Bayesian inference for independent cluster point processes

M.N.M. van Lieshout[1,2], **C. Lu**[1]

[1] Stochastics, Centrum Wiskunde & Informatica, The Netherlands
[2] Mathematics of Operations Research, University of Twente, The Netherlands
`c.lu@utwente.nl`

Spatial and spatio-temporal cluster point processes have been extensively studied, with much of the literature assuming Poisson structures for parent, child, or both processes due to their straightforward probability density functions. Frequentist and Bayesian approaches have been developed to infer such models. However, in practice, assuming a Poisson parent process often results in an overabundance of clusters, while assuming a Poisson child process may not adequately capture real-world complexities, as illustrated by the arson fire data example. In this paper, we propose a two-step Bayesian inference framework for independent cluster point processes that are not restricted by Poisson assumptions. Specifically for the fire data, we introduce a model incorporating a repulsive prior and a shifted-Poisson Gaussian child. We develop a Markov chain Monte Carlo method to estimate cluster states given fixed model parameters and a Monte Carlo Expectation-maximization method to estimate the model parameters themselves. We show the convergence of the proposed approach and validate its effectiveness through a simulation study and the application to the real fire data. The Bayesian framework we present is flexible and can accommodate a variety of parent and child processes provided that their probability functions have an analytical form. Furthermore, from a practical perspective, our approach enables to predict the probability of being-forming clusters in a spatio-temporal manner.

Bayesian algorithms for parameter estimation for Gibbs point processes based on partly missing data

N. Gillot[1], R. S. Stoica[1], A. Särkkä[3], D. Gemmerlé[2]

[1] Université de Lorraine, CNRS, Inria, IECL, F-54500 Nancy France
[2] Université de Lorraine, CNRS, IECL, F-54500 Nancy France
[3] Department of Mathematical Sciences, Chalmers University of Technology & the University of Gothenburg, Sweden
`nathan.gillot@univ-lorraine.fr`

In order to characterise galactic patterns, new models and inference procedures for point processes with interaction were introduced in [1, 2]. In this talk, we will present some further developments of this work.

Here, we concentrate on the special case where the process takes place on the bounded region $W = A \cup B$ but it can be observed only in the region A. In this situation, the likelihood of the considered point process cannot be derived from the available observations, unless simulated data is produced via Markov Chain Monte Carlo (MCMC) procedures in the region B, where direct observations are not available. This operation increases the general computational cost while the convexity of the likelihood can not be guaranteed.

In order to overcome this drawback, we propose to use an Approximate Bayesian Computation (ABC) framework in order to deal with this problem.

The ABC framework we propose allows a theoretical construction of Metropolis-Hastings dynamics that samples from the joint distribution of the un-observed pattern and the parameter of interest, conditionally on the observed data. The theoretical properties of the proposed dynamics enable the construction of different approximate algorithms that exhibit convergence properties. The proposed method is applied to simulated and real data. Conclusions and perspectives are depicted.

For a review of MCMC and ABC inference methods for spatial patterns characterisation through Gibbs point processes with interactions, and references, we recommend [3].

References

[1] N. Gillot, R. S. Stoica, D. Gemmerlé. *Study the galaxy distribution characterisation via bayesian statistical learning of spatial marked point processes.* In Proceedings RING. Preprint HAL, hal-04163649, 2023.

[2] N. Gillot, R. S. Stoica, A. Särkkä, D. Gemmerlé. *Spatial point proces modelling and bayesian inference for large data sets.* . In Proceedings RING. Preprint HAL, hal-04645186, 2024.

[3] R. S. Stoica. *Random Patterns and Structures in Spatial Data.* Chapman and Hall, 2025.

Bayesian Inference for Neyman-Scott Point Processes with Anisotropic Clusters

J. Dvořák[1], **E. Ewers**[2], T. Mrkvička[3], C. Redenbach[2]

[1] Faculty of Mathematics and Physics, Charles University, Czech Republic
[2] Department of Mathematics, RPTU University Kaiserslautern-Landau, Germany
[3] Faculty of Agriculture and Technology, University of South Bohemia, Czech Republic
`emily.ewers@rptu.de`

There are few inference methods available to accommodate covariate-dependent anisotropy in point processes. To address this, we propose an extended Bayesian MCMC approach for Neyman-Scott cluster processes. We focus on anisotropy and inhomogeneity in the offspring distribution. Our approach provides parameter estimates as well as significance tests for the covariates and anisotropy through credible intervals which come with the posterior distributions. Additionally, it is possible to check the assumption of constant direction of clusters or constant elongation of the ellipse. We demonstrate the effectiveness of this approach through a simulation study for a Thomas process and an application to real-world data.

Combining Efforts: A modelling framework for Integrating Citizen Science and Survey Data using INLA

J. Belmont[1], S. Martino[2], F. Lindgren[3], H. Rue[3],J. Illian[1]

[1] School of Mathematics & Statistics, University of Glasgow, Glasgow, UK
[2] Department of Mathematical Sciences, NTNU, Trondheim, Norway
[3] School of Mathematics, University of Edinburgh, Edinburgh, UK
[4] King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
`Jafet.BelmontOsuna@glasgow.ac.uk`

Over the last decade, opportunistic data collected through citizen science (CS) programs have proven to be a valuable and cost-effective resource for monitoring wildlife populations. However, the absence of standardized sampling protocols in these initiatives introduces analytical challenges, as data collection is often influenced by observer preferences, accessibility, and areas of public interest. In contrast, standardized surveys are carefully structured to minimize these biases but provide much lower spatiotemporal coverage. Integrating these two valuable data sources could enhance the accuracy and quality of species distribution predictions. Yet, reliably combining them remains challenging due to differences in study design, spatial and temporal coverage, and potential sampling biases. In this work, we introduce a new approach to integrating opportunistic citizen science (CS) observations with data from planned surveys while accounting for observational errors inherent in both sources. The proposed integrated model is fitted using the INLA framework to jointly analyse the geographical locations where a species has been reported by participating volunteers and the detection/non-detection data from robust designed surveys. By integrating CS data with survey data using robust modelling techniques, we can enhance our understanding of species distributions.

Complex observation processes - animal movement modelling and software

J. Illian[1]

[1] School of Mathematics and Statistics, University of Glasgow, Scotland
`janine.illian@glasgow.ac.uk`

Data on the movement of animals in space and time are collected in a number of contexts, including ecological research, animal conservation, agriculture and wildlife management. These data, in combination with an appropriate statistical analysis approach, can provide insight into the animals' behaviour, for example regarding habitat use, migration routes, or responses to environmental change.

Movement data are collected using a variety of technologies , resulting in different types of observation processes, each with their respective advantages and disadvantages. I particular, the data never provide information on all elements of the animals' movement behaviour and statistical methodology needs to account for the observation processes and potentially infer the information not directly contained in the data from the information contained in the data.

The talk discusses two different data collection approaches of animal movement data — telemetry data of the individual location of animals and data derived from passive acoustic receiver arrays. We discuss modelling approaches through case studies, focussing also on software implementations, in particular `inlabru` and `patter`.

References

[1] A. Baddeley, E. Rubak, and R. Turner. *Spatial point patterns: methodology and applications with R*, volume 1. CRC press Boca Raton, 2016.

Building spatial population dynamics models using the SPDE approach

L. Blackman[1], J. Belmont Osuna[1], P. Bouchet[2], E Jones[2], F. Lindgren[3], J. Illian[1]

[1] School of Mathematics and Statistics, University of Glasgow, UK
[2] Biomathematics and Statistics Scotland, UK
[3] School of Mathematics, The University of Edinburgh, UK
`2966362b@student.gla.ac.uk`

There are currently limited options for spatial population dynamics models that can effectively utilize ecological survey data, with most mechanistic models requiring detailed demographic data. To address this gap, we propose integrating biological differential equations into the Integrated Nested Laplace Approximations (INLA) framework using the Stochastic Partial Differential Equation (SPDE) approach.
By extending the iterative INLA method for estimating the parameters of an SPDE [1], we aim to incorporate the logistic growth equation as a spatio-temporal random effect in an inhomogeneous Poisson Process. This model could estimate and predict demographic processes from count data, and is a significant step towards combining mechanistic and statistical population models. Future extensions could include more complex and biologically realistic dynamics, such as multispecies systems using the Lotka-Volterra equations.

References

[1] R. Anderka, M. Deisenroth, and S. Takao. *Iterated INLA for state and parameter estimation in nonlinear dynamical systems.* arXiv preprint arXiv:2402.17036, 2024.

Comparison of resampling methods in the case of small samples

Salme Kärkkäinen, Lasse Moisio, Jesse Sorvali

Department of Mathematics and Statistics, University of Jyväskylä
`salme.karkkainen@jyu.fi`

In the current on-going research, we evaluated resampling methods in the case of classification methods. In the literature often, the so-called repeated hold-out method with division 50/50 is used in practice. Our goal is therefore to compare its performance with other resampling methods in the case of smaller sample sizes. We compare an expected generalization error (Err) with error rates estimated by resampling methods.

The comparison is performed by the simulation experiments based on the characteristics of benthic macroinvertebrate image data. From the real data, we formed three set-ups: 2 largest classes and 2 best features, 8 largest classes and 8 best features, and 50 classes and 8 best features. In simulation, we used priors and class-specific multivariate Gaussian distributions estimated from the real data and simulated 100 data sets for each set-up. This leads to the natural choice of quadratic discriminant analysis (QDA) to be used as a reference classifier. We compare QDA with linear discriminant analysis (LDA), K-nearest neighbour (KNN), neural networks (NN), random forests (RF), and support vector machines with radial and linear kernel (SVMR, SVML). With the single K-fold cross-validation (CV), the values K=2, 3, 5,10, N, and with repeated CV, K=5,10, and 5 repetitions were applied. For the repeated hold-out, 50%, 20%, and 10% of the test set and 50 repetitions and for the bootstrap, 50 bootstrap samples were used.

When compared with Errs of all set-ups and classifiers, we found that the most biased error rates are typically obtained by repeated hold-out with division 50/50, a leave-one out bootstrap estimate, and in some cases by the bootstrap estimator 0.632. When considering the least biased error rates of Errs over smallest data sets of three set-ups and all classifiers, we found most often repeated hold-out with division 90/10, N-CV, the bootstrap estimator 0.632 and repeated 10-CV.

Impacts of Spatiotemporal Scale on Habitat Preference: Modelling Cattle Movement with GF-iSSA

M. R. Morton[1,2], M. K. Schneider[1], C. M. Pauler[1], J. B. Illian[2]

[1] Agroscope, Grazing Systems, 8046 Zürich, Switzerland
[2] School of Mathematics and Statistics, University of Glasgow, Glasgow, G12 8QQ, Scotland
`m.morton.2@research.gla.ac.uk`

Telemetry tracking of animal movement has become a standard practice in some agricultural settings, with aims of optimising land-use to improve economic and ecological sustainability, as well as animal welfare. With technological advancement in both data collection and analysis, the spatial scale and temporal frequency of environmental and tracking data have become increasingly fine. However, technical limitations still necessitate trade-offs, prompting the questions: what is the appropriate resolution for the system and associated research aims? How do inferences change with scale?

In this talk, we investigate these questions in the context of habitat preferences of cattle in Swiss mountain rangelands. We compare two modelling approaches using the INLA framework for approximate Bayesian inference with R package inlabru: a Log Gaussian Cox Process (LGCP) model of the overall distribution of movement steps, and Gaussian Field Integrated Step Selection Analysis (GF-iSSA). GF-iSSA is a recently introduced method for modelling highly spatiotemporally correlated tracking data in continuous space [1]. It relies on a complex integration scheme to enable computationally efficient modelling and account for restrictions in animal mobility. Here, we explore the features and limitations of this approach when inferring habitat preference from tracking data collected in an agricultural setting with heterogeneous terrain.

References

[1] R. Arce Guillen, F. Lindgren, S. Muff, T.W. Glass, G.A. Breed, and U.E. Schlägel. *Accounting for unobserved spatial variation in step selection analyses of animal movement via spatial random effects.* Methods in Ecology and Evolution, 14 (2023) 2639–2653. https://doi.org/10.1111/2041-210x.14208.

Fractals and Point patterns: Detecting Glaucoma

J. Henderson[1], B.M. Davis[2], **H.J. Mitchell**[1]
[1] Mathematical Sciences Research Centre, Queen's University Belfast, United Kingdom
[2] UKRI-STFC Central Laser Facility, Rutherford Appleton Laboratory, United Kingdom
`h.mitchell@qub.ac.uk`

Glaucoma is a complex neurodegenerative disease and the leading cause of irreversible blindness globally [1]. The disease results in the degeneration of retinal ganglion cells (RGCs) which ultimately leads to blindness. Currently there is no cure but there are several effective treatments which aim to prevent further loss of RGCs. Early diagnosis is therefore important but it is complicated due to the high plasticity of the central nervous system and prolonged diagnosis periods, during which subtle changes in disease progression may go undetected. However, recent advances in adaptive-optics imaging is enabling clinicians to visualise subsets of the RGC population at the mircroscopic resolution. The images obtained of the RGC population can be treated as point patterns. This research focuses on the use of point pattern techniques and fractal measures [2] applied to these RGC point patterns to detect disease-related changes.

References

[1] B.M. Davis, L. Crawley, M. Pahlitzsch, F. Javaid and M.F. Cordeiro. *Glaucoma: the retina and beyond.* Acta neuropathologica. 2016.

[2] F.P. Agterberg. *Fractals and spatial statistics of point patterns.* Journal of Earth Science, 2013.

Creating non-reversible rejection-free samplers by rebalancing skew-balanced Markov jump processes

E. Jansson[1], M. Schauer[1], **R. Seyer**[1], A. Sharma[1]

[1] Department of Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, Sweden
`rubense@chalmers.se`

Markov chain sampling methods form the backbone of modern computational statistics. However, many popular methods are prone to random walk behavior, i.e., diffusion-like exploration of the sample space, leading to slow mixing that requires intricate tuning to alleviate. Non-reversible samplers can resolve some of these issues. We introduce a device that turns jump processes that satisfy a skew-detailed balance condition for a reference measure into a process that samples a target measure that is absolutely continuous with respect to the reference measure. This sampler is rejection-free, non-reversible and continuous-time. As an example, we apply the device to Hamiltonian dynamics discretized by the leapfrog integrator, resulting in a rejection-free non-reversible continuous-time version of Hamiltonian Monte Carlo (HMC). We prove the geometric ergodicity of the resulting sampler under certain convexity conditions, and demonstrate its qualitatively different behavior to HMC through numerical examples.

Cox–Ingersoll–Ross Models for Dynamics on Linear Networks

Ali Sharifi[1], Mehdi Moradi[1]

[1] Department of Mathematics and Mathematical Statistics, Umeå University, Sweden

`ali.sharifi@umu.se; mehdi.moradi@umu.se`

We present an extension of the classical Cox–Ingersoll–Ross (CIR) diffusion that is tailored to phenomena evolving along linear networks such as roadways, power lines or river systems. By embedding the square-root process in a one-dimensional spatial framework, the proposed model captures the state variable's intrinsic local dynamics and the spatial heterogeneity that characterises network-constrained environments. We establish the existence and uniqueness of the process, show that all relevant moments remain uniformly bounded, and prove a strong Hölder-type regularity for its sample paths. For inference, we build a likelihood-based procedure on an Euler–Maruyama discretisation and demonstrate that the resulting estimators are strongly consistent under mild moment conditions.

To accommodate abrupt regime changes that frequently occur in practice, we further introduce a Markov-modulated CIR (MMCIR) formulation in which the drift and volatility parameters switch according to an underlying finite-state Markov chain. Closed-form updates for the maximum-likelihood estimators are obtained via an EM algorithm, and their almost-sure convergence is established. Simulation studies confirm the theoretical findings, while applying traffic-flow data on an urban road network illustrates the practical improvement in predictive performance over classical homogeneous diffusions. The proposed framework thus provides a flexible and statistically sound tool for analysing stochastic dynamics that are confined to linear network structures.

Validating methods for uncertainty propagation in two-stage Bayesian models - with application on linking climate and dengue in the Philippines

S. J. Villejo[1,3], S. Martino[2], J. Illian[1]

[1] School of Mathematics and Statistics, University of Glasgow, UK
[2] Department of Mathematical Sciences, Norwegian University of Science and Technology, Norway
[3] School of Statistics, University of the Philippines Diliman, Philippines
`stephen.villejo@glasgow.ac.uk`

A two-stage modeling framework has the advantage of being more computationally efficient than a joint modelling approach when the first-stage model is already complex in itself, and avoids the potential problem of unwanted feedback effects [?, ?]. Two common ways of doing two-stage modeling are the crude plug-in method and the posterior sampling method. The former ignores the uncertainty in the first-stage model, while the latter can be computationally expensive. This work validates the correctness of the two aforementioned approaches by testing the self-consistency property of Bayesian models, via the simulation-based calibration method [?]. We then show a data application on linking climate and dengue fever, which is an infectious disease caused by dengue arbovirus and commonly transmitted by *Aedes* mosquitoes [?]. The first stage fits the climate models, and the second stage fits the health model with dengue incidence as response and the climate predictions from the first stage as inputs. We postulate a Bayesian spatio-temporal model and use the integrated nested Laplace approximation (INLA) approach for inference [?]. Results show that temperature is positively related with dengue, but very hot conditions tend to have negative impact on dengue incidence. Moreover, the relationship between rainfall and dengue varies in space, depending on the climate type of the area. For areas with uniform and low variation in the amount of rainfall all year round, rainfall is negatively associated with dengue, while for areas with pronounced dry and wet season, rainfall is positively related with dengue.

A magnitude-weighted model evaluation score for earthquake forecasting

J. Jansson Valter[1], A. Arjon[2], F. Schoenberg[2]
[1] Department of Mathematical Sciences,
Chalmers University of Technology & Unversity of Gothenburg, Sweden
[2] Department of Statistics & Data Science, UCLA, Los Angeles, USA
`juljans@chalmers.se`

Earthquakes can be modelled with spatio-temporal point processes, such as the Epidemic-Type-Aftershock-Sequence (ETAS) model. These models are great at predicting aftershocks, that is smaller earthquakes following a larger earthquake, but they are not much better than a simple homogeneous Poisson model at predicting large earthquakes [1]. In seismologists' efforts to anticipate future earthquakes, the ability accurately to forecast the largest earthquakes is of primary importance, not only for short-term response such as preparation of emergency personnel and disaster relief, but also for longer-term preparation in the form of building codes, urban planning, and earthquake insurance.

The task of forecasting large earthquakes is difficult, partly because the events are rare, and some even argue that earthquakes are inherently unpredictable [2]. Another caveat is that current model evaluation scores, like the log-likelihood score, does not take earthquake magnitude into account when assessing the model, since it is not weighting the larger magnitudes any more heavily than the smaller ones [1]. To combat this, a new model evaluation score called the Q statistic, which rewards models for accurately estimating the largest 5% of earthquakes, was recently introduced [1]. The goal is that, using the Q statistic, new models can be found in the future which are better at forecasting large earthquakes.

In this talk, we present the Q statistic, giving an overview of its statistical properties in the context of earthquake modelling using spatio-temporal point processes. Moreover, the Q statistic is evaluated for several different earthquake models, fitted on the California earthquake data from years 2012-2017.

References

[1] Frederic Schoenberg and Danijel Schorlemmer. Critical questions about csep, in the spirit of dave, yan, and ilya. *Seismological Research Letters*, 95(6):3617–3625, 2024.

[2] Robert J. Geller, David D. Jackson, Yan Y. Kagan, and Francesco Mulargia. Earthquakes cannot be predicted. *Science*, 275(5306):1616–1616, 1997.

Efficient non-parametric estimation of variable productivity Hawkes processes

S. Phillips[1], F. Schoenberg[1]

[1] Department of Statistics and Data Science, University of California, Los Angeles
`sophiep@ucla.edu`

The Hawkes self-exciting point process provides a natural model for clustered event data, where events generate offspring according to a productivity function. While the original Hawkes model assumed constant productivity across points [1], allowing the productivity to vary as a function of time or other covariates is useful in applications such as seismology and epidemiology. We compare methods for estimating a time-varying productivity in Hawkes models and demonstrate applications to infectious disease modeling. Based on the analytic maximum likelihood estimator proposed in [4], we introduce a regularized estimator for pointwise productivities, along with a closed-form least-squares estimator that enables fast and stable performance for binned data [2]. We apply these methods to U.S. Covid-19 case data from 2020–2022 to estimate the effective reproduction rate R_t and forecast outbreaks in incidence [3].

References

[1] A. Hawkes. Spectra of some self-exciting and mutually exciting point processes *Biometrika*, 58(1), 1971. DOI: 10.1093/biomet/58.1.83.

[2] S. Phillips and F. Schoenberg. Efficient non-parametric estimation of variable productivity Hawkes processes. *Journal of Applied Statistics*, 1–18, 2024. DOI: 10.1080/02664763.2024.2426019.

[3] S. Phillips, G. Mohler, and F. Schoenberg. Detection of surges of SARS-Cov-2 using nonparametric Hawkes models. *Epidemics*, accepted 25 Mar. 2025.

[4] F. Schoenberg. Nonparametric estimation of variable productivity Hawkes processes. *Environmetrics*, 33(6), 2022. DOI: 10.1002/env.2747.

Tessellated spatial Poisson point process models

N. D'Angelo[1]

[1] Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy
`nicoletta.dangelo@unipa.it`

Recent developments in local models for point processes have highlighted the need for flexible methodologies to capture spatial heterogeneity [2, 6, 4]. In this work, we introduce a novel framework that extends segmented regression models to spatial contexts. Our approach is based on a two-step procedure: first, a spatial segmentation algorithm [1] in employed to geographically weighted regression estimates. The resulting subdomains delineate regions within the study area where constant parameters can be assumed. Subsequently, log-linear Poisson models are fitted within the identified non-overlapping regions, to estimate region-specific parameters, enabling the application of classical inferential procedures, such as hypothesis testing on regression parameters. This framework supports multiple applications, functioning either as a classic segmented regression model — where one or more covariates may exhibit distinct segmentation patterns — or as a local spatial modeling tool. The latter assumes the presence of a tessellation of the space, partitioning it into regions where "global" log-linear models provide valid descriptions of the underlying point process. Furthermore, it can be conceptually extendable to the geostatistical framework. To validate the proposed methodology, we conduct simulation studies and apply our proposal to two case studies with the fire and seismic data analysed in [5] and [3].

References

[1] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282.

[2] Baddeley, A. (2017). Local composite likelihood for spatial point processes. *Spatial Statistics*, 22:261–295.

[3] D'Angelo, N., Siino, M., D'Alessandro, A. and Adelfio, G. (2022). Local spatial log-Gaussian Cox processes for seismic data. *AStA Adv Stat Anal*, 106:633–671.

[4] D'Angelo, N., Adelfio, G., and Mateu, J. (2023). Locally weighted minimum contrast estimation for spatio-temporal log-gaussian Cox processes. *Computational Statistics & Data Analysis*, 180:107679.

[5] D'Angelo, N., Albano, A., Gilardi, A. and Adelfio, G. (2025). Non-separable spatio-temporal Poisson point process models for fire occurrences. *Environ Ecol Stat*, 32:347–381.

[6] Siino, M., Adelfio, G., Mateu, J. and Chiodi, M. (2017). Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity. *Stoch Environ Res Risk Assess*, 31:1633–1648.

Demystifying spatial confounding

E. Dupont[1], I. Marques[2], T. Kneib[3]
[1] Department of Mathematical Sciences, University of Bath, U.K.
[2] Department of Statistics, The Ohio State University, U.S.A.
[3] Chair of Statistics and Campus Institute Data Science, University of Göttingen, Germany
`eahd20@bath.ac.uk`

In spatial regression models, spatial random effects are included to approximate unmeasured spatial variation in the response variable. However, as spatial random effects are typically not independent of the covariates in the model, this can lead to significant bias in covariate effect estimates of interest. This fundamental problem, which makes covariate effect estimation in any spatial regression potentially unreliable, is referred to as spatial confounding. In recent years, there has been much interest in spatial confounding, not least because the most established methods for dealing with the problem were proven to be ineffective. However, research into the topic has sometimes led to puzzling and seemingly contradictory results. Here, we develop a broad theoretical framework that brings mathematical clarity to the mechanisms of spatial confounding, providing explicit and interpretable analytical expressions for the resulting bias. From these, we see that it is a problem directly linked to spatial smoothing, and we can identify exactly how the features of the model and the data generation process affect the size and occurrence of bias. Our framework can be used to understand and generalise existing results on spatial confounding, including dependency on spatial frequency information and suggested methods for bias adjustment.

References

[1] E. Dupont, I. Marques, I., and T. Kneib. *Demystifying spatial confounding.* arXiv:2309.16861, 2023.

Modeling marked Poisson point processes with real-valued continuous marks

M. Tarantino[1], O. Cronie[2], N. D'Angelo[1], G. Adelfio[1]

[1] Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy
[1] Department of Mathematical Sciences, Chalmers University of Technology & Uinversity of Gothenburg, Sweden
`marco.tarantino01@unipa.it`

In the analysis of spatial point patterns with associated real-valued marks, incorporating the influence of marks into the estimate of the intensity function can lead to a more accurate understanding of the underlying process. Standard models either rely on strong distributional assumptions about the marks to incorporate their effect on the estimated intensity [1], or exclude them from the fitting procedure, studying the marks through marked summary statistics to understand their relationship with the spatial distribution of the points [2]. In this article, we address this issue by proposing two approaches, one parametric and one semi-parametric, to model the intensity of marked point patterns with real-valued marks in two spatial dimensions without making any distributional assumption. Both methods allow us to estimate the effect of the mark on the intensity of the process. Furthermore, we can estimate the density of the mark across the observed window. We show that by including mark information in the model, if the real-valued mark has an impact on the intensity, we can obtain equal or even better estimate of the intensity of the process with respect to unmarked models, while giving an additional information about the process. Both approaches are tested through simulations and applications on real data.

References

[1] Y. Ogata. *Statistical models for earthquake occurrences and residual analysis for point processes.* Journal of the American Statistical association. Taylor & Francis, 1988.

[2] M. Eckardt, M.Moradi. *Marked spatial point processes: current state and extensions to point processes on linear networks.* Journal of Agricultural, Biological and Environmental Statistics. Springer, 2024.

Minimax Estimation of the Structure Factor of a Point Process

G. Mastrilli[1 2]
[1] Univ Rennes, Ensai, CNRS, CREST-UMR 9194, F-35000 Rennes, France
[2] Inria, Paris, France
`gabriel.mastrilli@ensai.fr`

Spectral methods have recently gained attention in spatial statistics as an alternative to direct-space techniques for second-order inference of spatial point processes. Instead of analyzing second-order summary statistics like the pair-correlation function or the Ripley K-function, one can estimate the Bartlett's spectral measure, also known in physics as the structure factor S. Several estimators of S have been introduced [1, 2, 3]. But, without a reference minimax rate of convergence, it remains unclear whether better estimators could be constructed.
We address this gap by proving the following minimax result: for estimating a structure factor of Hölder regularity β, the worst-case error of any estimator is at least the expected number of observed points raised to the power $-\beta/(2\beta+d)$. We then construct multi-taper estimators that achieve this optimal rate and further show that they exhibit exponential concentration. However, these estimators are oracle in the sense that they depend on unknown parameters related to the smoothness of S. To obtain a practical estimator, we propose a cross-validation method based on random thinning and decorrelation between distinct frequencies, which performs well in practice.

References

[1] D. Hawat, G. Gautier, R. Bardenet, and R. Lachièze-Rey. On estimating the structure factor of a point process, with applications to hyperuniformity. *Statistics and Computing*, vol. 61, no. 33, 2023.

[2] T. A. Rajala, S. C. Olhede, J. P. Grainger, and D. J. Murrell. What is the Fourier transform of a spatial point process? *IEEE Transactions on Information Theory*, 2023.

[3] J. Yang and Y. Guan. Fourier analysis of spatial point processes. *arXiv preprint*, arXiv:2401.06403, 2024.

Estimating the hyperuniformity exponent of spatial point processes

G. Mastrilli[1,2], B. Błaszczyszyn[2], **F. Lavancier**[1]

[1] Univ Rennes, Ensai, CNRS, CREST, France
[2] Inria, ENS/PSL Univ, Paris, France
`frederic.lavancier@ensai.fr`

A point process Φ is hyperuniform if it exhibits a slower growth in the variance of the number of points at large scales compared to Poisson point processes. Initially conceptualized in statistical physics, hyperuniform systems have attracted significant interest due to their unique position between perfect crystals, liquids, and glasses. Hyperuniformity has gained attention in various applied contexts, offering insights into phenomena ranging from DNA and the immune system to active matter theory, urban systems, ices, rock dispersion on Mars, hydrodynamics, avian photoreceptors, and cosmology. Despite this, there is a lack of method to detect and quantify hyperuniformity. Under certain conditions, hyperuniformity is equivalent to the structure factor S of Φ satisfying $S(k) \sim |k|^\alpha$ as $|k| \to 0$, where the exponent $\alpha > 0$ characterises the degree of hyperuniformity. Estimating α faces the challenge of estimating the structure factor S for small frequencies. In this contribution, we construct a multi-scale, multi-taper estimator of α, based on just one realization of Φ. Our approach shortcuts the step of estimating $S(k)$ for small but non zero frequencies k. We analyse its (expanding window) asymptotic behavior, enabling the construction of an asymptotic confidence interval for α. The method is illustrated on simulated data and applied to the analysis of hyperuniformity in a real dataset of marine algae.

Topological data analysis for random sets

V. Gotovac Đogaš[1], M. Mandarić[1]

[1] Department of Mathematics, Faculty of Science, University of Split, Croatia
`vgotovac@pmfst.hr`

This work paves the way for a methodology for detecting outliers and testing the goodness-of-fit of random sets using topological data analysis. Our approach is based on building a filtration using the sublevel sets of the signed distance function and analyzing various summary functions derived from the persistence diagrams of the resulting persistent homology. Outlier detection is performed using functional depth measures [1] applied to these summary functions. To evaluate goodness-of-fit, we use global envelope tests [3] with summary statistics serving as test statistics. The methodology is supported by a simulation study based on random set germ-grain models and is also applied to real-world data from histological images of mastopathic and breast cancer tissue.

References

[1] I. Gijbels and S. Nagy. *On a general definition of depth for functional data*, Stat Sci **32**, 630—639, 2017.

[2] V. Gotovac Đogaš, M. Mandarić. *Topological data analysis for random sets and its application in detecting outliers and goodness of fit testing* , arXiv:2404.04702, 2024.

[3] M. Myllymäki, T. Mrkvička, R. Grabarnik, H. Seijo and U. Hahn. *Global envelope tests for spatial processes*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **9**, 381–404, 2016.

Optimizing Sample Size and Survey Design for Plant Disease Surveillance using Gibbs Marked Point Processes

F. d'Alayer[1], M. Cendoya[2], E. Gabriel[1], S. Soubeyrand[1]

[1] INRAE BioSP, Avignon
[2] ivia Valencià
`francois.dalayer@inrae.fr`

Accurately detecting plant diseases in a spatial context is crucial for timely intervention and effective management. Traditional sample size determination methods ensure a high probability of detecting a disease if it is present but typically assume spatial independence of disease occurrences. However, many plant pathogens exhibit spatial aggregation or repulsion due to environmental factors or host distribution.
In this study, we extend the classical approach to sample size calculation [1] by incorporating spatial dependencies modeled through Gibbs marked point processes. This framework allows us to explicitly account for interactions between infected and non-infected locations, leading to a more realistic representation of disease spread. We introduce a novel criterion of sample efficiency to assess the robustness of our method across various spatial patterns. Our results demonstrate that, even under aggregation or repulsion, the adapted sample size calculations maintain a high probability of detection while optimizing the allocation of sampling efforts.
Beyond sample size determination, we leverage the Gibbs model to develop a survey design strategy aimed at minimizing the expected time to first detection.
Our findings provide a methodological advancement for plant disease monitoring, offering insights and practical guidelines for designing efficient sampling schemes under spatially structured uncertainty. These results are particularly relevant for policymakers and epidemiologists seeking to improve early warning systems and optimize resource allocation in plant disease surveillance.

References

[1] European Food Safety Authority (EFSA). *A framework to substantiate absence of disease: the risk based estimate of system sensitivity tool (RiBESS) using data collated according to the EFSA Standard Sample Description - An example on Echinococcus multilocularis.* EFSA Supporting Publications.

Approximate likelihood estimation for Gibbs point processes

M. Rost[1], O. Cronie[1]

[1] Department of Mathematical Sciences, Chalmers University of Technology & Uinversity of Gothenburg, Sweden
`mathisr@chalmers.se`

When fitting a model to data, one would ideally like to use maximum likelihood estimation, due to its nice statistical properties. Unfortunately, the likelihood function of a general Gibbs point process is typically not tractable, due to the associated normalizing constant. This has led to the development of a range of alternative methods, such as Takacs-Fiksel estimation (including its special case pseudolikelihood estimation) and Point Process Learning.

Leveraging recent probabilistic results for Gibbs processes, in this talk we present an approach to perform approximate likelihood estimation for Gibbs processes. Specifically, we show that the likelihood function can be expressed completely in terms of the Papangelou conditional intensity, which is typically known and tractable. This new likelihood representation involves an infinite series expansion, and we discuss different ways of approximating it, and thereby the likelihood function. We further discuss how this plays out in certain models and compare it to the state-of-the-art.

Conformal novelty detection for replicate point patterns with FDR or FWER control

Christophe A. N. Biscio[1], Adrien Mazoyer[2], **Martin V. Vejling**[1]

[1] Department of Mathematical Sciences, Aalborg University
[2] Institut de Mathématiques de Toulouse, Université de Toulouse
`mvv@math.aau.dk`

Monte Carlo tests are widely used for computing valid p-values without requiring known distributions of test statistics. When performing multiple Monte Carlo tests, it is essential to maintain control of the type I error, but it is not trivial how to provide false discovery rate guarantees while maintaining the power. In this talk, we highlight that multiple Monte Carlo testing is an instance of conformal novelty detection [1, 2], thereby paving the way to constructing efficient testing procedures with false discovery rate guarantees. The performance is investigated in the context of global envelope tests [3] for replicate point pattern data through a simulation study and an application to a sweat gland data set. The preprint is available on arXiv [4].

References

[1] Stephen Bates and Emmanuel Candès and Lihua Lei and Yaniv Romano and Matteo Sesia. *Testing for outliers with conformal p-values*, volume 51. The Annals of Statistics, 2023.

[2] Ariane Marandon and Lihua Lei and David Mary and Etienne Roquain. *Adaptive novelty detection with false discovery rate guarantee*, volume 52. The Annals of Statistics, 2024.

[3] Mari Myllymäki and Tomáš Mrkvička and Pavel Grabarnik and Henri Seijo and Ute Hahn. *Global envelope tests for spatial processes.* Journal of the Royal Statistical Society, Series B, 2017.

[4] Christophe A. N. Biscio and Adrien Mazoyer and Martin V. Vejling. *Conformal novelty detection for replicate point patterns with FDR or FWER control.* arXiv, 2025.

Nonparametric isotropy test for spatial point processes using random rotations

C. Fend[1], C. Redenbach[1]

[1] Department of Mathematics, RPTU University Kaiserslautern-Landau, Germany
`chiara.fend@rptu.de`

In spatial statistics, point processes are often assumed to be isotropic meaning that their distribution is invariant under rotations around the origin. Most statistical tests for the null hypothesis of isotropy found in the literature are based either on asymptotics or on Monte Carlo simulation of a parametric null model. However, fitting a parametric null model is challenging in case of anisotropy and asymptotic distributions are only available to a limited extent.

In [1], we present a novel nonparametric and computationally cheap test for the hypothesis of isotropy for stationary point processes. Our test is based on resampling the Fry points which are the pairwise difference vectors of the observed point pattern. The resampling itself is achieved by using random rotation techniques. We investigate empirical levels and powers of the proposed test in a simulation study for a diverse set of point process models. Anisotropic point process models are hereby obtained either via the geometric anisotropy mechanism or through oriented clusters where the orientation distribution is given by a von-Mises-Fisher distribution. Finally, a real data set of amacrine cells is tested for isotropy.

References

[1] C. Fend and C. Redenbach. *Nonparametric isotropy test for spatial point processes using random rotations.* Spatial Statistics (64), 2024.

Functional summary statistics and testing for independence in multi-type point processes on the surface of 3D shapes

S. Ward[1], **E. A. K. Cohen**[1], N. M. Adams[1]

[1] Department of Mathematics, Imperial College London, London, United Kingdom
`e.cohen@imperial.ac.uk`

Recent advances in 3D biological imaging technologies now allow spatial point patterns of protein molecules to be observed directly on a cell's outer membrane, where the underlying geometry is complex and must be respected for principled statistical inference. In prior work, we developed functional summary statistics for point patterns on the surfaces of 3D convex shapes [1], leading to new insights into E. coli outer membrane assembly [3]. In this work, we extend these methods to the multi-type setting. We begin by developing functional summary statistics for multi-type homogeneous and inhomogeneous point processes on the sphere, and then generalise them to convex 3D surfaces, assuming a known bijective mapping from the shape to the sphere. To support inference in the inhomogeneous case, we employ a plug-in estimator for the intensity function of a spatial point process on a manifold [2]. We demonstrate how these statistics can be used to test for independence between the component processes, with particular focus on methods for generating samples from the null distribution. We conclude with a discussion on extending the framework to a class of non-convex shapes. This talk is based on work presented in [4].

References

[1] S. Ward, E. A. K. Cohen, and N. Adams, "Testing for complete spatial randomness on 3-dimensional bounded convex shapes," *Spatial Statistics*, vol. 41, 2021.

[2] S. Ward, H. S. Battey, and E. A. K. Cohen, "Estimation of the intensity function of a spatial point process on a Riemannian manifold," *Biometrika*, vol. 110, 2023.

[3] S. Kumar, P. Inns, S. Ward, V. Lagage, J. Wang, R. Kaminska, S. Uphoff, E. A. K. Cohen, G. Mamou, and C. Kleanthous, "Immobile lipopolysaccharides and outer membrane proteins differentially segregate in growing Escherichia Coli," *Proceedings of the National Academy of Sciences*, vol. 122, no. 10, 2025.

[4] S. Ward, E. A. K. Cohen, and N. M. Adams, "Functional summary statistics and testing for independence in multi-type point patterns on the surface of three dimensional convex shapes," *Spatial Statistics*, vol. 67, 2025.

Semi-parametric Markov model for Multi-type Point Processes with Applications to Industry Agglomeration

I. T. Jensen[1], J. F. Coeurjolly[2], R. Waagepetersen[1]

[1] Department of Mathematical Sciences, Aalborg University, Denmark
[2] Laboratory Jean Kuntzmann, Department DATA, Université Grenoble Alpes, France
`itj@math.aau.dk`

We propose a semi-parametric Markov model for multi-type point processes. The model includes a parametric part, which describes the effect of spatial covariates and interactions within and between point types, and a non-parametric part, which accounts for unobserved variables that affect all types equally. The specification of the model allows for repulsive, attractive, and Poisson processes. Similar efforts in the past has either lacked the semi-parametric aspect [1] or required attraction between points [2, 3]. Parameters are estimated with a composite likelihood setup, which produces a consistent and asymptotically Gaussian estimator. The covariance matrix of the maximum composite likelihood is estimated thus allowing inference on the parameters. We apply the method to Dutch shop data, and are able to identify interactions between shop types while accounting for spatial covariates, which has not been possible with the approaches previously employed in the economics literature.

References

[1] I. Ba and J.F. Coeurjolly *Inference for low- and high-dimensional inhomogeneous Gibbs point processes* Scandinavian Journal of Statistics, 2022, 50:3, 993–1021, DOI: 10.1111/sjos.12616

[2] K. Hessellund and G. xu and Y. Guan and R. Waagepetersen *Second-order semi-parametric inference for multivariate log Gaussian Cox processes* Journal of the Royal Statistical Society Series C, 2022, 71, 244–268, DOI: 10.1111/rssc.12530

[3] K. Hessellund and G. xu and Y. Guan and R. Waagepetersen *Semiparametric Multinomial Logistic Regression for Multivariate Point Pattern Data* Journal of the American Statistical Association, 2022, 117:539, 1500–1515, DOI: 10.1080/01621459.2020.1863812

Composite likelihood inference for space-time point processes

A. Jalilian[1], **F. Cuevas-Pacheco**[2], G. Xu[1], R. Waagepetersen[4]

[1] Razi University, Iran & Lancaster Ecology and Epidemiology Group, Lancaster University, United Kingdom
[2] Departamento de Matemáticas, Universidad Técnica Federico Santa María, Chile
[3] University of Miami, United States of America
[4] Department of Mathematical Sciences, Aalborg University, Denmark
`francisco.cuevas@usm.cl`

The dynamics of a rain forest is extremely complex involving births, deaths and growth of trees with complex interactions between trees, animals, climate, and environment. We consider the patterns of recruits (new trees) and dead trees between rain forest censuses. For a current census we specify regression models for the conditional intensity of recruits and the conditional probabilities of death given the current trees and spatial covariates. We estimate regression parameters using conditional composite likelihood functions that only involve the conditional first order properties of the data. When constructing assumption lean estimators of covariance matrices of parameter estimates we only need mild assumptions of decaying conditional correlations in space while assumptions regarding correlations over time are avoided by exploiting conditional centering of composite likelihood score functions. Time series of point patterns from rain forest censuses are quite short while each point pattern covers a fairly big spatial region. To obtain asymptotic results we therefore use a central limit theorem for the fixed timespan - increasing spatial domain asymptotic setting. This also allows us to handle the challenge of using stochastic covariates constructed from past point patterns. Conveniently, it suffices to impose weak dependence assumptions on the innovations of the space-time process. We investigate the proposed methodology by simulation studies and an application to rain forest data. [1].

References

[1] J. Abdollah, F. Cuevas-Pacheco, Xu. Ganggang, and R. Waagepetersen *Composite likelihood inference for space-time point processes (2025)*, volume 81. Biometrics.

Marked Spatio-temporal Models for Intensity Estimation on Linear Networks: the case of Ambulance Interventions in Milan

E. Delvino[1], **A. Gilardi**[1], R. Borgoni[1]

[1] Department of Economics, Management and Statistics, University of Milano - Bicocca, Italy

`andrea.gilardi@unimib.it`

Ambulance interventions represent critical events that demand immediate medical assistance and prompt dispatch of emergency response vehicles. Analysing their spatial and temporal distribution is essential for optimizing resource allocation and identify high-risk areas. Furthemore, ambulance interventions also represent a classical example of spatial point patterns occurring on linear networks, an increasingly popular type of processes which are constrained to lie on a restricted spatial domain which, in this case, corresponds to a street network [1, 3].

This work proposes a semi-parametric spatio-temporal marked Poisson point process model to study the intensity function of all emergency calls recorded in the street network of Milan from 2015 to 2017, dividing the events according to the intervention time [2]. More precisely, we assume that the ambulance dispatches recorded over the network with mark m during a one-hour interval t represent a realisation of a inhomgeneous marked Poisson process with intensity function $\lambda_m(t, s)$. The key for model development is assuming that the intensity function can be decomposed into the product of two components, say

$$\lambda_m(t, s) = \mu_m(t) g_m(s),$$

representing the temporal ($\mu_m(t)$) and spatial ($g_m(s)$) dimensions of the process, respectively. As already mentioned, the marks denote the ambulance access times, distinguishing between responses below and above five minutes. The temporal component was modelled via a semi-parametric Poisson regression with deterministic temporal covariates. More precisely, we included a linear term to capture the year-to-year variations, the hourly and weekly effects were smoothed using cyclic cubic regression splines, and the daily trends were included using dummy variables and interactions. The spatial model has been estimate using a parametric approach that incorporates covariates related to the road type, the traffic flow, and census data related to buildings and population. Both components are estimated separately for each mark m and they are combined in a multiplicative form to obtain a temporally-varying estimate of the intensity function. Our preliminary results highlight that central areas tend to have rapid emergency responses, whereas peripheral areas show longer waiting times, indicating spatial disparities in emergency service efficiency. These effects are not constant in time and are less evident during the night and the early morning hours.

Acknowledgments

This study was carried out within the MOST—Sustainable Mobility National Research Center and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1033 17/06/2022,CN00000023).

References

[1] Baddeley, A., Nair, G., Rakshit, S., McSwiggan, G. and Davies, T.M., 2021. Analysing point patterns on networks - A review. *Spatial Statistics*, 42, p.100435.

[2] Eckardt, M., Moradi, M. Marked Spatial Point Processes: Current State and Extensions to Point Processes on Linear Networks. *JABES* 29, 346–378 (2024).

[3] Gilardi, A., Borgoni, R. and Mateu, J., 2024. A nonseparable first-order spatiotemporal intensity for events on linear networks: An application to ambulance interventions. *The Annals of Applied Statistics*, 18(1), pp. 529-554.

A kernel-based test for the first-order separability of spatial-temporal point processes

M. Ghorbani[1], N. Vafaei[1], M. Myllymäki[2]

[1] Department of Engineering Sciences and Mathematics, Luleå University of Technology, Sweden

[2] Natural Resources Institute Finland (Luke), Helsinki, Finland.

`mohammad.ghorbani@ltu.se`

We present an innovative statistical test designed to assess the first-order separability of a spatio-temporal point process. Our proposed test employs block permutations and a novel test statistic that incorporates a machine learning technique known as the Hilbert-Schmidt independence criterion. To enhance the practicality of the criterion, we apply the kernel trick. The block permutations are designed to maintain the second-order structure of the point pattern, disrupting it only at the block borders. This design enables the application of our test to a general spatio-temporal point process, which may exhibit small-scale clustering or regularity. We investigated the empirical level of the block permutation based tests with the new and two previously proposed test statistics for clustered and regular point processes, represented in our study by log Gaussian Cox processes and determinantal point processes. By comparing our results with those obtained from a previously proposed permutation-based test, we confirmed the effectiveness of our method in terms of significance level, power, and notably computational cost. We applied the test to real-world datasets, namely the UK's 2001 foot and mouth disease epidemic and varicella data from Valencia.

Keywords: Block permutation, feature space, Kernel mean embedding, Machine learning, Reproducing kernel Hilbert space, Separability of intensity function, Spatio-temporal point processes.

Computing forest structural diversity maps using point process statistics

M. Myllymäki[1], T. Rajala[1], J. Kostensalo[2], A. Kangas[2], P. Packalen[1], L. Mehtätalo[2], S. Tuominen[1], M. Kuronen[1]

[1] Natural Resources Institute Finland (Luke), Helsinki, Finland
[2] Natural Resources Institute Finland (Luke), Joensuu, Finland
`mari.myllymaki@luke.fi`

Measuring large patterns of trees with location and size information is time consuming and costly. Thus, traditionally mapped tree data from forests were available only from rather small areas. These days remote sensing techniques, particularly airborne laser scanning (ALS), allow detection of dominant trees and, thus, mapping them at large areas. Using ALS, optical imagery, and field data, it is also possible to predict undetected trees and tree attributes. The outputs are large scale tree maps, i.e., inhomogeneous point patterns of trees with diameter, height, volume and species marks [1, 2].

Local forest biodiversity hotspots are small areas within a landscape or a single stand. Such hotspots are characterized by a high variability in, e.g., tree species composition, size distribution or spatial pattern, in comparison to the surrounding areas. Traditionally remote sensing has been used predict different indices quantifying forest structure using a fixed spatial resolution only. When data are instead available at a single-tree level, more enhanced analysis of forest structure is possible using point process statistics. We give examples of species, size and structural diversity maps [3].

References

[1] J. Kostensalo, L. Mehtätalo, S. Tuominen, P. Packalen, and M. Myllymäki (2023). Recreating structurally realistic tree maps with airborne laser scanning and ground measurements. *Remote Sensing of Environment* **298**(1), 113782. DOI: 10.1016/j.rse.2023.113782

[2] J. Kostensalo, P. Packalen, M. Kuronen, L. Mehtätalo, S. Tuominen, and M. Myllymäki (2025) Large-Scale Tree-Level Mapping of Forest Structure Including Species Type with Remote Sensing Data and Ground Measurements. http://dx.doi.org/10.2139/ssrn.5124085

[3] T. Rajala, A. Kangas, and M. Myllymäki (2025). Computing Maps of Forest Structural Diversity: Aggregate Late. http://dx.doi.org/10.2139/ssrn.5119277.

Spatial summaries of nuclei distribution in stained tissue images and their prognostic strength in survival analysis on patients with melanoma

H. Häbel[1], H. Eriksson[2]

[1] Department of Learning, Informatics, Management and Ehtics, Karolinska Institutet, Sweden

[2] Hanna Eriksson, Department of Oncology-Pathology, Karolinska Institutet, Sweden

`henrike.habel@ki.se`

Analysis of stained tissue images is commonly performed in cancer diagnosis and prognosis, where immunohistochemistry (IHC) is a valuable tool for characterizing cells or biomarkers of interest within the tissue. The extracted characteristics may include summaries of nuclei intensity, vessel diameter, and stained area fraction. Usually, there are several IHC markers available, and the question arises which of them may be associated with the outcome (for example overall survival or cancer-specific survival) [1]. For the investigation of their prognostic strength, image-assisted survival analysis can be conducted, where IHC marker specific characteristics are included in a regression model along with other prognostic factors such as age, sex, and tumor specific characteristics. However, such IHC characteristics usually do not describe the spatial distribution of the point patterns of nuclei. In my talk, I will discuss challenges and benefits of including summaries of the spatial distribution of nuclei as prognostic factors. In particular, I will propose how spatial indices can be derived from the so-called L-function [2] on the example of patients with stage III melanoma.

References

[1] S. Corvigno, A. Mezheyeuski, L. Martin De La Fuente, S. Westbom-Fremer, J.W. Carlson, J. Fernebro, E. Åvall-Lundqvist, P. Kannisto, I. Hedenfalk, S. Malander, C. Rolny, H. Dahlstrand, A. Östman. *High density of stroma-localized CD11c-positive macrophages is associated with longer overall survival in high-grade serous ovarian cancer.* Gynecologic Oncology, 159 (3): 860-868, 2020.

[2] H. Häbel, A. Balazs, M. Myllymälki. *Spatial analysis of airborne laser scanning point clouds for predicting forest structure.* Mathematical and Computational Forestry & Natural-Resource Sciences (MCFNS), [S.l.], 13 (1): 15-28. 2021.

Random fractal models for analyzing second-order properties of stationarity point processes on linear networks

F. J. Rodríguez-Cotés[1], J. F. Díaz-Sepúlveda[1], R. Giraldo[2] and J. Mateu[3]

[1] Departamento de Estadística, Universidad Nacional de Colombia, Medellín, Colombia
[2] Departamento de Estadística, Universidad Nacional de Colombia, Bogotá, Colombia
[3] Department of Mathematics, Universitat Jaume I, Castellón, Spain
`frrodriguezc@unal.edu.co`

The statistical modeling of second-order characteristics in point processes often commences with testing the hypothesis of spatial randomness. We address the challenge of assessing complete randomness within the geometric framework of point processes on linear networks, where the conventional properties of a point process undergo alterations and data visualization becomes less intuitive. In the case of planar scenarios, traditional goodness-of-fit tests rely on quadrat counts and distance-based methods. As an alternative approach, we propose a novel statistical test of randomness based on the fractal dimension, which is calculated using the box-counting method. This provides a robust inferential perspective, offering a departure from the more commonly descriptive application of this method. Additionally, it enables the discrimination between clustered and inhibitory behaviors of point patterns. We assess the performance of our methodology through a simulation study and the analysis of a real dataset. The results bolster the efficacy of our approach, presenting it as a viable alternative to the computationally more demanding classical distance-based strategies.

References

[1] A. Baddeley, A. Nair, S. Rakshit, G. McSwiggan and T.M. Davies. *Analysing point patterns on networks—A review*, Spatial Statistics, **42**, 100435, 2021.

[2] K. Falconer. *Fractal geometry: mathematical foundations and applications*, John Wiley & Sons Chichester, 2013.

[3] Y. Caballero, R. Giraldo and J. Mateu. *A spatial randomness test based on the box-counting dimension*, AStA Advances in Statistical Analysis, **106(3)**, 499-524.

Pseudo-spectra of multivariate inhomogeneous spatial point processes

J. Yang[1], Q.-W. Ding[1], J. Shin[2]

[1] Institute of Statistical Science, Academia Sinica, Taiwan
`junhoyang@stat.sinica.edu.tw`
[2] School of Mathematics, Statistics and Data Science, Sungshin Women's University, South Korea

In this presentation, we propose a spectral method for multivariate inhomogeneous spatial point processes. A key ingredient is utilizing the asymptotic behavior of the periodogram. The periodogram is an asymptotically unbiased estimator of the spectrum of a second-order stationary point process. By extending this property, we show that under inhomogeneity, the expectation of the periodogram also converges to a matrix-valued function, which we refer to as the pseudo-spectrum. The pseudo-spectrum shares similar properties with the spectrum of stationary processes and can be interpreted using local parameters. We derive a consistent estimator of the pseudo-spectrum through kernel smoothing and propose two bandwidth selection methods. The performance and utility of our frequency domain methods are illustrated through simulation studies and a real data analysis of rainforest data.

Attractive coupling of determinantal point processes using non symmetric kernels

A. Poinas[1]

[1] Laboratoire de Mathématiques et Applications, University of Poitiers, France
`arnaud.poinas@univ-poitiers.fr`

Determinantal point processes (or DPP for short) are commonly used to model datasets of repulsive point patterns. DPPs are defined through a symmetric function, called their kernel, from which we can derive their main statistical properties such as their likelihood or their k-th order intensity function. The symmetry of the kernel of DPPs is an important property. It is used to establish their theoretical results and it is also the cause of their repulsive nature. Now, there is a natural way of constructing a coupling of determinantal point processes as a DPP itself in order to model marked point processes with repulsion between points of the same mark and between points of different marks.

While most works on DPPs only consider symmetric kernels, it is possible to define these processes using a non symmetric kernel, allowing for some positive dependency. Although, not much is known about how this change affects the usual properties of DPPs.

In this talk, we study these DPPs with generic kernels and we determine how the common results on DPPs with symmetric kernels generalize in this case. We also show how to use these point processes to construct attractive couplings of DPPs in order to model marked point processes with repulsion between points of the same mark but attraction between points of different marks.

Statistical properties for critical point processes

J. Chevallier[1], **J-F. Coeurjolly**[1], R. Waagepetersen[2]

[1] Laboratoire Jean Kuntzmann, Université Grenoble Alpes, France
[2] Aalborg University, Denmark
`jean-francois.coeurjolly@univ-grenoble-alpes.fr`

Spatial point patterns emerge in many applications: in forestry (e.g. locations of plants/trees), environmental sciences (e.g. lightning strike impacts in France), economis (locations of banks), astrophysics (e.g. sunspots),... Even in the stationary case (which is th context of the present study), many models exist in the literature to model clustering effects or repulsion between points/objects, such as Cox, Gibbs or determinantal point processes to cite a few. I will present a "new" class of stochastic models, simply obtained as critical points obtained from a Gaussian random fields. The topic of zeroes or critical points of (Gaussian) random fields has given rise to a huge literature, ... and I would say essentially, in the probabilistic community. This is why by "new" I mean that I'll be focusing on properties and problems useful for statistical applications. These points deal with understanding precisely first and second moments (and more generally higher-order intensity functions), problems related to the simulation of such models and asymptotic results for functionals of these point processes, which include the natural intensity estimate and a slightly modified version of the Ripley's K function.