Simulation and Estimation of Lévy Random Fields

Katja Ickstadt
Department of Statistics
University of Dortmund
D-44221 Dortmund, Germany
ickstadt@statistik.uni-dortmund.de

Robert L. Wolpert
Institute of Statistics and Decision Sciences, Duke University
211c Old Chem, Box 90251
Durham, NC 27708-0251, USA
wolpert@isds.duke.edu

1. Introduction

Lévy processes, such as the Poisson, gamma, and one-sided stable processes, are useful for building hierarchical models for a wide range of applications in spatial statistics and in time series. Examples in spatial statistics include a class of Poisson/gamma random field models, in which the Poisson intensity is given by a kernel mixture of a gamma random field (Wolpert and Ickstadt, 1998a). In Best, Ickstadt and Wolpert (2000) these models are generalized to a class of point process regression models useful for applications in spatial epidemiology.

The analysis of hierarchical process models is typically based on Markov chain Monte Carlo methods which rely on the generation of stochastic processes or random fields. Here we present an efficient numerical algorithm, the so-called Inverse Lévy Measure (ILM) algorithm, for simulating Lévy processes and fields.

We also discuss an application, remote sensing of forest heights in Queensland, Australia, by airborne laser sensors, in which we seek to identify which Lévy process is generating the data we observe.

2. The Inverse Lévy Measure algorithm

Scale and dilation mixtures $Y_t \equiv \sum u_j X_{\nu_j t}$ of a Poisson process X_t have the following characteristic function:

$$\begin{aligned} E[e^{i\omega Y_t}] &= e^{\sum_j (e^{i\omega u_j} - 1)\nu_j t} \\ &= e^{\iint_{\mathbb{R}_+ \times (0,t]} (e^{i\omega u} - 1)\nu(du\, ds)} \qquad \star \end{aligned}$$

generalizing that of a Poisson process, i.e., $E[e^{i\omega X_t}] = e^{(e^{i\omega}-1)t}$.

The Lévy Khinchine formula states that every increasing, stationary stochastic process with independent increments is of form $\star$ with Lévy measure $\nu(du\,ds)$ on $\mathbb{R}_+ \times \mathbb{R}_+$. An example is the gamma process with characteristic function $E[e^{i\omega Y_t}] = (1 - i\omega\beta)^{-\alpha t}$ and Lévy measure $\nu(du\,ds) = \alpha u^{-1} e^{-\beta u}\,du\,ds$.

This suggests a representation of the process $Y(ds)$ as $Y(ds) \equiv \int_{\mathbb{R}_+} u\,H(du\,ds)$, i.e., as a Poisson process in one higher dimension ($\mathbb{R}_+ \times \mathbb{R}_+$) with mean measure $E[H(du\,ds)] = \nu(du\,ds)$. This representation is the key idea for the following ILM algorithm, which generates a Lévy process $\Lambda(ds)$ on a space $\mathcal{S}$.

Let $\Pi(ds)$ be the uniform measure on a one- or two-dimensional space $\mathcal{S}$, $\nu(du\,ds) = \nu(u,s)\,du\,\Pi(ds)$ the Lévy measure and $\tau(u,s) = \int_u^\infty \nu(x,s)\,dx$.

- Generate independent draws σ_m from $\Pi(ds)$.
- Generate the event times of a standard Poisson process by sampling independent standard exponential random variables T_i and set $\tau_m \equiv \sum_{i \le m} T_i$.
- Set $\upsilon_m \equiv \inf[u \ge 0 : \tau(u,\sigma_m) \le \tau_m]$ by setting υ_m to the inverse Lévy measure $\tau(\cdot,\sigma_m)^{-1}(\tau_m)$.
- Set $\Lambda(ds) \equiv \sum_{m \le M} \upsilon_m \delta_{\sigma_m}(ds)$.

The algorithm can be generalized in various ways: The space $\mathcal{S}$ can be of arbitrary dimension or a discrete set; $\Pi(ds)$ can be any probability measure on $\mathcal{S}$, not necessarily the uniform measure; and the algorithm is also applicable for generating inhomogeneous Lévy processes.

A detailed description of the ILM algorithm, the underlying theory, examples, and Splus code for generating Lévy processes and random fields are given in Wolpert and Ickstadt (1998b). For the one-dimensional gamma process with uniform measure for $\Pi(ds)$ the algorithm was introduced by Bondesson (1982).

3. The data

The data are part of a large data set about a plantation site for gumpie messmate 128km north northwest of Brisbane, Queensland, Australia. In order to assess the biomass of the area, a helicopter flies over certain paths of the site at about 70kph at an altitude of 100m and sends laser pulses straight down (about 175 per second). Laser heights are measured and are used to estimate tree heights. These in turn can be taken to estimate the biomass.

In the following we investigate one such flight path of length 280m. Figure 1 shows the laser heights (dots) and the imputed ground level (line). The ground level is estimated using a moving minimum based on 30 measurements. Subtracting the ground heights from the actual

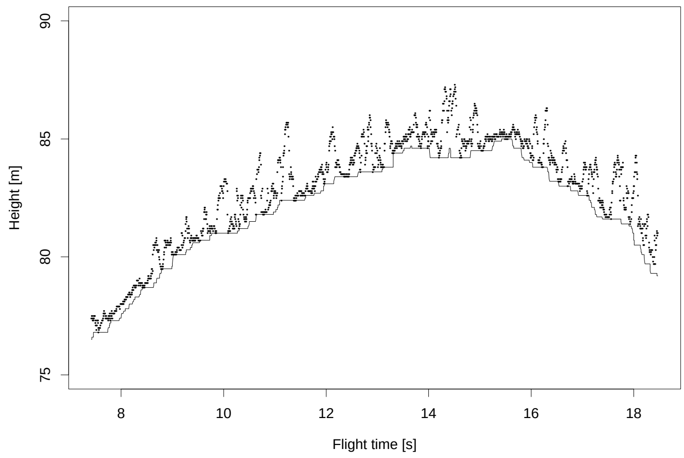

Figure 1: Laser heights (dots) and imputed ground level (line)

measurements leads to estimated values of the laser reflections or tree heights; a histogram of these heights with values up to four meters is given in figure 2.

4. A first result

As a first goal we intend to describe the tree heights by a Lévy process. For the given data set, a (homogeneous) gamma process based on a gamma distribution $Ga(\alpha, \beta)$ with mean α/β and variance α/β^2 is a suitable model. Assuming independence between measurements, maximum likelihood estimation of the parameters α and β yields $\hat{\alpha} = 1.53$ and $\hat{\beta} = 1.88$. The corresponding gamma density is added to the histogram of the laser reflections in figure 2 (solid line). Thus, the tree heights identified by the helicopter can be modelled by a Lévy process with 175 jumps per second, each jump height following a $Ga(1.53, 1.88)$ distribution.

5. Conclusion and future work

The ILM algorithm is a flexible tool for generating inhomogeneous Lévy processes in arbitrary dimensions. It is useful for simulation as well as for inference based on Lévy process models.

The analysis presented is a first step towards a Bayesian posterior analysis of the biomass in the plantation site using Markov chain Monte Carlo methods based on the ILM algorithm.

Acknowledgement

The authors would like to thank Robert Denham, Queensland University of Technology, Australia, and Christian Witte, Department of Natural Resources, Queensland, Australia, for providing the data set and for helpful discussions.

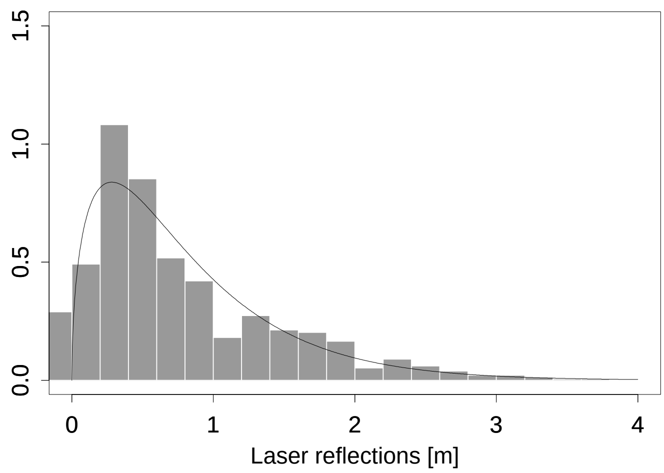

Figure 2: Histogram of the laser reflections (tree heights) and $Ga(1.53, 1.88)$ density

REFERENCES

Best, N.G., Ickstadt, K., and Wolpert, R.L. (2000). Spatial Poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American Statistical Association*, **95**, 1076–1088.

Bondesson, L. (1982). On simulation from infinitely divisible distributions. *Advances in Applied Probability*, **14**, 855–869.

Wolpert, R.L. and Ickstadt, K. (1998a). Poisson/gamma random field models for spatial statistics. *Biometrika*, **85**, 251–267.

Wolpert, R.L. and Ickstadt, K. (1998b). Simulation of Lévy random fields. In Dey, D.; Müller, P. and Sinha, D., eds.: *Practical Nonparametric and Semiparametric Bayesian Statistics*, volume 133 of Lecture Notes in Statistics, 227–242, Springer-Verlag, New York.