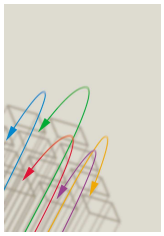
BELGIAN SCIENCE POLICY



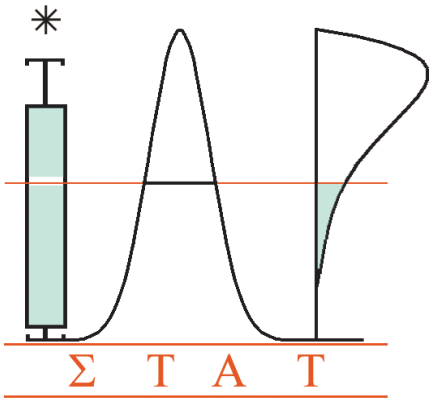
Progress Report 2011

IAP-Network in Statistics

Contract P6/03

May 10, 2012

http://www.stat.ucl.ac.be/IAP/PhaseVI



Contents

1 Accomplished Research Projects

1.1 Introduction

The research network consists of 5 Belgian partners and 4 European partners, which are given in Table 1 below.

Abbreviation	Partner
UCL	Université catholique de Louvain
KUL-1	Katholieke Universiteit Leuven 1
KUL-2	Katholieke Universiteit Leuven 2
UG	Universiteit Gent
UH	Universiteit Hasselt
UJF	Université Joseph Fourier
EMC	Erasmus Medical Center
USC	Universidad de Santiago de Compostela
LSHTM	London School of Hygiene and Tropical Medicine

Table 1: *Belgian and European partners of the network.*

The research project has been built up around five work packages. Table 2 below gives the *main* contributors to each work package and indicates per package the partner that is coordinating the work.

Work package	Contributing partners
WP1: Multivariate data with qualitative constraints	UCL, KUL-1*, UH, UJF, EMC, USC
WP2: Temporally and spatially related data	UCL*, KUL-1, UG, UJF, EMC, USC
WP3: Incomplete data	UCL, KUL-1, KUL-2, UG, UH*, USC, LSHTM
WP4: Data with latent heterogeneity	UCL, KUL-1, KUL-2*, UH, UJF, EMC, USC
WP5: Highdimensional and compound data	UCL, KUL-1, KUL-2, UG*, UH, UJF, EMC, USC

Table 2: *Main contributors per work package, and coordinating partner per work package (indicated with a *).*

In the subsections below we describe the progress that has been made in the various work packages. For each of the work packages we indicate interactions with research results in other packages. The references mentioned in the text can be found at the end of this report.

1.2 Work package 1: Multivariate data with qualitative constraints

Boundaries, frontiers, and efficiency and productivity analysis

During the year 2011, research on the topic of estimation of a support of a (univariate or multivariate) density function, and of estimation of a more general support continued.

A particular support function is a frontier function or production function, with applications in e.g. economics. Among the classical estimators for a frontier function are the Data Envelopment (DEA) estimator and the Free Disposal Hull (FDH) estimator. The study of bootstrap procedures for DEA estimators has been carried out by researchers at the UCL group. The bootstrap can be based on a complex "double smoothing" procedure (smoothing of the estimated frontier and of a joint multivariate density in the input/output space) or by using subsampling. See for example Simar and Wilson (2011a,b,c), and Kneip, Simar and Wilson (2011). A discussion of stochastic FDH and DEA estimators for frontier analysis is in Simar and Zelenyuk (2011).

Issues of robustness in nonparametric partial frontier modeling are discussed in Daouia and Gijbels (2011a). The concept of extremiles was exploited in Daouia and Gijbels (2011b) when estimating frontier cost models.

Nonparametric and semiparametric estimation of curves and surfaces, and estimation under qualitative constraints

Unknown functions and surfaces are not always smooth, and may show different types of irregularities in different regions. In Desmet and Gijbels (2011) local linear techniques are used to develop a method for estimation of curves that exhibit jump and peak irregularities.

Another approach to deal with inference for curves with irregularities is to use penalized regression techniques with an appropriate choice of the penalty function. In Antoniadis, Gijbels and Nikolova (2011) nonparametric regression estimation is considered in extended generalized linear models using penalized splines with non-quadratic penalties. This unified approach allows to study the optimization problem and asymptotic properties of the resulting estimator for a whole class of penalties. Penalized splines estimation and various grouped regularization techniques in additive varying coefficient models are discussed in Antoniadis, Gijbels and Lambert-Lacroix (2011). Antoniadis, Gijbels and Verhasselt (2011, 2012) focus on variable selection in additive models and varying coefficient models using penalized splines estimation combined with the nonnegative garrote method.

Often the interest is not only in estimating a mean function, but also in estimating a variance function, or more generally a dispersion function. Gijbels and Prosdocimi (2012) illustrate the importance of estimating both, mean and dispersion function, in an analysis of data concerning induced abortion in Italy. This analysis also motivated the study of robust methods in this context.

When modeling dependencies through copulas, estimation of constrained copulas comes into play when studying classes of specific dependence structures (such as e.g. positive quadrant dependence). Gijbels and Sznajder (2011a,b) study estimation of copulas under the qualitative constraints of positive quadrant dependence and left tail decreasingness, among others, in their papers on testing for specific dependence structures.

Nonparametric and semiparametric testing and estimation procedures

When studying multivariate data, it is often of crucial importance to know if groups of observations (e.g. defined dy different treatments) differ in their relative dispersions. Gijbels and Omelka (2011) develop new tests for homogeneity of dispersions.

A particular class of semiparametric modeling is the class of single-index models. Single-index modeling is used to study conditional probabilities in two-way contingency tables in Geenens and Simar (2011).

Nonparametric estimation for functional data analysis is a challenging research area. A new semimetric for such analysis is proposed in Timmermans, Delsol and von Sachs (2011). Slaets, Claeskens and Hubert (2012) discuss phase and amplitude-based clustering techniques for functional data.

Multivariate data, robust analysis and nonparametric inference

Many univariate robust estimators are based on quantiles. Several outlier detection methods use a robust measure of skewness such as the medcouple. The latter is defined in terms of quantiles. The use of smoothing techniques in order to reduce the mean squared error of such quantile-based estimators is the aim of the paper by Hubert, Gijbels and Vanpaemel (2011).

A review on robust statistics for outlier detection is provided in Rousseeuw and Hubert (2011).

Croux, Gijbels and Prosdocimi (2012) consider extended generalized additive models, and develop methods for robust estimation of mean and dispersion functions in this multivariate setting. Lambert-Lacroix and Zwald (2011) study robust regression using Huber's criterion and an adaptive lasso penalty.

Modelling and measuring of dependencies and copula functions

Quite some research efforts in the network the past year dealt with the study of modeling dependencies through copula functions. In particular researchers from KUL-1, UCL and UHasselt are involved in this research. In October 2011 Dominik Sznajder (KU Leuven) defended a Ph.D. thesis on this topic. Denuit and Mesfioui (2011) focus on the Archimedean family of copulas, and compare conditional distributions derived under these.

A way to model conditional dependencies is through conditional copula functions. Nonparametric estimation of conditional copulas has been studied in Veraverbeke, Omelka and Gijbels (2011), Gijbels, Veraverbeke and Omelka (2011) and Gijbels, Omelka and Veraverbeke (2011), whereas semi-parametric estimation of conditional copulas using local likelihood ideas and local polynomial fitting was dealt with in Abegaz, Gijbels and Veraverbeke (2011).

Extreme-value copulas are used for studying dependencies involved in extremal events. Nonparametric estimation of an extreme-value copula is discussed in Gudendorf and Segers (2011), whereas tests for extreme-value dependence for multivariate copulas has been developed in Kojadinovic, Segers and Yan (2011). A Bayesian approach to bivariate extremes can be found in Guillotte, Perron and Segers (2011). Among the most known elliptical copulas are the Gaussian and t-copulas. The tails of correlation mixtures of such copulas are studied in Manner and Segers (2011), and herein applied to a case study involving financial time series.

Interactions with other Work packages

Copulas are used to model dependencies between variables. Time or spatial dependencies are dependencies of a special nature, and as such there is a clear connection with WP2. When dealing with current status data, censored data or missing data in WP3, semi-and nonparametric techniques, such as kernel estimation, local polynomial fitting, estimation techniques for location/scale models are also used. In WP1 methods for selecting variables in flexible regression models have been developed. Selection of important variables among a large set of variables is a topic that interacts with some topics in WP5.

1.3 Work package 2: Temporally and spatially related data

Taking its motivation from modelling complex financial or psychological processes this work package focuses on the development and investigation of feasible models for temporally and spatially related data, in discrete or continuous times and in one or higher dimensions. A particular emphasis is on non- and semi-parametrically efficient methods, on-line estimation and automatic model building under non standard assumptions such as violations of stationarity or homogeneity. The main achievements of the research network for temporally and/or spatially related data can be subdivided into the categories: complex univariate time series data, multivariate time series data, continuous time models and spatially related data.

Main contributions in the context of univariate correlated data have been made in case of non stationarity due to structural breaks or time-varying spectra. In the presence of local stationarity the estimation of the time-varying long memory parameter is considered in Roueff and von Sachs (2011). In Eichler et al. (2011) dynamic factor models are fitted to non-stationary time series. On the other hand using local linear/polynomial fitting Croux *et al.* (2011) derive methods for robust forecasting in non-stationary time series while non-parametric estimation of a spectral density with improved estimation at the peaks is dealt with in Desmet and Gijbels (2011). Local linear fitting techniques (in a non time series setup) are among the methods studied in detail and applied in various contexts in WP1. In case of stationarity new developments include statistical inference for time series with irregular spectra, for their extremes, under censoring and for non linear processes. Density estimation and associated problems in nonparametric regression for dependent data where the observations do not necessarily come from a linear process are discussed in Johannes and Subba Rao (2011). Censoring of time series data is another problem encountered with real data. In nonparametric location-scale regression models Van Keilegom et al. (2011) consider the estimation based on censored data, while Cetinyurek and Lambert (2011) assume interval censored data. Empirical likelihood confidence intervals for dependent duration data are developed in El Ghouch et al. (2011). An important aspect was the development and investigation of feasible models for complex financial or psychological temporally related data. The non-parametric estimation of the volatility function is, for example, considered in Monsalve-Cobis et al. (2011). Statistical models and methods for dependence in insurance data are studied in Van Keilegom and Veraverbeke (2011). Important contributions are made in case of multivariate temporally related data. Correlation between multiple series have been modelled and statistically analysed. Manner and Segers

(2011) study tails of correlation mixtures of elliptical copulas while Veraverbeke et al. (2011) and Gijbels et al. (2011) consider the estimation of a conditional copula and association measures. Multivariate time series can also be studied from the perspective of their largest values or spikes. Gudendorf and Segers (2011) estimate non-parametrically an extreme-value copula in arbitrary dimensions, while Kojadinovic et al. (2011) derive large-sample tests of extreme-value dependence for multivariate copulas and Einmahl et al. (2011) propose an M-estimator for tail dependence in arbitrary dimensions. New methods are developed in continuous time models including Bayesian inference. Hunt and Devolder (2011) study semi-Markov regime switching interest rate models and minimal entropy measure. A versatile Bayesian hierarchical model with random effects for the system parameters is introduced in Lodewyckx et al. (2011). Multilevel or hierarchical Ornstein-Uhlenbeck process models have been studied in Oravecz and Tuerlinckx (2011). There are clear links between this work and WP4. New developments concern dynamic conditional correlation model, multivariate extensions of volatility models and modelling of extremes of multiple series. Bauwens et al. (2011) model, for example, multivariate volatility of electricity futures. Finally, detection and modelling spatial correlation has been investigated. There is a growing interest in improving the level of knowledge of spatial and spatio-temporal processes using spectral techniques (cf. GonzÃ¡lez-Manteiga and Crujeiras (2011)). Semi-parametric fractal Gaussian random fields are studied in Ruiz-Medina and Crujeiras (2011) and Crujeiras and Ruiz-Medina (2011). In Antoniadis et al. (2011) or González-Quintela et al. (2011) applications to real data are provided.

1.4 Work package 3: Incomplete data

The work on incomplete data can be captured under the broad headings of: (1) complex modeling approaches for missing data; (2) sensitivity analysis tools; (3) censored survival data; and (4) general incomplete data structures.

First, for complex modeling approaches for missing data, work has been done to propagate the proper use of incomplete data methodology in an applied context. Bunouf, Grouin, Molenberghs, and Koch (2011) is an example of such, bringing together colleagues from academia, the biopharmaceutical industry and regulatory authorities, from both sides of the Atlanic Ocean. Molenberghs, Kenward, Verbeke and Teshome (2011) developed a semi-parametric theory based on pseudo-likelihood rather than the more conventional generalized estimating equations. They included singly robust and doubly robust alternatives. It is a collaboration between UH, KUL-2, and LSHTM. Semi-parametric methodology, also in view of missing covariates, was the point of attention in Creemers *et al* (2011). Multiple imputation as a modeling tool for incomplete data received attention from Birhanu *et al* (2011).

Second, sensitivity analysis tools have been developed to assess the impact of untestable assumptions about the incomplete data distribution, given the observed data, on important scientific conclusions. Contributions were made by Poleto *et al* (2011ab) and Creemers *et al* (2011).

Third, we turn to the category of censored survival data. Several asymptotic results on quantiles have been generalized by Veraverbeke (2010) to the situation of censored data in an extended Koziol-Green model. In this model, there is covariate information and dependence between the survival times and censoring times described by an Archimedean copula. Further study on nonpara-

metric estimation of copulas has been carried out in cooperation with Irène Gijbels (KU Leuven) and Marek Omelka (Charles University Prague). In Gijbels, Veraverbeke and Omelka (2011) and Veraverbeke, Omelka, and Gijbels (2011), this is done for copulas and association measures in the presence of a single covariate. In Gijbels, Omelka, and Veraverbeke (2012), the previous results are extended to multivariate and functional covariates. A bootstrap procedure has been established in Omelka, Veraverbeke, and Gijbels (2012). Abegaz, Gijbels, and Vervarbeke (2012) estimated conditional copulas in a semi-parametric way, starting from a parametric family with parameter depending on the covariate. Janssen, Swanepoel, and Veraverbeke (2012, obtained asymptotic results for a smooth version of the empirical copula estimator where the smoothing is done with Bernstein polynomials. Van Keilegom and Veraverbeke (2011) is an invited discussion on a paper on statistical models and methods for dependence in insurance data. Biard, Loisel, Macci, and Veraverbeke (2010) studied some asymptotic results for risk measures in insurance. A famous estimator in the context of censored data is the Kaplan-Meier estimator. A brief review and discussion of this estimator is provided in Gijbels (2011).

In the classical Koziol-Green, Braekers and co-workers introduced copula functions at various levels to obtain different more flexible semi-parametric models for informative censoring. On the one hand, they considered an extension of the Koziol-Green model under dependent censoring while on the other hand, they introduced flexibility in the Koziol-Green proportional hazards assumption by a parametric copula family. For both models, they derived asymptotic results as uniform consistency and weak convergence for the derived semi-parametric distribution estimators. Next, they developed a goodness-of-fit test for the flexibility in these models. For left-censored data, they developed a zero-inflated Cox regression model in which the zero-inflated probability is modeled parametrically and the positive response variables are investigated through a Cox regression model. In this semi-parametric model, they proved the consistency and the asymptotic normality of both the finite and infinite dimensional estimators for the different parameters. Hereby, they also provide estimators for the variance of these estimators since a standard plug-in estimator is not possible due to the lack of a closest form. For unbalanced clustered multivariate survival data, frailty models are commonly used to model the time until an event in which the association within a cluster is taken into account. By exploiting properties of Archimedean copula functions, Braekers and colleagues developed copula-based alternatives for the frailty model. Hereby, they are able to lift the present practical limitations for copula models in regard to small balanced cluster sizes. This work is done together with Leen Prenen and Luc Duchateau, bringing together colleagues from UH and UGent. An extension of the previous work was done for unbalanced hierarchical clustered survival data. Hereby, the team developed a new copula model by using hierarchical, nested Archimedean copula functions. This model is compared to other models for hierarchical clustered survival models. This work is joint between Roel Braekers, Leen Prenen, Paul Janssen, Candida Geerdens, Luc Duchateau and Klaartje Goethals. In a recurrent event setting, they developed a copula model for the sequentially observed gap times between the different events. Hereby, they derived asymptotic properties for the semi-parametric estimator of the bivariate distribution function of the gap times. This work is done together with Jacobo De Una-Alvarez (UVigo).

Fourth, general incomplete data structures have received attention. Verbeke and Molen-

berghs (2011) studied arbitrariness of models for augmented and coarse data, thereby emphasizing incomplete-data and random-effects models. Molenberghs, Kenward, Aerts, Verbeke, Tsiatis, and Davidian (2011) studied the effect of random sample size, such as in incomplete data, sequential trials, and random cluster size, on inferential properties and connected the concepts of ignorability, ancillarity, completeness, separability, and degeneracy. It is a unifying theory, bringing together seemingly disparate concepts.

In the area of joint modeling of longitudinal and survival data, Sattar, Weissfeld and Molenberghs (2011) studied the joint occurrence of incomplete data and censoring.

1.5 Work package 4: Data with latent heterogeneity

In many situations, statistical models are used that assume the presence of latent, unobservable, structures to explain the variability observed in the data. The fact that those structures, can never be observed poses particular problems with respect to checking model assumptions. Several attempts have been made to extend the traditional models which usually assume the distribution of the outcomes to belong to the exponential family, conditionally on the latent variables which are most often assumed normally distributed. For example, Molenberghs and Verbeke (KUL2, 2011) have proposed a generalization for time-to-event data. Also, Creemers et al. (UH, KUL2, LSHTM, 2011) proposed a generalized shared-parameter model to model missingness when analysing longitudinal data. Related to the fact that random effects are unobserved is that usually the model fitted is the marginal model, obtained by integrating over the random effects. This implies that, strictly speaking, the parameters no longer retain their random-effects interpretation but should be considered as parameters in a marginal model. For example, Pryseley A. et al. (KUL2 & UH, 2011) discuss the estimation of negative variance components in models for Gaussian and non-Gaussian data, as well as the way such negative variance components should be interpreted.

When random-effects models are used for non-Gaussian data, non-linear link functions are used to relate the expectation to the latent variables. This implies that the marginal log-likelihood method can no longer be derived analytically. Many approximation methods have been proposed in the statistical literature, some of which perform very poorly in some contexts. To avoid this, Molenberghs, Verbeke, and Iddi (UH & KUL2, 2011), proposed a pseudo-likelihood method allowing the fitting of complex models for large datasets. The basic idea is to partition the dataset in sub samples and to combine the results obtained from fitting submodels to the various data sets. Related methodology has been applied by Molenberghs et al. (UH, KUL2, LSHTM, 2011) for the analysis of incomplete data. Alternatively, estimation problems can sometimes be solved by fitting the models in a Bayesian context. For example, Zhang et al. (KUL2 & EMC, 2011) used Bayesian latent variable models for spatially correlated tooth-level binary data in caries research.

Also, some contributions have been made with respect to developing software for the analysis of models with latent structure. A key example thereof is the work by Molas and Lesaffre (KUL2 & EMC 2011) who developed the R package HGLMMM for the fitting of hierarchical generalized linear models.

The fact that models with latent structures are applicable in many different contexts becomes apparent in the many publications in various settings where such models have been applied. Van

Beirendonck et al. (KUL2, 2011) and Van de Perre et al. (KUL2, 2011) used mixed models to account for the clustering of animals within farms. In two different publications, Diya et al. (KUL2 & EMC, 2011) applied random effects models to correct for the clustering of nurses within hospital units, when investigating the relationship between in-hospital mortality or re-admission into the intensive care nursing unit on one hand, and nurse staffing levels on the other. Thilakarathne et al. (KUL2, 2011) used random effects to account for the correlation between genes in several applications in the context of statistical genetics and bioinformatics. Finally, Wellens et al. (KUL2, 2011), and Agbaje et al. (KUL2 & EMC, 2011), to study interrater and intrarater variability in various contexts.

1.6 Work package 5: Highdimensional and compound data

In a joint collaboration effort between the UGent team, the Flemish Institute for Biotechnoloy (VIB), John Hopkins School of Public Health (Baltimore, USA) and the KULeuven, new techniques based on wavelet based functional mixed models for the analysis of tiling arrays (Clement et al., 2012) were developed and also on improved base-calling methods for the Roche 454 sequence. The collaboration with the VIB and the KULeuven has also resulted in a very powerful test for differential gene expression (Thas et al., 2012). The research team remains also active in applying cutting edge statistical methods to important problems in biomarker discovery (Hollevoet et al., 2012; De Ruyck et al., 2011), high-dimensional genomics (De Roy et al., 2012, Van Pottelberge et al., 2011) and environmental sciences (Staelens et al., 2012).

In ^{18}O-labeled mass spectra, two samples, obtained for two different biological conditions, are processed in the same spectrum. One of the samples is labeled with ^{16}O, the other with ^{18}O . Thus, the molecules of the heavy oxygen labeled sample become, on average, 4 mass units heavier. This allows distinguish them from the molecules from the other sample. As a result, the relative abundance of particular molecules can be obtained in the same mass spectrum, which allows to remove the between spectrum variability from the comparison. We have developed a series of statistical models, which allow analyzing data from such spectra, while taking into account various design factors like, e.g., incomplete labeling. The results have been published in a series of papers:

In peptide centric mass spectrometry, detection of the signal of interest, i.e., detection of peaks that are related to a peptide, is an important issue. For an effective signal extraction, the concept of the isotopic distribution of a peptide is very useful. The use of the isotopic distribution for signal detection requires efficient methods for the computation of the distribution, which is not a trivial task due to its combinatorial nature. We have reviewed solutions proposed for this problem and developed a new approach, which is more computation- and memory-efficient (Valkenborg et al., 2012; Claesen et al., 2012).

Microarrays are a very common and popular method of analysing gene expression. Thus, methods of analysis of microarray data are still of much interest. We have investigated the possibility of modelling the data by assuming a Laplace distribution based linear model, because the Laplace distribution seems to fit microarray data better than the normal distribution (Van Sanden and Burzykowski, 2011).

One of the possible uses of microarrays is discovery of biomarkers that could serve for dis-

ease diagnosis, treatment selection, etc. This application raises numerous issues, related to, e.g., large number of potential biomarkers, criteria for selecting the good candidates, etc. We have investigated statistical approaches that could be used to this aim (Van Sanden et al., 2012)

In many situations one observes a lot of variables, and a selection of the most important variables is a key task. Variable selection in flexible regression models (additive models and varying coefficient models) is accomplished in Antoniadis, Gijbels and Verhasselt (2011, 2012) using a combination of the nonnegative garotte technique and P-splines estimation. This results into a powerful variable selection method, for which the performance in comparison with other recent variable selection procedures has been illustrated.

In a flexible framework Van Deun, Wilderjans, van den Berg, Antoniadis and Van Mechelen (2011) studied sparse simultaneous component based data integration.

2 Network Activities

2.1 Web site and newsletter

All activities of the IAP-statistics network can be followed very closely from our web site. The address of the web site is

`http://www.stat.ucl.ac.be/IAP/PhaseVI`.

The web site contains e.g. the following information:

- Our logo
- Call for applications
- Description of the project
- List of scientific personnel working under the IAP project
- Downloadable member list
- Research activities (workshops, seminars, short courses,...)
- Downloadable technical reports, list of publications and list of books written my members of the network
- Annual reports and reports of scientific meetings
- Contact details

In addition an electronic newsletter is sent out every month to all IAP-members. In this newsletter, new activities (seminars, short courses, meetings, ...) are announced and a link to the appropriate web page is added for more details. The newsletter also contains a link to the updated list of publications and technical reports of the network.

2.2 Scientific meetings

2.2.1 Annual workshop

The annual IAP workshop of 2011 was organized by the UCL group, and took place on Friday 25 November. The goal of this workshop was to focus on major breakthroughs in the network over the last 5 years, and also on important and promising future research directions the network likes to take a lead in in the near future. The main organizers of the workshop were Jan Johannes and Ingrid Van Keilegom. A total of 74 researchers participated to the workshop, There were sessions on 'Dependence and copulas', 'High-dimensional data', 'Hierarchically structured data' and 'Instrumental variables regression'. The introductions given at the start of each session, which were meant to help the PhD students to better understand the scientific talks of each thematic session, were very much appreciated not only by the PhD students but also by more senior members of the network. More detailed information about the workshop can be found on the webpage:

`http://www.uclouvain.be/375357.html`.

2.2.2 Meetings

The following meetings were organized by the network in 2011:

- The University of Hasselt (UH) organized on May 19-20, 2011 a conference on 'Recent Advances in Statistics and Probability'. The conference was in honor of Noël Veraverbeke, on the occasion of his retirement. There were sessions on 'Survival Analysis', 'Nonparametric methods' and 'Probability and stochastic processes', and the list of invited speakers included four IAP members: L. Duchateau (UG), I. Gijbels (KUL1), W. González-Manteiga (USC) and I. Van Keilegom (UCL).

- EMC organized its yearly Biostatistics Symposium in Spring, which attracted about 80 participants. The subject was joint modelling of survival and repeated measurements.

- June 16-18 : the IWFOS (International Workshop on Functional and Operatorial Statistics) was co-organized by the European partner of Santiago de Compostela (together with two other universities) in Santander (Spain) partly under the heading of the IAP network. There were 87 participants from 17 countries. It is important to note that the lecture notes of the conference, which have been published Springer in the book 'Recent Advances in Functional Statistics and Related Topics', mentioned the IAP logo on the first page of the book. It is a nice example of how the international visibility of the network is put forward by members of the network. In addition, the international journal TEST will devote a special number to IWFOS contributions.

  The objective of the conference was to highlight the major trends in different areas of statistics with infinite dimension through the exchange of ideas and the promotion of collaboration between researchers from different countries. It aimed at contributing to future developments of the fields.

  More detailed information can be found on the webpage of the conference: `http://eio.usc.es/pub/iwfos/`.

2.3 Organization of the network: administrative meeting

The annual administrative meeting took place on 25 November 2011 in Louvain-la-Neuve, during the annual workshop. The meeting was attended by : E. Ceulemans (KUL1), L. Duchateau (UG), K. Faes (UH), I. Gijbels (KUL1), J. Johannes (UCL), T. Snijders (follow-up), I. Van Keilegom (UCL), N. Veraverbeke (UH), G. Verbeke (KUL2). The participants to this administrative meeting discussed issues related to past and future scientific activities organized by the network, scientific collaborations in the network, work valorization (such as web page, reports, ...), network organization, management and visibility. A detailed report of this meeting was sent to all participants, promoters of the network, members of the follow-up committee and C. Lejour and V Feys (from Belspo).

2.4 Collaborations, working groups and seminars

2.4.1 Collaborations

The IAP network is working on a broad range of research topics in statistics. There is a large number of scientific collaborations within the network, as can be seen from the list of technical reports and publications (see Section 3, and in particular Subsection 3.10, where all joint technical reports and publications are collected). Below, we mention a few examples of ongoing collaborations between members of different teams of the network.

- Members of the team of KUL-1 (I. Gijbels, K. Van Deun, I. Van Mechelen) collaborate intensively with members of the team of UJF (A. Antoniadis, S. Lambert-Lacroix), and this on various topics: variable selection, grouped regularization techniques and sparse methods for data integration. This led to several joint papers.

- Members of the teams of UHasselt (N. Veraverbeke) and KUL-1 (I. Gijbels), and international collaborators (M. Omelka, Prague University; and A. F. Abegaz, University of Groningen, both former postdoctoral researchers of the IAP-network), among others, continued their collaboration on the study of copulas in statistical inference.

- P. Janssen (UH), L. Duchateau (UG), C. Legrand (UCL) and PhD students of their respective research groups are working together on projects related to frailty models and competing risks in survival analysis.

- I. Van Keilegom (UCL), R. Crujeiras Casais (USC), P. Janssen (UH) and L. Duchateau (UG) joined forces from 4 partner universities of the network to work on projects related to the modeling and the analysis of spatial data that are subject to censoring. Both theoretical, methodological and applied research is carried out by the group.

- There are very strong collaborations between KUL-2, EMC and UH, primarily through the joint appointments of staff members:

  – E.Lesaffre at KUL-2 and EMC

  – G.Molenberghs at KUL-2 and UH

  Also, the collaboration between KUL-2 and UH has been formalized in I-BioStat, the interuniversity institute of biostatistics and statistical bioinformatics.

- UH, KUL-2, and LSHTM are involved in a Taylor & Francis book project on missing data, together also with North Carolina State University.

- There is lots of collaboration between P. Eilers at EMC and members of the group at UCL, namely with Ph. Lambert on density estimation and smoothing, and with B. Govaerts on chemometric subjects.

- Members from EMC (P. Eilers) and UH (N. Hens) work together on modeling of contact network frequencies and on the analysis of serological measurements.

- C. M. Cadarso Suárez (USC) collaborates with the KUL-2 group (G. Molenberghs), with the UH group (C. Faes) and with the UCL group (I. Van Keilegom) on two projects in the area of neural activity, and multistate models in survival analysis.

2.4.2 Working groups

Below are a few examples of active working groups in the network. They are an important tool to stimulate interactions between network partners, and to stay informed of the research achievements of other partners of the network.

- *Frailty models*
  The working group on frailty models and related models with members from UH, UG, UCL, and KUL-2 continues to meet on a regular basis. The main research topics discussed in this working group are related to competing risks and to transformation models for survival data (where the transformed cumulative hazard is modeled as a function of covariates in a linear way and where such models are extended to time-varying covariates and frailties).

- *Bioinformatics*
  I. Van Mechelen (KUL-1) and G. Verbeke (KUL-2) are both partners in a KUL Center of Excellence for computational systems biology. Their teams have made several methodological contributions (of immediate relevance for WP5) in the context of statistical bioinformatics, which have lead to new biological and medical insights.

- *Variable selection and sparseness*
  Members of the KUL-1 (in particular I. Gijbels and A. Verhasselt) and of UJF (A. Antoniadis) have extensive collaborations on the development of semi- and nonparametric methods for complex data, such as heavy noisy data. One of the aspects of the work is exploring regularization techniques. These collaborations are situated in WP1 and on the interface between WP1 and WP4.

- *Copulas*
  The working group on 'Modeling dependencies and inference based on copulas' consisting of members of KUL-1 and UH has already led to several joint publications of members of these two universities.

- There are several working groups between KUL-2 and UH:
  - Sensitivity analysis for incomplete data
  - Bioinformatics and statistical genetics
  - Surrogate markers

  PhD students, post-docs, and staff members from both universities regularly attend these meetings which take place approximately three times per year.

- *Goodness-of-fit tests*
  C. Heuchenne and I. Van Keilegom (UCL) have extensive collaborations with members of

the USC partner (W. González-Manteiga and R. Crujeiras Casais) on goodness-of-fit tests in (semi)-parametric regression, when the data are or are not subject to right censoring. They also work together on a project dealing with ROC-curves in regression.

2.4.3 Seminars

Each of the participating partners organizes on a regular basis statistics seminars at their universities. Announcements of these seminars are sent out to most Belgian statisticians, including those participating in the network.

Apart from the regular statistics seminars at the universities involved, several seminars have been organized by the network itself, around central themes of the network. They are on some occasions given by members of the network, in order to foster research interactions and exchange of ideas. These seminars are indicated by a star (*).

- *Anthony Davison (EPFL, Lausanne), who is member of the follow-up committee of our IAP network, has been awarded the Francqui Chair 2011. He has given a series of lectures on 'Likelihood theory' and 'Statistics of extremes'. The host university is the University of Hasselt (UH). The lectures took place at UH, KUL-1 and UCL. There were 10 lectures in total between March 28 and May 5, 2011. The organizing committee consisted of the following four members of the network : P. Janssen, promoter (UH), I. Gijbels (KUL-1), I. Van Keilegom (UCL) and N. Veraverbeke (UH).

- Anastasios Tsiatis (North Carolina State University) has been awarded the 2010-2011 Princess Lilian Foundation Visiting Professorship. The promoters (hosts) are Geert Molenberghs (UH and KUL-2), Geert Verbeke (KUL-2) and Marc Aerts (UH). During his visit, he has given lectures at KUL-2, UH and UCL on May 18, 19, 23, 25 and 27, 2011. He has also provided guidance to the doctoral students of KUL-2 and UH.

- April 29, 2011: Carla Moreira (University of Vigo, Spain), 'Kernel density estimation with double truncated data', at UCL

- *April 29, 2011: Rawane Samb (UCL), 'Nonparametric estimation of the residual density', at UCL

- November 4, 2011: Yvik Swan (Université Libre de Bruxelles), 'Univariate and multivariate Chen-Stein characterizations – a parametric approach', at UCL

- *November 4, 2011: Philippe Lambert (UCL), 'Nonparametric additive models for interval-censored data', at UCL

- December 2, 2011: Natalie Neumeyer (University of Hamburg, Germany), 'Some specification tests in nonparametric quantile regression', at UCL

- Elisa María Molanes López (Universidad Carlos III de Madrid, Spain), 'The importance of being diagnosed (by Youden index)', at USC

- *Maria José Rodríguez Álvarez (USC), 'A new semiparametric ROC regression approach based on direct methodology. Application to endocrine data', at USC
- *Beatriz Pateiro López (USC), 'Multivariate uniformity tests', at USC

2.5 Short courses

Several short (intensive) courses have been organized. These courses were intended for all members of the network, and in particular (but not exclusively) for the PhD-students. The announcements were each time sent out to all members and posted on the web site. No (or reduced) registration fees were required for IAP-members.

A list of the short courses organized during the working year 2011 is given below.

- April 15, 2011: Short course on 'Nonparametric methods for ROC curves', by Juan Carlos Pardo Fernández (University of Vigo, Spain), at USC
- April 27, May 2, 2011: Short course on 'Single index models', by Olivier Lopez (Université Paris VI, France), at UCL, jointly organized by the FNRS Graduate School in Statistics and the IAP network
- May 24, 2011: Short course on 'Continuous time models in financial econometrics', by Jun Yu (Singapore Management University), at UCL
- November, 2011: Short course on 'Using LaTeX and Beamer', by Beatriz Pateiro López (USC), at X Congreso Galego de Estatística e Investigación de Operacións
- December 15, 2011: Short course on 'An introduction to weak dependence techniques', by Paul Doukhan (Université de Cergy Pontoise, France), at UCL

2.6 PhD jury committees

Below are a few examples of IAP members that took part in the PhD jury at other universities of the network. This participation is a very useful way to get familiar with the research carried out at other groups of the network. A complete list of PhD theses currently in preparation in the network can be found on the website:

`http://www.stat.ucl.ac.be/IAP/PhaseVI/research_theses.html`

The list mentions (among others) which members of partner universities of the network take part in the PhD committee. This participation is a very useful way to get familiar with the research carried out at other groups of the network. The website also contains a list of defended theses in the network since 2007.

- Paul Eilers (EMC) was member of the PhD jury of Réjane Rousseau, UCL, who defended her thesis in June 2011.
- Paul Janssen (UH) was member of the PhD jury of Kukatharmini Tharmaratnam (KUL-2), who obtained her PhD degree in July 2011.

- Noël Veraverbeke (UH) was member of the PhD jury of Dominik Sznajder (doctoral student financed by the network), KUL-1. Dominik Sznajder obtained his PhD degree in October 2011.

- Geert Molenberghs (UH, KUL-2) is member of the guidance committee of Inge Vrinssen (PhD student KUL).

- Irène Gijbels (KUL-1) is member of the guidance committee of Majda Talamakrouni (PhD student UCL).

2.7 Prizes obtained by network members

- G. Molenberghs (KUL-2 and UH), M. Aerts (UH), and G. Verbeke (KUL-2) were the promotors of the 2010-2011 Princess Lilian Foundation visiting professorship for Prof. Anastasios A. Tsiatis (University of North Carolina, USA).

- G. Molenberghs (KUL-2 and UH) and G. Verbeke (KUL-2) received the Excellence-In-Continuing-Education Award for 2011, selected by the Advisory Committee on Continuing Education of the American Statistical Association, for the short course 'Foundations and recent advances in longitudinal and incomplete data,' taught at the Joint Statistical Meetings (American Statistical Association, Institute of Mathematical Statistics, International Biometrics Society, and the Statistical Society of Canada), Miami, U.S.A., July 30-31, 2011.

- The paper

  > Gijbels, I. and Verhasselt, A. (2010). Regularisation and P-splines in generalised linear models. *Journal of Nonparametric Statistics*, Volume **22**, Issue 3, 271–295;

  received the Student Paper Award 2010, of the Journal of Nonparametric Statistics. A. Verhasselt received the award at the occasion of the Joint Statistics Meeting, Miami, Florida, in August 2011. She presented the paper at that meeting.

- The conference paper

  > De Brabanter, K., De Brabanter, J., Gijbels, I., Suykens, J.A.K., and De Moor, B. (2011). New developments in kernel regression with correlated errors.

  received a Best Poster Award, at the Graybill 2011 conference, Fort Collins, Colorado, USA.

- I. Van Keilegom (UCL) was selected as co-editor of the Journal of the Royal Statistical Society - Series B for a period of 4 years (2012-2015).

3 Technical Reports and Publications

Below we provide the scientific output related to the IAP-statistics network. We give both the technical reports and the publications of network members in 2011 :

- Technical Reports: These are manuscripts that have been written in 2011, and have been submitted for publication to an international journal. The reports are also available on our web site:

  `http://www.stat.ucl.ac.be/IAP/PhaseVI/publication_tr.html`.

  Each Technical Report has a number of the form TR11xxx, and we mention these reference numbers below. The web site also contains the pdf-file of many of the Technical Reports.

- Refereed publications: We list all published papers in international journals in 2011 (with refereeing system). We make the distinction between published papers and papers in press. See also the IAP-Statistics Reprints Series on our web site:

  `http://www.stat.ucl.ac.be/IAP/PhaseVI/publication_reprint.html`,

  for the published papers (reference numbers are of the form R11xxx). The papers in press have a label of the form RP11xxx.

- Non-refereed publications: We also include (an incomplete list of) papers that have been published without undergoing a peer review. The reference numbers are of the form NR11xxx (for the published ones) and NRP11xxx (for the ones in press).

- Books: These are books written by members of the network, that are published by international editors. They can also be found on the webpage

  `http://www.stat.ucl.ac.be/IAP/PhaseVI/publication_books.html`

  (reference numbers are of the form B11xxx and BP11xxx).

Below we list the research output of the IAP-network for each of the categories described above. We start with separate lists for each partner in the network, followed by a list of the technical reports and publications that are co-signed by researchers from at least two different groups from the network.

3.1 Université catholique de Louvain, UCL

3.1.1 Technical reports

[1] Autin, F., Freyermuth, J.M. and R. von Sachs, Combining thresholding rules: a new way to improve the performance of wavelet estimators, 2011. TR11023.

[2] Autin, F., Freyermuth, J.M. and R. von Sachs, Block-Threshold-Adapted Estimators via a maxiset approach, 2011. TR11019.

[3] Autin, F., Freyermuth, J.-M. and R. von Sachs, Ideal denoising within a family of tree-structured wavelet estimators, 2011. TR11002.

[4] Badin, L., Daraio, C. and L. Simar, Explaining Inefficiency in Nonparametric Production Models: the State of the Art, 2011. TR11035.

[5] Badin, L., Daraio, C. and L. Simar, How to Measure the Impact of Environmental Factors in a Nonparametric Production Model?, 2011. TR11021.

[6] Bauwens, L., Hafner, C. and D. Pierret, Multivariate volatility modeling of electricity futures, 2011. TR11015.

[7] Bereswill, M. and J. Johannes, On the effect of noisy observations of the regressor in a functional linear model, 2011. TR11042.

[8] Bertrand, A. and C. M. Hafner, On heterogeneous latent class models with applications to the analysis of rating scores, 2011. TR11030.

[9] Bocart, F. and C. M. Hafner, Econometric analysis of volatile art markets, 2011. TR11031.

[10] Bouezmarni, T., El Ghouch, A. and A. Taamouti, Bernstein Estimator for Unbounded Density Copula, 2011. TR11029.

[11] Chen, S.X. and I. Van Keilegom, Estimation in semiparametric models with missing data, 2011. TR11010.

[12] Comte, F. and J. Johannes, Adaptive functional linear regression, 2011. TR11041.

[13] Daniel, B.C., Hafner , C.M., Manner, H. and L. Simar, Asymmetries in Business Cycles and the Role of Oil Production, 2011. TR11034.

[14] Daouia, A. and B.U. Park, On Projection-Type Estimators of Multivariate Isotonic Functions, 2011. TR11039.

[15] Daouia , A., Gardes, L. and S. Girard, On kernel smoothing for extremal quantile regression, 2011. TR11033.

[16] Dahlke, M., Jay Breidt, F., Opsomer, J. and I. Van Keilegom, Nonparametric endogenous post-stratification estimation, 2011. TR11004.

[17] Delsol, L. and I. Van Keilegom, Semiparametric M-Estimation with Non-Smooth Criterion Functions, 2011. TR11044.

[18] Denuit, M., Eeckhoudt, L. and H. Schlesinger, When Ross meets Bell: the linex utility function, 2011. TR11016.

[19] Devolder, P. and H. Tassa, Solvency capital, inflation and time horizon in pension liabilities, 2011. TR11018.

[20] Devolder, P., Solvency requirement for long term guarantee: risk measure versus probability of ruin, 2011. TR11017.

[21] Einmahl, J.H.J., Krajina, A. and J. Segers, An M-estimator for tail dependence in arbitrary dimensions, 2011. TR11007.

[22] Florens, J.P., Simar, L. and I. Van Keilegom, Frontier estimation in nonparametric location-scale models, 2011. TR11032.

[23] Gaddah, A. and R. Braekers, A Goodness-of-fit test for a flexible copula Koziol-Green model. *Submitted for publication*, 2012. TR12001.

[24] Gudendorf , G. and J. Segers, Nonparametric estimation of multivariate extreme-value copulas, 2011. TR11020.

[25] Haedo, C. and M. Mouchart, A stochastic independence approach for different measures of global specialization, 2011. TR11008.

[26] Heuchenne, C., Laurent, S., Legrand, C. and I. Van Keilegom, Likelihood based inference for semi-competing risks, 2011. TR11024.

[27] Hunt, J. and P. Devolder, Semi-Markov regime switching interest rate models and minimal entropy measure, 2011. TR11012.

[28] Hunt, J. and P. Devolder, A semi-Markov regime switching extension of the Vasicek model, 2011. TR11011.

[29] Jaeger, J. and P. Lambert, Bayesian generalized profiling estimation in hierarchical linear dynamic systems, 2011. TR11001.

[30] Johannes, J. and R. Schenk, Adaptive estimation of linear functionals in functional linear models, 2011. TR11040.

[31] Johannes, J. and M. Schwarz, Adaptive Gaussian inverse regression with partially unknown operator, 2011. TR11036.

[32] Kojadinovic, I., Segers, J. and J. Yan, Large-sample tests of extreme-value dependence for multivariate copulas, 2011. TR11014.

[33] Mouchart, M., Russo, F. and G. Wunsch, Inferring causal relations by modelling structures, 2011. TR11009.

[34] Müller, U. and I. Van Keilegom, Efficient parameter estimation in regression with missing responses, 2011. TR11028.

[35] Noh, H., El Ghouch, A. and I. Van Keilegom, Quality of fit measures in the framework of quantile, 2011. TR11027.

[36] Noh, H., El Ghouch, A. and I. Van Keilegom, On assessing model adequacy in linear quantile regression, 2011. TR11026.

[37] Pigeon, M., Antonio, K. and M. Denuit, Individual Loss Reserving with the Multivariate Skew Normal Model, 2011. TR11065.

[38] Rotolo, F., Legrand, C.and I. Van Keilegom, Simulation of clustered multi-state survival data based on a copula model, 2011. TR11043.

[39] Samb, R., Heuchenne, C. and I. Van Keilegom, Estimation of the error density in a semiparametric transformation model, 2011. TR11025.

[40] Simar, L., Vanhems, A. and P.W. Wilson, Statistical Inference for DEA Estimators of Directional Distances, 2011. TR11037.

[41] Simar, L. and V. Zelenyuk, To Smooth or Not to Smooth? The Case of Discrete Variables in Nonparametric Regressions, 2011. TR11045.

[42] Timmermans, C., Delsol, L. and R. von Sachs, Using Bagadis in nonparametric functional data analysis: predicting from curves with sharp local features, 2011. TR11022.

[43] Vanhems, A. and I. Van Keilegom, Semiparametric transformation model with endogeneity: a control function approach, 2011. TR11013.

[44] Van Keilegom, I. and N. Veraverbeke, Statistical models and methods for dependence in insurance data, 2011. TR11003.

3.1.2 Refereed publications (published)

[45] Autin, F., Freyermuth, J.M. and R. von Sachs, Ideal denoising within a family of tree-structured wavelet estimators. *Electronic Journal of Statistics*, **5**, 829-855, 2011. R11211.

[46] Braekers, R. and A. Gaddah, Flexible Modelling in the Koziol-Green Model by a Copula function. *Communications in Statistics: Theory and Methods*, **40**, 1218-1235, 2011. R11178.

[47] Cetinyurek, A. and P. Lambert, Smooth estimation of survival functions and hazard ratios from interval-censored data using Bayesian penalized B-splines. *Statistics in Medicine*, **30**, 75-90, 2011. R11004.

[48] Daouia, A. and I. Gijbels, Estimating frontier cost models using extremiles. In *Festschrift in honor of Leopold Simar*, Editors: P. Wilson, I. Van Keilegom. Springer-Verlag: Berlin, Heidelberg, ISBN 978-3-7908-2348-6. Chapter 4, pages 65-81, 2011. R11202.

[49] Daouia, A. and I. Gijbels, Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, **161**, 147-165, 2011. R11201.

[50] Davydov, Y. and S. Liu, Transformations des lois multivariées à queue régulière. *Revue Roumaine de Mathématiques Pures et Appliquées*, 55, 6, 483-492, 2011. R11017.

[51] Denuit, M., Haberman, S. and A. Renshaw, Longevity-indexed life annuities. *North American Actuarial Journal*, 15, 97-111, 2011. R11024.

[52] Denuit, M. and M. Mesfioui, Dispersive effect of cross-aging with archimedean copulas. *Statistics and Probability Letters*, 81, 1407-1418, 2011. R11023.

[53] Denuit, M., Eeckhoudt, L. and M. Menegatti, A note on subadditivity of zero-utility premiums. *ASTIN Bulletin*, 41, 239-250, 2011. R11022.

[54] Eichler, M., Motta, G. and R. von Sachs, Fitting dynamic factor models to non-stationary time series. *Journal of Econometrics*, 163, 51-70, 2011. R11013.

[55] El Ghouch, A., Van Keilegom, I. and I. W.McKeague, Empirical likelihood confidence intervals for dependent duration data. *Econometric Theory*, **27**, 178-198, 2011. R11006.

[56] Gaddah, A. and R. Braekers, An extension of the Koziol-Green model under dependent censoring. *Journal of Nonparametric Statistics*, **23**, 439-453, 2011. R11177.

[57] Geenens, G. and L. Simar, Single-index modeling of conditional probabilities in two-way contingency tables. *Statistics*, 45, 451-478, 2011. R11029

[58] Gonzàlez-Manteiga, W., Pardo-Fernàndez, J.C. and I. Van Keilegom, ROC curves in nonparametric location-scale regression models. *Scandinavian Journal of Statistics*, 38, 169-184, 2011. R11010.

[59] Gschlossl, S., Schoenmaekers, P. and M. Denuit, Risk classification in life insurance: Methodology and case study. *European Actuarial Journal*, 1, 23-41, 2011. R11021.

[60] Gudendorf, G. and J. Segers, Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *Journal of Multivariate Analysis*, **102**, 37-47, 2011. R11003.

[61] Guillotte, S., Perron, F. and J. Segers, Non-parametric Bayesian inference on bivariate extremes. *Journal of the Royal Society. Series B, Statistical Methodology*, 73, 3, 377-406, 2011. R11011.

[62] Kneip, A., Simar, L. and P.W. Wilson, A Computational Efficient, Consistent Bootstrap for Inference with Non-parametric DEA Estimators. *Computational Economics*, 38,483-515, 2011. R11030.

[63] Kojadinovic, I., Segers, J. and J. Yan, Large-sample tests of extreme-value dependence for multivariate copulas. *The Canadian Journal of Statistics*, 39, 4, 97-111, 2011. R11025.

[64] Linton, O., Mammen, E., Perch Nielsen, J. and I. Van Keilegom, Nonparametric regression with filtered data. *Bernoulli*, **17**, 1, 60-87, 2011. R11008.

[65] Lambert P., Smooth and semi- and nonparametric Bayesian estimation of bivariate densities from bivariate histogram data. *Computational Statistics and Data Analysis*, **55**, 429-445, 2011. R11005.

[66] Lambert, P., Perelman, S. pestieau, P. and J. Schoenmaeckers, Health Insurance Coverage and Adverse Selection. *The Individual and the Welfare State. Life histories in Europe*, Chap. 20, Part 3, 225-231, 2011. R11034.

[67] Lambert, Ph., Comments on: Inference in multivariate Archimedean copula models. *Test*, 20, 2, 284-286, 2011. R11033.

[68] Lambert, Ph., Nonparametric additive location-scale models for interval censored data. *Statistics and Computing, Springer* Science+Business Media, LLC, 2011

[69] Manner, H. and J. Segers, Tails of correlation mixtures of elliptical copulas - Insurance: Mathematics and Economics. *Journal of Statistical Planning and Inference*, **48**, 153-160, 2011. R11002.

[70] Motta, G., Hafner, C. and R. von Sachs, Locally stationary factor models: identification and nonparametric estimation. *Econometric Theory*, **27**, 6 , 2011. R11007.

[71] Mouchart, M., Russo, F. and G. Wunsch, Inferring causal relations by modelling structures. *Statistica*, 4, 412-432, 2011. R11015.

[72] Pigeon, M. and M. Denuit, Composite Lognormal-Pareto model with random threshold. *Scandinavian Actuarial Journal*, 177-192, 2011. R11020.

[73] Roueff, F. and R. von Sachs, Locally stationary long memory estimation. *Stochastic Processes and their Applications*, **121**, 813-844, 2011. R11009.

[74] Russo, F., Wunsch, G. and M. Mouchart, Inferring causality through counterfactuals in observational studies - Some epistemological issues. *Bulletin de Methodologie Sociologique*, 111, 43-64, 2011. R11014.

[75] San Martín, E., Jara, A., Rolin, J.-M. and M. Mouchart, On the Bayesian nonparametric generalization of IRT-type models. *Psychometrika*, **76**, 3,385-409, 2011. R11012.

[76] Schubert, T. and L. Simar, Innovation and export activities in the German mechanical engineering sector: an application of testing restrictions in production analysis. *Journal of Productivity Analysis*, **36**, 55-69, 2011. R11029.

[77] Segers, J., Comments on: Inference in multivariate Archimedean copula models. *Test*, 20, 2, 281-283, 2011. R11039.

[78] Simar, L. and P.W. Wilson, Performance of the bootstrap for DEA estimators and iterating the principle. *Handbook on Data Envelopment Analysis*, **10**, 241-271, 2011. R11035.

[79] Simar, L. and P.W. Wilson, Two-Stage DEA: Caveat Emptor. *Journal of Productivity Analysis*, **36**, 205-218, 2011. R11031.

[80] Simar, L. and P.W. Wilson, Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis*, **36**, 33-53, 2011. R11027.

[81] Simar, L. and V. Zelenyuk, Stochastic FDH/DEA estimators for frontier analysis. *Journal of Productivity Analysis*, **36**, 1-20, 2011. R11026.

[82] Timmermans , C., Delsol , L. and R. von Sachs, Bases Giving Distances. A New Semimetric and its Use for Nonparemetric Functional Data Analysis. *Recent Advances in Functional Data Analysis and Related Topics*, Contributions to Statistics, Physica-Verlag, 307-313, 2011. R11018.

[83] Trufin, J., Albrecher, H. and M. Denuit, Properties of risk measures derived from ruin theory. *Geneva Risk and Insurance Review*, **36**, 174-188, 2011. R11019.

[84] Van Keilegom, I. and N. Veraverbeke, Discussion on Statistical models and methods for dependence in Insurance data by Haug, S., Klüppelberg C. and Peng L. *Journal of the Korean Statistical Society*, **40**, 155-157, 2011. R11175.

[85] Van Keilegom, I., De Uña-Álvarez, J. and L. Meira-Machado, Nonparametric location-scale models for censored successive survival times. *Journal of Statistical Planning and Inference*, **141**, 1118-1131, 2011. R11001.

[86] Varron, D. and I. Van Keilegom, Uniform in bandwidth exact rates for a class of kernel estimators. *Annals of the Institute of Statistical Mathematics*, **63**, 1077-1102, 2011. R11016.

3.2 Katholieke Universiteit Leuven, KUL-1

3.2.1 Technical reports

[87] Antoniadis, A., Gijbels, I. and Lambert-Lacroix, S., Penalized estimation in additive varying coefficient models using grouped regularization, 2011. TR11038.

[88] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines, 2011. TR11053.

[89] Croux, C., Fried, R., Gijbels, I. and K. Mahieu, Robust forecasting of non-stationary time series, 2011. TR11054.

[90] De Brabanter, K., De Brabanter, J., Gijbels, I. and B. De Moor, Derivative estimation with local polynomial fitting, 2012. TR12004.

[91] De Spiegeleer, J. and W. Schoutens, Pricing Contingent Convertibles: A Derivatives Approach, 2011. TR11057.

[92] Dhaene, J., Dony, J., Forys, M., Linders, D. and W. Schoutens, FIX - The fear index. Measuring market fear, 2011. TR11055.

[93] Dhaene, J., Linders, D., Schoutens, W. and D. Vyncke, The herd behavior index: A new measure for systemic risk in financial markets, 2011. TR11056.

[94] Gijbels, I. , Omelka, M. and N. Veraverbeke, Multivariate and functional covariates and conditional copulas, 2012. TR12002.

[95] Gijbels, I. and M. Omelka, Testing for homogeneity if multivariate dispersions using dissimilarity measures, 2011. TR11058.

[96] Gijbels, I. and D. Sznajder, Positive quadrant dependence testing and constrained copula estimation, 2011. TR11059.

[97] Gijbels, I. and D. Sznajder, Testing tail monotonicity by constrained copula estimation, 2011. TR11060.

[98] Helsen, W.F., De Neve, J., D'Hooghe, P., Gijbels, I., Gram, M., Hubert, M., Peers, K., Probst, S. and J. Van Winckel, A Ménage à Trois: Injury Prevention, Medical Attention and Time Loss Injuries, 2011. TR11061.

[99] Hubert, M., Dierckx, G. and D. Vanpaemel, Detecting influential data points for the Hill estimator in Pareto-type distributions, 2011. TR11062.

[100] Hubert, M., Gijbels, I. and D. Vanpaemel, Reducing the Mean Squared Error of quantile-based estimators by smoothing, 2011. TR11063.

[101] Omelka, M., Veraverbeke N. and I. Gijbels, Bootstrapping the conditional copula, 2012. TR12003.

[102] Verboven, S., Hubert, M., and P. Goos, Robust preprocessing and model selection for spectral data, 2011. TR11064.

3.2.2 Refereed publications (published)

[103] Antoniadis, A., Gijbels, I. and M. Nikolova,Penalized likelihood regression for generalized linear models with nonquadratic penalties. *The Annals of the Institute of Statistical Mathematics*, Volume **63**, Issue 3, 585-615, 2011. R11197.

[104] Beirlant J., Boniphace E. and G. Dierckx, Generalized sum plots. *REVSTAT-Statistical Journal*, **9**, 181–198, 2011. R11198.

[105] Beirlant J., Dierckx G. and A. Guillou, Bias-reduced estimators for bivariate tail modelling. *Insurance: Mathematics & Economics*, **49**, 18-26, 2011. R11199.

[106] Beirlant, J., Geffray S. and A. Guillou, Estimation of the Bias of the Maximum Likelihood Estimators in an Extreme Value Context. *Communications in Statistics - Theory and Methods*, **40**, 3959-3971, 2011. R11200.

[107] Ceulemans, E., Timmerman, M. E., and H. A. L. Kiers. The CHull procedure for selecting among multilevel component solutions. *Chemometrics and Intelligent Laboratory Systems* , **106**, 12-20, 2011. R11036.

[108] Daouia, A. and I. Gijbels, Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, **161**, 147-165, 2011. R11201.

[109] Daouia, A. and I. Gijbels, Estimating frontier cost models using extremiles. In *Festschrift in honor of Leopold Simar*, Editors: P. Wilson, I. Van Keilegom. Springer-Verlag: Berlin, Heidelberg, ISBN 978-3-7908-2348-6. Chapter 4, pages 65-81, 2011. R11202.

[110] De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., and I. Partchev. The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software*, **39**, 1-28, 2011. R11037.

[111] Desmet, L. and I. Gijbels, Curve fitting under jump and peak irregularities using local linear regression. *Communications in Statistics, Theory and Methods*, Volume **40**, Number 22, 4001-4020, 2011. R11203.

[112] Engelen, S. and M. Hubert, Detecting outlying samples in a parallel factor analysis model. *Analytica Chimica Acta*, **705**, 155-165, 2011. R11204.

[113] Facon, B., Magis, D., Nuchadee, M.-L., and P. De Boeck. Do Raven's Colored Progressive Matrices function in the same way in typical and clinical populations? Insights from the intellectual disability field. *Intelligence*, **39**, 281-291, 2011. R11038.

[114] Frederickx, S., and I. Van Mechelen. Identifying the situational triggers underlying avoidance of communication situations and individual differences therein. *Personality and Individual Differences*, **52**, 438-443, 2012. R11039.

[115] Gijbels, I., Veraverbeke, N. and M. Omelka, Conditional copulas, association measures and their applications. *Computational Statistics and Data Analysis*, **55**, 1919-1932, 2011. R11121.

[116] Gijbels, I., The Kaplan-Meier estimator. In *International Encyclopedia of Statistical Science*, 1st Edition, Lovric, Miodrag (Ed.), Springer, ISBN: 978-3-642-04897-5. Part 11, pages 709-710, 2011. R11205.

[117] Gijbels, I. and I. Prosdocimi, Smooth estimation of mean and dispersion function in extended Generalized Additive Models with application to Italian Induced Abortion data. *Journal of Applied Statistics*, Volume **38**, Issue 11, 2391-2411, 2011. R11206.

[118] Janssen, R., and M. Crauwels. Content and student factors in mastering environmental studies - nature in primary education: Evidence from a national assessment in Flanders (Belgium). *Journal of Biological Education*, **45**, 20-28, 2011. R11040.

[119] Lodewyckx, T., Kim, W., Lee, M. D., Tuerlinckx, F., Kuppens, P., and E.-J. Wagenmakers. A tutorial on Bayes factor estimation with the product space method. *Journal of Mathematical Psychology*, **55**, 331-347, 2011. R11041.

[120] Lodewyckx, T., Tuerlinckx, F., Kuppens, P., Allen, N. B., and L. Sheeber. A hierarchical state space approach to affective dynamics. *Journal of Mathematical Psychology*, **55**, 68-83, 2011. R11042.

[121] Madan, D. and W. Schoutens, Conic Finance and the Corporate Balance Sheet. *International Journal of Theoretical and Applied Finance*, Publisher: World Scientific, **14**, 587-610, 2011. R11207.

[122] Magis, D., Béland, S., and G. Raîche. A test-length correction to the estimation of extreme proficiency levels. *Applied Psychological Measurement*, **35**, 91-109, 2011. R11043.

[123] Magis, D., and P. De Boeck. Identification of differential item functioning in multiple-group settings: A multivariate outlier detection approach. *Multivariate Behavioral Research*, **46**, 733-755, 2011. R11044.

[124] Magis, D., and G. Raîche. catR: an R package for computerized adaptive testing. *Applied Psychological Measurement*, **35**, 576-577, 2011. R11045.

[125] Magis, D., and G. Raîche. On the relationships between Jeffreys modal and weighted likelihood estimation of ability under logistic IRT models. *Psychometrika*, **77**, 163-169, 2012. R11046.

[126] Magis, D., Raîche, G., Béland, S., and P. Gérard. A generalized logistic regression procedure to detect differential item functioning among multiple groups. *International Journal of Testing*,**11**, 365-386, 2011. R11047.

[127] Masol, V. and W. Schoutens, Comparing alternative Levy base correlation models for pricing and hedging CDO tranches. *Quantitative Finance*, **11**, 763-773, 2011. R11208.

[128] Oravecz, Z., and F. Tuerlinckx. The linear mixed model and the hierarchical Ornstein-Uhlenbeck model: Some Equivalences and differences. *British Journal of Mathematical and Statistical Psychology*, **64**, 134-160, 2011. R11048.

[129] Oravecz, Z., Tuerlinckx, F., and J. Vandekerckhove. A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods*, **16**, 468-490, 2011. R11049.

[130] Partchev, I., and P. De Boeck. Can fast and slow intelligence be differentiated? *Intelligence*, **40**, 23-32, 2012. R11050.

[131] Schoutens W., and G. Van Damme, The beta-variance gamma model. *Review of Derivatives Research*, **14**, 263-282, 2011. R11210.

[132] Slaets, L., Claeskens, G. and M. Hubert, Phase and amplitude-based clustering for functional data. *Computational Statistics and Data Analysis*, *56*, 2360-2374, 2012. R12007.

[133] Thorrez, L., Laudadio, I., Van Deun, K., Quintens, R., Hendrickx, N., Granvik, M., Lemaire, K., Schraenen, A., Van Lommel, L., Lehnert, S., Aguayo-Mazzucato, C., Cheng-Xue, R., Gilon, P., Van Mechelen, I., Bonner-Weir, S., Lemaigre, F., and F. Schuit. Tissue-specific disallowance of housekeeping genes: The other face of cell differentiation. *Genome Research*, **21**, 95-105, 2011. R11051.

[134] Rousseeuw, P. and M. Hubert, Robust statistics for outlier detection. *WIREs Data Mining & Knowledge Discovery*, **1**, 73-79, 2011. R11209.

[135] Vande Gaer, E., Ceulemans, E., Van Mechelen, I., and P. Kuppens. The CLASSI-N method for the study of sequential processes. *Psychometrika*, **77**, 85-105, 2012. R11052.

[136] Vandekerckhove, J., Tuerlinckx, F., and M.D. Lee. Hierarchical diffusion models for two-choice response times. *Psychological Methods*, **16**, 44-62, 2011. R11053.

[137] Van Deun, K., Wilderjans, T.F., van den Berg, R.A., Antoniadis, A., and I. Van Mechelen. A flexible framework for sparse simultaneous component based data integration. *BMC Bioinformatics*, **12**, 448, 2011. R11054.

[138] Van Mechelen, I., and A.K. Smilde. Comparability problems in the analysis of multiway data. *Chemometrics and Intelligent Laboratory Systems*, **106**, 2-11, 2011. R11055.

[139] Van Nijlen, D., and R. Janssen. Measuring mastery across grades: An application to spelling ability. *Applied Measurement in Education*, **24**, 367-387, 2011. R11056.

[140] Veraverbeke, N, Omelka, M. and I. Gijbels, Estimation of a conditional copula and association measures. *Scandinavian Journal of Statistics*, **38**, 766-780, 2011. R11172.

[141] Wan, L., Ogrinz, B., Vigo, D., Bersenev, E., Tuerlinckx, F., van den Bergh, O., and A. Aubert. Cardiovascular autonomic adaptation to long-term confinement during a 105-day simulated Mars mission. *Aviation, Space and Environmental Medicine*, **82**, 711-716, 2011. R11057.

[142] Wilderjans, T.F., Ceulemans, E., Van Mechelen, I., and D. Depril. ADPROCLUS: A graphical user interface for fitting additive profile clustering models to object by variable data matrices. *Behavior Research Methods*, **43**, 56-65, 2011. R11058.

[143] Wilderjans, T.F., Ceulemans, E., Van Mechelen, I., and R. A. van den Berg. Simultaneous analysis of coupled data matrices subject to different amounts of noise. *British Journal of Mathematical and Statistical Psychology*, **64**, 277-290, 2011. R11059.

3.2.3 Refereed publications (in press)

[144] Abegaz, F., Gijbels, I. and N. Veraverbeke, Semiparametric estimation of conditional copulas. *Journal of Multivariate Analysis*, (in press), 2012. RP12001.

[145] Aelvoet, W., Terryn, N., Molenberghs, G., De Backer, G., Vrints, C. and M. van Sprundel, Do inter-hospital comparisons of in-hospital, acute myocardial infarction case-fatality rates serve the purpose of fostering quality improvement? An evaluative study. *BMC Health Services Research*, (in Press), 2011. RP11035.

[146] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines. *Journal of Computational and Graphical Statistics*, (to appear), 2011. RP12002.

[147] Ceulemans, E., Kuppens, P., and I. Van Mechelen. Capturing the structure of distinct types of individual differences in the situation-specific experience of emotions: The case of anger. *European Journal of Personality*, (in press), 2011. RP11013.

[148] Cho, S.-J., Partchev, I., and P. De Boeck. Parameter estimation of multiple item response profiles model. *British Journal of Mathematical and Statistical Psychology*, (in press), 2011. RP11014.

[149] Croux, C., Gijbels, I. and I. Prosdocimi, Robust estimation of mean and dispersion functions in extended generalized additive models. *Biometrics*, (to appear), 2011. RP12003.

[150] De Boeck, P., Cho, S.-J., and M. Wilson. Explanatory secondary dimension modelling of latent DIF. Applied Psychological Measurement. De Roover, K., Ceulemans, E., and Timmerman, M. E. (in press). How to perform multiblock component analysis in practice. *Behavior Research Methods*, (in press), 2011. RP11015.

[151] De Roover, K., Ceulemans, E., Timmerman, M. E., and P. Onghena. A clusterwise simultaneous component method for capturing within-cluster differences in component variances and correlations. *British Journal of Mathematical and Statistical Psychology*, (in press), 2011. RP11016.

[152] De Roover, K., Ceulemans, E., Timmerman, M. E., Vansteelandt, K., Stouten, J., and P. Onghena. Clusterwise simultaneous component analysis for analyzing structural differences in multivariate multiblock data. *Psychological Methods*, (in press), 2011. RP11017.

[153] Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., and E.-J. Wagenmakers. Testing theories of post-error slowing. *Attention, Perception, and Psychophysics*, (in press), 2011. RP11018.

[154] Fung, W., He, X., Hubert, M., Portnoy, S. and H. Wang, Editorial for the Special Issue on quantile regression and semiparametric methods. *Computational Statistics & Data Analysis*, (to appear), 2012. RP12004.

[155] Gijbels, I. and I. Prosdocimi, Flexible mean and dispersion function estimation in extended Generalized Additive Models. *Communications in Statistics - Theory and Methods*, Special Issue on *Statistics for Complex Problems*, (to appear), 2012. RP12005.

[156] Guillaume, F. and W. Schoutens, Calibration risk: Illustrating the impact of calibration risk under the Heston model. *Review of Derivatives Research*, (to appear), 2011. RP11057.

[157] Kadengye, D. T., Cools, W., Ceulemans, E., and W. Van den Noortgate. Simple imputation methods versus direct likelihood analysis for missing item scores in multilevel educational data. *Behavior Research Methods*, (in press), 2011. RP11019.

[158] Magis, D., and P. De Boeck. A robust outlier approach to prevent Type I error inflation in DIF. *Educational and Psychological Measurement*, (in press), 2011. RP11020.

[159] Magis, D., and B. Facon. Angoff's Delta method revisited: improving DIF detection under small samples. *British Journal of Mathematical and Statistical Psychology*, (in press), 2011. RP11021.

[160] Partchev, I., De Boeck, P., and R. Steyer. How much power and speed is measured in this test? *Assessment*, (in press), 2011. RP11022.

[161] Schepers, J., and I. Van Mechelen. A two-mode clustering method to capture the nature of the dominant interaction pattern in large profile data matrices. *Psychological Methods*, (in press), 2011. RP11023.

[162] Vanbrabant, K., Kuppens, P., Braeken, J., Demaerschalk, E., Boeren, A., and F. Tuerlinckx. A relationship between verbal aggression and personal network size. *Social Networks*, (in press), 2011. RP11024.

[163] Vigo, D.E., Ogrinz, B., Wan, L., Bersenev, E., Tuerlinckx, F., van den Bergh, O., and A. Aubert. Sleep-wake differences in heart rate variability during a 105-day simulated mission to Mars. *Aviation, Space and Environmental Medicine*, (in press), 2011. RP11025.

[164] Wilderjans, T.F., Ceulemans, E., and P. Kuppens. Clusterwise HICLAS: A generic modeling strategy to trace similarities and differences in multi-block binary data. *Behavior Research Methods*, (in press), 2011. RP11026.

[165] Wilderjans, T.F., Ceulemans, E., and I. Van Mechelen. The SIMCLAS model: Simultaneous analysis of coupled binary data matrices with noise heterogeneity between and within data blocks. *Psychometrika*, (in press), 2011. RP11027.

3.2.4 Books (published)

[166] De Spiegeleer, J. and W. Schoutens, *Contingent Convertible Coco-Notes: Structuring & Pricing*. Publisher: Euromoney Books. London, 2011. B11007.

[167] De Spiegeleer, J. W. and Schoutens, *The Handbook of Convertible Bonds: Pricing, Strategies and Risk Management*. Chichester: Wiley. ISBN: 978-0-470-68968-4, 2011. B11006.

3.3 Katholieke Universiteit Leuven, KUL-2

3.3.1 Refereed publications (published)

[168] Agbaje, J., Mutsvari, T., Lesaffre, E., Declerck, D. (2011). Examiner performance in calibration exercises compared with field conditions when scoring caries experience. *Clinical Oral Investigations.* , 2011. R11081.

[169] Alqerban, A., Jacobs, R., Fieuws, S., Nackaerts, O., The SEDENTEXCT Project Consortium, Willems, G. (2011). Comparison of 6 cone-beam computed tomography systems for image quality and detection of simulated canine impaction-induced external root resorption in maxillary lateral incisors. *American Journal of Orthodontics and Dentofacial Orthopedics*, **140** (3), e129-e139, 2011. R11087.

[170] Assam, P., Tilahun, A., Alonso, A., Molenberghs, G. (2011). Information-theoretic approach to surrogate markers evaluation with failure time endpoints. *Lifetime Data Analysis*, **17** (2), 195-214, 2011. R11072.

[171] Birhanu, T., Molenberghs, G., Sotto, C., and M.G. Kenward, Doubly robust and multiple-imputation-based generalized estimating equations. *Journal of Biopharmaceutical Statistics,* **21**, 202-205, 2011. R11145.

[172] Borgermans L., Goderis G., Van Den Broeke C., G.Verbeke, Carbonez A., Ivanova A., Mathieu C., and Heyrman J. (2011), 'Patients' experiences with patient-centered care are associated with documented outcome of care indicators for diabetes: findings from the Leuven Diabetes Project,' *International Journal of Care Pathways*, **15**, 65-75, 2011. R11064.

[173] Creemers A., Hens N., Aerts M., Molenberghs G., G.Verbeke, and Kenward M.G. (2011), 'Generalized shared-parameter models and missingness at random,' Statistical Modelling, 11, 279-310, 2011. R11068.

[174] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and M.G. Kenward Generalized shared-parameter models and missingess at random. *Statistical Modelling,* **11**, 270-310, 2011. R11068.

[175] Decramer, M., Molenberghs, G., Liu, D., Celli, B., Kesten, S., Lystig, T., Tashkin, D., UPLIFT investigators (2011). Premature discontinuation during the UPLIFT study. *Respiratory Medicine*, **105** (10), 1523-1530, 2011. R11071.

[176] de Geus, H., Bakker, J., Lesaffre, E., le Noble, J. (2011). Neutrophil gelatinase-associated lipocalin at ICU admission predicts for acute kidney injury in adult patients. *American Journal of Respiratory and Critical Care Medicine*, **183** (7), 907-14, 2011. R11078.

[177] De Leeneer, K., De Schrijver, J., Clement, L., Baetens, M., Lefever, S., De Keulenaer, S., Van Criekinge, W., Deforce, D., Van Nieuwerburgh, F., Bekaert, S., Pattyn, F., De Wilde, B., Coucke, P., Vandesompele, J., Claes, K., Hellemans, J. (2011). Practical Tools to Implement Massive Parallel Pyrosequencing of PCR Products in Next Generation Molecular Diagnostics. *PLoS One*, **6** (9), art.nr. e25531, -., 2011. R11085.

[178] Diya, L., Van den Heede, K., Sermeus,W., Lesaffre, E. (2011). The Use of "Lives Saved" Measures in Nurse Staffing and Patient Safety Research: Statistical Considerations. *Nursing Research*, **60** (2), 100-106, 2011. R11080.

[179] Diya, L., Van den Heede, K., Sermeus, W., Lesaffre, E. (2011). The relationship between inhospital mortality, readmission into the intensive care nursing unit and/or operating theatre and nurse staffing levels. *Journal of Advanced Nursing* (epub ahead of print), art.nr. 10.1111/j.1365-2648.2011.05812.x., 2011. R11074.

[180] Hernandez, M., Sorsa, T., Obregn, F., Tervahartiala, T., Valenzuela, M., Pozo, P., Dutzan, N., Lesaffre, E., Molas, M., Gamonal, J. (2011). MMP-13 enhances MMP-9 activation during progression of chronic periodontitis. *Journal of Clinical Periodontology.* , 2011. R11084.

[181] Lima, A., van Bommel, J., Sikorska, K., van Genderen, M., Klijn, E., Lesaffre, E., Ince, C., Bakker, J. (2011). The relation of near-infrared spectroscopy with changes in peripheral circulation in critically ill patients. *Critical Care Medicine*, **39** (7), 1649-1654, 2011. R11076.

[182] Lingsma, H., Steyerberg, E., Lesaffre, E. (2011). Logistic random effects regression models: a comparison of statistical packages for binary and ordinal outcomes. *BMC Medical Research Methodology*, **11**, art.nr. 77, -., 2011. R11077.

[183] Molas, M., Lesaffre, E. (2011). Hierarchical Generalized LinearModels: The R Package HGLMMM. *Journal of Statistical Software*, **39** (13), 1-20, 2011. R11079.

[184] Molenberghs, G. and G. Verbeke, On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference*, **141**, 861-868, 2011. R11220.

[185] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, **21**, 187-206, 2011. R11013.

[186] Molenberghs G., G.Verbeke, and Iddi S. (2011), 'Pseudo-likelihood methodology for partitioned large and complex samples,' *Statistics and Probability Letters*, **81**, 892-901, 2011. R11062.

[187] Pryseley, A., Tchonlafi, C., Verbeke, G., and G. Molenberghs, Estimating negative variance components from Gaussian and non-Gaussian data: a mixed models approach. *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011.

[188] Pryseley A., Tchonlafi C., G.Verbeke, and Molenberghs G. (2011), 'Estimating negative variance components from Gaussian and non-Gaussian data: A mixed models approach,' *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011.

[189] Sotto, C., Beunckens, C., Molenberghs, G. and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2011. R11147.

[190] Star, H., Thevissen, P., Jacobs, R., Fieuws, S., Solheim, T., Willems, G. (2011). Human dental age estimation by calculation of pulp/tooth volume ratios yielded on clinically acquired cone beam computed tomography images of monoradicular teeth. *Journal of Forensic Sciences*, **56** (Suppl 1), S77-S82, 2011. R11089.

[191] Thilakarathne, P., Clement, L., Lin, D., Shkedy, Z., Kasim, A., Talloen, W., Versele, M., Verbeke, G. (2011). The Use of Semi-parametric Mixed Models to Analyze PamChip Peptide Array Data: an Application to an Oncology Experiment. *Bioinformatics*, **27** (20), 2859-2865, 2011. R11067.

[192] Thilakarathne, P., Verbeke, G., Engelen, K., Marchal, K., Lin, D. (2011). Identifying Differentially Expressed genes in the Absence of Replication. *International Journal of Bioinformatics Research and Applications*, **7** (3), 2011. R11070.

[193] Truyers, C., Buntinx, F., De Lepeleire, J., De Hert, M., vanWinkel, R., Aertgeerts, B., Bartholomeeusen, S., Lesaffre, E. (2011). Incident somatic comorbidity after psychosis: results from a retrospective cohort study based on Flemish general practice based data. *BMC Family Practice*, **12** (1), 132, 2011. R11073.

[194] Van Beirendonck S., Driessen B., G.Verbeke, and Geers R. (2011) 'Behaviour of piglets after castration with and without carbon dioxide anaesthesia,' *Journal of Animal Science*, **89**, 3310-3317, 2011. R11063.

[195] Van der Borght K., Van Craenenbroeck E., Lecocq P., Van Houtte M., Van Kerckhove B., Bacheler L., G.Verbeke, and van Vlijmen H. (2011), 'Cross-validated stepwise regression for identification of novel non-nucleoside reverse transcriptase inhibitor resistance associated mutations,' *BMC Bioinformatics*, **12**, 386, 2011. R11065.

[196] Vandenbroeck, S., De Geest, S., Dobbels, F., Fieuws, S., Stalmans, I., Zeyen, T. (2011). Prevalence and Correlates of Self-Reported Non-adherence with Eye Drop Treatment: The Belgian Compliance Study in Ophthalmology (BCSO). *Journal of Glaucoma*, **20** (7), 414-421, 2011. R11086.

[197] Vandenbroeck, S., De Geest, S., Dobbels, F., Fieuws, S., Stalmans, I., Zeyen, T. (2011). The Prevalence and Correlates of Self-reported Non-adherence: The Belgian Compliance Study in Ophthalmology (BCSO). *Journal of Glaucoma*, **20** (7), 414-421, 2011. R11088.

[198] Van de Perre V., Driessen B., Van Thielen J., G.Verbeke, Geers R. (2011), 'Comparison of pig behaviour when given a sequence of enrichment objects or a chain continuously,' *Animal Welfare*, **20**, 641-649, 2011. R11066.

[199] Van de Voorde, P., Sabbe, M., Tsonaka, R., Rizopoulos, D., Calle, P., De Jaeger, A., Lesaffre, E., Matthys, D. (2011). The long-term outcome after severe trauma of children in Flanders (Belgium): a population-based cohort study using the International Classification of Functioning-related outcome score. *European Journal of Pediatrics*, **170** (1), 65-73, 2011. R11083.

[200] Vangeneugden, T., Molenberghs, G., Verbeke, G. and C. Demétrio, Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, **38**, 215-232, 2011. R11183.

[201] Wellens, N., Van Lancker, A., Flamaing, J., Gray, L., Moons, P., Verbeke, G., Boonen, S., Milisen, K. (2011). Interrater reliability of the interRAI Acute Care. *Archives of Gerontology and Geriatrics*, **19** (Epub ahead of print: DOI:10.1016/j.archger.2011.07.005), 2011. R11069.

[202] Verbeke, G. and G. Molenberghs, Arbitrariness of models for augmented and coarse data, with emphasis on incomplete-data and random-effects models. *Statistical Modelling*, **10**, 391-419, 2011. R11181.

[203] Zhang, Y., Todem, D., Kim, K., Lesaffre, E. (2011). Bayesian latent variable models for spatially correlated tooth-level binary data in caries research. *Statistical Modelling*, **11** (1), 25-47, 2011. R11082.

[204] Zicari, F., Van Meerbeek, B., Debels, E., Naert, I., Lesaffre, E. (2011). An up to 3-year controlled clinical trial comparing the outcome of glass fibre posts and composite cores with goldalloy bases posts and cores for the restoration of endodontically treated teeth. *International Journal of Prosthodontics*, **24** (4), art.nr. 2270, 363-372, 2011. R11075.

3.3.2 Refereed publications (in press)

[205] Kassahun, W., Neyens, T., Molenberghs, G., Faes, C. and G. Verbeke, Modeling overdispersed longitudinal binary data from the Jimma longitudinal studies using a combined beta and normal random-effects model. *Archives of Public Health*, (in Press), 2011. RP11051.

[206] Luts, J., Molenberghs, G., Verbeke, G., Van Huffel, S. and J.A.K. Suykens, A mixed effects least squares support vector machine model for classification of longitudinal data. *Computational Statistics and Data Analysis*, (in Press), 2011. RP11038.

[207] Molenberghs, G., Kenward, M.G., Aerts, M., Verbeke, G., Tsiatis, A.A., Davidian, M. and D Rizopoulos D, On random sample size, ignorability, ancillarity, completeness, separability, and degeneracy: sequential trials, random sample sizes, and missing data. *Statistical Methods in Medical Research,* 2011, in press. RP11002.

[208] Verbeke, G., Fieuws, S., Molenberghs, G. and M. Davidian, The analysis of multivariate longitudinal data: A review. *Statistical Methods in Medical Research*, (in Press), 2011. RP11037.

3.3.3 Non-refereed publications (published)

[209] Sunaert P., Bastiaens H., Nobels F., Feyen L., Vermeire E., G.Verbeke, De Maeseneer J., Willems S., and De Sutter A. (2011) 'Verbeteren regionale acties de zorg voor diabetespatienten ?,' *Huisarts Nu*, **40**, 332-338, 2011. NR11001.

3.4 Universiteit Gent, UG

3.4.1 Refereed publications (published)

[210] Baele, G., van de Peer, Y. and S. Vansteelandt, Context-dependent codon partition models provide significant increases in model fit in atpB and rbcL protein-coding genes. *BMC Evolutionary Biology*, **11**, Article 145, 2011. R11111.

[211] Belachew, T., Hadley, C., Lindstrom, D., Getachew, Y., Duchateau, L. and P. Kolsteren, Food insecurity and age at menarche among adolescent girls in Jimma Zone Southwest Ethiopia: a longitudinal study. *Reproductive biology and endocrinology*, **9**, 125, 2011. R11127.

[212] Bekaert, M., Timsit, J.-F., Vansteelandt, S., Depuydt, P., Vésin, A., Garrouste-Orgeas, M., Decruyenaere, J., Clec'h, C., Azoulay, E. and D. Benoit, Attributable Mortality of Ventilator Associated Pneumonia: A Reappraisal Using Causal Analysis. *American Journal of Respiratory and Critical Care Medicine*, **184**, 1133-1139, 2011. R11116.

[213] Bekaert, K.M., Tuyttens, F.A.M., Duchateau, L., De Brabander, H.F., Aluwe, M., Millet, S., Vandendriessche, F. and L. Vanhaecke, The sensitivity of Flemish citizens to androstenone: Influence of gender, age, location and smoking habits. *Meat Science*, **88**, 548-552, 2011. R11133.

[214] Baecke P. and D. Van den Poel, Data augmentation by predicting spending pleasure using commercially available external data. *Journal of intelligent information systems* **36**, 367-383, 2011. R11125.

[215] Baecke P. and D. Van den Poel, Improving purchasing behavior predictions by data augmentation with situational variables. *International journal of information technology & decision making*, **9**, 853-872, 2011. R11126.

[216] Best, D.J., Rayner, J.C.W. and O. Thas, Easily Applied Tests for the Rayleigh Distribution. *Shakhya series B*, **72**, 254-263, 2011. R11097.

[217] Bowden, J. and S. Vansteelandt, Mendelian randomisation analysis of case-control data using Structural Mean Models. *Statistics in Medicine*, **30**, 678-694, 2011. R11116.

[218] Buyze, J. and E. Goetghebeur, Crossover studies with survival. *Statistical Methods in Medical Research*, **25, 1-18**, 2011. R11090.

[219] Clement, L., De Beuf, K., Thas, O., Vuylsteke, M., Irizarry, R. and C. Crainiceanu, Fast Wavelet Based Functional Models for Transcriptome Analysis with Tiling Arrays. *Statistical Applications in Genetics and Molecular Biology*, **11** (1), article 4, 2011. R11100.

[220] Comté, L., Vansteelandt, S., Rode, R.A. and B. Vrijens, Estimation of HIV treatment-efficacy by combining structural nested mean models with pharmacokinetic models of antiretroviral drug exposure. *Statistics and its Interface*, **4**, 511-520. R11115.

[221] De Bock, K. W. and D. Van den Poel, An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction. *Expert systems with applications* **38**, 12293-12301, 2011. R11123.

[222] De Boeck, B. and Thas, O. and J.P. Ottoy, Reweighted Smooth Tests of Goodness of Fit. *Communications in Statistics-Theory and Methods*, **40**, 1651-1670, 2011. R11094.

[223] De Boeck, B. and Thas, O., Rayner, J.C.W. and D.J. Best, Smooth Tests For the Gamma Distribution. *Journal of Statistical Computation and Simulation*, **81**, 843-855. IF: 0.215, 2011. R11095.

[224] De Decker, S., Gielen, I.M.V.L., Duchateau, L., Lang, J., Dennis, R., Corzo-Menendez, N., van Bree, H.J.J., Van Soens, I., Binst, D.H.A.R., Waelbers, T., and L.M.L.M. Van Ham, Intraobserver and interobserver agreement for results of low-field magnetic resonance imaging in dogs with and without clinical signs of disk-associated wobbler syndrome. *J. Am. Vet. Med. Assoc.*, **238**, 74-80, 2011. R11141.

[225] De Decker, S., Gielen, I.M.V.L., Duchateau, L., Corzo-Menendez, N., van Bree H.J.J., Kromhout, K., Bosmans, T. and L.M.L. Van Ham, Intraobserver, interobserver, and intermethod agreement for results of myelography, computed tomography-myelography, and low-field magnetic resonance imaging in dogs with disk-associated wobbler syndrome. *J. Am. Vet. Med. Assoc*, **238**, 1601-1608, 2011. R11134.

[226] De Ruyck, K., Sabbe, N., Oberije, C., Vandecasteele, K., Thas ,O., De Ruysscher, D., Lambin, P., Van Meerbeeck, J., De Neve, W. and H. Thierens, Development of a Multi-Component Prediction Model for Acute Esophagitis in Lung Cancer Patients Receiving Chemo-Radiotherapy. *International Journal of Radiation Oncology, Biology, Physics*, **81**, 537-544. IF: 4.503, 2011. R11102.

[227] Fischer, K, Goetghebeur, E., Vrijens, B. and I. R. White, A structural mean model to allow for noncompliance in a randomized trial comparing 2 active treatments. *Biostatistics*, **12**, 247,257, 2011. R11091.

[228] Goetghebeur, E., Van Rossem, R., Baert, K., Vanhoutte, K., Boterberg, T., Demetter, P., De Ridder, M., Harrington, D., Peeters, M., Storme, G., Verhulst, J., Vlayen, J., Vrijens, F., Vansteelandt, S. and W. Ceelen, Quality Insurance of rectal cancer - phase 3: statistical methods to benchmark centers on a set of quality indicators. KCE report 161C, 114 pages, 2011. R11114.

[229] Goetghebeur, E., Causal Inference: Sense and Sensitivity versus Prior and Prejudice. Nova Acta Leopoldina, Abhandlungen der Deutschen Akademie der Naturforscher Leopoldina, Computermodelle in der Wissenschaft - zwischen Analyse, Vorhersage und Suggestion, Vortrage anlasslich der Jahresversammlung vom 2. bis 4. Oktober 2009 zu Halle (Saale), ISSN 0369-5034, 47-64, 2011. R11115.

[230] Hoffmann, T.J., Vansteelandt, S., Lange, C., Silverman, E.K., DeMeo, D.L. and N.M. Laird, Combining Disease Models to Test for Gene-Environment Interaction in Nuclear Families. *Biometrics*, **67**, 1260-1270, 2011. R11110

[231] Janssen, P. and L. Duchateau, Comments on: Inference in multivariate Archimedean copula models. *TEST*, **20**, 271-275, 2011. R11132.

[232] Lievens, A., Van Aelst, S., Van den Bulcke, M. and E. Goetghebeur, Enhanced analysis of real-time PCR data Nucleic Acids Research, 1-15, 2011. R11092.

[233] Lievens, A., Van Aelst, S., Van den Bulcke, M., and E. Goetghebeur, "Enhanced analysis of real-time PCR data by using a variable efficiency model: FPK-PCR," *Nucleic Acids Research*, **40** (2): e10, 2011. R11117.

[234] Lipinska, U., Hermans, K., Meulemans, L., Dumitrescu, O., Badiou, C., Duchateau, L., Haesebrouck, F., Etienne, J. and Lina, G., Panton-Valentine Leukocidin Does Play a Role in the Early Stage of Staphylococcus aureus Skin Infections: A Rabbit Model. *PLOS ONE 6*, e22864, 2011. R11131.

[235] Martinussen, T., Vansteelandt, S., Gerster, M. and J.v.B. Hjelmborg, Estimation of direct effects for survival data using the Aalen additive hazards model. *Journal of the Royal Statistical Society - Series B*, **73**, 773-788, 2011. R11113.

[236] Möberg, K., De Nobele, S., Devos, D. , Goetghebeur, E., Segers, P., Trachet, B., Vervaet, C., Renard, M., Coucke, P., Loeys, B., De Paepe, A. and J. De Backer, The Ghent Marfan trial

: a randomized, double-blind placebo controlled trial with β-blockers. *International Journal of Cardiology.*, 2011. R11093.

[237] Pey, P., Vignoli, M., Haers, H., Duchateau, L., Rossi, F. and Saunders, J.H., Contrast-enhanced ultrasonography of the normal canine adrenal gland. *Vet. Radiol. Ultrasound*, **52**, 560-567, 2011. R11130.

[238] Pezeshki, A., Stordeur, P., Wallemacq, H., Schynts, F., Stevens, M., Boutet, P., Peelman, L.J., De Spiegeleer, B., Duchateau, L., Bureau, F. and C. Burvenich, Variation of inflammatory dynamics and mediators in primiparous cows after intramammary challenge with Escherichia coli. *Vet. Res.*, **42**, 15, 2011. R11139.

[239] Prinzie A. and D. Van den Poel, Modeling complex longitudinal consumer behavior with Dynamic Bayesian networks: an Acquisition Pattern Analysis application. *Journal of intelligent information systems* **36**, 283-304, 2011. R11124.

[240] Quataert, P., Verschelden, P., Janne Elise, B., Verbeke, G., Goetghebeur, E. and F. Ollevier, A diagnostic modelling framework to construct indices of biotic integrity: a case study of fish in the Zeeschelde estuary (Belgium). *Estuarine Coastal and Shelf Science*, **94**, 222-233, 2011. R11128.

[241] Rayner, J.C.W., Thas, O. and D.J. Best, Smooth Tests of Goodness of Fit. *WIREs Computational Statistics*, **3**, 397-406, 2011. R11096.

[242] Rodriguez, P.Y.D., Coddens, A., Del Fava, E., Abrahantes, J.C., Shkedy, Z., Martin, L.O.M., Munoz, E.C., Duchateau, L., Cox, E. and B.M. Goddeeris, High prevalence of F4(+) and F18(+) Escherichia coli in Cuban piggeries as determined by serological survey. *Trop. Anim. Health Prod.*, **43**, 937-946, 2011. R11135.

[243] Sjolander, A. and S. Vansteelandt, Doubly robust estimation of attributable fractions. *Biostatistics*, **12**, 112-121, 2011. R11108.

[244] Stevens, M.G.H., Peelman, L.J., De Spiegeleer, B., Pezeshki, A., Van de Walle, G.R., Duchateau, L. and C. Burvenich, Differential gene expression of the toll-like receptor-4 cascade and neutrophil function in early- and mid-lactating dairy cows. *JDS*, **94**, 1277-1288, 2011. R11136.

[245] Thorleuchter D., Van den Poel D. and Prinzie A. (2011). Analyzing existing customers' websites to improve the customer acquisition process as well as the profitability prediction in B-to-B market. *Expert systems with applications* **39**, 2597-2605, 2011. R11122.

[246] Van Aelst, S., Vandervieren, E., and G. Willems, A Stahel-Donoho estimator based on huberized outlyingness, *Computational Statistics and Data Analysis*, **56** (3), 531-542, 2011. R11118.

[247] Van Aelst, S., and G. Willems, Robust and Efficient One-way MANOVA Tests, *Journal of the American Statistical Association*, **106** (494), 706-718, 2011. R11119.

[248] Van Aelst, S., Vandervieren, E., and G. Willems, Stahel-Donoho Estimators with Cellwise Weights, *Journal of Statistical Computation and Simulation*, **81** (1), 1-27, 2011. R11120.

[249] VanderWeele, T.J. and S. Vansteelandt, A weighting approach to causal effects and additive interaction in case-control studies: marginal structural linear odds models. *American Journal of Epidemiology*, **174**, 1197-1203, 2011. R11112.

[250] Van Hoecke, H., Van Cauwenberge, P., Thas, O. and J. Watelet, The ARIA Guidelines in Specialist Practice: a Nationwide Survey. *Rhinology*, **48**, 28-34. IF: 2.182, 2011. R11099.

[251] Van Pottelberge, G., Mestdagh, P., Bracke, K., Thas, O., van Durme, Y., Joos, G., Vandesompele, J. and G. Brusselle, MicroRNA Expression in Induced Sputum of Smokers and Patients with Chronic Obstructive Pulmonary Disease. *American Journal of Respiratory and Critical Care Medicine*, **183, 898-906**. IF: 10.689, 2011. R11098.

[252] Vansteelandt, S. and N. Keiding, G-computation: lost in translation? *American Journal of Epidemiology*, **173**, 739-742, 2011. R11109.

[253] Vansteelandt, S., Bowden, J., Babanezhad, M. and E. Goetghebeur, On Instrumental Variables Estimation of Causal Odds Ratios. *Statistical Science*, **26**, 403-422, 2011. R11114.

[254] Verbeken, M., Suleman, S., Baert, B., Vangheluwe, E., Van Dorpe, S., Burvenich, C., Duchateau, L., Jansen, F.H. and B. De Spiegeleer, Stability-indicating HPLC-DAD/UV-ESI/MS impurity profiling of the anti-malarial drug lumefantrine. *Malaria Journal* **10**, 51, 2011. R11138.

[255] Vignoli, M., Barberet, V., Chiers, K., Duchateau, L., Bacci, B., Terragni, R., Rossi, F. and J.H. Saunders, Evaluation of a manual biopsy device, the 'Spirotome', on fresh canine organs: liver, spleen, and kidneys, and first clinical experiences in animals. *Eur. J. Cancer Prev.*, **20**, 140-145, 2011. R11137.

[256] Yewhalaw, D., Wassie, F., Steurbaut, W., Spanoghe, P., Van Bortel, W., Denis, L., Tessema, D.A., Getachew, Y., Coosemans, M., Duchateau, L. and N. Speybroeck, Multiple Insecticide Resistance: An Impediment to Insecticide-Based Malaria Vector Control Program. *PLOS ONE 6*, e16066, 2011. R11140.

3.4.2 Refereed publications (in press)

[257] De Bock, K. W. and D. Van den Poel, Reconciling performance and interpretability in customer churn prediction using ensemble learning based on generalized additive model. *Expert systems with applications*, (to appear), 2011. RP11008.

[258] De Roy, K., Clement, L., Thas, O., Wang, Y. and N. Boon, Flow cytometry for fast microbial community fingerprinting. *Water Research*, (to appear). IF: 4.966, 2011. RP11004.

[259] Hollevoet, K., Nackaerts, K., Thas, O., Thimpont, J., Germonpre, P., De Vuyst, P., Bosque, L., Legrand, C., Kellen, E., Kishi, Y., Delanghe, J., and J. van Meerbeeck, The Effect

of Clinical Covariates on the Diagnostic and Prognostic Value of Soluble Mesothelin and Megakaryocyte Potentiating Factor. CHEST. (DOI 10.1378/chest.11-0129). IF: 6.519, (to appear), 2011. RP11007.

[260] Möberg, K., De Nobele, S., Devos, D., Goetghebeur, E. , Segers, P., Trachet, B., Vervaet, C., Renard, M., Coucke, P., Loeys, B., De Paepe, A. and J. De Backer, The Ghent Marfan trial: a randomized, double-blind placebo controlled trial with beta-blockers. *International Journal of Cardiology,* (in press), 2011. RP11006.

[261] Rayner, J.C.W., Thas, O., Pipelers, P. and Beh, E. Calculating Bivariate Orthonormal Polynomials by Recurrence. *Australian and New Zealand Journal of Statistics*, (to appear). IF: 0.618, 2011. RP11001.

[262] Staelens, J., Wuyts, K., Adriaenssens, S., Van Avermaet, P., Buysse, H., Van den Bril, B., Roekens, E., Ottoy, J.P., Verheyen, K., Thas, O. and E. Deschepper, Trends in atmospheric nitrogen and sulphur deposition in northern Belgium. *Atmospheric Environment*, (to appear), 2011. RP11002.

[263] Thas, O., Clement, L., Rayner, J.C.W., Carvalho, B. and Van Crieckinge, W. An Omnibus Consistent Adaptive Percentile Modified Wilcoxon Rank Sum Test with Applications in Gene Expression Studies. *Biometrics*, (to appear). IF: 1.764, 2011. RP11002.

[264] Thas, O., De Neve, J., Clement, L. and J.P. Ottoy, Probabilistic Index Models (Read paper, with discussion). *Journal of the Royal Statistical Society*, Series B, **74**, 1-29, (to appear). IF: 3.5, 2011. RP11003.

[265] Van Rompaye, B., Jaffar, S., Goetghebeur, E. (2011). Cox models for estimation in cause-specific survival analysis faced with misclassified cause-of failure. *Epidemiology*, (in press), 2011. RP11009.

3.5 Universiteit Hasselt, UH

3.5.1 Technical reports

[266] Gaddah, A. and R. Braekers, A Goodness-of-fit test for a flexible copula Koziol-Green model, 2012. TR12001.

[267] Gijbels, I. , Omelka, M. and N. Veraverbeke, Multivariate and functional covariates and conditional copulas, 2012. TR12002.

[268] Omelka, M., Veraverbeke N. and I. Gijbels, Bootstrapping the conditional copula, 2012. TR12003.

3.5.2 Refereed publications (published)

[269] Adriaenssens, N., Coenen, S., Versporten, A., Muller, A., Minalu Ayele, G., Faes, C., Vankerckhoven, V., Aerts, M., Hens, N., Molenberghs, G. and H. Goossens, European Surveillance of

Antimicrobial Consumption (ESAC): Outpatient macrolide, lincosamide and streptogramin use in Europe (1997-2009). *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi37-vi45, 2011. R11190.

[270] Adriaenssens, N., Coenen, S., Versporten, A., Muller, A., Minalu Ayele, G., Faes, C., Vankerckhoven, V., Aerts, M., Hens, N., Molenberghs, G. and H. Goossens, European surveillance of antimicrobial consumption (ESAC): Outpatient quinolone use in Europe (1997-2009). *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi47-vi56, 2011. R11191.

[271] Adriaenssens, N., Coenen, S., Versporten, A., Muller, A., Minalu Ayele, G., Faes, C., Vankerckhoven, V., Aerts, M., Hens, N., Molenberghs, G. and H. Goossens, European Surveillance of Antimicrobial Consumption (ESAC): Outpatient antibiotic use in Europe (1997-2009). *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi3–vi12, 2011. R11187.

[272] Biard, R., Loisel, S., Macci, C. and N. Veraverbeke, Asymptotic behavior of the finite-time expected time-integrated negative part of some risk processes. *Journal of Mathematical Analysis an Applications*, **367**, 535-549, 2011. R11176.

[273] Birhanu, T., Molenberghs, G., Sotto, C. and M.G. Kenward, Doubly robust and multiple-imputation-based generalized estimating equations. *Journal of Biopharmaceutical Statistics,* **21**, 202-205, 2011. R11145.

[274] Braekers, R. and A. Gaddah, Flexible Modelling in the Koziol-Green Model by a Copula function. *Communications in Statistics: Theory and Methods*, **40**, 1218-1235, 2011. R11178.

[275] Coenen, S., Adriaenssens, N., Versporten, A., Muller, A., Minalu, G., Faes, C., Vankerckhoven, V., Aerts, M., Hens, N., Molenberghs, G. and H. Goossens, European Surveillance of Antimicrobial Consumption (ESAC): Outpatient use of tetracyclines, sulfonamides and trimethoprim, and other antibacterials in Europe (1997–2009). *Journal of Antrimicrobial Chemotherapy*, **66 Suppl**, vi57-vi70, 2011. R11192.

[276] Cornelissen, T., Molenberghs, G., Yperman, J., Schreurs, S. and R. Carleer, A statistical data-processing methodology of Py-GC/MS data for the simulation of flash co-pyrolysis reactor experiments. *Chemometrics and Intelligent Laboratory Systems*, **110**, 12-128, 2011. R11196.

[277] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and M.G. Kenward, Generalized shared-parameter models and missingess at random. *Statistical Modelling,* **11**, 270-310, 2011. R11068.

[278] Faes, C., Molenberghs, G., Hens, N., Coenen, S. and H. Goossens, Analysing the composition of outpatient antibiotic use: a tutorial on compositional data analysis. *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi89-vi94, 2011. R11194.

[279] Gaddah, A. and R. Braekers, An extension of the Koziol-Green model under dependent censoring. *Journal of Nonparametric Statistics*, **23**, 439-453, 2011. R11177.

[280] Gijbels, I., Veraverbeke, N. and M. Omelka, Conditional copulas, association measures and their applications. *Computational Statistics and Data Analysis*, **55**, 1919-1932, 2011. R11121.

[281] Iddi, S. and G. Molenberghs, A combined overdispersed and marginalized multilevel model. *Computational Statistics and Data Analysis*, **56**, 1944-1951, 2011. R11195.

[282] Janssen, P. ,Swanepoel, J. and N. Veraverbeke, Large sample behavior of the Bernstein copula estimator. *Journal of Statistical Planning and Inference*, **142**, 1189-1197, 2012. R12006.

[283] Minalu, G., Aerts, M., Coenen, S., Versporten, A., Muller, A., Adriaenssens, N., Beutels, P., Molenberghs, G., Goossens, H. and N. Hens, Application of mixed-effects models to study the country-specific outpatient antibiotic use in Europe: A tutorial on longitudinal data. *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi79-vi87, 2011. R11193.

[284] Molenberghs, G. and G. Verbeke, On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference*, **141**, 861-868, 2011. R11220.

[285] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, 21, 187-206, 2011. R11013.

[286] Molenberghs G., G.Verbeke, and Iddi S. (2011), 'Pseudo-likelihood methodology for partitioned large and complex samples,' *Statistics and Probability Letters*, **81**, 892-901, 2011. R11062.

[287] Mwambi, H., Ramroop, S., Shkedy, Z. and G. Molenberghs, An approach to estimating the force of infection and the rate of recovery fom Respiratory Syncytial Virus (RSV) data. *Statistical Methods in Medical Research*, **20**, 551-570, 2011. R11180.

[288] Poleto, F., Molenberghs, G., Paulino, C.D., and J.M. Singer, Sensitivity analysis for incomplete continuous data. *Test*, **20**, 589-606, 2011. R11179.

[289] Poleto, F.Z., Paulino, C.D., Molenberghs, G. and J.M. Singer, Inferential implications of over-parameterization: a case study in incomplete categorical data. *International Statistical Review*, **79**, 92-113, 2011. R11184.

[290] Pryseley, A., Tchonlafi, C., Verbeke, G., and G. Molenberghs, Estimating negative variance components from Gaussian and non-Gaussian data: a mixed models approach. *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011.

[291] Roca Pardiñas, Javier, Cadarso Suárez, Carmen María, Pardo Vázquez, JL, Leborán, V, Molenberghs, G., Faes, C., Acuña, Carlos, Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*. **30 (14)**, 1695-1711, 2011. John Wiley & Sons, Ltd., 2011. R11156.

[292] Rodriguez, P.Y.D., Coddens, A., Del Fava, E., Abrahantes, J.C., Shkedy, Z., Martin, L.O.M., Munoz, E.C., Duchateau, L., Cox, E. and B.M. Goddeeris, High prevalence of F4(+) and F18(+) Escherichia coli in Cuban piggeries as determined by serological survey. *Trop. Anim. Health Prod.*, **43**, 937-946, 2011. R11135.

[293] Sattar, A., Weissfeld, L.A. and G. Molenberghs, Analysis of non-ignorable missing and left-censored longitudinal data using weighted random effects tobit model. *Statistics in Medicine*, **30**, 3167-3180, 2011. R11186.

[294] Sotto, C., Beunckens, C., Molenberghs, G. and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2011. R11147.

[295] Vangeneugden, T., Molenberghs, G., Verbeke, G. and C. Demétrio, Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, **38**, 215-232, 2011. R11183.

[296] Van Keilegom, I. and N. Veraverbeke, Discussion on Statistical models and methods for dependence in Insurance data by Haug, S., Klüppelberg C. and Peng L. *Journal of the Korean Statistical Society*, **40**, 155-157, 2011. R11175.

[297] Van Nieuwenhuyze, J., Loosveldt, G. and G. Molenberghs, A Method for Evaluating Mode Effects in Mixed-mode Surveys. *Public Opinion Quarterly*, **74**, 1027-1045, 2011. R11185.

[298] Veraverbeke, N, Omelka, M. and I. Gijbels, Estimation of a conditional copula and association measures. *Scandinavian Journal of Statistics*, **38**, 766-780, 2011. R11172.

[299] Verbeke, G. and G. Molenberghs, Arbitrariness of models for augmented and coarse data, with emphasis on incomplete-data and random-effects models. *Statistical Modelling*, **10**, 391-419, 2011. R11181.

[300] Versporten, A., Coenen, S., Adriaenssens, N., Muller, A., Minalu Ayele, G., Faes, C., Hens, N., Vankerckhoven, V., Aerts, M., Hens, N., Molenberghs, G. and H. Goossens, European Surveillance of Antimicrobial Consumption (ESAC): Outpatient penicillin use in Europe (1997-2009). *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi13-vi23, 2011. R11188.

[301] Versporten, A., Coenen, S., Adriaenssens, N., Muller, A., Minalu, G., Faes, C., Vankerckhoven, V., Aerts, M., Hens, N., Molenberghs, G. and H. Goossens, European Surveillance of Antimicrobial Consumption (ESAC): Outpatient cephalosporin use in Europe (1997-2009). *Journal of Antimicrobial Chemotherapy*, **66 Suppl**, vi25-vi35, 2011. R11189.

[302] Vieira, A.M.C., Leandro, R.A., Demétrio, C.G.B. and G. Molenberghs, Double generalized linear model for tissue culture proportion data: A Bayesian perspective. *Journal of Applied Statistics*, **38**, 1717-1731, 2011. R11182.

[303] Wellens, N., Van Lancker, A., Flamaing, J., Gray, L., Moons, P., Verbeke, G., Boonen, S., Milisen, K. (2011). Interrater reliability of the interRAI Acute Care. *Archives of Gerontology and Geriatrics*, **19** (Epub ahead of print: DOI:10.1016/j.archger.2011.07.005), 2011. R11069.

3.5.3 Refereed publications (in press)

[304] Abegaz, F., Gijbels, I. and N. Veraverbeke, Semiparametric estimation of conditional copulas. *Journal of Multivariate Analysis*, (in press), 2012. RP12001.

[305] Aelvoet, W., Terryn, N., Molenberghs, G., De Backer, G., Vrints, C. and M. van Sprundel, Do inter-hospital comparisons of in-hospital, acute myocardial infarction case-fatality rates serve the purpose of fostering quality improvement? An evaluative study. *BMC Health Services Research*, (in Press), 2011. RP11035.

[306] Bunouf, P., Grouin, J.-M. and G. Molenberghs, Analysis of binary outcome derived from frequent longitudinal data: Application to daily pain evaluation. *Statistics in Medicine*, (in Press), 2011. RP11046.

[307] Bunouf, P., Grouin, J.-M., Molenberghs, G. and G. Koch, Analyzying frequently observed longitudinal data: Application to daily pain evaluation in clinical trials. *Journal of the Royal Statistical Society, Series A*, (in Press), 2011. RP11040.

[308] Creemers, A., Aerts, M., Hens, N. and G. Molenberghs, A nonparametric approach to weighted estimating equations for regression analysis with missing covariates. *Computational Statistics and Data Analysis*, (in Press), 2011. RP11042.

[309] Decramer, M., Molenberghs, G., Celli, B., Kesten, S., Lystig, T., Mehra, S. and D.P. Tashkin, behalf of the UPLIFT investigators, Premature discontinuation during the Understanding the Potential Long-term Impacts on Function with Tiotropium (UPLIFT) study: significance to efficacy results. *Respiratory Medicine*, (in Press), 2011. RP11039.

[310] Ghebretinsae, A.H., Faes, C., Molenberghs, G., De Boeck, M. and H. Geys, A Bayesian, generalized frailty model for comet assays. *Journal of Biopharmaceutical Statistics*, (in Press), 2011. RP11044.

[311] Kassahun, W., Neyens, T., Molenberghs, G., Faes, C. and G. Verbeke, Modeling overdispersed longitudinal binary data from the Jimma longitudinal studies using a combined beta and normal random-effects model. *Archives of Public Health*, (in Press), 2011. RP11051.

[312] Lin, D., Tilahun, A., Cortiñas, J., Shkedy, Z., Molenberghs, G., Goelhmann, H.W.H., Talloen, W., and L. Bijnens, Comparison of methods for the selection of genomic biomarkers. *International Journal of Data Mining and Bioinformatics*, 2011. RP11041.

[313] Luts, J., Molenberghs, G., Verbeke, G., Van Huffel, S. and J.A.K. Suykens, A mixed effects least squares support vector machine model for classification of longitudinal data. *Computational Statistics and Data Analysis*, (in Press), 2011. RP11038.

[314] Mallinckrodt, C.H., Molenberghs, G., Rubergh, S.J., Persinger, C., Sashegyi, A. and S.R. Lindborg, A portfolio-based approach to optimize proof of concept clinical trials. *Journal of Biopharmaceutical Statistics*, (in Press), 2011. RP11049.

[315] Milanzi, E., Alonso, A. and G. Molenberghs, Ignoring overdispersion in hierarchical loglinear models: Possible problems and solutions. *Statistics in Medicine*, (in Press), 2011. RP11047.

[316] Molenberghs, G., Kenward, M.G., Aerts, M., Verbeke, G., Tsiatis, A.A., Davidian, M. and D Rizopoulos D, On random sample size, ignorability, ancillarity, completeness, separability,

and degeneracy: sequential trials, random sample sizes, and missing data. *Statistical Methods in Medical Research,* 2011, in press. RP11002

[317] Neyens, T., Faes, C. and G. Molenberghs, A generalized Poisson-gamma model for spatially overdispersed data. *Spatial and Spatio-temporal Epidemiology*, (in Press), 2011. RP11045.

[318] Tomsin, K., Mesens, T., Molenberghs, G. and W. Gyselaers, Venous pulse transit time in normal pregnancy and pre-eclampsia. *Reproductive Sciences*, (in Press), 2011. RP11043.

[319] Van Nieuwenhuyze, J., Loosveldt, G. and G. Molenberghs, A method to evaluate mode effects on the mean and variance of a continuous variable in mixed mode surveys. *International Statistical Review*, (in Press), 2011. RP11046.

[320] Van Puyenbroeck, K., Hens, N., Coenen, S., Michiels, B., Beunckens, C., Molenberghs, G., Van Royen, P. and P. Verhoeven, Efficacy of daily intake of L. casei Shirota on respiratory symptoms and influenza vaccination immune response: a randomized double-blind placebo-controlled trial in healthy elderly nursing home residents *American Journal of Clinical Nutrition*, (in Press), 2011. RP11050.

[321] Vanstreels, L., Molenberghs, G. and J.U. Voigt, Controversies in stroke prevention. *Acta Cardiologica*, (in Press), 2011. RP11052.

[322] Verbeke, G., Fieuws, S., Molenberghs, G. and M. Davidian, The analysis of multivariate longitudinal data: A review. *Statistical Methods in Medical Research*, (in Press), 2011. RP11037

3.6 Université Joseph Fourier, UJF–LMC–IMAG

3.6.1 Technical reports

[323] Antoniadis, A., Gijbels, I. and Lambert-Lacroix, S., Penalized estimation in additive varying coefficient models using grouped regularization, 2011. TR11038.

[324] Antoniadis, A. and I. Gijbels and A. Verhasselt, Variable selection in varying coefficient models using P-splines, 2011. TR11053.

[325] Antoniadis, A. and X. Brossat, J. Cugliari and J.-M. Poggi, Multivariate Functional Time Series Clustering, 2011. TR11052.

[326] Giacofci, M., Lambert-lacroix, S., Marot, G. and F. Picard, Wavelet-based clustering for mixed-effects functional models in high dimension, submitted, 2011. TR11005.

[327] Lambert-Lacroix, S. and F. Letué, Partial Least Squares and Cox model with application to gene expression, 2011. TR11006.

[328] Touzani, S., *Response surface methods based on analysis of variance expansion for sensitivity analysis*, Appied Mathematics, PhD thesis, University of Grenoble, 20 April 2011, Grenoble, France, 2011. TR11050.

3.6.2 Refereed publications (published)

[329] Antoniadis, A., Gijbels, I. and M. Nikolova,Penalized likelihood regression for generalized linear models with nonquadratic penalties. *The Annals of the Institute of Statistical Mathematics*, Volume **63**, Issue 3, 585-615, 2011. R11197.

[330] Antoniadis, A., Gijbels I. and M. Nikolova, Penalized likelihood regression for generalized linear models with nonquadratic penalties. *Annals of the Institute of Mathematical Statistics*, **63**, 3, 585-616, 2011. R11168.

[331] Aseervatham, S., Antoniadis, A., Burlet, M., Denneulin, Y. and E. Gaussier, A sparse version of the ridge logistic regression for large-scale text categorization. *Pattern Recognition Letters*. **32**, 2, 101-106, 2010. R11014.

[332] Aseervatham, S., Antoniadis, A., Burlet, M., Denneulin, Y. and E. Gaussier, Régression logistique et catégorisation de textes, in *Modeles probabilistes pour l'accès à l'information textuelle*, Editors: E. Gaussier and F. Yvon, Hermes Science, 97-120, 2011. R11169.

[333] Boshnakov, G. and S. Lambert-Lacroix, Periodic Levinson-Durbin algorithm for entropy maximisation, *Computational Statistics and Data Analysis*, **56**, 15-24, 2011. R11170.

[334] Lambert-Lacroix, S. and L. Zwald, Robust regression through the Huber's criterion and adaptive lasso penalty. *Electronic Journal of Statistics*, **5**, 1015-1053, 2011. R11171.

3.6.3 Refereed publications (in press)

[335] Antoniadis, A., Helbert, C., Prieur, C., and L. Viry, Spatio-temporal prediction for West Africa monsoon, *Environmetrics*, (in press), 2011. RP11028.

[336] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines. *Journal of Computational and Graphical Statistics*, (to appear), 2012. RP12002.

[337] Cenac, P., Maume-Deschamps, V. and C. Prieur, Some multivariate risk indicators : minimization by using a Kiefer–Wolfowitz approach to the mirror stochastic algorithm, *Stat. Risk. Model.*, (in Press), 2011. RP11034.

[338] Di Bernardino, E., Laloë, V. Maume-Deschamps and C. Prieur, Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory, *ESAIM-Probab. Stat.*, (in Press), 2011. RP11033.

[339] Quaine, F., Paclet F., Letué, F. and F. Moutet, Force sharing and neutral line during finger extension tasks, *Hum. Mov. Sci.*, (in press), 2011. RP11031.

[340] Van Deun, K., Wilderjans, T., van den Berg, R.A., Antoniadis, A. and I. Van Mechelen, A flexible framework for sparse simultaneous component based data integration, *BMC Bioinformatics*, (in press), 2011. RP11030.

3.7 Erasmus Medical Center, EMC

3.7.1 Refereed publications (published)

[341] Atzberger, C. and P.H.C. Eilers, Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. *International Journal of Remote Sensing*, **32**, 3689-3709, 2011. R11212.

[342] Atzberger, C. and P.H.C. Eilers, A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America. *International Journal of Digital Earth*, **4**, 365-386, 2011. R11213.

[343] Hendrickx, D.M., M.M.W.B. Hendriks, P.H.C. Eilers, A.K. Smilde and H.C.J. Hoefsloot, Reverse engineering of metabolic networks, a critical assessment. *Molecular Biosystems*, **7**, 511-520, 2011. R11214.

[344] Kempen, G.T.H. van, H.T. van der Leest, R.J. van den Berg, P.H.C. Eilers and R.H.S. Westerink, Three Distinct Modes of Exocytosis Revealed by Amperometry in Neuroendocrine Cells. *Biophysical Journal*, **100**, 968-977, 2011. R11215.

[345] Marx, B.D., P.H.C. Eilers and B. Li, Multidimensional single-index signal regression. *Chemometrics and Intelligent Laboratory Systems*, **109**, 120-130, 2011. R11216.

[346] Rooi, J.J. de and P.H.C. Eilers, Deconvolution of pulse trains with the L_0 penalty. *Analytica Chimica Acta*, **705**, 218-226, 2011. R11217.

[347] Uh, H.-W. and P.H.C. Eilers, Haplotype Estimation from Fuzzy Genotypes Using Penalized Likelihood. *PLoS ONE* **6**, 2011. R11218.

[348] Verbeek, S., P.H.C. Eilers, K. Lawrence, R.C.M. Hennekam and F.G.A. Versteegh, Growth charts for children with Ellis-van Creveld syndrome. *European Journal Of Pediatrics*, **170**, 207-211, 2011. R11219.

3.7.2 Refereed publications (in press)

[349] Eilers, P.H.C., Sea Level Trend Estimation by Seemingly Unrelated Penalized Regressions. *Proceedings of the 26th International Workshop on Statistical Modelling*, 2011. RP11059.

[350] Rippe, R.C.A., and P.H.C. Eilers, Segmented smoothing with an L_0 penalty. *Proceedings of the 26th International Workshop on Statistical Modelling*, 2011. RP11060.

[351] Rooi, J.J. de and P.H.C Eilers, Mixture models for baseline estimation. *Chemometrics and Intelligent Laboratory Systems*, on-line, 2011. RP11058.

[352] Rooi, J.J. de and P.H.C. Eilers, Using text mining tools to compose structure priors for inferring gene networks. *Proceedings of the 26th International Workshop on Statistical Modelling*, 2011. RP11061.

[353] Schnabel, S.K., P.H.C. Eilers and F.A. van Eeuwijk, Optimal time scaling for plant growth analysis. *Proceedings of the 26th International Workshop on Statistical Modelling*, 2011. RP11062.

3.8 Universidad de Santiago de Compostela, USC

3.8.1 Technical reports

[354] Berrendero, J.R., Cuevas, A. and B. Pateiro-López, Multivariate uniformity tests: the distance to boundary method for the case of unknown support, 2011. TR11047.

[355] García-Portugués, E., Crujeiras, R. and W. González-Manteiga, Exploring wind direction and SO2 concentration by circularlinear density estimation, 2011. TR11049.

[356] Oviedo, M and M. Febrero-Bande, Utilities for Statistical Computing in Functional Data Analysis: The R package fda.usc, 2011. TR11048.

3.8.2 Refereed publications (published)

[357] Agra-García, P., Fernández-Casado, R., Crujeiras, R., Vázquez-Abal, M.E., Otero-Espinar, M.V., Cortés-Ayaso, A. (2011). Una nueva forma de vivir las Matemáticas: Matcampus 2010. *La Gaceta de la RSME*. **14**, 333-342, 2011. R11160.

[358] Berrendero, J.R., Cuevas, A., Pateiro-López, B., A multivariate uniformity test for the case of unknown support. *Statistics and Computing*. **22(1)**, 259-271, 2012. R12001.

[359] Carballo-Quintás, M., Martínez-Silva, I., Cadarso Suárez, C., Álvarez-Figueiras, M.M., Ares Pena, FJ, López Martín, E., A study of neurotoxic biomarkers, c-fos and GFAP after acute exposure to GSM radiation at 900 MHz in the picrotoxin model of rat brains. *Neurotoxicology*. **32**, 478-494, 2011. R11157.

[360] Crujeiras, R., Cortés-Ayaso, A., Fernández-Casado, R., Agra-García, P., Vázquez-Abal, M.E., Otero-Espinar, M.V., Matcampus 2010. *Gamma*. **10**, 31-37, 2011. R11161.

[361] Crujeiras, R., Prieto A, Discussion on "An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach". *Journal of the Royal Statistical Society, Series B.* **73**, 473-398, 2011. R11162.

[362] Diz, P., Pacheco, J.J., Otero Cepeda, Xosé Luis, Velazco, C., Tomás-Carmona, I., Relvas, M, Evaluation of partial-mouth recording systems of gingival parameters in a poruguese adult population. *Journal of Public Health Dentistry*, 2011. R11166.

[363] Faraldo-García, A., Santos-Pérez, S., Crujeiras, R., Labella-Caballero, T., Soto-Varela, A., Influence of age and gender in the sensory analysis of balance control. *European Archives of Oto-Rhino-Laryngology and Head & Neck*. **269**, 673-677, 2012. R12003.

[364] Ferraty F, González-Manteiga, W., Martínez Calvo, A. and P. Vieu, Presmoothing in functional linear regression. *Statistica Sinica*. **22**, 69-94, 2012. R12002.

[365] González-Quintela A., Fernández-Conde S., Alves-Pérez, M.T., Campos J., López-Ratón, M., Puerta R, Monte R, F Gude, Temporal and spatial patterns in the incidence of alcohol withdrawal syndrome in a defined community. *Alcohol*. **45 (2)**, 105-111, 2011. R11167.

[366] González-Manteiga, W., Crujeiras, R., Recent advances in spatio-temporal stochastic modelling. *Environmetrics*. **23**, 1-2, 2012. R12005.

[367] Gonzàlez-Manteiga, W., Pardo-Fernàndez, J.C. and I. Van Keilegom, ROC curves in nonparametric location-scale regression models. *Scandinavian Journal of Statistics*, 38, 169-184, 2011. R11010.

[368] Martínez Calvo, A. and W. González-Manteiga, Bootstrap in functional linear regression. *Journal of Statistical Planning and Inference*. **141**, 453-461, 2011. Elsevier Science Bv., 2011. R11152.

[369] Martín Acero, R., Fernández-del Olmo, M., Sánchez, J.A., Otero Cepeda, X.L., Aguado X., Rodríguez, F.A., Reliability of squat and countermovement jump tests in children 6-8 years of age. *Pediatric Exercise Science*. **23**, 151-160, 2011. R11163.

[370] Monsalve-Cobis, A.E., González-Manteiga, W., Febrero-Bande, M., Goodness of test for interest rate models: An approach based on empirical processes. *Computational Statistics and Data Analysis*. **55 (12)**, 3073-3092, 2011. R11158.

[371] Otero Cepeda, X.L., Franco, J., Martínez Insúa, A., Jimenez, P., Santana Penin, U.A., Maxillary changes and occlusal traits in crania with artificial fronto-occipital deformation. *American Journal of Physical Anthropology*, 2011. R11165.

[372] Otero Cepeda, X.L., Devesa, A., Otero, A., Vilarnovo, Z., Casteleiro, N., Castañón, N., Quintana A., Reimunde, P., Devesa, J., Effects of growth hormone (GH) replacement and cognitive rehabilitation in patients with cognitive disorders after traumatic brain injury. *Brain Injury*. **25(1)**, 65-72, 2011. R11164.

[373] Pata, M.P., Kneib T., Cadarso Suárez, C.M., Lustres-Pérez, V. and E. Fernández Pulpeiro, Categorical Structured Additive Regression for Assessing Habitat Suitability in the Spatial Distribution of Mussel Seed Abundance. *Environmetrics*. **23 (1)**, 75-84, 2012. R12004.

[374] Rodríguez-Álvarez, M.X., Garcia-Tahoces, P., Cadarso Suárez, C.M., Lado, M.J., Comparative study of ROC regression techniques. Applications for the computer-aided diagnostic system in breast cancer detection. *Computational Statistics & Data Analysis*. **55 (1)**, 888-902, 2011. Elsevier Science Bv., 2011. R11153.

[375] Rodríguez-Álvarez, M.X., Roca Pardiñas, J., Cadarso Suárez, C.M., A new flexible direct ROC regression model - Application to the detection of cardiovascular risk factors by anthropometric measures. *Computational Statistics and Data Analysis*. **55 (12)**, 3257-3270, 2011. R11159.

[376] Reyes Santías F, MX Rodríguez-Álvarez, Cadarso Suárez, C.M., Estimating hospital production functions through flexible. *Mathematical and Computer Modelling*. **54 (7-8)**, 1760-1764, 2011. Pergamon-Elsevier Science Ltd., 2011. R11154.

[377] Roca Pardiñas, J., Cadarso Suárez, C.M., Pardo Vázquez, JL, Leborán, V, Molenberghs, G., Faes, C., Acuña, C., Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*. **30 (14)**, 1695-1711, 2011. John Wiley & Sons, Ltd., 2011. R11156.

[378] Ruiz-Medina, M.D., Crujeiras, R., Minimum contrast parameter estimation for fractal random fields based on the wavelet periodogram. *Communications in Statistics: Theory and Methods*. **40**, 1-15, 2011. Taylor & Francis Inc., 2011. R11155.

3.8.3 Refereed publications (in press)

[379] Berrendero, J.R., Cuevas, A. and B. Pateiro-López, Testing uniformity for the case of a planar unknown support. *The Canadian Journal of Statistics*, (to appear), 2011. RP11011.

[380] Cuevas, A., Fraiman, R. and B. Pateiro-López, On statistical properties of sets fulfilling rolling-type conditions. *Advances in Applied Probability*, (to appear), 2011. RP11010.

[381] Fraiman, R. and B. Pateiro-López, Quantiles for finite and infinite dimensional data. *Journal of Multivariate Analysis*, (to appear), 2011. RP11014.

[382] Faraldo-García, A., Santos-Pérez, S., Crujeiras, R., Soto-Varela, A. and T. Labella-Caballero, Comparative study of computerized dynamic posturography and the Sway Star system in healthy subjects. *Acta Oto-Laryngologica*, (to appear), 2011. RP11012.

[383] Gayoso-Diz P., Otero-Gonzalez A., MX Rodríguez-Álvarez, F Gude, Cadarso Suárez, C.M., García F. and A. De Francisco, Insulin resistance index (HOMA-IR) levels in a general adult population: Curves percentile by gender and age. The EPIRCE study. *Diabetes Res Clin Pract*, (in press), 2011. RP11015.

[384] Pateiro-López, B. and A. Rodríguez-Casal, Recovering the shape of a point cloud in the plane. *TEST*, (to appear), 2011. RP11013.

[385] Reyes Santías F, Martínez Calvo, A. and C.M. Cadarso Suárez, Applying a simulation model to manage waiting lists for hospital inpatient activity in an EU region. *Mathematical and Computer Modelling*, (in press), 2011. RP11016.

3.8.4 Books (published)

[386] Chernick, M.R., González-Manteiga, W., Crujeiras, R., Barrios, E.B., *Bootstrap methods*. International Encyclopedia of Statistica Science, 2011. B11016.

[387] Crujeiras, R., Ruiz-Medina, M.D., *Wavelet-based minimum contrast estimation of linear Gaussian random fields*. Recent Advances in Functional Data Analysis and Related Topics. Physica-Verlag (Springer), 2011. B11004.

[388] González-Manteiga, W., Crujeiras, R., *A general view of the goodness-of-fit tests for statistical models*. Modern Mathematical Tools and Techniques in Capturing Complexity, 2011. B11002.

[389] González-Manteiga, W. and P. Vieu, *Methodological Richness of Functional Data Analysis*. Learning and Data Science, 2011. B11003.

[390] Roca Pardiñas, J., Lado, M.J., Cadarso Suárez, C.M. and P. Garcia-Tahoces, *Reducing False Positives in a CAD Scheme for Detecting Breast Microcalcificacions: A Cuantitative Study with GAMs. Cancer Prevention*, 2011. B11005.

3.9 London School of Hygiene and Tropical Medicine, LSHTM

3.9.1 Technical reports

[391] Carpenter, J.R., Roger, J.H. and M.G. Kenward,Relevant, Accessible Sensitivity Analyses Using Multiple Imputation, 2011. TR11046.

3.9.2 Refereed publications (published)

[392] Birhanu, T., Molenberghs, G., Sotto, C., and M.G. Kenward, Doubly robust and multiple-imputation-based generalized estimating equations. *Journal of Biopharmaceutical Statistics,* **21**, 202-205, 2011. R11145.

[393] Carpenter, J.R., Goldstein, H. and M.G. Kenward, REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types. *Journal of Statistical Software,* **45**, 1-14, 2011. R11149.

[394] Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and M.G. Kenward Generalized shared-parameter models and missingess at random. *Statistical Modelling,* **11**, 270-310, 2011. R11068.

[395] Kenward, M.,G. and and G. Rosenkranz, Joint modelling of outcome, observation time and missingness. *Journal of Biopharmaceutical Statistics*, **21**, 252-262, 2011. R11146.

[396] Lee, Y., Mangseok, N., and M.G. Kenward, Robust estimation of dropout models using hierarchical likelihood. *Journal of Statistical Computation and Simulation*, **81**, 693-706, 2011. R11148.

[397] Molenberghs G., Kenward, M.G., Verbeke, G. and T. Birhanu, 'Pseudo-likelihood estimation for incomplete data,' *Statistica Sinica*, **21**, 187-206, 2011. R11013.

[398] Petersen, I., Carpenter, J., Welch, C., Walters, K., Morris, R., White, I., Marston, L. and I. Nazareth, Missing Data Imputation in Electronic Health Databases: Development of a Longitudinal Model for Cardiovascular Risk Factors. *Pharmacoepidemiology and Drug Safety*, **20**, 231-232, 2011. R11150.

[399] Rucker, G., Carpenter, J.R. and G. Schwarzer, Detecting and adjusting for small-study effects in meta-analysis. *Biometrical Journal*, **53**, 351-368, 2011. R11151.

[400] Sotto, C., Beunckens, C., Molenberghs, G. and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2011. R11147.

3.9.3 Refereed publications (in press)

[401] Bartlett, J.W., Frost, C. and J.R. Carpenter, Multiple imputation models should incorporate the outcome in the model of interest. *Brain*, 2011, in press. RP11004.

[402] Daniel, R.M. and M.G. Kenward, A method for increasing the robustness of multiple imputation. *Computational Statistics and Data Analysis,* 2011, in press. RP11001.

[403] Daniel, R.M., Kenward, M.G., Cousens, S. and B. de Stavola, Using directed acyclic graphs to guide analysis in missing data problems. *Statistical Methods in Medical Research*, 2011, in press. RP11003.

[404] Molenberghs, G., Kenward, M.G., Aerts, M., Verbeke, G., Tsiatis, A.A., Davidian, M. and D. Rizopoulos, On random sample size, ignorability, ancillarity, completeness, separability, and degeneracy: sequential trials, random sample sizes, and missing data. *Statistical Methods in Medical Research,* 2011, in press. RP11002

3.10 List of joint publications

3.10.1 Technical reports

[405] Antoniadis, A., Gijbels, I. and Lambert-Lacroix, S., Penalized estimation in additive varying coefficient models using grouped regularization, 2011. TR11038. (KUL1, UJF)

[406] Antoniadis, A. and I. Gijbels and A. Verhasselt, Variable selection in varying coefficient models using P-splines, 2011. TR11053. (KUL1, UJF)

[407] Gaddah, A. and R. Braekers, A Goodness-of-fit test for a flexible copula Koziol-Green model, 2012. TR12001. (UH, UCL)

[408] Gijbels, I. , Omelka, M. and Veraverbeke, N. (2012). Multivariate and functional covariates and conditional copulas, 2012. TR12002. (UH, KUL1)

[409] Omelka, M., Veraverbeke N. and I. Gijbels, Bootstrapping the conditional copula, 2012. TR12003. (UH, KUL1)

3.10.2 Refereed publications (published)

[410] Antoniadis, A., Gijbels I. and M. Nikolova, Penalized likelihood regression for generalized linear models with nonquadratic penalties. *Annals of the Institute of Mathematical Statistics*, **63**, 3, 585-616, 2011. R11168. (KUL1, UJF)

[411] Antoniadis, A., Gijbels, I. and M. Nikolova,Penalized likelihood regression for generalized linear models with nonquadratic penalties. *The Annals of the Institute of Statistical Mathematics*, Volume **63**, Issue 3, 585-615, 2011. R11197. (KUL1, UJF)

[412] Birhanu, T., Molenberghs, G., Sotto, C. and M.G. Kenward, Doubly robust and multiple-imputation-based generalized estimating equations. *Journal of Biopharmaceutical Statistics,* **21**, 202-205, 2011. R11145. (UH, LSHTM)

[413] Braekers, R. and A. Gaddah, Flexible Modelling in the Koziol-Green Model by a Copula function. *Communications in Statistics: Theory and Methods*, **40**, 1218-1235, 2011. R11178. (UH, UCL)

[414] Creemers A., Hens N., Aerts M., Molenberghs G., G.Verbeke, and Kenward M.G. (2011), 'Generalized shared-parameter models and missingness at random,' Statistical Modelling, 11, 279-310, 2011. R11068. (UH, KUL2, LSHTM)

[415] Daouia, A. and I. Gijbels, Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, **161**, 147-165, 2011. R11201. (KUL1, UCL)

[416] Daouia, A. and I. Gijbels, Estimating frontier cost models using extremiles. In *Festschrift in honor of Leopold Simar*, Editors: P. Wilson, I. Van Keilegom. Springer-Verlag: Berlin, Heidelberg, ISBN 978-3-7908-2348-6. Chapter 4, pages 65-81, 2011. R11202. (KUL1, UCL)

[417] Gaddah, A. and R. Braekers, An extension of the Koziol-Green model under dependent censoring. *Journal of Nonparametric Statistics*, **23**, 439-453, 2011. R11177. (UH, UCL)

[418] Gijbels, I., Veraverbeke, N. and M. Omelka, Conditional copulas, association measures and their applications. *Computational Statistics and Data Analysis*, **55**, 1919-1932, 2011. R11121. (KUL1, UH)

[419] Gonzàlez-Manteiga, W., Pardo-Fernàndez, J.C. and I. Van Keilegom, ROC curves in non-parametric location-scale regression models. *Scandinavian Journal of Statistics*, 38, 169-184, 2011. R11010. (USC, UCL)

[420] Janssen, P. and L. Duchateau, Comments on: Inference in multivariate Archimedean copula models. *TEST*, **20**, 271-275, 2011. R11132. (UH, UG)

[421] Molenberghs, G., Kenward, M.G., Verbeke, G. and B. Teshome Ayele, Pseudo-likelihood estimation for incomplete data. *Statistica Sinica*, **21**, 187-206, 2011. R11013. (UH, KUL2, LSHTM)

[422] Molenberghs, G. and G. Verbeke, On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference*, **141**, 861-868, 2011. R11220. (UH, KUL2)

[423] Molenberghs G., G.Verbeke, and Iddi S. (2011), 'Pseudo-likelihood methodology for partitioned large and complex samples,' *Statistics and Probability Letters*, **81**, 892-901, 2011. R11062. (UH, KUL2)

[424] Pryseley, A., Tchonlafi, C., Verbeke, G., and G. Molenberghs, Estimating negative variance components from Gaussian and non-Gaussian data: a mixed models approach. *Computational Statistics and Data Analysis*, **55**, 1071-1085, 2011. R11011. (UH, KUL2)

[425] Quataert, P., Verschelden, P., Janne Elise, B., Verbeke, G., Goetghebeur, E. and F. Ollevier, A diagnostic modelling framework to construct indices of biotic integrity: a case study of fish in the Zeeschelde estuary (Belgium). *Estuarine Coastal and Shelf Science*, **94**, 222-233, 2011. R11128. (KUL2, UG)

[426] Roca Pardiñas, Javier, Cadarso Suárez, Carmen María, Pardo Vázquez, JL, Leborán, V, Molenberghs, G., Faes, C., Acuña, Carlos, Assessing neural activity related to decision-making through flexible odds ratio curves and their derivatives. *Statistics in Medicine*. **30 (14)**, 1695-1711, 2011. John Wiley & Sons, Ltd., 2011. R11156. (UH, USC)

[427] Rodriguez, P.Y.D., Coddens, A., Del Fava, E., Abrahantes, J.C., Shkedy, Z., Martin, L.O.M., Munoz, E.C., Duchateau, L., Cox, E. and B.M. Goddeeris, High prevalence of F4(+) and F18(+) Escherichia coli in Cuban piggeries as determined by serological survey. *Trop. Anim. Health Prod.*, **43**, 937-946, 2011. R11135. (UH, UG)

[428] Sotto, C., Beunckens, C., Molenberghs, G. and M.G. Kenward, MCMC-based estimation methods for continuous longitudinal data with non-random (non)-monotone missingness. *Computational Statistics and Data Analysis*, **55**, 301-311, 2011. R11147. (UH, KUL2, LSHTM)

[429] Vangeneugden, T., Molenberghs, G., Verbeke, G. and C. Demétrio, Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, **38**, 215-232, 2011. R11183. (UH, KUL2)

[430] Van Keilegom, I. and N. Veraverbeke, Discussion on Statistical models and methods for dependence in Insurance data by Haug, S., Klüppelberg C. and Peng L. *Journal of the Korean Statistical Society*, **40**, 155-157, 2011. R11175. (UH, UCL)

[431] Veraverbeke, N, Omelka, M. and I. Gijbels, Estimation of a conditional copula and association measures. *Scandinavian Journal of Statistics*, **38**, 766-780, 2011. R11172. (KUL1, UH)

[432] Verbeke, G. and G. Molenberghs, Arbitrariness of models for augmented and coarse data, with emphasis on incomplete-data and random-effects models. *Statistical Modelling*, **10**, 391-419, 2011. R11181. (UH, KUL2)

3.10.3 Refereed publications (in press)

[433] Abegaz, F., Gijbels, I. and N. Veraverbeke, Semiparametric estimation of conditional copulas. *Journal of Multivariate Analysis*, (in press), 2012. RP12001. (UH, KUL1)

[434] Aelvoet, W., Terryn, N., Molenberghs, G., De Backer, G., Vrints, C. and M. van Sprundel, Do inter-hospital comparisons of in-hospital, acute myocardial infarction case-fatality rates serve

the purpose of fostering quality improvement? An evaluative study. *BMC Health Services Research*, (in Press), 2011. RP11035. (KUL1, UH)

[435] Antoniadis, A., Gijbels, I. and A. Verhasselt, Variable selection in varying coefficient models using P-splines. *Journal of Computational and Graphical Statistics*, (to appear), 2011. RP12002. (KUL1, UJF)

[436] De Roy, K., Clement, L., Thas, O., Wang, Y. and N. Boon, Flow cytometry for fast microbial community fingerprinting. *Water Research*, (to appear). IF: 4.966, 2011. RP11004. (KUL2, UG)

[437] Kassahun, W., Neyens, T., Molenberghs, G., Faes, C. and G. Verbeke, Modeling overdispersed longitudinal binary data from the Jimma longitudinal studies using a combined beta and normal random-effects model. *Archives of Public Health*, (in Press), 2011. RP11051. (UH, KUL2)

[438] Luts, J., Molenberghs, G., Verbeke, G., Van Huffel, S. and J.A.K. Suykens, A mixed effects least squares support vector machine model for classification of longitudinal data. *Computational Statistics and Data Analysis*, (in Press), 2011. RP11038. (UH, KUL2)

[439] Molenberghs, G., Kenward, M.G., Aerts, M., Verbeke, G., Tsiatis, A.A., Davidian, M. and D Rizopoulos D, On random sample size, ignorability, ancillarity, completeness, separability, and degeneracy: sequential trials, random sample sizes, and missing data. *Statistical Methods in Medical Research,* 2011, in press. RP11002. (UH, KUL2, LSHTM)

[440] Thas, O., De Neve, J., Clement, L. and J.P. Ottoy, Probabilistic Index Models (Read paper, with discussion). *Journal of the Royal Statistical Society*, Series B, **74**, 1-29, (to appear). IF: 3.5, 2011. RP11003. (KUL2, UG)

[441] Van Deun, K., Wilderjans, T., van den Berg, R.A., Antoniadis, A. and I. Van Mechelen, A flexible framework for sparse simultaneous component based data integration, *BMC Bioinformatics*, (in press), 2011. RP11030. (KUL1, UJF)

[442] Verbeke, G., Fieuws, S., Molenberghs, G. and M. Davidian, The analysis of multivariate longitudinal data: A review. *Statistical Methods in Medical Research*, (in Press), 2011. RP11037. (UH, KUL2)