An introduction to geostatistical analysis of spatio-temporal data with R



Workshop 1

An introduction to geostatistical analysis of spatio-temporal data with R

Organizers: Denis Allard, BioSP, INRA; Liliane Bel, AgroParisTech; Edith Gabriel, Université d'Avignon; Thomas Opitz, BioSP, INRA; Eric Parent, AgroParisTech

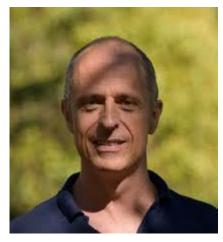
Where: TBA, Lancaster University

When: July 4th, 8:30 am - 17:30 pm

The organizers

Eric Parent, AgroParisTech, Paris







Liliane Bel, AgroParisTech, Paris

Denis Allard, BioSP, INRA (Avignon)



Edith Gabriel, University of Avignon



Thomas Opitz, BioSP, INRA (Avignon)

04/07/2017

SPAT 2017 Workshop 1

Why this workshop ?

Space-time models and methods are essential in modern statistics with applications in:

environment, ecology, forestry, climate, health, air pollution, hydrogeology, crime data analysis, road crashes, etc. [all from SPAT 2017 programme]

But, space-time statistics is a bit intimidating at first



Some particular points about Space-Time statistics

- Involves (very) large datasets: difficult to store, to read, to manipulate, to analyse
- ✓ We don't know how to explore the data for choosing a model/method
- ✓ Spatio-temporal or tempo-spatial?
- Likelihood (almost) impossible to compute
- ✓ Are there R packages? Which one should we use? Are there important features missing?

- We identified the need to illustrate the geostatiscal analysis of spatiotemporal data
- On a realistic dataset the whole workflow for
- From reading the data to do predictions and validate the model

✓ Ftc...

Working sessions and production of a review paper

RESSTE Réseau MIA RESSTE		RESeau Statistiques pour données Spatio-TEmporelles			
Bienvenue	Lancement: Oct 2014	Estimation: Mai 2015	Estimation & Visualisation: Nov 2015	Atelier d'Avignon, Avril 2016	
Modèles hiérarchiques - Mai 2017					
User login		Bienvenue			

Username Password Create new account Request new password

Log in

Les progrès de l'instrumentation, des systèmes électroniques embarqués et de l'imag génèrent de très grandes quantités de données, géolocalisées et répétées dans le tem nombre de variables, parfois assez fortement dépendantes, sont disponibles. Les déf données spatio-temporelles sont multiples :ils concernent à la fois la visualisation de définition de modèles pertinents, les méthodes d'estimation pour ceux-ci et la mise e méthodes pour des ensembles de données de (très) grande taille.

Le réseau RESSTE (RESeau Statistiques pour données Spatio-TEmporelles) proposcientifique autour des modèles, méthodes et algorithmes pour les données spatio-te

fédérant des statisticiens de différents horizons et des modélisateurs des autres disci traiter de ce type de données, en confrontant les approches et points de vue des diffé statistiques. Il a pour ambition de contribuer à une plus large diffusion des méthode

spatio-temporelles, auprès des statisticiens et des modélisateurs. A moven terme, il s nouveaux fronts de recherche en statistiques spatio-temporelles répondant aux enje

dans 19 équipes (voir la liste complète ici) et est animé par un bureau de six personn

les scientifiques, et de favoriser les collaborations entre équipes de recherche. RESSTE est financé depuis 2014 par le département MIA de l'INRA. Il compte à ce j Soumission

Journal de la Société Française de Statistique

Analyzing spatio-temporal data with R: Everything you always wanted to know - but were afraid to ask

Titre: Donnees spatio-temporelles avec R: tout ce que vous avez toujours voulu savoir sans jamais avoir osé le demander

RESSTE Network¹

Abstract: We present an overview of (geo-)statistical models, methods and techniques for the analysis and prediction of continuous spatio-temporal processes residing in continuous space. Various approaches exist for building statistical models for such processes, estimating their parameters and performing predictions. We cover the Gaussian process approach, very common in spatial statistics and geostatistics, and we focus on R-based implementations of numerical procedures. To illustrate and compare the use of some of the most relevant packages, we treat a real-world application with high-dimensional data. The target variable is the daily mean PM₁₀ concentration predicted thanks to a chemistrytransport model and observation series collected at monitoring stations across France in 2014. We give R code covering http://journal-sfds.fr Online soon at

http://informatique-mia.inra.fr/resste

Denis Allard, BioSP, MIA

Organization of the workshop

- Session 1: Handling and importing large spatio-temporal data using structured objects; projection coordinate systems for geolocated data. (Eric)
- Session 2: Visualizing data according to their temporal, spatial or spatiotemporal structures.
 (Edith)

Lunch

- Session 3: Statistical inference for spatio-temporal models: method of moments; maximum likelihood, pairwise composite likelihoods. (Liliane)
- Session 4: Prediction and validation. (Thomas)

One of the outputs: review of R packages

Cov models Estimation Packages Data. Kriging Big N STDF, STSDF 4 classes: Separable, fit.StVariogram krigeST Moving gstat or STIDF data Product-Sum, Metric, Weighted Least Sq. Linear model Neigh Sum Metric. structure from only. for the trend. only. Geom. anisotropy No NAS. spacetime. Weighted Least Sq. Pairwise CompRandFld Space \times Time Many classes: Kri data design. Separable, Comp. Lik. No trend model. Comp. Lik., No NAs at all. Product-Sum, Porcu, Hypothesis Tapering. Tapering., Gneiting, ... Chordal and Use of spam testing. Geodesic dist. RFsp: extension Comprehensive: RFinterpolate Use of spam RandomFields RFfit Tapering from sp Product-Sum, Weighted Least Sq., incl. trend package Gneiting, Porcu, Max. Lik. modeling mixtures, ...

TABLE 5. Most important R packages for spatio-temporal analysis, along with their main features and limitations.

And now ?

We will make all material avalailable

- Access to the full dataset
- Workshop presentations
- Link to extended material at http://informatiquemia.inra.fr/resste
- Mail list: sending updated Rmd, link to the paper
- Send email to <u>denis.allard@inra.fr</u> to be added to the list