RESEARCH ARTICLE

All that glitters is not gold: Relational events models with spurious events

Cornelius Fritz[1]* [ID], Marius Mehrl[2], Paul W. Thurner[2] and Göran Kauermann[1]

[1]Department of Statistics, LMU Munich, Munich, Germany and [2]Geschwister Scholl Institute of Political Science, LMU Munich, Munich, Germany
*Corresponding author. E-mail: cornelius.fritz@stat.uni-muenchen.de

Guest Editor: Carter Butts

Abstract

As relational event models are an increasingly popular model for studying relational structures, the reliability of large-scale event data collection becomes more and more important. Automated or human-coded events often suffer from non-negligible false-discovery rates in event identification. And most sensor data are primarily based on actors' spatial proximity for predefined time windows; hence, the observed events could relate either to a social relationship or random co-location. Both examples imply spurious events that may bias estimates and inference. We propose the Relational Event Model for Spurious Events (REMSE), an extension to existing approaches for interaction data. The model provides a flexible solution for modeling data while controlling for spurious events. Estimation of our model is carried out in an empirical Bayesian approach via data augmentation. Based on a simulation study, we investigate the properties of the estimation procedure. To demonstrate its usefulness in two distinct applications, we employ this model to combat events from the Syrian civil war and student co-location data. Results from the simulation and the applications identify the REMSE as a suitable approach to modeling relational event data in the presence of spurious events.

Keywords: data augmentation; longitudinal network analysis; relational event model; spurious events

1. Introduction

In recent years, event data have become ubiquitous in the social sciences. For instance, interpersonal structures are examined using face-to-face interactions (Elmer & Stadtfeld, 2020). At the same time, political event data are employed to study and predict the occurrence and intensity of armed conflict (Fjelde & Hultman, 2014; Blair & Sambanis, 2020; Dorff et al., 2020). Butts (2008a) introduced the Relational Event Model (REM) to study such relational event data. In comparison to standard network data of durable relations observed at specific time points, relational events describe instantaneous actions or, put differently, interactions at a fine-grained temporal resolution (Borgatti et al., 2009).

However, in some contexts there arise problems regarding the reliability of event data. While data gathered from, for example, direct observations (Tranmer et al., 2015) or parliamentary records (Malang et al., 2019) should prove unproblematic in this regard, other data collection methods may be prone to spurious events, that is events that are recorded but did not actually occur as such. For instance, data collection on face-to-face interactions relies on different types of sociometric badges (Eagle & Pentland, 2006) for which a recent study reports a false-discovery rate of the event identification of around 20% when compared to video-coded data

(Elmer et al., 2019). Political event data on armed conflict, in contrast, are generally collected via automated or human coding of news and social media reporting (Kauffmann, 2020). Spurious events may arise in this context if reports of fighting are wrong, as may be the case for propaganda reasons or due to reporters' reliance on rumors, or when fighting took place between different belligerents than those named. Such issues are especially prevalent in machine-coded conflict data where both false-positive and false-discovery rates of over 60% have been reported (King & Lowe, 2003; Jäger, 2018). However, even human-coded data suffer from this problem (Dawkins, 2021; Weidmann, 2015).

This discussion suggests that specific types of event data can include unknown quantities of spurious events, which may influence the substantive results obtained from models such as the REM (Butts, 2008a) or the Dynamic Actor-Oriented Model (Stadtfeld et al., 2017; Stadtfeld, 2012). We thus propose a Relational Events Model with Spurious Events (REMSE) as a method that allows researchers to study relational events from potentially error-prone contexts or data collections methods. Moreover, this tool can assess whether spurious events are observed under a particular model specification and, more importantly, whether they influence the substantive results. The REMSE can thus serve as a straightforward robustness check in situations where the researcher, due to their substantive knowledge, suspects that there are spurious observations and wants to investigate whether they distort their empirical results.

We take a counting process point of view where some increments of the dyadic counting processes are true events, while others may be attributed to spurious events, that is, exist due to measurement error. This decomposition results in two different intensities governing the two respective types of events. The spurious events are described by a spurious-event intensity that we specify independently of the true-event intensity of true events. We present the model under the assumption that the spurious events are purely random. Therefore, we can model the respective intensity solely as a constant term. However, more complex scenarios involving the specification of exogenous and endogenous covariates for the spurious-event intensity are also possible. In general, we are however primarily interested in studying what factors drive the intensity of true events. We model this intensity following Butts (2008a), but the methodology is extendable to other model types such as Stadtfeld et al. (2017); Vu et al. (2015); DuBois et al. (2013); Perry & Wolfe (2013) or Lerner et al. (2021).

This article is structured as follows: We begin in Section 2 by introducing our methodology. In particular, we lay out the general framework to study relational event data proposed by Butts (2008a) in Section 2.1 and introduce an extension to this framework, the REMSE, to correct for the presence of spurious events in the remainder of Section 2. Through a simulation study in Section 3, we investigate the performance of our proposed estimator when spurious events are correctly specified and when they are nonexistent. We then apply the proposed model in Section 4 to analyze fighting incidents in the Syrian civil war as well as social interaction data from a college campus. A discussion of possible implications and extensions for the analysis of events concludes the article in Section 5.

2. A Relational Event Model with Spurious Events
2.1 Modeling framework for relational events

We denote observed events in an event stream $\mathcal{E} = \{e_1, \ldots, e_M\}$ of M elements. Each object $e \in \mathcal{E}$ consists of a tuple encoding the information of an event. In particular, we denote the two actors of an event by $a(e)$ and $b(e)$ and the time of the event with $t(e)$. For simplicity of notation, we omit the argument e for $a()$ and $b()$ when no ambiguity exists and write a_m for $a(e_m)$, b_m for $b(e_m)$, and t_m for $t(e_m)$ $\forall m \in \{1, \ldots, M\}$. Stemming from our application cases, we mainly focus on undirected events in this article; hence, the events $e = (a, b, t)$ and $\tilde{e} = (b, a, t)$ are equivalent in our framework. Note however that the proposed method also generalizes to the directed case.

We denote the set of actor-tuples between which events can possibly occur by $\mathcal{R}$, where, for simplicity, we assume that $\mathcal{R}$ is time-constant.

Following Perry & Wolfe (2013) and Vu et al. (2011a), we assume that the events in $\mathcal{E}$ are generated by an inhomogeneous matrix-valued counting process

$$\mathbf{N}(t) = (N_{ab}(t)|(a, b) \in \mathcal{R}), \tag{1}$$

which, in our case, is assumed to be a matrix-valued Poisson process (see Daley & Vere-Jones 2008 for an introduction to stochastic processes). Without loss of generality, we assume that $\mathbf{N}(t)$ is observed during the temporal interval $\mathcal{T}$, starting at $t = 0$. The cells of (1) count how often all possible dyadic events have occurred between time 0 and t; hence, $\mathbf{N}(t)$ can be conceived as a standard social network adjacency matrix with integer-valued cell entries (Butts, 2008b). For instance, $N_{ab}(t)$ indicates how often actors a and b have interacted in the time interval $[0, t]$. Therefore, observing event $e = (a, b, t)$ constitutes an increase in $N_{ab}(t)$ at time point t, that is $N_{ab}(t - h) + 1 = N_{ab}(t)$ for $h \to 0$. We denote with $\boldsymbol{\lambda}(t)$ the matrix-valued intensity of process $\mathbf{N}(t)$. Based on this intensity function, we can characterize the instantaneous probability of a unit increase in a specific dimension of $\mathbf{N}(t)$ at time point t (Daley & Vere-Jones, 2008). We parameterize $\boldsymbol{\lambda}(t)$ conditional on the history of the processes, $\mathcal{H}(t)$, which may also include additional exogenous covariates. Hence, $\mathcal{H}(t) = (\mathbf{N}(u), X(u)|u < t)$, where $X(t)$ is some covariate process to be specified later. Note that we opt for a rather general characterization of Poisson processes, including stochastic intensities that explicitly depend on previous events. We define the intensity function at the tie-level:

$$\lambda_{ab}(t|\mathcal{H}(t), \vartheta) = \begin{cases} \lambda_0(t, \alpha) \exp\{\theta^\top s_{ab}(\mathcal{H}(t))\}, & \text{if } (a, b) \in \mathcal{R} \\ 0, & \text{else} \end{cases} \tag{2}$$

where $\vartheta = (\alpha^\top, \theta^\top)^\top = \text{vec}(\alpha, \theta)$ is defined with the help of a dyadic operator $\text{vec}(\cdot, \cdot)$ that stacks two vectors and $\lambda_0(t, \alpha)$ is the baseline intensity characterized by coefficients α, while the parameters θ weight the statistics computed by $s_{ab}(\mathcal{H}(t))$, which is the function of sufficient statistics. Based on $s_{ab}(\mathcal{H}(t))$, we can formulate endogenous effects, which are calculated from $(N(u)|u < t)$, exogenous variables calculated from $(X(u)|u < t)$, or a combination of the two which results in complex dependencies between the observed events. Examples of endogenous effects for undirected events include degree-related statistics like the absolute difference of the degrees of actors a and b or hyperdyadic effects, for example, investigating how triadic closure influences the observed events. In our first application case, exogenous factors include a dummy variable whether group a and b share an ethno-religious identity. Alternatively, one may incorporate continuous covariates, for example, the absolute geographic distance between group a and b.

We give graphical representations of possible endogenous effects in Figure 1 and provide their mathematical formulations together with a general summary in Appendix A. When comparing the structures in the first row with the ones in the second row in Figure 1, the respective sufficient statistic of the event indicated by the dotted line differs by one unit. Its intensity thus changes by the multiplicative factor $\exp\{\theta_{endo}\}$, where θ_{endo} is the respective parameter of the statistic if all other covariates are fixed. The interpretation of the coefficients is, therefore, closely related to the interpretation of relative risk models (Kalbfleisch & Prentice, 2002).

Previous studies propose multiple options to model the baseline intensity $\lambda_0(t)$. Vu et al. (2011a, 2011b) follow a semiparametric approach akin to the proportional hazard model by Cox (1972), while Butts (2008a) assumes a constant baseline intensity. We follow Etezadi-Amoli & Ciampi (1987) by setting $\lambda_0(t, \alpha) = \exp\{f(t, \alpha)\}$, with $f(t, \alpha)$ being a smooth function in time
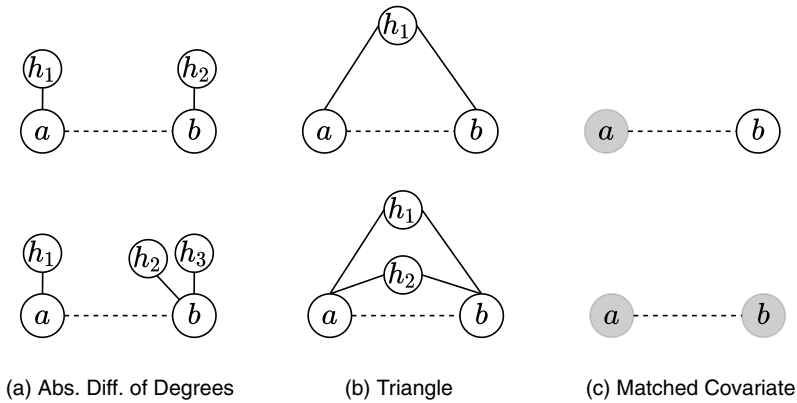
Figure 1. Graphical illustrations of endogenous and exogenous covariates. Solid lines represent past interactions, while dotted lines are possible but unrealized events. Node coloring indicates the node's value on a categorical covariate. The relative risk of the events in the second row compared to the events in the first row is $\exp\{\theta_{end}\}$ if all other covariates are fixed, where θ_{end} is the coefficient of the respective statistic of each row.

parameterized by B-splines (de Boor, 2001):

$$f(t, \alpha) = \sum_{k=1}^{K} \alpha_k B_k(t) = \alpha^\top B(t), \tag{3}$$

where $B_k(t)$ denotes the kth B-spline basis function weighted by coefficient α_k. To ensure a smooth fit of $f(t, \alpha)$, we impose a penalty (or regularization) on α which is formulated through the a priori structure

$$p(\alpha) \propto \exp\{-\gamma \alpha^\top \mathbf{S} \alpha\}, \tag{4}$$

where γ is a hyperparameter controlling the level of smoothing and $\mathbf{S}$ is a penalty matrix that penalizes the differences of coefficients corresponding to adjacent basis functions as proposed by Eilers & Marx (1996). We ensure identifiability of the smooth baseline intensity by incorporating a sum-to-zero constraint and refer to Ruppert et al. (2003) and Wood (2017) for further details on penalized spline smoothing. Given this notation, we can simplify (2):

$$\lambda_{ab}(t|\mathcal{H}(t), \vartheta) = \begin{cases} \exp\{\vartheta^\top \mathcal{X}_{ab}(\mathcal{H}(t), t)\}, & \text{if } (a, b) \in \mathcal{R} \\ 0, & \text{else,} \end{cases} \tag{5}$$

with $\mathcal{X}_{ab}(\mathcal{H}(t), t) = \text{vec}(B(t), s_{ab}(\mathcal{H}(t)))$.

2.2 Accounting for spurious relational events

Given the discussion in the introduction, we may conclude that some increments of $\mathbf{N}(t)$ are true events, while others stem from spurious events. Spurious events can occur because of coding errors during machine- or human-based data collection. To account for such erroneous data points, we introduce the REMSE.

First, we decompose the observed Poisson process into two separate matrix-valued Poisson processes, that is $\mathbf{N}(t) = \mathbf{N}_0(t) + \mathbf{N}_1(t) \, \forall \, t \in \mathcal{T}$. On the dyadic level, $N_{ab,1}(t)$ denotes the number of true events between actors a and b until t, and $N_{ab,0}(t)$ the number of events that are spurious. Assuming that $N_{ab}(t)$ is a Poisson process, we can apply the so-called thinning property, stating that two separate processes that sum up to a Poisson process are also Poisson processes (Daley & Vere-Jones 2008). A graphical illustration of the three introduced counting processes,
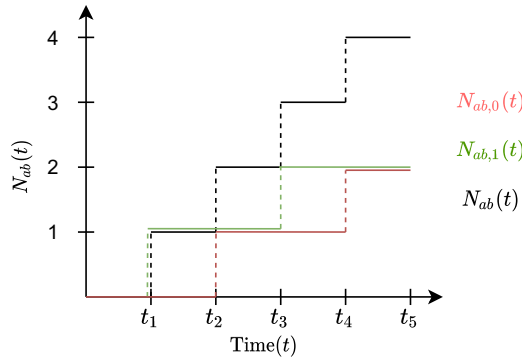
Figure 2. Graphical illustration of a possible path of the counting process of observed events ($N_{ab}(t)$) between actors a and b that encompasses spurious ($N_{ab,0}(t)$) and true events ($N_{ab,1}(t)$).

$N_{ab,0}(t)$, $N_{ab,1}(t)$, and $N_{ab}(t)$, is given in Figure 2. In this illustrative example, we observe four events at times t_1, t_2, t_3, and t_4, although only the first and third constitute true events, while the second and fourth are spurious. Therefore, the counting process $N_{ab}(t)$ jumps at all times of an event, yet $N_{ab,1}(t)$ does so only at t_1 and t_3. Conversely, $N_{ab,0}(t)$ increases at t_2 and t_4.

The counting processes $\mathbf{N}_0(t)$ and $\mathbf{N}_1(t)$ are characterized by the dyadic intensities $\lambda_{ab,0}(t|\mathcal{H}_0(t), \vartheta_0)$ and $\lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta_1)$, where we respectively denote the history of all spurious and true processes by $\mathcal{H}_0(t)$ and $\mathcal{H}_1(t)$. This can also be perceived as a competing risks setting, where events can either be caused by the true-event or spurious-event intensity (Gelfand et al., 2000). To make the estimation of θ_0 and θ_1 feasible and identifiable (Heckman & Honoré, 1989), we assume that both intensities are independent of one another, which means that their correlation is fully accounted for by the covariates. Building on the superpositioning property of Poisson processes, the specification of those two intensity functions also defines the intensity of the observed counting process $N_{ab}(t)$. In particular, $\lambda_{ab}(t|\mathcal{H}(t), \vartheta) = \lambda_{ab,0}(t|\mathcal{H}_0(t), \vartheta_0) + \lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta_1)$ holds (Daley & Vere-Jones, 2008).

The true-event intensity $\lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta_1)$ drives the counting process of true events $\mathbf{N}_1(t)$ and only depends on the history of true events. This assumption is reasonable since if erroneous events are mixed together with true events, the covariates computed for actors a and b at time t through $s_{ab}(\mathcal{H}(t))$ would be confounded and could not anymore be interpreted in any consistent manner. We specify $\lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta_1)$ in line with (2) at the dyadic level by:

$$\lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta) = \begin{cases} \exp\{\vartheta_1^\top \mathcal{X}_{ab,1}(\mathcal{H}_1(t), t)\}, & \text{if } (a, b) \in \mathcal{R} \\ 0, & \text{else.} \end{cases} \tag{6}$$

At the same time, the spurious-event intensity $\lambda_{ab,0}(t|\mathcal{H}_0(t), \vartheta_0)$ determines the type of measurement error generating spurious events. One may consider the spurious-event process as an overall noise level with a constant intensity. This leads to the following setting:

$$\lambda_{ab,0}(t|\mathcal{H}_0(t), \vartheta_0) = \begin{cases} \exp\{\alpha_0\}, & \text{if } (a, b) \in \mathcal{R} \\ 0, & \text{else.} \end{cases} \tag{7}$$

The error structure, that is, the intensity of the spurious-event process, can be made more complex, but to ensure identifiability, $\lambda_{ab,0}(t|\mathcal{H}_0(t), \vartheta_0)$ cannot depend on the same covariates as $\lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta)$. We return to the discussion of this point below and focus on model (7) for the moment.

2.3 Posterior inference via data augmentation

To draw inference on $\vartheta = \text{vec}(\vartheta_0, \vartheta_1)$, we employ an empirical Bayes approach. Specifically, we will sample from the posterior of ϑ given the observed data. Our approach is thereby comparable to the estimation of standard mixture (Diebolt & Robert, 1994) and latent competing risk models (Gelfand et al., 2000).

For our proposed method, the observed data are the event stream of all events $\mathcal{E}$ regardless of being a real or a spurious event. To adequately estimate the model formulated in Section 2, we lack information on whether a given event is spurious or not. We denote this formally as a latent indicator variable $z(e)$ for event $e \in \mathcal{E}$:

$$z(e) = \begin{cases} 1, & \text{if event } e \text{ is a true event} \\ 0, & \text{if event } e \text{ is a spurious event} \end{cases}$$

We write $z = (z(e_1), \ldots, z(e_M))$ to refer to the latent indicators of all events and use z_m to shorten $z(e_m)$. Given this notation, we can apply the data augmentation algorithm developed in Tanner & Wong (1987) to sample from the joint posterior distribution of (Z, ϑ) by iterating between the I Step (Imputation) and P Step (Posterior) defined as:

$$\text{I Step: Draw } Z^{(d)} \text{ from the posterior } p(z | \vartheta^{(d-1)}, \mathcal{E});$$

$$\text{P Step: Draw } \vartheta^{(d)} \text{ from the augmented } p(\vartheta | z^{(d)}, \mathcal{E}).$$

This iterative scheme generates a sequence that (under mild conditions) converges to draws from the joint posterior of (ϑ, Z) and is a particular case of a Gibbs' sampler. Each iteration consists of an Imputation and a Posterior step, resembling the Expectation and Maximization step from the EM algorithm (Dempster et al., 1977). Note, however, that Tanner & Wong (1987) proposed this method with multiple imputations in each I Step and a mixture of all imputed complete-data posteriors in the P Step. We follow Little & Rubin (2002) and Diebolt & Robert (1994) by performing one draw of Z and ϑ in every iteration, which is a specific case of data augmentation. As Noghrehchi et al. (2021) argue, this approach is closely related to the stochastic EM algorithm (Celeux et al., 1996). The main difference between the two approaches is that in our P Step, the current parameters are sampled from the complete-data posterior in the data augmentation algorithm and not fixed at its mean as in Celeux et al. (1996). Consequently, the data augmentation algorithm is a proper multiple imputation procedure (MI, Rubin, 1987), while the stochastic EM algorithm is improper MI (see Noghrehchi et al., 2021). We choose the data augmentation algorithm over the stochastic EM algorithm because Rubin's combination rule to get approximate standard errors can only be applied to proper MI procedures (Noghrehchi et al., 2021).

In what follows, we give details and derivations on the I and P Steps and then exploit MI to combine a relatively small number of draws from the posterior to obtain point and interval estimates for ϑ.

Imputation-step: To acquire samples from $Z = (Z_1, \ldots, Z_M)$ conditional on $\mathcal{E}$ and ϑ, we first decompose the joint density by repeatedly applying the Bayes theorem:

$$p(z | \vartheta, \mathcal{E}) = p(z_M, \ldots, z_1 | \vartheta, \mathcal{E})$$

$$= \prod_{m=1}^{M} p(z_m | z_1, \ldots, z_{m-1}, \vartheta, \mathcal{E}). \tag{8}$$

The distribution of z_m conditional on $z_1, \ldots, z_{m-1}, \vartheta$ and $\mathcal{E}$ is:

$$Z_m | z_1, \ldots, z_{m-1}, \vartheta, \mathcal{E} \sim \text{Bin}\left(1, \frac{\lambda_{a_m b_m, 1}(t_m | \mathcal{H}_1(t_m), \vartheta_1)}{\lambda_{a_m b_m, 0}(t_m | \mathcal{H}_0(t_m), \vartheta_0) + \lambda_{a_m b_m, 1}(t_m | \mathcal{H}_1(t_m), \vartheta_1)}\right). \tag{9}$$

Note that the information of $z_1, \ldots z_{m-1}$ and $\mathcal{E}$ allows us to calculate $\mathcal{H}_1(t_m)$ as well as $\mathcal{H}_0(t_m)$. By iteratively applying (9) and plugging in $\vartheta^{(d)}$ for ϑ, we can draw samples in the I Step of $Z = (Z_1, \ldots, Z_M)$ through a sequential design that sweeps once from Z_1 to Z_M. The mathematical derivation of (9) is provided in Appendix B.

Posterior step: As already stated, we assume that the true-event and spurious-event intensities are independent. Hence, the sampling from the complete-data posteriors of ϑ_0 and ϑ_1 can be carried out independently. In the ensuing section, we therefore only show how to sample from $\vartheta_1|z, \mathcal{E}$, but sampling from $\vartheta_0|z, \mathcal{E}$ is possible in the same manner. To derive this posterior, we begin by showing that the likelihood of $\mathcal{E}$ and z with parameter ϑ_1 is the likelihood of the counting process $\mathbf{N}_1(t)$, which resembles a Poisson regression. Consecutively, we state all priors to derive the desired complete-data posterior.

Given a general z sampled in the previous I Step and $\mathcal{E}$, we reconstruct a unique complete path of $\mathbf{N}_1(t)$ by setting

$$N_{ab,1}(t) = \sum_{\substack{e \in \mathcal{E}; \\ z(e)=1, t(e) \le t}} \mathbb{I}(a(e) = a, b(e) = b) \; \forall \, (a, b) \in \mathcal{R}, \; t \in \mathcal{T}, \tag{10}$$

where $\mathbb{I}(\,\cdot\,)$ is an indicator function. The corresponding likelihood of $\mathbf{N}_1(t)$ results from the property that any element-wise increments of the counting process between any times s and t with $t > s$ and arbitrary actors a and b with $(a, b) \in \mathcal{R}$ are Poisson distributed:

$$N_{ab,1}(t) - N_{ab,1}(s) \sim \text{Pois} \left(\int_s^t \lambda_{ab,1} \left(u | \mathcal{H}_1(u), \vartheta_1\right) du \right). \tag{11}$$

The integral in (11) is approximated through simple rectangular approximation between the observed event times to keep the numerical effort feasible, so that the distributional assumption simplifies to:

$$Y_{ab,1}(t_m) = N_{ab,1}(t_m) - N_{ab,1}(t_{m-1}) \sim \text{Pois} \left((t_m - t_{m-1}) \lambda_{ab,1} \left(t_m | \mathcal{H}_1(t_m), \vartheta_1\right) \right) \tag{12}$$

$$\forall \, m \in \{1, \ldots, M\} \text{ with } z_m = 1 \text{ and } (a, b) \in \mathcal{R}.$$

We specify the priors for α_1 and θ_1 separately and independent of one another. The prior for α_1 was already stated in (4). Through a restricted maximum likelihood approach, we estimate the corresponding hyperparameter γ_1 such that it maximizes the marginal likelihood of z and $\mathcal{E}$ given γ_1 (for additional information on this estimation procedure and general empirical Bayes theory for penalized splines see Wood 2011, 2020). Regarding the linear coefficients θ_1, we assume flat priors, that is $p(\theta_1) \propto k$, indicating no prior knowledge.

In the last step, we apply Wood's (2006) result that for large samples, the posterior distribution of ϑ_1 under likelihoods resulting from distributions belonging to the exponential family, such as the Poisson distribution in (12), can be approximated through:

$$\vartheta_1|z, \mathcal{E} \sim N \left(\hat{\vartheta}_1, \mathbf{V}_1 \right). \tag{13}$$

Here, $\hat{\vartheta}_1$ denotes the penalized maximum likelihood estimator resulting from (12) with the extended penalty matrix $\tilde{\mathbf{S}}_1$ defined by

$$\tilde{\mathbf{S}}_1 = \begin{bmatrix} \mathbf{S}_1 & \mathbf{O}_{p \times q} \\ \mathbf{O}_{p \times q} & \mathbf{O}_{q \times q} \end{bmatrix}$$

with $\mathbf{O}_{p \times q} \in \mathbb{R}^{p \times q}$ for $p, q \in \mathbb{N}$ being a matrix filled with zeroes and $\mathbf{S}_1$ defined in accordance with (4). For $\vartheta_1 = \text{vec}(\alpha_1, \theta_1)$, let p be the length of α_1 and q of θ_1. The penalized likelihood is

then given by:

$$\ell_p(\vartheta_1; z, \mathcal{E}) = \ell(\vartheta_1; z, \mathcal{E}) - \gamma_1 \vartheta_1^\top \tilde{\mathbf{S}}_1 \vartheta_1, \tag{14}$$

which is equivalent to a generalized additive model; hence, we refer to Wood (2017) for a thorough treatment of the computational methods needed to find $\hat{\vartheta}_1$. The variance matrix in (13) has the following structure:

$$\mathbf{V}_1 = \left(\mathcal{X}_1^\top \mathbf{W}_1 \mathcal{X}_1 + \gamma_1 \tilde{\mathbf{S}}_1 \right)^{-1}.$$

Values for γ_1 and $\hat{\vartheta}_1$ can be extracted from the estimation procedure to maximize (14) with respect to ϑ_1, while $\mathcal{X}_1 \in \mathbb{R}^{(M|\mathcal{R}|) \times (p+q)}$ is a matrix whose rows are given by $\mathcal{X}_{ab,1}(\mathcal{H}_1(t_m), t_{m-1})$ as defined in (6) for $m \in \{1, \ldots, M\}$ and $(a, b) \in \mathcal{R}$. Similarly, $\mathbf{W}_1 = \text{diag}(\lambda_{ab,1}(t|\mathcal{H}_1(t), \vartheta_1);$ $t \in \{t_1, \ldots, t_M\}, (a, b) \in \mathcal{R})$ is a diagonal matrix.

For the P Step, we now plug in $z^{(d)}$ for z in (13) to obtain $\hat{\vartheta}_1$ and $\mathbf{V}_1$ by carrying out the corresponding complete-case analysis. In the case where no spurious events exist, the complete estimation can be carried out in a single P Step. In Algorithm 1, we summarize how to generate a sequence of random variables according to the data augmentation algorithm.

Multiple imputation: One could use the data augmentation algorithm to get a large amount of samples from the joint posterior of (ϑ, Z) to calculate empirical percentiles for obtaining any types of interval estimates. However, in our case this endeavor would be very time-consuming and even infeasible. To circumvent this, Rubin (1976) proposed multiple imputation as a method to approximate the posterior mean and variance. Coincidentally, the method is especially successful when the complete-data posterior is multivariate normal as is the case in (13); thus, only a small number of draws is needed to obtain good approximations (Little & Rubin, 2002). To be specific, we apply the law of iterative expectation and variance:

$$\mathbb{E}(\vartheta|\mathcal{E}) = \mathbb{E}(\mathbb{E}(\vartheta|\mathcal{E}, z)|z) \tag{15}$$

$$\text{Var}(\vartheta|\mathcal{E}) = \mathbb{E}(\text{Var}(\vartheta|\mathcal{E}, z)|z) + \text{Var}(\mathbb{E}(\vartheta|\mathcal{E}, z)|z). \tag{16}$$

Next, we approximate (15) and (16) using a Monte Carlo quadrature with K samples from the posterior obtained via the data augmentation scheme summarized in Algorithm 1 after a burn-in period of D iterations:

$$\mathbb{E}(\vartheta|\mathcal{E}_{\text{obs}}) \approx \frac{1}{K} \sum_{k=D+1}^{D+K} \hat{\vartheta}^{(k)} = \bar{\vartheta} \tag{17}$$

$$\text{Var}(\vartheta|\mathcal{E}_{\text{obs}}) \approx \frac{1}{K} \sum_{k=D+1}^{D+K} \mathbf{V}^{(k)} + \frac{K+1}{K(K-1)} \sum_{k=D+1}^{D+K} \left(\hat{\vartheta}^{(k)} - \bar{\vartheta} \right) \left(\hat{\vartheta}^{(k)} - \bar{\vartheta} \right)^\top$$
$$= \bar{\mathbf{V}} + \bar{\mathbf{B}}, \tag{18}$$

where $\hat{\vartheta}^{(k)} = \text{vec}\left(\hat{\vartheta}_0^{(k)}, \hat{\vartheta}_1^{(k)} \right)$ encompasses the complete-data posterior means from the kth sample and $\mathbf{V}^{(k)} = \text{diag}\left(\mathbf{V}_0^{(k)}, \mathbf{V}_1^{(k)} \right)$ is composed of the corresponding variances defined in (13). We can thus construct point and interval estimates from relatively few draws of the posterior based on a multivariate normal reference distribution (Little & Rubin, 2002).

Result: $(\vartheta^{(1)}, z^{(1)}), ..., (\vartheta^{(D)}, z^{(D)})$
Set: $\vartheta^{(0)}$ to be the posterior mean of the *true* and *spurious* events, which are
   sampled randomly from the observed events with equal probability
for $d \in \{1, ..., D\}$ **do**
   Imputation Step: Sample $Z^{(d)} | \vartheta^{(d-1)}, \mathcal{E}$
   for $m \in \{1, ..., M\}$ **do**
      • Sample $Z_m^{(d)} | z_1^{(d)}, ..., z_{m-1}^{(d)}, \vartheta^{(d-1)}, \mathcal{E}$ according to (9)
      • If $z_m^{(d)} = 1$ update $s_{ab}(\mathcal{H}_1(t_m)) \ \forall \ (a, b) \in \mathcal{R}$

   end
   Posterior Step: Sample $\vartheta^{(d)} | z^{(d)}, \mathcal{E}$

      • Reconstruct $\boldsymbol{N}_0(t)$ and $\boldsymbol{N}_1(t) \ \forall \ t \in \mathcal{T}$ from $z^{(d)}$ and $\mathcal{E}$ according to (10)
      • Obtain $\hat{\vartheta}_0$ and $\boldsymbol{V_o}$ by maximizing the penalized Poisson likelihood
        given in (12) (only for $\boldsymbol{N}_0(t)$ instead of $\boldsymbol{N}_1(t)$)
      • Sample $\vartheta_0^{(d)} | z^{(d)}, \mathcal{E} \sim N(\hat{\vartheta}_0, \boldsymbol{V_o})$
      • Obtain $\hat{\vartheta}_1$ and $\boldsymbol{V_1}$ by maximizing the penalized Poisson likelihood
        given in (12)
      • Sample $\vartheta_1^{(d)} | z^{(d)}, \mathcal{E} \sim N(\hat{\vartheta}_1, \boldsymbol{V_1})$

end

Algorithm 1: Pseudo-Code to obtain D samples from the data augmentation algorithm.

3. Simulation study

We conduct a simulation study to explore the performance of the REMSE compared to a REM, which assumes no spurious events, in two different scenarios, including a regime where measurement error is correctly specified in the REMSE and one where spurious events are instead nonexistent.

Simulation design: In S=1,000 runs, we simulate event data between $n = 40$ actors under known true and spurious intensity functions in each example. For exogenous covariates, we generate categorical and continuous actor-specific covariates, transformed to the dyad level by checking for equivalence in the categorical case and computing the absolute difference for the continuous information. Generally, we simulate both counting processes $\mathbf{N_1(t)}$ and $\mathbf{N_0(t)}$ separately and stop once $|\mathcal{E}_1| = 500$.

The data-generating processes for true events is identical in each case and given by:

$$\lambda_{ab,1}(t | \mathcal{H}_1(t), \vartheta_1) = \exp\{-5 + 0.2 \cdot s_{ab, \text{Degree Abs.}}(\mathcal{H}_1(t)) \qquad \text{(DG 1-2)}$$
$$+ 0.1 \cdot s_{ab, \text{Triangle}}(\mathcal{H}_1(t)) - 0.5 \cdot s_{ab, \text{Repetition Count}}(\mathcal{H}_1(t))$$
$$+ 2 \cdot s_{ab, \text{Sum cont.}}(\mathcal{H}_1(t)) - 2 \cdot s_{ab, \text{Match cat.}}(\mathcal{H}_1(t))\},$$

where we draw the continuous exogenous covariate (cont.) from a standard Gaussian distribution and the categorical exogenous covariates (cat.) from a categorical random variable with seven possible outcomes, all with the same probability. Mathematical definition of the endogenous and exogenous statistics are given in Appendix A. In contrast, the spurious-event intensity differs across regimes to result in correctly specified (DG 1) and nonexistent (DG 2) measurement errors:

$$\lambda_{ab, 0}(t | \mathcal{H}_0(t), \vartheta_0) = \exp\{-2.5\} \qquad \text{(DG 1)}$$

Table 1. Result of the simulation study for the REMSE and REM with the two data-generating processes (DG 1, DG 2). For each DG and covariate, we note the AVE (AVerage Estimate), RMSE (Root-Mean-Squared Error), and CP (Coverage Probability). We report the average Percentage of False Events (PFE) for each DG in the last row

		REMSE			REM		
	Coefs.	AVE	RMSE	CP	AVE	RMSE	CP
DG 1 (PFE: 4.819 %)							
Intercept	−5.0	−4.936	0.337	0.944	−3.510	1.523	0.003
Degree abs	0.2	0.198	0.009	0.940	0.168	0.033	0.018
Triangle	0.1	0.101	0.019	0.949	0.094	0.019	0.932
Repetition	−0.5	−0.494	0.035	0.946	−0.385	0.120	0.039
Cov. cont.	2.0	1.982	0.101	0.951	1.557	0.453	0.003
Cov. cat.	−2.0	−1.986	0.246	0.952	−1.594	0.461	0.515
$\widehat{PFE}$ (in %)		4.835					
DG 2 (PFE 0 %)							
Intercept	−5.0	−5.040	0.286	0.954	−5.027	0.281	0.955
Degree abs	0.2	0.201	0.008	0.958	0.201	0.008	0.956
Triangle	0.1	0.102	0.018	0.955	0.102	0.018	0.952
Repetition	−0.5	−0.505	0.030	0.969	−0.504	0.030	0.964
Cov. cont.	2.0	2.009	0.087	0.952	2.006	0.086	0.948
Cov. cat.	−2.0	−2.007	0.231	0.952	−2.004	0.230	0.952
$\widehat{PFE}$ (in %)		0.0001					

$$\lambda_{ab,0}(t|\mathcal{H}_0(t), \vartheta_0) = 0 \qquad\qquad \text{(DG 2)}$$

Given these intensities, we follow DuBois et al. (2013) to sample the events.

Although the method is estimated in a Bayesian framework, we can still assess the frequentist properties of the estimates of the REMSE and REM. In particular, the average point estimate (AVE), the root-mean-squared error (RMSE) and the coverage probabilities (CP) are presented in Table 1. The AVE of a specific coefficient is the average over the posterior modes in each run:

$$\text{AVE} = \frac{1}{S} \sum_{s=1}^{S} \bar{\vartheta}_s,$$

where $\bar{\vartheta}_t$ is the posterior mean (17) of the tth simulation run. To check for the average variance of the error in each run, we further report the RMSEs of estimating the coefficient vector ϑ:

$$\text{RMSE} = \sqrt{\frac{1}{S} \sum_{s=1}^{S} (\bar{\vartheta}_s - \vartheta)^\top (\bar{\vartheta}_s - \vartheta)},$$

where ϑ is the ground truth coefficient vector defined above. Finally, we assess the adequacy of the uncertainty quantification by computing the percentage of runs in which the real parameter lies within the confidence intervals based on a multivariate normal posterior with mean and variance

Table 2. Descriptive information on the two analyzed data sets

	Conflict Event Data in 4.1	Co-location Events in 4.2
Source	ACLED	MIT Human Dynamics Lab
	(Raleigh et al., 2010)	(Madan et al., 2012)
Observational period	2017:01:01 - 2019:01:01	2008:11:01 - 2008:11:04
Number of actors	68[1]	58
Number of events	4,362	2,489[2]

given in (17) and (18). According to standard statistical theory for interval estimates, this coverage probability should be around 95% (Casella & Berger, 2001).

Results: DG 1 shows how the estimators behave if the true and false intensities are correctly specified. The results in Table 1 suggest that the REMSE can recover the coefficients from the simulation. On the other hand, strongly biased estimates are obtained in the REM, where not only the average estimates are biased, but we also observe high RMSEs and violated CP.

In the second simulation, we assess the performance of the spurious event model when it is superfluous. In particular, we investigate what happens when there are no spurious events in the data, that is, all events are real, and the intensity of $N_{ab,2}(t)$ is zero in DG 2. Unsurprisingly, the REM allows for valid and unbiased inference under this regime. But our stochastic estimation algorithm proves to be robust as for most runs, the simulated events were at some point only consisting of true events. In other words, the REMSE can detect the spurious events correctly and is unbiased if none occur in the observed data.

For both DG 1 and DG 2, the PFE estimated by the REMSE closely matches the observed one whereas the REM, by constraining it to zero, severely underestimates the PFE in DG 1. In sum, the simulation study thus offers evidence that the REMSE increases our ability to model relational event data in the presence of measurement error while being equivalent to a standard REM when spurious events do not exist in the data.

4. Application

Next, we apply the REMSE on two real-world data sets motivated by the types of event data discussed in the introduction, namely human-coded conflict events in the Syrian civil war and co-location event data generated from the Bluetooth devices of students in a university dorm[3]. Information on the data sources, observational periods, and numbers of actors and events is summarized in Table 2. Following the above presentation, we focus on modeling the true-event intensity of the REMSE and limit the spurious-event intensity to the constant term. Covariates are thus only specified for the true-event intensity. In our applications, the samples drawn according to Algorithm 1 converged to a stationary distribution within the first 30 iterations. To obtain the reported point and interval estimates via MI, we sampled 30 additional draws. Due to space restrictions, we keep our discussions of the substantive background and results of both applications comparatively short.

4.1 Conflict events in the Syrian civil war

In the first example, we model conflict events between different belligerents as driven by both exogenous covariates and endogenous network mechanisms. The exogenous covariates are

selected based on the literature on inter-rebel conflict. We thus include dummy variables indicating whether two actors share a common ethno-religious identity or receive material support by the same external sponsor as these factors have previously been found to reduce the risk of conflict (Popovic, 2018; Gade et al., 2019). Additionally, we include binary indicators of two actors being both state forces or both rebel groups as conflict may be less likely in the former but more likely in the latter case (Dorff et al., 2020).

Furthermore, we model endogenous processes in the formation of the conflict event network and consider four statistics for this purpose. First, we account for repeated fighting between two actors by including both the count of their previous interactions as well as a binary indicator of repetition, which takes the value 1 if that count is at least 1. We use this additional endogenous covariate as a conflict onset arguably comprises much more information than subsequent fighting. Second, we include the absolute difference in a and b's degree to capture whether actors with a high extent of previous activity are prone to engage each other or, instead, tend to fight less established groups to pre-empt their rise to power. Finally, we model hyperdyadic dependencies by including a triangle statistic that captures the combat network's tendency towards triadic closure.

Given that fighting should be a relatively obvious event, one may wonder why conflict event data may include spurious observations. This is because all common data collection efforts on armed conflict cannot rely on direct observation but instead use news and social media reporting. Spurious events thus occur when these sources report fighting which did not actually take place as such. In armed conflict, this can happen for multiple reasons. For instance, pro-government media may falsely report that state security forces engaged with and defeated rebel combatants to boost morale and convince audiences that the government is winning. Social media channels aligned with a specific rebel faction may similarly claim victories by its own forces or, less obviously, battles where a rival faction fought and suffered defeat against another group. In war-time settings, journalists may also be unable or unwilling to enter conflict areas and thus base their reporting on local contacts, rumors, or hear-say. Finally, spurious observations may arise here when reported fighting occurred but was attributed to the wrong belligerent faction at some point in the data collection process. From a substantive perspective, it is thus advisable to check for the influence of spurious events when analyzing these data.

Table 3 accordingly presents the results of an REM and the REMSE. Beginning with the exogenous covariates, belligerents are found to be less likely to fight each other when they share an ethno-religious identity or receive resources from the same external sponsor. In contrast, there is no support for the idea that state forces exhibit less fighting among each other than against rebels in this type of internationalized civil war, whereas different rebel groups are more likely to engage in combat against one another. Furthermore, we find evidence that endogenous processes affect conflict event incidence. The binary repetition indicator exhibits the strongest effect across all covariates, implying that two actors are more likely to fight each other if they have done so in the past. As indicated by the positive coefficient of the repetition count, the dyadic intensity further increases the more they have previously fought with one another. The absolute degree difference also exhibits a positive effect, meaning that fighting is more likely between groups with different levels of previous activity. And finally, the triangle statistic's positive coefficient suggests that even in a fighting network, triadic closure exists. This may suggest that belligerents engage in multilateral conflict, attacking the enemy of their enemy, in order to preserve the existing balance of capabilities or change it in their favor (Pischedda, 2018).

This discussion holds for both the results of REM and REMSE. Their point estimates are generally quite similar in this application, suggesting that spurious events do not substantively affect empirical results in this case. That being said, there are two noticeable differences between the two models. First, the coefficient estimates for the binary indicator of belligerents having fought before differs between the two models. In the REM, it implies a multiplicative change of $\exp\{4.911\} = 135.775$ while for the REMSE, it is estimated at $\exp\{5.059\} = 157.433$. While both models thus identify this effect to be positive and significant, it is found to be stronger when

Table 3. Combat events in the Syrian civil war: Estimated coefficients with confidence intervals noted in brackets in the first column, while the Z values are given in the second column. The results of the REMSE are given in the first two columns, while the coefficients of the REM are depicted in the last two columns. The last row reports the estimated average Percentage of False Events (PFE)

	REMSE		REM	
	Coef./CI	Z Val.	Coef./CI	Z Val.
Intercept	−10.047	−102.124	−9.944	−115.723
	[−10.24,−9.854]		[−10.112,−9.775]	
Degree abs	0.03	12.677	0.03	17.295
	[0.026,0.035]		[0.027,0.034]	
Repetition count	0.009	45.28	0.009	56.342
	[0.009,0.01]		[0.009,0.01]	
First repetition	5.052	54.321	4.911	64.946
	[4.87,5.235]		[4.763,5.059]	
Triangle	0.074	10.229	0.073	18.989
	[0.06,0.089]		[0.065,0.08]	
Match ethno-religious Id.	−0.387	−5.225	−0.393	−5.852
	[−0.532,−0.242]		[−0.525,−0.262]	
Match rebel	0.159	3.27	0.171	4.381
	[0.064,0.255]		[0.094,0.247]	
Match state force	−0.087	−0.721	−0.077	−0.723
	[−0.323,0.149]		[−0.287,0.132]	
Common sponsor 1	−0.218	−2.51	−0.227	−2.957
	[−0.388,−0.048]		[−0.378,−0.077]	
$\widehat{PFE}$ (in %)	1.1		0	

spurious events are accounted for. Second, the two models differ in how precise they deem estimates to be. This difference is clearest in their respective Z values, which are always farther away from zero for the REM than the REMSE. As a whole, these results nonetheless show that spurious events have an overall small influence on substantive results in this application. The samples from the latent indicators z also indicate that only approximately 1% of the observations, about 50 events, are on average classified as spurious events. These findings offer reassurance for the increasing use of event data to study armed conflict.

4.2 Co-location events in university housing

In our second application, we use a subset of the co-location data collected by Madan et al. (2012) to model when students within an American university dorm interact with each other. These interactions are deduced from continuous (every 6 minutes) scans of proximity via the Bluetooth signals of students' mobile phones. Madan et al. (2012) used questionnaires to collect a host

of information from the participating students. This information allows us to account for both structural and more personal exogenous predictors of social interaction. We thus include binary indicators of whether two students are in the same year of college or live on the same floor of the dorm to account for the expected homophily of social interactions (McPherson et al., 2001). In addition, we incorporate whether two actors consider each other close friends[4]. Given that the data were collected around a highly salient political event, the 2008 US presidential election, we also incorporate a dummy variable to measure whether they share the same presidential preference and a variable measuring their similarity in terms of interest in politics (Butters & Hare, 2020). In addition, we include the same endogenous network statistics here as in section 4.1. These covariates allow us to capture the intuitions that individuals tend to socialize with people that they have interacted with before, are not equally popular as they are, and they share more common friends with (Rivera et al., 2010). Compared to the first application, sources of spurious events here are more evident as students may not actually interact with but be physically close to and even face each other, for example, riding an elevator, queuing in a store, or studying in a common space.

We present the results in Table 4. Beginning with the exogenous covariates, we find that the observed interactions tend to be homophilous in that students have social encounters with people they live together with, consider their friends, and share a political opinion with. In contrast, neither a common year of college nor a similar level of political interest are found to have a statistically significant effect on student interactions. At the same time, these results indicate that the social encounters are affected by endogenous processes. Having already had a previous true event is found to be the main driver of the corresponding intensity, hence having a very strong and positive effect. Individuals who have socialized before are thus more likely to socialize again, an effect that, as indicated by the repetition count, increases with the number of previous interactions. Turning to the other endogenous covariates, the result for absolute degree difference suggests that students a and b are more likely to engage with each other if they have more different levels of previous activity, suggesting that, for example, popular individuals attract attention from less popular ones. As is usual for most social networks (Newman and Park 2003), the triangle statistic is positive, meaning that students "socialize" with the friends of their friends.

As in the first application, the REM and REMSE results presented in Table 4 are closely comparable but also show some differences. Again, the effect estimate for binary repetition, at $\exp\{2.715\} = 15.105$, is higher in the REMSE than in the REM ($\exp\{2.615\} = 13.667$) while Z values and confidence intervals obtained in the REM are substantially smaller in the REM than in the REMSE. In the co-location data too, the results are thus not driven by the presence of spurious events but accounting for these observations does affect results to some, albeit rather negligible, extent. This is the case even though the average percentage of spurious events here is comparatively high at 3%. That leaving out the corresponding 81 events yielded similar estimates may indicate that spurious events were mainly observed at the periphery of the interaction network and hardly affected the behavior in the network's core. More generally, these results may assuage concerns over sensor data reliability (see Elmer et al., 2019).

5. Discussion

In summary, this paper extends the relational event framework to handle spurious events. In doing so, it offers applied researchers analyzing instantaneous interaction data a useful tool to explicitly account for measurement errors induced by spurious events or to investigate the robustness of their results against this type of error. Our proposed method controls for one explicit measurement error, namely that induced by spurious events. The simulation study showed that our approach can detect such false events and even yield correct results if they are not present. Still, we want to accentuate that numerous other types of measurement error may be present when one

Table 4. Co-location Events in University Housing: Estimated coefficients with confidence intervals noted in brackets in the first column, while the Z values are given in the second column. The results of the REMSE are given in the first two columns, while the coefficients of the REM are depicted in the last two columns. The last row reports the estimated average Percentage of False Events (PFE)

	REMSE		REM	
	Coef./CI	Z Val.	Coef./CI	Z Val.
Intercept	−10.077	−124.905	−10.012	−139.269
	[−10.235,−9.919]		[−10.153,−9.871]	
Degree abs	0.025	5.361	0.025	6.369
	[0.016,0.035]		[0.017,0.032]	
Repetition count	0.066	27.263	0.065	29.988
	[0.061,0.07]		[0.061,0.069]	
First repetition	2.714	42.024	2.615	44.704
	[2.587,2.84]		[2.501,2.73]	
Triangle	0.049	6.597	0.049	8.109
	[0.035,0.064]		[0.037,0.061]	
Match floor	0.117	2.197	0.123	2.439
	[0.013,0.221]		[0.024,0.222]	
Match presidential pref	0.195	4.374	0.188	4.499
	[0.108,0.282]		[0.106,0.27]	
Match year	−0.003	−0.051	−0.012	−0.236
	[−0.109,0.104]		[−0.112,0.088]	
Dyad. friendship	0.157	3.145	0.15	3.15
	[0.059,0.254]		[0.057,0.243]	
Sim. interested in politics	−0.018	−0.74	−0.021	−0.917
	[−0.064,0.029]		[−0.065,0.024]	
$\widehat{PFE}$ (in %)	3.264		0	

analyses relational events, which we disregard in this article. For instance, true events may be missing. These false negatives, for example, unreported conflict events between different belligerents, are difficult to tackle because of a lack of information.

We explicitly recommend the use of the REMSE as a method for checking robustness. When substantive knowledge suggests the presence of spurious events, the REMSE can be used to assess whether REM results hold when accounting for them. Spurious events may be common in datasets which come from sensors or are coded from journalistic sources, as discussed above, and more generally seem credibly present in data that are based on secondary sources instead of direct observation. Spurious events also occur and may possibly be more influential, in data where relations are directed, the model we introduce accordingly also generalizes to directed event data. Especially for politically contentious data, where some events may be openly claimed to be false, the REMSE offers a possibility to adjudicate whether overall findings depend on such contested observations.

But also where the data content is non-political, it is recommendable to check how common and influential false observations are. We provide replication code implementing the REMSE for this purpose.

When specifying the REMSE, two aspects require caution so that identifiability is ensured. First, given we know which events are spurious, our model simplifies to a competing risk model; thus, the identifiability issues discussed in Heckman & Honoré (1989) or Tsiatis (1975) apply. For this reason, we presented our model under the assumption of independence between the true-event and spurious-event intensities. Second, the particular specification of the covariates might also affect the identifiability of the model. This may occur when one assumes complex dependencies of spurious events and exogenous covariates are unavailable, or the prior information about the coefficients is too weak. For the model specification employed in this article, this is not an issue due to the simple form of the spurious-event intensity as long as at least one exogenous or endogenous term has a nonzero effect on the true-event intensity. For more complex models, one may use multiple starting values of the data augmentation algorithm or formulate more informative priors for θ_1 and possibly θ_0.

Our latent variable methodology can also be extended beyond the approach presented here. A straightforward refinement along the lines of Stadtfeld & Block (2017) would be to include windowed effects, that is, endogenous statistics that are only using history ranging into the past for a specific duration, or exogenous covariates calculated from additional networks to the one modeled. The first modification could also be extended to separable models as proposed in Fritz et al. (2021). A relatively simplistic version of the latter type of covariate was incorporated in Section 4.2 to account for common friendships but more complex covariates are possible. This might be helpful, for instance, when we observe proximity and e-mail events between the same group of actors. Moreover, with minor adaptions, the proposed estimation methodology could handle some of the exogenous or endogenous covariates having nonlinear effects on the intensities.

Finally, the framing of the simultaneous counting processes may be modified and their number extended. To better understand the opportunities our model framework entails, it is instructive to perceive the proposed model as an extension to the latent competing risk model of Gelfand et al. (2000) with two competing risks. For time-to-event data, one could thus employ an ego-centric version[5] of our model for model-based clustering of general duration times, which could prove to be a valuable tool for medical applications. Or our proposed methodology could be conceived as a general tool to correct for additive measurement errors in count data and extend it to spatial data analysis to be used in settings described in Raleigh et al. (2010).

Competing interests. None.

Notes

1 We include all actors that, within the two-year period, participated in at least five events. To verify their existence and obtain relevant covariates, we compared them first to data collected by Gade et al. (2019) and then to various sources including news media reporting. We omitted two actors on which we could not find any information as well as actor aggregations such as "rioters" or "syrian rebels."

2 To capture new events instead of repeated observations of the same event, we omit events where the most recent previous interaction between a and b occurred less than 20 min before.

3 The full replication code is provided under the following repository: https://github.com/corneliusfritz/Replication-Code-All-that-Glitters.

4 We symmetrized the friendship network, that is, if student a nominated student b as a close friend, we assume that the relationship is reciprocated.

5 See Vu et al. (2011b) for further information on egocentric models.

References

Bauer, V., Harhoff, D., & Kauermann, G. (2021). A smooth dynamic network model for patent collaboration data. In *AStA advances in statistical analysis*.

Blair, R. A., & Sambanis, N. (2020). Forecasting civil wars: Theory and structure in an age of "big data". *Journal of Conflict Resolution*, *64*(10), 1885–1915.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892–895.

Butters, R., & Hare, C. (2020). *Polarized networks? New evidence on American voters' political discussion networks*. Political Behavior.

Butts, C. T. (2008a). A relational event framework for social action. *Sociological Methodology*, *38*(1), 155–200.

Butts, C. T. (2008b). Social network analysis: a methodological introduction. *Asian Journal of Social Psychology*, *11*(1), 13–41.

Casella, G., & Berger, R. (2001). *Statistical inference*. Belmont: Brookls/Cole.

Celeux, G., Chauveau, D., & Diebolt, J. (1996). Stochastic versions of the EM algorithm: An experimental study in the mixture case. *Journal of Statistical Computation and Simulation*, *55*(4), 287–314.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187–202.

Daley, D. J., & Vere-Jones, D. (2008). *An introduction to the theory of point processes: Volume I: Elementary Theory and Methods*. New York: Springer.

Dawkins, S. (2021). The problem of the missing dead. *Journal of Peace Research*, *58*(5), 1098–1116. doi:10.1177%2F0022343320962159.

de Boor, C. (2001). *A practical guide to splines*. New York: Springer.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Diebolt, J., & Robert, C. P. (1994). Estimation of finite mixture distributions through bayesian sampling. *Journal of the Royal Statistical Society: Series B (Methodological)*, *56*(2), 363–375.

Dorff, C., Gallop, M., & Minhas, S. (2020). Networks of violence: Predicting conflict in nigeria. *The Journal of Politics*, *82*(2), 476–493.

DuBois, C., Butts, C. T., McFarland, D., & Smyth, P. (2013). Hierarchical models for relational event sequences. *Journal of Mathematical Psychology*, *57*(6), 297–309.

Eagle, N., & Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, *10*(4), 255–268.

Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*(2), 89–102.

Elmer, T., Chaitanya, K., Purwar, P., & Stadtfeld, C. (2019). The validity of RFID badges measuring face-to-face interactions. *Behavior Research Methods*, *51*(5), 2120–2138.

Elmer, T., & Stadtfeld, C. (2020). Depressive symptoms are associated with social isolation in face-to-face interaction networks. *Scientific Reports*, *10*(1), 1009.

Etezadi-Amoli, J., & Ciampi, A. (1987). Extended hazard regression for censored survival data with covariates: A spline approximation for the baseline hazard function. *Biometrics*, *43*(1), 181–192.

Fjelde, H., & Hultman, L. (2014). Weakening the enemy: A disaggregated study of violence against civilians in Africa. *Journal of Conflict Resolution*, *58*(7), 1230–1257.

Fritz, C., Thurner, P. W., & Kauermann, G. (2021). Separable and semiparametric network-based counting processes applied to the international combat aircraft trades. *Network Science*, *9*(3), 291–311. doi:10.1017/nws.2021.9.

Gade, E. K., Hafez, M. M., & Gabbay, M. (2019). Fratricide in rebel movements: A network analysis of syrian militant infighting. *Journal of Peace Research*, *56*(3), 321–335.

Gelfand, A. E., Ghosh, S. K., Christiansen, C., Soumerai, S. B., McLaughlin, T., & J (2000). Proportional hazards models: A latent competing risk approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *49*(3), 385–397.

Heckman, J. J., & Honoré, B. E. (1989). The identifiability of the competing risks model. *Biometrika*, *76*(2), 325–330.

Jäger, K. (2018). The limits of studying networks via event data: Evidence from the ICEWS dataset. *Journal of Global Security Studies*, *3*(4), 498–511.

Kalbfleisch, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data*, Wiley Series in Probability and Statistics, Hoboken: Wiley.

Kauffmann, M. (2020). Databases for defence studies. In: D. Deschaux-Dutard (Ed.), *Research methods in defence studies* (pp. 129–150). London: Routledge.

King, G., & Lowe, W. (2003). An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design. *International Organization*, *57*(3), 617–642.

Lerner, J., Lomi, A., Mowbray, J., Rollings, N., & Tranmer, M. (2021). Dynamic network analysis of contact diaries. *Social Networks*, *66*, 224–236.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken: John Wiley & Sons, Inc.

Madan, A., Cebrian, M., Moturu, S., Farrahi, K., & Pentland, A. S. (2012). Sensing the health state of a community. *IEEE Pervasive Computing*, *11*(4), 36–45.

Malang, T., Brandenberger, L., & Leifeld, P. (2019). Networks and social influence in European legislative politics. *British Journal of Political Science*, *49*(4), 1475–1498.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444.

Newman, M. E., & Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *68*(3), 8.

Noghrehchi, F., Stoklosa, J., Penev, S., & Warton, D. I. (2021). Selecting the model for multiple imputation of missing data: just use an IC!. *Statistics in Medicine*, *40*(10), 2467–2497.

Perry, P. O., & Wolfe, P. J. (2013). Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society. Series B (Methodological)*, *75*(5), 821–849.

Pischedda, C. (2018). Wars within wars: Why windows of opportunity and vulnerability cause inter-rebel fighting in internal conflicts. *International Security*, *43*(1), 138–176.

Popovic, M. (2018). Inter-rebel alliances in the shadow of foreign sponsors. *International Interactions*, *44*(4), 749–776.

Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing ACLED: An armed conflict location and event dataset. *Journal of Peace Research*, *47*(5), 651–660.

Rivera, M. T., Soderstrom, S. B., & Uzzi, B. (2010). Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. *Annual Review of Sociology*, *36*(1), 91–115.

Robins, G., Snijders, T., Wang, P., Handcock, M., & Pattison, P. (2007). Recent developments in exponential random graph (p∗) models for social networks. *Social Networks*, *29*(2), 192–215.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Ruppert, D., Wand, M., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge: Cambridge University Press.

Stadtfeld, C. (2012). *Events in social networks: A stochastic actor-oriented framework for dynamic event processes in social networks*, Ph. D. thesis. KIT.

Stadtfeld, C., & Block, P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*(14), 318–352.

Stadtfeld, C., Hollway, J., & Block, P. (2017). Dynamic network actor models: Investigating coordination ties through time. *Sociological Methodology*, *47*(1), 1–40.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, *82*(398), 528–540.

Tranmer, M., Marcum, C. S., Morton, F. B., Croft, D. P., & de Kort, S. R. (2015). Using the relational event model (rem) to investigate the temporal dynamics of animal social networks. *Animal Behaviour*, *101*(6), 99–105.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(1), 20–22.

Vu, D., Pattison, P., & Robins, G. (2015). Relational event models for social learning in MOOCs. *Social Networks*, *43*, 121–135.

Vu, D. Q., Asuncion, A. U., Hunter, D. R., & Smyth, P. (2011a). Continuous-time regression models for longitudinal networks. In *Proceedings of the 25th Annual Conference on Neural Information Processing Systems 2011, NIPS* 2011 (pp. 2492–2500).

Vu, D. Q., Asuncion, A. U., Hunter, D. R., & Smyth, P. (2011b). Dynamic egocentric models for citation networks. In *Proceedings of the 28th International Conference on Machine Learning*. In *ICML* (pp. 857–864).

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

Weidmann, N. B. (2015). On the accuracy of media-based conflict event data. *Journal of Conflict Resolution*, *59*(6), 1129–1149.

Wood, S. N. (2006). On confidence intervals for generalized additive models based on penalized regression splines. *Australian and New Zealand Journal of Statistics*, *48*(4), 445–464.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(1), 3–36.

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. Boca Raton: CRC Press.

Wood, S. N. (2020). Inference and computation with generalized additive models and their extensions. *Test*, *29*(2), 307–339.

A. Definition of undirected network statistics

As REMs for undirected events are so far sparse in the literature, there are no standard statistics that are commonly used (one exception being Bauer et al., 2021). Thus we define all statistics based on prior substantive research (Rivera et al., 2010; Wasserman & Faust, 1994) and undirected statistics used for modeling static networks (Robins et al., 2007). Generally, nondirected statistics have to be invariant to swapping the positions of actor *a* and *b*. For the following mathematical definitions, we denote the set of all actors by $\mathcal{A}$.

For degree-related statistics, we include the absolute difference of the degrees of actors a and b:

$$s_{ab,\,Degree\,Abs.}(\mathcal{H}(t)) = |\sum_{h\in\mathcal{A}}(\mathbb{I}(N_{ah}(t^-) > 0) + \mathbb{I}(N_{ha} > 0)(t^-))$$

$$-\sum_{h\in\mathcal{A}}(\mathbb{I}(N_{bh}(t^-) > 0) + \mathbb{I}(N_{hb}(t^-) > 0))|,$$

where t^- is the point-in-time just before t. Alternatively, one might also employ other bivariate functions of the degrees as long as they are invariant to swapping a and b, such as the sum of degrees. When simultaneously using different forms of degree-related statistics, collinearities between the respective covariates might severely impede the interpretation.

To capture past dyadic behavior, one can include $N_{ah}(t^-)$ directly as a covariate. Since the first event often constitutes a more meaningful action than any further observed events between the actors a and b, we additionally include a binary covariate to indicate whether the respective actors ever interacted before, leading to the following endogenous statistics:

$$s_{ab,\,Repition\,Count}(\mathcal{H}(t)) = N_{ah}(t^-)$$

$$s_{ab,\,First\,Repition}(\mathcal{H}(t)) = \mathbb{I}(N_{ab}(t^-) > 0).$$

Hyperdyadic statistics in the undirected regime are defined as any type of triadic closure, where actor a is connected to an entity that is also connected to actor b:

$$s_{ab,\,Triangle} = \sum_{h\in\mathcal{A}}\mathbb{I}(N_{ah}(t^-) > 0)\mathbb{I}(N_{bh}(t^-) > 0)+$$

$$\mathbb{I}(N_{ha}(t^-) > 0)\mathbb{I}(N_{bh}(t^-) > 0)+$$

$$\mathbb{I}(N_{ah}(t^-) > 0)\mathbb{I}(N_{hb}(t^-) > 0)+$$

$$\mathbb{I}(N_{ha}(t^-) > 0)\mathbb{I}(N_{hb}(t^-) > 0)$$

Finally, actor-specific exogenous statistics can also be used to model the intensities introduced in this article. We denote arbitrary continuous covariates by $x_{a,\,cont}\ \forall\ a \in \mathcal{A}$. On the one hand, we may include a measure for the similarity or dissimilarity for the covariate through:

$$s_{ab,\,Sim.\,cont} = |x_{a,\,cont} - x_{b,cont}|$$

$$s_{ab,\,Dissim.\,cont} = \frac{1}{|x_{a,\,cont} - x_{b,\,cont}|}.$$

For multivariate covariates, such as location, we only need to substitute the absolute value for any given metric, for example, Euclidean distance. In other cases, it might be expected that high levels of a continuous covariable result in higher or lower intensities of an event:

$$s_{ab,\,Sum\,cont} = x_{a,\,cont} + x_{b,\,cont}.$$

Which type of statistic should be used depends on the application case and the hypotheses to be tested. Categorical covariates, that we denote by $x_{a,cat}\ \forall\ a \in \mathcal{A}$, can also be used to parameterize the intensity by checking for equivalence of two actor-specific observations of the variable:

$$s_{ab,\,Match\,cat} = \mathbb{I}(x_{a,\,cat} = x_{b,\,cat}).$$

Besides actor-specific covariates also exogenous networks or matrices, such as $x_{Network} \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{A}|}$, can also be incorporated as dyadic covariates in our framework:

$$s_{ab, Dyad. Network} = \frac{x_{Network, ab} + x_{Network, ba}}{2},$$

where $x_{Network,ab}$ is the entry of the ath row and bth column of the matrix $x_{Network}$. Extensions to time-varying networks are straightforward when perceiving changes to them as exogenous to the modeled events (Stadtfeld & Block, 2017).

B. Mathematical Derivation of (9)

For $m \in \{1, \ldots, M\}$, let $Y_{a_m b_m, 1}(t_m) = N_{a_m b_m, 1}(t_m) - N_{a_m b_m, 1}(t_{m-1})$ be the increments of the latent counting process of true events between the time points t_m and t_{m-1}, where we additionally define $t_0 = 0$ without the loss of generality. We observe $\mathcal{E}$; hence, we can reconstruct the respective increment $Y_{a_m b_m}(t_m) = N_{a_m b_m}(t_m) - N_{a_m b_m}(t_{m-1}) = Y_{a_m b_m, 0}(t_m) + Y_{a_m b_m, 1}(t_m)$, where $Y_{a_m b_m, 0}(t_m)$ is the increment of the spurious-event counting process. The second equality holds since by design the sum of increments of the processes counting the true and false events is the increment of the observed counting process, that is $N_{ab}(t) = N_{ab,0}(t) + N_{ab,1}(t)$. To sample from $Z_m | z_1, \ldots, z_{m-1}, \mathcal{E}$, note that $Z_m = Y_{a_m b_m, 1}(t_m) | Y_{a_m b_m}(t_m)$ holds. Heuristically, this means that if we know that one of the two thinned counting processes jumps at time t_m, the probability of the jump being attributed to $N_{a_m b_m, 1}(t)$ is the probability that the mth event is a true event. For the increments of the involved counting processes, we can then use the properties of the Poisson processes and the fact that the intensities are piecewise constant between event times to derive the following distributional assumptions $\forall\, m = 1, \ldots, M$:

$$Y_{a_m b_m, 0}(t_m) | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta \sim \text{Pois}\left(\delta_m \lambda_{a_m b_m, 0}(t_m | \mathcal{H}_0(t_m), \vartheta_0) \right) \tag{B1}$$

$$Y_{a_m b_m, 1}(t_m) | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta \sim \text{Pois}\left(\delta_m \lambda_{a_m b_m, 1}(t_m, \mathcal{H}_1(t_m), \vartheta_1) \right) \tag{B2}$$

$$Y_{a_m b_m}(t_m) | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta \sim \text{Pois}\Big(\delta_m \big(\lambda_{a_m b_m, 0}(t_m | \mathcal{H}_0(t_m), \vartheta_0)$$
$$+ \lambda_{a_m b_m, 1}(t_m | \mathcal{H}_1(t_m), \vartheta_1) \big) \Big), \tag{B3}$$

where we set $\delta_m = t_m - t_{m-1}$. We can now directly compute the probability of $Z_m = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta$:

$$p(z_m = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta) = p(Y_{a_m b_m, 1}(t_m) = 1 | Y_{a_m b_m}(t_m) = 1, z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)$$

$$= \frac{p(Y_{a_m b_m, 1}(t_m) = 1, Y_{a_m b_m}(t_m) = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}{p(Y_{a_m b_m}(t_m) = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}$$

$$= \frac{p(Y_{a_m b_m, 1}(t_m) = 1, Y_{a_m b_m, 2}(t_m) = 0 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}{p(Y_{a_m b_m}(t_m) = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}$$

$$= \frac{p(Y_{a_m b_m, 1}(t_m) = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}{p(Y_{a_m b_m}(t_m) = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}$$

$$\times \frac{p(Y_{a_m b_m, 2}(t_m) = 0 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}{p(Y_{a_m b_m}(t_m) = 1 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)}$$

$$= \frac{\lambda_{a_m b_m, 1}(t_m | \mathcal{H}_1(t_m, \vartheta_1))}{\lambda_{a_m b_m, 0}(t_m | \mathcal{H}_0(t_m), \vartheta_0) + \lambda_{a_m b_m, 1}(t_m | \mathcal{H}_1(t_m), \vartheta_1)}$$

In the last row we plug in (B1) and (B2) for the probabilities in the numerators and (B3) in the denominator to prove claim (9). The calculation for $p(z_m = 0 | z_1, \ldots, z_{m-1}, \mathcal{E}, \vartheta)$ is almost identical to the one shown here.