PROGRAMME AND ABSTRACTS

2nd International Workshop on Computational and Financial Econometrics

http://www.dcs.bbk.ac.uk/cfe08/

and

First Workshop of the ERCIM Working Group on Computing & Statistics

http://www.dcs.bbk.ac.uk/ercim08/

University of Neuchâtel, Switzerland
June 19-21, 2008

Address:

Universite de Neuchâtel
UniMail
Rue Emile-Argand 11
CH-2007 Neuchâtel
Switzerland

ERCIM
European Research Consortium for Informatics and Mathematics

SCHEDULE

All lectures take place at the UniMail building (Rue Emile-Argand 11), University of Neuchâtel.

Thursday, 19th June 2008

08:50 - 09:00	Opening (Room: GGA)
09:00 - 09:55	Plenary Talk (Herman K. Van Dijk)
09:55 - 10:15	Coffee Break
10:15 - 12:15	Parallel Sessions B
12:15 - 14:00	Lunch Break
14:00 - 16:00	Parallel Sessions C
16:00 - 16:25	Coffee Break
16:25 - 17:20	Plenary Talk (Oliver Linton)
18:00 - 19:30	Reception

Friday, 20th June 2008

09:00 - 11:00	Parallel Sessions E
11:00 - 11:20	Coffee Break
11:20 - 12:15	Plenary Talk (Bernard Philippe)
12:15 - 14:00	Lunch Break
14:00 - 16:00	Parallel Sessions G
16:00 - 16:25	Coffee Break
16:25 - 18:30	Parallel Sessions H
20:00	Conference Dinner

Saturday, 21st June 2008

09:00 - 11:00	Parallel Sessions I
11:00 - 11:20	Coffee Break
11:20 - 13:20	Parallel Sessions J
13:20 - 15:00	Lunch Break
15:00 - 17:00	Parallel Sessions K
17:00 - 17:20	Coffee Break
17:20 - 18:15	Plenary Talk (Michael Berry)
19:30	Fondue Dinner

SPECIAL MEETINGS by invitation to group members

- COST meeting, Thursday 19th of June, Room GB1, 16:30 - 17:20.
- COST meeting, Friday 20th of June, Room GB1, 08:00 - 08:55.
- CSDA Editorial Board Lunch-Meeting , Friday 20th of June, 12:30 - 14:30.

GENERAL INFORMATION

Lecture Rooms

The paper presentations will take place at the UniMail, University of Neuchâtel. There are ten lecture rooms. Three of them (GGA, GPA and GB1) are in the Chemistry building, while the other seven are in the main building of UniMail. There will be signs indicating the location of the lecture rooms. Please ask for assistance and directions at the registration desk.

The plenary talks will take place in the lecture room GGA (Chemistry building), and will last 55 minutes including questions. Each session will be 2 hours long. Chairs are requested to keep the session on schedule. Papers should be presented in the order in which they are listed in the programme for the convenience of attendees who may wish to switch rooms mid-session to hear particular papers. In the case of a no-show, please use the extra time for a break or a discussion so that the remaining papers stay on schedule.

Presentation instructions

The lecture rooms will be equipped with a PC, a computer projector and in most cases an overhead projector. The session chairs should obtain copies of the talks in a USB stick before the session starts (use the lecture room as the meeting place), or obtain the talks by email prior to the conference beginning. Presenters must deliver the files with the presentation in PDF (Acrobat) or PPT (Powerpoint) format on a USB memory stick to the session chair ten minutes before each session.

The PC in the lecture rooms should be used for presentations. The session chairs should have a laptop for backup.

Swiss plugs/power outlets are different from those in the rest of Europe, including Germany. We cannot provide adapters, so please do not forget to take your adapters if needed.

Internet

There will be access to PCs connected to the Internet at the main entrance of the UniMail. The wireless Internet connection is also freely available at UniMail.

Messages

You may leave messages for each other on the bulletin board by the registration desks.

SUPPORTERS

ERCIM (European Research Consortium for Informatics and Mathematics)

Journal of Computational Statistics & Data Analysis

Elsevier

The Society for Computational Economics

International Association for Statistical Computing

COST Action IC0702 SoftStat

Department of Computer Science, University of Neuchâtel, Switzerland

FINRISK (Financial Valuation and Risk Management), Switzerland

Philips Moris International

Banque Cantonale Neuchâteloise.

the explanatory power of competing Multivariate GARCH specifications, as well as to analyse empirical properties of some bivariate processes on the Warsaw Stock Exchange.

#74: **Aggregation of vector ARMA processes: some further results**

Presenter: Giacomo Sbrana@United Nations, USA

The contemporaneous aggregation of ARMA processes has received considerable attention in the econometric literature. More recently, due to the focus on modelling and forecasting aggregate variables in the Euro-area, many empirical papers have addressed the issue of choosing between aggregated ARMA versus national specific ARMA models. This paper attempts to shed further light on the characteristics of the aggregate ARMA processes. It is shown that the parameters of the macro process can be expressed as direct functions of the micro parameters provided that the micro processes can be expressed as a VMA(1) system. More specifically, summing up across N moving average processes of order one leads to a moving average process of order one whose parameters are exact functions of the micro parameters. Therefore, the forecasting properties of the aggregate process can be easily recovered from the structure of the VARMA generation process. Moreover, the forecasting performances of aggregate and disaggregate predictors are compared. Furthermore, Monte Carlo simulations show that the small sample properties for the aggregate process are particularly good.

#69: **Confidence bands for VAR forecast paths**

Presenter: Anna Staszewska@University of Lodz, Poland

The problem of forecasting from vector autoregressive models has attracted considerable attention in the literature. The most popular non-Bayesian approaches use large sample normal theory or the bootstrap to evaluate the uncertainty associated with the forecast. The literature has concentrated on the problem of assessing the uncertainty associated with the prediction for a single time period. The present paper considers the more important but much less studied problem of how to assess the uncertainty when the forecasts are done for a succession of periods. It describes and evaluates bootstrap methods for constructing confidence bands for forecast paths. Similar methods have already been employed in the construction of confidence bands for impulse response paths. The bands are constructed from forecast paths obtained in bootstrap replications with an optimisation procedure used to find the envelope of the most concentrated paths. The methods are evaluated by means of Monte Carlo experiments performed on a range of DGPs.

ES16 Room: GB1 STATISTICS WITH INCOMPLETE DATA Chair: Gil Gonzalez-Rodriguez

#83: **Pruning decision trees with fuzzy concepts**

Presenter: Matthias Steinbrecher@Otto-von-Guericke-University Magdeburg, Germany
Co-authors: Rudolf Kruse

Decision trees have become a widely spread method of analysing data in business organisations. One reason for this acceptance can be attributed to the intuitive comprehensibility of the results. Since the data-generating processes may change over time, one has to deal with a forest of decision trees that share one part of their structure and differ in another part. A critical task is to select those subtrees that are not only contained in some minimum portion of all trees but also identify all subtrees that exhibit a certain temporal behavior that the user is interested in. We present a method of specifying the behavior of the temporal changes of interest with means of fuzzy, i.e., linguistic concepts. The user is allowed to postprocess the accustomed decision trees with likewise intuitive and interpretable linguistic descriptions.

#84: **Aggregate loss models: a nonparametric approach**

Presenter: Ricardo Cao@Universidade da Coruna, Spain
Co-authors: Juan Vilar, Maria Concepcion Ausin

This paper describes a nonparametric approach to make inference for aggregate loss models in the insurance framework. We assume that an insurance company provides a historical sample of claims given by claim occurrence times and claim sizes. Furthermore, information may be incomplete as claims may be censored and/or truncated. In this context, the main goal of this work consists in fitting a probability model for the total amount that will be paid on all claims during a fixed future time period. In order to solve this prediction problem, we propose a new methodology based on nonparametric estimators for the density functions with censored and truncated data, the use of Monte Carlo simulation methods and bootstrap resampling. An alternative Bayesian approach, using a semiparametric model based on Coxian distributions, has been also proposed. The developed methods are useful to obtain the best strategy in different insurance decision problems. The proposed procedures are illustrated with a real data set provided by the insurance department of an international commercial company.

#85: **A possible extension of upper and lower probabilities to the case of fuzzy random variables**

Presenter: Wolfgang Trutschnig@Vienna University of Technology, Austria

In many situations a real-valued random variable X can not be observed precisely but it may be possible to observe a random interval I or a fuzzy random variable X* that contains the true random variable with probability one. The question arises, in which way the distribution of the random variable X can be approximated by means of the observable random interval or the observable fuzzy random variable. In case of a random interval one possible answer is given by Dempster's well-known lower and upper

probabilities. It will be shown that DempsterÂ´s concept can easily be extended to the case of fuzzy random variables, which yields so-called fuzzy-valued probabilities. The main properties of these generalized probabilities will be mentioned and it will be shown that as in the classical, real-valued case these probabilities can be regarded as limits of relative frequencies. Furthermore related open problems will be stated.

#99: **Hypothesis testing about the means of fuzzy random variables.**

Presenter: . . . . . Ana Colubi@University of Oviedo, Spain
Co-authors: . . . . . M. Angeles Gil

Fuzzy random variables (FRVs for short) model random mechanisms associate with each experimental outcome imperfect values. That is, FRVs are considered to manage random experiments involving imprecisely-valued characteristics, so that data available from the experimental performance can be properly described by means of fuzzy sets. FRVs are particularly useful in handling the imprecision underlying many real-life perceptions, classifications and judgements. In the last years an inference statistical methodology concerning fuzzy random variables (especially estimation and testing about the fuzzy mean value) has been developed. In this communication we focus on the problem of testing about the fuzzy mean of an FRV. Methods for the one-sample, two-sample and multi-sample (ANOVA) problem are discussed. The procedures are illustrated by means of real-life example and empirical studies are discussed.

#102: **Fuzzy techniques in the analysis of distributions of real random variables.**

Presenter: . . . . . Gil Gonzalez-Rodriguez@European Center for Soft Computing, Spain

Some random variables which are traditionally modelled as ordinal variables or simply coded as a real-valued one can be intuitively and more expressively modelled by means of fuzzy random variables. This is the case of the forest fire index, which is usually fixed to range in a discrete scale from 1 to 5 (1 meaning risk absence and 5 being associated with the maximum risk). The different nature of the extreme values along with the lack of precision underlying the discretization process suggest the possibility of representing them by means of adequate fuzzy sets capturing these features. On the other hand, the empirical comparison of the hypothesis testing about means for fuzzy random variables and real-valued ones, allows us to conclude that in most of the cases conclusions are definitely more powerful for the first ones. Motivated by these two arguments, some suitable fuzzifications of real-valued random variables have been suggested. Especially interest are paid to those for which the fuzzy mean value of the transformed original variable fully characterizes its distribution. In this communication, some of these fuzzifications are presented and a discussion about their statistical implications is considered.

CS01 Room: E003 TIME SERIES AND FINANCIAL ECONOMETRICS Chair: Christian Francq

#15: **Asymptotic properties of sample inverse autocorrelations under weak assumptions**

Presenter: . . . . . Ahmed El Ghini@Universite Lille 3, France

The inverse autocorrelation function has been widely used in time series literature. The function is defined via the inverse of the spectral density, and plays an important role in the identification and estimation of ARMA models. Most of the studies assume that the innovations of the linear time series are independent and identically distributed. On the other hand, there are many applications in which this strong assumption is either questionable or clearly inadequate. For example, most empirical time series in business and finance exhibit nonlinearity and conditional heteroscedasticity. This paper is devoted to asymptotic properties of the estimates of the inverse autocorrelation function derived by the orthogonality method. Under weak dependence assumptions, it is established that the estimates are consistent and asymptotically normal. An application to linear processes with GARCH innovations is discussed. A Monte Carlo study illustrates the theoretical results and shows that the method performs well in finite samples, for various kinds of non-linear processes. An application to a real financial data is provided.

#9: **Econometric asset pricing modelling**

Presenter: . . . . . Fulvio Pegoraro@Banque de France, France
Co-authors: . . . . . Henri Bertholon, Alain Monfort

The purpose of this paper is to propose a general econometric approach to asset pricing modelling based on three main ingredients : (i) the historical discrete-time dynamics of the factor representing the information, (ii) the Stochastic Discount Factor (SDF), and (iii) the discrete-time risk-neutral (R.N.) factor dynamics. Retaining an exponential-affine specification of the SDF, its modelling is equivalent to the specification of the factor loading vector and of the short rate, if the latter is neither exogenous nor a known function of the factor. In this general framework, we distinguish three modelling strategies: the Direct Modelling, the Risk-Neutral Constrained Direct Modelling and the Back Modelling. In all the approaches we study the internal consistency constraints, implied by the absence of arbitrage opportunity (AAO) assumption, and the identification problem. We also propose interpretations of the factor loading vector in terms of market price of risk. The general modelling strategies are applied to two important cases: security market models and term structure of interest rates models. In the context of security market models, we show the relevance of our methods for various kinds of specifications: switching regime models, stochastic volatility models, Gaussian and Inverse Gaussian GARCH-type models (with or without regime-switching). In the interest rates modelling context, we consider several illustrations: VAR modelling, Switching VAR modelling and Wishart modelling. We also propose, using a Gaussian VAR(1) approach, an example of joint modelling of geometric returns, dividends and short rate. In these contexts we stress the usefulness

Thursday 19.06.2008 14:00-16:00 Parallel Session C

ES06 Room: GGA COMPUTATIONAL METHODS FOR MIXTURES Chair: Sylvia Fruehwirth-Schnatter

#19: **Finite mixture model diagnostics using the bootstrap**

Presenter: Bettina Grun@Wirtschaftsuniversitat Wien, Austria
Co-authors: Friedrich Leisch

The EM algorithm provides a common framework for maximum likelihood estimation of finite mixture models. The fitted models can differ with respect to the component specific models and may also allow for concomitant variables to model the component weights. The use of resampling methods to analyze finite mixture models fitted with the EM algorithm is appealing because the bootstrap similarly to the EM algorithm constitutes a common framework for these models. In addition standard asymptotic theory can sometimes be not directly applied to finite mixtures due to violation of regularity conditions. In this talk we will outline various possibilities to use bootstrapping for model diagnostics such as for determining the number of components, checking model identifiability and analyzing the stability of induced clusters. The application of the diagnostic tools is demonstrated on several examples.

#15: **A flexible prior distribution for markov switching autoregressions with student-t errors**

Presenter: Philippe Deschamps@University of Fribourg, Switzerland

This paper proposes an empirical Bayes approach for Markov switching autoregressions that can constrain some of the state-dependent parameters (regression coefficients and error variances) to be approximately equal across regimes. By flexibly reducing the dimension of the parameter space, this can help to ensure regime separation and to detect the Markov switching nature of the data. The permutation sampler with a hierarchical prior is used for choosing the prior moments, the identification constraint, and the parameters governing prior state dependence. The empirical relevance of the methodology is illustrated with an application to quarterly and monthly real interest rate data.

#14: **On mixture of Kalman filtering and learning**

Presenter: Hedibert Lopes@University of Chicago, USA
Co-authors: Nicholas Polson, Carlos Carvalho, Michael Johannes

In this paper we present a novel particle filtering and learning strategy for a wide class of state space models that can be represented as mixture of dynamic linear models (DLMs). These methods provide samples from the joint posterior distribution of states and parameters, in a sequential fashion, avoiding the burden of hard to converge MCMC samplers. Our methodology provides an extension to the mixture of Kalman filters and naturally incorporates nonlinearities in the state dynamics. We use conditional sufficient statistics for parameter learning and we extend this approach to state filtering whenever possible. We provide two applications. First, a dynamic factor switching model which illustrates the efficiency gains over traditional methods. Second, we analyze a nonlinear model that has been extensively considered in the pure filtering literature, where we also add sequential parameter learning.

#3: **Bayesian estimation of finite mixtures of univariate and multivariate skew-normal and skew-t distributions**

Presenter: Sylvia Fruehwirth-Schnatter@Johannes Kepler University Linz, Austria

Skewnormal and skew-t densities, both for univariate as well as multivariate data sets, have been introduced with the goal of capturing skewness and kurtosis without loosing unimodality of the fitted distribution. Very recently, finite mixtures of such densities have been introduced for the purpose of robust clustering. Rather little work has been done on efficient statistical estimation of such mixtures and in the present paper Bayesian inference is carried out. For mixtures of univariate and multivariate skewnormal densities, we develop MCMC estimation based on data augmentation and Gibbs sampling. The first step of data augmentation and the Gibbs sampler is based on the standard procedure of drawing classification using the skew components densities. To carry out parameter estimation within each component, we use a second step of data augmentation based on a stochastic representation of the uni- and the multivariate skewnormal density in terms of a random-effect models with truncated normal random effects. This allows drawing the parameters from standard density. This MCMC scheme is extended to univariate and multivariate mixtures of skew-t densities through a third step of data augmentation based on representing the t-density as scale mixtures of normals.

ES17 Room: GB1 PROBABILISTIC METHODS IN LEARNING PROBLEMS Chair: Ana Colubi

#67: **Probabilistic fuzzy systems in financial modelling**

Presenter: Uzay Kaymak@Erasmus University Rotterdam, Netherlands

Probabilistic fuzzy systems (PFS) are semi-parametric models in which a linguistic description of the system behaviour encoded by the fuzzy rules can be combined with the statistical properties of the data. Mathematically, they are related to neural networks, support vector machines, kernel models and Parzen window density estimators. They are interesting as they allow the modeler to focus on the experts' information coded linguistically while the model itself has strong theoretical grounding. All common fuzzy models can be extended to their probabilistic fuzzy equivalents. In this paper we study probabilistic fuzzy equivalents of Mamdani fuzzy systems and zero-order Takagi-Sugeno fuzzy systems. We consider their main characteristics and discuss how they can be

applied in financial modelling. In particular, attention is paid to value-at-risk estimation by using probabilistic fuzzy systems. A sequential approach is proposed for determining the model parameters, where the location of the antecedent membership functions is determined by using fuzzy clustering while maximum likelihood parameter estimation is used for determining the probability parameters of the PFS. The validity of the VaR models obtained is evaluated by using a statistical back-testing method (Kupiec test) based on failure rates.

#49: **Parameter-free feature selection with mutual information**

Presenter: . . . . . Michel Verleysen@Universite Catholique de Louvain, Belgium
Co-authors: . . . . . Damien Francois

Machine learning of high-dimensional data faces the curse of dimensionality, a set of phenomena that limit the performance of the tools. Many limitations come directly from the representation of the data, and not from the analysis tool. It is therefore needed to reduce the data dimensionality. There are basically two ways to do this: either to select features among the original variables, or to project the latter on new ones. Although more general and thus more powerful in theory, projecting features induces a loss of interpretability. On the contrary, by selecting original features, one can come back to the application and interpret which are the relevant factors for the analysis; this is important advantage in many applications. This paper shows how to use Mutual Information (MI) for feature selection. In practice, the MI criterion has to be estimated and the search for possible feature subsets restricted for computation time reasons. It is shown how to use resampling and permutation tests to select optimal parameters for the estimator, and to stop the search procedure in a sound way. It is also shown how to design an estimator of feature subset relevance inspired from the mutual information criterion, with the supplementary advantage to restrict the estimation to a two-dimensional problem.

#87: **Fuzzy text mining and digital obesity**

Presenter: . . . . . Trevor Martin@University of Bristol, UK
Co-authors: . . . . . Yun Shen

The phrase *digital obesity* summarises a range of problems arising from our propensity to generate and retain a rapidly growing volume of data, at web-scale as well as at corporate and personal scales. Much of this data is in text form, but is effectively wasted unless we can find and use the *right* data when needed. Statistical methods help to a degree, but tend to *average out* useful information, as well as suffering from a mismatch between the precisely defined terms used by formal models and the far more subtle and expressive terms used in human communication. Humans communicate using language where the majority of concepts are fuzzy, defined by common usage rather than by necessary and sufficient conditions. The success of fuzzy control is one example where fuzzy set theory enabled computers to work with ill-defined terms such as *hot* and *slow* rather than precise values. Fuzziness enables computers to work with ill-defined concepts, leading to more effective use of text-based information in business and other situations. Although the input information is rarely complete (and may be incorrect) the approximately correct solutions are generally sufficient as well as being easier to compute and understand.

#96: **The estimation of prediction error for neural networks: a simulation study.**

Presenter: . . . . . Simone Borra@University of Rome ”Tor Vergata”, Italy
Co-authors: . . . . . Agostino Di Ciaccio

One fundamental problem in statistics is that of obtaining an accurate estimate of the prediction error, i.e. the expected loss on future observations, of a learning algorithm trained on the available sample data. This problem has particular relevance every time a very large sample is not available, the underlying distribution is not known and you need to evaluate the prediction error of a non-parametric model which could overfit data. The simplest estimator of prediction error is the Apparent Error defined as the average of the loss function on the training data-set. Apparent error usually produces an optimistic estimate of prediction error because it uses the same data both for training and for evaluation of the model. Using powerful non-linear models, as Neural Networks, it is possible to obtain very small values of Apparent Error, just including more parameters in the model. A way to evaluate the prediction error of the model is to estimate the Optimism, defined as the expected difference between the prediction error and the Apparent Error on new training data, adding it to the Apparent Error. We considered several approaches to prediction error estimation for Neural Networks. In particular, estimators based on Cross-Validation (as Leave-one-out, K-fold cross-validation) and Repeated Cross-Validation (obtained averaging a set of cross-validation estimates on different random split); estimators based on non-parametric Bootstrap (as 0.632 bootstrap and the modified version 0.632+ to take into account situations of severe overfit) and parametric Bootstrap (where the Optimism is proportional to a covariance term estimated by Bootstrap). Using an extensive simulation approach we were able to compare the estimators with respect to different characteristics of data. We considered a regression problem with 1000 data generating distributions showing different level of non-linearity and signal/noise ratio. In each population, we drew 30 samples on which we trained two different NN, calculating also all estimators of prediction error. We generated also a very large sample from each population, to obtain a reliable estimation of the true prediction error for each NN. Finally, we compared all prediction error estimators on the bases of bias and variability. We obtained some interesting suggestions about the efficiency of the different prediction error estimators with respect to the s/n ratio and the neural network complexity.

#114: **Application of neural networks and support vector machines to pricing European options**

Presenter: . . . . . Chris Charalambous@University of Cyprus, Cyprus

Artificial Neural networks (ANN), as discipline, studies the information processing capabilities of networks made up of simple processors which are in some way connected with different strengths (weights) like the living neurons of the brain. During the

training phase the connecting weights are modified so that the network output matches the required response (target values) as closely as possible. The multilayer neural network is the most widely used type of neural networks. It consists of an input layer a number of hidden layer and an output layer. There are several issues involved in designing a multilayer neural network: Selecting the number of hidden layers, the number of neurons in each hidden layer, finding the global solution. Support Vector Machine (SVM) models are close cousin to classical multilayer neural networks and overcome the above limitations of neural networks. The weights of the network are found by solving a positive definite quadratic programming problem, rather than by solving a non convex, unconstrained minimization problem as in standard neural network training. In this talk we will present an overview of both ANNs and SVMs and explore their performance in pricing European options. Furthermore, we will show how neural networks can be used to modify some input parameters of an analytical option pricing model so that its pricing performance can greatly be improved.

ES22 Room: GPA ROBUSTNESS WITH HIGH DIMENSIONAL DATA Chair: Stefan Van Aelst

#7: **Fast bootstrap for robust Hotelling tests**

Presenter: Ella Roelant@Ghent University, Belgium
Co-authors: Stefan Van Aelst, Gert Willems

If we want to estimate the location and scatter of a multivariate data set with outliers, the sample mean and sample covariance matrix will no longer be satisfactory as they can be extremely sensitive to outliers. As robust alternatives, we will focus on S-estimators and MM-estimators which are efficient, positive breakdown estimators. Inference for robust estimators is often based on the asymptotic distribution of these estimators. However, as asymptotic estimates may be inaccurate, the bootstrap can be used as an alternative approach. Unfortunately, the standard bootstrap procedure is non-robust and computationally demanding. Both these problems are resolved by the fast and robust bootstrap (FRB) procedure. This method exploits the property that robust estimators such as S- and MM-estimators can be written as the solution of a system of smooth fixed-point equations. We consider a robust version of the one-sample and two-sample Hotelling test by using S- or MM-estimators instead of the empirical mean and covariance matrix. The FRB can then be used to mimic the distribution of the test statistic and critical values can be determined through the quantiles of the recalculated statistics. Simulations show good performance and illustrate that the bootstrap can outperform the asymptotic variance approach.

#11: **Stahel-Donoho estimators with cellwise weights**

Presenter: Gert Willems@Ghent University, Belgium
Co-authors: Ellen Vandervieren, Stefan Van Aelst

The Stahel-Donoho estimator is a well-known affine equivariant robust estimator of multivariate location and scatter. It is defined as a weighted mean and covariance, where each point receives a weight in function of a measure of its *outlyingness*. This measure is based on the one-dimensional projection in which the point is most outlying, the underlying idea of which is that every multivariate outlier must be a univariate outlier in some projection. Points with large outlyingness should then receive small weights. In the sense that the entire point is either downweighted or not, all components of the point are treated in the same way irrespective of their *responsibility* for the outlyingness. Here we investigate an adaptation of the Stahel-Donoho estimator where we allow separate weights for each component. The idea is to start from the outlyingness of the point as measured in the original Stahel-Donoho procedure. Subsequently, for each point, we attempt to identify to what extent each variable is contributing to the outlyingness and we use this information to adjust the outlyingness and corresponding weight in a componentwise manner. By adapting the estimator in this way, we are giving up affine equivariance but we may gain efficiency.

#6: **Outlyingness weighted quadratic covariation**

Presenter: Kris Boudt@K. U. Leuven, Belgium
Co-authors: Christophe Croux, Sebastien Laurent

Quadratic covariation is a natural estimator for the volatility of a multivariate price process. It is consistently estimated by the sum of outer products of high-frequency returns. The Realized BiPower Covariation (RBPCov) is often used to estimate the quadratic covariation of the continuous component of the price diffusion. This paper introduces the Realized Outlyingness Weighted Quadratic Covariation (ROWQCov) as an alternative to the RBPCov. The new estimator equals a weighted sum of outer products of high-frequency returns and downweights returns that, because of jumps or other reasons, are outliers under the Brownian SemiMartingale (BSM) model. Under this model the ROWQCov is consistent for the integrated covariance matrix. Besides robustness to jumps, the new estimator also enjoys the desirable properties of positive semidefiniteness and affine equivariance. Our Monte Carlo study suggests that, at all sampling frequencies, the new estimator is more efficient than the RBPCov under the BSM model, consistent under the BSM model with jumps in cases where the RBPCov is no longer consistent and its robustness to jumps is preserved under temporal aggregation. We illustrate this new method on 15-minute return series of the EUR/USD and GBP/USD exchange rates.

#64: **On non-parametric robust quantile regression by support vector machines**

Presenter: Andreas Christmann@University of Bayreuth, Germany
Co-authors: Ingo Steinwart

We consider the non-parametric estimation of quantile functions by support vector machines (SVMs) based on the pinball loss

#31: **Autocorrelation based tests for vector error correction models with uncorrelated but non independent errors**

Presenter: Hamdi Raissi@University Lille 3, France

We consider in this paper the estimation and test-of-fit for vector error correction models with non independent innovations. The asymptotic properties of the residual sample autocorrelations are derived. It is shown that the asymptotic distribution can be quite different for models with iid innovations and models in which the innovations are non independent. Consequently, the usual chi-square distribution does not provide an adequate approximation of the distribution of the Box-Pierce goodness-of-fit portmanteau statistic in the presence of non independent innovations. We thus propose a modified portmanteau test whose asymptotic distribution is a weighted sum of independent chi-squared random variables. We also propose a modified Lagrange multiplier test. Monte Carlo experiments illustrate the finite sample performance of the different tests.

#32: **Regenerative block empirical likelihood for Markov chains**

Presenter: Hugo Harari-Kermadec@Universite Paris-Dauphine, France

Empirical likelihood is a powerful semi-parametric method leading to estimation, test and confidence intervals. Many extensions of this method have been proposed in recent years. However they essentially focus on an i.i.d. setting. In the case of dependent data, the empirical likelihood method cannot be directly applied on the data but rather on blocks of consecutive data catching the dependence structure. Generalization of empirical likelihood based on the construction of blocks of increasing nonrandom length have been proposed for time series satisfying mixing conditions. Following some recent developments in the bootstrap literature, we propose a generalization for a large class of Markov chains, based on small blocks of various lengths. Our approach makes use of the regenerative structure of Markov chains, which allows to construct blocks which are almost independent.

ES18 Room: GB1 INTELLIGENT DATA ANALYSIS Chair: Christian Borgelt

#56: **Learning from data with soft class labels using mixture models and belief functions**

Presenter: Etienne Come@Universite de Technologie de Compiegne, France
Co-authors: Thierry Denoeux, Latifa Oukhellou, Patrice Aknin

This paper addresses classification problems in which the class membership of training data is only partially known. Each learning sample is assumed to consist in a feature vector x_i and an imprecise and/or uncertain *soft* label m_i defined as a Dempster-Shafer basic belief assignment over the set of classes. This framework thus generalizes many kinds of learning problems including supervised, unsupervised and semi-supervised learning. Here, it is assumed that the feature vectors are generated from a mixture model. Using the General Bayesian Theorem, an extension of Bayes' theorem in the belief function framework, we derive a criterion generalizing the likelihood function. A variant of the EM algorithm dedicated to the optimization of this criterion is proposed, allowing us to compute estimates of model parameters. Experimental results demonstrate the ability of this approach to exploit partial information about class labels.

#72: **Probabilistic noise clustering as M-estimators**

Presenter: Frank Klawonn@University of Applied Sciences BS/WF, Germany

Probabilistic or fuzzy clustering approaches use weights to assign data to clusters. Depending on the parametrisation of the clusters and the distance function used for clustering, various cluster shapes are possible ranging from simple spherical clusters to clusters described by hyperplanes or quadrics. Probabilistic clustering can be viewed as a class of M-estimators known from robust statistics. Although various investigations concerning robustness issues in probabilistic clustering are available, recent developments have not been taken into account. Especially in the context of noise clustering establishing a close connection to robust regression can be established leading to new interesting weighting functions to control the properties of the robust estimator.

#75: **Multi-criteria ant feature selection in intelligent classification**

Presenter: Joao M. C. Sousa@Technical University of Lisbon, Instituto Superior Tecnico, Portugal
Co-authors: Susana Vieira

This paper proposes a multi-criteria ant colony optimization (ACO) algorithm for feature selection using intelligent (neural or fuzzy) classifiers. The proposed algorithm deals with the feature selection problem as a multi–criteria problem with a single objective function. The two criteria considered are the size of the subset of features (features cardinality), and the performance of the classifier, which is build based on the selected features. A pheromone matrix is used for each criterion, and different heuristics for the two criteria are used. In order to study the influence of the parameters and to establish the most suitable values for such parameters, the ANalysis Of the VAriance (ANOVA) statistical method is used. Experiments show the significance of parameters concerning the classification error and the number of features. The performance of the proposed multi-criteria algorithm is compared to the performance of an ant feature selection algorithm based only on one criterion (improving classification performance). The results show the advantage of using the multi-criteria algorithm.

#97: **Accelerating fuzzy clustering**

Presenter: Christian Borgelt@European Center for Soft Computing, Spain

Extensions of earlier work on an approach to accelerate fuzzy clustering by transferring methods that were originally developed to speed up the training process of (artificial) neural networks are presented. The core idea of this approach is to consider the

difference between two consecutive steps of the alternating optimization scheme of fuzzy clustering as providing a gradient. This gradient may be modified in the same way as the gradient of (artificial) neural network back propagation is modified in order to improve the training. Even though these modifications are, in principle, directly applicable, carefully checking and bounding the update steps can improve the performance and can make the procedure more robust. In addition, this talk provides a much more detailed experimental evaluation that is based on cluster comparison measures, which can nicely be used to study the convergence speed.

CS05 Room: B217 MARKOV-SWITCHING MODELS FOR FINANCIAL RETURNS Chair: Alessandra Amendola

#14: **Bayesian estimation of a Markov-switching threshold GJR model**

Presenter: David Ardia@University of Fribourg, Switzerland

A Bayesian estimation of a Markov-switching threshold GJR(1,1) model is proposed. The specification is based on a regime-switching model with parallel asymmetric GJR models where asymmetries are centered at free threshold parameters. The model aims at determining (i) whether structural breaks are present within the GARCH dynamics; (ii) whether GARCH asymmetries (i.e., leverage effects) are present, and if they are different between the regimes; (iii) if the threshold parameters (i.e., locations of bad news) are similar between the regimes. The MCMC estimation scheme allows a fully automatic Bayesian estimation of the model and thus, avoids the difficult task of choosing and tuning a sampling algorithm. The presence of two distinct volatility regimes is shown in an empirical application to SMI log-returns. Moreover, the results indicate no difference between asymmetries and locations of the asymmetry for highly volatile and tranquil periods. The performance of the model is compared to a single-regime specification and document a better fit and an improvement of the forecasting ability for the Markov-switching model.

#106: **Forecasting volatility under fractality, regime-switching, long memory and Student-t innovations**

Presenter: Leonardo Morales-Arias@University of Kiel, Germany
Co-authors: Thomas Lux

In this paper we examine the forecasting performance of volatility models that incorporate features such as long (short) memory, regime-switching and multifractality along with two competing distributional assumptions of the error component, i.e. Normal vs. Student-t. Our precise contribution is twofold. First, we introduce a new model to the family of Markov-switching multifractal models of asset returns (MSM), namely, the Markov-switching multifractal model of asset returns with student-t innovations (MSM-t). This model is an extension of the MSM model with normal innovations and can be estimated via Maximum Likelihood or GMM. We investigate the in-sample as well as the out-of-sample performance of this model via Monte Carlo simulations and compare it vis-a-vis other existing models (MSM, GARCH and GARCH-t). Second, we perform a comprehensive in-sample and out-of-sample cross-sectional analysis of the MSM models (binomial MSM, binomial MSM-t, lognormal MSM, lognormal MSM-t) as well as other competing volatility models (GARCH, GARCH-t, FIGARCH and FIGARCH-t). Our cross-sections consist of all-share equity portfolios, bond indices and portfolios of real estate at the country level. Furthermore, we investigate whether there is an improvement upon singular forecasts when optimally combining forecasts obtained from the different models at hand.

#30: **Modeling international financial returns with a multivariate regime switching copula**

Presenter: Alfonso Valdesogo Robles@CORE-Universite Catholique de Louvain, Belgium

In order to capture observed asymmetric dependence in international financial returns, we construct a multivariate regime-switching model of copulas. We model dependence with one Gaussian and one canonical vine copula regime. Canonical vines are constructed from bivariate conditional copulas and provide a very flexible way of characterizing dependence in multivariate settings. We apply the model to returns from the G5 and Latin American regions, and document two main findings. First, we discover that models with canonical vines generally dominate alternative dependence structures. Second, the choice of copula is important for risk management, because it modifies the Value at Risk (VaR) of international portfolio returns.

#124: **Joint forecasts of Dow Jones stocks under general multivariate loss function**

Presenter: Matei Demetrescu@Goethe University Frankfurt, Germany
Co-authors: Tansel Alp

Univariate asymmetric loss functions have been considered in the forecasting literature. But it is not clear what general conditions multivariate loss functions should fulfill; and there is no simple asymmetric multivariate loss function available, either. Our contributions are as follows. We suggest a flexible class of multivariate loss functions based on suitable combinations of univariate loss functions. To estimate the forecast distributions of daily returns of 30 DJIA stocks, we employ a state-of-the-art multivariate GARCH model. It easily copes with large number of series while allowing for non-ellipticity, fat tails and tail-dependence. Based on Engle's DCC GARCH, the model employs multivariate affine normal inverse Gaussian distributions as conditional probability laws. Thus, the number of parameters to be estimated simultaneously does not depend on the number of series. We check the finite-sample properties of multi-step quasi-ML estimators for our non-Gaussian model. These behave roughly the same as the multi-step estimators for the Gaussian DCC model in sample sizes typically available for most financial time series. We fit our model with daily data from 2002 to 2007 (keeping data from 2008 for out-of-sample analyses), and use a parametric bootstrap procedure to derive point forecasts under a number of different multivariate loss functions. We also find CCC-type models to be inferior in terms of mean forecast loss.