POINT PROCESS, SPATIAL-TEMPORAL

1. Introduction.

A spatial-temporal point process (also called space-time or spatio-temporal point process) is a random collection of points, where each point represents the time and location of an event. Examples of events include incidence of disease, sightings or births of a species, or the occurrences of fires, earthquakes, lightning strikes, tsunamis, or volcanic eruptions. Typically the spatial locations are recorded in three spatial coordinates, e.g. longitude, latitude, and height or depth, though sometimes only one or two spatial coordinates are available or of interest. Figure 1 is an illustration of a realization of a spatial-temporal point process with one spatial coordinate depicted. Figure 2 displays some point process data consisting of micro-earthquake origin times and epicenters in Parkfield, California, between 1988 and 1995, recorded by the High-Resolution Seismographic station Network (Nadeau et al., 1994). Figure 3 displays the centroids of wildfires occurring between 1876–1996 in Los Angeles County, California, recorded by the Los Angeles County Department of Public Works (times of the events not shown).

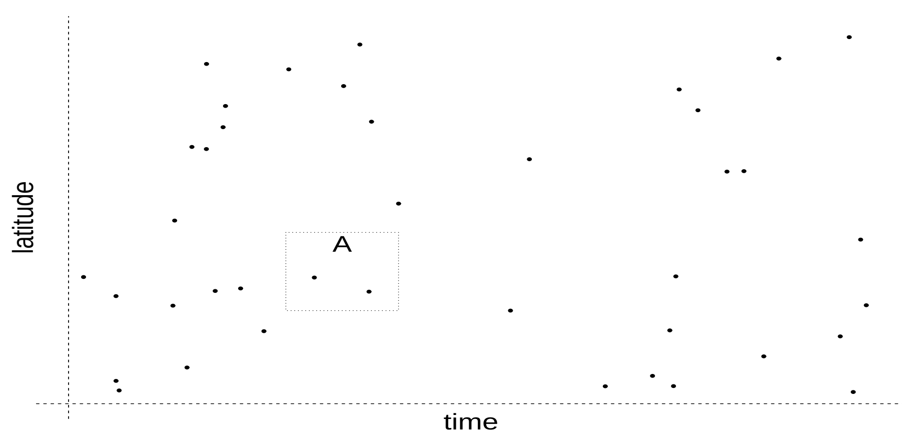

Figure 1: Spatial-temporal point process

2. Characterizations.

A spatial-temporal point process N is mathematically defined as a random measure on a region $S \subseteq \mathbf{R} \times \mathbf{R}^3$ of space-time, taking values in the

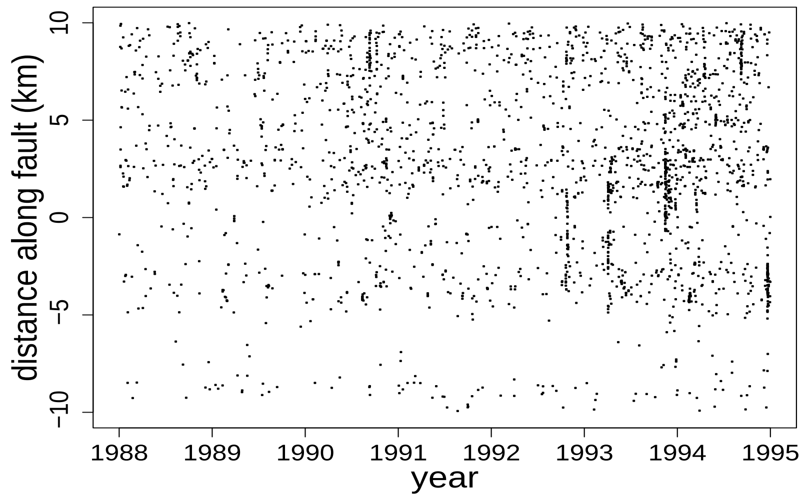

Figure 2: Epicenters and times of Parkfield microearthquakes, 1988–1995

non-negative integers $\mathbf{Z}^+$ (or infinity). In this framework the measure $N(A)$ represents the number of points falling in the subset A of S. For the set A in Figure 1, for example, the value of $N(A)$ is 2. Attention is typically restricted to points in some time interval $[T_0, T_1]$, and to processes with only a finite number of points in any compact subset of S.

Traditionally the points of a point process are thought to be indistinguishable, other than by their times and locations. Often, however, there is other important information to be stored along with each point. For example, one may wish to analyze a list of points in time and space where a member of a certain species was observed, along with the size or age of the organism, or alternatively a catalog of arrival times and locations of hurricanes along with the amounts of damage attributed to each. Such processes may be viewed as *marked* spatial-temporal point processes, i.e. random collections of points, where each point has associated with it a further random variable called a mark.

Much of the theory of spatial-temporal point processes carries over from that of spatial point processes. However, the temporal aspect enables a nat-

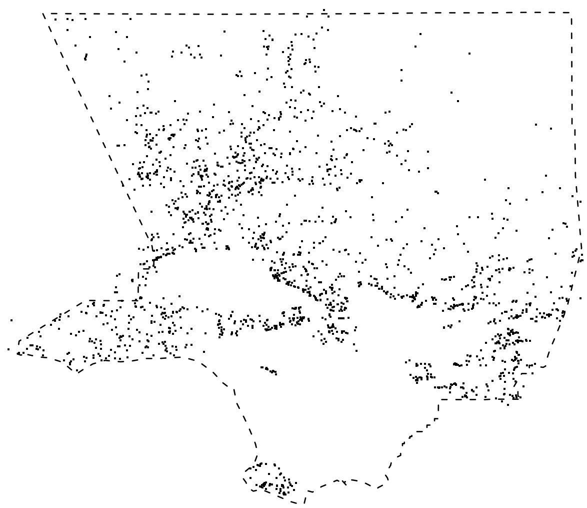

Figure 3: Centroids of recorded Los Angeles County wildfires, 1878–1996

ural ordering of the points that does not generally exist for spatial processes. Indeed, it may often be convenient to view a spatial-temporal point process as a purely temporal point process, with spatial marks associated with each point. Sometimes investigating the purely temporal (or purely spatial) behavior of the resulting marginalized point process is of interest.

The spatial region of interest is often a rectangular portion of $\mathbf{R}^2$ or $\mathbf{R}^3$, but not always. For the data in Figure 2, for example, the focus is on just one spatial coordinate, and in Figure 3 the region of interest is Los Angeles County, which has an irregular boundary. Cases where the points are spatially distributed in a sphere or an ellipse are investigated by Brillinger (1997) and Brillinger (2000). When the domain of possible spatial coordinates is discrete (e.g. a lattice) rather than continuous, it may be convenient to view the spatial-temporal point process as a sequence $\{N_i\}$ of temporal point processes which may interact with one another. For example, one may view the occurrences of cars on a highway as such a collection, where N_i represents observations of cars in lane i.

Any analytic spatial-temporal point process is uniquely characterized by its associated conditional rate process λ (Fishman and Snyder, 1976). $\lambda(t, x, y, z)$ may be thought of as the frequency with which events are expected to occur around a particular location (t, x, y, z) in space-time, conditional on the prior history H_t of the point process up to time t. Note that in the statistical literature (e.g. Daley and Vere-Jones, 1988; Karr, 1991), λ is more commonly referred to as the conditional intensity rather than conditional rate. However, the term intensity is also used in various environmental sciences, e.g. in describing the size or destructiveness of an earthquake, so to avoid confusion, the term rate may be preferred.

Formally, the conditional rate $\lambda(t, x, y, z)$ associated with a spatial-temporal point process N may be defined as the limiting conditional expectation

$$\lim_{\Delta t, \Delta x, \Delta y, \Delta z \downarrow 0} \frac{E[N\{(t, t+\Delta t) \times (x, x+\Delta x) \times (y, y+\Delta y) \times (z, z+\Delta z)\}|H_t]}{\Delta t \Delta x \Delta y \Delta z},$$

provided the limit exists. Some authors instead define λ as

$$\lim_{\Delta t, \Delta x, \Delta y, \Delta z \downarrow 0} \frac{P[N\{(t, t+\Delta t) \times (x, x+\Delta x) \times (y, y+\Delta y) \times (z, z+\Delta z)\} > 0|H_t]}{\Delta t \Delta x \Delta y \Delta z}.$$

For orderly point processes (processes where $\lim_{|A| \downarrow \emptyset} P\{N(A) > 1\}/|A| = 0$ for interval A), the two definitions are equivalent. λ is a predictable process whose integral, C (called the compensator), is such that $N - C$ is a martingale. There are different forms of conditioning corresponding to different types of martingales; see Kallenberg (1983), Merzbach and Nualart (1986), or Schoenberg (1997).

3. Models.

The behavior of a spatial-temporal point process N is typically modelled by specifying a functional form for $\lambda(t, x, y, z)$, which represents the infinitesimal expected rate of events at time t and location (x, y, z), given all the observations up to time t. Although λ may be estimated nonparametrically (Diggle 1985; Guttorp and Thompson, 1990; Vere-Jones, 1992), it is more common to estimate λ via a parametric model.

In general, $\lambda(t, x, y, z)$ depends not only on t, x, y, z but also on the times and locations of preceding events. When N is a Poisson process, however, λ is deterministic; i.e. $\lambda(t, x, y, z)$ depends only on t, x, y, and z. The simplest model is the stationary Poisson, where the conditional rate is constant:

$\lambda(t, x, y, z) = \alpha$, for all t, x, y, z. In the case of modeling environmental disturbances, this model incorporates the idea that the risk of an event is the same at all times and locations, regardless of where and how frequently such disturbances have occurred previously. Processes that display substantial spatial heterogeneity, such as earthquake epicenters, are sometimes modelled as stationary in time but not space.

Stationary spatial-temporal point processes are sometimes described by the second order parameter measure $\rho(t', x', y', z')$ which measures the covariance between the numbers of points in spatial-temporal regions A and B, where region B is A shifted by (t', x', y', z'). For example, Kagan and Vere-Jones (1996) explore models for ρ in describing spatial-temporal patterns of earthquake hypocenters and times. For a self-exciting (equivalently clustered or underdispersed) point process, the function ρ is positive for small values of t', x', y', and z'; N is self-correcting (equivalently inhibitory or overdispersed) if instead the covariance is negative. Thus the occurrence of points in a self-exciting point process is associated with other points occurring nearby in space-time, whereas in a self-correcting process, the points have an inhibitory effect.

Self-exciting point process models are often used in epidemiology and seismology to model events that are clustered together in time and space. A commonly used form for such models is a spatial-temporal generalization of the Hawkes model, where $\lambda(t, x, y, z)$ may be written as

$$\mu(t, x, y, z) + \int_{T_0}^{t} \int_{x'} \int_{y'} \int_{z'} \nu(t - t', x - x', y - y', z - z') dN(t', x', y', z').$$

The functions μ and ν represent the deterministic background rate and clustering density, respectively. Often μ is modelled as merely a function of the spatial coordinates (x, y, z), and may be estimated non-parametrically as in Ogata (1998). When observed marks m associated with each point are posited to affect the rate at which future points accumulate, this information is typically incorporated into the function ν, i.e.

$$\begin{aligned} \lambda(t, x, y, z) &= \\ \mu(t, x, y, z) &+ \int_{T_0}^{t} \int_{x'} \int_{y'} \int_{z'} \nu(t - t', x - x', y - y', z - z', m) dN(t', x', y', z', m). \end{aligned}$$

A variety of forms have been given for the clustering density ν. For instance, in modeling seismological data with two spatial parameters (x and y) and a mark (m) indicating magnitude, Musmeci and Vere-Jones (1992) introduced explicit forms for ν, including the diffusion-type model

$$\nu(t, x, y, m) = \frac{C}{2\pi\sigma_x\sigma_y t} \exp\left\{\alpha m - \beta t - (x^2/\sigma_x^2 + y^2/\sigma_y^2)/2t\right\}.$$

Ogata (1998) investigated the case where

$$\nu(t, x, y, m) = \frac{K_0 \exp\{\alpha(m - m_o)\}}{(t + c)^p (x^2 + y^2 + d)^q},$$

as well as a variety of other models. Several other forms for ν were suggested by Rathbun (1993), Kagan (1991); see Ogata (1998) for a review.

Sometimes λ is modelled as a product of marginal conditional intensities

$$\lambda(t, x, y, z) = \lambda_1(t)\lambda_2(x, y, z),$$

or even

$$\lambda(t, x, y, z) = \lambda_1(t)\lambda_2(x)\lambda_2(y)\lambda_4(z).$$

These forms embody the notion that the temporal behavior of the process is independent of the spatial behavior, and in the latter case that furthermore the behavior along each of the spatial coordinates can be seen as independent. In such cases each of the functions λ_i may be estimated individually; see e.g. Rathbun (1993) or Schoenberg (1997). Occasionally one subdivides the spatial region into a finite number of subregions and fits temporal point process models to the data within each subregion. In such a case the conditional intensity may be written

$$\lambda(t, x, y, z) = \sum_i \lambda_1(t) 1_i(x, y, z),$$

where 1_i are indicator functions denoting the relevant region. An example is in Zheng and Vere-Jones (1994). Introduction of interactions between different subregions are incorporated into this model by Lu, Harte, and Bebbington (1999).

For further remarks on modeling and examples see Vere-Jones and Thomson (1984) and Snyder and Miller (1991).

4. Estimation and inference

The parameter vector θ for a model with conditional rate $\lambda(t, x, y, z; \theta)$ is usually estimated by maximizing the log-likelihood function

$$\begin{aligned} L(\theta) &= \int_{T_0}^{T_1} \int_x \int_y \int_z \log\{\lambda(t, x, y, z; \theta)\} dN(t, x, y, z) \\ &- \int_{T_0}^{T_1} \int_x \int_y \int_z \lambda(t, x, y, z; \theta) dz dy dx dt. \end{aligned}$$

Asymptotic properties of the maximum likelihood estimator $\hat{\theta}$ have been derived under various conditions, along with formulas for standard errors; see e.g. Rathbun and Cressie (1994). Alternatively, simulations may be useful for obtaining approximate standard errors and for other types of inference.

The estimated conditional rate $\lambda(t, x, y, z; \hat{\theta})$ can be used directly for prediction and risk assessment. See Fishman and Snyder (1976) and Brillinger (1982), for example.

Spatial-temporal point processes may be evaluated via residual analysis, as described in Schoenberg (1997). One typically selects a spatial coordinate and rescales the point process in that direction. If the z-coordinate is chosen, for example, then each point (t_i, x_i, y_i, z_i) of the observed point process is moved to a new point $(t_i, x_i, y_i, \int_{z_0}^{z_i} \lambda(t_i, x_i, y_i, z; \hat{\theta}) dz)$, where z_0 is the lower boundary in the z-direction of the spatial region being considered. The resulting rescaled process is stationary Poisson if and only if the model is correctly specified (Schoenberg, 1999). Hence a useful method for assessing the fit of a point process model is to examine whether the rescaled point process looks like a Poisson process with unit rate. Several tests exist for this purpose, see e.g. Ripley (1979) or Heinrich (1991).

References.

Brillinger, D.R. (1982). Seismic risk assessment: some statistical aspects. *Earthq. Predict. Res.* **1**, 183–195.

Brillinger, D.R. (1997). A particle migrating randomly on a sphere. *J. Theoretical Probability* **10**, 429–443.

Brillinger, D.R. (2000). Some examples of random process environmental data analysis. in: *Handbook of Statistics*, Vol. 18 (eds. C. R. Rao and P. K. Sen). Amsterdam, North-Holland, pp. 33–56.

Daley, D., and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes.* Springer, Berlin.

Diggle, P. (1985). A kernel method for smoothing point process data *Appl. Stat.* **34**, 138-147.

Fishman, P. M. and Snyder, D. L. (1976). The statistical analysis of space-time point processes. IEEE Trans. Inf. Theory IT-22(3), 257-274.

Guttorp, P., and Thompson, M. (1990). Nonparametric estimation of intensities for sampled counting processes. *J. Royal Stat. Soc. B* **52**, 157-173.

Hawkes, A., and Adamopoulos, L. (1973). Cluster models for earthquakes – regional comparisons. *Bull. Int. Statist. Inst.* **45** (3), 454-461.

Heinrich, L. (1991). Goodness-of-fit tests for the second moment function of a stationary multidimensional Poisson process. *Statistics* **22**, 245-278.

Kagan, Y.Y. (1991). Likelihood analysis of earthquake catalogs. *J. Geophys. Res.* **106**, 135–148.

Kagan, Y. and Vere-Jones, D. (1996). Problems in the modelling and statistical analysis of earthquakes, in: *Lecture Notes in Statistics (Athens Conference on Applied Probability and Time Series Analysis)*, **114**, C.C. Heyde, Yu.V. Prorohov, R. Pyke, and S.T. Rachev, eds., New York, Springer, pp. 398–425.

Kallenberg, O. (1983). *Random Measures, 3rd ed.* Akademie-Verlag, Berlin.

Karr, A. (1991). *Point Processes and Their Statistical Inference, second ed.* Dekker, NY.

Lu, C., Harte, D., Bebbington, M. (1999). A linked stress release model for historical Japanese earthquakes: coupling among major seismic regions. *Earth, Planets, Space*, **51** (9), 907-916.

Merzbach, E., and Nualart, D. (1986). A characterization of the spatial Poisson process and changing time. *Ann. Probab.* **14**, 1380-1390.

Musmeci, F. and Vere-Jones, D. (1992). A space-time clustering model for historical earthquakes, *Ann. Inst. Statist. Math.* **44**, 1-11.

Nadeau, R., Antolik, M., Johnson, P.A., Foxall, W., and McEvilly, T.V. (1994). Seismological studies at Parkfield III: microearthquake clusters in the study of fault-zone dynamics. *Bull. Seis. Soc. Amer.* **84**(2), 247–263.

Ogata, Y. (1998). Space-time point process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50**(2), 379–402.

Rathbun, S.L. (1993). Modeling marked spatio-temporal point patterns. *Bull. Int. Statist. Inst.* **55**, Book 2, 379–396.

Rathbun, S., and Cressie, N. (1994). Asymptotic properties of estimators for the parameters of spatial inhomogeneous Poisson point processes. *Adv. Appl. Prob.* **26**, 122-154.

Ripley, B. (1979). Tests of 'randomness' for spatial point patterns. *J. Royal Stat. Soc. B* **41**, 368-374.

Schoenberg, F. (1997). *Assessment of Multi-dimensional Point Processes.* Ph.D. Thesis, University of California, Berkeley.

Schoenberg, F. (1999). Transforming spatial point processes into Poisson processes. *Stoch. Proc. Appl.* **81**, 155-164.

Snyder, D. L. and Miller, M. I. (1991). *Random Point Processes in Time and Space.* Wiley, New York.

Vere-Jones, D. (1992). Statistical methods for the description and display of earthquake catalogs. In *Statistics in the Environmental and Earth Sciences*, A. Walden and P. Guttorp, eds. Edward Arnold, London, pp. 220-246.

Vere-Jones, D. and Thomson, P. J. (1984). Some aspects of space-time modelling, in *Proc. Twelfth International Biometrics Conference*, Tokyo, pp. 265–275

Zheng, X. and Vere-Jones, D. (1994). Further applications of the stochastic stress release model to historical earthquake data. *Tectonophysics* **229**, 101–121.

Frederic Paik Schoenberg
David R. Brillinger
Peter Guttorp