TIES 2002
Annual Conference

COLLECTION OF ABSTRACTS

June 18-22, 2002

Facoltà di Ingegneria, Università di Genova
Villa Cambiaso, Via Montallegro 1
Genova, Italy

The International ENVIRONMETRICS Society - TIES

Acknowledgements

The Organizers of TIES 2002 are grateful to all the Authors whose contributions have been accepted for presentation at the 2002 Environmetrics Conference in Genova.
The contributions, presented here in a rigorously alphabetic order, cover a wide range of topics, all of them in the most advanced areas of each topic.
The participation of scientists from a large number of different countries shows the importance that a young science (Environmetrics) has gained worldwide.

We are indebted to our colleagues Fedele Greco, Michele Scagliarini, Giorgio Tortarolo and Carlo Trivisano for their constant and enthusiastic assistance in all the phases of the conference preparation.

Finally, the financial support received from Università di Genova, Regione Liguria and J. Wiley Editor is gratefully acknowledged.

Organizers

Daniela Cocchi
Vincenzo Dovì
Sylvia Esterby

Sponsoring Organizations:

The Committe on Probability and Statistics in the Physical Sciences of the Bernoulli Society for Mathematical Statistics and Probability

Società Italiana di Statistica

The International Environmental Modelling and Software Society

ASA Statistics and Environment Section

Associazione Italiana di Ingegneria Chimica

Linear Modelling of Monthly Temperature Data

Teresa Alpuim
University of Lisbon
Bloco C2. Piso 2 Cidade Universitária. Campo Grande.
1700 Lisboa, Portugal
Telephone 351-21 7500021
Email talpuim@fc.ul.pt

In this talk we will use linear models to detect changes in Lisbon temperature. A set of monthly average temperatures, from 1856 to 1999 will be analysed. The model will take into account the seasonal variation trough the year, the increase in temperature, allowing for different rates of increase corresponding to each month of the year, and a cyclic component with a period of about 42 years. It will be seen that the model fits well the data but the residuals present a moderate time correlation. In view of this, recent theory developed by Alpuim and El-Shaarawi allowing optimal estimation and statistical inference in linear models with time correlated residuals will be applied to this set of temperature measurements. A parsimonious model including seasonal variation, a different time trend for each month of the year and a cyclic component, will be derived with the help of trigonometric polynomials and F-tests adjusted to the case of correlated residuals. A final test will be performed to show rigorously that there are significant differences in the annual rates of increase of temperature for each month of the year. Consequently, we conclude that not only the climate is getting warmer but the shape of the curve that describes the seasonal variation of temperature is changing.

Statistical modelling of pollutants concentration in the surface waters of River Ave Basin

Teresa Alpuim[1], **Arminda Manuela Gonçalves**[2]
[1]University of Lisbon, Portugal [2]University of Minho, Portugal
Departamento de Matemática Universidade do Minho Pólo de Azurém
4810-058 Guimarães, Portugal
Telephone 351 253 510433/351 253 418445
Fax 351 253 510401
Email **mneves@math.uminho.pt**

The hydrological basin of the Ave River is located north of Portugal and has an approximate area of 1390 km2. Its main route is the Ave River, which runs along 93,5 km, from its source to its mouth. This is a fertile area in agriculture and forests and is also an important industrial centre especially for textile manufacturing. There are about 340 registered industrial units, 230 of these are of the textile and garment sector. Many discharges, with no previous treatment are carried out in the waters, especially those related to industrial effluents from the textile industries. One of the main reasons for the extreme pollution of the waters is the fact that the social economical development of the region and the construction of substructures are out of phase. Consequently, were the industries are actually affected by the bad quality of the water they use and which they themselves polluted. However, only in 1990, a solution was approved for this "most polluted area";. The Integrated System for Cleaning the Ave Valley has been established and is made up of three subsystems, which include interceptors along the Ave River and its main streams flowing into drainage points and Stations for Treatment of Residuary Waters (ETAR). Through a special network for monitoring the quality of water has been able to control the quality of the water of this basin since 1978. Selecting a total of 20 sample points, in stations that gather information, has carried out this task. This selection is done monthly and is based on analytical, physical, chemical and bacteriological parameters in order to obtain a picture of the quality of the waters of this basin. It's the analysis of this stations the support of the research which we are doing. Our research program has the following aims: -develop methods for statistically modelling the concentration of pollutants in the surface water of this basin; -organize and collect the large quantity of observable data available, in particular, try to obtain indexes, which inform in a quick and clear way, the general population, on the quality of this water; -study and evaluate the performance of the ETAR's recently installed. The statistical methods employed use time series, spatial statistics and analysis of multivariable data.

Maximum ozone concentration forecasting by functional nonparametric approaches

German Aneiros-Perez, Herve Cardot[1], Graciela Estevez-Perez, Frdric Ferraty[1], **Philippe Vieu**[2]
[1]INRA Toulouse, Biometrie et Intelligence Artificielle
[2]Laboratoire de Statistique et Probabilites
Universitè Paul Sabatier
31062 Toulouse cedex, France

Prediction of maximum ozone concentration s of great importance, especially to ward o .the population and allow the authorities to take preventive measures soon enough. Ozone concentration and meteorological variables are now observed each hour or each ten minutes so that we get curves as covariates. Much work has been done in the statistical community to propose effective models for predicting ozone concentration one day ahead but there has been much less effort to study models that take account of the functional nature of theses data. We study here a new non-linear model based on kernel estimators that can take account of the functional characteristics of the data via the
estimation of a semi norm used to measure the distance between two functional observations. This approach s used to forecast maximum ozone concentration in Toulouse (France).We compare t to more classical technics and t gives promising results.

Human encroachment in National Parks in South Africa

M.W. Barry[1], B. Reyers[2], A.S. van Jaarsveld[1], K.I. Meikeljohn[3]
[1]Centre for Environmental Studies, University of Pretoria, [2]Dept. of Zoology and Entomology, University of Pretoria, [3]Dept of Geography and Geoinformatics, University of Pretoria
CFES University of Pretoria Pretoria 0002, South Africa
Telephone 27-12-4204283
Fax 27-12-4203210
Email **mwbarry@zoology.up.ac.za**

One of the main aims of protected areas is to separate biodiversity contained within their borders from threatening processes on the outside. These external threats have an impact on biodiversity within protected areas and much work has focused on determining these impacts and their effects on biodiversity. An area that has received little attention, however, is the study of human impacts within protected areas. Tourists and staff, as well as the infrastructure they require, have an impact on ecosystems in protected areas. Little is known of the size and implications of these anthropogenic disturbances that continue to expand as tourism develops. The present study assessed the human encroachment in several parastatal conservation areas in South Africa, including the Kruger National Park and Addo Elephant National Park. Collated data included remote sensed land cover, staff and tourist numbers, resource utilisation, roads, pylons and waste production dating back to 1980. These data were analysed using a variety of landscape metrics and other techniques that allowed for the determination of anthropogenic encroachment within these conservation areas. Results indicate an increase in both staff and tourist numbers with an associated increase in resource utilisation and waste production. Infrastructure development has contributed to increased levels of habitat fragmentation, as is evidenced by the time series analysis of remote sensed images. Extrapolation of these results of human encroachment questions the sustainability of tourism development and conservation within National Parks in South Africa.

Describing the air quality in Rome

Sabrina Bellantone, **Giovanna Jona Lasinio,** Simona Sorrenti
Dipartimento di Statistica, Probabilità e Statistiche Applicate
Università di Roma "La Sapienza", Italy
Telephone 39-0649910473
Email **giovanna.jonalasinio@uniroma1.it**

In this study we analyse the air quality in Rome using data collected by the monitoring network of the town's municipality in the years 1998 and 1999. We first describe the pollutants behaviour by means of air pollution indices as proposed in Bruno and Cocchi (2002) , according to this approach we use all the stations included in the monitoring network regardless if they collect all the same type of pollutants. In this way we are able to describe the general "state" of the air quality. A second step is to propose a use of the same indices in order to describe the spatial variation of CO, NO2, O3 and PM10 simultaneously using data collected only by stations that register the same type of pollutants.
As a third step our work proceed by building a classification of the air quality of "the day" based on the same indices using neural network (classical cluster analysis didn't give valid results); an LVQ architecture has been chosen as the best with these data. This classification is the base to attempt a day by day prediction of the air quality using neural network. The LVQ architecture gives excellent results.
Special attention is then given to Benzene which has been studied separately in order to verify if the political intervention taken in the year 1998 was effective . The classification using neural network highlighted how in 1999 the level of Benzene was strongly reduced compared to the previous year.
Essential References
F. Bruno, D. Cocchi, A unified strategy for building simple air quality indices, Environmetrics, 13, 3, pp. 243-261.
W.R. Ott Environmental indices: theory and practice, Ann Arvor Science, Michigan 1978
M. Smith. Neural networks for statistical modeling. Van Nostrand Reinhold 1993.

Modelling precipitation at dense networks in large geographical areas

Enrica Bellone
National Center for Atmospheric Research (NCAR)
P.O. Box 3000 Boulder,
CO 80307-3000, USA
Telephone +1 303 497 1724
Fax +1 303 497 1333
Email enrica@ucar.edu

The assessment of the effects of climate change and climate variability in several fields -- e.g. crop yields -- requires simulations of realistic precipitation fields under current and altered scenarios. Nonhomogeneous hidden Markov models (NHMMs) provide a relatively simple framework for simulating precipitation at multiple rain gauge stations conditional on synoptic atmospheric patterns. The example presented here involves 175 rain gauges that cover the South-East region of the United States. Given the size of the network, direct application of the NHMM methodology would be extremely computationally intensive, if not infeasible. To overcome this problem, simulations of spatially coherent precipitation fields are obtained in three stages. The first step consists of identifying different geographical regions. Then NHMM fits are performed separately for each region. The third step involves re-connecting the various regions and simulating precipitation data that reproduces the spatial dependence observed in the data.

Is the Earth's climate changing?

Presidents Invited Lecture
Lennart Bengtsson
Max Planck Institute for Meteorology, Hamburg, Germany and Environmental Systems Science Centre, University of Reading, UK.
Deutsches Klimarechenzentrum Bundesstrasse 55
20146 Hamburg, Germany
Telephone 040 / 41173 - 349
Email bengtsson@dkrz.de

The climate of the Earth has over the last 500.000 years undergone major climate changes. So far at least 4 major glaciation cycles have been identified from ice core measurements at Vostok, Antarctica. The glaciation periods dominate some 80% of the time, interrupted by short interglacials lasting some 20.000 years or so. Mankind has presently spent some 10.000 years in such an interglacial. During the glacial periods the amount of land ice were some 25 milj. km3 more than now covering significant areas of the higher latitudes.
On much shorter timescales of several hundred years, the climate in certain regions at least have been reported to undergo variations. The warm episode during the early medieval time and the cold periods during 1600-1850 called the Little Ice Age in Europe, are examples of such events.
During the 20th century we can identify two periods of warming, one during the first part of the century and another during the last 25 years. A period of colder climate/weather occurred in the 1960s and 1970s. A reliable observing system making it possible to analyse the global atmosphere accurately with high time and space resolution as a matter of fact only exist for the last 25 years and for the Northern Hemisphere extra tropics for some 50 years. Surface observations are available since the late 19th century covering some 50% of the Earth and in some limited regions 50-100 further back in time. Consequently our knowledge of climate and climate change is strongly dependent on indirect information and the analysis of incomplete and partly less accurate meteorological records. There is a broad scientific agreement that the long term climate changes of the ice ages are caused by variations in the Earth´s orbit around the sun, while the changes on centennial time scales are open to debate. Our theoretical knowledge of climate is nowadays mainly coming from the study of numerical models of the Earth climate system. Such models have undergone an enormous development in recent decades partly as a consequence of computer performance development. The models are further systematically used in weather and short term climate prediction such as for El Niño. Model improvements and more accurate observations from satellite based systems have led to significant improvements in predictive skill in recent years, thus increasing the confidence in climate models. Based on advanced model studies the indications are that climate variations during the last thousand years or so are dominated by internal chaotic processes in the climate system. Such internal processes can create significant climate anomalies over several decades. Possible variation in solar irradiation cannot be ruled out, but so far only indirect data exist. Major volcanic eruptions do also contribute but hardly beyond a few years. There are strong indications that the warming in the first part of the last century, which was geographically limited, was caused by natural internal processes in the climate system. However, this appears not to be the case for the present warming, which has a different and more widespread distribution. Here the indications are that this warming is driven by the increasing concentration of greenhouse gases.
In my lecture I intend to give a broad review of our present understanding of climate, discuss results from climate models, analyse their interpretation and examine conditions for climate change during the 21st century.

Estimating Upper Ocean Currents in the Labrador Sea

Thomas Bengtsson[1], Ralph Milliff[2], Richard Jones[2]
[1]Geophysical Statistics Project National Center for Atmospheric Research Climate and Global Dynamics Division
[2]Colorado Research Associates
Email **tocke@ucar.edu**

The upper ocean dynamics of the Labrador Sea in winter have implications for understanding the massive air-sea exchanges of heat, water vapor, and momentum that drive the global thermohaline circulation. Through the brief and violent process of ocean deep convection, this global-scale overturning circulation (sometimes referred to as the "great ocean conveyor") sequesters the air-sea interface physics and chemistry from the surface for hundreds and even thousands of years. To the extent that subducted parcels of surface fluid can be identified "along the conveyor", they can provide information on trace species (e.g. [CO_2]) and surface properties (e.g. inferred temperatures) from past climates.
The Labrador Sea is one of a very few places on the globe where the ocean deep convection process is observed to occur. In the pre-conditioning of the upper ocean to set-up ocean deep convection, we are interested in the convergences of momentum in the upper ocean current field. To this end, ocean current information is obtained from time series of in-situ drifting buoy position reports that are communicated in near-real time via satellite. From these data we wish to estimate the temporal and spatial scales of ocean eddies associated with inertial processes. An approximation of a continuous-time stochastic differential equation is employed to model buoy-drift, and the Kalman Filter is used to evaluate likelihoods and estimate parameters.

An experimental catchment for first flush flows in the town of Genova

C. Berretta, I. Gnecco, P. La Barbera, L. G. Lanza, A. Molini
University of Genova Department of Environmental Engineering
1 Montallegro
16145 Genova, Italy
Telephone 010 3532485
Fax 010 3532481
Email **berrettachristian@yahoo.it**

Runoff pollution and the relationship between pollutant load and runoff have been investigated in the framework of the first flush quality issue in the urban area of Genoa. A first flush flow monitoring device has been specifically constructed, which consists of an auto-sampler connected with a tipping bucket device for flow measurements. Two monitoring systems of this type have been installed in the experimental catchment of Villa Cambiaso (University of Genoa) for sampling separately roof and road runoff. Both water quality and flow measurements were collected from January to April 2002 by analysing the following parameters: total suspended solids, COD, NH4+, pH and heavy metals in dissolved form. Cumulative load curves were produced for all monitored water quality characteristics to analyse the occurrence and nature of the pollutant load on the drainage system due to first flush flows. The first flush phenomenon was observed for total suspended solids, COD and NH4+. Among the heavy metals analysed, significant concentrations of Zn, Pb and Cu were detected in road runoff due to the presence of a parking lot in the experimental catchment. Noticeable concentrations of Zinc were found in roof runoff, due to the construction material of gutters. In the monitored events the influence of the initial rainfall intensity and cumulated depth on the occurrence of the phenomenon was observed; the entity of the first flush strongly depends on the duration of the antecedent dry weather period. Complete characterisation of the first flush phenomenon will require an extended monitoring period at least over a complete hydrological cycle. In light of the present experience, it is proposed to install the monitoring system in the densely urbanised historical centre of Genoa. The final objective is to determine the first flush runoff volume that is required to design effective detention tanks for efficient treatment.

Empirical predictions of the Elbe oxygen deficit using data from different sources

Laurent Bertino[1], Wilhelm Petersen[2], Hans Wackernagel[1]
[1]Centre de Géostatistique, ENSMP, Fontainebleau, France,[2]Institute for Coastal Research, GKSS, Geesthacht, Germany
35 rue Saint Honore F-
77300 Fontainebleau, France
Telephone +33 1 64 69 47 77
Fax +33 1 64 69 47 05
Email **bertino@cg.ensmp.fr**

The river Elbe is one of the largest in central Europe. During their course from the Czech Republic to the North Sea through Germany, the Elbe and its tributaries cross large industrial and agricultural regions and the urban areas of, among others, Berlin, Prague and Hamburg. The river estuary in Hamburg harbour suffers from intermittent oxygen deficits (below 3mg/l) caused by the high loads of nutrients coming from upstream pollutions. Some constraints have been imposed on the activity of Hamburg harbour to manage the problem, but an efficient policy of this section of the river needs precise short-term forecasts of the oxygen deficit. The biological mechanisms controlling the oxygen balance have been clearly identified (Schroeder, 1997) but an accurate coupled hydrodynamical, chemical and biological model of the river is still missing. The complexity of biogeochemical processes in the Hamburg region under tidal influence makes this task very complex. The goal of the demonstration study is (1) to assess the usefullness of linear statistics for prediction of the oxygen deficit and (2) to provide guidelines for further biogeochemical modeling. Transport times are of crucial importance to model chemical reactions but they vary according to river runoff. A modified temporal coordinate is then proposed to determine them experimentally by cross-covariance analysis of time series of a tracer variable. The diffusion process in the Elbe hydrodynamics is put into evidence at that stage. Helicopter data (spatially rich) and data from fixed stations (with high temporal frequency) are combined in an empirical orthogonal functions approach to link the spatial features of the oxygen deficit to upstream measurements that can be used for operational forecasting.

Short term ground-level ozone forecasting with extreme value methods

Silvano Bordignon[1], **Francesco Pauli[2]**,
[1]Dip. di Scienze Statistiche, Univ. di Padova, [2]Dip. di Scienze Economiche e Statistiche, Univ. di Trieste
PO Box 80115 3508 TC Utrecht
Telephone +31 30 2533051
Email **francesco.pauli@econ.univ.trieste.it**

Ground-level ozone forecasting on a timely basis is of great concern for public environmental agencies. Typically a forecast should be available at least for one-day in advance, since the time required to prepare emission reduction measures is at least one day and preferentially a few days, depending on the logistic. The recent trend on forecasting models for ozone concentrations relies mainly on nonlinear modelling the relationship between ozone, precursors and meteorological conditions (see Bordignon et al. (2002) and the references therein for some examples). Unfortunately all these models provide ozone forecasts in term of conditional means and, consequently, they do not exhibit particularly good performances in forecasting high pollution episodes. Since extreme ozone concentrations are the events one is mainly interested, in this work we explore the possibility of one-day-ahead forecasting ozone concentrations by extreme value methods. First, we focus on measures of tail dependence to evaluate the extremal dependence behaviour of consecutive days ozone concentrations. Second, multivariate parametric models for extremal dependence, recently proposed by Coles and Pauli (2001), are introduced and inference procedures based on maximum likelihood and Bayesian methods are considered. Then predictive conditional densities from these models are obtained. Finally, the previous models are employed to forecast one-day-ahead daily maxima of ozone data taken from the air quality monitoring network of the Padova district in Northeast Italy. REFERENCES Bordignon S., Gaetan C. and Lisi F. (2002), Nonlinear models for ground-level ozone forecasting, Dip. di Scienze Statistiche, Un. di Padova (submitted). Coles S.G. and Pauli F. (2001), Models and inference for uncertainty in extremal dependence, Biometrika (to appear).

On the Use of Hierarchical Models in Method Comparison Studies

Alessandra R. Brazzale[1], Alberto Salvan[2], Stefano Roletti[3]
[1]Institute for Biomedical Engineering, National Research Council, Padova -[2]Laboratory of Environmental Hygiene, Italian National Institute of Health, Rome and Institute for Biomedical Engineering, National Research Council, Padova - [3]A.R.P.A. Piemonte, Dipartimento di Ivrea - SETIL workgroup (C. Magnani, coordinator)
LADSEB-CNR, Corso Stati Uniti 4,
I-35127 Padova (PD), Italy
Telephone +39 049 829.5761
Fax +39 049 829.5763
Email **alessandra.brazzale@ladseb.pd.cnr.it**

BACKGROUND: One of the objectives of the SETIL study (an epidemiological multi-centric case-control study which is currently in progress in Italy) is to investigate the role of extremely low frequency magnetic fields as a possible risk factor for some types of childhood cancer, including leukemia. Commercially available magnetic field meters are used to measure residential exposure of the cases and controls who enter the study. OBJECTIVES: Our aim is to promote the use of hierarchical models for assessing reliability and agreement of two or more types of measurement devices. We illustrate the approach by using repeated calibration data available from the periodical calibration of the magnetic field meters used in the SETIL study. METHODS: A linear model with nested random effects was fitted to the meter calibration data accounting for the different aspects of the calibration protocol while maintaining a straightforward interpretation of the parameters. RESULTS: Both meters proved to provide reliable measures of the true magnetic flux density being generated. Differences in overall accuracy were, however, assessed. CONCLUSIONS: The approach is very flexible and can easily be tuned to the different needs arising in the measurement agreement framework. It can be seen as an extension of the common practice of retrieving a measure of agreement from one-way random effects models.

Structural analysis of effectors of the oncogenic Ras protein

Marcus Brunnert
Department of Statistics, University of Dortmund
Marcus Brunnert Fachbereich Statistik Universität Dortmund
D-44221 Dortmund, Germany
Telephone 049 231 7555401
Fax 049 231 7555303
Email **brunnert@statistik.uni-dortmund.de**

The expanding knowledge about protein structures is necessary for the understanding of human cell functions. Two statistical methods concerning the analysis of protein structures will be discussed in the background of the oncogenic risk of Ras mutations. A modified statistical classification method of protein structures using state-space models will be presented. In this context, a modified filtering algorithm is used to calculate a classification criterion. The second statistical method enables the search for an optimal structure for an amino-acid sequence using a score function. This score function evaluates the hypothetical structure of an amino-acid sequence. Both statistical methods are applied to effectors of the Ras proteins. Ras is a protein that can switch on or switch off important cell processes referring to the regulation of cell growth. The oncogenic variants of the Ras protein can be found in 20-30% of human tumors.

A simple nonseparable space-time covariance model for ozone

Francesca Bruno[1], Peter Guttorp[2], Paul Sampson[2], Daniela Cocchi[1]
[1]Dipartimento di Scienze Statistiche, Università di Bologna,
[2]Department of Statistics, University of Washington - Seattle
Dipartimento di Scienze Statistiche
Via delle Belle Arti 41,
40126 Bologna, Italy
Telephone 390512098271
Fax 39051232153
Email **bruno@stat.unibo.it**

Geostatistical spatio-temporal models are usually characterized by having to specify not only space and time components, but also spatio-temporal interaction components of variation. In the literature there are several papers that consider the simplest hypothesis of separability between space-time covariance functions. This means that the space-time covariance structure is fully determined by the product of pure temporal covariance component and a simple spatial covariance structure. We proprose a new model for space-time covariance structure in which space-time nonseparability arises because of temporal nonstationarity (in scale). In particular, we apply the model to tropospheric ozone data in some sites of Emilia Romagna.

Modeling short-term health effects using air quality indices: an application to Bologna data (Italy)

Francesca Bruno, Rossella Miglio
Dipartimento di Scienze Statistiche "P.Fortunati", Università degli Studi di Bologna
Via delle Belle Arti 41,
40126 Bologna
Italy
Email **bruno@stat.unibo.it**

Epidemiological studies over the past decade suggest the existence of a persistent statistical association between increases in daily air pollutant concentration and increases in mortality. Generalized additive models for Poisson data are typically used to model daily mortality, adjusting for non-linear trends, and including smoothed or parametric functions and temporal lags for meteorological and pollution variables.

Lately, some interest is focused on developing appropriate statistical methods to evaluate the health effects of each air pollutant when pollutant acting alone (no other) and in case of pollutant acting in the presence of a variety of mixtures of other air pollutants. For this purpose, some studies examine the effects of a specified pollutant stratifying by the level of other pollutants and others include grouping (or combining) pollutants as defined by source profiles, grouping (or combining) pollutants as defined by principal component or factor analysis.

Our proposal is to consider as pollutant combination an air quality index. This index will be considered alone or with a single pollutant. Generalized additive models using air quality indices (Bruno and Cocchi, 2002) as predictor with (or in alternative to) single pollutant are applied on Bologna data for the period 1996-1999. For each pollutant three different scenarios are considered: a linear effect, two linear terms constrained at 100 index value and a dicomization of the index with respect to the 100 value. We have also considered models where each pollutant is present in conjunction with the index in the three different previously described forms. Model adequacy was checked by the analysis of the partial autocorrelation function and the analysis of residuals.

Heavy metals in tsp and pm10 measured in an industrial site of southern Italy: source characterization and dependence on meteorological conditions

R. Caggiano[1] **M. D'Emilio[2]**, M. Ragosta[2] M. Macchiato[3], S. Sabia[3]
[1]Istituto di Metodologie Avanzate di Analisi Ambientale CNR - Tito Scalo (PZ), Italy, [2]INFM - Dipartimento di Ingegneria e Fisica dell'Ambiente - Università degli Studi della Basilicata - Potenza, Italy, [3]INFM – Dipartimento di Scienze Fisiche -Università Federico II; Napoli, Italy
Contrada Macchia Romana
85100 Potenza, Italy
Telephone 0971/427267
Fax 0971/427271
Email **demilio@na.infn.it**

In this paper, we present the preliminary results concerning the study of metal content in Total Suspended Particulate and in PM10. Particularly we focus our attention on the characterization of source profiles. We measure daily atmospheric concentrations of particulate and its content of nine heavy metals (Cd, Cr, Cu, Fe, Mn, Ni, Pb and Zn) in a site located at industrial area of Tito Scalo (Basilicata, Southern Italy). The monitoring survey started in April 1997. Furthermore we measure in continuous meteorological data (temperature, rainfall, wind speed and wind direction); an automatic system collect all the data and give us, on daily scale, average ambient temperature, amount of precipitation, resultant wind vector, prevailing wind direction. In order to investigate the metal behaviour in particulate matter of different size and to better characterize the relationships among particulate size, metals and emission sources, we apply a multivariate procedure aimed to point out the relationships among different metals. The statistical method allows us to identify the correlation patterns and, consequently, to define the profiles of sources. Particularly we analyse the correlation structure of metal concentrations in TSP and in PM10 in different meteorological conditions for pointing out if in different data sets we may observe different abatement effects and different anomalous emission distribution.

An emission inventory at regional level: results and uncertainties

Stefano Caserini, Anna Fraccaroli, Anna Maria Monguzzi, Marco Moretti, Angelo Giudici, Cinzia Pastorello
Fondazione Lombardia per l'Ambiente,
Piazza Diaz 7,
20123 Milano,Italy
Email: stefano.caserini@polimi.it

In order to assess the role of pollutant load to atmosphere, the INEMAR emission inventory considers different sources (point, area, biogenic, road transport), and organizes all information needed for their estimation: activity indicators, emission factors, other statistical data necessary on spatial and temporal distribution of emissions.
Since the optimal methodology to set up an emission inventory, that implies direct quantification of all emissions of different source typologies for both areas and time of interest through measurements, can be used exclusively for some specific pollutants typologies and sources, emissions evaluation is mainly carried out by means of an indicator characterizing the activity of the source and its emission factor, specific of the industrial process and of the adopted flue gas control system. The reliability of this evaluation, based on a linear relation between the source activity and the emission, depends on the precision of emission factors.

COPERT methodology proposed by CORINAIR has been used for road transport emissions, with the basic distinction between non-urban driving (highway and other main roads) and urban driving. Traffic is the major contributor to atmospheric emissions for a great number of municipalities.
The database is now characterized by more than 1 million of records, regarding 18 pollutants (acidifiers as SO_2, NO_x, NH_3, CO, ozone precursors NMVOC, greenhouse gases as CO_2, CH_4, N_2O; PM_{10}, heavy metals and dioxin have also been considered), for each of the 1546 municipalities of the Lombardy region, for 198 source activities (referring to 11 groups and 45 sub-categories of the Corinair SNAP 97 nomenclature) and for 20 typologies of fuels.
A detailed emission data collection allows the spotting out of contributions of different sources to local and global air pollution, and consequently the assessment of more proper intervention strategies.

An automatic device for dynamic calibration of tipping bucket rain gauges

D. Cerisola, **A. Vallebona**, P. La Barbera, L. Lanza, A. Molini, L. Stagi
DIAM - Dept. of Environmental Engineering - University of Genova
Via Montallegro 1
16145 Genova, Italy
Telephone 390103532123
Fax 390103532481
Email vallenet@libero.it

The measurement of liquid precipitation at the ground is affected by different sources of both systematic and random errors, mainly due to wind, wetting and evaporation induced losses. The tipping-bucket rain gauge (TBR) is also known to underestimate rainfall at higher intensities because of the rain water amount lost during the tipping movement of the bucket, the so-called systematic mechanical error. Though easily avoidable by means of dynamic calibration, such errors produce biases in rainfall measurement that are commonly neglected in the management of many operational monitoring networks. An automatic mechanical device controlled by software has been developed for the dynamic calibration of TBRs. The operating principle is based on the generation of constant water flows, reproducing different rainfall intensities by means of varied water head and diameter orifices. A suitable calibration curve is obtained for each analysed instrument, which can be used to correct the recorded data. The influence of the systematic mechanical error on the estimation of the statistics of extreme events has been also analysed. Since systematic mechanical errors basically apply at the highest rain rates, correction of rain records and the assessment of the influence on rainfall statistics can only be performed at fine resolution in time, where intense rain rates may be commonly observed even during precipitation events totalising low to intermediate rain depths at larger scales. This has been achieved using a disaggregation random cascade algorithm applied to hourly historical rainfall series. Dissaggregated data have been corrected by means of the derived calibration curves, then re-aggregated at the original time-scale. The errors induced on the most common statistical tools used to characterise extreme rainfall, for e.g. design purposes, have been quantified based on the above methodology for precipitation time series recorded in the Liguria region of Italy.

Uncertainty in Ocean/Climate Models

Peter Challenor
Southampton Oceanography Centre
Southampton SO14 3ZH, UK
Thelephone +44 23 80596413
Fax +44 23 80596400
Email **P.Challenor@soc.soton.ac.uk**

Large numerical models are now ubiquitous in oceanography and climate studies. As such models move from being used simply for understanding the physical processes involved to the formation of policy, for example on greenhouse warming, it is important that we address the problem of what are the uncertainties on the output from such models. Comparison with data can give us a measure of model error, but even here to get the necessary quantity and coverage of data we are limited to measurements of the atmosphere and the ocean surface from satellites. Such information is useful but limited. We give an example which shows where and how large our uncertainties are but which cannot give us any information on where they come from. An alternative approach is to take the uncertainties on our model inputs (including internal model parameters) and propagate these through the model equations. The mathematics is too complex to do this theoretically and Monte Carlo methods are impractical due to the sheer size of the models involved – we are lucky if we get more than a single run of a state of the art model even using some of the fastest computers in the world. We will present an alternative based on work carried out at the Universities of Sheffield and Durham in the UK. Here, a statistical approximation to the model output, called an emulator, is created. This non-linear model (and its associated uncertainty) is used in place of the full model, either to theoretically propagate uncertainty or in a Monte Carlo simulation. We illustrate these methods with a 'toy' model of the Earth's climate and show how they could be scaled up to full-sized climate simulations, including the problems to be overcome.

Cumulative Risk Assessment Based on Quantitative Responses

James J. Chen[1], Yi-Ju Chen[1], Linda K. Teuschler[2], Glenn Rice[2], Karen Hamernik[2], Alberto Protzel[2], **Ralph L. Kodell[1]**
[2]USEPA; [1]USFDA.
Division of Biometry and Risk Assessment
3900 NCTR Road Jefferson,
Arkansas (AR) 72079 USA
Telephone 870-543-7008
Fax 870-543-7662
Email **rkodell@nctr.fda.gov**

The Relative Potency Factor (RPF) approach is a procedure used to normalize and combine different toxic potencies among a group of chemicals selected for cumulative risk assessment. The RPF method implicitly assumes that the slopes of the dose-response functions of the chemicals are all equal, i.e., that the relative potency is constant over all doses. Unless a constant-relative-potency constraint is enforced in the estimation of slopes, then the estimated dose-response of the mixture will depend on the choice of the index chemical. In this presentation, a procedure is proposed for cumulative assessment of risk from exposure to multiple chemicals that are assumed to have a common mode of action. Two classification algorithms are proposed to cluster the chemicals into subclasses such that the chemicals in the same subclass have a common slope. The joint response is estimated by fitting the dose-response model of the mixture under dose addition. Chemicals within subclasses are first combined using simple dose addition (constant relative potency), and then subclasses of chemicals are combined using a general form of dose addition (non-constant relative potency). Thus, the proposed method allows one to estimate the joint toxic response for chemicals having different dose-response slopes. Although the proposed method will give the same predicted mean response for toxic responses measured on a quantitative scale regardless of the choice of index chemical, when transformed to a risk scale the prediction will depend on the variance of the index chemical. If a homogeneous-variance model is imposed, then all choices of index chemicals will give the same predicted risk. A data set consisting of six hypothetical pesticide chemicals is used to illustrate the proposed procedure.

Weighted jackknife method in small-area estimation with an environmental application

Shijie Chen, **Partha Lahiri**
University of Nebraska- Lincoln
Lincoln, NE, USA
Email: plahiri@survey.umd.edu

Estimation of various environmental variables for small geographical areas is a challenging task because of lack of adequate samples to measure these variables for such small areas. In this paper, we develop an empirical best linear unbiased prediction (EBLUP) method which combines the primarysource of information on the environmental variable of interest with other related auxiliary information available from various secondary sources.
We estimate the mean square error of the proposed EBLUP using a weighted jackknife method which captures all sources of variabilities.
We demonstrate the usefulness of our method for a real-life small-area estimation problem involving environmental variables.

Mortality and air pollution: a dynamic generalized linear modelling approach

Monica Chiogna, Carlo Gaetan
Dipartimento di Scienze Statistiche Universita' di Padova – Italy
Via C.Battisti 241 35121 Padova, Italy
Telephone +39 049 8274183
Fax +39 049 8274170
Email **monica@stat.unipd.it**

We propose to model short term pollutant exposure effects on health by using dynamic generalized linear models. The time series of count data are modelled by a Poisson distribution having mean driven by a latent Markov process; estimation is performed by the extended Kalman filter and smoother. This modelling strategy allows us to take into account possible overdispersion and time-varying effects of the covariates. The ideas are illustrated by re-analysing data on the relationship between daily non accidental deaths and air pollution in the city of Birmingham, Alabama.

Small area solutions for the analysis of pollution data

Daniela Cocchi, Enrico Fabrizi, Carlo Trivisano
Dipartimento di Scienze Statistiche, Università di Bologna
Via delle Belle Arti 41
40126 Bologna, Italy
Telephone +39 051 2098234
Fax +39 051 232153
Email **cocchi@stat.unibo.it**

The term "small area" usually refers to a small geographic region or to any subset of a population. The usual finite population estimators of small area characteristics, as means or totals, show very large variances when the small area includes only few sample units, rising the need for more stable estimators. Small area modelling in finite populations has been extensively studied under the normal hypothesis. Solutions are usually found within the Empirical Bayes and the Hierarchical Bayes contexts.
Many models employed in environmental risk assessment needs small area estimation techniques, whenever local aggregates are involved (disease mapping, disease clustering,…)
Pollutants emission inventories are built by using both direct measurements of large point sources and indirect model-based estimates of diffuse sources. Indirect emission estimates for each pollutant are obtained by multiplicating a measure of "activity" related to the pollutant by an emission factor. Indirect estimates are characterized by high uncertainty and much effort has been devoted to its assessment. Nonetheless, most of them focus on the evaluation of the uncertainty due both to the estimation of the emission factor and the rigid multiplicative model.
In our talk we discuss some developments which do not need the normal assumption. In the special case of pollutants emissions inventories we focus on measures of activity which must be estimated on small samples. We propose small area models in order to obtain more stable estimates for these aggregates. The use of these estimators may in fact have a good impact on the reduction of the overall uncertainty associated to emission estimates.

Hierarchical modeling of extreme rainfall

Daniela Cocchi, **Michele Scagliarini**
Dipartimento di Scienze Statistiche, Università di Bologna
Via delle Belle Arti 41
40126 Bologna, Italy
Telephone +39 051 2098234
Fax +39 051 232153
Email **scagliar@stat.unibo.it**

Many real situations require to estimate the extremal behavior of an environmental process over a specified geographical region. A typical application is the study of extremal rainfall behavior within a catchment region, useful for the evaluation and prevention of hydrological risks. Researchers are often interested in a model holding for the whole region, though in practice only repeated pointwise rainfall measurements from a network of sites over the region are available. This paper proposes a modeling strategy for the extremal behavior of the aggregated process, starting from the pointwise data in the presence of auxiliary information at the site level. Situations occur typically where, besides the pointwise time series data, also cross section information about the different sites is available. The whole area is not considered as homogeneous and the existence of different sub-areas is hypothesized. This structure of information suggests a hierarchical formulation of the statistical model.

A new model for the inference of population characteristics from experimental data using uncertainties in data. Application to interlaboratory studies

Wim Cofino[1], D.E. Wells[2]
[1]Institute for Inland Water Management
and Waste Water Treatment (RIZA)
P.O. Box 17, 8200 AA Lelystad, The Netherlands
Tel.: (31 320) 298 633
Fax: (31 320) 297 642
email: W.Cofino@riza.rws.minvenw.nl
[2] Scottish Office, Fisheries Research Station, Marine Laboratory,
P.O. Box 101, Victoria Road, Aberdeen, United Kingdom.

A new model to make inferences about population characteristics from experimental datasets is presented. It has been developed in analogy with quantum mechanics. To each datapoint an 'observation measurement function' (OMF) is associated. This function encompasses the observed value and the uncertainty estimate thereof. The set of OMF's span a space in which 'population measurement functions' (PMF) are constructed. The PMF's can be associated with modes of the distribution. Knowledge of the PMF's enable the population characteristics to be computed. The model provides thus the different modes of the distribution and for each mode the expectation value, the standard deviation and a percentage indicating the fraction of observations encompassed (or 'how representative the particular mode is for the entire dataset'). An implementation of the model which does not require uncertainty estimates is provided too. In this poster, the model is elaborated and applied to the evaluation of interlaboratory studies. It has, however, a much wider generic application. It is demonstrated that the model can cope with asymmetric, strongly tailing and multimodal distributions and that it is superior to existing techniques (e.g. robust statistics).

Reference:
Cofino, W.P., I. van Stokkum, D.E. Wells, R. A.L. Peerboom, F. Ariese, A new model for the inference of population characteristics from experimental data using uncertainties in data. Application to interlaboratory studies, *Chemometrics and Intelligent Laboratory Systems* 53 (2000) 37-55

Monitoring Tailor Made III – Results of an International Workshop on information for sustainable water management

Wim Cofino
Institute for Inland Water Management
and Waste Water Treatment (RIZA)
P.O. Box 17
8200 AA Lelystad, The Netherlands
Tel.: (31 320) 298 633
Fax: (31 320) 297 642
email: W.Cofino@riza.rws.minvenw.nl

Autumn 2000, the international workshop Monitoring Tailormade was organised for the third time. The previous workshops focussed on the design and implementation of monitoring programmes and on the assessment of data. The third workshop had a broader scope. Present day watermanagement, as elaborated in Integrated Water Resource Management (IWRM) or Comprehensive Water Resource Management, acknowledges that water issues cannot be dealt with in an isolated fashion, but require to be considered in conjunction with social and economic issues. As a consequence, the information base has to be broadened too – integrated water resource management requires information over the water resources in relation to information from the social and economic domains. The benefits of measures in one area have to be coupled to the effects (costs) in the other areas, so that decision makers can make proper trade-offs. Indicators play an important role – indicators encapsulate the information in a meaningful way enabling the trade-offs to be made. Indicators also can constitute an effective means of communication between different stakeholders. Indicators, however, provide several problems as e.g scale mismatches, in particular for indicators derived by different disciplines, loss of information upon aggregation, mathematical problems related to the spatial or temporal aggregation of data, and uncertainty issues. For a proper making of the trade-offs, it is preferred that the indicators are assembled in an overall framework which gives insight in cause-effect chains. To this end, the D,P,S,I,R framework appears to have a wide support. Finally, institutional arrangements involving multidisciplinary teams are beneficial in elaborating the indicator framework and carrying out the assessments.
In the presentation, the outcome of the Monitoring Tailormade III workshop will be presented discusiing the items touched upon above.

Statistical Experiences of Estimating Bio-availability and Toxicity of Metals to Plants

Ray Correll, Daryl Stevens, Mary Barnes and Mike McLaughlin
[1] CSIRO PMB 2, GLEN OSMOND, S.A. 5064, AUSTRALIA

Critical concentrations of metals in soils are often derived by toxicity experiments where metals are added to soils and plant response determined, to create a dose-response function. Critical concentrations are then derived from effect concentrations where the plant experiences a 5% (EC5) or 10% (EC10) growth reduction. Doses may be expressed either as total or extractable metal concentrations, but the former is more common. Experimental and biological variations lead to imprecision in the EC5 and EC10 values. Furthermore, selection of the appropriate initial metal dose rates is problematic and a poor spread of observations may exacerbate imprecision. This paper describes options for determining appropriate dose rates to maximise precision of effect concentration values, and determine if expressing metal doses as extractable concentrations improves precision.

The amount of metal extractable with NH4NO3 may be a better measure of the available metal than is the total metal. In some trials, the extractable and total concentrations of the metal in the pots are recorded. The curves associated with each measure are not always parallel, which implies difficulty in extrapolating from total to available and EC50 to EC10. The error structure presents an interesting challenge in estimating the confidence intervals.

The Ozone Hole: Spatial Trend in a Massive, Global Dataset

The J. Stuart Hunter Lecture
Noel Cressie,
The Ohio State University, Columbus, OH, USA

As technology progresses, the availability of massive environmental data with global spatial coverage has become quite common. An example of such data is Total Column Ozone (TCO) remotely sensed from a satellite. In their raw form, these data are often spatially (and temporally) dense, but irregular. However, for practical use, the data are typically aggregated on a space-time grid at a given resolution. The resolution of the spatial grid that covers the entire globe needs to be sufficiently fine to be of use in answering a large variety of environmental questions, but there is a practical drawback of creating massive datasets that can be difficult to manage. The problem considered here is one of detecting large-scale spatial trend at a given time point (actually, in a given time interval). We propose a sequential aggregation method, producing different levels of coarser (spatial) resolution data and, at the same time, preserving both the local information content and the locations of the raw data. Each dataset of coarser resolution is used to estimate the large-scale trend in the data. In estimating the large-scale trend, we consider different parameterizations of a smooth spatial trend on the sphere, all linear in the data and satisfying the topological constraints imposed by the sphere. These parameterizations include spherical harmonics, tensor products of splines, and a spatial-covariance-based method. Each trend type is fitted to coarser resolution data and the fit is used to predict at the finest resolution, where comparison can be made to the original fine-resolution data. Results are obtained on the relative losses incurred by using different trend types and coarser-resolution data. The research presented in this talk is joint with Gardar Johannesson, graduate student at The Ohio State University.

The Need of Statistical Methodological Tools in the Compilation of Air Emission Inventories

Riccardo De Lauretis[1], Daniela Romano[2]
[1]ANPA (Agenzia Nazionale Protezione Ambientale)
[2]ANPA (Agenzia Nazionale Protezione Ambientale) consultant
via Vitaliano Brancati 48,
00144 Roma, Italy
Telephone +39 06 5007 2928
Fax +39 06 5007 2986
Email **delauretis@anpa.it**

The aims of a National Air Emission Inventory are mainly the monitoring of atmospheric emissions from all sources on a national (sub-national) scale, the verification of compliance with national and international protocols, the effects of environmental policy. Different statistical methodologies can be applied to fulfil the objectives, improving and assessing the quality of the inventory . The process of air emissions estimation implies at first the choice of emission factors and activity data; thus it is strictly related to the assessment of data with regard to their uncertainty. Statistical methods can be used for qualitative and quantitative analysis of the uncertainty; in particular, classical analysis, e.g. Montecarlo and bootstrap simulation models, and fuzzy analysis. Since the contribution of the sectors to the national total differs considerably, the analysis of key sources, as sources that need particular attention for data improvements and quality control, is very important. Ranking and classification methods, as well as non-parametric methods, can be applied to this extent; sensitivity analysis can also be used to evaluate key parameters. National estimates are sometimes used to estimate emissions at local level. Multivariate analysis and classification methods can be applied to individuate proxy variables and to improve top-down methodologies. The study of national time series, as well as the analysis of different scenarios and projections, can be carried out by trend analysis. The effects of environmental policy can be assessed and policy targets for emission reduction evaluated; thus better policies in terms of cost-benefits results, by means of optimisation methods, can be proposed or addressed. The application of all these methods guarantees a high quality emission inventory and a systematic quality assurance and quality control approach to be applied in inventory development in order to improve the reliability.

Welcome Address

Marco Del Borghi
Head of DICHEP, Università di Genova, Italy

This Environmetrics Conference is the last of a series of Conferences, the first of which was held in Cairo fourteen years ago. On that occasion the foundations of the International Environmetrics Society were laid.
It was felt that the Society might act as a veritable forum for a large number of scientists and engineers working in the field of "Applications of quantitative methods to environmental problems".
The decisive momentum to the creation of the Society was given by the insight of Prof. El-Shaarawi who was the first President of the Society and whom we are particularly glad to welcome today in Genoa.
The rapid growth in the number of members and the prestige gained by the Society worldwide show how far sighted and far reaching that decision was.
The official Journal of the Society (it has the name as the Society itself and is published by Wiley r (who is one of the sponsors of this Conference), covers today some of the most significant contributions in this sector and has become the most authoritative voice in "quantitative environmental methods".
One can easily realize the increasing importance of the thematic areas which form the core of the Environmetrics Society by simply looking through the titles of the contributions of this Conference. From climate change to ozone layer depletion, from the Kyoto accord to human health and epidemiology, all of them presented by the most distinguished scholars worldwide.
Due to lack of time I can mention but a few of them. I shall limit myself to the plenary lectures held by prof. Lennart Bengtsson on climate changes (which will inaugurate the Conference), by prof. Noel Cressie on the estimation of the ozone layer and by prof. Jan Pasman on human health risk assessment. The panel on the perspectives of quantitative methods in the environmental sciences will be a convenient epitome for the whole Conference and will point to future activities of the Society.
The University of Genoa (in particular the Faculty of Engineering) has always had an active part in the activities of the International Environmetrics Society. There have been contributions from our Faculty to nearly all Environmetrics Conferences held so far, our staff members have passionately taken part in the life of the Society.
Therefore I am particularly happy to welcome all participants (from over twenty countries) and I am proud that our Faculty has the possibility of hosting for five days the most advanced forum on quantitative methods for the environment.
My most sincere wishes for a fruitful and productive work in the next days.

On the dynamic scaling of rainfall fields

C. De Michele, P. Bernardara, N.T. Kottegoda, R. Rosso
DIIAR, Politecnico di Milano, Milan, Italy

The concepts of scaling and multiscaling, developed in the last twenty years, provide a powerful framework for studying spatial and temporal variability of numerous hydrological processes. Recently, the introduction of the dynamic scaling concept and its evidence in rainfall fields allows a link between the spatial and temporal scales and consequently a simplification of the mathematical tractability of the problem. On this route, the paper discusses the space-time scaling properties of rainfall fields. This issue is important in the identification of the statistical properties of rainfall used in the building of rainfall generators and downscaling models of rainfall. In particular, the paper considers the spatial and temporal variability of second order moments of the cumulated rainfall field. It shows how to link the scaling exponent in time with the scaling exponent in space through the dynamic scaling exponent. A self-affine equation explains the space-time variability of the cumulated rainfall surface. This is tested using two clusters of rainfall, large mesoscale cells, selected from the GARP Atlantic Tropical experiment, GATE.

An application of wavelet transform for the comparative study of two different prospection magnetic techniques

M. D'Emilio[1], M. Ragosta[1], D. Chianese[2], E. Rizzo[2] M. Lanfredi[3], V. Lapenna[3] Macchiato M. [4]
[1]INFM - Dipartimento di Ingegneria e Fisica dell'Ambiente - Università degli Studi della Basilicata Potenza, Italy, [2]Dipartimento di Ingegneria e Fisica dell'Ambiente - Università degli Studi della Basilicata - Potenza, Italy, [3]Istituto di Metodologie Avanzate di Analisi Ambientale CNR [4]Dipartimento di Scienze Fisiche -Università Federico II Napoli, Italy
Contrada Macchia Romana
85100 Potenza, Italy
Telephone 0971/427267
Fax 0971/427271
Email demilio@na.infn.it

Different monitoring techniques generally have different detection limits and are influenced in a different way by the environmental conditions in which they are used. Here we present a scale-space study on the inter-dependence among measurements obtained by means of two different magnetic techniques used for soil monitoring. The first one, susceptibility monitoring, is an active prospecting method based on the measure of the ratio of induced magnetisation to an applied low magnetic field (of the same order of the Earth's magnetic field). The other one, the magnetometric technique, is a passive prospecting method based on the study of the local variability of the Earth's magnetic field. These two techniques exploit different physical mechanisms and measure different variables which are related to soil composition. Both the methods are considered among the more promising non invasive techniques to monitor heavy metals in the soil. The data obtained by means of these techniques often shows non-stationary and transient dynamics. In fact, soil anomalies or discontinuities may introduce these kind of dynamics. Such local features are poorly represented or entirely missed when data analyses are based upon traditional methods, such as the standard Fourier Transform. We use wavelet transform to characterise the scale-space properties of experimental data. Owing to the ability of wavelets to represent both long and short range dynamics, we can make no preliminary assumptions about the stationarity or continuity of the data. Rather, wavelet transform may help to detect local correspondences between measurements obtained from the two techniques, which might be related to the presence of different structures in subsurface soil.

Non Parametric Tests and Confidence Regions for Intrinsic Diversity Profiles of Ecological Populations

Tonio Di Battista, Stefano A. Gattone
Università di Chieti
Viale Pindaro,42
65122 Pescara, Italy
Telephone +39 0854537551
Fax +39 085 4537542
Email **dibattis@dmqte.unich.it**

Evaluation of diversity profile is an useful tool for ecologists in order to quantify the diversity of biological communities. Measures of diversity profile can be expressed as function of the unknown abundance vector. Thus, the estimators and related confidence regions and test of hypotheses involve aspect of multivariate analysis. In this setting, using a suitable sampling design, inference is developed assuming an asymptotic specific distribution of the profile estimator. However, in a biological framework, ecologists work with small sample size and the use of any probability distribution is hazardous. Assuming that the sample belongs to the family of replicated sampling design, we show that the diversity profile estimator can be expressed as a linear combination of the ranked abundance vector estimator. Owing to this result we are able to develop a nonparametric approach based on bootstrap in order to build balanced simultaneous confidence sets and test of hypotheses for diversity profile. Finally, the proposed procedure is applied on the avian populations of four parks in Milan, Italy.

A Way to Model Periodicity in Environmental Modeling

Boyan Dimitrov[1], **Mohammed A. El-Saidi[2]**, Zohel Khalil[3]
[1]Kettering University, Michigan, USA, [2]Ferris State University, Michigan, USA, [3]Concordia University, Canada
College of Business
119 South Street Big Rapids, MI 49307-2284, USA
Thelephone 231-591-2479
Fax 231-591-3521
Email **Mohammed_El-Saidi@ferris.edu**

In this paper we follow a pattern of development of the scheme of classical series of Bernoulli trials subject to constant evolution. Related reasons, construction of random variables, probability distributions, processes, models, results and modeling power are briefly discussed. Areas of application are selectively presented, and directed mostly to environmental studies, risk analysis and some related technical problems. The pattern of this development, growth of new ideas, new results and directions of further research studies and applications are the main goals of this work.

Managing CO_2 emissions in the chemical industry

Vincenzo G. Dovì, Marina S. Fumagalli
Dipartimento di Ingegneria Chimica e di Processo "G.B. Bonino"
Università di Genova, Via Opera Pia 15
16145 Genova, Italy
Telephone: +39 010 3532921
Fax: +39 010 3532586
Email: **dovi@istic.unige.it**

The compliance with the targets of CO_2 emission reductions, foreseen by the Kyoto accord, requires significant contributions from the three most notable CO_2 producers: power stations, transports and process industry.
The management of CO_2 emissions in the chemical industry is the goal of this note.
Targets are country based, whereas the plants (products) of a Company may be scattered (marketed) worldwide. This gives rise to a complex overall strategy, which may include both technical and commercial measures for the attainment of the overall targets.

Modelling Contaminant Age Relationship in Fish Tissue in the Great Lakes

Abdel H. El-Shaarawi
National Water Research Institute and Department of Statistics University of British Columbia, Canada

The Canadian and US governments have advisories in place to limit the consumption of the Great Lakes fish. Most advisories use fish length or age as a surrogate variable for the concentration of the contaminants in fish. The approach commonly employed is to divide the range of the surrogate variable into a set of successive intervals such that some intervals are identified as no consumption categories (high risk); others as restricted consumption (limited risk) and the remaining intervals as risk free. It is important then to model the relationship between the surrogate variable and the contaminants concentrations. The model allows the estimation of the: 1) concentration of the contaminants for a specified value of the surrogate variable and its precision, and 2) probability that the concentration of the contaminants falling in a certain interval as a function of the surrogate variable.

The aim of this paper is to present a model for the concentration of polychlorinated biphenyl (PCB) in lake trout from Lakes Superior and Ontario waters using age as a surrogate variable. The model highlights the contrasts and similarities between the evolving patterns of the PCB accumulation in whole fish tissue from the two lakes. Classical and bootstrapping techniques are implemented to make inferences about various quantities of interest.

Pointing in the Right Direction: Using Indicators of System Behaviour to Assess Sustainability

Louise Erasmus[1] Albert S. van Jaarsveld[2]
[1]Centre for Environmental Studies, [2]Conservation Planning Unit
Department of Zoology and Entomology
University of Pretoria Pretoria, 0002 South Africa
Telephone +27 12 420 4048
Fax +27 12 420 3210
Email **lerasmus@zoology.up.ac.za**

Indicators are routinely used to measure and assess system sustainability. Although such indicators are useful for measuring the status quo of a system, they lack in predictive ability of system sustainability over long time frames. This latter ability is a requirement, per definition, for sustainability assessment. An ability to address changes in system sustainability over time requires the continuous reassessment of system indicators. The combination of these state indicators with indicators of predicted system behaviours could enhance the comprehensive utility of indicators for assessing system sustainability over time. In order to demonstrate a simple method for deriving indicators of system dynamics, this study employed historical environmental and developmental data of selected countries. A two-dimensional correlation plot was drawn using axes corresponding to 1) environmental impact and 2) socio-economic development. Coordinates on these two axes were used as measures of the status quo of environmental impact and socio-economic development respectively for each country. Vectors describing historical trends in environmental impact and socio-economic development were subsequently derived for each country. These vectors describe historical trends in terms of the rate and direction of change in environmental impact and socio-economic development, and serve as predictors of probable future system behaviours. The combination of such indicators of system behaviour with state indicators can be used to construct more comprehensive sustainability assessments.

The Perspective of Quantitative Science in the Debate about Environmental Degradation

Panel Session
Sylvia R. Esterby (Organizer)
Okanagan University College, Canada

Panelists: N. Cressie[1], V. Dovì[2], P. Guttorp[3], T. Jakeman[4], P. Sullivan[5]
[1]The Ohio State University, Columbus, OH, USA; [2]Dipartimento di Ingegneria Chimica e Di Processo, Università di Genova, Italy; [3]NRCSE, University of Washington, USA; [4]Centre of Resource and Environmental Studies, Australia; [5]Department of Applied Mathematics, University of Western Ontario, Canada

It may seem logical to assume that, if scientists obtain defensible estimates of the current status of environmental quality and make defensible predictions about the future degradation of environmental quality, decision makers will use this information in the formulation of environmental policy that will protect the environment from further degradation. Of course this is over simplistic, as evident, for example from the current situation regarding reduction of emissions under the Kyoto protocol. This being so, what position should environmental scientists take in these complex and political questions of environmental protection? This question raises key issues: defensibility, inclusion and visibility of environmetricians, collaborative and integrative research, successes and failures, and differences between environmental compartments and levels of policy making.

The provision of defensible estimates and predictions is extremely important so that the case for environmental protection is not hurt by apparent contradictions. This involves the exercise of great diligence in view of the prevalent situation of incomplete information and makes it natural to think that environmetricians should take the leadership in the discussion about the level of uncertainty in the scientific results reported. In turn, the question arises as to how environmetricians become part of the information gathering team and ensure that they are heard in the discussions. The diligence referred to above may imply that investigations are conducted in a collaborative manner, bringing together scientists and methodologies from the pertinent disciplines. Each environmental protection issue, *ie.* the environmental compartment and the level at which policy is being made, will bring some different considerations and possibly different chances of successfully impacting decisions.

The above comments are intended to provide some starting points for the discussion. The varied and wide experience of the panellists, together with the anticipated floor discussion, should bring us more insight into how we can have greater impact upon environmental decision-making.

Environmental and Accident Risk connected to Road Tunnel Fires: Experimental and theoretical investigation

B. Fabiano, **F. Currò**, E. Palazzi, R. Pastorino
DIChePChemical and Process Engineering Department "G.B. Bonino;,
University of Genoa, via Opera Pia 15,
16145 Genova, Italy
Telephone 103532917
Fax 103532586
Email **chempla@unige.it**

Road traffic and especially heavy goods traffic has continually increased over many years but, notwithstanding the existence of broad national and international regulations on the topic, recent accidents involving tunnels show that further research is still needed to assess environmental and accident risk. On the other side, due to the sharp reduction of vehicle pollutant emissions, the ventilation of road tunnel is more and more determined by the need of controlling smoke in case of fire. Even if considerable data are available in the literature on tunnel fires, a closer examination reveals that the majority of these data are not of adequate quality or detail to allow effective model development and validation.

On these basis, the first step of this research was carried out on a real scale, utilizing a highway tunnel to simulate the evolving scenarios following an accident with fire development. The aim was to investigate the quality of the tunnel environment both up- and downstream of a severe fire. The experimental runs were carried out utilizing different combustion materials and cars. In particular, following aspects were studied in depth: longitudinal and transverse temperature gradients in presence and in absence of fire-extinguishing automatic systems; evolution of the heuristic pollutant during the fire.

In the second phase of the research, a laboratory scale tunnel was designed and realized "ad hoc", based on a critical evaluation of the theory of similarity applied to fire events. Several experiments were performed, considering three ventilation techniques and the most important modelling parameters of ventilation rate and heat-release rate. The focus was on the modelling of the phenomenon and the verification of the effectiveness of preventive and protective measures to reduce both environmental and accident risk.

Combining the results of full scale and laboratory tests, practical recommendations are drawn concerning, from one side, general measures relating to the tunnel and the traffic (e.g. ventilation design and operation), on the other side specific measures to reduce the consequence of accidents and implement emergency response management. Fundamental understanding of the combination of different phenomena affecting the air flow, such as piston effect of vehicles; chimney effect in presence of tunnel slope; net difference between injected and exhausted flows etc. and the validity of extrapolating small-scale results to larger scales are finally discussed.

Nonlinear Statistical Sensitivity Analysis of Environmental Models with Application on Heavy Metals Adsorption from Contaminated Wastewater

Alessandro Fassò[1], Emilio Porcu[1], Antonio Esposito[2], A.P. Reverberi[2], Francesco Vegliò[2]
[1]Department of Engineering, University of Bergamo, [2]Department of Chemical and Process Engineering "G.B. Bonino", University of Genova
Via Marconi 5, 24044 Dalmine BG, Italy
Telephone +39 338 8026616
Fax +39 035 562 779

In this paper we discuss some linear and nonlinear heteroskedastic models for global sensitivity analysis. We than introduce a simple mathematical model (LEM - Local Equilibrium Model) describing heavy metal adsorption processes from industrial waste waters. To do this, fixed bed column has been considered as continuous contacting configuration with low cost adsorption material. In particular, this unit operation can be described by a set of PDE equations. Using a first large Monte Carlo experiment, we then carried out a global sensitivity analysis of the model parameters to get insight into the physical model. Moreover, using a second large Monte Carlo experiment, a local sensitivity analysis has been carried out to determine the uncertainty of the output system using an experimental error associate with the model variable values.

Two-stage estimation of coverages with second-phase corrections

Lorenzo Fattorini, Marzia Marcheselli, **Caterina Pisani**
Dipartimento di Metodi Quantitativi Università di Siena
P.zza S.Francesco 8
53100 Siena, Italy
Telephone 0577/232750
Fax 0577/232626
Email **pisani4@unisi.it**

A two-stage estimation of coverages of k land categories partitioning a study area is considered. In the first stage a sample of N points is selected according to a spatial sampling and the selected points are assigned to the k categories on the basis of satellite imagery or various types of aerial photos. Simple random sampling, stratified sampling and sistematic sampling of points are considered. The theoretical results suggest the use of stratified sampling as the best design-based strategy at first stage. Since misclassifications may take rise at this stage, then, in the second stage, a sample of points is selected to be visited and correctly classified on the ground. A stratified sampling of the N points is considered in which strata are determined by the aerial classification. The statistical properties of the resulting estimator are derived and a conservative estimator of its variance-covariance matrix is proposed. The procedure is compared with a previously proposed procedure on the basis of a simulation study performed on some artificial populations.

Environmental Assessment Based on Multiple Indicators[⊕]

Jerzy A. Filar[1], N. Phillip Ross[2], Ming-Lu Wu[*3]
[1]CIAM University of South Australia, [2]US EPA, [3]City University of Hong Kong

ABSTRACT: Environmental assessment is key to any environmental protection policy and is now among the highest priorities in socioeconomic development around the world. The present paper focuses on the application of multiple criteria decision making (MCDM) techniques to assessing the effectiveness of US environmental policies for the reduction of toxic releases. We first introduce the US toxic releases problem and the *Toxic Releases Inventory* (TRI) data base. Then, we describe in detail an MCDM approach called the *technique for order preference by similarity to ideal solution* (TOPSIS) that is quite intuitive and easy to understand and use. We address the question of how to assign suitable weights to the environmental indicators to reflect their relative importance in the assessment process. We use entropy, the degree of variation of each indicator, as the base to determine their importance weights. Next, we apply the MCDM technique to assess the state and movement of US toxic releases of priority chemicals constituting the "33/50 program". Finally we present some conclusions and suggestions for future research in this area, including the formation of better environmental indicators and the uses of other assessment techniques.
Key Words: Environmental Assessment; Multiple Criteria Decision Making; Importance Weights

⊕ This study has been supported by a Research Assistance Grant from the US EPA. The help of the staff at Center for Environmental Statistics in obtaining the data used in this study is hereby gratefully acknowledged.

* Dr. Wu was a visiting scholar at the School of Mathematics, University of South Australia between October and November of 1997 when this analysis was done.

Air emission inventories, needs for relevant statistics

Jean-Pierre Fontelle
CITEPA
10, rue du Faubourg Poissonnière
75010 – Paris, France
Telephone +33 1 44 83 68 83
Fax +33 1 40 22 04 83
Email **jean-pierre.fontelle@citepa.org**

Nowadays, Air Pollution is not only a topic of interest for scientists or criteria of luxury for developed societies. It is more and more an integrated component of the current life. The presentation will focus especially on the knowledge of substances emitted into the atmosphere and the link with statistics. The estimations of emissions has to be considered in various frames from local to international levels and deals as well with scientists, local policy as national and international ones : e.g. international conventions of the United Nations and their protocols such as those of Göthenburg (Long Range Transbundary Air Pollution) and Kyoto Greenhouse Gas Effect) . Moreover, the implementation of EU directives (Integrated Pollution and Prevention Control and National Emission Ceilings, calls for similar requests. The presentation will explicit the needs for relevant socio-economic statistics which are absolutely necessary with regard to consistency, transparency, confidentiality, timeliness and accuracy for the preparation of air emission inventories. It will also demonstrate the impact of these requests on international commitments.

Properties of quartile estimators for rainfall enhancement experiments

Tressa L. Fowler
National Center for Atmospheric Research
P.O. Box 3000 Boulder,
CO 80307-3000, USA
Telephone (303)497-8373
Email **tressa@ucar.edu**

Quartile estimates have been used in both the South Africa and Coahuila, Mexico cloud seeding experiments to compare the distributions of the seeded and unseeded thunderstorms. Quartiles were chosen as the estimator of choice over the mean for their robustness to departures from normality. However, quartile estimators are really a special case of percentile estimators, which are not unique. Several methods of computing percentiles from a sample are in use. Different software packages use different methods and sometimes a single software package employs different methods in different functions. Many of them fail to document or inform the user of the choice of method. All methods are asymptotically equivalent. However, on small and/or skewed samples of data, the choice of method can make a large difference in the resulting estimate of the quartile. Cloud seeding samples are both small and skewed. For these types of data, the estimate of the quartile is NOT robust to the choice of method. If cloud seeding experiments are going to continue to employ quartile estimators, the estimator needs to be determined a priori. Otherwise, the estimator that produces the best results post facto may be chosen. Also, published results may not be reproducible by independent evaluators who are unaware of the issues surrounding quartile estimators. An empirical study is undertaken to examine the properties of the different estimators on small, skewed samples of data, such as those resulting from cloud seeding experiments. These properties are detailed and recommendations for future experiments are made. Additionally, permutation tests are used to determine p-values for the differences in the quartiles. The effect of the choice of estimator on the resulting p-value is also assessed.

Statistical assessment of air quality numerical models

Montserrat Fuentes
North Carolina State University
Campus Box 8203 NCSU Department of Statistics North Carolina State University
Raleigh, NC 27695
Telephone 919 515 1921
Fax 919 515 1169
Email **fuentes@stat.ncsu.edu**

Evaluation of physical models for air quality applications is crucial to assist in control strategy selection. The high risk of getting the wrong control strategy has costly economic and social disbenefits, which will be increasingly important for the proposed new multi-pollutant standards and combined controls strategies. The objective comparison of modeled concentrations with observed field data provides a means for assessing model performance. However, for dry deposition fluxes and concentrations of air pollutants there is a very limited supply of evaluation data sets. Here we develop a formal method for evaluation of numerical models, this method can be successfully implemented even when the field measurements are taken at a very sparse set of monitoring stations. We specify a simple model for both, the output of the air quality models and the observations in the ground, in terms of the unobserved ground truth, and estimate the model in a Bayesian way. This yields solutions to the model evaluation and bias removal problems simultaneously. It provides improved evaluation of air quality numerical models via the posterior predictive distribution of the ground observations, and enables us to remove the bias in the numerical models output by estimating additive and multiplicative bias parameters in the model. We apply our methods to data on SO_2 concentrations.

Fractional Vector Autoregressive Models for Environmental Data

Mariagrazia Granturco, **Eugenia Nissi**
Università D'Annunzio di Chieti
65127 Pescara, Italy
Telephone 085/68496
Fax 085/4537542
Email **nissi@dmqte.unich.it**

Some pollutants highlight a tendency to persist in the atmosphere for long time, so their temporal pattern could be explained by fractional ARIMA models for strongly dependent data; moreover, considering their strong correlation, a vector autoregressive model for multiple time series could be suitable. The aim of this paper is to propose the use of Fractional Vector Autoregressive models in order to jointly model principal pollutants observed in urban areas. In order to emphasize the advantages of this class of model as regard to classical Vector Autoregressive Models, an application to real data sets is proposed.

Incorporation of concomitant information and deterministic models into statistical trend tests

Anders Grimvall
Department of Mathematics, Linköping University,
SE-58183 Linköping, Sweden
Telephone +46 13 281482 (281455)
Fax +46 13 100746
Email angri@mail.mai.liu.se

Detection of trends in environmental data was long regarded as a matter of separating seemingly persistent changes from more rapid fluctuations in univariate time series of data. The low frequency components were attributed to anthropogenic activities, whereas the remaining components were assumed to represent natural fluctuations. During the past decade there has been a significant increase in the number of articles where different types of concomitant information is incorporated into statistical tools for trend analysis. In this review paper we show how such information can be incorporated into non-parametric, semi-parametric and parametric procedures. The non-parametric techniques, which include different types of Mann-Kendall tests, are normally focused on the statistical significance of observed changes in the state of the environment. The semi-parametric and parametric procedures, on the other hand, are frequently oriented towards signal extraction or variance reduction that can facilitate the interpretation of collected environmental data without providing a rigorous statistical test. In the latter category, we can also find several procedures in which process-oriented deterministic models are employed to facilitate the detection and attribution of environmental change. The research on climate change can represent an area in which the process-oriented models are used directly to separate anthropogenic signals from natural fluctuations and to explore what kind of changes that might be detected in measured data. Air and water quality studies represent areas in which it may be more advantageous to first derive reduced forms or meta-models of the process-oriented models, and then incorporate such simplified models in procedures for trend detection.

Exploration of trends in atmospheric temperatures

Howard Grubb
School of Applied Statistics, The University of Reading
PO Box 240, READING, RG6 6FN, U.K.
Telephone (0118) 931 6573
Fax (0118) 975 3169
Email H.J.Grubb@reading.ac.uk

Trends in climatic time series often form the starting point for studies into global change. Long temperature time series are used here to illustrate ideas surrounding the exploration and preliminary
estimation of trends. These trends might be in aggregate or component series, as well as in derived measures of these series, such as their extreme behaviour and other functionals.

Ecosystem Management via Interacting Models of Political and Ecological Processes

Timothy C. Haas
School of Business Administration University of Wisconsin at Milwaukee
P. O. Box 742 Milwaukee, WI 53201, U.S.A.
Telephone 414 229 4360
Fax 414 229 6957
Email **haas@uwm.edu**

The decision to implement environmental protection options is a political one. Depending on the political mechanism operating, a country may or may not heed the most persuasive scientific analysis of an ecosystem's future health. A predictive understanding of the political processes that result in ecosystem management decisions may help guide the formulation of ecosystem management policy. To this end, a web-based management tool is being developed that is supported by a stochastic, temporal model of how political processes influence and are influenced by ecosystem processes. This model consists of temporally interacting influence diagrams that represent the political mechanisms that govern the decision making of a country's president, environmental protection agency, legislature, and rural inhabitants. These within-country decision making influence diagrams interact with influence diagrams of international environmental protection organizations and a spatio-temporal influence diagram of the ecosystem that the country is affecting. As an example, a set of interacting influence diagrams is used to represent both the decisions made to manage the cheetah population in the countries of Kenya, Tanzania, and Uganda -- and the population dynamics of cheetah living in the range enclosed by these countries (the Serengeti). This model is used to simulate a sequence of ecosystem management decisions and to compute the probability of cheetah extinction at the end of this decision sequence. A method for statistically estimating all parameters in this model is described that optimally combines data on some of the model's outputs with expert-derived judgements of parameter values. This estimation method is non-Bayesian and uses the expert-derived values to impose constraints on the solution so that estimates can always be computed no matter what the ration is of number of parameters to number of observations.

Density Estimate and its Application to Analysis of Temperature series

Ivana Horova, Jiri Zelinka, Rudolf Brazdil, Marie Budikova
Masaryk University,Brno,Czech Republic
Department of Applied Mathematics
Faculty of Science Janackovo
nam.2a 662 95 Brno, Czech Republic
Telephone +420 5 41 32 12 51
Fax +420 5 41 21 03 37
Email **horova@math.muni.cz**

Nonparametric density estimates attempt to reconstruct the probability density from which a random sample has come, using the sample values and as few assumptions as possible about the density. These methods are smoothing operations on the sample distribution. Methods of kernel estimates represent one of the most effective nonparametric methods. These methods are simply enough which makes the numerical calculation easy and fast and also the possibilities for mathematical analysis of properties of obtained estimates are very good. We employed the automatic procedure for the selection of the bandwidth, the kernel and the order of the kernel. The procedure is used for analysis air temperature fluctuation for series of Central England and Prague- Klementinum in the periods 1661-2000 and 1771-2000,respectively.

Spatial covariance modelling in a complex coastal domain by multidimensional scaling

Gudmund Høst, Anders Loland
Norwegian Computing Center
P.O. Box 114,
Blindern N-0314 Oslo, Norway
Telephone (+47) 22 85 25 00
Fax (+47) 22 69 76 60
Email **gudmund.host@nr.no**

In aquatic studies, spatial interactions may be both easier to interpret and to quantify by using water distance than by using geographic distance. The water distance is the shortest path between those two sites that may be traversed entirely over water. One problem is that water distances may be non-Euclidean, and thus covariance and variogram functions are not necessarily valid when using the water distance as distance metric. Another problem is that the computation of water distances for a large set of spatial locations is computationally expensive. Our alternative is a computationally efficient method for calculation of a Euclidean approximation to water distances. The first step of the method is to define a triangular grid covering the complex domain in interest. Using this triangular grid, we pre-compute approximate water distances using a graph search algorithm. These water distances are then approximated by multidimensional scaling, giving a Euclidean space. Finally, we use linear interpolation to move the data locations into the new Euclidean space. By using this method, subsequent computations of water distances between any locations can be done very fast and the method leads to a theoretically valid spatial covariance model. We apply our method on herring data from the Vestfjord system in Northern Norway. The application suggests that the geometry can be represented accurately by a two-dimensional deformed space.

Estimation of the anthropogenic impact on nutrient loads carried by the Elbe River

Mohammed Hussian[1], Anders Grimvall [1], W. Petersen [2]
[1]Department of Mathematics, Linköping University, Sweden, [2]Institute of Hydrophysics, GKSS Research, Germany
Telephone +46 13 28 1496
Fax +46 13 100746
Email **mohus@mai.liu.se**

The political and economic reforms in Eastern Europe that culminated in the early 1990s implied dramatic changes in the pressure on the environment. In the Elbe River Basin, the reunification of Germany was accompanied by a substantial decrease in direct emissions to water. Furthermore, the nitrogen surplus in the agriculture, i.e. the difference between the amount of nitrogen added to the topsoil and the amount removed through harvesting, dropped to less than half of its previous level in 1990. We examined how the cited interventions in the drainage area influenced the load of nitrogen and phosphorus carried by the Elbe River to the North Sea. In particular, we tried to clarify the anthropogenic impact on the loads of nitrogen and phosphorus by removing the fluctuations that could be attributed to temporal variation in water discharge. The results of our study showed that, after flow-normalisation, the loads of total nitrogen and NO3-N at sampling sites upstream and downstream of Hamburg exhibited a smooth downward trend that started around 1990 and has persisted since then. The decrease in NH4-N was more rapid and completed in the mid 1990s, whereas the anthropogenic signals in the loads of total phosphorus and orthophosphate varied between the sampling sites and were remarkably weak at the sampling sites located downstream of Hamburg. The statistical procedures we employed involved locally weighted regression to expand the observed series of concentration data to complete series of daily values and a semiparametric regression technique to model the time-varying relationships between nutrient loads and water discharge that were used for flow-normalisation.

Some nonparametric tests in meteorological time series

Jana Jureckova[1], Jan Picek[2]
[1]Charles University in Prague, [2]Technical University in Liberec, Czech Republic
Sokolovska 83,
CZ-186 75 Prague 8, Czech Republic
Thelephone +420 2 2191 3285
Fax + 420 2 22323316
Email **jurecko@karlin.mff.cuni.cz**

A general class of tests on the parameters of autoregressive models, based on new concept of autoregression rank scores was constructed by Hallin and Jureckova (1999). Among these tests, the test on the order of autoregression was applied to climatic data in Hallin et al. (1997) and Kalvova et al. (2000) and the test of independence of two autoregressive time series was applied to the temperature and mortality data in Hallin et al. (1999), among others. Some other tests will be also mentioned. [1]Hallin,M.and Jureckova,J.(1999). Optimal tests for autoregressive models based on autoregression rank scores. Ann. Stat.27, 1385-1414. [2]Hallin,M.,Jureckova,J.,Kalvova,J.,Picek,J. and T.Zahaf (1997). Non-parametric tests in AR models with applications to climatic data. Environmetrics 8, 651-660. [3]Hallin,M.,Jureckova,J.,Picek,J.and T.Zahaf (1999).Nonparametric tests of independence of two autoregressive time series based on autoregression rank scores. J.Statist.Plan.Infer.75,319-330. [4]Kalvova,J.,Jureckova,J.,Picek,J.and I.Nemesova (2000). On the order of autoregressive (AR)model in temperature series. Meteorologicky casopis 3,19-23.

Evidence for climate variables having heavy-tailed distributions

Richard W. Katz
National Center for Atmospheric Research
NCAR, BOX 3000,
BOULDER, CO 80307, USA
Telephone 303-497-8114
Fax 303-497-8125
Email **rwk@ucar.edu**

The question of whether the distributions of climate variables have "heavy" (i.e., Pareto-like) tails is important, both for estimation of the likelihood of extreme events and for detection of trends in extremes corresponding to climate change. In the present talk, the methodology of the statistics of extremes is employed to detect and model heavy-tailed distributions. In particular, the Poisson-generalized Pareto model (or "peaks over threshold" approach) is applied to daily climate time series, with the exceedances of a high threshold being modeled as a Poisson process and the excesses over this threshold by the generalized Pareto distribution. One example is considered in detail, involving a 100-yr time series of daily precipitation amount at Fort Collins, CO, USA. This data set has been of recent interest, because of a flood that occurred on 28 July 1997. With such a long record, the evidence for a heavy-tailed distribution turns out to be relatively strong. Annual cycles in the precipitation parameters are explicitly modeled, reducing the possibility that this heavy tail is indirectly induced in annual maxima.

Peoples' willingness to pay for environmental conservation in Tabriz, Iran: a preview of social and economic determinants

Ali Mohammad Khorshiddoust
Physical Geography Department, Faculty of Humanities and Social Sciences, Tabriz University, Tabriz 51664, Iran
Telephone 0098411 3805815
Fax 0098411 3356013
Email **khorshid@tarbizu.ac.ir**, and **environ111m@yahoo.com**

The main objective of this research was to study different social, and economic factors affecting peoples' willingness to pay (WTP) and their attitudes toward the environmental conservation in Tabriz, a metropolitan city being located on north west of the country. The research carried out, was consisted of two main parts, namely, the examination of existing environmental, social, and economic conditions of the study area, and the analysis of survey and interview results. The contingent valuation method (CVM) was applied in order to estimate the amount of WTP. The results showed immense amounts of WTP and found some interesting factors determining peoples' participation in environmental conservation.

Combining hyperspectral data and ground surveys for the assessment of non-wood goods and services of forests

Michael Köhl, **Wolfgang Stümer**, Bernhard Kenter
Dresden University of Technology
Chair of Forest Biometrics and Computer Sciences
Wilsdrufferstr. 18
D-01737 Tharandt, Germany
Telephone: +49 35203 3816-13;
Fax: +49 35203 3816-28
Email: koehl@forst.tu-dresden.de

Forest resources assessments are in a dilemma: on one hand survey budgets are cut; on the other hand information needs, especially on non-wood goods and services (NWGS), are increasing. Thus, new methodological approaches have to be developed that decrease the survey cost by simultaneously increasing the information provided. From a statistical point of view, it is well known that the utilisation of auxiliary data such as remote sensing data can increase cost-efficiency of surveys. From a practical point of view, there is, however, the contradiction that remote sensing data provide a rough classification of different forest patterns while the assessment of NWGS such as biological diversity or protective function of forests requires intensive field-based assessments. This paper describes methods developed for the assessment of NWGS by combination of high resolution hyperspectral data and ground surveys and reviews the results in terms of information needs.

Two different kind of results have to be provided by forest resource assessments: (1) statistics (e.g. total values, mean values or proportions) and (2) mapped information. The increase of cost-efficiency for providing statistical key-parameters was realised by the application of a two-phase sampling design. Hyperspectral remote sensing data are used for stratification purposes and combined with intensive ground surveys. Mapped output is realised by relating terrestrial samples to the spectral information of pixels via the "k nearest neighbour" (kNN) -method. For the entire set of pixels without associated ground assessments the k nearest neighbours in the spectral image space are determined among those pixels which coincide with the location of field samples. For the k nearest neighbours the values of attributes assessed on the ground are weighted by the distance in the spectral image space and assigned to the respective pixels for which no terrestrial information is available. The kNN-procedure results in maps that show the spatial distribution of attributes assessed on the ground in the resolution of the remote sensing data. Both, the two-phase sampling design and the kNN-method were adapted for the handling of data on nominal and ordinal scales.

Non homogeneous Hidden Markov Models for categorical space-time series and an application to pollution data from the monitoring network in Rome

Francesco Lagona
Università di Roma Tre
Dept. of Social Sciences University of Roma Tre
via Corrado Segre 4, 00146 ROME, ITALY
Thelephone 39-6-5517-6400
Fax 39-6-5517-6248
Email **lagona@uniroma3.it**

Categorical series of pollutant concentrations arise when data are discretized in a number of classes according to threshold values of interest (e.g., limit values mandated by institutional directives). If the aim is to provide exposure in a metropolitan area of interest, these data are often in the form of multivariate datasets providing measurements of pollutants and atmospheric variables taken hourly at multiple stations. When modeling categorical, multivariate, space-time pollutants series, the typical issue is to estimate a parsimonious model, still capable to capture the dependencies across time, space and pollutants. The focus of the talk is on the use of Hidden Markov Chain (HMC) models for multivariate categorical data. The basic assumption underlining HMC models is that the distribution of the observed data at each time depends on the (hidden) state of a latent Markov chain and that the conditional distribution of the observed data is independent of past data, given the current state. In this talk, a multivariate random field is assumed for the observation process and the transition probabilities of the chain are modeled as functions of atmospheric variables (hence the chain is non homogeneous). Under this framework we can model different hypotheses of conditional independencies across space, time and pollutants. Parameters are estimated through an E-M algorithm, while a Viterbi algorithm is implemented to reconstruct the state sequence. The procedure can be easily generalized to include (pollutant) missing value estimation. The model is fitted to the space-time data series of two pollutants measured hourly at three monitoring stations in Rome over a two-year period.

Nonparametric statistics for deterministic dynamical models

Salim Lardjane
Université de Bretagne-Sud, Applied Statistics Laboratory
Iup Gsi Lorient 2 Rue Coat-St-Haouen
56100 Lorient, France
Telephone 33297880543
Fax 33297880551
Email **salim.lardjane@univ-ubs.fr**

In this paper, we show how nonparametric statistical methods can be used rigorously to investigate the properties of dynamical models encountered in ecology and population dynamics. The applications are based on theoretical results about the asymptotic properties of nonparametric statistical procedures in the case of deterministic dynamical systems and non i.i.d. random variables, and are illustrated by a set of simulations.

Cluster Modelling of Small Area Health Data in Environmental Risk Assessment

Andrew Lawson
University of Aberdeen,UK
37 Errol Road Invergowrie Dundee
UK DD2 5BX
Telephone 44-1382-562440
Email a.lawson@maths.abdn.ac.uk

In this presentation the basic foci of small health data analysis are reviewed and their interrelation with environmental risk assessment is considered. In particular the area of cluster detection is one where there are strong links with environmental issues, specifically environmental epidemiology. Examples of the analysis of the spatial distribution of relative risk in the vicinity of putative health hazards such as incinerators or landfill sites, are many. Here I will focus on some newer developments in the area of cluster analysis where the prior focus (e.g. a putative source|) is not known. The approach described examines the use of Bayesian pseudolikelihood models for clustering. Unlike some recent proposals for the analysis of clustering, these models do not require the use of birth-death reversible jump MCMC and so are easier to interpret. These models are very flexible and have close links to nonparametric approaches to smoothing risks. I will present some results from simulation studies and also a data example will be examined.

Opening Address

Roberto Levaggi
Ministry of the Environment, Regione Liguria, Italy

The knowledge of environmental conditions, the forecast of effects of human actions on the different components of the environment, the comprehension of the process through which these same actions exert their influence on the quality of air, water, soil, on ecosystems and on climate, must be, and really they are, an absolutely necessary reference to fix the strategies in defense of the environment, and consequently, of the environmental policy.
Nowadays this is true (and it will be more and more in the future) for the territorial levels this policy is referred to, from the world level to the local one, passing through the European and national levels.
The importance of the "measure" of environmental parameters in the field of laws and programmatic acts of European Union is obvious.
With the term "measure" of environmental parameters we mean not only the improvement of the environmental quality through direct measures, but also the more and more frequent use of the techniques of modelling and of the methods of objective evaluation.
We can think, for example, to the set of rules about the "evaluation and quality management system of air". It provides that the direct measure of the quality of air must be supported by the techniques of modelling, in order to foresee the efficiency of the interventions through which we can obtain the respect of the limits.
The recent signature of the Kyoto protocol by the European countries makes us understand how important and necessary is to reconcile environmental policies with other policies, that are important for the socio-economic development.
The European laws and the consequent national laws point out the problem of proving through scientific methods the effects of the single choices in the field of energy, of transport, of building and so on.
If we really want to contribute to the achievement of the purposes of Kyoto, we must "measure" the strategies and actions we must follow in the field of energy, transport, building.
For what I have said so far and for the issues that will be dealt with in the next days, I wish to all of you a successful conference and hope that the improvement of the know how can help us to make better choices to defend the environment.

A Model for Structuring the Quantitative Assessment of Public Health and Environmental Risks of Toxic Substances

Plenary Invited Lecture
Saul M. Lemkowitz[1], **Hans J. Pasman[2]**
[1]Faculty of Applied Sciences, Delft University of Technology, the Netherlands, [2]Defence Research, Netherlands Organisation for Applied Research, Delft, the Netherlands"

Abstract: Modern industrial societies produce, consume, and distribute huge amounts of a great number of chemicals, many of which are highly toxic if exposed to in too high concentration. Examples include toxic substances produced intentionally, such as hydrogen fluoride and phosgene, which are later employed as inputs to create useful products like lead-free gasoline, plastics and pharmaceuticals. Other examples are toxic substances unintentionally produced, such as dioxins, which are formed as by-products from combustion processes and are inadvertently dispersed into the environment. It is essential for societies to be able to quantitatively assess toxic risks of substances. Such assessment is necessary in order to establish priorities in risk reduction policy and to check whether such policy is working properly. Quantitatively assessing toxic risks to society is, however, an extremely complex process, and consequently much confusion still exists about how to
execute it properly. We have developed a model that structures the entire process of quantitative risk assessment of toxic substances. The utility of the model will be illustrated by applying it to real examples. These will include dioxins, a serious public health risk, and acidifying substances such as sulphur and nitrogen oxides, which generate important risks to the environment. It should be stressed that various forms of the model are being successfully applied in the Netherlands for quantitative assessment of toxic risks.

Structural model for the entire process of quantitative assessment of public health and environmental risks from toxic substances.

Variance reduction for trend analysis of hydrochemical data in brackish waters

Claudia Libiseller, Anders Nordgaard
Dept. of Mathematics/ Div. of Statistics, Linköping University
SE-58183 Linköping, Sweden
Telephone +46- 13-281499
Fax +46-13-100746
Email **cllib@mai.liu.se**

The concentrations of nutrients and other substances in a water body can be strongly influenced by random fluctuations in the mixing of waters of different origin. Hence, the water quality at given sites can exhibit a large temporal variation that makes it difficult to extract anthropogenic signals from collected data. In this paper, we examine how the human impact on nutrient concentrations in such water bodies can be clarified by replacing conventional time series or geostatistical approaches by trend detection techniques in which we analyse the variation in nutrient concentrations with salinity and time. To test the significance of extracted trends we propose a non-parametric and a parametric approach. The non-parametric trend test is based on a weighted sum over all Mann-Kendall statistics computed for different salinity levels. The parametric approach is derived from a regression model in which the concentration of the nutrient is regarded as dependent variable and time (months) as explanatory variable. Since the residuals obtained from this model are most likely serially dependent and non-normally distributed a bootstrap technique is applied to assess the parameters.

Are car-free days effective? A case study in Palermo using Dynamic Linear Models

Gianfranco Lovison, Daria Mendola
Dipartimento di Scienze Statistiche e Matematiche "S.Vianelli"
Università di Palermo, Italia
Email lovison@unipa.it, daria.mendola@tin.it

It is a common opinion that increasing vehicular traffic is one of the main causes of air pollution in urban areas. Yet, actual evaluation of the effectiveness of limitations to vehicular circulation is a complex process because of the presence of multiple interactions between meteorological variables (like temperature, solar radiation, strength and direction of wind, etc.) and monitoring network variables (like location of the monitoring stations, width of the streets, height of surrounding buildings, closeness of emission sources, etc.) In Italy, environmental regulations laws compel public administrators to take appropriate measures when pollutant concentrations in urban areas exceed the alarm thresholds. The emergency remedy consists of closing to traffic, partially or totally, the streets where increased levels of pollutants have been registered. Although this is a widespread practice in Italian cities, there are no systematic studies on the effectiveness of such measures of traffic limitation. In particular, little is known about what should be the optimal duration of these restrictions, what the real benefits are, and, if any are present, how long these
benefits last. Can statisticians help in answering these issues? Can statisticians support policy makers to set up priorities in risk reduction policies for air pollution control? Can air quality predictions be useful for adopting appropriate environmental management strategies in advance?
In this talk we study the effectiveness of vehicle circulation restrictions on the reduction of air pollution. We do it by means of models in the VARMAX class (Vectorial AutoRegressive Moving Average models with exogenous variables). Specifically, we estimate a VARX model with p=4 response variables (daily concentrations of SO2, NO2, CO, Pm10) and r=3 explanatory variables (humidity percentage, temperature and days of closing to traffic), and use it to evaluate the impact of traffic limitations, within an intervention analysis approach. Owing to the high spatial variability of concentrations of air pollutants we focus in particular on a single monitoring station located in the center of the city of Palermo, in the middle of an area often interested by both high levels of
pollutants concentration and restrictions of vehicular traffic.

Indexing Environmental Quality Indicators: A Multiple Criteria Decision-making Approach

Ranjan Maitra[1], N. Phillip Ross[2], Bimal Sinha[3]
[1]Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD, USA. [2]United States Environmental Protection Agency, [3]UMBC.
1000 Hilltop Circle Baltimore, MD 21250, USA.
Telephone 410-455-2436
Fax 410-455-1066
Email **maitra@math.umbc.edu**

Environmental data quality indicators come from several sources and often represent different aspects of the environment. Combining these indicators to come up with an overall index is important for planning and policy and also to understand different environmental impact. We present here a data-driven multiple criteria decision-making (MCDM) approach. The methodology is implemented on data on toxic releases in the 50 U. S. states as well as on air, water and land quality in both the US and 106 countries of the world. Further analysis is done on homogeneous clusters of the U. S. air, water and land data.

Multivariate non-parametric tests of trend when the data are incomplete

Alvo Mayer, Juncheol Park
[1]University of Ottawa, Ottawa, Canada
585 King Edward Ave Ottawa,
Ontario Canada K1N 6N5
Telephone 613 562-5800 ext 3487
Fax 613 562-5776
Email **malvo@science.uottawa.ca**

In environmental and medical studies, multivariate data are often recorded over regular time intervals and examined for monotone increasing or decreasing trends in one or more of the variables. Dietz and Killeen (1981) proposed a non-parametric test based on the Kendall measure of correlation and applied it to medical data. In this paper, we are concerned with situations when the data are partially incomplete. New test statistics based on the Spearman and Kendall correlation coefficients are proposed which are shown to be asymptotically chi squared. Results from a limited simulation study reveal that in most situations, the proposed test statistic performs better than its counterpart which deletes the missing data.

Implementing Regional Scale Monitoring in the Pacific Northwest Region of the US

Dan McKenzie
US EPA
200 SW 35th St Corvallis, OR 97333, USA
Telephone 541-754-4625
Fax 541-754-4716
Email **mckenzie.dan@epamail.epa.gov**

Vast forest greeted the early non-native explorers and settlers as they traversed the Pacific Northwest, currently the states of Washington, Oregon and Northern California, west of the Cascade Mountains. This natural resource supported an ever-expanding logging industry, reaching a peak of 5 billion board ft harvest from Federal lands during the 1980s. Intensive silviculture on both public and private forests dramatically changed the forested landscape. In the early 1990s, the courts ruled on several lawsuits, primarily addressing the Northern Spotted Owl and its habitat, which challenged the federal land management practices, concluding that changes were required to meet environmental statues and laws. In 1994, the Northwest Forest Plan (NFP) was completed and the Secretaries for the Departments of Agriculture and Interior signed a Record of Decision that provided new standards and guidelines for ecosystem management on 24 million acres of public land. Six federal departments and agencies established an interagency and intergovernmental framework to achieve a coordinated implementation of the NFP. Court challenges to the NFP resulted in a ruling the plan valid, but cautioned that to remain valid, the monitoring had to be implemented. This presentation addresses three of the NFP monitoring components, the Northern Spotted Owl (NSO) and Aquatic Riparian Effectiveness Monitoring program (AREMP), and strategic surveys for rare and uncommon species. The NSO monitoring program is founded upon detailed demographic studies at eight representative sites. AREMP is evaluating a probabilistic survey approach to assess watershed condition. Information on rare and uncommon species is utilizing a systematic grid. The experience of implementing regional scale interagency monitoring program and the usefulness of monitoring results and lessons learned in each of these approaches will be presented.

Two methods for quickly and accurately constructing forest/non-forest classifications from satellite imagery and inventory data

Ronald E. McRoberts
North Central Research Station
U. S. Forest Service
1992 Folwell Avenue.
Saint Paul, Minnesota 55108 USA
Telephone: 651-649-5174;
Fax: 651-649-5285
Email: rmcroberts@fs.fed.us

Landsat Thematic Mapper satellite imagery and forest inventory plot data were used to construct forest/non-forest maps for a heavily forested and a sparsely forested area in Minnesota, USA. The satellite imagery was classified using a logistic regression model and a simple k-Nearest Neighbors technique to obtain percent forest cover maps. The maps were compared visually to a similar map obtained from a large-scale, national land cover classification effort. The maps were also compared with respect to their utility as a means of stratification for the purpose or reducing the variance of estimates of mean percent forest cover. Maps obtained with both classification methods were obtained quickly and easily, produced forest/non-forest classifications of inventory plots that were 80-90 % accurate, and reduced variances of stratified estimates by factors in excess of 3.

Environmental modeling using GLM - possibilities and limits

Jaroslav Michalek
Department of Applied Mathematics, MU Brno, Czech Republic
MU Brno, Janáckovo
nám. 2a 662 95 Brno,
Czech Republic
Telephone + 420 5 41321251
Fax + 420 5 41210337
Email **michalek@math.muni.cz**

The use of Generalized Linear Models (GLM) for statistical analysis of environmental data became very popular in the end of twentieth century when Nelder and Wedderbrun in 1972 established the theoretical background of GLM. Nowadays the GLM create a basic and very important part of many well known and famous statistical software packages and this way the statistical analysis based on GLM is available for many researchers. But still when using GLM one should respect several mathematical assumptions for getting serious results with real information from the data. To verify these assumptions is sometimes very difficult from many reasons often it is the small sample size. For the most typical example of such assumptions let us mention the problem of unknown distribution of response variable, choice of the unknown link function, selection of the most suitable predictors, variance is not a function of the mean and behaves independently of the mean. In the contribution there will be discussed what statistical conclusions can be expected if the assumptions for GLM based analysis are violated or are not correct. Presented conclusions will be illustrated on simulated data and on real environmental data.

A probability density function for concentration in relative dispersion of a contaminant plume in the atmosphere

R. J. Munro, P. C. Chatwin, **N. Mole**
Dept. of Applied Mathematics, University of Sheffield
Hicks Building,
Sheffield S3 7RH, U.K.
Telephone +44 114 2223772
Fax +44 114 2223739
Email **n.mole@sheffield.ac.uk**

Observations of the dispersion of a contaminant plume in the atmospheric boundary layer, obtained using a Lidar, are analysed in a (moving) coordinate frame relative to the instantaneous centre of mass of the plume. To improve the estimates of relative dispersion statistics, maximum entropy inversion is used to remove noise from the Lidar concentration profiles before carrying out the analysis. A parametric form is proposed for the probability density function (pdf) of concentration, consisting of a mixture of a beta distribution and of a generalised Pareto distribution. This pdf allows for the possibility of both unimodal and bimodal distributions, and is shown to give a satisfactory fit to observations from a range of positions relative to the source. The variation of the fitted parameters with crossplume location is analysed, and the maximum possible concentration is found to decrease away from the plume centre. The corresponding pdf in the absolute (fixed) frame of reference is obtained by convolving the relative frame pdf with the pdf of the position of the plume centre of mass. The results are compared with the absolute frame pdf obtained directly from the observations.

Comparative Patterns of Habitat Richness and Rarity for Pennsylvania Vertebrates

W. L. Myers[1], G. P. Patil[2], C. Taillie[2]
[1]Center for Remote Sensing of Earth Resources, Penn State University, USA; [2]Center for Statistical Ecology and Environmental Statistics, Penn State University, USA

Keywords**:** Biodiversity, Habitat, Spatial Pattern, Richness, Rarity, Vertebrates, Ecosystem Health.

The GAP Analysis biodiversity assessment for Pennsylvania has mapped potential habitats for all (breeding) vertebrate species at landscape scale. A regional habitat importance index (RHII) has been formulated to give emphasis for joint occurrence of species having rare and/or especially localized habitat. Vertebrate taxa have been segregated for analysis according to general life histories, including both aquatic and terrestrial forms. Comparative analysis of spatial patterns in terms of occurrence, consistency, and complexity are conducted for several vertebrate groups. Healthy ecosystems are indicated by coherent patterns of richness encompassing also multiple instances of rare habitats; whereas rare habitats in species-poor areas are indicative of ecosystem distress syndrome. Locally complex spatial patterns are often indicative of habitat fragmentation processes, but can be particularly informative in terms of important corridors of connectivity. Substantial inconsistency of indications can be observed, however, between ecologically divergent taxonomic groups. This may be especially evident between substantially terrestrial versus aquatically dependent groups, or between those needing grassland versus forest habitats.

Detecting aberrations in waterborne infections surveillance data

Elena N. Naumova[1] , Ian B. MacNeill[2]
[1] Tufts University School of Medicine, Boston, MA, [2] University of Western Ontario, London, Canada

Over the past 15 years many states in the USA have initiated surveillance to monitor emerging waterborne infections. Surveillance systems are thought to be useful for identifying cases of waterborne diseases, for detecting outbreaks, and for understanding the risk factors associated with disease incidence. To identify aberrations in systematically collected data, several statistical approaches have been applied.

In our previous study, we evaluated the temporal and spatial variations in cases of giardiasis and cryptosporidiosis, reported passively to the Massachusetts Department of Public Health for the period 1993-1996, and assessed the relationship of those variations to the source of drinking water. The analysis consisted of three parts: 1) the analysis of temporal variations; 2) the analysis of spatial variations by examining the impact of a water supply category on the disease incidence; and 3) the analysis of temporal variations in giardiasis reporting with respect to a water supply category. To evaluate the spatial variation, we linked each reported case to a created database indicating the source of residential water supply. During a suspected outbreak of cryptosporidiosis in Worcester in 1995, a significant increase in reported cases was also observed in the Boston area, with the average number of reported cases per week exceeding the pre-outbreak level by 50-fold. Following the suspected outbreak, weekly rates of reported giardiasis also increased in the Boston and Worcester metropolitan areas.

Using the above mentioned analysis as an example, this presentation aims to indicate potential problems, which may arise when standard techniques for detecting aberrations are applied to waterborne infections surveillance data. We will also indicate potential situations where well-established statistical techniques can be used to improve the review of surveillance data and quality of data interpretation. In particular, we will focus on situations that can be described as a change-point problem.

Multi-resolution analysis of the temporal evolution of polar stratospheric ozone and the detection of vortex breakpoint

Philippe Naveau, Philippe Keckhut, Alain Hauchecorne, Leah Goldfarb
Institut Pierre Simon Laplace, Service d'Aeronomie, France
Service d'Aeronomie - IPSL BP 3,
91371 Verrieres-le-Buisson CEDEX, France
Telephone +1 64 47 43 13 OR +1 69 08 77 24
Fax +1 1 69 20 29 99 OR +1 69 08 77 16
Email **naveau@aerov.jussieu.fr**

To understand and model the temporal variability of stratospheric ozone, it is necessary to accurately attribute the influence of air transport between poles and mid-latitudes polar on such ozone time series. To perform this task, statistical tools that take into account of the different time scales present in the ozone evolution are necessary. In addition, an automated detection procedure is needed to find the breaking point of the polar vortex that separates two different regimes driven by different physical processes. In this presentation, a statistical analysis based on wavelet decomposition is implemented to find characteristics in stratospheric polar ozone measured from different locations from 1980 to 1992. By decomposing the small temporal scales from large ones through our wavelet analysis, a more meaningful description of ozone variation is proposed. Our main focus is to estimate the impact of the polar vortex on such time series. First, an extraction method based on the multi-resolution analysis is presented to detect the annual timing of the vortex breakpoint. Second, a change-point statistical model is proposed to describe the non-stationarity of the ozone time series and the distributions of ozone during winter and spring are described. This wavelet approach provides a very good estimator of the timing of the vortex breakpoint and allows for a more objective estimation of ozone behavior before and after such breakpoints. Finally, the multi-resolution analysis of the potential vorticity (a proxy of airmass origin) at 375 and 475 K is also implemented and compared to the ozone decomposition.

Time series of air pollutants

Jiri Neubauer, Jaroslav Michalek, Frantisek Bozek
Masaryk University, Brno
Brnenska 1 682 01 Vyskov, Czech Republic
Telephone 420737439287
Email **neubauer@email.cz**

The topic of the article is the description of multivariate time series of air pollutants. These time series contain values of carbon monoxide (CO), nitrogen oxides (NOx), sulphur dioxide (SO2) and dust. The values in time series are based on all day averages of air pollutants which were measured in the town of Vyskov in the Czech Republic from 1.1. 1996 to 31.12.2001. The data shown are measured for the needs of air pollution monitoring in given locality. The article deals with the description of these time series and their relations, the forecasting of future values and extreme values. For the purpose of elaborating time series methods of time and spectral domain were used.

Using wavelets for decomposing and forecasting air pollution time series

Orietta Nicolis, Alessandro Fassò
University of Bergamo
Facoltà di Ingegneria
Via Marconi, 5
24044 Dalmine (Bergamo), Italy
Telephone 3471268815
Fax 35562779
Email **orietta.nicolis@univr.it**

The main aim of the work is to analyse and forecast intra-daily air pollutiom time series by wavelet transforms. Just as the periodogram produces an ANOVA decomposition of the energy of a signal into different Fourier frequencies, the wavelet periodogram decompose the energy into different scale components. For this property, the wavelet tool is particulary adapt to decompose a signal in his cyclical components. In this research work we propose a simple strategy aimed at improving time prediction accuracy, combining wavelet analysis with a (linear and non linear) forecasting models. First a wavelet transform is used to decompose the air pollution time series in varying scales and temporal resolution. The latter provide a sensible decomposition of the data so that the undelying temporal structure of the original time series became more tractable. Than we test different forecasting models to be applied on each resolution scales. This methodology is applied to recent intra-daily air pollution time series of Bergamo area.

Wavelet transform as a tool in climate variability studies

D. Novotna[1], N. Datsenko [2], D. Sonechkin[2]
[1]Institute of Atmospheric Physics, Prague, Czech Republic, [2]Hydrometeorological Research Centre of Russia, Moscow, Russia
Institute of Atmos. Phys., Acad. of Sci. of the Czech Rep. Bocni II/1401,
14131 Praha, Czech Republic
Telephone 420-2-72016043
Fax 420-2-72763745
Email **nov@ufa.cas.cz**

200-year time series of surface air temperature from central Europe were analyzed by continuous wavelet transform (WT) to study the climate variability. The time series were decomposed by WT into two components - the statistically stationary oscillations with periods up to 30 years and the low-frequency oscillations with periods comparable with the lengths of the analyzed time series. WT of particular time series and WT of their pair differences were compared. The stationary oscillations found in the different time series are not statistically distinguishable for the stations whose geographical distances do not exceed 400 km. The low-frequency oscillations are the same for the series investigated.

Environmental Indicators Initiative - Keeping the Quality In

Barry D. Nussbaum
Office of Environmental Information, U.S. Environmental Protection Agency

The U.S. Environmental Protection Agency's current Environmental Indicator Initiative has focused on developing and presenting information which is useful for the public and decision makers. Frequently the analysis of the data to produce the resulting information must undergo different quality checks depending on its use and audience. Perhaps the most important element of the quality check is the simple review of whether the underlying data are suitable for the desired indicator

Detection and Delineation of Critical Areas for Assessment and Management at Landscape Scales using Cellular Synoptic Data

G. P. Patil[1], W. L. Myers[2], C. Taillie[2]
[1]Center for Statistical Ecology and Environmental Statistics, Penn State University, USA;
[2] Center for Remote Sensing of Earth Resources, Penn State University, USA;

Keywords: Hotspots and Hotspot Areas, Critical Areas and Corridors, Biocomplexity, Regionalized Ecosystem Health, Ecosystem Distress Syndrome, Adjacency Relations, Cellular Tesselations, Connected Components, Zonation Tree, Spatial Scan Statistics, Echelons.

Echelons frame local values of synoptically mapped environmental indicators in regional context for comparative purposes and objective analysis of complex hierarchies in spatial variation across landscapes. The environmental indicators are considered as surface variables in virtual (or real) topographies. Echelons are structural entities consisting of peaks, foundations of peaks, foundations of foundations, and so on in an organizational hierarchy. It is natural to cast the echelon hierarchy as a dendrogram, from which profiles of spatial complexity can be obtained and "principal families" determined as contiguous areas of criticality from perspectives of either pronounced ecosystem health or pronounced ecosystem distress. Echelons have proven effective for elucidating concentration and connectivity of biodiversity, complexity of landscape change induced by factors such as wildland fire, pattern of propagation for urban sprawl, etc. Contemporary study of human disease as a component of ecosystem health entails a spatial scan statistic for detecting geographic clusters of disease and other responses that are significantly elevated with respect to the regional setting. In conjunction with the spatial scan statistic, echelon analysis can more clearly delineate zones of elevated intensity for focus of investigation.

Prioritization and Ranking of Watersheds Based on Watershed Indicators Without Having to Integrate Indicators for Multi-Criterion Decision Support

G. P. Patil
Center for Statistical Ecology and Environmental Statistics
The Pennsylvania State University, University Park. PA, USA

Keywords**:** Ranking and prioritization of Watersheds, Composite Indicators, Cumulative Rank Frequency Operator, Echelon Analysis of Geographic Posets, Hasse Diagram, Linearization of Partial Orders, Combinational MCMC.

To prioritize and rank means to linearize. Rather than derive a composite index, we will prioritize without having to integrate the indicators. This is now possible, and the approach is relatively novel and innovative. We have developed it for nationwide prioritization for UNEP with land, air and water indicators measuring the human environment interface at national level. For another example, a landscape atlas published by U.S. EPA (1997) considers 33 indicators of ecological condition on 123 watersheds (7-digit HUCs) of the Mid-Atlantic region and attempts to rank the watersheds using clustering and quintile-frequency methods. We address the question of ranking a collection of objects when a suite of indicator values is available for each member of the collection. The objects can be represented as a cloud of points in indicator space, but the different indicators (coordinate axes) typically convey different comparative messages and there is no unique way to rank the objects. A conventional solution is to assign a composite numerical score to each object by combining the indicator information in some fashion. Consciously or otherwise, every such composite involves judgments (often arbitrary or controversial) about tradeoffs or substitutability between indicators. Rather than trying to impose a unique ranking, we take the view that the relative positions in indicator space determine only a partial ordering and that a given pair of objects may not be inherently comparable. Working with Hasse diagrams of the partial order, we study the collection of all rankings that are compatible with the partial order and arrive at the ranking and prioritization, using cumulative rank frequency operator specially developed for the purpose.

Scale Invariance of Regional Wet and Dry Durations of Rainfields: a Diagnostic Study

Harry Pavlopoulos[1], Vijay K. Gupta[2]
[1]Department of Statistics, Athens University of Economics and Business, [2]Department of Civil & Environmental Engineering Cooperative Institute for Research in Environmental Sciences
University of Colorado
76 Patission Str., GR-10434 Athens, Greece
Email hgp@aueb.gr

Most of the recent work on rainfall data analyses and modelling has focused on either spatial or temporal variability. In this paper the structure of rainfall intermittence in space and time is investigated. Using a series of TOGA-COARE radar scans converted to maps of pixel rain rate over a tropical oceanic region of size 240x240 Km^2, regionally scale-invariant behavior of the probability distributions of wet and dry epoch durations is explored. Durations of wet and dry epochs are estimated by lengths of wet and dry spells respectively in time series of spatially averaged rain rate over sampled square sub regions of spatial scales ranging from 120 Km to 2 Km. The investigation is based on sample quantiles and sample moments of the underlying marginal probability distributions, focusing on the behavior of their tails and their variation with respect to spatial scale. We find that sample tail quantiles and sample moments of wet durations exhibit *power-law* multiscaling, while sample tail quantiles and sample moments of dry durations exhibit *exponential* multiscaling across the above range of scales. These findings provide new statistical diagnostic tools for validation of spatio-temporal models for rainfields.

Predicting aboveground biomass using field data and high resolution spectral imaging data

Edzer J. Pebesma, Steven M. de Jong
Dept. of Physical Geography Utrecht University
3508 TC Utrecht, The Netherlands
telephone: +31 30 2533051;
fax: +31 30 2531145
Email: **e.pebesma@geog.uu.nl**

In July of 1997, an experimental flight with the Digital Airborne Imaging Spectrometer (DAIS7915) was carried out in the Peyne catchment area in Southern France. At the same time, field biomass estimates for 83 sites were collected. A first analysis of these data showed that multiple linear regression of log-transformed above biomass on a selection of five of the 79 available wavelength bands, combined with kriging interpolation of the residual variation, was reasonably succesful for spatial prediction of aboveground biomass. However, multiple collinearity of the image information played a large role during, and also after, the variable subset selection. In this presentation we will discuss this problem, and also compare the multiple regression approach to alternatives like ridge regression and principal component regression.

Probabilistic Quantitative Ecological Risk Assessment Modelling using AQUARISK

Jayasuriya Perera[1], **RayCorrell[2],** John Twining[1]
[1]Australian Nuclear Science & Technology Organisation, Australia, [2]Commonwealth Scientific and Industrial Research Organsiation, Australia.
Private Bag No.2,
Glen Osmond SA 5064, Australia
Telephone +61 8 8303 8769
Fax +61 8 8303 8763
Email **Ray.Correll@csiro.au**

A computer program to assess risk to aquatic communities due to chemical pollution in water bodies – AQUARISK, has been developed at Australian Nuclear Science and Technology Organisation (ANSTO) with technical support from Commonwealth Scientific and Industrial Research Organisation (CSIRO). The program uses water quality data (WQD) ie concentrations of chemical pollutants and dose response data (DRD) ie dose response of aquatic species to chemical pollutants data to derive risk to aquatic communities. The program fits Lognormal and Burr Type III probability distributions to the WQD and DRD to derive to derive maximum likelihood parameters (MLE) of the distributions. The user can select any combination of these two distributions for the risk analysis. In the case of Burr Type III distribution, Nedler Mead simplex algorithm was used to maximise the log-likelihood function to derive MLE parameters. Critical values of hazardous concentrations (HCu,v values) i.e. hazardous concentration affecting only u per cent of species at v percent confidence level can be determined using Aldenberg and Slobb's method of extrapolation constants for lognormal distribution of DRD and using a Bootstrapping technique for Burr Type III distribution. The program estimates the degree to which the water quality measurement data are likely to exceed reference values such as HC values. The convolution of the hypothesised distributions of WQD and DRD is used to estimate the degree of overlap between WQD and DRD and thereby assess the risk to aquatic community from each chemical as an average percentage of species affected. This is calculated in two different methods (i) numerical integration and (ii) Monte Carlo simulation. The extent of reduction in pollution to achieve an agreed degree of reduction of risk to aquatic communities can be estimated and provides the required link between environmental engineering management and specific downstream impacts.

Assessment of Primary Production by Statistical Analysis of Water-Quality Data

Wilhelm Petersen[1], Ulrich Callies[2]
[1]Institute for Coastal Research, [2]GKSS Research Centre
Max Planck Str. 1
D-21502 Geesthacht, Germany
Telefone +49 4152 87-2358
Fax +49 4152 87-2366
Email **wilhelm.petersen@gkss.de**

Time series of weekly water-quality data at a station on the Elbe River (Schnackenburg from 1984–2000) were subjected to principal component analysis (PCA). Considering the amplitudes of composite patterns of variables is a step towards a process-oriented interpretation of water-quality data. One specific objective was to investigate the impact of improved water quality after the German reunification in 1990 on primary production and the oxygen budget. To discriminate anthropogenic signals from natural fluctuations a separation of the impact of discharge was attempted based on a linear regression approach. A dominant pattern of co-variation in the residual data could be attributed to biological activity (primary production). The most relevant variables of this 'biomode' are oxygen saturation, pH and ortho-phosphate. We conclude that multivariate statistical analysis of water-quality data can help to estimate primary production when direct observations of algal concentrations are missing. In the years from 1998 – 2000 the trend of the 'biomode' indicates an increased load of oxygen consuming biomass caused by enhanced primary production in the middle stretches of the Elbe River which corresponds with the observation of more severe oxygen deficits in the tidal section of the river.

Kriging and mixed effects models: some connections and a case study on soil degradation

Alessio Pollice
Dipartimento di Scienze Statistiche Università degli Studi di Bari
Via C. Rosalba n.53
70124 Bari, Italy
Telephone ++39 080 5049243
Fax ++39 080 5049147
Email **apollice@dss.uniba.it**

The effectiveness of the use of mixed effects models for estimation and prediction purposes in spatial statistics for continuous data is reviewed in the classical and Bayesian frameworks by comparison to standard Geostatistical techniques. A case-study about the degradation produced by tillage on some typic southern italian soils is also provided.

Analysis of Particulate Fine Matters in Lombardia Region, Italy: a Geostatistical Approach

Emilio Porcu, Marco Geraci
University Milano Bicocca
Piazza dell'Ateneo Nuovo,1 Milano, Italy
Telephone 3494712702
Email **emilio.porcu@unimib.it**

The aim of this work is to give a description of the distribution of the Particulate Fine Matters (PM10)in the Lombardia region, Italy. As known, this pollutant is particularly dangerous and diffused in the region of interest. The problem is to extend partial information, given by stations located in all the territory, to global information, regarding the map. In other words, the problem is the passage from the discrete to the continuous case. To obtain our purpose some preliminary operations are necessary, such as the analysis of the phenomenon's variability on the plane; nevertheless, it could also be important the analysis of perturbing factors, such as wind, which could influence our case of study, in terms of direction and spatial trend. For these characteristics, and other ones omitted for the sake of sintesis, the phenomenon should be analyzed with geostatistical methods. The instrument used for the analysis of correlation in space is the variogram, which represents the first step of a procedure conducting to a representation of the correlation in space with a theoretical model. Kriging and Cokriging methods will be used to interpolate data and build the map, in such a way as to evaluate potential risk areas for the case considered.

Statistical Issues for Geostatistical Modeling of Soil Moisture

Stephen L. Rathbun, Henry Lin
The Pennsylvania State University
Department of Statistics 326 Thomas Building
Pennsylvania State University University Park, PA 16802, USA
Telephone (814) 863 7395
Fax (814) 865 1348
Email **rathbun@stat.psu.edu**

Soil moisture plays a pivotal role in regulating global climate, hydrological cycles, carbon sequestration, ecosystem function, and environmental quality. Spatio-temporal variation in soil moisture is regulated by a complex milieu of interacting pedological, climatic, topographic, and vegetative processes. We propose a hierarchical spatio-temporal model for such interactions. The joint distribution of all relevant variables is written as a product of physically motivated conditional distributions in which the data for each variable is modeled as a function of the realizations of previously modeled variables in the hierarchy. Feedback mechanisms can be incorporated by allowing dependence on the past values of any variable. In this paper, we shall present preliminary results, illustrating our approach on data from the Little Washita watershed in Oklahoma, USA. These data include point samples for soil, moisture, soil temperature, and climate, Digital Elevation Models for topographic data, soil maps, and land use/ land cover maps derived from satellite imagery. Future plans call for integrating these data with remote sensing data from the airborne Electronically Scanned Thinned Array Radiometer.

Normalisation of environmental air quality data using spatio-temporal analysis

Renata Romanowicz, Peter C. Young, Patrick Brown, Peter Diggle
Lancaster University, UK
Environmental Science Department Lancaster University Lancaster,
LA1 4YQ UK
Telephone + 1524 593894
Fax + 1524 593985
Email R.Romanowicz@lancaster.ac.uk

This paper investigates to what extent it is possible to infer the level of human impact on the environment by removing natural fluctuations in air quality data (nitric oxide NO, nitric dioxide NO2 and ozone O3 concentrations) measured at 23 stations around Paris. In addition, the subsequent prediction of the spatio-temporal pattern of changes in NO, NO2 and O3 concentrations near Paris should help to identify the major sources of a pollutant and may be useful in various ways: for example as an aid to the regulatory bodies who are responsible for the control of pollution in the area. Analysis of the time and spatial variability of observations of NO, NO2 and O3 in the Paris area is divided into: (i) Time series analysis of AirParif data; (ii) Development of combined spatial and temporal analysis techniques using NO observations from 19 stations; (iii) Derivation of a temporal model of O3-NOx (NO+NO2) variations. The first part describes the application of Dynamic Harmonic Regression (DHR) model to non-stationary time series. The DHR model is used to interpolate the data sets of NO concentrations over periods where measurements are missing and to decompose the time series into trend and harmonic components. The short-term (two year data set) analysis is performed for 19 stations round Paris. The results of this analysis are used in further spatio-temporal analysis of the data. Also a set of 9 years of observations is analysed for one of the stations to investigate the long-term variations of the concentrations. The spatio?temporal analysis of NO data consists of four steps: preliminary analysis of spatial relations within the data sets, development of a spatio - temporal model for the trend; derivation of the model for trend and first weekly components and derivation of the full model, consisting of a trend and two weekly components. The results of the analysis indicate that the simple spatio-temporal model consisting of trend and noise efficiently represents spatio-temporal variations in the data. The aim of the third part of this study is the analysis of the relation between ozone and NOx concentrations in the area round Paris. The technique developed here may be subsequently used to build a more general model incorporating temperature, hydrocarbons and wind speed observations, which were not used at this stage. The present analysis consists of: (i) development of a non-linear dynamic relation between ozone and NOx concentrations; and (ii) spatial analysis of the distribution of pollutants in the Paris area. The results of this work suggest that a non-linear dynamic relation between NOx and ozone data can be identified, robust both in time and space. The under - prediction of the ozone peaks indicates that some additional information on hydrocarbons is needed to improve the predictions.

A geostatistical approach to assessment of regional air quality models

Paul D. Sampson, Peter Guttorp
University of Washington
Department of Statistics Box 354322
Seattle, WA 98195-4322, USA
Telephone 1-206-685-6785
Fax 1-206-616-9443
Email **pds@stat.washington.edu**

Under the US Clean Air Act, states are responsible to maintain air quality. When the standards are violated, the state is required to submit a plan for improving air quality to the Environmental Protection Agency. For sufficiently severe violations, the state must develope a numerical air quality model, demonstrate that it describes historical data well, and that the proposed plans will result in improvements according to the model. Here we focus on how to demonstrate that the model reflects current data. Since most air quality models are gridded, and data are collected at points, it is necessary to somehow make the two data sets compatible. We propose to use geostatistical tools to estimate grid square averages from data, taking into account spatial and temporal inhomogeneities. The methods are applied to data from California's Central Valley, collected in 1990 in connection with the SARMAP model assessment exercise.

Identification of Signal Peptides Using a Hidden Markov Model: its Application to a Lytic Hydrolase of Biotechnological Importance

M. Santos[1], **L. Sousa**[2], M. A. Amaral Turkman[2], W. Urfer[3]
[1]Dep of Plant Biology - Lisbon University, [2]Dep of Statistics and Operation Research - Lisbon University, [3]Dep of Statistics - Dortmund University
DEIO Ed. C2 - Piso 2
1749-016 Lisboa, Portugal
Telephone +351 21 7500413
Fax +351 21 7500081
Email **lisetes@fc.ul.pt**

The good quality of the wine is ensured by a proper fermentation. The bacterium *Oenococcus oeni* is an important player in this process due to its ability to convert malic acid into lactic acid, in the presence of high ethanol concentrations, thus leading to a reduction in wine acidity. Bacteriophages such as fOg44 can destroy *O. oeni* cells, impairing malolactic fermentation. Successful bacteriophage attack depends on a phage product, a lytic hydrolase known as lysin. The fOg44 lysin (Lys44) belongs to a family of lysosymes with the capacity to cleave 6-O-acetylated peptidoglycans such as the present in the cell walls of bacterial pathogens such as *Staphylococcus aureus*, which are not hydrolysed by other enzymes. Our studies on the fOg44 lysin may therefore give insights on how to prevent oenococcal lysis and, at the same time, may lead to its use as a food preservative due to its action on bacterial pathogenic contaminants. Unlike most lysins, the N-terminal region of the *Oenococcus oeni* phage fOg44 lysin (Lys44) seems to function as an export signal. Known as signal peptide, the export signal is an N-terminal peptide typically between 15 and 40 amino-acids long, which is cleaved from the mature part of the protein during translocation across the membrane. To decide about the existence of a signal peptide in Lys44, a hidden Markov model is applied. For proteins, a hidden Markov model consists of a number of states connected by transition probabilities. A distribution over the 20 amino-acids is associated with each state. The model indicates, with high probability, the presence of a signal peptide with cleavage site between residues 27 and 28. Experimental data confirm this prediction and thus the suitability of the model.

How to test information redundancy of environmental detectors in the time domain

Emma Sarno
Dipartimento di Economia e Statistica, Università della Calabria, Italy
Telephone 0984 492426
Fax 0984 492421
Email **e.sarno@unical.it**

The increasing interest towards environmental themes has made the question of the informational efficiency of monitoring networks a rather topical issue. There is a vast literature on how networks should be designed and evaluated. Nevertheless, proposals often arise from spatial statistics and sometimes they overlook the temporal dimension of such data, which are actually collected over time. Recently, Costanzo and Sarno (2000) considered this point explicitly. They proposed to model data generating processes corresponding to each monitoring site by ARIMA models. Subsequently, they measured the structural discrepancies between models by means of the Autoregressive distance (Piccolo, 1990), that is linked to the forecast function. They located sites that, having maximum distance from the remaining, provided information which could be only poorly foreseen by them. However, their criterion was merely descriptive. To overcome this restriction, here we consider a probabilistic framework. Our aim is to find significant differences between models, since any lack of significance is equivalent to information redundancy. Hence, by applying a sequence of tests where the null hypothesis is that data were generated by the same model, we will state the optimal dimension of an existing network for a given amount of information, starting from the number of significant tests. Hourly data of atmospheric carbon monoxide from an air monitoring network will be used as case-study.

Identification of meteorologically induced variations of water quality using an over-parameterised mechanistic model

Mirco Scharfe, Ulrich Callies
GKSS Research Centre
Max-Planck-Str. 1,
21502 Geesthacht, Germany
Telephone +49 4152 87 1876
Fax +49 4152 87 2818
Email **mirco.scharfe@gkss.de**

Water quality data (chlorophyll_a, oxygen, silica) that have been collected at the station Weir Geesthacht on the River Elbe (Germany) are analysed to identify that proportion of observed variability which can be attributed to time dependent meteorological forcing. The specific approach is based on running a simple mechanistic water quality model (WAMPUM) which has been developped at GKSS. WAMPUM considers a series of water packages moving downstream independently. Therby all dispersion processes are neglected. The model is set up in such a way that meteorological forcing is the only time dependent input. Even though the model concept is very simple, Monte Carlo experiments prove the model being significantly over-parameterised. A local uncertainty analysishas been carried out to analyse how minimising the mean deviation of model output from observations may reduce uncertainty about some linear combinations of model input parameters while other parameter combinations remain completely uncontrolled by data. The reduction of model uncertainty by model calibration is assessed by analysing the eigenvectors of the posterior model input parameter covariance matrix. Characteristic features for different years are discussed. Monte Carlo experiments are carried out to illustrate the extent to which results of the local uncertainty anlaysis are informative with respect to finite changes of model parameters.

Some statistical problems in turbulent diffusion

Tom Schopflocher
Ph.D. Research Scientist Chemical Computing Group, Inc.
1010 Sherbrooke Street West, Suite 910 Montreal, Quebec
H3A 2R7 Canada
Telephone (514) 393-1055 ext.46
Fax (514) 874-9538
Email **tom@chemcomp.com**

Suppose a miscible scalar contaminant is introduced into a turbulent flow (consider the pollutant emitted from a smokestack, for example). A difficult topic in fluid mechanics is how one should monitor the concentration level of this (potentially harmful) substance. Certainly, this is a problem of great practical importance with regard to risk and hazard analysis. Concentration of contaminant is a random variable, characterized by having large fluctuations about its mean. Ideally, one would like to monitor these fluctuations using the fundamental physics of the system. Unfortunately, due to the non-closure problem that is omnipresent in the study of turbulence, it is impossible to solve for the PDEs that govern both the velocity field and the concentration level. While it is often straightforward to use conventional time series methods to describe the measurements of contaminant dispersion, these techniques have usually been considered *ad hoc*, and therefore unlikely to provide any insight into the physics of the problem.
The dispersion of contaminant in a turbulent flow is governed by two distinct processes. Convective forces act at fast time scales, but do not contribute to molecular mixing. Molecular diffusion, while a very slow process relative to convection, is the only agent available to reduce contaminant concentration. This being said, it is reasonable to believe the underlying PDF is a mixture. In this talk, I apply the two-state Markov process methodology of Hamilton to model some high-quality, stationary, wind-tunnel data of CSIRO. Although the analysis is preliminary, it is promising. At the very least, it is interesting to compare these results to those obtained when one assumes an *iid* sequence and fits various forms of mixtures directly to the data.

Air quality assessment in UK and Republic of Ireland using non parametric methods

Marian Scott[1], Ron Smith[1], Marco Giannitrapani[2]
[1]CEH, Edinburgh, UK, [2]University of Glasgow and Palermo
University of Glasgow
Dept of Statistics, University of Glasgow
Glasgow G12 8QW, UK
Thelephone 00 44 141 330 5125
Fax 00 44 141 330 4814
Email **marian@stats.gla.ac.uk**

Atmospheric pollutants such as sulphur dioxide, ammonia and ozone have been monitored over a spatially distributed European network for the past 20-30 years. During this time, a number of international conventions and protocols have been put in place to effect a reduction in the emissions of such gases. It is therefore of interest to policy makers to demonstrate (if possible) that the reductions in emissions can be linked to decreasing atmospheric concentrations and deposits at vegetation and ground level. Measurement data from the EMEP (European Monitoring Environment Programme) will be used to illustrate some trend analysis techniques for identification and quantification of any trend but also to highlight difficulties and problems which may occur.

Estimation of bioavailability of environmental oestrogens in population models

Silvia Selinski
Department of Statistics University of Dortmund
44221 Dortmund Germany
Telephone +49-(0)231-755 5918
Fax +49-(0)231-755 5303
Email **selinski@statistik.uni-dortmund.de**

Potentially adverse human and environmental hormone mimetic effects of environmental oestrogens are a matter of current concern. Environmental oestrogens belong to the so-called endocrine active compounds (EACs) which alter signalling processes of the endocrine system leading to a broad range of effects during foetal and postnatal development, puberty, adulthood, and aging. A number of synthetic chemicals, so-called xenoestrogens, and several plant-derived compounds, so-called phytoestrogens, are known to have weak oestrogenic activity. The present study is part of the risk assessment of the weak environmental oestrogens daidzein, p-tert-octylphenol, and bisphenol A. The isoflavone daidzein is an important phytoestrogen with respect to dietary exposure (soy beans and soy products). P-tert-octylphenol and bisphenol A are industrial chemicals. The toxicokinetics and the bioavailability of these three substances in female DA/Han rats after oral and single intravenous application were investigated by the use of population models.

KEYWORDS: Hierarchical models, population models, endocrine active compounds, environmental oestrogens, toxicokinetics, bioavailability, daidzein, bisphenol A, p-tert-octylphenol.

Evaluation of standards using a mixed model

Eric P. Smith, Alyaa Zahran, Mahmoud Mahmoud, Keying Ye
Department of Statistics Virginia Tech.
Blacksburg VA 24061-0439
Telephone 540-231-7929
Fax 540-231-3863
Email **epsmith@vt.edu**

A common approach in the assessment of water quality is to collect data at a site over time and compare the measurements with a standard. The approach that is used for testing is weakened by small sample sizes. However, it is common that the data are collected as part of a network. For example, data is collected at numerous sites within a watershed. By combining the measurements into a model we can borrow strength from neighbors and hopefully improve the assessment of a single site. We evaluate this approach for testing impairment for data arising under the Clean Water Act in the United States.

Evaluation of model uncertainty in ecological risk assessment using Bayesian Model Averaging

Eric P.Smith, Keying Ye,
Department of Statistics, Virginia Tech.

Variable selection is one of the most important and controversial issues in modern data analysis. In the study of relationships between biological communities and environmental conditions, variable selection is especially important as it guides decisions about environmental management. Using a case study from the Great Lakes we use Bayesian Model Averaging (BMA) to select interesting subsets of environmental variables (such as water chemistry, depth, etc), that can affect the abundance of benthic microinvertebrates. We implement BMA for a multivariate technique called Canonical Correspondence Analysis (CCA) and use its results to represent sites, species and selected environmental variables on ordination diagrams (biplots) along with "error bars" representing uncertainty due to both sampling variability and model selection. BMA provides data analysts with an efficient tool for discovering promising models and obtaining estimates of their posterior probabilities via Markov chain Monte Carlo (MCMC). These probabilities are further used as weights for model averaged predictions and estimates of the parameters of interest. As a result, variance components due to model selection can be estimated and accounted for, contrary to the practice of conventional data analysis.

The development of statistical uncertainty assessments for air pollution effects

Rongvald I. Smith
Centre for Ecology and Hydrology,Edinburgh
Bush Estate Penicuik Midlothian
EH26 0QB, UK
Telephone +44 (0)131 445 4343
Fax +44 (0)131 445 3943
Email **ris@ceh.ac.uk**

Acidification of remote and fragile ecosystems in the UK still occurs from the gaseous emissions of sulphur and nitrogen compounds despite the general improvement in the European pollution climate. Future government policy depends on the predicted modelled effects of this pollution, primarily from rain and cloud deposition. These deposition effects are normally produced from deterministic models and there is a need to generate statistically acceptable measures of uncertainty on the model outputs. This is not a complex modelling situation, but there are a number of issues which do not have simple solutions. These include the spatial and temporal modelling of observed concentration data, the development of sensitivity analysis for the model, the use of measured surrogate data, and the spatial scale dependence of the results. Both the development of the statistical analysis and the presentation of output for policy development will be considered.

Environmental impact from dust generated by refuse materials: a new approach

C. Solisio[1], V. Garerri[2], A. Del Borghi[1], M. Pescetto[2]
[1]Dipartimento di Ingegneria Chimica e di Processo "G.B.Bonino", Università di Genova
Via Opera Pia 15, Genova, Italy
Tel: +39 –010- 353 2587
Fax: +39 –010- 353 2586
Email: **solisio@unige.it**
[2] Provincia di Savona, Italy

The transport and storage of refuse materials costitute a severe environmental problem in several coastal areas where major harbor facilities are located. In spite of their significant environmental impact, these activities have received little attention from both the scientific community and the regulatory authorities.
We have chosen to tackle this problem using a multidisciplinary approach, which includes materials science, air transport of dust particles, determination of optimal scheduling and management of transport and storage activities.
Some preliminary results obtained in the coastal area of the Savona harbor are presented and discussed.

Review of a Strategy for the Prediction of the PDF of Scalar Concentrations

Paul J. Sullivan
Department of Applied Mathematics
The University of Western Ontario
London, Canada N6A 5B7
Email pjsul@uwo.ca

The importance of the PDF (and particularly the high concentration tails) in addressing many relevant environmental issues motivates this talk. The intractable differential equations that describe the evolution of the PDF and the experimental difficulties involved in its measurement are reviewed. An alternative formulation in terms of the moment evolution equations is observed to be less intractable than that for the PDF, however, it is further shown that a low parameter PDF that accurately represents the high concentration tail behavior is beneficial here both from an experimental and theoretical predictive viewpoint.
A simple, physically based, closure scheme is shown to be useful in determining the low order moments required for the parametric PDF described above. An illustration using plume data from a well-controlled, grid turbulence, wind tunnel data is described.

The San Francisco Estuary Regional Monitoring Program: management implications and benefits of redesign.

Karen Taberski[1], Don Stevens[2], Sarah Lowe[3], Bruce Thompson[3], Jay Davis[3], Rainer Hoenicke[4]
[1]San Francisco Bay Regional Water Quality Control Board, [2]Oregon State University, [3]San Francisco Estuary Institute, [4]California Resources Agency
1515 Clay Street,
#1400 Oakland, CA 94612, USA
Telephone 510 622 2424
Fax 510 622 2460
Email KMT@rb2.swrcb.ca.gov

The San Francisco Estuary Regional Monitoring Program was established in 1993 to monitor water quality in the San Francisco Estuary. The program was established as a unique collaborative effort between the state water quality regulatory agency, the dischargers from the area and the San Francisco Estuary Institute. The monitoring program was developed to measure contaminant concentrations and toxicity in water and sediment and to measure tissue accumulation in bivalves. The design of the program has been deterministic. Station locations were chosen based on spatial coverage, distance from a point source of contamination (to reflect ambient conditions) and at sites where there was historical data. Stations were located along the "spine of the Estuary";, mostly in deeper areas. In 1997, a thorough external review of the program was conducted. Through the review process, new ideas to improve and redesign the program were developed: 1) segments of the Estuary have been redefined based on analysis of water quality data rather than geography, 2) the monitoring design has been changed from a deterministic to a statistically random design and, 3) power analysis was conducted to determine the power of having a particular number of stations per segment. The benefits of this design to the regulatory agency are that: 1) segmentation that may be used for water quality regulation is based on water quality factors, 2) data will be truly representative of the waterbody, and segments of the waterbody, so that there is increased confidence in designations of "impairment"; and, 3) there is increased spatial coverage, allowing for analysis of both shallow and deep portions of the Estuary.

Selecting weights of satellite image and ancillary information in k-NN estimation: a genetic algorithm approach

Erkki Tomppo, Merja Halme
Head of Finnish National Forest Inventory - The Finnish Forest Research Institute
Unioninkatu 40 A FIN-
00170 Helsinki, Finland
Telephone 358-9-85705340
Fax 358-9-625308
Email: erkki.tomppo@metla.fi

The non-parametric k-nn multi-source estimation method is widely studied by forest inventory groups when utilizing satellite images and field data (Tomppo, 1991, Nilsson 1997; Eisele 1997; Tomppo et al. 1999; Gjertsen et al. 2000, Franco-Lopez et al. 2001, Tomppo et al. 2001). The method makes it possible to estimate all inventory variables at the same time, their number usually being more than 100. Pixel-level root mean square errors may be high (Tokola et al. 1996), but can be reduced by means of a relocation of sample plots (Halme and Tomppo 2001). When the area of interest belongs to several vegetation zones, a spectral vector of the remote sensing data may correspond to several different ground data vectors. Thus an important decision is the choice of the pixel-dependent geographical area from which the nearest field plots (in the spectral space) for each pixel are selected. The weighting of different spectral variables is an obvious problem when computing the distance in the spectral space. A new method is presented where information about large scale forest variation is supplemented to the variables of k-nn estimation. The weights of the large scale forest variables and image variables for the distance metric are computed by means of a genetic algorithm. The tests with practical forest inventory data show that the method performs noticeably better than other applications of k-nn estimation methods in forest inventory.

The Advanced-Maximum-Linkage-Clustering-Approach and its application to aerial photographs and erosion data

Lars Tschiersch, Matthias Zerbst
Department of Statistics, University of Dortmund, Germany
Lars Tschiersch Universität Dortmund Institut für Wirtschafts- und Sozialstatistik
Vogelpothsweg 78
44221 Dortmund, Germany
Telephone +49 231 755 5735
Email su0491@studserver.uni-dortmund.de

The advanced maximum-linkage-algorithm (AMLA) is a derivative of the maximum-linkage-algorithm (MLA) given by Zerbst (2001). AMLA produces clusterings which have a good separation between the builded classes. To reach the separation, centroids were calculated, which are the basis for classification. The separation between the classes is considered large, if this is also true for the so far calculated centroids. The choice/selection of centroids having a large distance to each other are also guaranteed using the MLA. AMLA is improved in this way that the underlying frequency structure of the data can be adapted individually to the problem under consideration. Therefore an additional parameter ??[0, 1] is introduced. The parameter gives the degree to which the frequency structure is regarded. Zerbst, M. (2001): Die pixelbasierte Clusterung von Luftaufnahmen im Rahmen von Erosionsuntersuchungen. Dissertation, Universität Dortmund, Dortmund, Germany.

Interpolating counts of exceedance of thresholds for NOx and ozone air quality data.

Jan van de Kassteele[1], Alfred Stein[1], M. Velders[2], Arnold L.M. Dekkers[2]
[1]Mathematical and Statistical Methods Group, Wageningen University, [2]RIVM, Bilthoven, The Netherlands.
Dijkgraaf 4-
13B 6708 PG Wageningen, The Neterlands
Telephone +31 317 424092
Email jan.vandekassteele@mat.dpw.wau.nl

In the Netherlands concentrations of pollutants in the air (NO, NO2, Ozone) are measured at a large number of sites to obtain information about the air quality. Regulations in the Netherlands and the European Union require that certain thresholds are not, or maximally a limited number of times, exceeded. For policy makers and health risk analysis it is therefore useful to have information in the form of maps about the number of days per year where a certain component exceeds a threshold value (norm), instead of a yearly average concentration. The chance of exceeding that norm, so not the actual values, may be even more interesting. In the Netherlands, RIVM operates the national air quality monitoring network which provides a dense dataset of observations of NO and NO2 (sum is NOx) and O3. From these, maps are compiled about the air quality by interpolating the observations. The interpolation of these spatial correlated data can be done by kriging. In this case however, we are confronted with count data, which do not obey the Gaussian assumptions on which the kriging procedure has been originally based. To overcome this problem, a recently developed technique, which is putting the kriging procedure in the framework of generalized linear models (Diggle et al., 1998), is implemented and applied. The first results of the method will be discussed. Simulations will be applied to obtain information on chances of exceedances.

Bayesian variable selection in multinomial probit models with application to spectral data

Marina Vannucci
Department of Statistics, Texas A&M University
Phone: +1-979-845-0805
Fax: +1-979-845-3144
Email mvannucci@stat.tamu.edu

Here we focus on classification problems where the number of predictors substantially exceeds the sample size and propose a Bayesian variable selection approach to multinomial probit models. Motivated by the binary model with latent variables, we consider multivariate extensions to the case of more than two categories and use latent variables to specialize the general distributional setting to the linear model with Gaussian errors. We then apply Bayesian variable selection techniques that use natural conjugate prior distributions. A posteriori we perform inference on the marginal distribution of single models using MCMC methods and truncated normal and student-t sampling techniques.
We present an application in chemometrics where we aim at classify three wheat varieties based on 100 near infra-red absorbances. Other possible applications are in cancer classification based on microarray data.

A Pairwise Framework for Geostatistics

Cristiano Varin [1], Gudmund Høst[2]
[1]Department of Statistical Sciences, University of Padua (Italy), [2] Norwegian Computing Center (Norway)
Dipartimento di Scienze Statistiche
Via Cesare Battisti 241/243
35121 Padova, Italy
Telephone 347 5353625
Fax 39 049 827 4170
Email sammy@stat.unipd.it

Spatial Generalized Linear Mixed Models (GLMMs) have proved to be an appealing tool for geostatistic purposes (Diggle, Tawn & Moyeed, 1998). Unfortunately, the cost for the flexibility in Generalized Linear Mixed Models is the formidable complexity of the likelihood and the predictive distribution which involve intractable high dimensional integrals. In the last years several authors have proposed solutions to overcome this problem but the developed techniques (like MCMC and MCEM algorithms) are computationally expensive and seem to be inadequate for large datasets, common in spatial/temporal and remote sensing applications. For this reason it is of interest to investigate alternative strategies for fitting and forecasting in spatial GLMMs. In this direction goes the pairwise likelihood, a pseudo likelihood formed by composition of valid likelihoods for pairs of observations. Pairwise likelihood is appealing since it dramatically reduces the computational burden as compared to full likelihood approaches and it produces estimators with reasonable theoretical properties. In this project we propose an EM type algorithm for fitting GLMMs by pairwise likelihood, and we extend this pairwise framework to the forecasting problem, building a simulation based predictor where only pairs of observations are involved: the pairwise predictor. The pairwise approach can be extended to general n-tuples of observations providing a flexible class of inference and forecast methods. This gives a possibility to select a method to suit both the problem at hand and the available computational power.

Added value of a (hierarchical) bootstrap model in environmental standard setting

Frederik A.M. Verdonck[1], Olivier Thas[1], Joanna Jaworska[2], Peter A. Vanrolleghem[1]
[1]Department of Applied Mathematics, Biometrics and Process Control (BIOMATH), Ghent University, [2]Procter & Gamble, ETC
Coupure Links 653,
B-9000 Gent, Belgium
Telephone +32 9 264 5937
Fax +32 9 264 6220
Email **frederik.verdonck@biomath.rug.ac.be**

Several toxicity data of species can be collected and used to set safe environmental standards in order to protect organisms and communities from adverse effects due to exposure to a certain chemical. These toxicity data typically have a hierarchical structure. Different species have different sensitivities towards a chemical (inter-species variability). Each species is tested in several laboratories (inter-laboratory variability). Each laboratory uses several individuals, each individual having its own sensitivity (intra-species variability). In current practices, summary statistics (such as the mean) are taken on each level. As a result, a lot of information is discarded. Hierarchical models can be used to account for all these levels of variability. Alternatively, one could ignore the hierarchy and just consider all available data points into a single distribution. In this presentation, the added value of an extended version of the bootstrap method, ?weighted hierarchical bootstrapping? is compared to a non-hierarchical model. Both scientific accuracy and feasibility of practical use are considered.

Data-Analysis of Environmental Air Pollutant Monitoring Systems in Europe

Kristina Voigt[1], Gerhard Welzl[1], Rainer Brüggemann[2]
[1]GSF - National Research Center for Environment and Health, Institute of Biomathematics and Biometry,
[2]Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Department of Ecohydrobiology, Department of Biostatistics, 85764 Neuherberg, Germany
Email kvoigt@gsf.de

Public access to information about the environment is being strengthened across Europe. The concept of public's right to information gives the basis for the access to environmental information. Mathematical and statistical groups should actively influence the interpretation of environmental and chemical data given in and extracted from the Internet in order to receive effective and correct results.
In this paper the quality of air pollutant monitoring systems in the 15 European member states is analyzed. For pragmatic reasons only the capitals are looked upon. Comprehensive data on environmental monitoring programs concerning air pollutants like ozone (O_3), nitrogen dioxide (NO_2), nitrogen oxide (NO), carbon dioxide (CO_2) and carbon monoxide (CO), and sometimes suspended dust, benzene and other environmental chemicals are available on the free Internet. As different monitoring information systems exist in the European member states a comparison of these systems with their pros and cons is of great interest to the public. Environmental air pollutant monitoring systems in the capitals of the 15 EEC member countries (objects) are evaluated by applying 5 evaluation criteria for the differentiation of these systems. The scores run from 0 = insufficient, 1=medium, to 2 = excellent.
Different data-analysis methods will be applied, the Hasse diagram technique, a method derived from discrete mathematics and the Partially Ordered Scalogram Analysis with Coordinates (POSAC) method, a multivariate statistics' approach. These methods are compared and advantages and disadvantages will be elaborated.
The aim of this data-analysis is the evaluation of the publicly available air quality monitoring systems in Europe with their pros and cons. This might help the interested public to find and understand the information given on the Internet. Furthermore our evaluation approach might give some recommendations for an improvement of the air quality monitoring systems.

Looking at ozone station data and model output in the framework of external drift kriging

H. Wackernagel[1], C. Lajaunie[1] C. Roth[2] N. Blond[3], R.Vautard[3]
[1]Centre de Géostatistique, Ecole des Mines de Paris; [2]Airparif; [3]Laboratoire de Météorologie Dynamique, Ecole Polytechnique
Centre de Géostatistique, 35 rue Saint Honoré F
77305 Fontainebleau, France
Telephone +33-1 64 69 47 60
Email wackernagel@cg.ensmp.fr

The problem of correcting model output using station data has been classically solved in meteorology under the denomination of "objective analysis", by performing a simple kriging of the difference between station data and model output. The external drift kriging offers an equivalent, yet more direct approach to objective analysis. After examining the relation between objective analysis and universal kriging in the framework of external drift kriging, results using Airparif station data and Chimere model output are presented and discussed.

Use of Landscape and Land Use Parameters for Classification and Characterization of Watersheds in the Mid-Atlantic across Five Physiographic Provinces

D. H. Wardrop[1], J. A. Bishop[1], M. Easterling[1], W. L. Myers[2], G. P. Patil[3], and C. Taillie[3]
[1]Co-operative Wetlands Center, the Pennsylvania State University, USA, [2]Center for Remote Sensing of Earth Resources, Penn State University, USA; [3]Center for Statistical Ecology and Environmental Statistics, the Pennsylvania State University, University Park. PA, USA

Keywords: Atlantic Slope Consortium, Mid-Atlantic Region, Characterization of Watersheds, Classification of Watersheds, Environmental and Ecological Indicators, Landscape Parameters, Land Use Parameters.

The Atlantic Slope Consortium (ASC) is a project designed to develop and test a set of indicators in coastal systems that are ecologically appropriate, economically reasonable, and relevant to society. The suite of indicators will produce integrated assessments of the condition, health and sustainability of aquatic ecosystems based on ecological and socioeconomic information compiled at the scale of estuarine segments and small watersheds. The research mandate of the ASC project is the following:Using a universe of watersheds, covering a range of social choices, we ask two questions:

- How "good" can the environment be, given those social choices?
- What is the intellectual model of condition within those choices, i.e., what are the causes of condition and what are the steps for improvement?

As a basis for compiling ecological indicators, a watershed classification system was required for the experimental design. The goal was to develop approximately five categories of watersheds for each physiographic province, utilizing landscape and land use parameters that would be predictive of aquatic resource condition. All 14-digit Hydrologic Unit Code (HUC) watersheds in the Mid-Atlantic region would then be classified according to the regime. Five parameters were utilized for the classification: three land cover categories, consisting of forested, agricultural, and urban, median slope or median elevation, and total variance of land covers in 1-km-radius circles positioned on all stream convergence points in a specified 14-digit HUC watershed. Cluster analysis utilizing these five parameters resulted in approximately five well-defined watershed classes per physiographic province. The distribution of all watersheds in the Mid-Atlantic region across these categories provides a unique report on the probable condition of watersheds in the region.

Wavelet Analysis of Climate-Related Flood Variability

Brandon Whitcher[1], Shaleen Jain[2]
[1]Geophysical Statistics Project, National Center for Atmospheric Research, [2]NOAA-CIRES Climate Diagnostics Center.
P.O. Box 3000, Boulder, CO 80307-3000
Telephone +1 303 497 1709
Fax +1 303 497 1333
Email whitcher@ucar.edu

The footprint of recent climate variability and change is being seen as trends and changing severity of hydrologic extremes (floods and droughts). The observed year-to-year variations and longer-term trends in flood magnitudes have important implications for floodplain management, reservoir operations, resource allocation, and emergency planning. Consequently, an understanding of the climate-related flood risk and the role of some key climate precursors is critical toward developing robust, risk aversive strategies for water resources management and operations. In this work, we use a wavelet-based framework to study the observed variability in western US floods, and explore the overlapping time scales for potential links between large-scale climate processes and floods. Wavelet correlation provides a convenient scale-by-scale decomposition of the association between floods and climate processes. Working with high-frequency data, such as daily observations, both short and long time scales may be analyzed and compared. The decomposition of a process into coefficients associated with a specific time and scale also allows flexible partitioning of the multi-scale content into different regimes. This wavelet-based framework is discussed in the context of understanding important scales of flood-climate correlations, as well as probing the relative role of some key large-scale climate processes (such as, the Madden-Julian Oscillation and El Nino/Southern Oscillation).

A clustering judging criterion and its application to aerial photographs

Matthias Zerbst, Lars Tschiersch
Universität Dortmund Institut für Wirtschafts- und Sozialstatistik, Vogelpothsweg 78
44221 Dortmund, Germany
Telephone +49 231 755 5735
Email zerbst@statistik.uni-dortmund.de

Building homogenous classes is one of the main goals in clustering. Homogeneity can be measured by the intra class- variance (Bock, 1998). Especially in erosion projects but also in other applications as well as the separation between the built classes is important as the homogeneity of the classes. Special cluster methods can be used to reach this aim, for instance the maximum-linkage-algorithm (Zerbst, 2001) or the advanced maximum-linkage-algorithm (Tschiersch, 2002). To judge the separation quality of such clusterings, the shortest distances between all centroids are considered. Zerbst (2001) showed that the arithmetic mean over all distances isn?t good enough for judging selectivity. Therefore the concentration centroid minimum distance criterion is proposed in this paper. This criteroin is based on the ratio of the integral over the minimal distances and the Gini coefficient. It also judges the class separation independent from the underlying data situation. Bock, H.-H. (1998): Clustering and neural networks. In: Rizzi, A.; Vichi, M.; Bock, H.-H. (eds): Advances in data science and classification. Springer, Heidelberg, 265-278. Zerbst, M. (2001): Die pixelbasierte Clusterung von Luftaufnahmen im Rahmen von Erosionsuntersuchungen. Dissertation, Universität Dortmund, Dortmund, Germany. Tschiersch, L.; Zerbst, M. (2002): The advanced-maximum-linkage clustering algorithm Technical Report, Department of Statistics, University of Dortmund, Dortmund, Germany.

TABLE OF CONTENTS