Workshop on Stochastic Weather Generators

Avignon, 17-19 September, 2014

Scientific and Organizing Committee :
Pierre Ailliot, Denis Allard, Edith Gabriel, Valérie Monbet, Philippe Naveau, Peter Thomson

Foreword

This workshop on Stochastic Weather Generators is organized in Avignon by the Biostatistics and Spatial Processes Unit (INRA) and statistics group from the Department of Mathematic, University of Avignon. It is the follow-up event of the workshop on stochastic weather generators organized in Roscoff, June 2012.

This workshop is twofold. First, it is intended to provide an overview of state-of-the-art research in the statistical modelling of climate variables with a simulation point of view. Second, it intends to promote scientific discussions to foster new ideas and create collaborative links among attendees.

We thank all contributors for their financial support (INRA, MIA, INRIA, EDF and ECCOREV). We thank the University of Avignon for its material support.

We wish to thank Sylvie Jouslin for her highly valuable contribution to the preparation of this workshop.

The organization committee

Participants

Pierre Ailliot
Mathematics, Université de Brest
pierre.ailliot@univ-brest.fr

Nawal Akrour
LATMOS, IPSL
Nawal.Akrour@latmos.ipsl.fr

Denis Allard
BioSP, INRA
allard@avignon.inra.fr

Anastassia Baxevani
Mathematics and Statistics, University of Cyprus
baxevani@ucy.ac.cy

Aurélien Bechler
UMR 518 MIA INRA/AgroParisTech
aurelien.bechler@agroparistech.fr

Patrick Bertuzzi
AgroClim, INRA
pbertuzzi@avignon.inra.fr

Julie Bessac
Mathématiques, IRMAR-Université Rennes 1
julie.bessac@univ-rennes1.fr

Delphine Blanke
Laboratoire de Mathématiques, Université d'Avignon
delphine.blanke@univ-avignon.fr

Florent Bonneu
Laboratoire de Mathématiques, Université d'Avignon
florent.bonneu@univ-avignon.fr

Marc Bourotte
BioSP, INRA
bourotte@paca.inra.fr

Trevor Carey-Smith
Climate Variability, NIWA, New Zealand
trevor.carey-smith@niwa.co.nz

Julie Carreau
HydroSciences, IRD
julie.carreau@univ-montp2.fr

Romain Chailan
Géosciences Montpellier, Université Montpellier 2
romain.chailan@gmail.com

Richard Chandler
Statistical Science, UCL
r.chandler@ucl.ac.uk

Sheng Chen
HHLY, IRSTEA, France
sheng.chen@irstea.fr

Jie Chen
Department of Construction Engineering, École de Technologie Supérieure, Québec
jie.chen@etsmtl.ca

Didier Dacunha Castelle
Mathématiques, UPS
didier.dacunha-castelle@math.u-psud.fr

Pascal Dkengne Sielenou
ETNA, IRSTEA
pascal.sielenou@irstea.fr

Nicolas Eckert
ETNA, IRSTEA
nicolas.eckert@irstea.fr

Edith Gabriel
Laboratoire de Mathématiques, Université d'Avignon
edith.gabriel@univ-avignon.fr

Marc Genton
Statistics, King Abdullah University of Science and Technology
marc.genton@kaust.edu.sa

Robin Girard
Department of Energy and Process, MinesParisTech
robin.girard@mines-paristech.fr

Yanira Guanche
IFREMER
yanira.guanche.garcia@ifremer.fr

William Kleiber
Applied Mathematics, University of Colorado at Boulder
william.kleiber@colorado.edu

Etienne Klein
BioSP, INRA
etienne.klein@avignon.inra.fr

Komlan Kpogo-Nuwoklo
CSN, IFREMER
kkpogonu@ifremer.fr

Christian Lantuéjoul
Geosciences, MinesParisTech
christian.lantuejoul@mines-paristech.fr

Marie Launay
AgroClim, INRA
mlaunay@avignon.inra.fr

Etienne Leblois
UR Hydrology-Hydraulique, Irstea
etienne.leblois@irstea.fr

Valérie Monbet
IRMAR, Université de Rennes 1
valerie.monbet@nuiv-rennes1.fr

Alexandre Mornet
ISFA / Allianz
alexandre.mornet@hotmail.fr

Philippe Naveau
LSCE, CNRS
`naveau@lsce.ipsl.fr`

Marco Oesting
UMR 518 MIA INRA/AgroParisTech
`marco.oesting@agroparistech.fr`

Thomas Opitz
BioSP, INRA
`opitz.thom@gmail.com`

Fabio Oriani
Chyn, Université de Neuchatel
`fabio.oriani@unine.ch`

Mehdi Othmani
Strategy-SEER-Prospective & Innovation, GDF SUEZ
`mehdi.othmani@gdfsuez.com`

Sylvie Parey
MFEE, EDF/R&D
`sylvie.parey@edf.fr`

Sylvain Pincebourde
Institut de Recherche sur la Biologie de l'Insecte, CNRS
`sylvain.pincebourde@univ-tours.fr`

Marc Prevosto
IFREMER
`Marc.Prevosto@ifremer.fr`

Damien Raynaud
LTHE – HMCI, CNRS
`damien.raynaud@ujf-grenoble.fr`

Valentin Resseguier
Fluminance / LOS, INRIA / IFREMER
`valentin.resseguier@inria.fr`

Pierre Ribereau
ICJ, Université de Lyon 1
`ribereau.pierre@gmail.com`

Igor Rychlik
Mathematical Sciences, Chalmers University of Technology
`igor.rychlik@chalmers.se`

Stephan Sain
IMAGe, NCAR
`ssain@ucar.edu`

Roman Schefzik
Institute for Applied Mathematics, Heidelberg University
`r.schefzik@uni-heidelberg.de`

Ying Sun
CEMSE Division, King Abdullah University of Science and Technology
`sunwardstat@gmail.com`

Peter Thomson
New Zealand, Statistics Research Associates
`peter@statsresearch.co.nz`

Glwadys Toutlemonde
I3M, Université Montpellier 2
`gwladys.toulemonde@univ-montp2.fr`

Pradeebane Vaittinada Ayar
LSCE, CNRS
`pvait@lsce.ipsl.fr`

Mathieu Vrac
LSCE, IPSL
`Mathieu.vrac@lsce.ipsl.fr`

Hans Wackernagel
Equipe de Géostatistique, MinesParisTech
`hans.wackernagel@mines-paristech.fr`

Pascal Yiou
LSCE, CNRS
`pascal.yiou@lsce.ipsl.fr`

Tuesday 16th September

19:30 *Welcome drink “La Cave des Pas Sages”* 41, rue des Teinturiers.

All talks will take place in Room 2E07. Access is in the historical building: take the large stairs up to 2nd floor. Then straight ahead in Eastern corridor.

Luggage room available nearby, in room 2E12.

Wednesday 17th September

8:20 – 8:50 Registration
8:50 – 9:00 Welcome address

9:00 – 10:30 **Simulating Precipitation**
Moderator: P. Naveau

- **E. Leblois** : Space-time simulation of intermittent rainfall with prescribed advection field: adaptation of the turning band method
- **N. Akrour** : Generation of Rainfall time series from micro to large scale
- **A. Baxevani** : Spatio-temporal precipitation generator with a censored latent Gaussian field

10:30 – 11:00 *Coffee break*

11:00 – 12:30 **Simulating wind conditions**
Moderator: P. Naveau

- **J. Bessac** : Markov-Switching AutoRegressive models for Cartesian components of wind fields in the North-East Atlantic
- **I. Rychlik** : Variability of wind speed encountered by a vessel
- **M. Oesting :** Conditional Modeling of Extreme Wind Gusts by Bivariate Brown-Resnick Processes

---

12:30 – 14:15 *Lunch at University restaurant*

14:15 – 15:45 **SWGs from a user prospective**
Moderator: D. Allard

- **S. Pincebourde** : The role of microclimates in climate change responses: ecologists need climatic data with high resolution
- **S. Parey** : A stochastic weather generator for temperature, examples of use and future developments
- **M. Othmani :** TBA

15:45 – 16:15 *Coffee break*

16:15 – 17:15 **Simulating Precipitation**
Moderator: D. Allard

- **Y. Sun** : A Stochastic Space-time Model for Intermittent Precipitation Occurrences
- **J. Carreau** : Assessing the impacts of the choice of spatial dependence structure for flood-risk rainfall

---

18:30 *Wine tasting: Meeting point, Place Pie, in front “Les Halles”*

21:00 *Light show “Luminescences” at the Palais des Papes.*
www.lesluminessences-avignon.com

Friday 19th September

9:00 – 10:30 **Multivariate and multisite Weather Generators**
Moderator: P. Thomson

- **M. Genton :** Cross-Covariance Functions for Multivariate Geostatistics
- **J. Chen** : Stochastic generation of precipitation and temperature: from single-site to multi-site
- **M. Bourotte** : A stochastic weather generator based on new spatio-temporal cross-covariance functions

10:30 – 11:00 *Coffee break*

11:00 – 12:30 **Multivariate and multisite Weather Generators, ensemble forecasts and diagnostic approaches**
Moderator: P. Thomson

- **R. Chandler** : Rglimclim: a multivariate, multisite daily weather generator for climate change impact studies
- **R. Schefzik** : Physically coherent probabilistic weather forecasts via ensemble copula coupling (ECC)
- **R. Girard** : Diagnostic approaches for scenarios of short-term wind power generation

12:30 – 14:00 *Lunch at University restaurant*

14:00 – 15:00 **General Discussion**

“SWG in 2020: roadblocks, challenges and future directions”

Thursday 18th September

9:00 – 10:30 **Downscaling**

Moderator: V. Monbet

- **S. Sain :** Impact of resolution in dynamic downscaling experiments
- **M. Vrac** : Conditional Stochastic Weather Generators for precipitation downscaling
- **A. Bechler** : Conditional Simulations of max-stable processes for the extreme downscaling

10:30 – 11:00 *Coffee break*

11:00 – 12:30 **Special statistical topics**

Moderator: V. Monbet

- **P. Thomson** : Estimating the slope and standard error of a long-term linear trend fitted to adjusted annual temperatures
- **T. Carey-Smith** : Stochastic rainfall seasonality: estimation and applications
- **W. Kleiber :** High Resolution Nonstationary Weather Simulation

12:30 – 14:15 *Lunch at University restaurant*

14:15 – 15:45 **Simulating Sea States / Physical model**

Moderator: P. Ailliot

- **Y. Guanche** : Climate-based multivariate simulation technique of sea states
- **K. Kpogo Nuwoklo:** Assessment in the form of metocean events of the swell climate in West Africa
- **V. Resseguier** : Reduced flow models from a stochastic Navier-Stokes representation

15:45 – 16:15 *Coffee break*

16:15 – 17:15 **Analogues and resampling methods**

Moderator: P. Ailliot

- **F. Oriani** : Daily rainfall simulation: reproducing high-order statistics with the Direct Sampling technique
- **P. Yiou** : Weather and event generators based on analogues of atmospheric circulation

19:45 *Workshop Dinner, « Le Bercail », Ile de la Barthelasse.*
Meeting at 19:15, Place de l'Horloge.

Space-time simulation of intermittent rainfall with prescribed advection field: adaptation of the turning band method

ETIENNE LEBLOIS(1), JEAN-DOMINIQUE CREUTIN(2)

(1) Irstea Lyon, F-69626 Villeurbanne Cedex, France. mail : etienne.leblois@irstea.fr

(2) CNRS / UJF-Grenoble 1 / G-INP / IRD, Laboratoire d'étude des Transferts en Hydrologie et Environnement, LTHE, UMR 5564, Grenoble, F-38041, France

Space-time rainfall simulation is useful to study questions like for instance the propagation of rainfall-measurement uncertainty in hydrological modeling. This talk adapts a classical Gaussian field simulation technique, the turning band method, in order to produce sequences of rainfall fields satisfying three key features of actual precipitation systems: i) the skewed point distribution and the space-time structure of non-zero rainfall (NZR); ii) the average probability and the space-time structure of intermittency; iii) a prescribed advection field. SAMPO (for Simulation of Advected Mesoscale Precipitations and their Occurrence) assembles various theoretical developments available from the literature. Coordinate changes introduce advection in the heart of the TBM. TBM outputs transformation into rainfall fields with a desired structure is controlled using Chebyshev-Hermite polynomial expansion. The intermittency is taken as a binary process statistically independent of the NZR. Example runs demonstrated the combined role of Taylor’s and advection velocities in the variability of rainfall as experienced by observers.

Nawal AKROUR, Aymeric CHAZOTTES, Sébastien VERRIER, Laurent BARTHES et Cécile MALLET

LATMOS - FRANCE

Abstract: Generation of Rainfall time series from micro to large scale

Rainfall is the result of many complex physical processes that induce particular features and make its observation complex. Simulation tools can be very helpful to generate synthetic rain fields able to represent “true” rain features at small spatial or temporal scales. These rain fields can be the starting point of many studies among which the test of the performances of a particular embedded observation device, beam filling effects, design of new systems, rain retrieval intercomparaison methods …

Due to its strong links with turbulence, precipitations have scale invariance properties leading to a power law energy spectrum. The Universal Multifractal Model (UMM) Schertzer and Lovejoy [1987] can be used to simulate geophysical fields with scaling properties. It is particularly well suited for various geophysical fields where turbulence is involved. In the case of rain, the problem is more complicated since the rain spectrum display several scaling regimes mainly due to the alternation of dry and wet/rainy periods (rain support) and high variability of rain. Indeed the UMM does not allow generate zero values and thus to represent dry periods. It follows that the use of the standalone UMM generates synthetic data that do not exhibit all the statistical properties of actual rainfall time series.

In this study, we propose a model based on UMM and allows generating synthetic rainfall time series at a fine resolution (15 seconds). In addition to zero values it displays the same statistical properties (multifractal properties, statistical distribution of rain intensity, events duration and volume, fractal co-dimension, multiple scaling regimes, autocorrelation) as actual measurements. The data used to develop the simulator were collected during 2 years in Palaiseau (France) at a resolution of 15 s.

Initially, we analyze the observed data in term of rain support. We find independence between the durations of rainy and dry periods. We are thus able to characterize rain time series as successive “rain events” spaced out by dry periods. It consequently leads to our simulation process of rainfall time series. The analysis of the observed rainy and dry periods’ probability distribution functions (pdf) shows a break at 5 minutes. To model dry and rainy periods, we have taken into account the presence of two behaviors. A Pareto distribution was used to model the short (<5 minutes) and long (>5minutes) durations. The simulated support showing the same fractal properties as the real one.

Then, UMM parameters are estimated on actual rain events. Events are simulated using UMM taking into account the average relation observed between intensities and durations of all events.

The resulting generator can simulate synthetic series of more than 2 years at a 15 seconds resolution that exhibit statistics properties at all scales that are very close to those observed. The necessary modifications to simulate synthetic rainfall series corresponding to other locations and other durations will be discussed.

Finally, in the last part the extension to the simulation of 2D rain maps is also discussed and some examples of synthetic maps are presented.

References:

Schertzer, D., S. Lovejoy, 1987: Physically based rain and cloud modeling by anisotropic, multiplicative turbulent cascades. J. Geophys. Res. 92, 9692-9714

Verrier S., C. Mallet, L. Barthès, (2011) Multiscaling properties of rain in the time domain, taking into account rain support biases, Journal of Geophysical Research-Atmospheres **116,** D20119

Spatio-temporal precipitation generator with a censored latent Gaussian field

Anastassia Baxevani, Department of Mathematics and Statistics, University of Cyprus

A daily stochastic spatio-temporal precipitation generator that yields precipitation realisations that are quantitative consistent, is described. The methodology relies on a latent Gaussian field that drives the precipitation occurrence and a transformed latent Gaussian field that models the intensity. At individual locations the model reduces to a Markov chain that describes the occurrence process and a composite of a gamma with a Generalized Pareto distribution for the precipitation intensity. Statistical parameters are estimated from data and extrapolated to locations and times with no direct observations using Delanay tesselation together with barycentric interpolation. One advantage of such a model is that stochastic generator parameters are readily available at any location and time of the year. The methodology is illustrated on a data set of daily rainfall values from Sweden. Performance of the model is judged through its ability to accurately reproduce a series of weather indices.

Markov-Switching AutoRegressive models for Cartesian components of wind fields in the North-East Atlantic

Julie Bessac[1], Pierre Ailliot[2], Valérie Monbet [1]

[1] *Institut de Recherche Mathématiques de Rennes, UMR 6625, Université de Rennes 1, France*
[2] *Laboratoire de Mathématiques de Bretagne Atlantique, UMR 6205, Université de Brest, France*

Several multi-site generators of (u, v) wind conditions are proposed in this work. A hidden Markov-Switching Vector AutoRegressive model is introduced and compared to unconditional and several conditional Vector AutoRegressive models with observed regime-switching. Various questions are explored such as the modeling of the regime in a multi-site context, the extraction of relevant classifications of wind data and the link between regimes extracted from the fitting of the hidden MS-AR model and large-scale weather types derived from descriptors of atmospheric circulation.

Variability of wind speed encountered by a vessel

IGOR RYCHLIK
Mathematical Sciences,
Chalmers University of Technology,
SE-412 96 Göteborg, Sweden.

Abstract

Wind speeds are modeled by means of a transformed Gaussian process, where a power transformation is applied. Its dependence structure is localized by introduction of time and space dependent parameters in the covariance function. The model has the advantage of having a relatively small number of parameters. These parameters have natural physical interpretation and are statistically fitted to represent variability of observed wind speeds in ERA Interim reanalysis data set. The model is validated using the on-board measured wind speeds on vessels sailing in Northern Atlantic.

Conditional Modeling of Extreme Wind Gusts by Bivariate Brown-Resnick Processes

Marco Oesting
INRA/AgroParisTech, France
Email: *oesting@agroparistech.fr*

Martin Schlather
University of Mannheim, Germany

Petra Friederichs
University of Bonn, Germany

Abstract

In order to incorporate the dependence between the spatial random fields of observed and forecasted maximal wind gusts, we propose to model them jointly by a bivariate Brown-Resnick process. As there is a one-to-one correspondence between bivariate Brown-Resnick processes and pseudo cross-variograms, stationary Brown-Resnick processes can be characterized by properties of the underlying pseudo cross-variogram. We particularly focus on the investigation of their asymptotic behavior and introduce a flexible parametric model both being interesting in classical geostatistics on their own. The model is applied to real observation and forecast data for 110 stations in Northern Germany. The resulting post-processed forecasts are verified.

The role of microclimates in climate change responses: ecologists need climatic data with high resolution

Sylvain Pincebourde, Institut de Recherche sur la Biologie de l'Insecte, IRBI CNRS UMR 7261, Parc Grandmont, 37200 Tours, France.

sylvain.pincebourde@univ-tours.fr

Organisms do not necessarily experience climatic conditions as measured at global scale by meteorological stations. Instead, they live in microhabitats with specific climatic conditions—the microclimates. For most species, microclimates can deviate substantially from global weather. Despite the complexity of the multiple interacting factors that determine those microclimates, the tools necessary to comprehend and model them are available today. The aim of this presentation is to show the necessity for integrating microclimates into ecological frameworks to increase our ability to forecast climate change impacts on ecological processes. Biophysical models used to compute the heat budget of organisms need to be run with climatic data, either historical or projected. However, the composite non-linearity in these models implies that we integrate climatic data with a high resolution both in space and in time. Stochastic weather generators are potentially ideal candidates to temporally downscale projected climatic data according to future global warming scenarios. My presentation reports the point of view of en ecologist, as a potential user of such SWGENs.

A stochastic weather generator for temperature, examples of use and future developments

S. Parey(2)*, T.T.Huong Hoang* (2)*, D. Dacunha-Castelle*(1)

(1) Laboratoire de mathématiques, Université Paris-Sud (Orsay)
(2) Electricité de France, Recherche & Développement (Chatou)

Since 2010, a new stochastic temperature generator has been proposed, improved and used at EDF/R&D. Developed in the framework of a phD thesis in collaboration with the mathematics laboratory of Paris 11 university, the generator is a Seasonal Functional Heterocedastic Auto-Regressive model designed after careful studies of the trends, seasonalities and extremes of observed temperature time-series. It is based on the simulation of the residuals, after removing trends and seasonalities in the mean and the variance, with a detailed treatment of the extremes. This allows simulations of as many temperature time-series as desired. The obtained time-series build a large sample of temperature similar to the observed ones, which allows making more robust inferences on extreme events like hot or cold spells. Furthermore, using suitable assumptions on the changes in mean, variance and seasonalities, a similarly large sample of temperature time-series for future climate can be produced and used to study the changes in extremes. After a description of the model and its validation, different examples of use will be presented, together with some perspectives for its improvement. Then, the very first results of preliminary studies toward a coupled temperature and rainfall generator, still with a special focus on extremes, will be discussed.

A Stochastic Space-time Model for Intermittent Precipitation Occurrences

Ying Sun[1] and Michael L. Stein[2]

June 7, 2014

Abstract

Modeling a precipitation field is challenging due to its intermittent and highly scale-dependent nature. Motivated by the features of high-frequency precipitation data from a network of rain gauges, we propose a threshold space-time t random field (tRF) model for 15-minute precipitation occurrences. This model is constructed through a space-time Gaussian random field (GRF) with random scaling varying along time or space and time. It can be viewed as a generalization of the purely spatial tRF, and has a hierarchical representation that allows for Bayesian interpretation. The randomness of the scaling process increases the variability across realizations from the GRF, and is shown to capture the variability of the precipitation occurrence better than the threshold GRF model for the data we have considered. For model comparisons and diagnostics, we focus on evaluating whether models can produce the observed conditional dry and rain probabilities given the neighboring sites have rain or no rain. The conditional probabilities, along with the marginal rainfall probabilities, are used to summarize the variability of the precipitation occurrence in space and time, and useful graphical tools are developed for visualization purpose. Model fitting and validation are conducted by Monte Carlo simulation-based approaches, where the statistical efficiency is presented by visualizing a set of the conditional probabilities calculated from simulations of the fitted models.

Some key words: Conditional probabilities; Gaussian random fields; Monte Carlo methods; random scaling; spatio-temporal dependence; t random fields.

Short title: A Space-time Model for Precipitation Occurrences

---

[1] CEMSE Division, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia. E-mail: sunwardstat@gmail.com

[2] Department of Statistics, University of Chicago, Chicago, IL, 60637. E-mail: stein@galton.uchicago.edu

Assessing the impacts of the choice of spatial dependence structure for flood-risk rainfall

Julie Carreau and Christophe Bouvier, IRD HydroSciences Montpellier

June 19, 2014

The Cévennes mountain range lies at the North of the Mediterranean coast in France and peaks at 1699 m. The combination of both the complex orography and the Mediterranean sea influences explain the high variability of rainfall in the region and the occurrence of intense rainfall events which may trigger floods and landslide with dramatic human and material consequences. Hence, the analysis of extreme rainfall is a critical step in the design of civil engineering structures (e.g. dams, reservoirs and bridges) or for urban and landscape planning.

Flood risk is measured in terms of high return levels of river runoff. These return levels may be estimated empirically from long series of river runoff. Since these are rarely available, hydrologists often rely on the following procedure : a stochastic rainfall generator simulates long series of pseudo-rainfall observations which are fed to a rainfall-runoff model from which a long series of pseudo-river runoff is obtained. When using spatially-distributed rainfall-runoff models, spatially-distributed rainfall generators are required.

Our goal is to study the impact of the choice of spatial dependence structure of the rainfall generator on (1) spatial features of the rainfall fields such as the spatial average of the simulated rainfall field and on (2) the runoff high return levels. We work in a multivariate framework where each random variable is the rainfall at one site. We focus on the class of rainfall whose spatial average is greater than 50 mm since this might lead to flooding.

The Gaussian copula has been often used in stochastic rainfall generators and is taken as our benchmark model. We compare several multivariate dependence structures against the Gaussian copula such as the Student t copula, the Skew Normal and Skew t multivariate distributions and the Heffernan and Tawn conditional model for extremes. When the model does not specify the margins, we used either Gamma or semi-parametric margins with a Generalized Pareto in the upper tail of the distribution.

Preliminary results comparing the spatial average of rainfall simulated by the different generators show the following : symmetric dependence structures such as the Gaussian and Student t copulas are not able to reproduce the shape of the distribution of the observed spatial average of rainfall whereas asymmetric dependence structures such as the skew Normal or the Heffernan and Tawn model seem more appropriate. Only the latter model seems to be able to reproduce the upper tail of the distribution, that is the occurrence of heavy spatial rainfall with large spatial average.

The paper finally evaluates the bias on flood return levels in the Gardon at Anduze catchment (545 km2), when using the different dependence structures in the rainfall generator. The distributed event-based rainfall-runoff model SCS-LR has been previously calibrated for simulating the floods from the spatial generated rainfalls. As we use an event-based model, the initial saturation of the catchment must also be generated from the distribution of a wetness index. The results show that the bias in the estimation of the rainfalls increases when estimating the flood discharges or return levels.

This is joint work with : Gwladys Toulemonde, Luc Neppel and Etienne Leblois.

Impact of resolution in dynamic downscaling experiments

Steve Sain, IMAGe, National Center for Atmospheric Research

Understanding the role of model resolution and the interaction with model components is becoming an increasingly important aspect of climate modeling. In this work, I will present an analysis of a regional climate model experiment focusing on monthly precipitation and understanding the interaction between model resolution and convective parameterizations. I will also present an approach for examining the added value in downscaling efforts. Additionally, I will tie both of these approaches to efforts for spatial statistical models for large datasets.

Conditional Stochastic Weather Generators for precipitation downscaling

Mathieu Vrac
Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL)
mathieu.vrac@lsce.ipsl.fr
http://www.lsce.ipsl.fr/Pisp/58/mathieu.vrac.html

Global Climate Models (GCMs) are the key-tools to model the main climate variables (precipitation, temperature, pressure, etc.) and their future evolutions under various greenhouse gas emission scenarios. However, their spatial resolution of about 250 km is inappropriate for many impacts studies, requiring input data at regional (a few km) or even local (a few hundred meters) scale. It is then necessary to "downscale" the large-scale GCM outputs to values at the relevant resolution. If the physical approach is based on regional climate models (RCMs) that solve the main physical equations of the regional dynamics, the so-called "statistical" downscaling approach relies on statistical models or relationships linking large-scale data and local-scale observations. Among the various statistical methodologies (e.g., weather typing, transfer functions, model outputs statistics), the "Stochastic Weather Generators" (SWGs) appear probably as the most flexible since they are explicitly based on probability density functions (pdf).

In this talk, two SWGs will be presented to statistically downscale daily precipitation at rain gauges. If other climate variables may obviously be of interest (e.g., temperature or wind), the downscaling of precipitation is still a challenging task du to their strong spatial and temporal variability, the number of impact models for which it is the key variable and the fact that the society is very vulnerable to their extreme events. Both presented SWGs try to model the variability of daily precipitation (occurrences and intensity) through densities whose the parameters depend on some large-scale data or information.

Although, the results show the good behaviour of the proposed models, the extreme rainfall (i.e., the very large precipitation intensity) is underestimated for some rain gauges. We then tried to extend one of the SWGs by including extreme values modelling through a distribution from the so-called "extreme values theory". This extension will be presented and an illustration showing the main interest of this approach will be discussed.

CONDITIONAL SIMULATIONS OF MAX-STABLE PROCESSES FOR THE EXTREME DOWNSCALING

Aurélien Bechler [1,2,3] & Liliane Bel [1,2] & Mathieu Vrac [3]

[1] *INRA, UMR 518 Math. Info. Appli., F-75005 Paris, France and aurelien.bechler@agroparistech.fr*
[2] *AgroParisTech, UMR 518 Math. Info. Appli., F-75005 Paris, France and liliane.bel@agroparistech.fr*
[3] *LSCE-IPSL, Centre d'Etudes de Saclay, Orme des Merisiers, Bat. 701 91191 Gif-sur-Yvette, France and mathieu.vrac@lsce.ipsl.fr*

Abstract. For a few years, regional climate models are used to provide future scenarios of precipitation with increasingly higher resolution. However, this resolution is not yet sufficient to describe efficiently what happens at local scale. Methods of downscaling (Maraun et al., 2010) have been developed and allow us to make the link between these two levels of resolution. Nevertheless, the extreme behavior is not well described by these downscaling methods. We propose a two step methodology to solve this problem. The first one consists in building a statistical link-function between the high and low resolution variables at some given locations. Once this univariate downscaling is performed, the conditional simulation algorithm of max-stable processes proposed by Dombry et al. (2013) and adapted to the extremal t process (Bechler et al., 2014) enables us to get extreme precipitation distributions at any point of the map. Application is performed on precipitation data in the south of France where extreme (cevenol) events have major impact (e.g. floods).

Estimating the slope and standard error of a long-term linear trend fitted to adjusted annual temperatures

Peter Thomson
Statistics Research Associates Ltd, New Zealand

Brett Mullan Stephen Stuart
National Institute of Water and Atmospheric Research, New Zealand

Abstract

This paper is concerned with estimating the slope of a linear trend fitted to a long annual temperature time series that has been adjusted for level shifts at a known set of times or changepoints. It determines and evaluates the impacts of the adjustments on the statistical properties of the slope estimates. These impacts could be considerable, especially on the standard error of the slope which is inversely proportional to $T^{3/2}$ where T is the length of the time series. Any reduction in T, for whatever reason, may well lead to a significant loss of precision. Two approaches are considered. The first assumes that the data has been prior adjusted for level shifts using data from neighbouring stations and simple local adjustment methods. The second estimates the slope and level shifts directly from the original temperature time series available at a given location and does not use data from neighbouring stations. These two procedures are compared and their statistical properties developed including, in particular, formulae for the standard errors of the respective slope estimates. The procedures are applied to a time series of adjusted annual temperatures recorded at a site in Wellington, New Zealand.

Stochastic rainfall seasonality: estimation and applications

Trevor Carey-Smith (National Institute of Water and Atmospheric Research, New Zealand)
Peter Thomson (Statistics Research Associates Ltd, New Zealand)

A hidden seasonal switching model for daily rainfall over a region is described where season onset times are stochastic and can vary from year to year. The model allows seasons to occur earlier or later than expected and have varying lengths. This stochastic seasonal variation accommodates considerably more of the observed intra-annual rainfall variability than can be represented using seasonal models with standard fixed seasons.

Methods used to estimate season onset probabilities from multi-year and multi-site data are described, touching on the apparent variability in seasonality between, and sometimes within, relatively small spatial regions. Relations between season onset probabilities and large scale climate drivers relevant to New Zealand, such as the Southern Annular Mode and the strength/location of the westerly jet, are explored with a view toward creating synthetic rainfall time-series downscaled from general circulation models.

High Resolution Nonstationary Weather Simulation

William Kleiber, Applied Mathematics, University of Colorado at Boulder

Abstract: Stochastic weather generators (SWGs) are used in many scientific studies, including model downscaling, climate impact assessment and seasonal resource planning. The fundamental requirement of a stochastic weather generator is simulated realizations of plausible weather patterns. In recent years, focus has shifted to developing spatially-consistent SWGs. Unless the region of interest is relatively small, or has homogeneous topography, the SWG will necessarily require simulation of nonstationary spatial fields. However, high resolution simulation of a nonstationary process is difficult, typically requiring a Cholesky decomposition of a matrix whose dimension equals that of the desired simulation resolution. We introduce an approach to large, high resolution nonstationary process simulation by exploiting ideas very similar to Sampson and Guttorp (1992), relying on spatially deforming geographical space to achieve approximate stationarity, then using fast stationary simulation algorithms, followed by an inverse transformation back to the nonstationary plane. We illustrate the algorithm on simulated and real datasets.

CLIMATE-BASED MULTIVARIATE SIMULATION TECHNIQUE OF SEA STATES

Y. Guanche, R. Mínguez, F. J. Méndez, IFREMER, France

The design of maritime structures, offshore devices or any coastal process requires information on sea state conditions that influence its behavior during its life cycle. In the last decades, there has been a increasing development of sea databases (buoys, reanalysis, satellite) allowing an accurate description of the marine climate and its interaction with a given structure in terms of functionality and stability. However, these databases have a limited time length, and its appliance entails an associated uncertainty. To avoid this, synthetically generated time series of the variables (eg. H_s, T_p, θ, storm surge, wind speed) offer a practical alternative.

The present work (Guanche et al. (2013b)) proposes a hybrid methodology to deal with this issue. This methodology combines different techniques and algorithms in order to be able to take into account different time and space scales. It comprises three interrelated steps. In the first step, synthetic daily sea level pressures fields (DSLP) in the wave generation area, decomposed into principal components, are simulated by using the multivariate simulation technique proposed by Morales et al. (2010). During the second step, daily mean sea conditions (DMSC), clustered by K-means, are simulated by applying an autoregressive logistic model (Guanche et al. (2013a)) and taking into account the previously simulated DSLP as covariates. The third step consists of a modified version of the methodology proposed by Morales et al. (2010), to be used with hourly sea state (HSS) parameters and conditioned to the catalogue of synoptic DMSC patterns simulated in the previous step. Within the third step, the extremes characterization is carried out by fitting a Pareto distribution over a predetermined threshold according to Mínguez et al. (2012).

To show the proposed method a case of study has been considered using high resolution hindcast databases developed by IH Cantabria and NCEP-NCAR. The direct comparison between simulated and empirical time series confirms the ability of the developed methodology to generate multivariate hourly time series of sea states. The synthetic time series obtained are statistically consistent and keep the temporal dependence structure of the initial stochastic process.

References

Guanche, Y., Mínguez, R., Méndez, F.J., 2013. *Autoregressive logistic regression applied to atmospheric circulation patterns.* Climate Dynamics http://dx.doi.org/10.1007/s00382-013-1690-3.

Guanche, Y., Mínguez, R., Méndez, F.J., 2013. *Climate-based Monte Carlo simulation of trivariate sea states*. Coastal Engineering . 80 107-121.

Mínguez, R., Guanche, Y., Méndez, F.J., 2013. *Point-in-time and extreme-value probability simulation technique for engineering design.* Structural Safety 41, 29–36.

Morales, J.M., Mínguez, R., Conejo, A.J., 2010. *A methodology to generate statistically dependent wind speed scenarios.* Applied Energy 87, 843–855.

Assessment in the form of metocean events of the swell climate in West Africa

K. A. Kpogo-Nuwoklo[1], M. Olagnon [1], Z. Guédé [1], P. Ailliot[2], S. Arnault[3]

July 7, 2014

[1]Laboratoire Comportement des Structures en Mer, IFREMER, Brest, France.
[2] Laboratoire de Mathématiques de Bretagne Atlantique, Université de Bretagne Occidentale, Brest, France.
[3] LOCEAN/IPSL, Université Pierre et Marie Curie, Paris, France.

Abstract

Accurate estimation of the long-term occurrence probabilities of sea conditions is a major issue for the design of offshore structures and for the preparation of marine operations. Because the large number of parameters involved, that can only be done through climate statistical modeling. In this study, a new approach is proposed to model the wave climate. It relies on a partition of the time-sequences of sea states parameters in accordance to the oceanic and meteorological events (e.g.pressure lows, storms, ...) that are at the source of the phenomena. The advantage is to provide a structure with physical meaning and temporal coherence that allow to represent the occurrence joint probabilities. The final goal is to reconstruct a wave climate over arbitrary long durations for engineers' uses.

The first part of this study is dedicated to the identification and extraction of temporal sequences of wave systems parameters that are the signatures of remote storms (swell events). It is carried out with a method for image segmentation (Watershed algorithm). Then a parametric model is proposed and fitted to individual swell event based on physical considerations. The joint distributions of the swell events' parameters are provided as well as a stochastic model for the occurrence process of the events.

Reduced flow models from a stochastic Navier-Stokes representation

V. Resseguier, E. Mémin, B. Chapron
INRIA, Fluminance group and IFREMER, LOS

Navier-Stokes stochastic model

This article will use an Eulerian stochastic description of velocity and tracer evolution, as in [1]. Unlike classical stochastic methods, a random part is added to the Lagrangian velocity before any model derivation. The time differentiation of a trajectories X_t of a particle is:

$$dX_t \quad = \quad w(X_t,t)dt + \sigma(X_t,t)dB_t \tag{1}$$

where $\forall f \in (L^2(\mathbb{R}^d,\mu))^d$, the finite-variance function space, $\sigma(.,t)f \overset{\triangle}{=} \int_\Omega \breve{\sigma}(.,y,t)f(y)dy$. $(x,t) \mapsto \sigma(x,t)dB_t$ is a centered gaussian process with covariance $a(x,y)\delta(t-t')dt$ where $a(x,y) = \sigma(x)\sigma(y)^t = \int_\Omega \breve{\sigma}(x,z)\breve{\sigma}^t(y,z)dz$. Thanks to this decomposition, it is possible to derive a stochastic representation of the so-called Reynolds transport theorem, cornerstone of the Navier-Stokes equation derivation in the deterministic case. Similarly to the Newton 2^{nd} law, a dynamical balance, between the temporal differentiation of the stochastic momentum, ρdX_t, and general stochastic forces actions, is assumed. This leads to the following stochastic Navier-Stokes representation:

$$\rho\left(\frac{\partial w}{\partial t} + (w\cdot\nabla)w + f\times w\right) = \tau(w) + \rho g - \nabla p' + f_V(w), \tag{2}$$

$$\rho\left((w\cdot\nabla)\sigma + f\times\sigma\right)dB_t = -\nabla d\hat{p} + f_V(\sigma)dB_t, \tag{3}$$

$$\text{where} \left\{ \begin{array}{rcl} \forall k, \tau_k(w) & = & \frac{1}{2}\left(\nabla\cdot(\nabla\cdot(\rho a w_k))^t - \nabla\cdot(\nabla\cdot(\rho a))^t w_k - 2\rho\left((\nabla\cdot\sigma)\sigma^t\nabla\right)w_k\right), \\ f_V(g) & = & \mu\left(\nabla^2 g + \frac{1}{3}\nabla\left(\nabla\cdot g\right)\right). \end{array} \right.$$

Here, in order to explain the fluctuation of a the time-uncorrelated part of the velocity, the integral of the pressure along time, $\int_0^t pdt$ was replaced by a general continuous semi-martingale $\int_0^t (p'dt + d\hat{p})$. Expression (2) can be seen as a generalization of several classical turbulence model. One can wonder whether τ is dissipative, as in a theoretical 3D direct energy cascade ([2], [3]). This is the case, for instance, if ρ is assumed to be constant and $\nabla\cdot(\nabla\cdot a)^t \geqslant 0$. The last constraint can be forced, in practice, keeping the definite non negative structure of a in each point of Ω.

The knowledge of small- scale physical flow realizations allows estimating σ and a:

$$adt \quad = \quad \mathbb{E}((\sigma dB_t)(\sigma dB_t)^t) \text{ and } (\nabla\cdot\sigma)\sigma^t dt = \mathbb{E}((\nabla\cdot\sigma dB_t)(\sigma dB_t)^t). \tag{4}$$

The value of $a(x,t)$ can be, then, used in a large-scale simulation, ruled by equation (2). The estimation (4) can be performed, for instance, through model reductions and Galerkin projections, as explained later in this paper. Another interesting way of using (4) is through Monte-Carlo small-scale simulations, such as particule filtering, where particles correspond to several likely values of the small-scale velocity. A third

example of large eddies simulation methods would be to use directly statistical properties of small-scale velocity, inferred from statistical knowledge of spatial measurements errors, to compute a, without using (4). Some works on this subject are currently performed and will be part of a full paper. In all these methods, the estimation of a is a solution of a closure problem and the knowing of a can lead to a better simulation of the drift, through (2), or of a tracer transport, through the stochastic transport theorem ([1]).
One can also have the dual point of view with a down scaling approach, as, for instance, mixing diagnostics, which is a important current issue in Meteorology and Oceanography ([4]). The equation (2) and the averaged stochastic transport theorem ([1]) applied to large-scale geophysical data can give information on the likelihood of locally strong gradient of the tracer and or of the velocity. In more complex stochastic geophysical models, a and σ could be infered from the observed advection-diffusion of an active tracer. The diffusion tensor, σ, can also be inferred from w, as proposed in [1] , through the projection on the free-divergent space on (3), which leads to the resolution of a system of Poisson equation on $\breve{\sigma}(.,z,t)$.

Model reduction

The Proper Orthogonal Decomposition (POD) is a decomposition, of an observed multivariate field such as a velocity $u(x,t)$. It is similar to a Principal Components Analysis (PCA) in Statistics. In POD, a spectral analysis of time (or space) autocorrelation tensor of observed data, leads to the decomposition:

$$\forall (x,t) \in \mathbb{R}^d \times \mathbb{R},\ u(x,t) \approx \bar{u}(x) + \sum_{i=1}^{N} b_i(t)\phi_i(x), \tag{5}$$

where $\|b_1\|^2_{L^2([0,T])} > ... > \|b_N\|^2_{L^2([0,T])}$, $\bar{u}$ is the temporal mean of u, N the number of observed snapshots and $(\phi_i)_{1\leqslant i\leqslant N}$ is an orthonormal set of $L^2(\mathbb{R}^d)^d$ which spans the same space as $((u(.,t_i)-\bar{u})_{1\leqslant i\leqslant N})$. In the following, $\bar{u}$ will be noted ϕ_0 and $b_0 \overset{\triangle}{=} 1$. Then, since only the first temporal modes have significant energy, a second approximation is done : $\forall (x,t) \in \mathbb{R}^d \times \mathbb{R}, u(x,t) \approx \sum_{i=0}^{n} b_i(t)\phi_i(x)$ where $N > n$.
A Galerkin projection enables us to look for an approximate solution of a PDE. This approximate solution at a fixed time, $u(.,t)$, is assumed to live in a finite-dimensional sub-space, $Span(\phi_1,...,\phi_n)$, instead of an infinite-dimensional one. Then, the time-space evolution equation of u (a PDE) becomes equivalent to the time evolution equations (a finite set of coupled ODEs) of the coefficients of u on the orthonormal basis $(\phi_i)_{1\leqslant i\leqslant n}$.
In Fluid Mechanics, the less energetic modes have smaller characteristic time and spacial quantities, and, thus, can be interpreted as small-scale velocity. Since, small scales interacts a lot with the large-scale resolved modes, the classical Navier-Stokes equation cannot be projected on $Span(\phi_1,...,\phi_n)$. The solution presented by [1] is to decomposed u as follow : $udt = wdt + \sigma dB_t$ with $w = \sum_{i=0}^{n} b_i\phi_i$ (projection on the truncated subspace) and $\sigma dB_t = \sum_{i=n+1}^{N} b_i\phi_i dt$ (projection on the complementary "small-scale" subspace). The drift, w, follows the stochastic Navier Stokes equation (2). Projecting this equation along ϕ_i for each $i \in [\![1,n]\!]$ gives the evolution equation of $b \overset{\triangle}{=} (b_i)_{1\leqslant i\leqslant n}$. $\forall i \in [\![1,n]\!]$, $\frac{db_i}{dt} = i_i + l_{.i}^{t} b + b^t c_{..i} b$, where the coefficients $(i_i)_{1\leqslant i\leqslant n}, (l_{j,i})_{1\leqslant i,j\leqslant n}$ and $(c_{k,j,i})_{1\leqslant i,j,k\leqslant n}$ are computed through the integration over the whole space of the terms of (2).
To compute them, one has to estimate the tensor, a. Since only one realization of the small-scale velocity is

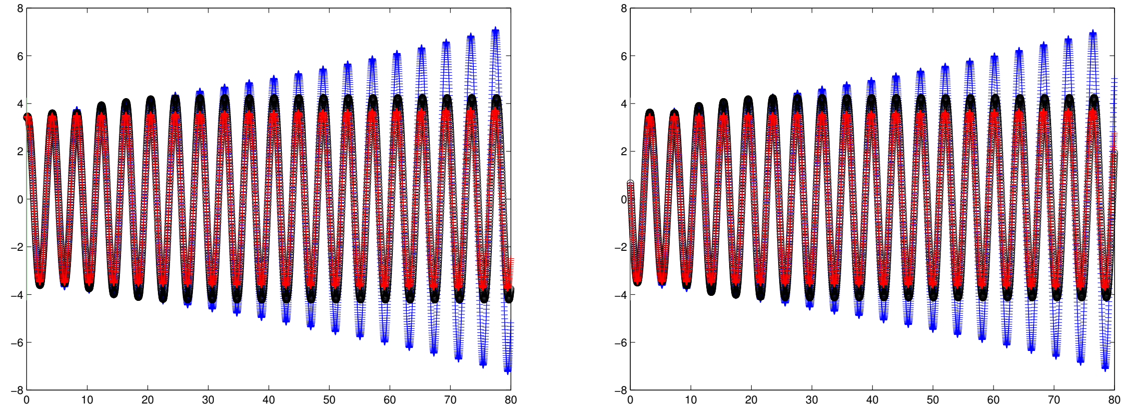

Figure 1: Reconstruction of the 2 first temporal modes (n=2) from a simulated 2D flow (Re=300). The black circles are the reference. The blue cross are computed with the deterministic reduced model whereas the red ones are computed with the stochastic one. The green line labels the common initial condition.

available, some time-ergodicity hypothesis is necessary, at least locally, to use (4). Otherwise, parametric and non-parametric estimation of $a(X_t, t)dt = d < X_t, X_t >$ are studied in literature. Existent non-parametric estimation assume σ either constant in time or in space. But, contrary to usual applications of this methods (finance), we have an Eulerian realization . And then, for all $x \in \mathbb{R}^d$, we can construct a spatially homogeneous local martingale $\tilde{X}_t^x \overset{\triangle}{=} \int_0^t \sigma(x,t)dB_t$. Its realization, $\int_0^t (u(x,t) - w(x,t))dt$, enables estimating, for all functions h_k, $\int h_k(t)a(x,t)dt = \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_i=0}^{T} h^k(t_i)(\tilde{X}_{t_{i+1}}^x - \tilde{X}_{t_i}^x)$. The functions h^k can be a real basis of $L^2([0,T])$ such as wavelets, as in ([5]), or a reduced one such as $(b_i/\|b_i\|_{L^2([0,T])})_{1 \leqslant i \leqslant n}$. Figure **??** shows that, if the decomposition of the tensor a on the reduced basis is well chosen, its influence can be impressive. Figure 2 shows first encouraging results with an estimated but constant in time, tensor.

The talk will present in details the mechanism of the stochastic Navier-Stokes derivation and of the considered model reduction approach. Experiments on the different estimation methods for the diffusion tensor are on going and a comparison of all these alternatives will be provided during the workshop.

References

[1] E. Mémin. Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, 2014.

[2] Geoffrey K Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation.* Cambridge University Press, 2006.

[3] Antti Kupiainen. Statistical theories of turbulence. *workshop Random media 2000*, June 2000.

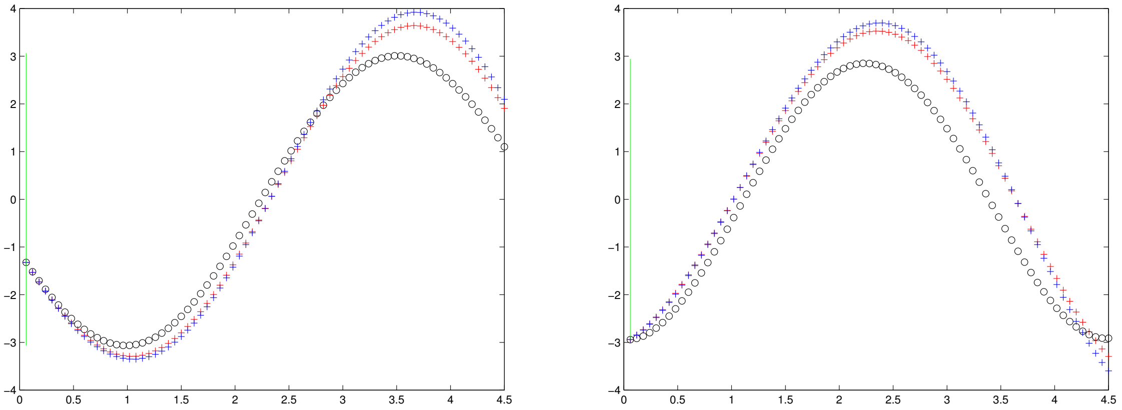

Figure 2: Reconstruction of the 2 first temporal modes (n=4) from a simulated 3D flow (Re=3900) ([6]). The black circles are the reference. The blue cross are computed with the deterministic reduced model whereas the red ones are computed with the stochastic one. The green line labels the common initial condition. Here, the tensor a is estimated, assuming that it is constant in time.

[4] Noboru Nakamura. A new look at eddy diffusivity as a mixing diagnostic. *Journal of the atmospheric sciences*, 58(24):3685–3701, 2001.

[5] Valentine Genon-Catalot, Catherine Laredo, and Dominique Picard. Non-parametric estimation of the diffusion coefficient by wavelets methods. *Scandinavian Journal of Statistics*, pages 317–335, 1992.

[6] Philippe Parnaudeau, Johan Carlier, Dominique Heitz, and Eric Lamballais. Experimental and numerical studies of the flow over a circular cylinder at reynolds number 3900. *Physics of Fluids*, 20(8):085101, 2008.

Daily rainfall simulation: reproducing high-order statistics with the Direct Sampling technique.

F.Oriani[1], R.Merothra[2], J.Straubhaar[1], G.Mariethoz[2], A.Sharma[2], P.Renard[1]

1. Centre for Hydrogeology and Geothermics, University of Neuchâtel.

2. School of Civil and Environmental Engineering, University of New South Wales.

Daily rainfall time-series simulation is a key topic for hydrological and climate science applications: the challenge is to simulate synthetic time-series honoring the reference statistics and persistence from the daily to the higher temporal scale.

We investigate the potential of a novel approach: the Direct Sampling (DS) [Mariethoz 2010], one recent algorithm belonging to the Multiple-Point statistics family. As opposed to the Markov Chain (MC) based techniques, which impose a specific time-dependency, the Direct Sampling aims at reproducing the overall statistical signature by resampling the training dataset and generating similar patterns at multiple scales. Can the DS reproduce the complexity of a daily time-series more accurately than a parametric model expressing a fixed time-dependency? The proposed technique is applied to rainfall records from different climate settings and it is compared to one of the most recent MC techniques. It is demonstrated that, without performing any optimization of the parameters, the Direct Sampling can generate extremely realistic daily rainfall time-series without the need of a complex prior structure.

Weather and event generators based on analogues of atmospheric circulation

Pascal Yiou & Ara Arakelian
Laboratoire des Sciences du Climat et de l'Environnement, (LSCE-IPSL)
pascal.yiou@lsce.ipsl.fr; ara.arakalian@lsce.ipsl.fr
http://www.lsce.ipsl.fr/Phocea/Pisp/index.php?nom=pascal.yiou

Analogues of atmospheric circulation have had numerous applications on weather prediction, climate reconstructions and detection/attribution analyses. A stochastic weather generator based on circulation analogues was recently proposed by Yiou (2014) to generate sequences of European temperatures. One of the features of this weather generator is that it preserves the spatial structure of the variables to be simulated. We also combined this approach of circulation analogues with a method of storm detection (Déroche et al., 2014) to generate a large catalogue (~600000 storms) of high impact extra-tropical storms, out of the ERA-Interim reanalysis data, covering 1989-present. We will present the gist of the method of circulation analogues, and two promising applications for weather generators (extra-tropical storms and the simulation of European cold spells).

References

Déroche, M.-S., Choux, M., Codron, F., and Yiou, P.: Three variables are better than one: detection of European winter windstorms causing important damages, Nat. Hazards Earth Syst. Sci., 14, 981-993, doi:10.5194/nhess-14-981-2014, 2014.
Yiou, P.: AnaWEGE: a weather generator based on analogues of atmospheric circulation, Geosci. Model Dev., 7, 531-543, doi:10.5194/gmd-7-531-2014, 2014.

Cross-Covariance Functions for Multivariate Geostatistics

Marc G. Genton[1,2] and William Kleiber[3]

June 16, 2014

Abstract

Continuously indexed datasets with multiple variables have become ubiquitous in the geophysical, ecological, environmental and climate sciences, and pose substantial analysis challenges to scientists and statisticians. For many years, scientists developed models that aimed at capturing the spatial behavior for an individual process; only within the last few decades has it become commonplace to model multiple processes jointly. The key difficulty is in specifying the cross-covariance function, that is, the function responsible for the relationship between distinct variables. Indeed, these cross-covariance functions must be chosen to be consistent with marginal covariance functions in such a way that the second order structure always yields a nonnegative definite covariance matrix. We review the main approaches to building cross-covariance models, including the linear model of coregionalization, convolution methods, the multivariate Matérn, and nonstationary and space-time extensions of these among others. We additionally cover specialized constructions, including those designed for asymmetry, compact support and spherical domains, with a review of physics-constrained models. We illustrate select models on a bivariate regional climate model output example for temperature and pressure, along with a bivariate minimum and maximum temperature observational dataset; we compare models by likelihood value as well as via cross-validation co-kriging studies. The talk closes with a discussion of unsolved problems.

Some key words: Asymmetry; Co-kriging; Multivariate random fields; Nonstationarity; Separability; Smoothness; Spatial statistics; Symmetry.

---

[1] CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia. E-mail: marc.genton@kaust.edu.sa

[2] Presenter

[3] Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526, USA. E-mail: william.kleiber@colorado.edu

Stochastic generation of precipitation and temperature: from single-site to multi-site

Jie Chen, François P. Brissette

Department of Construction Engineering, École de technologie supérieure, Université du Québec, 1100, Rue Notre-Dame Ouest, Montreal, QC., H3C 1K3, Canada
Email: jie.chen@etsmtl.ca

Stochastic weather generators are computer algorithms that can be used to generate climate variables with an arbitrary length and with no missing values. The statistical characteristics of the generated data are expected to be similar to those of the actual data for a given location. An important implementation of parametric stochastic weather generators has been their use as downscaling tools for climate change impact studies by perturbing their parameters based on climate change signals derived from global or regional climate models. According to the implemented location(s), stochastic weather generators can be classified into single- and multi-site models. The single-site weather generator is expected to generate climate variable for a specific site, while the multi-site model can simultaneously generate spatially correlated climate variables for multiple sites.

This work first presents a Matlab-based single-site weather generator (WeaGETS) for daily precipitation and temperature. Several options are provided by WeaGETS for generating precipitation and temperature. For example, the first-, second- and third-order Markov chains are provided to generate precipitation occurrence. Four probability distributions (exponential, gamma, skewed normal and mixed exponential) are available to generate daily precipitation amounts. Based on a corrected first-order linear autoregressive model, two schemes are used to generate maximum and minimum temperatures depended on whether they are conditioned on each other. A spectral correction approach is included for correcting the underestimation of inter-annual variability, which is a problem common to most weather generators. The single-site weather generator has recently been extended to simultaneously generate precipitation and temperature at multiple sites (MulGETS). The temporally independent but spatially correlated random numbers is used to generate the appropriate spatial correlation for temperature and precipitation.

The performance of both versions has been presented based on extensive testing at several locations under various climates. The impact of using multi-site weather generators in hydrology was also thoroughly investigated. Results clearly show that for large watersheds, multi-site weather generators are necessary to take full advantage of distributed hydrology models. The Matlab codes to both weather generators are freely available on the Mathworks file exchange website.

A STOCHASTIC WEATHER GENERATOR BASED ON NEW SPATIO-TEMPORAL CROSS-COVARIANCE FUNCTION

Marc Bourotte [1], Denis Allard [2] & Emilio Porcu [3]

[1] *bourotte@avignon.inra.fr*
[2] *allard@avignon.inra.fr*
[3] *emilio.porcu@usm.cl*

Multivariate space-time data are increasingly recorded in various scientific disciplines. When analyzing these data, one of the (geo)statistitian's goal is thus to describe the multivariate space-time dependencies. In a Gaussian framework, this necessitates to propose relevant models for multivariate space-time covariance functions, mathematically described as matrix-valued covariance functions for which nonnegative definiteness must be ensured. Separable models are straightforward to build, as well as models that separate out the time dimension or the multivariate aspect. Here, we propose a new fully non separable parametric class of cross-covariance functions for multivariate spatio-temporal Gaussian random fields. Space-time components belong to the (univariate) Gneiting class of space-time covariance functions, with a Matern covariance function in the spatial dimensions. In this class, the smoothness parameter and the scale parameter of the spatial Matern component can be different for each variable. The multivariate components follow the parametrization proposed by Apanasovich et al. (2012). We first show that this model is a valid matrix-valued covariance functions for multivariate space-time random fields. We then illustrate the model on a French dataset of climatic variables.

Rglimclim: a multivariate, multisite daily weather generator for climate change impact studies

Richard Chandler, Department of Statistical Science, University College London

Weather generators based on generalised linear models (GLMs) have been shown in intercomparison studies (e.g. Frost *et al*. 2011; Liu *et al*. 2013) to perform comparably with other state-of-the-art techniques in reproducing a wide range of "weather" properties that are of interest in impacts applications. Moreover, they provide a relatively straightforward route into multisite weather generation, with a proven ability to generate consistent weather sequences simultaneously at multiple spatial locations. To date however, most applications of GLMs in this context have been univariate with the exception of Furrer and Katz (2007) who considered precipitation and temperature simultaneously. There is increasing demand in many application areas for weather generators that are both multisite and multivariate in nature.

This talk will present Rglimclim (http://www.homepages.ucl.ac.uk/~ucakarc/work/glimclim.html) which is a multivariate, multisite weather generator based on GLMs. It is based on the Glimclim software package that has been widely used for univariate weather generation in many countries, but has been updated to allow for the simultaneous generation of multiple weather variables. A user interface in R (www.R-project.org) has also been written to simplify the processes of model fitting, checking and simulation. The flexibility of GLMs allows the structure of the generator to be determined by consideration of the physical relationships between variables, rather than by statistical convenience. This will be illustrated with an example of a weather generator constructed for the catchment of the river Thames in England, which has been guided by the structure of numerical weather prediction models.

References

Frost, A.J., S.P. Charles, B. Timbal, F.H.S. Chiew, R. Mehrotra, K.C. Nguyen, R.E. Chandler, J.L. McGregor, G. Fu, D.G.C. Kirono, E. Fernandez and D.M. Kent (2011). A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. *J. Hydrol*, **408**, 1-18, doi: 10.1016/j.jhydrol.2011.06.021.

Furrer, E.M., and R.W. Katz (2007). Generalized linear modelling approach to stochastic weather generators. *Clim. Res.* **34**, 129-144.

Liu, W., G. Fu, C. Liu and S.P. Charles (2013). A comparison of three multi-site statistical downscaling models for daily rainfall in the North China Plain. *Theor. Appl. Climatol.* **111**, 585-600, doi: 10.1007/s00704-012-0692-0.

Physically coherent probabilistic weather forecasts via ensemble copula coupling (ECC)

Roman Schefzik, Thordis Thorarinsdottir and Tilmann Gneiting
Institute for Applied Mathematics, University of Heidelberg

State of the art weather forecasts depend on ensemble prediction systems, which consist of multiple runs of dynamical numerical weather prediction models differing in the initial conditions and/or the parameterized numerical representation of the atmosphere. Statistical postprocessing of the ensemble forecasts can address biases and dispersion errors. However, current postprocessing approaches are mostly univariate and apply to a single weather quantity at a single location and for a single prediction horizon only. Such methods do not account for dependencies which are crucial in many applications.

To address this, we propose a tool called ensemble copula coupling (ECC), in which the postprocessed forecast ensemble inherits the spatial, temporal and inter-variable dependence pattern of the unprocessed raw ensemble. Essentially, this is achieved by using the empirical copula of the raw ensemble to aggregate samples from predictive distributions obtained by univariate postprocessing techniques.

We study several variants of ECC, discuss relationship to discrete copulas, and assess the predictive performance in an application to weather forecasts over Germany, using real data from the European Centre for Medium-Range Weather Forecasts (ECMWF).

Diagnostic approaches for scenarios of short-term wind power generation

Robin Girard, Department Energy and Processes, MinesParisTech

Scenarios of short-term wind power generation are becoming increasingly popular as input to multistage decision-making problems e.g. multivariate stochastic optimization and stochastic programming. It is intuitively expected that the quality of these scenarios can substantially impact the benefits from their use in decision-making. The scenario can be temporal scenario or more generally multivariate (temperature, wind, ...) spatio-temporal scenario. So far however, the verification of these scenarios is almost always focused on their marginal distributions for each lead time, each variable and each space location only, thus overlooking other interdependence structure.

Diagnostic approaches to their evaluation will be proposed and discussed in this talk. They are inspired by functional data analysis, event-based verification and hypothesis testing. An application to the evaluation of various sets of scenarios of short-term wind power generation demonstrates the interest of the proposed tools.