Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2006) 1698 – 1718
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

The analysis of marked point patterns evolving through space and time

Aila Särkkä[a], Eric Renshaw[b,*]

[a]Department of Mathematical Statistics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[b]Department of Statistics and Modelling Science, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH, UK

Received 12 May 2005; received in revised form 6 July 2006; accepted 7 July 2006
Available online 28 July 2006

Abstract

A maximum pseudo-likelihood approach has previously been developed for fitting pairwise interaction models to patterns generated by growth–interaction processes that are sampled at *fixed time* points. This approach is now extended, not only by estimating the parameters of the process *through time*, but also by employing least squares estimation since likelihood based approaches are much more computationally demanding. First, simple stochastic models are used to demonstrate that least squares methods are as powerful as likelihood-based approaches, as well as being mathematically and computationally simpler. The algorithm generates simulations of the deterministic growth–interaction and stochastic immigration–death process, and through these the parameter estimates are determined. Logistic and linear growth are then combined with (symmetric) disc-interaction and (asymmetric) area-interaction processes, and between them these generate a variety of mark–point spatial structures. A robustness study shows that the procedure works well in that the presence, structure and strength of a growth–interaction process can be determined even when an incorrectly presumed model is employed. Thus, the technique is likely to prove to be very useful in general practical applications where the underlying process generating mechanism is almost certain to be unknown. Finally, the procedure is applied to the analysis of a new Swedish pine forest data set for which tree location and diameter at breast height were recorded in 1985, 1990 and 1996.

Keywords: Area interaction; Disc interaction; Growth–interaction process; Least squares estimation; Linear growth; Logistic growth; Robustness; Space–time processes; Swedish pine data

1. Introduction

In recent years there has been a surge of interest in spatio-temporal modelling, due in no small measure to the large amount of data generated by environmental scientists studying pollution and global climate monitoring through geographical information systems, remote sensing platforms, monitoring networks and computer simulation models. As a consequence, substantial progress has been made in the parallel development of methods of analysis involving geostatistical, hierarchical and multivariate time series approaches, together with implementation of space–time dynamic models and point process models.

Geostatistical models (e.g., Kyriakidis and Journel, 1999) are used to study underlying continuous spatial processes which are observed only at a finite set of locations, and typically yield space–time trends of some phenomenon of

* Corresponding author. Tel.: +44 141 548 3591; fax: +44 141 552 2079.
E-mail addresses: aila@math.chalmers.se (A. Särkkä), eric@stams.strath.ac.uk (E. Renshaw).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2006.07.008

interest. Whilst multivariate time series methods (e.g., Bennett, 1999) involve the construction of specific space–time dynamic models. Brown et al. (2001), for example, consider a high-dimensional multivariate state-space time series model in which the cross-covariance structure is derived from the spatial context of the component series; as such its interpretation is essentially independent of the spatial locations at which the data are recorded. Algorithms are developed for estimating the parameters of various models by maximum likelihood, including a non-separable model due to Brown et al. (2000), and applied to a radar-rainfall data set. Both methodologies generate their own specific problems. Geostatistical approaches require complete specification of the joint space–time covariance structure, yet realistic covariance functions can be difficult to specify and implement. Whilst time-series methods specify dynamic models that are linked spatially, and it can be difficult to predict what happens at unmonitored sites because of the lack of a continuous spatial component in the model structure. However, geostatistical and time-series approaches can be combined in a statistical model that is both temporarily dynamic and spatially descriptive. Such *space–time dynamic models* have been considered by, for example, Goodall and Mardia (1994), Guttorp et al. (1994), Mardia et al. (1998), Meiring et al. (1998), and Wikle and Cressie (1999). Goodall and Mardia (1994) also provide an early review of spatio-temporal modelling, and outline an approach based on a general spatio-temporal state space process designed to model the evolution of spatial fields through time. Estimation is recursive, based on the Kalman filter, and is central to Mardia et al. (1998) who combine Kriging (spatial statistics) with the Kalman filter (multivariate time series) and thereby obtain a powerful modelling strategy called the Kriged Kalman filter. Wikle and Cressie (1999) also consider the spatio-temporal Kalman filter, using it to reduce dimension in the analysis of large spatio-temporal data sets. Wikle (2001) describes a parallel approach for modelling complicated dynamical spatio-temporal processes, based on a stochastic integro-difference equation, where the redistribution kernel is allowed to vary with space and/or time. Xu et al. (2005) utilise a similar approach within a hierarchical Bayesian framework. Whilst Stroud et al. (2001) model the mean function of the quantity of interest at each time point as a locally weighted mixture of regression surfaces which are then allowed to evolve through time. MacNab and Dean (2001) propose general additive mixed models for the analysis of geographic and temporal variability of mortality rates. Their aim is to identify temporal trends and produce series of smoothed maps from which spatial patterns of mortality risks can be monitored over time.

In environmental science the processes studied are often highly complex, involving multiple sources of data from a variety of different platforms, as well as various degrees of additional scientific knowledge. Such complexity is exacerbated by the difficulty of employing a full likelihood approach, though this can be ameliorated by using pseudo-likelihood procedures (e.g., Renshaw and Särkkä, 2001 (henceforth denoted as RS)). A related approximation technique involves extending Veccia's (1988) conditional density approach based on partial observations to space–time processes (Jones and Zhang, 1997); the nearest-neighbour conditioning vector is based on preliminary estimates of the space–time correlation function. Rather than specifying joint multivariate spatio-temporal covariance structures it may be much easier to factor the joint distributions into a series of conditional models and then link these together in a hierarchical (usually Bayesian) framework. Wikle (2003) reviews the use of such models, focussing on the spatio-temporal modelling of the intensity surface based on observations at fixed spatial locations. Hierarchical spatio-temporal models are also commonly employed in other disciplines, such as disease mapping. Waller et al. (1997), for example, extend the spatial models of Besag et al. (1991) to accommodate temporal effects, as well as space–time interactions, thereby providing a hierarchical framework for modelling regional disease rates over space and time. Whilst Knorr-Held and Richardson (2003) analyse space–time surveillance of meningoccocal disease for time-series of disease counts. Hierarchical space–time models have also been applied to mapping of the human brain by means of functional magnetic resonance images (Gössl et al., 2001). Lawson (2001) discusses modelling of spatio-temporal clustering in disease and space–time processes in relation to pollution sources and space–time scan statistics. Moreover, descriptive spatio-temporal statistical methods are also highly relevant in climatology (Wikle, 2002). Methods used include: evaluating empirical orthogonal functions; principal oscillation pattern analysis; spatio-temporal canonical correlation analysis; and space–time spectral analysis to infer relationships between spatial and temporal scales of variability.

Now most of the above studies are essentially motivated by the desire to make predictions of inferences about space–time structure based on sampling at fixed locations. In many ecological situations, however, the locations of the measurements play a fundamental role in the process generating mechanism, and so employing fixed location sampling strategies will result in a considerable loss of information. So if both the marks and their point locations develop in time, a far better approach would be to base estimation and inference strategies on spatio-temporal mark–point process