Estimation, diagnostics, and extensions of nonparametric Hawkes processes

庄 建倉 (Zhuang, Jiancang) モデリング研究系 准教授

【Abstract】

The Hawkes self-exciting model has become one of the most popular point-process models in many research areas in the natural and social sciences because of its capacity for investigating the clustering effect and positive interactions among individual events/particles. This poster discusses a general nonparametric framework for the estimation, extensions, and post-estimation diagnostics of Hawkes models.

【Introduction】

A Hawkes process consists of a series of discrete events that each stem from one of two subprocesses: the background subprocess or the clustering subprocess. The former is considered a Poisson process, which can be inhomogeneous in space and/or nonstationary in time, while the latter consists of events from the exciting effect of all the events that occurred in the past. Equivalently, each event, whether it is a background event or an excited event, triggers ('encourages') the occurrence of the future events according to some probability rules. For a spatiotemporal point process with no overlapping events (a simple temporal point process), the conditional intensity at a space-time location (t, x) is

$$\lambda(t,x) = \lim_{\delta_t \downarrow 0, ||\delta_x|| \downarrow 0} \frac{\Pr\{N([t, t+\delta_t) \times (x, x+\delta_x)) \mid \mathcal{H}_t\}}{\delta_t ||\delta_x||} = \mu(t,x) + \sum_{i: t_i < t} g(t-t_i, x-x_i)$$

where $\mathcal{H}_t$ denotes the observation history up time t but not including t.

Examples: *Linlin model* **(**Ogata, 1982**) ,** *space-time Epidemic Type Aftershock Sequence (ETAS)* (e.g., Ogata, 1998; Zhuang et al., 2004), *model for break-in burglary crimes* (Mohler et al. 2011), *epidemic forecasting for collected invasive meningococcal disease (IMD)* (Meyer et al., 2012, 2016), *model of social networks* **(**Fox, 2016; Zipkin, 2015), etc.

【Parametric estimation】

Likelihood function and MLE Given the observation data of a spatiotemporal parametric Hawkes model in a space-time window $S \times T$, the likelihood function can be written as

$$\log L(\cdot\,;\theta) = \sum_{i:(t_i,x_i)\in S\times T} \log \lambda(t_i, x_i; \theta) \;-\int_T\int_S \lambda(t,x;\theta)\,\mathrm{d}x\,\mathrm{d}t$$

Stochastic declustering

$$\varphi_j = \Pr\{\text{Event } j \text{ is a background event}\} = \frac{\mu(t_j, x_j)}{\lambda(t_j, x_j)}$$

$$\rho_{ij} = \Pr\{\text{Event } j \text{ is triggered by event } i\} = \frac{g(t_j - t_i, x_j - x_i)}{\lambda(t_j, x_j)}$$

Expectation-maximization algorithm

Treating the whole process as a missing data problem, the complete likelihood for the whole process is

$$\log L_{cmplt}(\cdot\,;\theta)$$

$$= \sum_{j=1}^{N} I(\eta_j = 0)\log\mu(t_i, x_j) + \sum_{j=1}^{N} I(\eta_j = i)\log g(t_j - t_i, x_j - x_i) - \int_T\int_S \lambda(t,x;\theta)\,\mathrm{d}x\,\mathrm{d}t$$

E-Step: For each step k, calculate $\varphi_j^{(k)}$ and $\rho_{ij}^{(k)}$ for $j = 1, 2, \cdots, n$ and $i = 1, 2, \cdots, j-1$.

M-Step: Maximize the expected log-likelihood:

$$\log L_{cmplt}^{(k+1)}(\cdot\,;\theta)$$

$$= \sum_{j=1}^{N} \varphi_j^{(k)} \log\mu(t_i, x_j) + \sum_{j=1}^{N} \rho_{ij}^{(k)} \log g(t_j - t_i, x_j - x_i) - \int_T\int_S \lambda(t,x;\theta)\,\mathrm{d}x\,\mathrm{d}t$$

to obtain the model parameters.

【Non-parametric estimation】

The nonparametric background rate μ can be estimated using a weighted kernel functions:

$$\hat{\mu}(t,x) = \sum \varphi_j\, Z_{\mathrm{x}}(x - x_j; h_{\mathrm{x}})\, Z_{\mathrm{t}}(t - t_j; h_{\mathrm{t}}),$$

where $Z_{\mathrm{t}}(\cdot, h_{\mathrm{t}})$ and $Z_{\mathrm{x}}(\cdot, h_{\mathrm{x}})$ are, respectively, the temporal and spatial kernels with bandwidths h_{t} and h_{x}.

The nonparametric triggering term is estimated by

$$\hat{\boldsymbol{g}}\,(t,x) = \frac{\sum_{ij}\rho_{ij}\, I(|t_j - t_i - t| < \delta_t)\, I(|x_j - x_i - x| < \delta_x)}{4\,\delta_t\,\delta_x \sum_{ij}\rho_{ij}}$$

where the denominator is for normalizing purposes.

Algorithm To avoid positive feedback, relaxation coefficients, ν and A, are introduced. We estimate

$$\lambda(t,x) = \nu\mu(t,x) + A \sum_{i:t_i<t} g(t - t_i, x - x_i)$$

through following steps.

(1) Stochastic declustering (Expectation): Calculate the background probability and triggering probabilities.

(2) Reconstruction (Maximization I). Estimate the nonparametric functions in the model using nonparametric methods such as kernel functions.

(3) Parametrization (Maximization II). Use the MLE method or EM algorithm to estimate the parameters in the parametric functions and the relaxation coefficients for the nonparametric functions.

【Extending and improving model formulations】

Point-source ETAS model to finite-source ETAS model

In the space-time ETAS model, all of the earthquake events are regarded as a point in space-time-magnitude domain. In fact, the rupture of each earthquake has a spatial extension on the earthquake fault. Adopting point source yields biased results. Guo, Zhuang and others (2015, 2017) developed a finite-source ETAS model to incorporate the spatial extensions of their ruptures. Each earthquake rupture consists of many small patches, and each patch triggers its own aftershocks isotopically and independently as a usual mainshock.

Spatial response function is separable from temporal response.

Point source: $f(x)$

Finite source:

$$f(x;\, S_i) = \int_{S_i} f(x-u)\,\tau_i(u)du$$

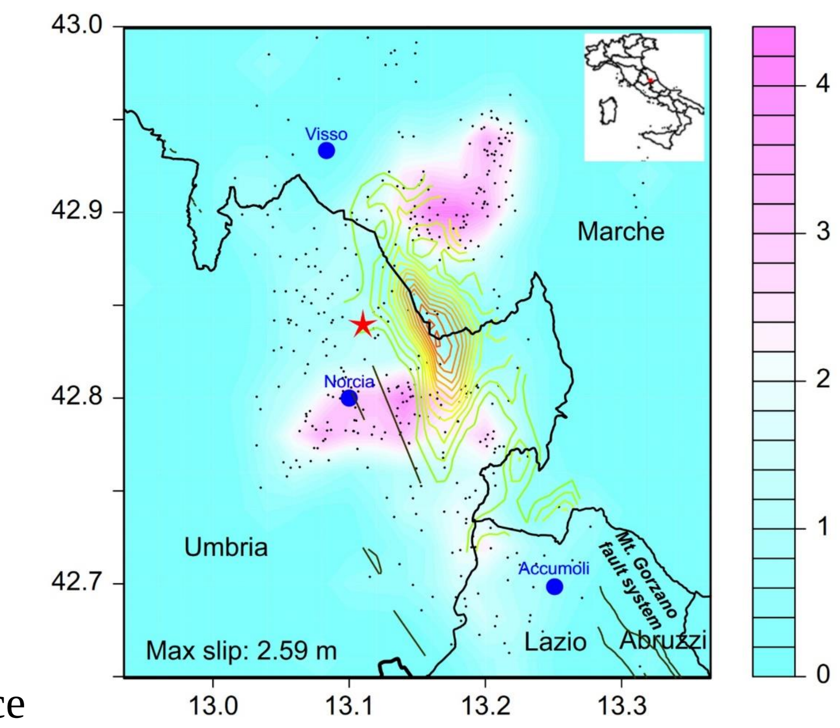
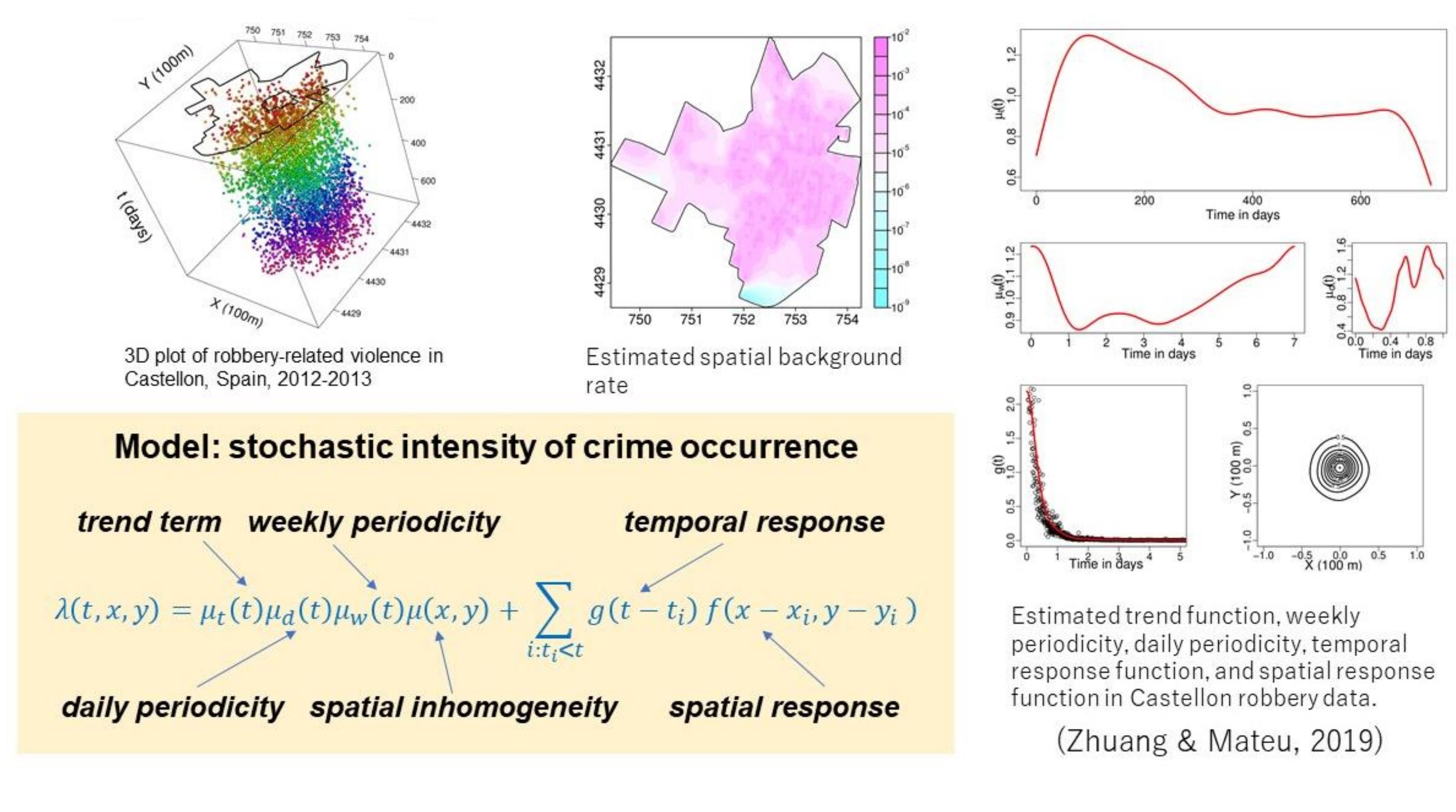


Figure 1. (c.f. Zhuang, 2019) Comparison between the pattern of the productivity of direct offspring along the rupture areas (contour images) inferred by the finite-source ETAS model and coseismic slip (contour lines) for the Norcia earthquake (2016-10-30, Mw 6.5).

Complexity in the background rate: Crime modelling

Past studies using crime data based on point processes do not consider the periodic components in the background rate. Zhuang and Mateu (2019) analyzed the robbery data in Castellon, Spain.



【Conclusion】

The process of using the Hawkes process to investigate the causal encouraging correlation among discrete events introduced in this study can be divided into 4 steps:

1. **Model design**. Design the model according to the features of the observation data, specifically the particular mathematical form of the Hawkes model by using the available empirical knowledge of the studied process.

2. **Estimation design**. Design the estimation according to the types of model formation, use the MLE method or the EM algorithm to estimate parametric model, and use stochastic reconstruction or Equation (2) to reconstruct the nonparametric components.

3. **Improvement**. Improve the estimation using kernel estimates or the Bayesian method.

4. **Diagnosing the new model**. The reconstruction method can be naturally used as a diagnostic tool to check whether it is possible to improve the model or not.