Statistiques Spatiale et Spatio-temporelle *JSTAR'07*

25 et 26 octobre 2007, Amphi J, Campus de Beaulieu, Rennes

PROGRAMME

- Pierre AILLIOT, *Université de Brest* : Space time modeling of precipitation using a hidden Markov model and censored Gaussian distributions
- Gilles GUILLOT, *AgroParisTech (Paris) & Chalmers University* : Statistical modeling in landscape genetics
- Xavier GUYON, *Université de Paris 1* : Variations et identification d'un champ gaussien sur R^2
- Valérie ISHAM, *University College London* : Models for nonstationary spatial-temporal rainfall fields
- Nathalie PEYRARD, *INRA (Toulouse)* : Long-range correlations improve understanding the influence of network structure in disease dynamics
- Olivier PERRIN, *Université de Toulouse 1* : Self-similarity and Lamperti Transformation for Random Fields
- Håvard RUE, *Norwegian Univeristy of Science and Technology (Trondheim)*: Bayesian analysis of latent Gaussian models without MCMC
- Jean-Luc STARCK, *Centre d'Etudes de SACLAY* : Morphological Component Analysis: From Images to Hyperspectral Data
- Rasmus WAAGEPETERSEN, *Aalborg University*: Statistical inference for inhomogeneous Cox processes

(http://www.math.univ-rennes1.fr/irmar/congres/)

Comité d'organisation: Jian-feng YAO (Responsable), David CAUSEUR, Dominique DEHAY, Patrick PÉREZ, Michel PIERRE, Valentin PATILEA

REMERCIEMENTS

Depuis 2004, les statisticiens rennais de l'IRMAR, du CREST/ENSAI et de l'IRISA/INRIA, organisent des Journées de Statistique de Rennes (JSTAR) autour d'un domaine de recherche choisi chaque année.

L'édition 2007 des Journées Statistiques de Rennes porte sur les *Statistiques Spatiale et Spatio-temporelle.*

Nous remercions les organismes suivants pour leur soutien financier à ces journées:

- Région Bretagne ;
- Rennes-Métroplole ;
- Conseil général de l'Ille et Vilaine ;
- Université de Rennes 1 ;
- ENS-Cachan (antenne de Bretagne) ;
- ENSAI ;
- INRIA-Rennes.

Les organisateurs:

Jian-feng YAO,
David CAUSEUR,
Dominique DEHAY,
Patrick PÉREZ,
Michel PIERRE,
Valentin PATILEA

Program

Amphi J, Building 28 (second floor)

Thursday October 25

13:30 - 14:00	Welcome
14:00 - 14:50	Jean-Luc Starck
14:50 - 15:40	Håvard Rue
15:40 - 16:00	Coffee break
16:00 - 16:50	Valérie Isham
16:50 - 17:40	Pierre Ailliot

Friday October 26

9:00 - 9:50	Rasmus Waagepetersen
9:50 - 10:40	Xavier Guyon
10:40 - 11:10	Coffee break
11:10 - 12:00	Olivier Perrin
12:00 - 14:00	Lunch
14:00 - 14:50	Gilles Guillot
14:50 - 15:40	Nathalie Peyrard
15:40 - 16:00	Coffee d'au revoir

- **Internet access:** the computer room is located in the Maths departement, Bulding 23: room 032 just on the right after the entrance.

  Each machine has a name of type: `tp-maths-nn` where `nn` is some number. Then use on this machine

  ```
   login :  569110nn
  passwd :  56911!!!
  ```

  For example, on the machine `tp-maths-49`, the login is `56911049`.

- A free lunch will be served on Friday 26, in the Caféteria, near Amphi J.

Jean-Luc Starck,Commissariat à l'Énergie Atomique (CEA)
Title: Morphological Component Analysis: From Images to Hyperspectral Data

The Morphological Component Analysis (MCA) is a a new method which allows us to separate features contained in an image when these features present different morphological aspects. MCA can be very useful for decomposing images into texture and piecewise smooth (cartoon) parts or for inpainting applications. MCA can also be extended to multichannel data (GMCA) and we show that GMCA is a very interesting alternative to standard blind source separation techniques. Finally we will see how GMCA can be used for different astronomical applications, especially for recovering the Cosmic Microwave Background from multichannel observations.

References:

1. http://jstarck.free.fr/IEEE07_GMCA.pdf
2. http://jstarck.free.fr/IEEE06_MMCA.pdf
3. http://jstarck.free.fr/ACHA05_Inpaint.pdf
4. http://jstarck.free.fr/AIEP04.pdf

Håvard Rue, Norwegian University of Science and Technology
Title: Bayesian analysis of latent Gaussian models without MCMC

In this talk, I will discuss some recent progress of analysing latent Gaussian models without the need for MCMC. This class of models is frequently used to analyse both count and survival data, both temporal and spatio-temporal data. We approach the problem of computing posterior marginals using integrated nested Laplace approximations, and obtain near instant inference and at the same time increased accuracy. I will also discuss how these approximations can be used to validate the model under study using cross-validation like measures, and how we can take advange of modern computer architecture to really speed up the computations.

The methodological part of this talk is described in the following report:

`http://www.math.ntnu.no/ hrue/reports/RueMartinoChopin_submission.pdf`

The remaining part is work in progress.

Valérie ISHAM, University College London
Title: Models for nonstationary spatial-temporal rainfall fields

For accurate determination of flood risk, hydrological rainfall-runoff models require long records of rainfall data at a much higher temporal and spatial resolution than are generally available, and this creates a need for models that can be simulated to provide artificial data. In this talk I will describe a programme of research carried out by an interdisciplinary team of statisticians from University College London and hydrologists from Imperial College, London. In this work we combine stochastic-mechanistic spatial-temporal models that represent the physical rainfall process, albeit in a highly idealised fashion, and generalised linear statistical models (GLMs). The former provide the high resolutions required, but are both temporally and spatially stationary by construction. For larger catchments it is important to allow for spatial heterogeneity, while for long-term design purposes it is necessary to allow adequately for temporal nonstationarities including possible climate change scenarios. Such nonstationarities can easily be incorporated in GLMs. The construction, fitting and validation of both types of model, together with work in progress on combining their strengths in a nonstationary continuous simulation, will be described.

References:

1. Chandler, R.E., Isham, V.S., Bellone, E., Yang, C. and Northrop, P.J. (2007) Space-time modelling of rainfall for continuous simulation. In Statistical Methods for Spatio-temporal Systems. Eds B Finkenstadt, L Held and V Isham. Boca Raton: Chapman and Hall/CRC, 177–215.

Pierre AILLIOT, Université de Brest
Title: Space time modeling of precipitation using a hidden Markov model and censored Gaussian distributions

The purpose of this talk is to develop a new space-time model for daily precipitation over mesoscale spatial areas. Such models have important applications. For example, they can be used as stochastic rainfall generators to provide realistic inputs to flooding, runoff and crop growth models, and also as components within general circulation models. A variety of stochastic models have been proposed in the literature, and among them the weather type models play an important role. The basic idea of these models consists of introducing an extra variable to describe the meteorological regime (weather type), and assume that this variable explains most of the space-time structure of the data.

In the model proposed in this talk, the weather types is introduced as a hidden Markov chain, and precipitation within weather types are described using censored power-transformed Gaussian distributions. The latter provide flexible and interpretable multivariate models for the mixed discrete-continuous variables that describe both precipitation, when it occurs, and no precipitation. Finally, the proposed model is a hidden Markov model, in which the hidden process has a discrete component (the weather type) and a continuous component (due to the censoring).

The parameters will be estimated using a Monte-Carlo algorithm and the fitted model will then be validated on rainfall data from a small network of stations in New Zealand encompassing a diverse range of orographic effects. We will show that it provides a better description of the spatial structure of precipitation than a more conventional HMM commonly used in the literature.

Rasmus Waagepetersen, Aalborg University
Title: Statistical inference for inhomogeneous Cox processes

Inhomogeneous spatial point processes find many applications in biology. In forest ecology for example, it may be of interest to study the relation between the intensity of locations of trees and environmental covariates describing topography and soil properties. A first choice of model would be an inhomogeneous Poisson process with intensity function given e.g. by a log linear model depending on the covariates. In practice, however, data are often clustered relative to an inhomogeneous Poisson process and more appropriate models are then Cox or cluster point processes.

For many common types of inhomogeneous Cox processes, the first and second order summary statistics (the intensity function and the K-function) have closed form expressions and this enables the construction of simple estimating functions for parameter estimation. We review such estimating functions and discuss how asymptotic properties of the parameter estimates may be obtained.

Likelihood-based inference for Cox processes is more complicated as the likelihood function can not be obtained in closed form. Using Markov chain Monte Carlo methods it is, however, possible to approximate the likelihood function and its derivatives. We discuss how simulation-based likelihood inference may be implemented in relation to a specific data example regarding positions of Minke whales in the North Atlantic.

References:

1. Møller, J. and Waagepetersen, R. (2007) Modern statistics for spatial point processes, Scandinavian Journal of Statistics, to appear.

2. Waagepetersen, R. (2007). An estimating function approach to inference for inhomogeneous Neyman-Scott processes, Biometrics, 63, 252-258.

3. Waagepetersen, R. (2007). Estimating functions for inhomogeneous spatial point processes with incomplete covariate data, submitted.

4. Waagepetersen, R. (2006). Estimating functions for inhomogeneous Cox processes. Proceedings of S4G, the 6th International conference on Stereology, Spatial Statistics and Stochastic Geometry, Prague, Czech Republic, June 26-29, 2006, pages 149-158.

Preprints are available at `http://www.math.aau.dk/~rw/publications.html`

Xavier GUYON, Université de Paris I
Title: Variations et identification d'un champ gaussien sur $\mathbb{R}^2$

La variation quadratique (VQ) d'un processus X indéxé par $\mathbb{R}$ est la limite, lorsqu'elle existe, de la suite $\{VQ_n(X) = \sum_1^n [X(\frac{i}{n}) - X(\frac{i-1}{n})]^2\}$ des sommes de carrés des accroissements linéaires de X sur une partition de plus en plus fine de $[0,1]$. Cette variation, classiquement étudiée si X est un processus gaussien ou une semi-martingale, permet d'identifier *p.s.* une fonctionnelle $\Psi(\theta)$ de θ, paramètre qui caractérise la loi de X. Par exemple, si X est le processus gaussien de Ornstein-Uhlenbeck de covariance $r(t) = \sigma^2 \exp -\alpha\,|t|$, $\alpha > 0$, cette VQ permet l'identification *p.s.* du produit $\sigma^2 \times \alpha$. Seule l'observation de X sur la partie bornée $[0,1]$ est nécessaire et le cadre est celui d'une asymptotique intensive. Mais, en dimension 1, là s'arrête la possibilité d'identification de θ.

L'objectif de cet exposé est de montrer que si X est un champ réel observé sur une partie continue bornée D de $\mathbb{R}^2$, l'outillage des variations quadratiques est beaucoup plus riche : variations curvilignes, variations superficielles, variations produits. De plus, ces variations peuvent dépendre d'un paramètre λ à notre disposition permettant d'enrichir la famille des fonctionnelles $\{\Psi_\lambda(\theta)\}$ utiles à l'identification. Si X est un champ gaussien, nous présentons des modèles où l'identification *p.s.* du paramètre θ de X via un jeu de ces VQ est possible : par exemple, pour le champ de Ornstein-Uhlenbeck de covariance $r(s,t) = \sigma^2 \exp -\{\alpha\,|s| + \beta\,|t|\}$, $\alpha, \beta > 0$, deux VQ linéaires et une VQ superficielle permettent d'identifier les 3 paramètres α, β et σ^2. D'autres exemples sont présentés.

Mots clés : variation quadratique, variation superficielle, champ gaussien, identification de paramètre, asymptotique intensive, déformation d'un champ stationnaire, drap brownien fractionnaire.

References:

1. GUYON X. [1987], Variations des champs gaussiens stationnaires : application à l'identification, *Proba. Th. Rel. Fields*, 75, 179-193.

2. GUYON X. ET PERRIN O. [2000], Identification of space defpormation using linear and superficial quadratic variations, *Stat. and Proba. Letters*, 47, 307-316.

3. COHEN S., GUYON X., PERRIN O. ET PONTIER M. [2004], Identification of an isometric transformation of a standard brownian sheet, *Jour. Stat. Plan. Inf.*, 136, 1317-1330.

4. COHEN S., GUYON X., PERRIN O. ET PONTIER M. [2006], Singularity functions for fractional processes : application to the fractional brownian shett, *Ann. I.H.P.*, 42, 187-205.

Olivier Perrin, Université de Toulouse 1
Title: Self-similarity and Lamperti Transformation for Random Fields

We define multi-self-similar random fields, that is, random fields that are self-similar component-wise. We characterize them, relate them to stationary random fields using a Lamperti-type transformation and study these stationary fields. We also extend the notions of local stationarity and local stationarity reducibility to random fields. Our work is motivated by applications arising from climatological and environmental sciences. We illustrate these new concepts with the fractional Brownian sheet and the Lévy fractional Brownian random field.

References:

1. Genton, M. G., Perrin, O., Taqqu, M. (2007): "Self-similarity and Lamperti transformation for random fields", Stochastic Models, Vol. 23, 397-411.

Gilles Guillot, AgroParisTech & Chalmers University
Title: Statistical modeling in landscape genetics

Recent developments in molecular markers and computational methods in statistics has led to the emergence of a new scientific field, landscape genetics, which is an amalgamation of population genetics and landscape ecology. This discipline aims to provide information on how environmental features influence gene flow, population structure and local adaptation.

To address some of these questions, ecologists collect skin samples of animals and genotype them at a few highly variable molecular markers. We proposed to analyze such geo-referenced molecular markers within different statistical models corresponding to different sets of assumptions about the dispersal of the animals of the species and their mating behaviors.

These models typically combine a spatial component based on classical spatial statistical models such as random fields and random tessellations, and a genetic component based on explicit biological assumptions about mating and dispersal process. They allow to make inference about the territories of populations and quantitative parameters about the dispersion of the animals. This requires heavy MCMC computations for inference. Models and algorihtms will be discussed with results from simulated and real data.

References:

1. G. Guillot, F. Santos, A. Estoup. Inference with Geneland, a sensitivity analysis to null alleles, isolation-by-distance and spatial sampling scheme. Submitted.

2. S. Ancelet, G. Guillot and O.Francois Hidden Markov random fields and the genetic structure of the Scandinavian brown bear population,. Journal de la Societe Francaise de Statistique, sous-presse, 2007.

3. O. François, S. Ancelet, G. Guillot. Bayesian Clustering using Hidden Markov Random Fields in Spatial Population Genetics. Genetics, 174: 805-816, 2006.

4. A. Coulon, Guillot G., Cosson J.-F., Angibault J.M.A., Aulagnier S. , Cargnelutti B., Galan M., Hewison A.J.M. Genetics structure is influenced by lansdcape features. Empirical evidence from a roe deer population. Molecular Ecology 15 1669-1679, 2006.

5. G. Guillot, Mortier, F., Estoup, A. Geneland : A program for landscape genetics. Molecular Ecology Notes, 5, 712-715, 2005.

6. G. Guillot, Estoup, A., Mortier, F. Cosson, J.F. A spatial statistical model for landscape genetics. Genetics, 170, 1261-1280, 2005

Nathalie Peyrard, Institut National de Recherche Agronomique (Toulouse)
Title: Long-range correlations improve understanding the influence of network structure in disease dynamics

Models of infectious diseases not conferring permanent immunity are characterized by a phase transition between extinction and persistence. A challenge in contemporary epidemiology is to understand how the geometry of a host's interaction network influences disease dynamics close to the critical point of such a transition. Here we address this challenge with the help of moment closures. Traditional moment closures, however, do not provide satisfactory predictions close to such critical points. We therefore introduce a new method for incorporating longer-range correlations into existing closures. Our method is technically simple, remains computationally tractable, and significantly improves the approximation's performance. Our extended closures thus provide an innovative tool for quantifying the influence of interaction networks on spatially or socially structured disease dynamics. In particular, we examine the effects of a network's clustering coefficient, as well as of new geometric measures, such as a network's square clustering coefficient. We compare the relative performance of different closures from the literature, with or without our long-range extension. In this way, we demonstrate that the normalized version of the Bethe approximation - extended to incorporate long-range correlations according to our method - is an especially good candidate for studying influences of network structure. Our numerical results highlight the importance of the clustering coefficient and the square clustering coefficient for predicting disease dynamics at low and intermediate values of transmission rate, and demonstrate the significance of path redundancy for disease persistence.

References:

1. Nathalie Peyrard, Alain Franc and Ulf Diekmann. Incorporating long-range correlations into second-order moment closures better approximates low-densities dynamics on a graph. Available as
`http://www.avignon.inra.fr/content/download/4814/77863/file/RR2006_18.pdf`

2. M. E. J. Newman. The structure and function of complex networks. SIAM Review 45, 167-256 (2003).
Available also as `http://arxiv.org/abs/cond-mat/0303516`