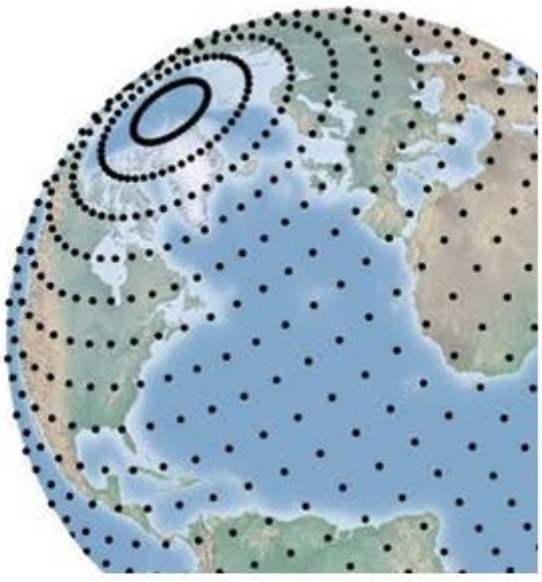
DAMES

Data analysis and modeling in Earth sciences

4th International Conference
Milan, Italy, 6 – 8 October 2014

Technical Program & Abstracts

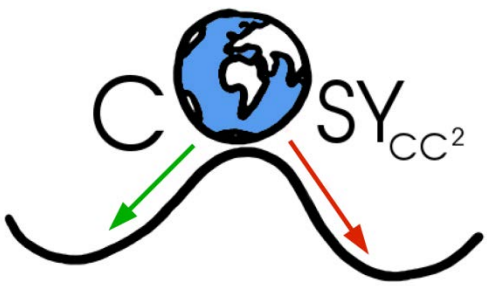


imati

PIK

Potsdam Institute for Climate Impact Research

This conference has been financially supported by the project “CoSy-CC^2: Complex Systems Complex Systems Approach to Understanding Causes and Consequences of Past, Present and Future Climate Change” funded by the German Federal Ministry for Education and Research (grant no. 01LN1306A).



Program schedule

Monday, 6 October 2014

9:30	*Registration*	
10:00	***Welcome remarks***	
10:15	**Johannes Werner**	Climate reconstructions – Challenges in space and time (Keynote lecture)
11:05		*Coffee Break*
11:35	**Dmitri Kondrashov**	Data-driven climate modeling and prediction
12:00	**Marianne Uhlig**	Potential predictability of decadal climate variability using relative entropy
12:25	**Andrej Ceglar**	Assessing the impact of climate variability on crop yields: implications for predictability on seasonal and decadal time scales
12:50	**Andrea Toreti**	A goodness-of-fit test for heavy-tailed distributions and its application to precipitation extremes
13:15	**Diego Barbero**	A statistical approach to the study of the shallow aquifer in the Piedmont region (NW Italy)
13:40		*Lunch Break*
15:15	**Kajsa Parding**	Influence of synoptic meteorological patterns on solar irradiance in northern Europe
15:40	**Shinya Nakano**	Modeling of typhoon trajectory patterns using the Gaussian process regression
16:05	***Poster Teasers***	Short oral summaries of posters
16:30-18:00	***Poster Session***	*and Coffee Break*

Tuesday, 7 October 2014

8:30	*Registration*	
9:00	**Philippe Naveau**	Statistical analysis of heavy rainfall in France via multivariate extreme value theory (Keynote Lecture)
9:50	**Anne Schindler**	On the internal variability of simulated precipitation
10:15	**Marc Wiedermann**	Northern hemisphere ocean-atmosphere coupling from an interacting climate network perspective
10:40	**Diego Guenzi**	Weather radars applications: innovative telecommunication systems and rainfall field reconstructions using a rain gauges network
11:05		*Coffee Break*
11:35	**Rosa Lasaponara**	Satellite fire monitoring based on low cost (Keynote Lecture)
12:25	**Rosaria Tondi**	HPC environment for massive earth data analysis: field examples at regional and continental scale
12:50	**Michael Leuenberger**	Application of Random Forest Algorithm for environmental data
13:15	**Shangguan Wei**	A comprehensive gridded global soil dataset
13:40		*Lunch Break*
15:15	**Bernard de Saedeleer**	Global temperature change: analysis of warming rates of Earth's climate
15:40	**Nikola Jajcay**	Effects of the 7-8-year cycle in daily mean air temperature (from European stations)
16:05	**Maria Caterina Bramati**	Ice wedge thermal variation in East Antarctica: a time series approach
16:30-18:00	***Poster Session***	*and Coffee Break*
19:30	Conference Dinner	

Wednesday, 8 October 2014

8:30	*Registration*	
09:00	**Alvaro Corral**	Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions (Keynote Lecture)
09:50	**Renata Rotondi**	Analysis of macroseismic fields using statistical data depth functions
10:15	**Cataldo Godano**	Spatial organization of foreshocks as a tool to forecast large earthquakes
10:40	**Mikhail Kanevski**	Unsupervised learning of spatial patterns of local seismicity
11:05		Coffee Break
11:35	**Reik Donner**	Different facets of dynamical complexity in the magnetosphere – A recurrence perspective
12:00	**Biagio di Mauro**	Spatial-temporal power law distribution of wildfires in Europe: a comparison between observed and simulated datasets
12:25	**Guillaume Lenoir**	Frequency and continuous time-frequency analysis for some unevenly sampled stochastic processes and implications for significance testing of time series
12:50	**Mario Lefebvre**	Improving the accuracy of river flow forecasts
13:15	**Lucky Imagbe**	Quantitative modelling of turbidity current flow and sedimentation processes, offshore Agadir basin, North West African margin
13:40		Lunch Break
15:15	**Raffaele Argiento**	Bayesian principal curve analysis to detect seismic faults
15:40	**Giada Adelfio**	ETAS model estimation with predictive measure
16:15	**Elisa Varini**	The Gompertz waiting time distributions of the stress release model in the Bayesian framework

16:40	Closing Remarks, Coffee Break & Farewell

Poster Presentations

1.	**Massimo Conforti**	Building a Vis-NIR spectral database for rapid determination organic carbon and nitrogen in forest soils of the southern Calabria (Italy)
2.	**Valeria Rago**	Modelling of landslide spatial hazard along a section of motorway of the Calabria region (Southern Italy) using a GIS-based statistical method
3.	**Felix Schindler**	The localized reduced basis multi-scale method
4.	**Nikola Jajcay**	Effects of the 7-8-year cycle in daily mean air temperature (from European stations)
5.	**Elisa Varini**	A state-space model for the analysis of seismic swarms
6.	**Diego Guenzi**	Permafrost research: monitoring through a sensors network at the 'Istituto Mosso' LTER site (Col d'Olen, Aosta)
7.	**Cataldo Godano**	Variability of the b value in the Gutenberg-Richter distribution
8.	**Shinya Nakano**	Optimization of the smoothness parameters in the Gaussian regression analysis for the modeling of typhoon trajectories
9.	**Michael Leuenberger**	Overview of machine learning applications in environmental data mining
10.	**Marc Wiedermann**	Coincidences of tree-ring responses, simulated vegetation patterns and climate extremes
11.	**Reik Donner**	Complex network analysis identifies impacts of the El Nino Southern Oscillation on moisture divergence in South America
12	**Andrea Toreti**	On the evaluation of climate model simulated precipitation extremes

Abstracts

Keynote presentations

Climate Reconstructions - Challenges in Space and Time

Johannes P. Werner and PAGES2k

Reconstructing past climate is more than just an academic exercise. Knowing about the evolution of past climate can help us not only validating climate models, but also show what the future could bring - things that happened in the past, like extreme droughts (Wetter et al. 2014), are valid states of the climate system that could occur again. It is also important to not only deriving a best estimate of past climate, but rather accounting for uncertainties in these climate reconstructions. Past climate has been recorded by human observations, in direct meteorological measurements and qualitative records, but mostly through natural biological or geochemical processes in what is called proxy records. These natural climate archives, such as tree rings or sediment layers, can then be used to estimate past climate.

We present an overview over some of the most widespread climate reconstruction methods, and show their current limits based on experiments over the Arctic. One of the challenges in that region is the sparseness of the data and its quality - not only from a signal content point of view, but also from uncertainties in the dating. Even layer counted data has uncertainties in the dating (Comboul et al. 2014), although it can sometimes be corrected for through cross-dating. We show how current methods deal with dating uncertainties, and how the more flexible (but sadly more demanding) Bayesian inference can help correctly propagating the uncertainties from the record to the final reconstruction product (Werner and Tingley, in prep.).

Comboul, M et al. (2014), Clim. Past. 10, 825–841, doi:10.5194/cp-10-825-2014
Werner, JP and MP Tingley (in prep)
Wetter, O et al. (2014), Clim. Change 125, 349-363, doi:10.1007/s10584-014-1184-2

Statistical analysis of heavy rainfall in France via multivariate extreme value theory

G. Marcon, S. A. Padoan, P. Naveau and P. Muliere

Analysing heavy rainfall in France is complex due to the high number of weather stations and the complexity of weather system patterns over the French territory. This leads to computational issues and classical Extreme Value Theory (ETV) cannot be directly applied. To bypass the computational hurdles, we perform a dimension reduction approach, based on EVT concepts, to create independent regional clusters. Within each cluster, we propose a nonparametric approach for estimating the maxima dependence function.

Satellite fire monitoring based on low cost

Rosa Lasaponara, Fortunato Desantis, Angelo Aromando, Antonio Lanorte

Wildfire represents one of the main disturbance factors throughout the world bringing profound transformations which affect ecosystems, landscapes and environments. Fires cause fuel consumption, production of smoke and ash, soil heating, alteration in vegetation structure and composition, soil erosion, changes in nutrient, micro-climate, hydrology, vegetation succession, habitat.
Remote sensing technologies can provide useful data to contrast fire phenomena, from risk estimation (including fuel mapping and fire spread simulation), fire detection to post fire monitoring and management (burn area mapping, fire severity, vegetation recovery, etc).
This paper presents an overview of satellite remote sensing of forest fire as well as the successful results obtained in the context of the FIRE-SAT project, focused on the operative application of the satellite based tools for fire monitoring. FIRE-SAT has been funded by the Civil Protection of the Basilicata Region in order to set up a low cost methodology for fire danger monitoring and fire effect estimation based on satellite Earth Observation techniques. To this aim, data available free of charge such as, Moderate Resolution Imaging Spectroradiometer (MODIS), ASTER, Landsat TM were used. Novel data processing techniques have been developed by researchers of the ARGON Laboratory of the CNR-IMAA for the operative monitoring of fire. FIRE-SAT system is made up of the following modules
(i) danger estimation to support and optimize fire fighting strategies from the alert to the management of resources including fire attacks.
(ii) mapping of burned areas and fire severity to assess the damage on vegetation and soil and support post fire management
(iii) estimation of landslide susceptibility after fire (conducted using both satellite time series and in situ analysis) to support the mitigation actions
(iv) assessment of post fire recovery (using satellite time series) to evaluate fire resilience and assess the post fire vegetation recovery capability.
The operative use of the integrated models we developed and tested from 2008 to 2014, pointed out that satellite data sets available free of charge as (MODIS, Landsat TM, ASTER) are very suitable for an effective, systematic (daily) and sustainable low-cost monitoring of large areas. FIRE-SAT system and tools we set up are capable to properly monitor spatial and temporal variations of fire susceptibility and provide useful information of both fire severity and post fire regeneration capability.

Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions

Alvaro Corral

Power-law distributions contain precious information about a large variety of physical processes. Although there are sound theoretical grounds for these distributions, the empirical evidence giving support to power laws has been traditionally weak. Recently, Clauset et al. have proposed a systematic method to find over which range (if any) a certain distribution behaves as a power law. However, their method fails to recognize true (simulated) power-law tails in some instances, rejecting the power-law hypothesis.

Moreover, the method does not perform well when it is extended to power-law distributions with an upper truncation.

We present an alternative procedure, valid for truncated as well as for non-truncated power-law distributions, based in maximum likelihood estimation, the Kolmogorov-Smirnov goodness-of-fit test, and Monte Carlo simulations. We will test the performance of our method on several empirical data which were previously analyzed with less systematic approaches. The databases include the half-lives of the radionuclides, the seismic moment of earthquakes worldwide and in Southern California, and a proxy for the energy dissipated by tropical cyclones. We find the functioning of the method very satisfactory.

A. Corral, F. Font, and J. Camacho, Noncharacteristic half-lives in radioactive decay, Phys. Rev. E 83, 066103 (2011).

A. Deluca and A. Corral, Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions, Acta Geophys. 61(6), 1351-1394 (2013).

Contributed talks

Monday, 6 October 2014

Data-driven Climate Modeling and Prediction

Dmitri Kondrashov, Mickael Chekroun, Honghu Liu and Michael Ghil

Comprehensive dynamical climate models aim to simulate past, present and future climate; more recently, they also attempt to predict climate on longer and longer time scales. These models, commonly known as general circulation models or global climate models (GCMs) represent a broad range of time and space scales and use a state vector that has many millions of degrees of freedom. Considerable work, both theoretical and data-based, has shown that much of the observed climate variability can be represented with a substantially smaller number of degrees of freedom. While detailed weather prediction out to a few days requires high numerical resolution, it is fairly clear that a major fraction of climate variance can be predicted in a much lower-dimensional phase space. Low-dimensional models (LDMs) can simulate and predict this fraction of variability, provided they are able to account for (i) linear and nonlinear interactions between the resolved high-variance climate components; and (ii) the interactions between the small number resolved components and the daunting number of unresolved ones. LDMs hold great promise to provide predictive understanding of natural hazards from the consequences of global warming and climate change.

We will present several applications by data-driven LDM approach for climate modeling and prediction, namely multi-level empirical model reduction (EMR), and compare it with existing methods. EMR aims to obtain a low-order nonlinear system of prognostic equations driven by stochastic forcing, and estimates both the dynamical operator and the properties

of the driving noise directly from observations or from a high-order model's simulation. We will also compare several methods to obtain dynamic variables from spatio-temporal datasets for the purpose of constructing data-driven LDM models, such as Principal Component Analysis (PCA) and Multichannel Singular-Spectrum Analysis (MSSA).

Potential Predictability of Decadal Climate Variability Using Relative Entropy

Marianne Uhlig, Hendrik Feldmann

The German research program MiKlip aims at the development of a decadal predictions system that could be used operationally by weather services. A module within MiKlip is dedicated to develop a regional downscaling system for the global predictions. One of the regional focuses of the downscaling experiments is Europe where global decadal predictions have shown skill.
Hindcast experiments for the period of 1961 to 2010 were performed with the Earth System Model MPI-ESM of the Max-Planck-Institute for Meteorology. For the regional downscaling over Europe the regional climate models (RCM) COSMO-CLM and REMO are applied to establish a regional ensemble for the CORDEX-EU domain.
A goal is to assess the prediction potential of several variables as well as the benefit for users. Decadal predictability in Europe originates from low frequency climate variations, especially in the North Atlantic. One possibility to quantify a variable's predictive potential could be to compare the distribution of a decadal forecast to a climatological one. How much and for how long the predicted probability density function (PDF) can be distinguished from the corresponding climatological PDF will then give a measure of predictability.
We use the relative entropy from information theory to measure the difference between the PDF's. The advantages of the relative entropy are that it is e.g. a non-symmetric measure, it is non-parametric and also able to account for non-linearities and is further simple to implement.

Assessing the impact of climate variability on crop yields: implications for predictability on seasonal and decadal time scales

A. Ceglar, A. Toreti and F. Dentener

The impact of climate variability on crop yield (grain maize and winter wheat), under current climate conditions and over several European regions, is here investigated. Crop yield time series, obtained from national statistical offices and covering the last 25 years, are analysed together with monthly cumulated precipitation, global solar radiation and mean temperature. Since crop yields are strongly influenced by changes and improvements in agro-management practices, a de-trending procedure is run. Then, a canonical powered partial least square regression is applied to identify climate variables with the highest impact on crop yield inter-annual variability. Those variables and the timing of impact during the growing season are used in a subsequent analysis aiming at characterising the large-scale atmospheric regimes presenting the largest risk for crop loss.
Results reveal remarkable spatial differences and highlight the importance of weather conditions during flowering for both grain maize and winter wheat. Finally, these findings

are discussed in the frame of seasonal crop yield forecasting and assessment of climate change impacts on decadal time scales.

A goodness-of-fit test for heavy-tailed distributions and its application to precipitation extremes

A. Toreti and G. Jogesh Babu

Extreme events, such as heavy precipitation, have a strong impact on both natural and human systems. Thus, it is essential to achieve a better characterisation and understanding of those phenomena. The natural statistical framework for the analysis of extreme values is the Extreme Value Theory. Since many issues in the available inference procedures still remain and reliability represents a key factor, a specific goodness-of-fit test for heavy-tailed distributions is here proposed. The procedure is proved to be valid and it is applied to daily global precipitation extremes simulated by a set of climate models of the Coupled Model Intercomparison Project Phase 5.

A statistical approach to the study of the shallow aquifer in the Piedmont region (NW Italy)

Barbero D., De Luca D. A., Forno M. G., Lasagna M., Magnea L.

We report on a statistical approach to the study of subsoil thermal data in order to establish if the Piedmont plain (NW Italy) is promising for geothermal purposes. The model we developed can be used during the preliminary stage of investigation for the exploitation of subsoil by means of low enthalpy geothermal plants.
As an example of our method, we performed a statistical analysis using groundwater thermometric data collected during the spring and autumn of 2009 in monitoring piezometers located in the Quaternary fluvial deposits hosting a shallow aquifer.
This study allowed us to determine the location of the homoeothermic surface, its depth range and the average value of the temperature with its associated uncertainty. This average value corresponds to the asymptotic behavior of the damped oscillations connected to seasonal temperature fluctuations found in the subsurface. The results we obtained are also in agreement with the solutions of the heat conduction equation with stationary boundary conditions.
Our method allows us to determine the statistical probability of finding a given groundwater temperature near the homeothermic surface, with a chosen confidence interval. This can be useful for example in planning to install “open loop” heat pump system. This methodology allows to significantly limit direct measurements, which would need to be widespread and repeated, and therefore expensive.

Influence of synoptic meteorological patterns on solar irradiance in northern Europe

Kajsa Parding, Beate Liepert, Laura Hinkelman, Thomas Ackerman, Jan Asle Olseth, Knut-Frode Dagestad

Solar radiation is the main external source of energy on Earth. It influences ecosystems directly via photosynthesis and indirectly by driving the hydrological cycle, atmospheric circulation and ocean circulation. Previous studies have reported decadal trends in observational records of surface shortwave (SW) irradiance around the world, too strong to be caused by varying solar output. These observed decadal trends have been dubbed "solar dimming and brightening" and are believed to be related to changes in atmospheric aerosols and cloud cover. Because the observed solar variability coincides with qualitative air pollution histories, the dimming and brightening have become almost synonymous with shortwave attenuation by anthropogenic aerosols. However, there are indications that atmospheric circulation patterns have influenced the dimming and brightening in some regions, e. g., Alaska and Scandinavia. In this work, we focus on the role of atmospheric circulation patterns in modifying shortwave irradiance in northern Europe.

To establish a connection between weather patterns and sunshine, empirical models of SW irradiance are fitted using a daily classification of European weather called Grosswetterlagen (GWL).
The GWL data set has the advantage of providing a more detailed and versatile account of atmospheric circulation compared to climate indices. We find that the observed decadal variations of SW irradiance in northern Europe are linked to changes in the frequency of anti-cyclonic and cyclonic weather patterns centred over Scandinavia and the North Atlantic. The observed changes in circulation patterns may be related to a northward shift in the North Atlantic storm track. However, there are also indications of a other factors such as varying aerosol emissions or small scale meteorological phenomena, playing an important role in the brightening of recent decades.

Modeling of typhoon trajectory patterns using the Gaussian process regression

S. Nakano, K. Suzuki, and G. Ueno

It is not ensured that the activity of typhoons is unchanged in the future, but it can be changed due to some factor like climage change. In order to assess the risk of tropical cyclone hazards in the future, it is important to consider the long term change of the activities of tropical cyclones. If we want to evaluate the risk of tropical cyclone hazards in a particular region, it is important to consider the long term change of typical trajectories of tropical cyclones. In this paper, the long term change of typical trajectories of typhoons is examined. In order to obtain the features of the typhoon trajectories, we employ a Gaussian process regression technique to obtain a model of typical typhoon velocity as a function of latitude, longitude, day of year, and Julian day. The characteristics of the long term change of the typical typhoon velocity are discussed on the basis of the result of the Gaussian process regression analysis.

Tuesday, 7 October 2014

On the internal variability of simulated precipitation

Anne Schindler, Andrea Toreti, Enrico Scoccimarro, Matteo Zampieri

Adequate knowledge of expected changes in precipitation is needed for planning as well as for mitigating and adapting to climate change. Potential changes in precipitation can be assessed by using climate model simulations under different scenarios. Yet the climate change signal can be obfuscated by natural variability. Here, we propose an approach that can be used in the attribution of the identified changes. By using long pre-industrial control simulations (in this exercise, a 330-year run of the CMCC Global Circulation Model) we can determine the shortest record length such that randomly chosen time periods of the same length cannot be statistically distinguished from each other. In the context of climate change assessment, this implies that any change simulated over a period of this length could be attributed to a change in forcing and not to natural variability.
For each land grid point, the empirical distribution over a given time period is compared with that of 99 (randomly chosen) periods of the same length. Should the Cramer-von Mises two sample test be unable to reject the hypothesis that the samples stem from the same distribution, then the periods are deemed statistically indistinguishable. Multiple testing is accounted for with the Walker test at the 5 % level. To also test for regional significance, we calculate the Benjamini-Hochberg false-discovery-rate for the 26 IPCC SREX regions.
We show that this minimum duration length depends on the season, the location and the extent of the region under investigation.

Northern hemisphere ocean-atmosphere coupling from an interacting climate network perspective

M. Wiedermann, J.F. Donges, R.V. Donner, D. Handorf, J. Kurths

In recent years extensive studies on the Earth's climate system have been carried out by means of advanced complex network statistics. However, most previous studies were constrained two conceptual restrictions: First, in most cases network measures have been computed without taking into account the topology of the discrete grid, regular or irregular, that climate data is typically stored on. Second, the great majority of recent studies have been focusing on single climatological fields located on surfaces parallel to or directly on the Earth's surface.

To address both issues, we introduce a novel approach for analyzing and quantifying interactions in the climate system by means of node-weighted interacting network measures. Two, so far disjoint, approaches, one for treating inhomogeneous and discrete node sampling and one for investigating interacting networks are combined and utilized for studying ocean-atmosphere coupling in the northern hemisphere. Specifically, we construct 18 coupled climate networks based on monthly data from the ERA 40 reanalysis, each consisting of two subnetworks. In all cases, one subnetwork represents sea-surface temperature (SST) anomalies while the other is based on the geopotential height (HGT) of

isobaric surfaces at different pressure levels up to the lower stratosphere. Our analysis reveals which isobaric layers show strong coupling with the dynamics of the oceans and where. By an exploratory investigation of the resulting interdependent network structure's connectivity, we identify well-known climatological phenomena such as the eddy driven jet stream in the Northern Atlantic and signatures of the Hadley circulation, especially in the northern Pacific. The analysis is performed separately for summer and winter months to identify key differences in the atmospheric dynamics. A strong coupling between the SST and HGT fields in the upper troposphere is detected during winter months and the corresponding local network measures reveal its spatial extent displaying well localized areas where the interaction between the two subsystems is strongest. For summer months the dynamical decoupling of the upper troposphere and lower stratosphere is well observed and the eddy driven jets are proven to be weaker compared to the winter months. Furthermore, our analysis reveals dynamic signatures which cannot be simply explained by the basic large-scale cellular structure of atmospheric dynamics, but reflect specific mechanisms of energy and heat transfer between the ocean and different atmospheric layers, which need to be further disentangled in future work.

Our method provides a general, unit-free, tool to investigate the interaction between two coupled network-like subsystems whose nodes carry different importance or weight. The method has also been successfully applied to investigate features of international trade and proves to be applicable to various tasks of data analysis beyond the climatological field of study.

Weather radars applications: innovative telecommunication systems and rainfall field reconstructions using a rain gauges network

D. Guenzi, F. Acquaotta, S. Barbero, R. Bechini, R. Boraso, V. Campana, R. Cremonini, S. Fratianni

Weather radar systems transmit raw data over the network, with 24h / 24h continuity and in any weather condition. They are often located in digital divide areas and, for this reason, radar measurements are sometimes processed partially on site and reduced in size to be adapted to the limited bandwidth available in those zones. This research project aims at studying, designing and implementing a data transferring and managing system, suitable to collect, transfer and process data from the weather radar, in order to better exploit the information addressed to the Functional Center. The main goal is to explore a methodology as much as possible standard and replicable, able both to convey the amount of raw information by means of local broadband connections and to process data through modern computing systems based on innovative technologies for the elaboration, record and analysis of the information.
For example, due to uncertainties affecting weather radar rainfall estimations, a common technique to reconstruct the rainfall field is to compare weather radar observations with ground measurements given by rain gauges. The metropolitan area of Turin, in Italy, offers the unique condition of a C-band polarimetric Doppler weather radar which is just 6 km far away from the downtown, in addition to a dense network made by tens of tipping-bucket rain gauges. In order to reduce biases and errors between ground measurements and weather radar rainfall estimations, this study shows an application of a reflectivity / rainfall

rate (Z-R) calibration technique applied on the native weather radar grid, taking the advantage of higher resolution obtained near the antenna.

HPC environment for massive earth data analysis: field examples at regional and continental scale

Rosaria Tondi, Alessandra Borghi, Carlo Cavazzoni

To obtain accurate and reliable estimations of the major lithological properties of the rock within a studied volume, geophysics uses the joint information provided by different geophysical datasets (e.g. gravimetric, magnetic, seismic). Representation of the different types of information entering the problem using probability density functions can provide the mathematical framework to formulate their combination. The maximum likelihood estimator of the resulting joint posterior probability density functions leads to the solution of the problem. However, one key problem appears to limit the use of this solver to an extensive range of real applications: information coming from potential fields that implies the presence of dense matrices in the resolving estimator.
Dense matrix systems rapidly challenge both the algorithms and the computing platforms, and are not suited to high-resolution 3D geophysical analysis. In this study, we propose a procedure that allows us to obtain fast and reliable solutions of the joint posterior probability density functions in the presence of large gravity datasets and using sophisticated model parametrization.
As it is particularly CPU consuming, this 3D problem makes use of the HPC environment to improve the performance and the accuracy of the simulations.
As field experiments we present the recovered 3D modeling of densities and seismic velocities beneath the Po plain (44°N - 46°N) (7.5°E - 12.5°E), the European continent (20°N - 90°N) (40°W - 70°E) and the Pacific Plate (90°N - 90°S) (121°E - 60°W).

Application of Random Forest Algorithm for Environmental Data

Michael Leuenberger and Mikhail Kanevski

Due to the large amount and complexity of data available nowadays in environmental sciences, we face to the need to apply robust and nonlinear algorithms allowing analyses and understanding of the phenomena under study. In the present research we investigate the potential of the machine learning algorithm - Random Forest (RF), for environmental data analysis and modelling. Recently this algorithm has gained a great popularity in a variety of data mining applications.

Developed by L. Breiman and A. Cutler, the Random Forest algorithm provides an ensemble of classification and regression trees. By a pseudo-random variable selection for each split node, the algorithm grows a variety of decision trees which return different results. A committee system votes (or averages) these results and assigns the predicted values to the unlabeled data within the validity domain. Furthermore, the algorithm provides the measure of the contribution of each variable. This measure can be used to detect and display the main factors affecting the phenomenon under study. Thus, RF algorithms have important

characteristics for natural hazards assessments: high quality predictions, ranking of independent variables according to their importance, estimation of prediction uncertainties and is computationally efficient.

The method was applied to different case studies, such as landslides, forest fires and permafrost. Results of the application of Random Forest algorithm show good estimations of the relevance of different environmental variables for predicting specific natural phenomenon. Finally, susceptibility maps were elaborated based on the selected variables.

A comprehensive gridded global soil dataset

Wei Shangguan, Yongjiu Dai, Qingyun Duan, Baoyuan Liu, and Hua Yuan

We developed a comprehensive, gridded Global Soil Dataset for use in Earth System Models (GSDE) and other applications as well. GSDE provides soil information including soil particle-size distribution, organic carbon, and nutrients, etc. and quality control information in terms of confidence level. GSDE is based on the Soil Map of the World and various regional and national soil databases, including soil attribute data and soil maps. We used a standardized data structure and data processing procedures to harmonize the data collected from various sources. We then used a soil type linkage method (i.e. taxotransfer rules) and the polygon linkage method to derive the spatial distribution of soil properties. To aggregate the attributes of different compositions of a mapping unit, we used three mapping approaches: area-weighting method, the dominant soil type method and the dominant binned soil attribute method. In the released gridded dataset, we used the area-weighting method as it will meet the demands of most applications. The dataset can be also aggregate to a lower resolution. The resolution is 30 arc-seconds (about 1 km at the equator). The vertical variation of soil property was captured by eight layers to the depth of 2.3 m.

Global temperature change: analysis of warming rates of Earth's Climate

Bernard De Saedeleer

Global temperature change is very often shown in the literature as a graph showing the temperature rise as a function of time, which contains lots of variability for many physical reasons. On the other hand, plenty of warming rates values derived somehow from that graph are mentioned, but which are often restricted to a few given periods of time, lacking a global view. The starting year chosen for the study is often rightly argued to play an important role for concluding to the sign of the rate (warming or cooling), but without explicit demonstration of this.

In this study, we perform a systematic analysis based on the GLOBAL Land−Ocean Temperature Index (LOTI) data provided by NASA GISS, by computing the warming rates using a linear regression between the middles of stepped lines (averages) of given duration. The full picture of the rates is given, whatever the duration of the averaging period is (from 5 to 30 years), and whatever the staring year is. All staircases are mainly rising.

In addition, it may be important to have the instantaneous warming rate at disposal, in order to achieve insightful conclusion on mechanisms in physical models. To extract that climatic signal, attempts to remove higher mode of variability than the 30-year duration defining Climate have been proposed, mainly by using moving averages over several years. But these do not properly remove the high frequency variations. So, the non-derivability of such curve is the stumbling block preventing from computing rigorously by derivation the instantaneous warming rate. Indeed, the wiggles would produce very noisy rates, switching yearly from positive to negative values, which is not physically sound for Climate studies, implying 30-years timescales.

We provide here decanal and centennial smooth curve fitted on the GISS data which have the advantage of being derivable. Analytic expressions are given, from which instantaneous warming rates can now be deduced by simple analytic derivation, which is then very clean as the high frequency noise has been removed.

By these two aspects of the study, this contribution aims at providing a comprehensive approach for deriving global warming rates. This rigorous computational analysis has the advantage to avoid any bias due to arbitrary choice of parameters, or due to visual distortion. The conclusion is that Earth's Climate is undoubtedly warming, but with a comprehensive view on the details on how strong this warming happen at several timescales: year, decade, century.

Effects of the 7-8-year cycle in daily mean air temperature (from European stations)

Nikola Jajcay, Jaroslav Hlinka, Milan Paluš

Recently, an information transfer from larger to smaller scales of the air temperature variability has been observed in daily mean surface air temperature (SAT) data as the influence of the phase of slow oscillatory phenomena with periods around 6-11 years on amplitudes of the variability characterized by smaller temporal scales from a few months to 4-5 years [1]. The strongest effect is exerted by an oscillatory mode with the period close to 8 years and its influence can be seen in 1-2 °C differences of the conditional SAT means taken conditionally on the phase of the 8-year cycle.
The size of this effect, however, changes in space and time. The changes in time are studied using sliding window technique, showing that the effect evolves in time, and during the last decades the effect is stronger and significant. Sliding window technique was used along with seasonal division of the data, and it has been found that the cycle is most pronounced in the winter season. Different types of surrogate data are applied in order to establish statistical significance and distinguish the effect of the 7-8-yr cycle from climate variability on shorter time scales.

[1] M. Palus, Phys. Rev. Lett. 112 078702 (2014)

Ice wedge thermal variation in East Antarctica: a time series approach

Raffi, R., Bramati, M.C., Baldassarre, A.

This research aims at studying the thermal variation of ice wedges at various depths. In particular, the analysis of the air, ground surface (GST), ice-wedge top and bottom temperatures are undertaken. The active layer depth is calculated through seasons and years using hourly data at three sites in northern Victoria Land: Baker Rocks, Boomerang Glacier and Mount Jackman. The recording period is from 2004 to 2013 at Baker Rocks and Boomerang Glacier, and from 2006 to 2013 at Mount Jackman.
Daily mean ground surface temperatures (DMGST) and daily mean air temperatures (DMAT) are highly correlated at Baker Rocks (r2=0.96), at Boomerang Glacier (r2 =0.95), and at Mount Jackman (r2=0.92) sites. This shows that the GST at measurement sites responds strongly to air temperature. Moreover, hourly ground surface temperature and DMGST are generally lower than the air temperature in the winter season, which shows the absence of a significant snow cover.
Standard deviations of the hourly temperature show that high temperature variability can exist over one month, with higher variability in winter than in summer. Frequent and large temperature fluctuations are common throughout winter with either a sharp drop or a rapid increase both in air and ground surface temperature. Variations of 25°C to 30°C were recorded over periods of one to four days.
The overall variability of temperatures is decomposed using spectral analysis in order to isolate seasonal effects from cycles and long term trends. The time series approach in the frequency domain is quite new in this field and it represents therefore the main contribution to the existing literature.

Wednesday, 8 October 2014

Analysis of macroseismic fields using statistical data depth functions

Agostinelli C., Rotondi R.

Modelling seismic attenuation is one of the most critical points in the seismic hazard assessment process. We consider the spatial distribution of the effects caused by an earthquake as expressed by the values of the macroseismic intensity recorded at various locations surrounding the epicentre. Considering the ordinal nature of the intensity, a way to show its decay with the distance is to draw curves - isoseismal lines - on the map bounding points of intensity not smaller than a fixed value. These lines are usually closed and nested curves around the epicentre with highly different shape because of the effect of ground conditions and of complexities in the rupture propagation. To forecast the damage scenario of a future earthquake we need to discover a common pattern on the basis of past events and to define a stochastic model which takes into account the various sources of uncertainty inherent in the phenomenon. An initial exploratory analysis can be performed through nonparametric statistical tools for functional data. Statistical data depth functions prove successful in the analysis of multivariate data sets, in particular deriving an overall centre and assigning ranks to the observed units. Recently, this method has been extended to

ordering of functions and trajectories. In particular, López-Pintado and Romo (2011) have proposed the half-region depth suited for functional data and for high dimensional data; its local version has been studied by Agostinelli (2014). In this work we propose the use of the local half-region depth to identify the attenuation trend of a set of macroseismic fields.
This research has been performed in the framework of the European project "Urban Prevention Strategies using Macroseismic and Fault sources" (UPStrat-MAFA).

Agostinelli (2014) Local Half Region for Functional Data, submitted
López-Pintado S. and Romo J. (2011) A half-region depth for functional data. Computational Statistics & Data Analysis, 55(4):1679-1695.

Spatial organization of foreshocks as a tool to forecast large earthquakes

E. Lippiello, W. Marzocchi, L. de Arcangelis, C. Godano

An increase in the number of smaller magnitude events, retrospectively named foreshocks, is often observed before large earthquakes. We show that the linear density probability of earthquakes occurring before and after small or intermediate mainshocks displays a symmetrical behavior, indicating that the size of the area fractured during the mainshock is encoded in the foreshock spatial organization. This observation can be used to discriminate spatial clustering due to foreshocks from the one induced by aftershocks and is implemented in an alarm-based model to forecast m>6 earthquakes. A retrospective study of the last 19 years Southern California catalog shows that the daily occurrence probability presents isolated peaks closely located in time and space to the epicenters of five of the six m>6 earthquakes. We find daily probabilities as high as 25% (in cells of size 0.04 3 0.04deg2), with significant probability gains with respect to standard models

Unsupervised Learning of Spatial Patterns of Local Seismicity

Mikhail Kanevski

The spatial distribution of earthquake epicentres demonstrates very complex patterns. There are many measures used to characterize either temporal (fractals, Fano and Allan factors, etc.) or spatial (fractals, correlation dimensions, Morisita index, etc.) clustering. In fact all these measures are of global character and some of them only give a characteristic size of the cluster (temporal or spatial). In this study we apply and analyse a local measure of clustering based on a sandbox counting (mass counting) and closely related to the correlation dimension. In local estimation of clustering each measurement point K (there are N points in total) is considered as a centre of clustering. By changing the radius R (R1,R2,…Rm) around the point K a number of events falling in the circle of the radius Ri is counting (m-is a number of different radii considered). Finally, N curves (objects) of m-dimensions are obtained. A global fractal dimension is estimated by computing an average slope of these curves at small distances. In this research a nonlinear unsupervised learning algorithm – self-organizing Kohonen maps are applied to find clusters in the m-dimensional space. After clustering different number of classes recognised by using k-means algorithm is considered and analysed. The preliminary results are quite promising and reveal the patterns

of spatial clustering in Swiss seismicity data. The extension of the approach, which is in progress, deals with marked and bivariate spatial point processes.

Different facets of dynamical complexity in the magnetosphere – A recurrence perspective

Reik Donner, Veronika Stolbova, Jonathan Donges, George Balasis, Marina Georgiou, Stelios Potirakis, Jürgen Kurths

Magnetic storms are among those natural phenomena that can have a global hazardous potential and seriously affect the proper functioning of infrastructures such as communication and electrical power generation and supply. In order to better understand the dynamical mechanisms underlying the emergence of such storms, the complex structure of fluctuations of magnetic field strength and its relationships with solar activity needs to be carefully analyzed. In this work, recurrence analysis is applied to one year of observations (2001) of the disturbance storm-time (Dst) index including several quiescence and activity phases. Our results reveal that recurrence characteristics provide excellent tracers for changes in the dynamical complexity along non-stationary records of geomagnetic activity. In particular, trapping time (characterizing the typical length of ``laminar phase'' in the observed dynamics) and recurrence network transitivity (associated with some generalized notion of effective attractor dimension) allow very good discrimination between activity (magnetic storms) and quiescence phases. Specifically, we find that the discriminatory skills are superior to many previously considered characteristics like Hurst exponent or symbolic dynamics based entropy concepts. Our results point to great potentials of recurrence-based measures for unveiling temporal changes in the dynamical complexity of the magnetosphere.

Spatial-temporal power law distribution of wildfires in Europe: a comparison between observed and simulated datasets

Di Mauro B.; Fava. F.; Frattini P.; Camia A.; Colombo R.; Migliavacca M.

Wildfire is a complex natural hazard that dissipates energy accumulated by forest and grassland through the cycles of growth and senescence. The occurrence of a wildfire across different landscapes is strictly related with both natural (e.g. temperature, precipitation) and anthropogenic (e.g. population density, fire policies) influences. The modeling of wildfires can be divided in two categories: the fire ecology and the statistical physics approach. The former aims to reproduce the interactions between the physical status of vegetation and forcing agents that leads to wildfires, the latter exploits emergent properties (i.e. power law scaling) of wildfire spatial-temporal distribution and models it as cellular automata.
In this research, we model wildfires monthly burned areas with a power law distribution across European ecoregions (Metzger et al. 2005) using both Ordinary Least Square (OLS) regression and Maximum Likelihood Estimation (MLE) (Clauset et al. 2009). The analysis is applied to observed (European Forest Fire Information System, EFFIS) and simulated (with Community Land Model, CLM) catalogs, in order to benchmark the process based land surface model that showed poor performances in the description of the interannual variability and severe fires season. Power law exponents have been recently interpreted as a

proxy of fire regimes in North America (Malamud et al. 2005); we produced European-scale maps of exponents and interpreted emerging patterns in terms of fire regimes and extreme events.

Frequency and Continuous Time-Frequency Analysis for Some Unevenly Sampled Stochastic Processes and Implications for Significance Testing of Time Series

Lenoir, Guillaume and Crucifix, Michel

Spectral analysis is a common approach for the interpretation of climate time series. In particular, the discrete Fourier transform, the Gabor transform and the continuous wavelet transform are widely used. The discretization procedure is traditionally done on a regular grid, however, numerous geoscientific data are only available with unevenly spaced time steps. Although a pragmatic solution may be to interpolate the data on a regular grid, this process may significantly affect the analysis. An alternative approach is to use the Lomb-Scargle periodogram to estimate the Fourier spectrum, but it has two main limitations:
- The spectrum of classical stochastic processes needs to be estimated by Monte-Carlo simulations, which may be time expensive.
- There is no rigorous theoretical equivalent for the continuous time-frequency transforms.
The objective of this contribution is to generalize the current framework for the analysis of unevenly sampled time series. Specifically, we extend the the Lomb-Scargle periodogram to some continuous time-frequency transforms. Then, we estimate the analytical spectrum of some stochastic processes under those transforms. The latter offers a significant gain on the computing cost associated with Monte-Carlo simulations.

Based on the previous results, we propose a rigorous framework to estimate the significance of Gabor or wavelet spectrum given unevenly sampled data. We proceed as follows:
- Define the model: Significance testing is always based on a model that needs to be properly defined.
- Define the parametric hypotheses H0 (null case) and H1 (alternative case), and attempt to find an appropriate statistical summary.
- Consider the correlations between neighbouring coefficients in the continuous transform.
- Reject some parts of the transform due to the Shannon-Nyquist Sampling Theorem.
Some examples of paleoclimate time series analysis are given. Finally, we investigate whether considering the interpolated data to get a constant time step and using the traditional tools may lead to significant errors.

Improving the accuracy of river flow forecasts

Mario Lefebvre

For short-term forecasts of river flows, various models based on one or two-dimensional stochastic processes have been considered. In order to improve their accuracy, precipitation forecasts should be incorporated into these models. Here, linear regression is used to forecast river flows, with and without the precipitation data. The model is applied to important rivers located in the United States.

Quantitative Modelling of Turbidity Current Flow and Sedimentation Processes, offshore Agadir basin, North West African Margin

Lucky Imagbe and Dave Waltham

A turbidity current suspension model is being developed to investigate flow and sedimentation processes in deep water basins.
The model seeks to constrain the most-recent field investigation in the Agadir basin, offshore Morocco and to determine the possibility of thin flows (5.0m-7.0m) being able to transport medium to coarse-grained sands of about 300 microns across the 35,000 km2 Agadir basin, without evidence of early deposition or sea floor erosion.
For simplicity, the quantitative suspension model being proposed is considered under three key sub-components including relating mean flow velocity to turbulence; relating turbulence to suspension and relating flow thickness to flow velocity.
The first component of the model which investigates the relationship between the flow velocity, and turbulence has been validated against three published flume experimental data, and the results show a reasonably close agreement with the calculated modelling results and therefore, could be further applied to simulate the field observations of Agadir basin as well as other analog deepwater turbidite systems.

Bayesian principal curve analysis to detect seismic faults

Raffaele Argiento and Alessandra Guglielmi

We consider the problem of detecting seismic faults from earthquake catalogs. From a modeling point of view we face this problem as a clustering of spatial point data with "curved" support. We follow a Bayesian nonparametric approach by considering a species sampling mixture model. Our first goal is to define a general/flexible class of distributions, such that can model data from clusters with non-standard shape. To this end, we extend the definition of principal curve given in (Tibshirani 1992) into a Bayesian framework. We propose a new hierarchical model, where the data in each cluster is parametrically distributed around the Bayesian principal curve, and the prior cluster assignment is given on the latent variables at the second level of the hierarchy according to a species sampling model. We illustrate the performance of our model on data coming from Italian CPTI catalogue.

ETAS model estimation with predictive measure

Giada Adelfio, Marcello Chiodi

A new method to estimate the space-time intensity of a ETAS-type process for seismic data is proposed, assuming the coexistence of both a nonparametric and parametric component. The method alternates a forward predictive likelihood (FLP) estimation for nonparametric background seismicity and maximum likelihood for parametric components of triggered seismicity, until convergence is reached. Examples of application are here proposed.

The Gompertz waiting time distributions of the stress release model in the Bayesian framework

Elisa Varini, Renato Rotondi

We consider the stress release model (SRM), a self-correcting marked point process popular in seismology and actuarial science. By assuming this model and conditional on the history of the process up to time t, we show that the waiting time of the next event from t follows a Gompertz distribution with a time-dependent shape parameter. Therefore simulation of point processes to characterize the waiting time distribution is no longer necessary. Some features of the model are also revealed: (1) Following a major event, the waiting time of the next event can be short with high probability; (2) A long period of time with no events reduces the uncertainty in the waiting time until the next event.
We illustrate an application of the SRM to a seismic sequence of 16 earthquakes occurred in Southern Italy from 1667 to 2002, with moment magnitude of at least 5.3. We fit the model to data by following the Bayesian approach because, despite the small data set, it still provides tools to measure the uncertainty of both parameters (posterior distribution) and forecast quantities (posterior predictive distribution). The poor prior knowledge in this application has led us to consider data-dependent vague proper prior distributions obtained through the Empirical Bayes method and by spanning the range of the marginal likelihood. By exploiting the Gompertz argument for the waiting time distributions, we perform both the retrospective and prospective forecasts of earthquake events for model validation and prediction of the next earthquake in the region, respectively.

Posters

Building a Vis-NIR spectral database for rapid determination organic carbon and nitrogen in forest soils of the southern Calabria (Italy)

Massimo Conforti, Raffaele Froio, Giorgio Matteucci, Gabriele Buttafuoco

Forest soils play an important role in the carbon stock; in particular, the evaluation of soil organic carbon (SOC) content is an important step for carbon sequestration studies. The project LIFE09 ENV/IT/078 Managing forests for multiple purposes: carbon, biodiversity and socio-economic wellbeing (ManFor C.BD.), among other objectives, is aimed to evaluate carbon sequestration taking into account forest management for conserving and enhancing carbon stocks, and increase carbon sequestration.
Conventional laboratory analyses for the determination of soil properties as organic carbon (OC) and nitrogen (N) are expensive and time-consuming. Visible-near infrared (Vis-NIR) spectroscopy in combination with chemometrics techniques is claimed to be a rapid, cost-effective and non-destructive method for measuring soil properties.
The objective of this study was to build Vis-NIR soil spectra database and develop prediction models for OC and N in a representative forest area of the Biogenetic Nature Reserve “Marchesale” located in Calabria region (south Italy).

A set of 265 soil samples were collected within the study area. Soil samples were air dried, sieved at 2 mm and analyzed to estimate OC and N content. Subsequently, the Vis-NIR reflectance of each soil sample was measured in laboratory, using an ASD FieldSpec IV 350-2500 nm spectroradiometer.
In order to develop models based on soil spectra and reference laboratory data of OC and N, partial least squares regression (PLSR) was used.
To evaluate the accuracy of the PLSR models, the dataset was randomly separated into two subsets: calibration set (70%, n=180) for developing the prediction model and validation set (30%, n=80) to test the models accuracy. Several calibration models were built and compared by cross-validation. The predictive ability of the cross-validation models was evaluated by the coefficient of determination (R2) and the root mean square error (RMSE) of calibration.
Results revealed a high level of agreement between measured and predicted values with high R2 and low RMSE values. The best calibration model obtained for OC show a R2 of 0.91 and a RMSE of 0.56% while for N, the R2 was equal to 0.81 and the RMSE to 0.50%.
Good results of validation were obtained for both OC (R2=0.87 and RMSE=0.76%) and N (R2=0.79 and RMSE=0.76 %).
The results indicate that Vis-NIR spectroscopy is a reliable alternative technique to determine OC and N in forest soils.
Finally, the spectral database reported in this study could be used to support soil survey in other areas of the Calabria region.

Modelling of landslide spatial hazard along a section of motorway of the Calabria region (Southern Italy) using a GIS-based statistical method

Valeria Rago, Massimo Conforti, Francesco Muto, Pasquale Versace

Landslides are one of the common hazard in Calabria (southern Italy). They are responsible both of direct and indirect damages and may cause loss of human life. Therefore, landslide hazard zonation is important in order to get mitigation measures and for planning of the territory.
This work presents a GIS-based bivariate statistical method to calculate the spatial probability of landslide occurance along a section of motorway “A3, Salerno-Reggio Calabria”, northern Calabria, where landslides are widespread. This study is included in a wider research project, named: PON01-01503, Landslides Early Warning-Sistemi integrati per il monitoraggio e la mitigazione del rischio idrogeologico lungo le grandi vie di comunicazione - aimed at the hydrogeological risk mitigation and at the early warning along the motorways.
The landslide spatial hazard evaluation started with geological and geomorphological analysis, based on air-photo interpretation and field survey, followed by processing and management of collected data through a Geographic Information System (GIS). Detailed landslide inventory map at the 1:10000 scale was carried out and a total of 835 landslides were mapped and the type of movement are represented mainly by slides and complex and subordinately flow. In order to estimate and validate landslide susceptibility map, the landslide inventory was randomly divided in two group. One group (LS-training set) was used to prepare susceptibility map and the second group (LS-validation set) to validate the susceptibility map. To evaluate the landslide susceptibility a bivariate landslide susceptibility

index (LSI) in GIS environment was used. Lithology, distance from faults, land use, slope, aspect, stream power index (SPI) and plan curvature were assumed as predisposing factors (PF). Intersection between layer of the LS-training set and each PF map allowed calculating the areal density of the LS-training set in each PF class.
Each PF map was then reclassified on the basis of the calculated weighting values.
In order to obtain the spatial hazard the weighting values were summed for each point of the study area by means overlay processes. The hazard map was classified into four classes: low, moderate, high and very high. The accuracy of the map was test crossing the group of the landslide of LS-validation set with the spatial hazard map and by using the computation of receiver operating characteristic (ROC) curve. The results obtained show the good reliability of the GIS-based statistical model to identifying landslide-prone areas. Finally, spatial hazard zonation represents an important step for landslide risk assessment.

The localized reduced basis multi-scale method

M. Ohlberger, F. Schindler

We are interested in the efficient and reliable numerical approximation of elliptic parametric multi-scale problems which consist of finding $p_h(mu) in V_h$, such that $b_h(p_h(mu) , q_h; varepsilon, mu) = l(q_h)$ for all $q_h in V_h$ for an a-priori given multi-scale parameter $varepsilon > 0 $ either in a multi-query context, where we want to solve for many parameters mu, or in a real-time context, where we have to solve for some parameters mu as fast as possible.
Model reduction using reduced basis (RB) methods is a well-established and reliable technique to reduce the computational complexity of parametric problems with respect to mu.
In the context of multi-scale problems, however, standard RB methods may become computationally too expensive.
The localized RB multi-scale (LRBMS) method was introduced in [F. Albrecht and B. Haasdonk and S. Kaulmann and M. Ohlberger, ``The localized reduced basis multiscale method", textit{Proceedings of Algoritmy 2012, Conference on Scientific Computing, Vysoke Tatry, Podbanske, September 9-14, 2012}, 393--403, (2012)] as a combination of model reduction and numerical multi-scale methods to overcome the shortcomings of classical RB methods.
We will present recent advances in the context of the LRBMS based on an efficient a-posteriori error estimator and present applications in the field of subsurface flow problems.

Granger causality estimate of information flow in temperature fields is consistent with wind direction

Nikola Jajcay, Jaroslav Hlinka, David Hartman, and Milan Paluš

Granger causality analysis is designed to quantify whether one time series is useful in forecasting another. We apply the time domain Granger causality analysis based on autoregressive processes to gridded daily surface air temperature data. For each grid-point pair, the direction and strength of the causal influence were computed with the one-day lag, effectively assessing the direction of the information flow in the temperature field. In order

to remove the influence of different distances of the grid-points in the original angularly regular grid of the NCEP/NCAR reanalysis, the data were transformed into an equidistant geodesic grid of 642 grid points. The strongest causalities have been found in the Northern Hemisphere's extratropics, where the temperature information is flowing eastward, in agreement with the prevailing westerlies. In contrast, only weak causalities have been observed in the tropics, which may be arising from higher spatio-temporal homogeneity.
In the second step, we quantitatively compared this estimate of information flow with the actual wind directions from NCEP/NCAR reanalysis data transformed onto the equidistant geodesic grid of 642 points. This was done for the surface layer and for the 850, 700, 500, 300 and 100hPa layers. The direction of the information flow matches the flow of the air masses, particularly well in the Northern Hemisphere's extratropics, i.e. for the strongest causalities. This agreement holds throughout the troposphere, slightly increasing with the height up to 500hPa level, then remains the same until bottom stratosphere. The agreement between the information flow in the air temperature field and the flow of air masses suggests the Granger causality as a suitable tools for constructing directed climate networks.

A state-space model for the analysis of seismic swarms

Elisa Varini, Yosihiko Ogata

It is widely recognized that earthquake clustering is a main feature of the seismicity and a seismic area can be affected by different types of earthquake clusters, such as aftershock sequences and swarms, due to its peculiar tectonic and volcanic environments. Different occurrence rates are expected to be observed in a sufficiently long period, each corresponding to and characterizing a different type of earthquake clusters.
We propose a probabilistic approach to model different types of earthquake clusters, also named states of the system, in order to identify and quantify them. To this end, we assume a state-space model (X,Y) in which the states of the hidden (unobserved) process X drive different realizations of the observed process Y. The earthquakes (observations) are first associated with a state and, conditioned on that state, follow an ETAS (Epidemic-Type Afetrshock-Sequence) point process; it follows that the hazard function of the observed process Y has explicit form conditionally to the state. The hidden state process X is assumed to be a pure jump Markov process.
The problem of the likelihood approximation is solved by particle filtering technique and parameter estimation is dealt with by Markov Chain Monte Carlo method in the Bayesian framework.
We analyse an earthquake sequence occurred off the east coast of Izu Peninsula (Japan) in 1998.

Permafrost research: monitoring through a sensors network at the 'Istituto Mosso' LTER site (Col d'Olen, Aosta Valley, Italy)

N. Colombo, S. Fratianni, D. Guenzi, F. Acquaotta, E. Giaccone, M. Giardino, L. Perotti, M. Freppaz, D. Godone, D. Said Pullicino, M. Martin, D. Viglietti, G. Viviano, F. Salerno, R. Balestrini, C. Delconte, G. Tartari, W. Alberto, A. Tamburini, M. Isaia, M. Gobbi, A. Merlone, S. Gruber

Permafrost is a specific ground thermal condition and rock glaciers are the most prominent geomorphic features of alpine permafrost. Atmospheric warming is likely to have strong impacts on permafrost, making cold areas at high elevation especially vulnerable. Therefore, climatic evolution and its impact on abiotic and biotic components of permafrost environments is a research topic of increasing importance. An integrated research and monitoring site on alpine permafrost has been set up in the active Col d'Olen rock glacier as a contribution to the Italian network of Long Term Ecological Research (LTER-Italia).
Climatic data from manned and automatic weather stations have been already collected in the study area and they will be updated and analysed in the future. Moreover, a network of portable instruments has been established on the rock glacier's body for collecting meteorological data, after a dedicated calibration to assess the uncertainties of the measures, in the context of the collaboration with MeteoMet2. The physical and chemical characteristics of the rock glacier outflow are under investigation by using temperature dataloggers and a multi-parametric probe spectro::lyser. Regular sampling of water (on weekly basis) has been also established for investigating major ion concentrations, trace elements and isotopic analyses. Moreover, the physical and chemical characteristics of fine-grained material are under analysis. The ground surface temperature monitoring will be conducted using Maxim iButton DS1922L mini-thermocrons and Hobo TidbiTv2 temperature loggers, regularly distributed on the rock glacier's surface and in a few selected surrounding sites. Total station and differential phase GPS will allow an accurate grid distribution and to acquire the coordinates of the dataloggers. In addition, high-resolution digital terrain models and thermal images of the rock glacier area will be obtained using a terrestrial laser scanner and an unmanned aerial vehicle (UAV). Finally, ground-dwelling arthropod colonisation of the rock glacier's body will be studied.

Variability of the b value in the Gutenberg-Richter distribution

C. Godano, E. Lippiello, L. de Arcangelis

The b value of the Gutenberg - Richter distribution is estimated as a function of a threshold magnitude m_{th} and it is found to depend on m_{th} for magnitudes larger than the completeness magnitude mc. We identify a magnitude interval $[m_c;m_m]$ where b is a decreasing function of m_{th} followed by a regime of increasing b for large magnitudes. This is a common feature of experimental catalogues for different geographic areas. The increase at large m_{th} is explained in terms of an upper magnitude cut-off in experimental catalogues due to finite size effects. We develop a rigorous mathematical framework to relate the decrease of b in the intermediate regime to the functional form of the distribution of the b values. We propose two hypotheses: The first is that the spatial and temporal variability of b leads to a b distribution peaked around its average value. The second is that mainshocks and aftershocks

are distributed according to the Gutenberg-Richter law with different b values, leading to a bimodal distribution of b. Simulated Epidemic Type Aftershock Sequences (ETAS) catalogues, generated according to this hypothesis, exhibit the same magnitude distribution of experimental ones. In alternative we cannot exclude the b dependence on m caused by magnitudes not homogeneously evaluated in a seismic catalogue. In the latter scenario our results provide the correction terms to the estimated magnitudes.

Optimization of the smoothness parameters in the Gaussian regression analysis for the modeling of typhoon trajectories

S. Nakano, K. Suzuki, and G. Ueno

We constructed a model of the typhoon's translation velocity as a function of latitude, longitude, day of year, and Julian day. This model is obtained from the typhoon best track data by using a Gaussian process regression analysis. In the Gaussian process regression, covariances between arbitrary two points must be given as smoothness parameters in advance. The optimal values of these smoothness parameters were determined by minimising the cross validation error to avoid the overfitting. The minimisation was achieved by the particle swarm optimisation. It was confirmed that the parameters determined by this method provided the reasonable velocity field pattern.

Overview of machine learning applications in environmental data mining

Michael Leuenberger and Mikhail Kanevski

The increase of environmental data of different nature: space-time monitoring, remote sensing images, in-situ measurements, etc., leads the geoscience community to develop and incorporate more robust and efficient methods, as well as new methodologies. Recent fundamental progress in machine learning algorithms (MLA) can contribute to the development of the field and can improve the results of the state of the art methods.
The aim of this study is to highlight and investigate the different issues that can occur when dealing with environmental data mining using contemporary machine learning algorithms. The major scientific problems can be formulated as follows: 1) development of a coherent and self-consistent methodology based on MLA for the recognition, modelling and prediction of structured patterns in environmental data; 2) construction and analysis of high dimensional input/feature spaces (in a validity domain) by using expert knowledge and application of a variety of features selection/features extraction algorithms; 3) ensemble learning of data for better predictability; 4) multi-task learning of multivariate data; 5) quantification of uncertainties for intelligent decision making process; 6) development of visual analytic tools for better understanding and communication of data and results. The poster presents some recent developments in machine learning which can give efficient answers to the mentioned problems. The applied part of the research deals with an application of the methodology and the methods developed for simulated, modeled and real environmental risk and natural hazard data: spatio-temporal pollution of the environment, landslides, forest fires, avalanches, permafrost and others.

Coincidences of tree-ring responses, simulated vegetation patterns and climate extremes

A. Rammig, M. Wiedermann, J. F. Donges, F. Babst, W. von Bloh, D. Frank, K. Thonicke and M. D. Mahecha

Climate extremes can trigger exceptional responses in terrestrial ecosystems, for instance by altering growth or mortality rates. Therefore one key question is which type of climate extremes may lead to substantial impacts on the carbon cycle and what is their spatial and temporal extent. We introduce here the method of coincidence analysis as a unit free measure to generally quantify the rate of simultaneously occurring extreme events between two sets of time series. We successfully apply this method to estimate impacts of extreme climatic events on European ecosystem productivity. We generate time series of ecosystem productivity by applying the Dynamic Global Vegetation Model LPJmL driven by the WATCH-ERA-Interim climatology. We then define different types of extremes in the climate time series such as heat events (e.g. 10% upper quantile of temperature) and drought events (e.g. 10% lower quantile of precipitation) during the growing season. These event time series are compared with time series of low net primary productivity (NPP, lower 10% quantile) as an indicator for the vulnerability of ecosystems to climate extreme events. For each pair of climate and ecosystem time series we obtain a rate of significantly co-occurring extreme events where the spatial extent of significantly reduced NPP values varies with the considered plant functional type (PFT) and climate variable. Additionally, we investigate the impact of simultaneously occurring heat and drought events on NPP and find that if such events take place the ecosystem's productivity is reduced even further. Our results imply that under an ever changing climate with an increasing rate of extreme events the ecosystem's production becomes significantly affected. Analyzing the combination of different extremes and their impact on different levels of ecosystem productivity (e.g. gross primary productivity, net ecosystem productivity or aboveground biomass) will lead to a deeper understanding of potential impacts of extreme events on the carbon cycle and help to estimate potential climate feedbacks from extreme events.

Complex network analysis helps to identify impacts of the El Nino Southern Oscillation on moisture divergence in South America

Niklas Boers, Reik V. Donner, Bodo Bookhagen, Jürgen Kurths

We investigate the temporal evolution of moisture divergence and its spatial clustering properties over South America. Our analysis focuses on dependencies on the El Nino Southern Oscillation. Moisture divergence is computed from daily reanalysis data of vertically integrated moisture flux provided by MERRA for the time period from 1979 to 2010. We use a sliding-window approach to construct a sequence of complex networks, each obtained from synchronization of events of strong positive (negative) moisture divergence, which we interpret as strong evapotranspiration (precipitation) events.
We make the following three key observations: i) Moisture divergence values over the Amazon rainforest are typically higher during positive ENSO periods (El Nino events). ii) The spatial coherence of strong positive (upwelling) events assumes a characteristic pattern of reduced coherence in this area during El Nino conditions. This influence of ENSO on moisture divergence and its spatial coherence is dominated by the El Nino events of 1982, 1987, and

1997. iii) The results on the clustering characteristics of the obtained climate networks qualitatively agree with the spatial distribution of connected regions with simultaneous events (i.e., events that occur at the same time), but provide a more detailed view on the spatial organization of strong atmospheric upwelling events. Interestingly, no comparable results are found for negative extremes of moisture divergence (strong precipitation events).

On the evaluation of climate model simulated precipitation extremes

A. Toreti, P. Naveau

Precipitation extremes represent a global threat for human and natural systems, especially in a climate change context. Therefore, it is essential to achieve a better understanding of those events and improve the current evaluation of climate model simulations. Here, we propose a statistical approach (based on a modified 2-sample Anderson-Darling statistic) that can be applied to address these issues. Statistical simulations as well as an application to gridded observations over Europe and eight climate model runs done in the framework of the Coupled Model Intercomparison Project Phase 5 (CMIP5; historical and scenarios RCP8.5 and 4.5) show the main features and the performance of the proposed method. The analysis over Europe reveals a lack of simple scaling relation between models and observations in the southern part of the domain and a significantly different tail behaviour in the whole region. The inter-model comparison shows a better agreement among models (i.e., a scaling factor seems to exist) although significant differences affect the modelled heavy-to-rare extremes. Concerning the projections for the 21st century, the majority of the investigated models shows an increase of the risk associated with precipitation extremes for large areas of the region, especially at the end of the century. Higher uncertainty characterises the areas where no significant differences are detected.

List of Participants

Adelfio, Giada	University of Palermo, Italy	giada.adelfio@unipa.it
Argiento, Raffaele	CNR-IMATI, Milan, Italy	raffaele@mi.imati.cnr.it
Barbero, Diego	University of Turin, Italy	diego.barbero@unito.it
Barnes, Solomon Ekow	Vukuzakhe High School, Durban, South Africa	solomonekowbarnes@yahoo.co.uk
Betro, Bruno	CNR-IMATI, Milan, Italy	bruno@mi.imati.cnr.it
Bodini, Antonella	CNR-IMATI, Milan, Italy	antonella.bodini@mi.imari.cnr.it
Bramati, Maria Caterina	Sapienza University, Rome, Italy	mariacaterina.bramati@uniroma1.it
Busnardo, Enrico	Universita degli Studi di Padova, Italy	enrico.busnardo@studenti.unipd.it
Ceglar, Andrej	EC JRC, Ispra, Italy	andrej.ceglar@jrc.ec.europa.eu
Conforti, Massimo	CNR-ISAFOM, Rende, Italy	massimo.conforti@isafom.cnr.it
Corral, Alvaro	CRM, Barcelona, Spain	acorral@crm.es
De Saedeleer, Bernard	Universite Catholique de Louvain, Belgium	Bernard.DeSaedeleer@uclouvain.be
Di Mauro, Biagio	University of Milan-Bicocca, Italy	b.dimauro@campus.unimib.it
Donner, Reik	Potsdam Institute for Climate Impact Research, Germany	reik.donner@pik-potsdam.de
Godano, Cataldo	Second University of Naples, Italy	godano.cataldo@unina2.it
Guenzi, Diego	Universita degli Studi di Torino, Italy	diego.guenzi@unito.it
Imagbe, Lucky	Royal Holloway University, London, UK	lucky.imagbe.2013@live.rhul.ac.uk
Jajcay, Nikola	Czech Academy of Sciences, Prague	jajcay@cs.cas.cz
Jha, Prakash	Ca' Foscari University of Venice, Italy	jprakash23@yahoo.com
Kanevski, Mikhail	University of Lausanne, Switzerland	Mikhail.Kanevski@unil.ch
Kondrashov, Dmitri	University of California, Los Angeles, USA	dkondras@atmos.ucla.edu
Kretschmer, Marlene	Potsdam Institute for Climate Impact Research, Germany	kretschmer@pik-potsdam.de
Lasaponara, Rosa	CNR-IMAA, Tito, Italy	rosa.lasaponara@imaa.cnr.it
Lefebvre, Mario	Ecole Polytechnique, Montreal, Canada	mlefebvre@polymtl.ca
Lenoir, Guillaume	Universite Catholique de Louvain, Belgium	guillaume.lenoir@uclouvain.be

Leuenberger, Michael	University of Lausanne, Switzerland	michael.leuenberger@unil.ch
Mistry, Malcolm	Ca' Foscari University of Venice, Italy	malcolm.mistry@unive.it
Nakano, Shinya	Institute for Statistical Mathematics, Tokyo, Japan	shiny@ism.ac.jp
Naveau, Philippe	CNSR-IPSL, Gif-sur-Yvette, France	naveau@lsce.ipsl.fr
Parding, Kajsa	University of Bergen, Norway	kajsa.parding@gfi.uib.no
Rago, Valeria	University of Calabria, Rende, Italy	valeria.rago@unical.it
Rotondi, Renata	CNR-IMATI, Milan, Italy	reni@mi.imati.cnr.it
Schindler, Anne	Meteoswiss, Zurich, Switzerland	anne.schindler@meteoswiss.ch
Schindler, Felix	WWU Munster, Germany	felix.schindler@wwu.de
Tondi, Rosaria	INGV, Bologna, Italy	rosaria.tondi@ingv.it
Toreti, Andrea	EC JRC, Ispra, Italy	andrea.toreti@jrc.ec.europa.eu
Uhlig, Marianne	Karlsruhe Institute of Technology, Germany	marianne.uhlig@kit.edu
Varini, Elisa	CNR-IMATI, Milan, Italy	elisa@mi.imati.cnr.it
Wei, Shangguan	Beijing Normal University, China	shanggv@bnu.edu.cn
Werner, Johannes	University of Bergen, Norway	johannes.werner@geo.uib.no
Wiedermann, Marc	Potsdam Institute for Climate Impact Research, Germany	marc.wiedermann@pik-potsdam.de
Zhou, Yu	MPI-PKS, Dresden, Germany	zhouyu@pks.mpg.de

A CNR-IMATI via Bassini 15, Milan

CD CONFERENCE DINNER Tuesday 19:30
Restaurant Il Calabrone 2, via Muzio Scevola 11

	RESTAURANTS (* also pizza)	**Address**	**Opening**
1	**Sunshine**	via Salieri 1	Sun-Thu
2 *	**Anthony**	via Pacini 67	Mon-Sat
3	**Osteria di Lambrate**	via degli Orombelli 18	Mon-Sat
4 *	**Strambio 6**	via Strambio 6	Sun-Fri
5 *	**Al Basilico Fresco**	viale Abruzzi 21	Mon-Sun
6	**Skuisito**	via Pacini 18	Mon-Sat
7	**Baia Chia**	via Bazzini 37	Mon-Sat
8 *	**La cuccuma**	via Pacini 26	Mon-Sun