Environ Ecol Stat (2009) 16:197–223
DOI 10.1007/s10651-007-0085-1

Point patterns of forest fire locations

Rolf Turner

Received: 1 February 2006 / Revised: 1 January 2007 / Published online: 30 January 2008

Abstract In this paper I demonstrate some of the techniques for the analysis of spatial point patterns that have become available due to recent developments in point process modelling software. These developments permit convenient exploratory data analysis, model fitting, and model assessment. Efficient model fitting, in particular, makes possible the testing of statistical hypotheses of genuine interest, even when interaction between points is present, via Monte Carlo methods. The discussion of these techniques is conducted jointly with and in the context of some preliminary analyses of a collection of data sets which are of considerable interest in their own right. These data sets (which were kindly provided to me by the New Brunswick Department of Natural Resources) consist of the complete records of wildfires which occurred in New Brunswick during the years 1987 through 2003. In treating these data sets I deal with data-cleaning problems, methods of exploratory data analysis, means of detecting interaction, fitting of statistical models, and residual analysis and diagnostics. In addition to demonstrating modelling techniques, I include a discussion on the nature of statistical models for point patterns. This is given with a view to providing an understanding of why, in particular, the Strauss model fails as a model for interpoint attraction and how it has been modified to overcome this difficulty. All actual modelling of the New Brunswick fire data is done only with the intent of illustrating techniques. No substantive conclusions are or can be drawn at this stage. Realistic modelling of these data sets would require incorporation of covariate information which I do not so far have available.

R. Turner (✉)
Department of Mathematics and Statistics, University of New Brunswick, Fredericton, Canada
e-mail: rolf@math.unb.ca

Keywords Covariates · Critical envelope · Geyer model · K-function · Kernel smoothing · Inhomogeneous K-function · Intensity · Interaction · Observation window · Papangelou conditional intensity function · Strauss model

1 Introduction

Phenomena encountered in the natural sciences often manifest themselves in the form of point patterns in two-dimensional space (the plane). Data observed in respect of such phenomena consist of, or include, the locations of the objects of study—plants, animals at a given instant, particles in a matrix, mineral deposits, nests, cells, and so on. Sometimes it is the case that aspects of the manner in which these points are arranged contain useful information or give some insight which helps to characterize the objects of study. This information, if present at all, is of necessity subtle and difficult to detect and even more difficult to extract and quantify.

The collection of forest fire locations (in a specified region over a specified time period) constitutes a planar point pattern. There has recently been increasing interest in applying the ideas of point pattern analysis to forest fire locations. Evidence of such interest includes the Fields workshop on "Forest Fires and Point Processes", 24–28 May 2005, and the BIRS workshop on "Forests, Fires, and Stochastic Modelling", 6–11 May 2006. Papers in this area include Podur et al. (2003); Wotton and Martell (2005); Preisler et al. (2004). Somewhat related papers include Chen and McAneney (2004); Cunningham and Martell (1973).

Podur et al. (2003) study patterns of lightning caused fires in Ontario for the years 1976–1998, with emphasis on a rectangular window in northwestern Ontario. They use the K-function (square root transformed to form the "L" function) to assess clustering, and kernel density smoothing to provide graphical depictions of the clustering. They conclude that there is clustering at distances of approximately 2.5° (latitude/longitude; roughly 200 km. in the area under study) and significant regularity at scales larger than 6°.

Interestingly their results show (p. 15) "no obvious relationship between lightning strike density and lightning-fire occurrence over a fire season, nor do they show any obvious relationship between the sum of DSR over the season and fire occurrence density." (The quantity DSR is "Daily Severity Rating", a composite measure of fire risk.) They do discern a relationship between the intensity of lightning caused fires and lightning strikes for days on which DMC ("duff moisture code", another measure of fire risk) exceeded 20. They also briefly consider the impact of topography and human population density on lightning caused fire occurrence. (Population density could conceivably have an impact via reporting or discovery rate.)

Wotton and Martell (2005) develop and fit a logistic model to estimate the probability of lightning fire occurrences in terms of a number of meteorological variables and indices relating to fire hazard (such as FFMC, i.e., fine fuel moisture code, and DMC). They work in terms of a discretized structure in which the study region (Ontario) is subdivided into 20×20 km. grid cells.

Preisler et al. (2004) use a spatio-temporal logistic model to estimate the probability of fire occurrence, again on the basis of weather variables and fire danger indices. They too use discretization, subdividing time × space 1 day × 1 × 1 km. voxels. They analyze the probabilities of "all" fires and of "large" fires (the latter being fires whose burn area exceeds 40.5 ha).

Chen and McAneney (2004) make use of the K-function (L function) in studying the pattern of homes destroyed when bush fires penetrate into urban areas in Australia. Cunningham and Martell (1973) study a model for the number of human caused fire occurrences in a given district, using FFMC as a predictor. They do not however consider the fires in terms of spatial pattern.

Except in the simplest cases (Poisson processes) the points of a spatial pattern will be mutually stochastically dependent. This dependence can be extremely complicated. To some extent each point must be considered in relation to every other point. This situation contrasts sharply with one-dimensional point processes in which the dimension is usually considered to be time, whereby points may be considered only in respect of their *past history*. Considering every point in relation to every other point is, to put it mildly, difficult. Theoretical models are fundamentally intractable from an analytical point of view. The only real exceptions to this are the Poisson models referred to above in which the dependence between points is in fact absent.

The simplest of all possible models is the constant intensity Poisson process, frequently referred to as the model of "complete spatial randomness" (CSR). Heretofore much statistical inference for spatial point patterns in the forestry context has been focussed on the often untenable null hypothesis of CSR. Such testing is "of limited scientific interest in itself" (Diggle 2003, p. 12; see also Baddeley et al. 2005) and appears to be done mainly because it is possible, or because "everybody else does it" rather than because it provides any real insight into the data. Usually the nature of the phenomenon under study, or a casual glance at a plot of the data, make it obvious that CSR is not a realistic option.

What is usually of more interest is the detection of trends in the intensity of fire locations, and determination of how (or whether) such trends are influenced by covariates. These covariates might include vegetation, other descriptors of terrain (such as elevation and slope), and proximity to concentrations of human population or to concomitants of human activity such as roads and railroads.

Interaction between points may in general be of some interest in its own right, but in the forest fire context the questions that might be asked about interaction would probably be very subtle and also very difficult to answer. More important is the impact of the presence of interaction on statistical inference concerning trends and their dependence upon covariates. If the process generating the pattern is inhomogeneous Poisson, then models may be fitted by means of maximum likelihood, and tests for dependence on covariates may reasonably be conducted via likelihood ratio tests. However, if there is interaction between the points these tests may be seriously misleading.

It is therefore important to be able to assess interaction in point patterns and to be able to model such interaction so that Monte Carlo testing for covariate dependence can be carried out effectively. A fundamental requirement for addressing such questions is a collection of plausible models for interactions. Describing interaction between forest fires is likely to be a challenging problem, and it is probable that new models for interaction may be required. However, formulating such models is consummately difficult, and is moreover fraught with theoretical peril.

A perfectly plausible model might easily turn out to be ill-defined. The primary example of this phenomenon is the Strauss model (Strauss 1975) which was originally proposed as a model for clustering but which turns out to be well-defined only when it takes the form of a model for repulsion or inhibition. It is important to understand the difficulties which arise here, and these will be discussed in Sect. 4.

In the following section (Sect. 2). I describe an interesting collection of forest fire patterns and my efforts to massage them into a form amenable to analysis. In Sect. 3. I discuss some exploratory data analysis including the formation of a naive estimate of an underlying spatial trend. Most readers will be familiar with the use of Ripley's K-function for investigating the nature of the interpoint interaction. When a spatial trend is present the use of the basic K-function is inappropriate. I describe a relatively recent innovation, the inhomogeneous K-function, which goes some way to alleviating this problem.

Section 4, as previously foreshadowed, contains a discussion of the nature of models for planar point processes. Two ideas (one of which appears to be efficacious, and one not) are suggested for modifying the Strauss process to accommodate modelling attraction between points. In Sect. 5. I give examples of model fitting, including an attempt to fit the "efficacious" model for clustering to one of the New Brunswick data sets, and the results of applying some diagnostic tools to the fits. Finally in Sect. 6. I suggest some alternative modelling strategies and summarize what has so far been learned.

2 The New Brunswick forest fire patterns

In this section I describe an interesting collection of data sets, kindly provided to me by the New Brunswick Department of Natural Resources. These data sets comprise complete records of all fires in New Brunswick which fell under the responsibility of this Department during the years 1987 to 2003, with—currently—the omission of 1988. (The 1988 data set lacks the specifications of the locations of the fires in terms of latitude and longitude. The locations are specified only in terms of a now disused grid system. Converting the grid locations to latitude and longitude would have to done by hand and promises to be extremely tedious.)

The data sets are relatively large and rather complex. They were communicated to me in the form of spreadsheets, one for each year. Each spreadsheet had 97 columns and between 286 and 654 rows (i.e., fires). The data required a good deal of reorganization, cleaning, and massaging to make them amenable to analysis from the point of view of spatial point patterns. There were inevitably many missing values and anomalous entries which had to be dealt with.

Much more remains to be done in the way of data cleaning and organizing. This task includes dealing with the large quantity of auxiliary information—95 variables in addition to the location (latitude and longitude) variables. Decisions will have to be made as to which of these variables bear some promise of providing fruitful avenues of investigation. Some of the variables (e.g., discovery date, size attained) are of obvious importance. Many others (e.g., "payroll", "fire boss") are obviously of no immediate interest. A substantial number require judicious consideration. Even more remains to be done in the way of acquiring appropriate information on a variety of covariates which might influence the prevalence of forest fires. So far (sad to say) the acquisition of such information remains on my to-do list.

The first step in organizing the data was to read the data sets from the spreadsheets into R (R Development Core Team 2005) data frames. This process was greatly facilitated by the `gdata` package written by Gregory R. Warnes which is available from the contributed packages repository of CRAN (2006). The next step was to extract locations of the fires. These were given in latitude and longitude, as in the following display:

```
  latitude longitude
1     4554      6731
2     4600      6732
3     4618      6544
4     4612      6730
5     4455      6660
6     4519      6707
.        .         .
.        .         .
.        .         .
```

Note that these numbers consist of degrees and minutes, juxtaposed, and that the longitude is in degrees *west* (causing the data to plot "from right to left" unless some adjustment is made). These coordinates had first to be translated into degrees expressed as decimal fractions—e.g., "4554" becomes 45.9—and then translated into New Brunswick stereographic projection coordinates (Thomson et al. 1977). These latter coordinates are designed to produce minimal distortion when a map is projected from an ellipsoidal surface (e.g., the Earth) onto the plane. They are also the coordinates in which the observation window was made available to me.

It cannot be overemphasized that in order to specify a point pattern properly it is necessary to specify an observation window. It is important to remember that there is information in "where the data aren't" as well as in where they are, and in order to know where the data aren't, it is necessary to know through what window one is looking at the data.

Here the window consists of a map of New Brunswick, which I obtained in the form of GIS "shapefiles". These were read into R and transformed into a "raw" window (consisting of a collection of polygons) using Roger Bivand's `maptools` package from CRAN (2006). These polygons were much too detailed to be of practical use in forming an observation window. There were 628 separate polygons (lakes and islands, in addition to "mainland New Brunswick") the largest of which had 138,398 edges. I therefore created two simplified windows, one being a "mask" and one being a simpler and smaller collection of approximating polygons.

A mask type window consists of an array of TRUE/FALSE values associated with a pixellation of a rectangular bounding box. In this instance I used a 500×500 pixellation which is relatively fine by the usual standards of point pattern analysis, but may actually be somewhat coarse for current purposes. This window may be seen for instance in the plot of the image depicting the "long term intensity" estimate in Fig. 4.

The simplified polygonal window was constructed using interactive graphical tools, written for this purpose in the R language. These tools plotted the raw boundary in small increments, with the relevant part of the mask window and the locations of the nearby fires (from all years) superimposed, and then made use of the `locator()` function to choose endpoints for the edges of the simplified polygon. The simplified polygon was reduced to 6 components: mainland New Brunswick, Deer Island, Campobello Island, Grand Manan Island, Isle Lameque, and Isle Miscou.

The use of the New Brunswick stereographic projection coordinate system resulted in having to deal with coordinates which are expressed as very large integers with a bewildering number of digits. Amongst other things, these huge numbers often created very untidy axis labels on graphs. I therefore decided to rescale the data in order to alleviate these problems. I made the width of the bounding box of the window equal to 1,000 units, and placed its lower left hand corner at the origin, i.e., at (0, 0). The

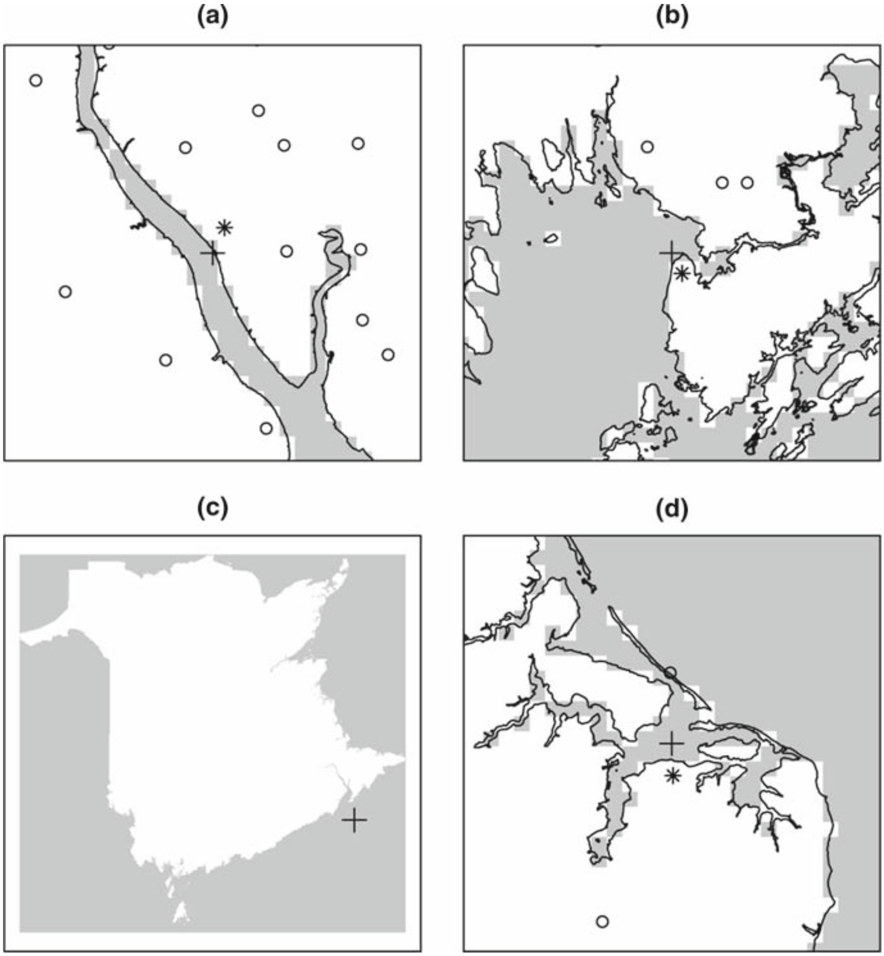

Fig. 1 Examples of points which plotted outside of the observation window

height of the bounding box was changed proportionately (resulting in a value rounded to 957). All of the forthcoming analyses will be expressed in terms of these transformed (nameless) coordinates.

Having constructed usable windows, I was then able to construct point pattern objects, that is, R objects “of class ppp”. (See Baddeley and Turner 2005). Plotting these patterns revealed that many points lay outside of the observation window which in theory should of course not happen. There are several reasons for this anomaly:

1. The simplification of the window (via discretization or the use of approximating—somewhat rough—polygons).
2. The relative coarseness of the fire locations which were given only to nearest minute which is of the order of 1 km.
3. Data entry errors.

I proceeded to “adjust” these “outsiders” by shifting points which appeared to be “mildly” out of place to nearby locations *inside* the window, and deleting points which appeared to be “wildly” out of place. (I assumed that these latter anomalies were due to data entry error.) Many outsider points were borderline between being “mildly” and “wildly” outside of the window. I such cases I mentally tossed a coin. Examples of the adjustments are shown in Fig. 1.

In this figure, the $+$ symbols indicate the initial, anomalous, positions of the points. The $*$ symbols indicate the positions to which they were shifted (in those cases—(a), (b), and (d)—for which such a shift made at least marginal sense). Case (c), in which the fire seems to be placed in the middle of the Bay of Fundy, clearly appears to be a case of data entry error. In cases (a) and (b) it seems plausible that the point is

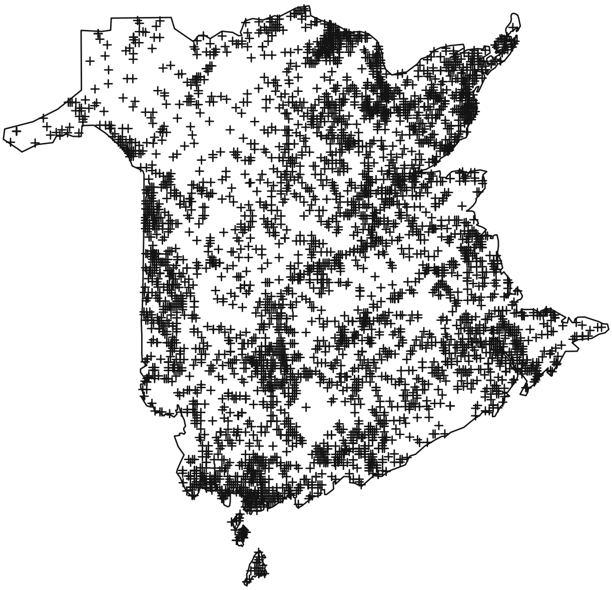

Fig. 2 Aggregate of all New Brunswick forest fires for all 16 available years

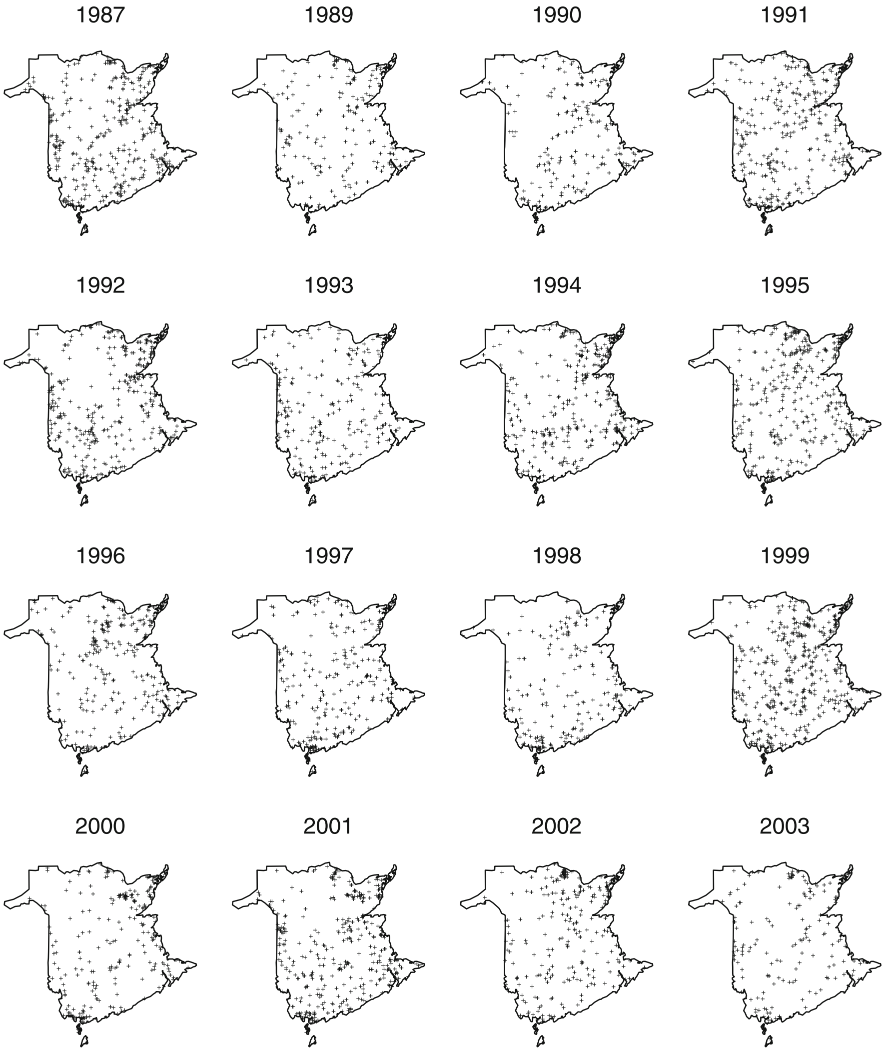

Fig. 3 New Brunswick forest fires, for all 16 available years, plotted individually

displaced due to the coarseness of the latitude and longitude coordinates. Case (d) is at best borderline.

The fires in the data sets were classified into categories “forest”, “grass”, “dump”, and “other”. For the purposes of this (very preliminary) analysis I restricted my attention to “forest” fires. It may well be the case that “grass” fires should be included with “forest” fires. Fuel type is usually reported as the fuel at the estimated point of origin, and many forest fires start in grass. However, for the time being, grass fires have been omitted. A plot of the aggregate of the forest fires from all 16 available years is shown in Fig. 2. A 4×4 array of plots of the forest fires year by year is shown in Fig. 3.

3 Exploratory data analysis

In the analysis of spatial point patterns underlying spatial trends are often of foremost interest. In the current context such trends should be modelled in terms of useful covariates, such as vegetation, terrain, and concentrations of human population or activity. Currently such covariates are not available to me. Instead of modelling trends in this way I have, mainly for purposes of illustration, adopted the procedure of estimating the long term trend for forest fires in New Brunswick, by smoothing the aggregate of the available data. Smoothing the aggregate should average out the year to year variations (in weather, human activity, etc.) that produce differences in the structure of the yearly patterns. What remains, we might argue, is (roughly) a long term spatial trend, provided that we assume stationarity of the behaviour of the trend with respect to time. Changing patterns of human activity (building of roads, deforestation, etc.) and climate change make the stationarity assumption rather dubious.

In analyzing the pattern of fires for a given year one of course requires a trend for that year, rather than the long term trend. One (very simplistic) way to proceed is to assume:

The trend for any given year is proportional to the long term trend. (1)

The proportionality constant is estimated as the ratio of the number of points in the pattern for the given year to the total number of points in the aggregate pattern. While such proportionality does not seem *a priori* completely implausible, it does appear to invite some skepticism. For instance in a given year there may be effects of the form “the northern part of the province is wetter than average, but the southern part is dryer than average”. Nonetheless one may, for want of a better expedient, proceed tentatively in this manner and examine the results.

I estimated the long term trend by applying a smoothing kernel (via the `density.ppp()` function from `spatstat`) to the superimposition of all 16 available point patterns. A major issue in applying `density.ppp()` is the choice of the “smoothing parameter” `sigma`. In this analysis. I set `sigma` somewhat arbitrarily to be equal to $0.05 \times |W|^{1/2}$ where $|W|$ denotes the area of the observation window i.e., I took `sigma` to be 5% of the edge length of a square of area equal to that of the window. Further exploration of the impact of the choice of `sigma` is clearly warranted, but criteria for its choice remain unclear. A plot of the resulting long term estimate is shown as an image in Fig. 4. I also experimented with `sigma` equal to $0.025 \times |W|^{1/2}$ and $0.10 \times |W|^{1/2}$. The impact of these different factors is discussed briefly at the end of this section.

Many questions of interest in the study of forest fires require a distinction to be made between forest fires that are due to “natural causes” and those that are caused by human activity. (See e.g., Cunningham and Martell 1973; Podur et al. 2003; Wotton and Martell 2005). The only “natural” cause recognized by the New Brunswick Department of Natural Resources (in common with most organizations with responsibility for dealing with forest fires) is *lightning*. Plots of the the aggregate of lightning caused fires and of the estimated long term spatial trend in these fires evinced a very different pattern from that of the totality of all forest fires. In future work with these data, lightning

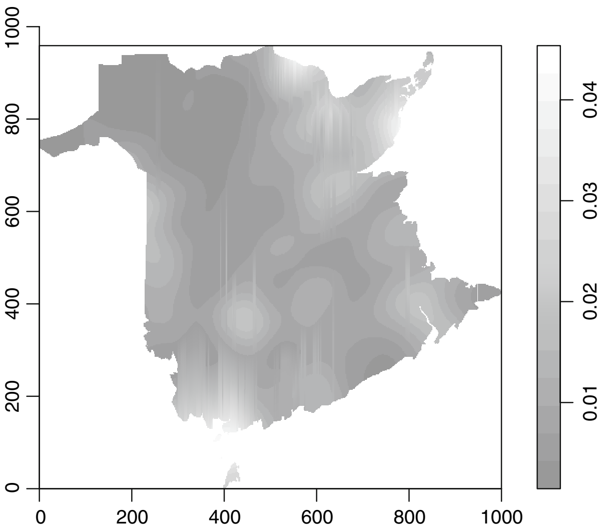

Fig. 4 Kernel smoothing estimate of long term trend

caused and human caused fires will need to be dealt with separately. However, for the purposes of the (preliminary) analyses conducted in this paper, I made no distinction between fires arising from the two causes.

The nature or behaviour of a point pattern may be thought of as comprising two components: trend (which has already been referred to), and dependence or interaction between the points of the patterns. The simplest manifestation of such interaction consists of either attraction (aggregation or clustering) or repulsion ("regularity") in the pattern. A useful step in analyzing a point pattern is to apply graphical tools which reveal information as to the nature of the interaction. A widely used tool for exploring the nature of interaction is Ripley's K-function (Ripley 1976, 1977, Diggle 2003, pp. 43 & 50, Cressie 1993, p. 639).

The basic idea in interpreting the K-function is that a constant intensity Poisson process (a process exhibiting "CSR") has a K-function equal to $K(r) = \pi r^2$. If there is attraction (with impact at distance r) then $K(r)$ is larger than it would be under CSR. Conversely, if there is repulsion then $K(r)$ is smaller than it would be under CSR.

What one deals with in practice is of course not that actual K-function of the process but rather an *estimate* of this function calculated from the point pattern under consideration. An issue that is worth emphasizing is that estimating the K Function is a more delicate problem than it appears to be. Edge effects (see Ripley 1988, chapter 3) are strongly biasing, and allowing for these effects is subtle. Some very clever people (e.g., Brian Ripley, Peter Diggle, Adrian Baddeley) have put a great deal of thought and effort into getting it right. The temptation to write one's own software to estimate the K function should be avoided. Software written by each of the aforementioned experts is readily available, specifically in the packages `spatial` (part of the `MASS` bundle), `splancs`, and `spatstat`, respectively. All of these packages are available from CRAN (2006).

For the purpose of illustration, we examine an estimate of the K-function for the New Brunswick fires year 2000 data. Figure 5 shows an estimate (`tran`) of the K-function as well as "`theo`", the theoretical value (under CSR), and a 0.05 significance level critical envelope (produced by the `envelope()` function from `spatstat`, using simulation). It is worth emphasizing that this is a *critical* envelope, and *not* a confidence envelope as it is often referred to by practitioners. Simulations are generated from the (usually unrealistic) "*null*" model (CSR) and the resulting envelope surrounds the K-function of the null model. If the K-function estimate for the observed pattern exhibits excursions outside of the envelope, then this constitute evidence against the null hypothesis of CSR.

The estimate "`tran`" is based upon the *translation* edge correction (Ohser 1983). Superimposed upon this plot is a K-function estimate "`bord`" based upon the *border* edge correction (Ripley 1988). This latter estimate is substantially different from the first. It is much rougher and descends far *below* the theoretical value at distances greater than about 60. However, the two estimates effectively agree at distances up to about 50. The reliability of the estimates at greater distances should be treated with caution. Note that the critical envelope shown was constructed using the translation edge correction and "`bord`" should not be compared with it.

The K-function estimate shown in Fig. 5 lies substantially above the critical envelope for distances up to about 200 units. On this basis one might be tempted to conclude that there is attraction between the points of the process. However, such a conclusion is unwarranted at this stage. There is indeed evidence against CSR, but the long term trend estimate shown in Fig. 4 displays striking evidence that there are

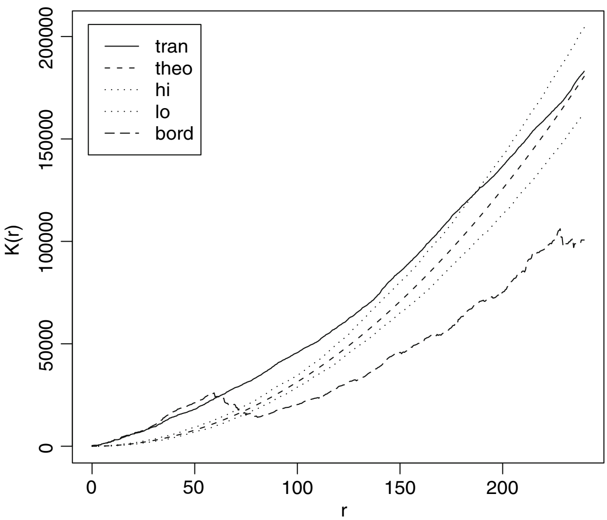

Fig. 5 Estimate of Ripley's K-function for the year 2000 data

several regions of high intensity. These regions could well explain the evidence of "attraction" evinced by the K-function estimate.

It is important to be aware that it is impossible, in theory and in practice, to distinguish rigourously between the concentrations of points due to trend, and concentrations due to (stochastic) interaction effects, on the basis of a single realization of a process. (See for instance Bartlett 1964.) In the current context we indeed have multiple realizations of the forest fire process. However, these are not i.i.d. realizations so there is no real replication available. Nonetheless an optimist might hope that the long term trend estimated from the multiple realizations gives information which at least partially separates trend and interaction effects.

Applying (tentatively!) assumption (1) makes available a trend estimate which in turn permits the use of the *inhomogeneous* K-function of Baddeley et al. (2000). A graph of an estimate of the inhomogeneous K-function for the year 2000 data, together with a 0.05 significance level critical envelope, is shown in Fig. 6. The range of distances at which the estimate is evaluated has been restricted to the short interval [0, 25] so as to make more of the fine detail apparent.

Examination of this plot reveals a hint of attraction at distances between 5 and 15 units, and what might be termed a "nugget" effect at distance 0. Plots over larger ranges of distance indicated that except for distances less than 15 units, the inhomogeneous K-function estimate stays within the critical envelope. For distances more than about 40 units, it remains close to but slightly above the lower bound of the envelope. Pursuit of an explanation for the "nugget" effect at 0 led to an examination of the nearest neighbour distances of the data, several of which turned out to be 0, showing that there are several coincident points in the pattern.

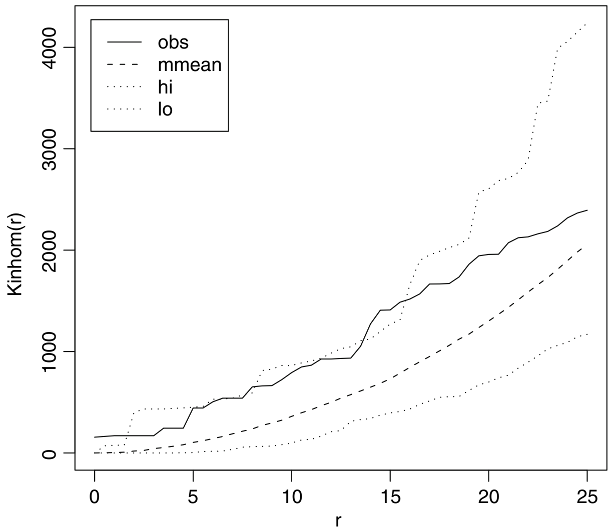

Fig. 6 Estimate of the inhomogeneous K-function for the year 2000 data with intensity taken to be proportional to the "long term trend" plotted in Fig. 4

The theory on which this sort of analysis is based makes the fundamental assumption that all points are distinct. This leaves one in a quandary as to what to do about the present violation of this assumption. For the purposes of the current study, I decided to eliminate points from coincident pairs. This is of course bad statistical practice—one should not modify the data to fit one's theory! On the other hand it's either that or develop a whole new theory, a rather Herculean task. Moreover, fires at 0 distance from each other, in the same year, would seem to be implausible from a practical point of view. It appears reasonable to proceed cautiously with modified data, so as to be able to exploit the power of existing tools for analysis.

The existence of coincident points is not peculiar to the year 2000 data. All of the data sets have a number of such points. There are also many pairs which appeared to be at anomalously small (though non-zero) distances as well. I decided to eliminate all such pairs for the purpose of this exploratory analysis. All further discussion will be in terms of the *modified* data sets. Any conclusions based on the forthcoming analyses should therefore be tempered with even more than the usual amount of caution, at least until a deeper understanding of the coincident point phenomenon is achieved. Possible reasons for the existence of coincident points should be explored with the supplier of these data sets, as a part of other data cleaning and verification tasks.

An inhomogeneous K-function plot was also constructed for the year 2000 *modified* data. This plot differed little from that for the unmodified data except that the "nugget effect" disappeared. Plots over larger distance ranges showed it to stay below the simulated mean level for distances greater than 40+ units, but to remain substantially above the lower bound of the critical envelope. The message from this is unclear. It would appear (see Sect. 5) that at least part of the problem is the assumption (1) that the annual trend is proportional to the long term trend.

Plots of the inhomogeneous K-functions, with critical envelopes, similar to that for the year 2000 data, were produced for all sixteen available years. The only plot which evinced any evidence of interaction was that for the year 1996. This plot seemed to indicate attraction or clustering for distances between about 15 and 80 units. The plots for the 1992 and 1994 data sets, like that for the year 2000 data set, track fairly close to the the lower edge of the critical envelope.

All of the plots involved in the foregoing discussion were based on a trend estimate proportional to the "long term trend", The latter trend was calculated using a smoothing parameter equal to 0.05 times the square root of the window area. Naturally, when a factor of 0.025 was used in place of 0.05, there was even less "evidence" for interaction. When a factor of 0.10 was used, there was "more" evidence.

In particular, for the year 2000 data there is little if any indication of attraction when a factor of 0.025 is used, merely a hint of attraction when 0.05 is used, and substantial indication of attraction at distances up to 25 units when 0.10 is used. For all three values of the factor the K-function graph tends to track the lower bound of the envelope for large values of distance; for the 0.025 value it goes well below the lower bound for distances greater than 100. On the basis of only an ad hoc trend formed by smoothing, and with no genuine replication of patterns, there appears to be no sensible way of determining the "right answer".

4 Models for point process interaction

The inhomogeneous K-function for the 1996 data seems to indicate some evidence of interaction in the form of attraction, so it would be interesting to fit a model capable of reflecting this property to this data set. The prototype model for point process interaction, appearing in the seminal work of David Strauss (1975) was indeed proposed as a model for attraction or "clustering". Unfortunately it turned out to be appropriate only for modelling *repulsion*. If one is to understand point process models, it is important to see clearly why this is so.

A point process, roughly speaking, is a mechanism which generates random patterns of points. Mathematically the word "random" means that these patterns are the values of a *random variable* which takes values in a "space" consisting of all possible point patterns. (We will not go into technical details here. Detailed discussion of this issue can be found in Kelley and Ripley (1976). See also Ripley and Kelley 1977, and Daley and Vere-Jones 2003, p. 123 ff.) From certain points of view, the most convenient descriptor of a random variable (such as a point process) is its probability density function, if it possesses one—and those processes in which we are interested do indeed possess one.

However, density functions for point processes are difficult to grasp intuitively. A closely related concept, which is somewhat easier to comprehend, is that of "Papangelou conditional intensity function" (see Papangelou 1974; see also, e.g., van Lieshout 2000, p. 39) which I shall denote by $\lambda(u, \boldsymbol{x})$. Roughly speaking $\lambda(u, \boldsymbol{x})|\Delta u|$ is the probability of observing a point of the process in a small neighbourhood Δu of u, conditional upon the rest of the process $\boldsymbol{x}$. The "rest of the process" simply means $\boldsymbol{x}$ with u removed, if u is a point of $\boldsymbol{x}$, and $\boldsymbol{x}$ if u is not a point of $\boldsymbol{x}$. Note that if $\boldsymbol{x}$ is a Poisson process with constant intensity λ then $\lambda(u, \boldsymbol{x}) = \lambda$ for all u, and more generally if $\boldsymbol{x}$ is an inhomogeneous Poisson process with intensity $\lambda(u)$, then $\lambda(u, \boldsymbol{x}) = \lambda(u)$.

The Papangelou conditional intensity function of a process is related (under certain theoretical restrictions) to the probability density function of the process (say $f(\boldsymbol{x})$) by

$$\lambda(u, \boldsymbol{x}) = \frac{f(\boldsymbol{x} \cup \{u\})}{f(\boldsymbol{x} \setminus \{u\})}.$$

It should be noticed that if $u \in \boldsymbol{x}$ then $\boldsymbol{x} \cup \{u\} = \boldsymbol{x}$, and if $u \notin \boldsymbol{x}$ then $\boldsymbol{x} \setminus \{u\} = \boldsymbol{x}$.

The function $f(\boldsymbol{x})$ takes the form $\alpha g(\boldsymbol{x})$ where $g(\boldsymbol{x})$ is explicitly representable in terms of model parameters and statistics calculated from $\boldsymbol{x}$. The constant α must have the property that $f(\boldsymbol{x})$ integrates to one, i.e.,

$$\frac{1}{\alpha} = \int_{\mathcal{X}} g(\boldsymbol{x})\, d\mu(\boldsymbol{x}) \tag{2}$$

where $\mathcal{X}$, is the space of all point patterns in W and μ is an appropriate measure on $\mathcal{X}$.

Except for Poisson processes, the integral on the right hand side of (2) cannot be calculated, so that it is impossible to express α explicitly in terms of the model parameters. It is this recalcitrance of α that makes spatial point process analysis such a difficult subject.

Most interest, in models involving non-trivial interaction, focuses on *pairwise interaction processes* of which the prototype is the Strauss process (Strauss 1975). A pairwise interaction model is expressed in terms of a (symmetric) pairwise interaction function $h(u, v) = h(v, u)$. In order that the process be stationary it is necessary that $h(u, v)$ depend only on the distance $|u - v|$ between u and v. (See Ripley 1988, p. 51.) The Papangelou conditional intensity function of a pairwise interaction process is equal to

$$\lambda(u, \boldsymbol{x}) = \beta \times \prod_{x_i \in \boldsymbol{x}} h(u, x_i)$$

where β is a constant—a parameter of the process. In general $h(u, v)$ will depend on other parameters.

For the Strauss process, the pairwise interaction function is

$$h(u, v) = \begin{cases} \gamma & \text{if } 0 < |u - v| < r \\ 1 & \text{if } |u - v| \geq r \end{cases}$$

where $0 \leq \gamma \leq 1$ and $r \geq 0$. The parameter r is called the interaction radius of the process. It is convenient to refer to a pair of points $\{u, v\}(u \neq v)$ such that $|u - v| < r$ as being an r-close pair. The conditional intensity function becomes

$$\lambda(u, \boldsymbol{x}) = \beta \times \gamma^{t(u, \boldsymbol{x})}$$

where $t(u, \boldsymbol{x}) = \#\{x_i \in \boldsymbol{x} : 0 < |u - x_i| < r\}$ and the density function becomes

$$f(\boldsymbol{x}) = \alpha \beta^{n(\boldsymbol{x})} \gamma^{s(\boldsymbol{x})}$$

where $n(\boldsymbol{x})$ is the number of points in $\boldsymbol{x}$ and $s(\boldsymbol{x})$ is the number of r-close pairs in $\boldsymbol{x}$.

If $\gamma = 1$ the process becomes simply a Poisson process with constant intensity β. If $\gamma = 0$ then there is zero probability of there being a point of the process within distance r of any other point of the process i.e., there is "total inhibition" at distances up to r. This special case of the Strauss process is called the "hard core" process. In general, for $0 < \gamma < 1$ there is inhibition at distances smaller than r but this inhibition is less than total.

One might reasonably expect that if one were to take γ to have a value greater than 1 then the model would determine a process in which there was a strong tendency for points to be at distance less than r from each other, that is to say a process which exhibited clustering. This does not however work: If $\gamma > 1$ then the integral of $g(\boldsymbol{x}) = \beta^{n(\boldsymbol{x})} \gamma^{s(\boldsymbol{x})}$ on the right hand side of (2) is infinite, whence the density function $f(\boldsymbol{x})$ does not exist.

We might then seek to adjust the Strauss process so as to come up with a model that encompasses the clustering phenomenon. To make the appropriate adjustment, we note that goes wrong if we try to use $\gamma > 1$ in the Strauss model is that there is a "blow-up" effect. Intuitively we may think along the lines that the more points of the pattern there are near u, the more there tend to be, resulting in an infinity of points anywhere that there is a point. To obtain a model that produces clustering we need to include attraction between points, and at the same time keep the "blow-up" effect under control.

One "obvious" way of doing this is to introduce inhibition at small interpoint distances. We define a pairwise interaction function

$$h(u, v) = \begin{cases} 0 & \text{if } 0 < |u - v| < r_1 \\ \gamma & \text{if } r_1 \leq |u - v| < r_2 \\ 1 & \text{if } |u - v| \geq r_2 \end{cases}$$

This interaction prevents points from piling up next to each other, and the resulting density is indeed well-defined, i.e., the "un-normalized density" $h(\boldsymbol{x})$ does indeed have a finite integral, for any (non-negative) value of γ.

The model so defined is called the "Strauss hard core model" in the "folk lore" of spatial point process theory. Although this model appears to be well known to the cognoscenti, I have been unable to find any explicit reference for it. For example, (Møller and Waagepetersen 2004, p. 86), proceed from the Strauss process to the "multiscale process" for which they cite Penttinen (1984). The multiscale process is also discussed by Takacs and Fiksel (1986). The multiscale process is a pairwise interaction process for which the interaction function is a step function, and the Strauss hard core process is clearly a special case. Takacs and Fiksel use the Strauss hard core model in their paper, without explicitly calling it that, and also use bivariate Strauss hard core models. Särkkä (1993) likewise uses bivariate Strauss hard core models, and this idea is further taken up in Baddeley and Turner (2000).

It would appear that the Strauss hard core model should provide a sensible means of modelling attraction. However, in practice the model seems to be highly unstable. For example I fitted a Strauss hard core model (with trend) to the year 2000 fire data, and then simulated a pattern from the fit. The pattern to which the model was fitted had 229 points; the simulated pattern (which one would expect to have a commensurate number of points) turned out to have 2679 points, mostly concentrated around hot spots of the trend in the Acadian Peninsula and between St. John and the Maine border.

Another way of preventing "blow-up" is to place an upper bound on the amount of influence that near neighbours can have. A model incorporating this idea, the "saturation model", was introduced by Geyer (1999). The density function for Geyer's model is made to depend upon quantities of the form $\min\{c, t(u, \boldsymbol{x})\}$ where c is a constant greater than or equal to 0. The parameter c is called the saturation parameter. The un-normalized density function for Geyer's model is indeed integrable for all non-negative γ and values of $\gamma > 1$ induce clustering.

Geyer (1999) also proves that processes satisfying this model have the pleasant Ruelle stability property which guarantees that they have "thermodynamic behaviour" and hence can be reliably simulated via the Metropolis–Hastings algorithm. The Geyer model appears to provide a reasonably effective means of modelling attraction in point patterns.

Notice that if c is very large, then $\min\{c, t(u, \boldsymbol{x})\}$ will be equal to $t(u, \boldsymbol{x})$ for all i. In such a case, for a *given* pattern, Geyer's model effectively reduces to the Strauss model as the saturation parameter is allowed to grow. Note that the γ of the (equivalent) Strauss model is equal to the square of the γ parameter in Geyer's model.

5 Fitting models to data

I now consider some examples of fitting models to forest fire patterns. I start out with the year 2000 data, for which I previously examined the inhomogeneous K-function estimate. This estimate produced no clear evidence of interaction between the points of the pattern, so I fitted a simple "trend only" (inhomogeneous Poisson) model, invoking assumption (1) to model the trend.

To fit this model I made use of the flexible modelling capabilities of the `spatstat` package. In this paper I only discuss the (quick and dirty) maximum pseudolikelihood (mpl) method of fitting, but it should be noted that `spatstat` is also capable of using the "one step improvement" on mpl developed by Huang and Ogata (1999).

Models to be fitted by the `spatstat` package must be expressed in terms of their Papangelou conditional intensity functions, given in exponential family form as

$$\lambda(u, \boldsymbol{x}) = \exp\{\phi^{\mathsf{T}} b(u) + \theta^{\mathsf{T}} S(u, \boldsymbol{x})\}$$

where $\phi^{\mathsf{T}} b(u)$ is the *trend* component of the conditional intensity and $\theta^{\mathsf{T}} S(u, \boldsymbol{x})$ is the *interaction* component. The modelling syntax of `spatstat` is based on this decomposition. This syntax is, as a result, closely analogous with that of the `glm()` function from R (R Development Core Team 2005) which is based in turn on the GLIM package (Royal Statistical Society GLIM Working Party 1999) with "trend" corresponding to "linear predictor", and "interaction" corresponding to "family". The modelling capabilities of `spatstat` directly estimate the exponential family parameters ϕ and θ; the functions $b(u)$ and $S(u, \boldsymbol{x})$ may depend on further parameters (termed "irregular") such as the interaction radius in the Strauss model. These parameters must be estimated by other means, possibly profile pseudolikelihood.

To model the trend on the basis of assumption (1) I include the long term trend (of which I have already obtained a completely specified estimate) in the model as an "offset" term. (This is generalized linear modelling terminology.) To conform to the exponential family structure, the offset must be the *log* of the existing trend estimate. Explicitly I assume that the conditional intensity function (which is actually "unconditional" since there is no interaction term) is equal to

$$\lambda(u, \boldsymbol{x}) = \lambda(u) = \beta\tau(u) = \exp\{\phi + \log(\tau(u)\}$$

where $\tau(u)$ is the (estimated) long term trend, β is the proportionality constant and $\phi = \log(\beta)$ is the only parameter to be estimated.
The `spatstat` syntax for fitting this model is

```
fit1 <- ppm(nbff.00,~offset(log(intens)),
        covariates=list(intens=intens))
```

where `intens` is the (non-parametric) estimate of the long term trend produced by `density.ppp()`, and `nbff.00` is the point pattern object of all New Brunswick forest fires in the year 2000.

The output from this fit is a large and complicated object, but the print method for objects of this class ("`ppm`") gives concise and interpretable output, the relevant feature of which is the "intercept" term, $\hat{\phi} = -2.917439$. The exponential of this term, 0.05407199, is the ratio of the (fitted) year 2000 intensity to the long term intensity. This is of course nearly identical to the ratio of the number of points in the (modified) year 2000 data set, to the total number of data points, i.e., 246/4599 = 0.05348989. The reason that it is not truly identical is "numerical"—integrals are being calculated via various numerical approximations. Observe in particular that the (calculated) integral of the long term intensity

```
 > summary(intens)$integral
[1] 4556.586
```

is a bit smaller than its theoretical value of 4,599, in part due to the fact that the kernels involved in the kernel smoothing estimator of trend are being integrated only over the observation window (rather than the whole plane) and hence are losing a small amount of "tail mass".

The output from the fit of course does not say whether the model being fitted is actually sensible. In other areas of statistics, the appropriateness of the model being fitted is often assessed via residual plots. Up until recently there was no appropriate concept of "residual" to associate with planar point process models. That situation changed with the publication of a paper by Baddeley et al. (2005) and the implementation of the concepts in that paper within the `spatstat` package. Thus residual plots of various sorts are now available for planar point process models. There remains however a great deal of effort to be expended in practicing with the use of these plots and building up an understanding of, and intuition for, their interpretation.

A quantile–quantile plot for the "trend only" model fitted to the year 2000 data is shown in Fig. 7. This plot gives a very strong message that the fit is not good. Since there seems little indication (from the inhomogeneous K-function) of interaction in the pattern the explanation would seem to be that the form of the trend is wrong. I decided to try a much more flexible form for the trend, using the `s()` (spline) function from the `mgcv` package. The fit with this form of trend is obtained by

```
fit <- ppm(nbff.00,~s(intens),
       covariates=list(intens=intens),use.gam=TRUE)
```

The quantile–quantile plot for this model is shown in Fig. 8. This plot appears to indicate that the fit is adequate. The conclusion would appear to be that at least for the year 2000 assumption (1) is *not* valid. That is, in a particular year the intensity of the forest fire process may go up in some areas, with respect to the long term average, and down in others. While not overwhelmingly surprising, this conclusion is, as remarked on page 205 not *a priori* completely obvious.

Other diagnostic plots for the "spline trend" model are shown in Fig. 9. These are harder to interpret. If the model fits well, then the image plots should "ideally" depict relatively flat surfaces for the smoothed residuals. The pattern of contours in

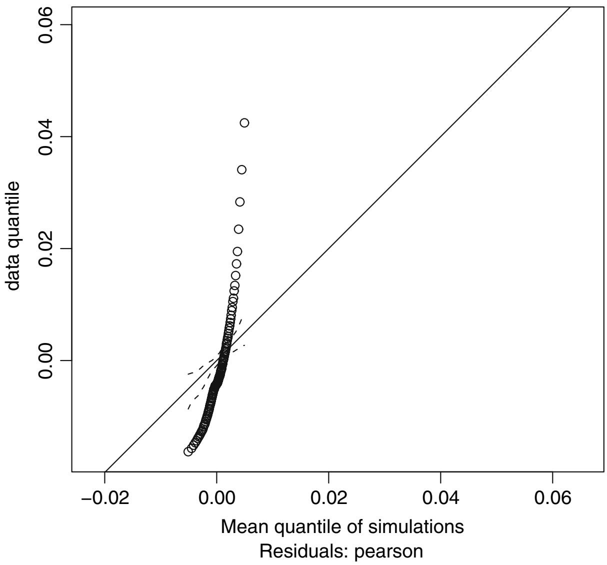

Fig. 7 The quantile–quantile plot for the inhomogeneous Poisson model with trend taken to be proportional to the "long term trend" plotted in Fig. 4

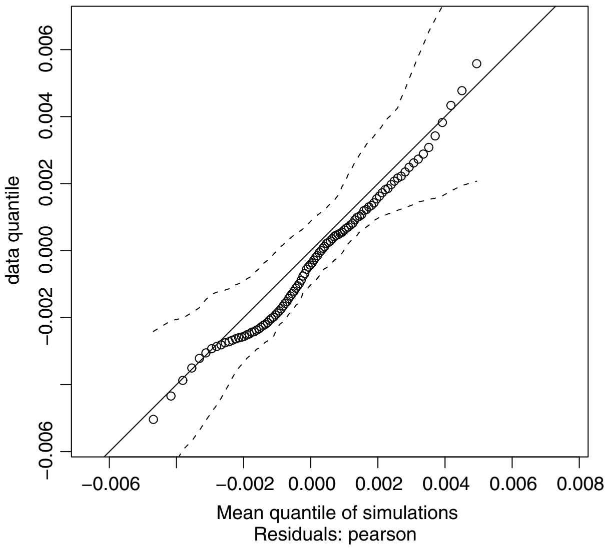

Fig. 8 The quantile–quantile plot for the model with trend equal to a smooth (spline) function of the long term trend, fitted to the year 2000 data

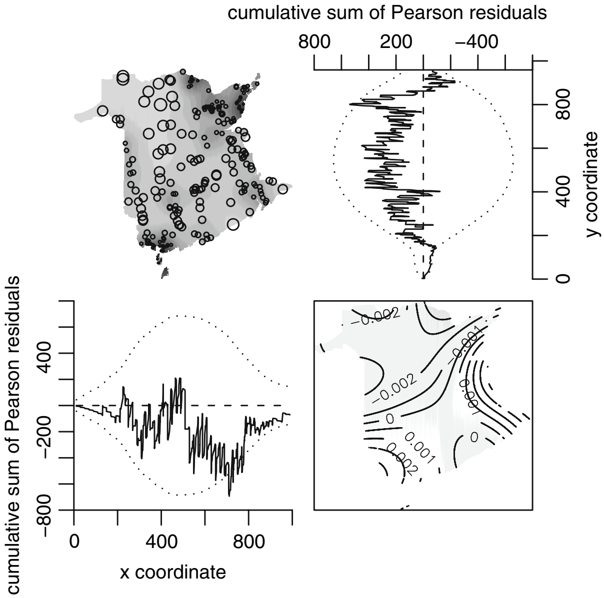

Fig. 9 Diagnostic plots for the model referred to in Fig. 8

the image plot in the lower right corner is complicated, but does not appear to have any particularly high or low spots, and might well be asserted to depict a surface with a "random" shape. In any case the quantile–quantile plot for this fit provides assurance that the residual surface is no less random than one might expect given that the model is appropriate.

The 1996 data set seems to be different from the other years in the available collection in that its inhomogeneous K-function estimate appears to indicate concentrations of points not explained by the trend based on assumption (1). Just why, in practical terms, 1996 is different from the other years is an interesting question in its own right, but one which I have not yet had a chance to explore. A trend only model, employing a smooth spline transform of the long term trend estimate, did not provide a good fit to this data set, unlike the case for the 2000 data. Hence there is some reason to believe that there is interaction, in the form of attraction or clustering, among the points of the 1996 pattern.

I attempted to model this apparent attraction in the 1996 pattern by means of the Geyer model discussed in Sect. 4. The Geyer model involves two irregular parameters, namely the interaction radius r and the saturation parameter c, which must be estimated somehow. I approached this problem via a (very rough) "profile pseudolikelihood" procedure. That is, I tried interaction radii varying from 10 to 80 in steps of 10 (the upper bound of 80 being chosen on the basis of the inhomogeneous K-function plot for these data). Likewise I considered values for c in the set $\{1, 2, \ldots, 10\}$. I then fitted models with interaction `Geyer(r,s)` for all 80 pairs of these `(r,s)` values, and extracted the log pseudolikelihood.

A plot of the resulting values is shown in Fig. 10. The $r = 10$ and $r = 20$ profiles appear to be very similar for $c \geq 5$. I tried several parameter pairs, including ($r = 10, c = 4$), ($r = 20, c = 5$) and ($r = 20, c = 8$). I fitted these models by means of R code such as the following :

```
fit.geyer<- ppm(nbff.96,~s(intens),inter=Geyer(10,4),
            covariates=list(intens=intens),use.gam=TRUE)
```

The quantile–quantile plot of the residuals looked best for the ($r = 10, c = 4$) pair. This plot is shown in Fig. 11. This figure still indicates considerable lack of fit, although it was a substantial improvement over the fit of a trend only model.

To get a feel for diagnostic techniques, it is helpful to apply them to simulated data where the right answer is known. To this end I simulated a point pattern from the foregoing fit of the "trend plus Geyer" model. I then fitted the "trend plus Geyer" model to the simulated data, and plotted the diagnostic and quantile–quantile plots that resulted. The quantile–quantile plot is shown in Fig. 12. This quantile–quantile plot makes it clear that what we can reasonably expect to get (if the form of the model is correct) is much better than what we actually get from the fit to the real 1996 data. It is interesting to note the striking sigmoid shape of the quantile–quantile plot—present despite the fact that the model is right. Presumably this is evidence that a sigmoid shape is acceptable provided that the plot stays within the critical band.

The estimate of γ from the `Geyer(10,4)` fit to the real 1996 data is $\hat{\gamma} = 1.4632$. This would appear to be substantially greater than 1 (in conformity with the apparent phenomenon of attraction between points). Whether it is significantly greater than 1 could be assessed by Monte Carlo inference, but there seems little point in my doing so, since I still haven't got the model right.

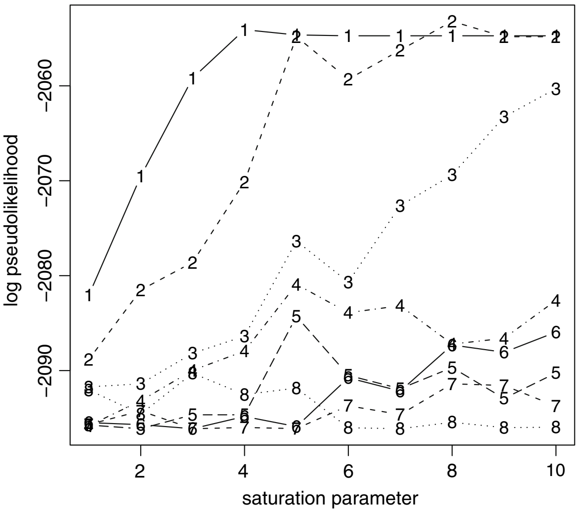

Fig. 10 Pseudolikelihood profiles for estimating values of of irregular parameters for the Geyer model for the 1996 data. The x-axis represents the saturation parameter. The profile labelled i is for interaction radius equal to $10 \times i$, for $i = 1, \ldots, 8$

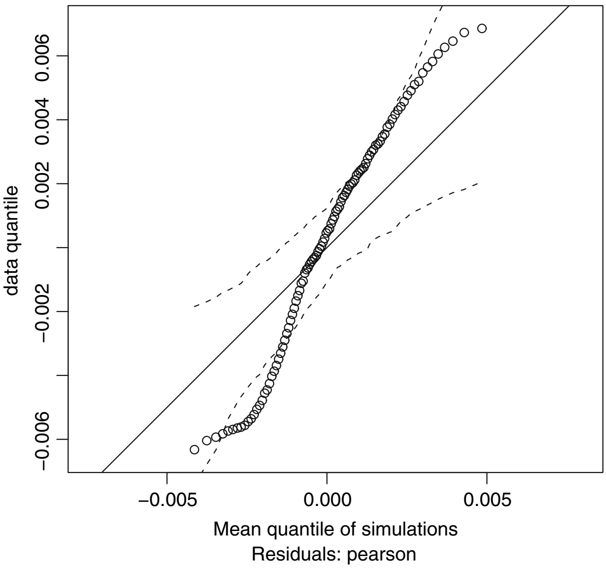

Fig. 11 Quantile–quantile plot for the smooth trend plus `Geyer(10,4)` model fitted to the 1996 data

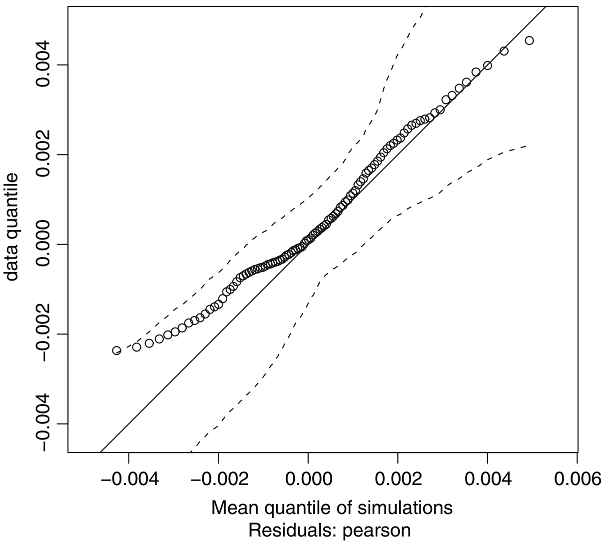

Fig. 12 Quantile–quantile plot for the smooth trend plus `Geyer(10,4)` model that was fitted to data simulated from this model

So what is the right model? That seems to be a "question for future work". In particular the *trend* should be modelled sensibly (rather than by assumption (1)). Other interaction models should be experimented with and the available set of models should be expanded. One type of model that holds some promise is the "multiscale process" (Møller and Waagepetersen 2004, p. 86) also called the "piecewise constant pairwise interaction" model, which is implemented by the function `PairPiece` in the `spatstat` package. As noted in Sect. 4 this generalizes the Strauss process by replacing a two-valued step function (with values γ and 1) by a multi-valued step function.

The difficulty is that this model is capable only of modelling repulsion between points; if any of the function values are greater than 1, the density function for the model does not exist (is not integrable). Developing a corresponding valid model may be possible, but will entail a substantial amount of work.

In the meantime, just for fun, I fitted the (meaningless) `PairPiece` model to the 1996 data. It is meaningless because several of the step function values do indeed turn out to be larger than 1:

```
fit <- ppm(nbff.96,~s(intens),inter=PairPiece(10*(1:8)),
       covariates=list(intens=intens),use.gam=TRUE)
fit
.....
Fitted interaction parameters gamma_i:
 [0,10) [10,20) [20,30) [30,40) [40,50) [50,60) [60,70)
 [70,80)
 1.8912 1.1813 1.1013 1.0516 0.9385 1.0865 0.9997 0.9832
```

The quantile–quantile plot for this fit is shown in Fig. 13. It indicates that in some sense the fit is better—for most of its extent the plot is very close to the theoretically ideal line. However, the lower end of the plot plunges down dramatically and protrudes beyond the lower edge of the critical envelope. Thus it appears that the lower tail of the residuals is too heavy, i.e., the lowest residuals are too (negatively) large. What is

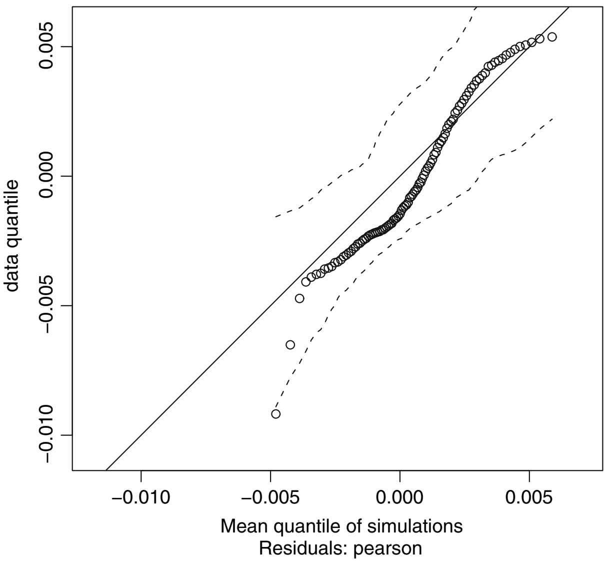

Fig. 13 Quantile–quantile plot for the (invalid) piecewise constant pairwise interaction model, with smooth trend, that was fitted the 1996 data

causing this phenomenon, and what sort of model might be used to accommodate it remain problems whose solutions are opaque to me at this time.

Note that there are extra subtleties associated with interpreting the quantile–quantile plot in Fig. 13. The critical envelope is based upon simulations from the fitted model. Since the fitted model is not a valid model, the simulation software projects the model onto the nearest valid model and simulates from that. In the current case, this makes the model effectively Poisson. This will have little impact on the conclusions however, since the fit is very close to "ideal" except for the droopy lower tail, which definitely indicates that something is not as it should be, irrespective of the critical envelope.

6 Alternative modelling strategies and conclusions

Clearly a great deal more remains to be done in analyzing the New Brunswick fire pattern data; the surface has not even been scratched as yet. What I can say on the basis of my analysis so far is that for most years there is no compelling evidence of interpoint interaction in the patterns. The situation is of course far from clear; evidence of interaction has been sought by means of the inhomogeneous K-function. I remain caught in the dilemma of how much to smooth the underlying trend estimate, which is all part of the over-riding dilemma of the indistinguishability of trend and interaction on the basis of a single realization of a spatial point process.

On the other hand, new diagnostic techniques indicate that the year 2000 pattern is adequately modelled in terms of a trend which is a smooth (spline) function of the long term trend estimate, with no interaction. The trouble with this argument is that the spline trend is too flexible, and too arbitrary. To have any hope of discerning reliably between trend effects and interaction effects, it is necessary to have a trend estimate that one can believe in. To produce such an estimate one needs to model the trend in terms of meaningful covariates (vegetation, terrain, etc.). Another possible class of covariates could be smooths of the patterns for the past, say, three years.

At the given level of smoothing of the "long term trend" estimate, the year 1996 stands out as being different. There seems to be substantial evidence of interpoint interaction for this data set. Why, in practical terms, 1996 is different in this respect from the other available years is an interesting question which should be pursued further. It may be the case that 1996 is anomalous because a substantial number of fires in that year were caused by a relatively small number of e.g., thunderstorms, or railway incidents, or activities of arsonists. It may also be the case that considerable theoretical development, to expand the scope of available models, will be required in order to obtain a satisfactory fit to the 1996 data.

Scope for a very different sort of modelling opens up if the available temporal information is used; so far I have made no attempt to do so. Each fire in the collection of New Brunswick data sets comes with a "discovery time" and an "out time". The discovery times could be used (as surrogates for occurrence times) in spatio-temporal modelling of the patterns. For the available temporal data to be useful, further data cleaning will be required: Some of the fires are recorded as having out times which precede their discovery times!

The spatio-temporal approach makes more sense than simply treating the patterns in a purely spatial manner. As Vere-Jones remarks (2006), referring to an observation of Whittle (1962), “space-time models are more likely to be insightful, to penetrate further into the physical process generating the picture in front of us at any given time, than a purely static analysis of that picture.” Once adequate spatio-temporal models are developed, temporal covariates, such as the weather history over, say, two or three weeks preceding the current time point, might provide useful explanatory power.

As proposed in Vere-Jones (2006) it may be possible to construct spatio-temporal models for forest fire patterns by adapting those already developed in the geophysical context for earthquake patterns. (See Ogata 1998, 1999, 2004; Daley and Vere-Jones 2003, p. 203 ff., p. 239 ff. for discussions of earthquake models.) On the other hand, it may be necessary to develop completely new models. There does not appear to have been a great deal of work done on spatio-temporal modelling of forest fire patterns so far, possibly in part for the reason that temporal information for forest fires is less readily available. Two exceptions are the papers of Peng et al. (2005) and Preisler et al. (2004).

A somewhat different approach to modelling forest fire patterns might be make use of Cox processes (see Daley and Vere-Jones 2003, p. 169, Diggle 2003, pp. 68–71, Cressie 1993, p. 657 ff., Stoyan et al. 1995, p. 154 ff., Møller and Waagepetersen 2004, Chap. 5). Briefly, these processes are inhomogeneous Poisson processes where the underlying trend is itself random (a Gaussian random field). These “doubly stochastic” processes (with, say, the random field varying continuously in time) seem intuitively plausible as models for forest fire data. The idea of using the Cox process model is interesting in theory at least. Fitting such processes presents substantial challenges however and it is not clear how much progress can be made.

Let me conclude by reiterating that the New Brunswick fire patterns constitute an interesting collection of data sets. A great deal remains to be done in the way of analyzing them. These data sets are now included in the `spatstat` package which is available from CRAN (2006).

Acknowledgements The author expresses his sincere gratitude to an anonymous referee whose cogent criticisms led to a substantial clarification of several issues dealt with in this paper. Thanks are also due to Prof. David Martell of the Faculty of Forestry, University of Toronto, who provided a great deal of useful advice. The data analyzed in this paper were kindly made available to the author by the New Brunswick Department of Natural Resources. The author’s research is supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Baddeley A, Møller J, Waagepetersen R (2000) Non- and semiparametric estimation of interaction in inhomogeneous point patterns. Stat Neerlandica 54(3):329–350

Baddeley A, Turner R (2000) Practical maximum pseudolikelihood for spatial point patterns. Australian & New Zealand J Stat 42:283–322

Baddeley A, Turner R (2005) Spatstat: an R package for analyzing spatial point patterns. Journal of Statistical Software 12(6):1–42. http://www.jstatsoft.org, ISSN:1548–7660

Baddeley A, Turner R, Møller J, Hazelton M (2005) Residual analysis for spatial point processes (with discussion). J Roy Stat Soc ser B 67(5):617–666

Bartlett MS (1964) The spectral analysis of two-dimensional point processes. Biometrika 51:299–311

Chen K, McAneney J (2004) Quantifying bushfire penetration into urban areas in Australia. Geophy Res Lett 31. doi: 10.1029/2004GL020244

CRAN (2006) Consolidated R Archive Network. http://www.r-project.org

Cressie NAC (1993) Statistics for Spatial Data (Revised ed.). Wiley, New York

Cunningham AA, Martell DL (1973) A stochastic model for the occurrence of man-caused forest fires. Canadian J Forestry Res 3:282–287

Daley DJ, Vere-Jones D (2003) An introduction to the theory of point processes, 2nd edn., Volume I: Elementary Theory and Methods. Springer, New York

Diggle PJ (2003) Statistical Analysis of Spatial Point Patterns, 2nd edn. Arnold, London

Geyer C (1999) Likelihood inference for spatial point processes. In: Barndorff-Nielsen OE, Kendall WS, van Lieshout MNM (eds) Stochastic Geometry: Likelihood and Computation, Number 80 in Monographs on Statistics and Applied Probability, Ch 3. Chapman & Hall/CRC, Boca Raton, pp 79–140

Huang F, Ogata Y (1999) Improvement of the maximum pseudo-likelihood estimators in various spatial statistical models. J Comput Graph Stat 8:510–530

Kelley FP, Ripley BD (1976) On Strauss's model for clustering. Biometrika 63:357–360

Møller JM, Waagepetersen R (2004) Statistical Inference and Simulation for Spatial Point Processes. Number 100 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton

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

Ogata Y (1999) Seismicity analysis through point-process modelling: a review. Pure and Appl Geophy 155:471–507

Ogata Y (2004) Space-time model for regional seismicity and detection of crustal stress changes. J Geoph Res 109:B03308, doi: 10.1029/2003JB002621

Ohser J (1983) On estimators for the reduced second moment measure of point processes. Math Operationsforschung und Stat, series Stat 14:63–71

Papangelou F (1974) The conditional intensity of general point processes and an application to line processes. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 28:207–226

Peng RD, Schoenberg RP, Woods J (2005) A space-time conditional intensity model for evaluating a wildfire hazard index. J Am Stat Asso 100(469):26–35

Penttinen A (1984) Modelling Interaction in Spatial Point Patterns: Parameter Estimation by the Maximum Likelihood Method. Number 7 in Jyväskylä Studies in Computer Science, Economics, and Statistics. University of Jyväskylä, Jyväskylä Finland

Podur J, Martell DL, Csillag F (2003) Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998. Ecol Model 164:1–20

Preisler HK, Brillinger DR, Burgan RE, Benoit JW (2004) Probability based models for estimation of wildfire risk. Int J Wildland Fire 13:133–142

R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0

Ripley BD (1976) The second-order analysis of stationary point processes. J Appl Prob 13:255–266

Ripley BD (1977) Modelling spatial patterns. J Roy Stat Soc 39:172–192

Ripley BD (1988) Statistical inference for spatial processes. Cambridge University Press, Cambridge

Ripley BD, Kelley FP (1977) Markov point processes. J London Math Soc 15:188–192

Royal Statistical Society GLIM Working Party (1999) The GLIM System, Release 4 Manual. Clarendon Press, Oxford

Särkkä A (1993) Pseudo-likelihood approach for pair potential estimation of Gibbs processes. Number 22 in Jyväskylä Studies in Computer Science, Economics and Statistics. University of Jyväskylä, Jyväskylä Finland

Stoyan D, Kendall WS, Mecke J (1995) Stochastic Geometry and its Applications, 2nd edn. John Wiley & Sons, Chichester

Strauss DJ (1975) A model for clustering. Biometrika 63:467–475

Takacs R, Fiksel Th (1986) Interaction pair-potential for a system of ant's nests. Biom J 28:1007–1013

Thomson DB, Mephan MP, Steeves RR (1977) The stereographic double projection. Technical Report 46, University of New Brunswick, Fredericton NB, Canada. http://gge.unb.ca/Pubs/Pubs.html

van Lieshout MNM (2000) Markov Point Processes and their Applications. Imperial College Press, London

Vere-Jones D (2006) Some models and procedures for space-time point processes. Preprint; to appear in Environmental and Ecological Statistics

Whittle P (1962) Topographic correlation, power-law covariance functions, and diffusion. Biometrika 49:305–314

Wotton BM, Martell DL (2005) A lightning fire occurrence model for Ontario. Canadian J Forestry Res 35:1389–1401

Author Biography

Rolf Turner is a Professor in the Department of Mathematics and Statistics at the University of New Brunswick. He is soon (as of January 2007) to take up a new position with “Starpath” at the University of Auckland. He has worked in time series, isotonic inference, generalized linear models, hidden Markov models, mixtures of regressions, and planar point processes. Much of his research effort is focussed on statistical computing, and in recent years he has contributed extensively to the development of the “spatstat” package in R (developed by Adrian Baddeley of the University of Western Australia).