–252. [[CrossRef](#)]
24. Kagan, Y.Y.; Jackson, D.D. New seismic gap hypothesis: Five years after. *J. Geophys. Res.* **1995**, *100*, 3943–3959. [[CrossRef](#)]
25. Zhuang, J.; Harte, D.S.; Werner, M.J.; Hainzl, S.; Zhou, S. Basic Models of Seismicity: Temporal Models. Community Online Resour. Stat. Seism. 2012. Available online: <http://www.corssa.org> (accessed on 21 June 2022). [[CrossRef](#)]
26. Utsu, T. Statistical study on the occurrence of aftershocks. *Geophys. Mag.* **1961**, *30*, 521–605.
27. Utsu, T.; Ogata, Y.; Mutsura, R.S. The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth* **1995**, *43*, 1–33. [[CrossRef](#)]
28. Schoenberg, F.P. On rescaled Poisson processes and the Brownian bridge. *Ann. Inst. Stat. Math.* **2002**, *54*, 445–457. [[CrossRef](#)]