

General Information	
<b>Name of Research Unit:</b>	(MAT-LVT-6) Centro de Estatística e Aplicações da Universidade de Lisboa
<b>Coordinator:</b>	Maria Antonia Conceição Abrantes Amaral Turkman
<b>Main Scientific Domain:</b>	Matemática
<b>Other Subdomains:</b>	n/a
Host Institutions	
<b>Leading Host Institution:</b>	Faculdade de Ciências - Universidade de Lisboa
<b>Other Institutions Involved:</b>	
Objectives & Achievements	
Unit Description	
<p>The Center of Statistics and Applications of the University of Lisbon (CEAUL-Centro de Estatística e Aplicações da Universidade de Lisboa), created in 1975, is a Unit I&amp;D, close to Department of Statistics and Operations Research (DEIO) of Faculty of Sciences of the University of Lisbon (FCUL:Host Institution), which offers physical facilities to CEAUL- infrastructures, installations, budget management and contribution of a technician - to the prosecution of CEAUL's activities in accordance with article 2 of the Regulation of Program of Financing Plurianual of I&amp;D (PFP I&amp;D); financial support is assured in the picture of PFP I&amp;D FCT, being the Foundation of FCUL the Management Institution.</p> <p>The Internal Organization of CEAUL's team (50 PhD Integrated (I) and 46 Collaborators (C) ) members in 31 December 2009) was systematized in 4 Research Groups (GI's), sharing similar scientific interests in Applied Probability, Statistical Inference, Probability Modelling and Data Analysis, but with diverse emphasis in methods and applications and accordingly with the Planning Subproject of each senior Principal Investigator. The groups were</p> <ul style="list-style-type: none"> <li>• GI 1: Order Statistics, Extremes and Applications (11 (I) + 8 (C)) Responsible Researcher: Maria Ivette Leal de Carvalho Gomes</li> <li>• GI 2: Probability, Modelling and Data Analysis (18 (I) + 13 (C)) Responsible Researcher: Dinis Duarte Ferreira Pestana</li> <li>• GI 3: Statistical Modelling in Environmental and Life Sciences (16 (I) + 22 (C)) Responsible Researcher: Maria Antonia Conceição Abrantes Amaral Turkman</li> <li>• GI 4: Applied Multivariate Data Analysis and Modelling (5 (I) + 3 (C)) Responsible Researcher: Maria Helena Bacelar Valente da Costa Nicolau</li> </ul> <p>In the last evaluation, as it was referred to in the last report, the International Panel has given a classification of VERY GOOD to the GLOBAL UNITY, although reports asymmetries between Groups' performances. In particular, for the group GI 4, there has been a strong recommendation about the integration of this group in the main strategy of CEAUL.</p> <p>As a consequence of the Panel recommendations, at a general meeting of the Scientific Committee in March 2009, it was set a schedule to restructure the Unity in 2010. The Scientific Committee met on 25th of February 2010 and approved a proposal by the scientific panel, supported by the Advisory Board, to extinguish the research group GI 4: "Applied Multivariate Data Analysis and Modelling" and integrate their members in one of the remaining 3 groups, according to their research interests. As such, upon approval by the FCT, there will be only the first three research groups mentioned above.</p> <p>The GI's are encouraged to interact between them in their activities, whenever that is appealing.</p> <p>The groups leaders, together with the elect Coordinator and Vice-Coordinator, are the core of the scientific panel whose responsibility is to monitor and inflect the main goals of the research unit, and to rule the investment policies.</p> <p>The Internal Regulation of CEAUL provides a Scientific Committee, which approves in a plenary meeting the general guidelines of Scientific and Financial management: a part of the budget for common expenses (Periodicals and Books for the Library, Common Software and Informatics Assistance as well as maintenance of WebPage) and the remaining is shared among the GI's, according to the number of PhDs in each GI.</p> <p>There is a scientific as well as a financial autonomy of the GIs.</p> <p>CEAUL has 3 associated FCT Projects during 2009: FCT/PPCDT/ MAT/58876/2004 "ERAS" PI: M. Ivette Gomes FCT/ PTDC/MAT/64924/2006"EXES PI: Laurentius de Haan FCT/ PTDC/MAT/64353/2006 "MEGA" PI: Prof. K.F.Turkman</p>	
General Objectives	
<p>This I&amp;D Unit plays a central role in Portugal as a Research Center in Probability, Statistics and its Applications, developing multidisciplinary activities that aim to carry out research in domains as diverse as Health and Life Sciences and Biology, Financial Risk and hazards posed by nature in Geophysics and Environment. Together with applied research, fundamental research is carried out on several areas: Probability, Stochastic processes, Statistics of Extremes, Computational Statistics and Simulation, Resampling Methodologies, Quality Control, Bayesian Statistics, Biostatistics, Statistics in Genomics and Proteomics, Survival Analysis, Clinical Trials, Time Series, Data Analysis, Sampling, Classification methods, Temporal and Spatial Statistics, Design of Experiments and Teaching of Statistics.</p> <p>The main objective of CEAUL is to contribute to the development of new theoretical framework and of methodological issues addressed to applications to real problems, making efforts to correct use of Statistics in Society and Academics. That is attained through different inter-linked activities, among which we mention Conferences and Seminars, Scientific Publications both in National and International peer review journals, Post-graduate courses (MSc, PhD), collaboration with researchers in other areas, and an updated Library to community.</p> <p>Another objective involves the links with National Statistical Institutions, namely National Institute of Statistics (INE) and Portuguese Statistical Society (SPE).</p> <p>Some of the senior members of CEAUL are deeply involved with the implementation of REVSTAT – Statistical Journal, an International Journal in the</p>	

area of Probability and Statistics in English.

Main objectives of GI1 were in the field of statistics of univariate, multivariate, multidimensional and spatial extremes, with special emphasis on their applications to Life Sciences, Environment, Risk, Insurance, Finance and Statistical Quality Control.

GI2 main interests have been in Meta analysis and applications in Medicine, Structure of population matrices and Principal Components Analysis, History of Science, Fractal issues in population models, scale and location problems. Analysis of binary longitudinal data, distance sampling models, Teaching of Statistics, mixture models, Analysis of messy data, Simulation studies of the Rényi's rarefaction.

The main objective of Group GI3 was to make contributions on specific probability and statistical issues in modeling environmental and biological data, as well as, to create a group with expertise in analyzing data coming from these fields. The areas of research were, spatial extremes ; hierarchical bayesian spatiotemporal analysis; stochastic modelling of dynamic systems; bayesian latent class models; non-linear time series; missing data mechanisms; bayesian classification methods; predictive quality of diagnostic tests; survival analysis and longitudinal models; bayesian analysis of allelic penetrance models for complex binary traits; ecological studies; modelling infectious diseases; syndromic surveillance; microarray data analysis.

### **Main Achievements during the year of 2009**

CEAUL as a whole has been increasing the scientific production. His members have been making a special effort in order to publish their work on the mainstream high quality journals in each field and to disseminate their main results in well recognized international meetings. As a consequence, during 2009 CEAUL had the following production at international level:

- 54 publications in peer review International Journals.
- 23 publications in International proceedings and 2 chapters of books
- 47 abstracts in international conferences and at the national level
- 4 books
- 8 publications in national periodicals
- 25 publications in national proceedings.

Members of the center have been involved in software development. An R package to estimate the baseline of mortality of influenza epidemics is already made available to the R community.

CEAUL members have not disregarded the national counterpart of their post-graduation teaching activities by involving themselves in the supervision of PhD and MSc students. During 2009, 6 Ph. D. students and 12 M. Sc. Students completed their theses under the supervision of integrated members of the center. Some members of the center co-supervised master students which completed their theses during 2009, but this was not accounted for, particularly because those theses are part of the program of integrated master degrees. This may justify the apparent drop in master supervisions relatively to previous years.

Besides the publication in the mainstream domains in the areas of Probability and Statistics, there has been some investment in communicating the importance of Statistics to general society and attracting young students to the area.

The CEAUL has been involved, through his members, in the organization of national (2) and international conferences (1), international workshops (1), international thematic sessions (2), seminars (6), one-day workshops (2) and advanced courses (5). These are dully described in each group productivity report.

The collaboration of CEAUL with SPE has been fruitful with the involvement of Daniel Paulino and Dinis Pestana in the production of the Portuguese Statistical Glossary and in the involvement of Fernando Rosado who acts as the editor of the SPE newsletter.

Members of CEAUL have kept deep international contacts and several members have been invited speakers at international conferences and workshops. International involvement is well described by each group coordinator as part of the group productivity.

### **Activities**

#### **Integrative/multidisciplinary activities during the year of 2009**

There has been a direct evidence of the interaction between the theoretical and applied components of the work. During 2009 members of CEAUL actively collaborated with researchers from several research institutions in different fields, resulting in international publications. The motivating problems of collaboration were on diverse areas such as: health, forest fires, chemical production, radioactivity, psychology and education. This is well documented in each research productivity group report.

Some members of CEAUL integrate the research teams of international projects, such as the DEEPFISHMAN project, regarding the Black scabbardfish abundance estimate in the Portuguese coastal waters and the projects on "Maternal Deaths and Severe Maternal Morbidity in Maputo City and Province" and "Support to Integrated Malaria Control in the Chókwè Region – Moçambique.

We also highlight bellow some of the CEAUL integrative/multidisciplinary activities in 2009, with reference to Links with Portuguese Institutions:

- Collaborations with:
- CEAUL / CIM

CEAUL is an associate of CIM (International Center for Mathematics), which promotes research in Mathematics and interdisciplinary conferences; In 2009 Cristina Rocha, from GI3, organized a one-day workshop on "Survival Analysis" (Lisbon November 2009), in collaboration with SPE. Invited speakers were: Giovanni Silva (GI3), Ana Abreu (University of Madeira), Guadalupe Gomez (University of Barcelona)

- CEAUL / SPE

SPE (Sociedade Portuguesa de Estatística) counts with the collaboration of CEAUL in several Integrative/multidisciplinary activities; in 2009 I Natario was a member of the organizing committee of the, "XVII Congresso da Sociedade Portuguesa de Estatística" (Sesimbra, October 2009): the short course "Análise de sobrevivência" was given by C Rocha and A L Papoila (GI3). D. Paulino, and D. Pestana are responsible for the Portuguese part in the ISI multilingual Statistics dictionary.

- CEAUL / INE



No. of Researchers Hired (LA)	0,00	0,00	0,00	0,00	0,00	0,00
<b>Balance</b>	0,00	0,00	0,00	0,00	0,00	0,00
No. of Researchers Hired (Ciência Programme)	0,00	0,00	0,00	1,00	0,00	1,00
No. of Researchers (FTE)	34,00	36,00	41,00	43,00	50,00	
Training Masters (Master thesis completed)	13,00	14,00	4,00	34,00	12,00	77,00
Training PhDs (PhD thesis completed)	5,00	9,00	16,00	11,00	6,00	47,00

#### Researchers Hired

Name	Start Date	End Date	Other Institution
Jorge Miguel Diamantino Miranda	01-04-2009	30-03-2010	

#### Technical Personnel Hired

Name	Start Date	End Date	Other Institution
No technical personnel found...			

#### Additional Comments

##### Additional Comments

Jorge Miguel Diamantino Miranda has been appointed by Ciência2008 Programme, which gave the opportunity to bilateral collaboration with another I&D Unity from FCUL, through a basic level BII\_Project "Spatial Heterogeneity of soil surface characteristics", announced in July 2008, with the objective of studying which soil characteristics are influential to the vegetation of the soil. The work started during 2009.

The Hired Grant Researcher Student BII\_2008 has been integrated in G13 for one year, J Miranda, through the Ciência2008 Programme and was supervised by V Andreozzi (G13), who interacts also with Centre for Environmental Biology (Professor C Cruz, CBA) FCUL, in Initiation to Research with Multidisciplinary activities Statistical Modelling and Environmental Biology.

#### Research Groups

Reference	Title / Principal Investigator
RG-LVT-6-948	<u>Order Statistics, Extremes and Applications</u> (MARIA IVETTE LEAL DE CARVALHO GOMES)
RG-LVT-6-949	<u>Probability, Modelling and Data Analysis</u> (Dinis Duarte Ferreira Pestana)
RG-LVT-6-951	<u>Statistical Modelling in Environmental and Life Sciences</u> (Maria Antonia Conceição Abrantes Amaral Turkman)
RG-LVT-6-1519	<u>Applied Multivariate Data Analysis and Modelling</u> (Maria Helena Bacelar Valente da Costa Nicolau)

Group Description	
<b>Title of Research Group:</b>	(RG-LVT-6-948) Order Statistics, Extremes and Applications
<b>Principal Investigator:</b>	MARIA IVETTE LEAL DE CARVALHO GOMES
<b>Main Scientific Domain:</b>	Matemática
<b>Group Host Institution:</b>	Faculdade de Ciências - Universidade de Lisboa
Funding, source, dates	
<b>Funding, source, dates</b>	
<p>ALL THE REFERRED AMOUNTS INCLUDE OVERHEADS TO PAY FFCUL</p> <p>1. FCT Base for Research Group Title: (RG-MATH-LVT-Lisboa-6-948) - Order Statistics, Extremes and Applications :</p> <p>Funding:</p> <p>corresponding to 2009:</p> <p>date: 3 April 2009 / 26.943,00 Eur</p> <p>date: 7 October 2009/ 9.472,20 Eur</p> <hr/> <p>2. Projects: The 2 following Projects have Principal Investigators from the Group Title: (RG-MATH-LVT-Lisboa-6-948) - Order Statistics, Extremes and Applications :</p> <p>FCT/PPCDT/ MAT/58876/2004</p> <p>"ERAS — EXTREMES, RISK, SAFETY and the ENVIRONMENT"</p> <p>Principal Investigator: M. Ivette Gomes</p> <p>Starting: January 1, 2007; Finishing: June 30, 2009</p> <p>Funding: 14280 Euros.</p> <p>FCT/ PTDC/MAT/64924/2006</p> <p>"EXES — SPATIAL EXTREMES"</p> <p>Principal Investigator: Laurentius de Haan</p> <p>Starting: July 2007; Finishing: 31 December 2010</p> <p>Funding: 60000 Euros</p>	
Objectives & Achievements	
<b>Objectives</b>	
<p>The objectives of the team for 2009 were mainly in the field of statistics of univariate, multivariate, multidimensional and spatial extremes, with special emphasis on their applications to Life Sciences, Environment, Risk, Insurance, Finance and Statistical Quality Control. Those objectives were related with the topics reproduced from the future research objectives in the 2008 Report:</p> <ol style="list-style-type: none"> <li>1) Adaptive estimation of univariate and multivariate parameters of extreme events. Computer-intensive methods, like the bootstrap, will be the basis of new methodologies of adaptive choice of thresholds mainly for second-order reduced-bias, now involving, with a high probability, the third-order behaviour of the underlying models.</li> <li>2) Due to the importance of random censored data in Life Sciences and Survival Analysis, one of the topics to be considered under this Research Group, we feel the need to develop adequate methodology in the area of extreme values, related with this type of data.</li> <li>3) Establishment of a link among recent developed methodologies, like the PORT methodology, with PORT standing for Peaks Over Random Thresholds, the MVRB, with MVRB standing for Minimum-Variance Reduced-Bias, generalized KERNEL and PWM methodologies, with PWM standing for Probability Weighted Moments.</li> <li>4) Development of new estimators of the endpoint of the underlying model, whenever finite and either univariate or multivariate. This is an extremely important parameter in the analysis of athletic events, related with the value of the ultimate world record, but also in all areas where tails are light or of an exponential type.</li> <li>5) Other important parameters under consideration will be the Value-at-Risk at any level <math>p</math>, the size of the loss occurred with small probability <math>p</math>, the adjustment coefficient, a rudimentary measure of risk in a collective of insurance risks, and the Weibull tail coefficient, the regular variation coefficient of the inverse failure rate function.</li> <li>6) In all the developed estimation procedures, we intend to consider, additionally to independent, identically underlying structures, diversified weak dependent structures as well as positively (or negatively) associated sequences of random variables. For dependent frameworks, we shall try to improve the estimation of the extremal index, of the pre-asymptotic extremal index and of the coefficient of asymptotic tail dependence.</li> <li>7) Development of a new methodology for assessing the adjustment of pRRARMAX (power Random Autoregressive for Maxima) processes to real data in the fields of finance, environment and biology. Inference for extremes of specific max-autoregressive processes, including the pRRARMAX, will be thus one of the topics under consideration.</li> </ol>	

8) All the new methodologies will be applied not only to simulated data but also to real data in the above mentioned areas, and some of them will be devised for the analysis of spatial data.

### Main Achievements

We refer separately the main achievements in the different topics of research:

**Adaptive Estimation of Parameters of Extreme Events.** The research on this topic began by the mid of 2009, and it is yet under development. Apart from a paper in international proceedings, [I4], we refer the abstract [I15], on the adaptive estimation of the extremal index, an important parameter in dependent frameworks, [I16], related with an invited conference at the 4th Workshop on Statistics Mathematics and Computation, and a paper presented at the XVI Congress of the Statistical Portuguese Society (SPE) (see [N24]). See also [N27], presented at the XVI SPE, and CEAUL 07/2009 preprint, available at <http://www.ceaul.fc.ul.pt/>

**Random Censoring and Biometrics.** In 2009, we have practically achieved no progresses in the fulcral part of this topic. A related paper, specifically in the area of Biometrics, is [P12].

**Linking PORT, MVRB, KERNEL and PWM methodologies for the extreme value index (EVI) estimation.** The articles [P2], [P11] and [N5] are essentially on MVRB methodologies, with a link to KERNEL estimation. A link between PORT and MVRB estimation has been achieved in [P10] and [N10]. On testing issues, we mention the articles [P7], [I8], [N14] and [N15]. Although not explicitly stated, the link between KERNEL and PORT can be found in [P6], where a new simple and efficient EVI-estimator has been introduced. The QUASI-PORT approach has been advanced in [I9]. It is also worth mentioning an invited communications at EVA 2009 (see [I12]), as well as another invited communication at the Workshop on the Jubilee of Professor J. Tiago Mexia (see [I17]). See also CEAUL 01/2009, 10/2009 and 27/2009 preprints, also available at CEAUL's website.

**Finite Endpoint Estimation and Estimation of other Parameters of Extreme Events.** As main achievements we refer [P1], on second-order reduced-bias high quantile estimation, as well as [P3], [P16], [P17] and [N18], in the area of max-semistability. See also [N9], for the analysis of data in athletics. Essentially in areas related with these topics, we also refer [P9], an invited book review. See also CEAUL 06/2009 preprint.

**Dependent Frameworks and Extremes of Stochastic Processes.** The main achievements on these topics are related with the articles [P4], [I3] and [N17], on EXPOS methods, [P8], where RCINAR processes have been studied, [I5], with a study of the mixed moment estimators in dependent frameworks, and [N26]. We also refer [I1], on modelling mortality, [I7] and [I19], on calibration methods, and [I11] and [N6], on pRARMAX processes. See also the preprints, CEAUL 03/2009, on this last topic. On random fields and spatial statistics, we mention [P14]. See also [N11], [N17], [N23], [N25] and [N26]. It is also worth mentioning an invited communication at EVA 2009, on estimating tail probabilities of the integral of a stochastic process (see [I13]) and on the estimation of extreme quantile regions for two dependent risks (see [I10]), as well as two invited communication at the Workshop on the Jubilee of Professor J. Tiago Mexia (see [I18] and [I19]). See also CEAUL 11/2009 preprint.

**8. Monte-Carlo Simulation and Data Analysis in the Fields of Extremes, Outliers, Forensic Statistics and Statistical Process Control.** In the area of control charts, we mention [P5], a paper on robust methodologies. See also [N8] and [N22]. The article [P15] and [I14] are publications in the related area of Reliability Theory. See also [N28]. The expected payoff to Internet auctions has been studied in [P13]. See also [I2], [I6], [N7], [N13] and [N19]. Regarding forensic statistics, see [N20], [N21], CEAUL 08/2009 and CEAUL 16/2009 preprints. Most of the material produced includes Monte-Carlo simulations, either in a small or even a large-scale.

The book, [N1], the papers [N2], [N3], [N4], published at Boletim SPE, the paper [N12], related with an invited conference at the XVI SPE meeting, and [N25] are transversal to all research developed by the team.

### Group Productivity

#### Publications in peer review Journals

[P1] Caeiro, F. & Gomes, M.I. (2009). Semi-parametric second-order reduced-bias high quantile estimation. *Test* 18:2, 392-413.

[P2] Caeiro, F., Gomes, M.I. & Henriques Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Commun. Statist. — Theory & Methods* 38:7, 1019-1040.

[P3] Canto e Castro, L. & Dias, S. (2009). Asymptotic distribution of certain statistics relevant to the fitting of max-semistable models. *Port. Math. (N.S.)* 66, No. 3, 401-412.

[P4] Cordeiro, C. & Neves, M.M. (2009). Forecasting time series with BOOT.EXPOS procedure. *Revstat* 7: 2, 135-149.

[P5] Figueiredo, F. & Gomes, M.I. (2009). Monitoring industrial processes with robust control charts. *Revstat* 7: 2, 151-170.

[P6] Fraga Alves, M.I., Gomes, M.I., de Haan, L. & Neves, C. (2009). Mixed Moment Estimators and Location Invariant Alternatives. *Extremes* 12, 149-185. DOI: 10.1007/s10687-008-0073-3, October 2008.

[P7] Fraga Alves, M.I., de Haan, L. & Neves, C. (2009). A test procedure for detecting super-heavy tails. *J. Statist. Planning and Inference* 139 (2), 213-227.

[P8] Gomes, D. & Canto e Castro, L. (2009). Generalized integer-valued random coefficient for a first order structure autoregressive (RCINAR) process. *J. Stat. Plann. Inference* 139, No. 12, 4088-4097.

[P9] Gomes, M.I. (2009). Nonparametric Analysis of Univariate Heavy-tailed Data: Research and Practice, by N. Markovich. *Wiley Series in Probability and Statistics*, 2007 (Book Review). *Biometrics* 65:1, 336-336.

[P10] Gomes, M.I. & Henriques-Rodrigues, L. (2009). High quantile estimation and the PORT methodology. *Revstat* 7:3, 245-264.

[P11] Gomes, M.I., Pestana, D.D. & Caeiro, F. (2009). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statistics and Probability Letters* 79, 295-303.

[P12] Guerreiro C.S., Ferreira P., Tavares L., Moura Santos P., Neves M.M., Brito M. and Cravo M. (2009). "Fatty acids, IL-6 and TNF- $\alpha$  polymorphisms: an example of nutrigenetics in Crohn's disease". *The American Journal of Gastroenterology* 104, 2241-2249.

[P13] de Haan, L., Vries, C.G. & Zhou, C. (2009). The expected payoff to Internet auctions. *Extremes* 12:3, 219-238.

[P14] Kabluchko, Z., Schlather, M. & de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*. Volume 37, Number 5, 2042-2065.

[P15] Reis, P & Canto e Castro, L. (2009). Limit Model For The Reliability of a Regular And Homogeneous Series-Parallel System, *REVSTAT – Statistical Journal*, Vol. 7:3, 227-243 2009.

[P16] Temido, M.G. & Canto e Castro, L. (2009). On a Second Order Condition for Max-Semistable Laws. *Pliska Studia Mathematica Bulgarica* 19, 97-

110, (2009).

[P17] Temido, M.G., Canto e Castro, L. & Dias, S. (2009). Tail Inference for a Law in Max-Semistable Domain of Attraction. *Pliska Studia Mathematica Bulgarica* 19 83-96, (2009).

#### Other publications International

##### PROCEEDINGS

[I1] Gomes, D., Nunes, C. & Canto e Castro, L. (2009). Cold and Heat waves: modelling mortality in Évora - Portugal. In *Proceeding of the 23rd International Workshop on Statistical Modelling* (Paul H. C. Eilers, editor); pp. 250-55.

[I2] Carmo, C., Moreira, J. & Miranda, C. La calidad de los resultados e el costo de la financiación bancaria, XV Congreso da Asociación Española de Contabilidad y Administración de Empresas (27 pp), Electronic publication: [http://www.aeca.es/pub/on\\_line/comunicaciones\\_xvcongresoaeaca/general.htm](http://www.aeca.es/pub/on_line/comunicaciones_xvcongresoaeaca/general.htm)

[I3] Cordeiro, C. & Neves, M.M. Bootstrapping EXPOS methods: a case study. In Chi Kin Chan et al. (eds.), *ISF 2009 Proceedings*, 67 (28 pages). See <http://forecasters.org/submissions09/CordeiroClaraISF2009.pdf>

[I4] Gomes, M.I. Mendonça, S. & Pestana, D.D. Adaptive reduced-bias tail index and value-at-risk estimation. In L. Sakalauskas et al. (eds.). *Applied Stochastic Models and Data Analysis — ASMDA 2009*, IMI & VGTU Editions, 41-44.

[I5] Gomes, M.I. & Miranda, C. Finite sample behaviour of the mixed moment estimator in dependent frameworks. In Luzar-Stifler et al. (eds.). *Proceedings of the ITI 2009*, University of Zagreb Editions, 237-242.

[I6] Gomes, M.I., Pestana, D., Sequeira, F., Mendonça, S. & Velosa, S. Uniformity of offsprings from uniform and non-uniform parents. In Luzar-Stifler et al. (eds.). *Proceedings of the ITI 2009*, University of Zagreb Editions, 243-248.

##### ABSTRACTS

[I7] Prata Gomes, D., Mexia, J.T. and Neves, M.M. Extremal behaviour of stationary processes: the calibration technique. In de Haan, L. and Turkman, K.F. (eds.). *Spatial Extremes, Theory and Applications* (SETA 2009).

In Colubi, A. et al. (eds.), *CFE'09 and ERCIM 2009: Book of Abstracts*, 2009:

[I8] Araújo Santos, P. & Fraga Alves, M.I. Conditional EVT for VaR estimation: comparison with a new independence test, 16.

[I9] Gomes, M.I. & Figueiredo, F. A quasi-PORT methodology for VaR based on second-order reduced bias estimation, 43.

In Breidt J. (ed.), *Graybill VIII/EVA2009 Book of Abstracts*, Colorado State Univ. editions, 2009:

[I10] Einmahl, J. & de Hann, L. Estimating extreme quantile regions for two dependent risks, 86.

[I11] Ferreira, M. & Canto e Castro, L. Modeling rare events through a pRARMAX process, 53.

[I12] Gomes, M.I. A quasi-PORT methodology for VaR: second-order reduced-bias estimation, 40.

[I13] de Haan, L., Ferreira, A. and Zhou, C. On estimating tail probabilities of the integral of a stochastic process, 86.

[I14] Reis, P. and Canto e Castro, L. Ultimate and penultimate models for the reliability of regular and homogeneous series-parallel systems, 93.

[I15] Neves, M.M. and Prata Gomes, D. A computational diagnostic for the threshold in the extremal index estimation, 92.

In Oliveira, T. and Oliveira, A. (ed.). *Program and Abstracts: 4th Workshop on Statistics, Mathematics and Computation*, Univ. Aberta Editions, 2009:

[I16] Gomes, M.I. Computational Methods in Statistics of Extremes: Adaptive Choice of Thresholds, 24.

In Pestana, D. et al. (eds). *Workshop em Estatística pela Jubilação do Professor João Tiago Mexia, Edições CMA/FCT-UNL*, 2009:

[I17] Gomes, M.I. & Henriques-Rodrigues, L. Classical and MVRB tail index estimators: asymptotic comparison at optimal levels, 7-8.

[I18] Neves, M.M. and Cordeiro, C. Exponential Smoothing and Resampling Techniques in Time Series Prediction, 9.

[I19] Prata Gomes, D., Tiago Mexia, J. and Neves, M.M. Extremal Behaviour of stationary processes: the calibration technique, 10.

#### Other publications National

##### BOOKS

[N1] Canto e Castro, L. & Graça Martins, M.E. (2009). *Estatística Descritiva com Excel*, 73-153, Edições INE.

##### NATIONAL JOURNALS

Boletim SPE, Primavera de 2009:

[N2] Fraga Alves, M.I. Investigação (em) estatística: CEAUL — Quem fomos, somos e para onde vamos, 40-49.

[N3] Gomes, M.I. Algumas reflexões avulsos sobre a Sociedade Portuguesa de Estatística, 50-52.

[N4] Rosado, F. Introduzindo Outliers, 60-64.

##### PROCEEDINGS

In Oliveira, I., M.M. et al. (eds.), *Estatística: Arte de Explicar o Acaso*, Edições SPE, 2009:

[N5] Caeiro, F. Gomes, M.I. & Pestana, D. Alguns resultados adicionais sobre a variância de um estimador de viés reduzido, 167-177.

[N6] Ferreira, M. & Canto e Castro, L. Um método de ajustamento do modelo pRARMAX, 255-266

[N7] Figueiredo, A.M., Figueiredo, F.O. & Pimenta-Monteiro, N. Partilha de rendas nos bancos privatizados em Portugal, 275-286.

[N8] Figueiredo, F.O. & Gomes, M.I. Estimação das normas de um processo na fase inicial de implementação de uma carta de controlo, 287-298.

[N9] Gomes, M.I. & Pestana, D. Caudas leves em desporto: estimação de parâmetros úteis, 307-318.

[N10] Henriques-Rodrigues, L. & Gomes, M.I. A metodologia PORT na estimação semi-paramétrica de um parâmetro de escala, 59-79.

[N11] Miranda, M.C. & Gomes, M.I. Comparação de estimadores do índice de cauda em estruturas dependents, 437-443.

[N12] Neves, M.M. & Santos, J.A. "Modelação e estimação não paramétrica e semiparamétrica, 59-79.

#### EXTENDED ABSTRACTS

In Braumann, C. et al. (eds.) XVI Congresso da Sociedade Portuguesa de Estatística: Programa e Livro de Resumos, Edições SPE, 2009:

- [N13] Alves da Rocha, C.D. & Ferreira, A. Uma aplicação da teoria de valores extremos a dados meteorológicos recolhidos em Évora, 369.
- [N14] Araújo Santos, P. & Fraga-Alves, M.I. Avaliação de modelos de previsão com um novo teste de independência, 222-223.
- [N15] Araújo Santos, P. & Fraga-Alves, M.I. EVT condicional na estimação de quantis em series heterocedásticas: um estudo comparativo, 237-239.
- [N16] Caeiro, F., Gomes, M.I. & Pereira, H. Comparação assintótica em níveis ótimos de estimadores do índice de cauda de viés reduzido, 100-101.
- [N17] Cordeiro, C. & Neves, M.M. Previsões usando BOOT.EXPOS e aplicação no tratamento de dados omissos, 153-154.
- [N18] Dias, S. & Canto e Castro, L. Estimação da componente fractal de leis max-semiestáveis, 202-203.
- [N19] Ferreira, M. & Canto e Castro, L. Ajustamento do modelo pRARMAX: uma generalização do método e uma aplicação a dados financeiros, 331-333.
- [N20] Figueira-Neves, M. A Estatística no tribunal, 353.
- [N21] Figueira-Neves, M. & Rosado, F. Bayes no tribunal, 354.
- [N22] Figueiredo, F. & Gomes, M.I. Cartas de controlo baseadas na função quantil, 114.
- [N23] Gomes, D., Temido, M.G. & Canto e Castro, L. Comportamento extremal de processos de valores inteiros RCINAR(1) generalizados, 276-277.
- [N24] Gomes, M.I. Estimação adaptativa MVRB: aplicação a dados financeiros, 207-208.
- [N25] Gonçalves, I. & Canto e Castro, L. Estudo de acontecimentos raros em contextos multivariados: da teoria à prática, 278.
- [N26] Miranda, M.C. & Gomes, M.I. Estimação do índice de cauda com redução do viés em contexto dependente — estudo comparativo com recurso ao método de Monte Carlo, 349-350.
- [N27] Prata Gomes, D., Mexia, J.T. & Neves, M.M. Comportamento extremal de sucessões estacionárias: aplicação da calibração linear na estimação do índice extremal, 235-236.
- [N28] Reis, P. & Canto & Castro, L. Modelos para fiabilidades de sistemas paralelo-série e série-paralelo de grande dimensão, 165-166.

#### Master and Ph.D. thesis completed

Ph. D. Thesis:

[PhD1] Henriques-Rodrigues, Lígia. "Estimação de Viés Reduzido em Estatística de Extremos". Universidade de Lisboa.

Local: Universidade de Lisboa

Orientador: M. Ivette Gomes.

M. Sc. Thesis:

[MSc1] José Francisco de Oliveira Perpétua Loff, "Estudo Epidemiológico da Aplicação em Portugal de uma Vacina contra o Vírus do Papiloma Humano", Mestrado em Matemática Aplicada às Ciências Biológicas, ISA/UTL, April 2009.

Orientador: M.M. Neves

[MSc2] Carla Daniela Moreira Bessa, "A Correlação Revisitada", Mestrado em Estatística Aplicada e Modelação, FEUP, December 2009.

co-Orientador: M.M. Neves

#### Organization of conferences

5. Organization of conferences (2000 ca.)

1. "3º Encontro, A Matemática nas Ciências Biológicas", ISA, January 4, 2009.

Organizer: M.M. Neves.

2. "XVI Jornadas de Classificação e Análise de Dados", April 2-4, 2009, Campus de Gambelas, Universidade do Algarve

M.M. Neves was a member of the Scientific Committee.

3. "Spatial Extremes, Theory and Applications (SETA 2009)", April 6-8, 2009, Lisboa. Laurens de Haan and Luísa Canto e Castro were among the organizers.

4. "4º Encontro, A Matemática nas Ciências Biológicas", ISA, July 14, 2009.

Organizer: M.M. Neves.

5. "6th Conference on Extreme Value Analysis — Probabilistic and Statistical Models and their Applications (EVA 2009)", Fort-Collins, Colorado, July, 22-26, 2009. M.I. Gomes was a member of the Scientific Committee and organizer of a Thematic Session on "Statistics of Extremal Events".

6. "TV Workshop on Statistics, Mathematics and Computation", Lisboa, 9-10 November, 2009.

Members of this Group in the Scientific Committee: M.I. Gomes & M.M. Neves.

7. Short course on Statistical Analysis with R. Took place in ISA and M Neves lectured part of the course

#### Industry contract research

6. Industry contacts research (2000 ca.)

We have kept on the contacts with the Portuguese Institute of Quality ("Instituto Português da Qualidade"), Portucel and Auto-Europa in order to introduce in the industry the recent developed robust methods in Statistical Process Control. Up to this date, there was not yet any kind of formal contract.

#### Internationalization

7. Internationalization (2000 ca.)



(Collaborative publication, Research, Graduate Training Networks or other forms of participation of the Research Group at the international level)

All members of this team have kept deep international contacts and are well-recognized internationally not only in the area of statistics of extremes, but also in the areas of spatial, environmental and financial statistics, risk assessment theory and statistical quality control. As a result of this, 3 members of the team, Ana Ferreira, M. Ivette Gomes and Laurens de Haan, have been invited participants in the Project RISK, RARE EVENTS AND EXTREMES, EPFL, Lausanne, Switzerland, July 1 until December 31, 2009 (<http://extremes.epfl.ch/>). There was collaborative publication with J. Einmahl, Z. Kabluchko, M. Schlather, C. de Vries and C. Zhou. The internationalization of the team was responsible for the number of invited lectures in 2009, either national or international: Luísa Canto e Castro (2.º Fórum e-Estatística – Numeracia e Cidadania, Lisboa), Ana Ferreira (SETA 2009, Lisboa; EPFL, Lausanne) Fernanda Figueiredo (seminar at FCUP; 1st Porto Meeting on Math. for Industry; ISSPC, Nantes; seminar at Universidade do Minho), M. Ivette Gomes (TUM, Munchen; Workshop on Statistics in Honor of Tiago Mexia Jubilee (WSHTMJ), UNL, Monte de Caparica; EVA 2009, Fort-Collins; ASMDA 2009, Vilnius; EPFL; FCG, Lisboa); Laurens de Haan (EPFL; Fundan Univ., China); Manuela Neves (WSHTMJ), UNL, Monte de Caparica). We also would like to emphasize the strong co-operation with researchers in the Erasmus School of Economics, Erasmus University Rotterdam and the Tinbergen Institute. The research contacts with the Katholieke Universiteit Leuven have also become stronger in recent years, and, jointly with Jan Beirlant, we are involved in the organization of an Invited Paper Meeting at ISI 2011, Dublin. Strong research contacts have also been kept with T.A. Buishand (Royal Netherlands Meteor. Inst.), Dejuan Li (Fundan Univ., China), Holger Drees (Saarland Univ., Germany), Jan Picek (Liberec Univ., Check Republic), Jap Geluk (The Petroleum Inst., United Arab Emirates), Jurg Husler (Bern University, Switzerland), Laszlo Markus (Eotvos Lorand Univ., Hungary), Liang Peng (Georgia Pol. Inst., USA) and Rolf Dieter-Reiss (Siegen Univ., Germany), among others. Also as a result of the internationalization of the team, we can mention the enormous number of reviews produced by members of the team in the most diversified journals

#### Government/Organization contract research

Nothing to report in this field

#### Future Research

##### Objectives

The main objectives during the nearby future are again in the field of statistics of univariate, multivariate, multidimensional and spatial extremes, with special emphasis on their applications to Life Sciences, Environment, Risk, Insurance, Finance, Statistical Quality Control, Outliers and Biometry. Topics in the area of Design of Experiments will also be of related interest. Among the topics to be dealt with, we mention:

- 1) The development of heuristic data-driven methods for threshold selections will be one of the most relevant topics of research in the nearby future. We shall also pay attention to more sophisticated adaptive choices of thresholds via the bootstrap methodology and other resampling techniques.
- 2) The areas of Statistical Quality Control (SQC), Experimental Design and Multivariate Data Analysis will be also fundamental areas of research in the nearby future. In the area of SQC, special emphasis will be given to the role of the quantile function in the building of robust control charts. Small balance samples will be used to study the power of randomization tests and the multivariate Statis approach will be used to deal with a possible restructuration of privatized firms.
- 3) Due to the great importance of random censored data in Life Sciences and Survival Analysis, one of the topics to be developpe by members of this Research Group, we feel the need to further develop new adequate methodology for the analysis of this type of data, not only in the field of extremes but also in areas like design of experiments and forensic statistics.
- 4) We shall proceed to a deeper asymptotic comparison, at optimal levels, of minimum-variance reduced-bias (MVRB) estimators of the extreme value index and to the development of new estimators of several relevant parameters of extreme events: the endpoint (an extremely important parameter in the analysis of athletic events, related with the value of the ultimate world record, but also in all areas where tails are light or of an exponential type, like most environmental data sets in the fields of hidrology and climatology), the Value-at-Risk at any level  $p$ , the size of the loss occurred with small probability  $p$ , the adjustment coefficient, a rudimentary measure of risk in a collective of insurance risks, the Weibull tail coefficient, the regular variation coefficient of the inverse failure rate function and the estimation of extreme multivariate quantile regions.
- 5) In all the developed methodology, we intend to consider, additionally to independent, identically underlying structures, diversified weak dependent structures as well as positively (or negatively) associated sequences of random variables. We shall also deal with diversified stochastic processes, including the pRARMAX processes and spatial processes. Exceedance probabilities of integrals of stochastic processes and extreme residual dependence for random vectors and processes will be two other topics under consideration in the nearby future.

##### Funding, source, dates

All the funding received during 2009 and reported in the section 1 of this group description report was transferred to 2010

The below-mentioned source funding was available during 2009:

FCT/PPCDT/ MAT/58876/2004

"ERAS — EXTREMES, RISK, SAFETY and the ENVIRONMENT"

Principal Investigator: M. Ivette Gomes

Starting: January 1, 2007; Finishing: June 30, 2009

Funding: 14,280 Euros.

FCT/ PTDC/MAT/64924/2006

"EXES — SPATIAL EXTREMES"

Principal Investigator: Laurentius de Haan

Starting: July 2007; Finishing: 31 December 2010

Funding: 60,000 Euros

From 2010 onwards, apart from FCT/ PTDC/MAT/64924/2006, members of this Research Group had the following application approved:

FCT/PTDC/MAT/101736/2008

Title: "EXTREMA: Statistics of Extremes in Today's World"

Principal Investigator: M. Ivette Gomes

Other responsible researchers: Luísa Canto e Castro, M. Isabel Fraga Alves and M. Manuela Neves,

Budget: 96,720 Euros

Starting date: January 1, 2010

Group Description	
<b>Title of Research Group:</b>	(RG-LVT-6-949) Probability, Modelling and Data Analysis
<b>Principal Investigator:</b>	Dinis Duarte Ferreira Pestana
<b>Main Scientific Domain:</b>	Matemática
<b>Group Host Institution:</b>	Faculdade de Ciências - Universidade de Lisboa
Funding, source, dates	
<b>Funding, source, dates</b>	
<p>ALL THE REFERRED AMOUNTS INCLUDE OVERHEADS TO PAY FFCUL</p> <p>FCT Base for Research Group Title: (RG-MATH-LVT-Lisboa-6-949) - Probability, Modelling and Data Analysis</p> <p>Funding:</p> <p>corresponding to 2009:</p> <p>date: 3 April 2009 / 16.839,40 Eur</p> <p>date: 7 October 2009 / 9.945,72 Eur</p>	
Objectives & Achievements	
<b>Objectives</b>	
<p>A• Meta analysis and applications in Medicine, namely in studies of desmoplastic melanoma. Uniformity issues related to combined p-values and random p-values. Reserach team: F. Sequeira, F. Brilhante, S. Mendonça, D. Pestana, L. Soares de Almeida, R. Vasconcelos, M. Malva, J. P. Martins.</p> <p>B• Structure of population matrices and Principal Components Analysis. Research team: J. L. Cadima and co-workers from ISA.</p> <p>C• History of Science — Research in Probability and Statistics in Portugal during the first Republic. Research team: R. Santos, S. Velosa, E. Graça Martins, D. Pestana, S. Maciel Pires, E. Mendes de Oliveira.</p> <p>D• Fractal issues in population models — analysis of the Gompertz model. Research team: J. L. Rocha, S. Aleixo, D. Pestana.</p> <p>E• Scale analysis using the independence of scale and location estimators. Research team: D. Pestana, M. Marques dos Santos, F. Brilhante, S. Mendonça, M. Malva, M. Reis, H. Iglésias Pereira.</p> <p>F• Analysis of binary longitudinal data, in view of medical applications. Research team: S. Cabral, H. Gonçalves.</p> <p>G• Further work on errors in distance sampling models, and modeling bird death rates in eolic plants. Research team: T. A. Marques, jointly with R. Bispo Carita (ISPA) and his co-workers at St.Andrews.</p> <p>H• Teaching of Statistics: Documentation for teachers, and collaboration with INE's Alea Project. Research team: E. Graça Martins, E. Mendes de Oliveira.</p> <p>I• Extensions of Dorfman's theory on analysis of combined blood samples: specificity and sensibility issues in quantitative tests; mixture models. Research team: D. Pestana, M. Felgueiras, R. Sousa, H. Iglésias Pereira, R. Vasconcelos.</p> <p>J• Analysis of messy data, and development of models for contingency tables. Research team: J. Paulo Martins, R. Vasconcelos, P. Semblano.</p> <p>K• Simulation studies of the Rényi's rarefaction, and applications to sampling theory. Research team: S. Aleixo, F. Diamantino, S. Mendonça, D. Pestana, H. Iglésias Pereira.</p> <p>L• Scientific translation in the area of Probability and Statistics, and development of an e-learning platform for professional translators for TRADULÍNGUAS. Research team: C. Ventura, S. Velosa, D. Pestana.</p>	
<b>Main Achievements</b>	
<p>Three research members of the group got their Ph. D. degree in 2009.</p> <p>A• Meta analysis and uniformity issues, combined p-values and random p-values, medical applications — [P2.13], [P2.20], [S2.3], [S2.4], [S2.5], [PHD2.3], [ON2.6], [ON2.12], [ON2.13], [ON2.14], [ON2.21] investigate the special absorbing properties of uniformity, relating them to entropy, and establish limitations to the techniques of computational augmented samples,</p> <p>B• Structure of population matrices and Principal Components Analysis — [P2.3], [P2.4], [P2.5], [ON2.15]</p> <p>C• History of Science — Research in Probability and Statistics in Portugal — Alguns "Clássicos da Probabilidade e da Estatística", Évora Congress on the History of Mathematics, D. Pestana and R. Santos.</p> <p>D• Fractal issues in population models — [P2.14], [P2.15], [P2.16], [P2.17], [P2.21], [S2.1], namely showing the entropy properties of betas, and the influence of skewness and kurtosis-like parameters on the dynamical issues involving beta(p,2) models.</p> <p>E• Scale analysis →→ [P2.2] with simile of ANOVA using spacings, for inference on the location parameter of exponentials. Work on the revised and augmente 2nd edition of Brilhante, Pestana, Rocha and Velosa book on inference on location and scale.</p> <p>F• Analysis of binary longitudinal data — [S2.6], [MSC2.2], [ON2.8], [ON2.9], [ON2.16], [ON2.18], with applications, and general work on mixed effects models.</p> <p>G• Further work on errors in distance sampling models, and modeling bird death rates in eolic plants — [P2.11], [P2.12], [ON2.11], [ON2.19]</p> <p>H• Teaching of Statistics — [S2.2], [ON2.4]</p> <p>I• Extensions of Dorfman's theory — [ON2.7], [ON2.22]</p> <p>K• Simulation studies of the Rényi's rarefaction, and applications to sampling theory — ON2.10].</p> <p>L• Scientific translation in the area of Probability and Statistics — collaboration with the firm Tradulínguas on e-learning for translators of scientific</p>	

material (D. Pestana, C. Ventura, S. Velosa)

The collaboration with Research Group 1, on inference procedures, continued [P2.9], [P2.10], [S2.7], [S2.8], [S2.9], [ON2.1], [ON2.2], [ON2.3]. Other inferential problems, namely on data transformations and design [PhD2.1], [ON2.20] and on mixtures [PhD2.2], [ON2.17], are under investigation.

## Group Productivity

### Publications in peer review Journals

[P2.1] Agra Coelho, C., and Pestana, D. (2009). Biography of João Tiago Praça Nunes Mexia, *Discussiones Mathematicae Probability and Statistics* 29, 95–113.

[P2.2] Brilhante, M. F., Mendonça, S. Pestana, D., and Rocha, M. L. (2009). Inference on the location parameter of exponential populations. *Discussiones Mathematicae Probability and Statistics* 29 (galley proof returned, in print)

[P2.3] Bruno-Soares, A.M., Cadima, J. & Matos, T.J.S. (2009). Predicting degradability parameters of diets for ruminants using regressions on chemical components. *Journal of the Science of Food and Agriculture* (accepted).

[P2.4] Cadima, J. & Jolliffe, I. (2009). On Relationships Between Uncentred and Column-Centred Principal Component Analysis, *Pakistan Journal of Statistics*, 25(4), 473-503.

[P2.5] Cadima, J., Calheiros, F.L. & Preto, I.P. (2009). The eigenstructure of block-structured correlation matrices and its implications for Principal Component Analysis. *Journal of Applied Statistics* (accepted).

[P2.6] Caneco, A., Grácio, C., Fernandes, S., J. Rocha, J. L., and Ramos, C. (2009). Synchronizability and Graph Invariants, *ISAST Transactions on Computers and Intelligent Systems*, 2, 47-52.

[P2.7] Caneco, A., Rocha J. L., and Clara Grácio (2009). Topological Entropy in the Synchronization of Piecewise Linear and Monotone Maps. Coupled Duffing Oscillators, *International Journal of Bifurcation & Chaos*, 19, 11 (2009), 1-14.

[P2.8] Caneco, A., Rocha J. L., and Clara Grácio (2009). Kneading Theory Analysis of the Duffing Equation, *Chaos, Solitons & Fractals*, 42 (2009), 1529-1538.

[P2.9] Gomes, M. I., Henriques-Rodrigues, L., Pereira, H., and Pestana, D. (2009). Tail index and second order parameters' semi-parametric estimation based on the log-excesses. *J. Statisti. Comput. and Simul.*

(accepted, DOI 10.1080/00949650902755178).

[P2.10] Gomes, M. I., Pestana, D., and Cairo, F. (2009). "A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator", *Statistics and Probability Letters* vol. 79, 295-303.

[P2.11] Marques, T. A., Thomas, L., Ward, J., DiMarzio, N. & P. L. Tyack (2009). Estimating cetacean population density using fixed passive acoustic sensors: an example with Blaineville's beaked whales. *The Journal of the Acoustical Society of America*. 125: 1982-1994.

[P2.12] Marques, T.A., Buckland, S.T., Andersen, M., Belikov, S., Boltunov, A. & Wiig, Ø. (2009). Estimating the Barents Sea polar bear subpopulation size. *Marine Mammal Science*. 25: 35-52

[P2.13] Pereira, C., Bernardo, M., Pestana, D., Costa Santos, J., Mendonça, M. C. (2009). Contribution of Teeth in Human Forensic Identification - Discriminant function sexing odontometrical techniques in Portuguese Population. *J. Forensic and Legal Medicine* 17, in print.

[P2.14] Pestana, D. D., Aleixo, S. M., and Rocha, J. L. (2009). The Beta(p,1) extensions of the random (uniform) Cantor sets. *Discussiones Mathematicae Probability and Statistics* 29 (galley proof returned, in print)

### Other publications International

#### CHAPTERS OF BOOKS

[P2.15] Aleixo, S., Rocha, J. L., and Pestana, D. (2009). Populational Growth Model Proportional to Beta Densities, in Peixoto, M. M; Pinto, A.A.; Rand, D. A. J. (Eds.), *Dynamics, Games and Science*, in honour of Maricio Peixoto and David Rand (in print).

#### PROCEEDINGS

[P2.16] Aleixo, S.M, Rocha, J. L., and Pestana, D.D. (2009). Populational Growth Models Proportional to Beta Densities with Allee Effect, *Proceedings of the BVP 2008, International Conference on Boundary Value Problems: Mathematical Models in Engineering, Biology and Medicine*, Espanha, American Institute of Physics, vol. 1124, 3-12.

[P2.17] Aleixo, S.M, Rocha, J. L., and Pestana, D.D. (2009). Dynamical Behaviour in the Parameter Space: New Populational Growth Models Proportional to Beta Densities. In Luzar-Stiffler, V., Jarec, I. and Bekic, Z. (eds.), *Proceedings of the ITI 2009, 31st International Conference on Information Technology Interfaces*, p. 213- 218. (IEEE: 10.1109/ITI.2009.5196082)

[P2.18] Caneco, A., Grácio, C., Fernandes, S., J. Rocha, J. L., and Ramos, C. (2009). Relationship Between the Behavior of Entropy, the Conductance and the Coefficient of Clustering of Networks Using Symbolic Dynamics, *Proceedings of the NOMA'09, International Workshop on Nonlinear Maps and their Applications*, 81-86, Urbino.

[P2.19] Caneco, A., Grácio, C., Fernandes, S., J. Rocha, J. L., and Ramos, C. (2009). Conditions for the Formation of Clusters Depending on the Conductance and the Coefficient of Clustering, *Proceedings of the INDS 2009, 2th International Workshop on Nonlinear Dynamics and Synchronization*, 12-16, Klagenfurt, Austria.

[P2.20] Gomes, M.I., Pestana, D.D., Sequeira, F., Mendonca, S. and Velosa, S. (2009). Uniformity of Offsprings from Uniform and Non-Uniform Parents. In Luzar-Stiffler, V., Jarec, I. and Bekic, Z. (eds.), *Proceedings of the ITI 2009, 31st International Conference on Information Technology Interfaces*, p. 243-248.

[P2.21] Pestana, D.D., Aleixo S.M. and Rocha, J. L. (2009). Hausdorff Dimension of the Random Middle Third Cantor Set. In Luzar-Stiffler, V., Jarec, I. and Bekic, Z. (eds.), *Proceedings of the ITI 2009, 31st International Conference on Information Technology Interfaces*, p. 279-284. (IEEE

10.1109/ITI.2009.5196094)

[P2.22] Gomes, M. I., Mendonça, S., and Pestana, D. (2009). Adaptive Reduced-Bias Tail Index and Value-at-Risk Estimation. The XIIIth International Conference "Applied Stochastic Models and Data Analysis" (ASMDA-2009) ISBN: 978-9955-28-463-5, Technika, Vilnius, p. 41-44.

#### Other publications National

##### PROCEEDINGS

[ON2.1] Caeiro, F., Gomes, M. I., e Pestana, D. (2009). "Alguns resultados adicionais sobre a variância de um estimador de viés reduzido do índice de cauda", In Oliveira, I., Correia, E., Ferreira, F., Dias, S., e Braumann, C. (eds.) Estatística: Arte de Explicar o Acaso, SPE, 167-177.

[ON2.2] Gomes, M. I., e Pestana, D. (2009). "Caudas leves em desporto: estimação de parâmetros úteis", In Oliveira, I., Correia, E., Ferreira, F., Dias, S., e Braumann, C. (eds.) Estatística: Arte de Explicar o Acaso, SPE, 307-318.

[ON2.3] Mendonça, S., Malva, M., e Pestana, D. (2009). "Expansões de Edgeworth e aproximações pré-assintóticas" In Oliveira, I., Correia, E., Ferreira, F., Dias, S., e Braumann, C. (eds.) Estatística: Arte de Explicar o Acaso, SPE, 429-436.

[ON2.4] Pestana, D. (2009). Representações Gráficas na Formação da Intuição, na Análise dos Dados e na Comunicação das Ideias. Centro de Filosofia da Ciência da Universidade de Lisboa. (in print).

[ON2.5] Santos, R. e Pestana, D. (2009). Distância entre os extremos de uma linha poligonal "ao acaso", In Oliveira, I., Correia, E., Ferreira, F., Dias, S., e Braumann, C. (eds.) Estatística: Arte de Explicar o Acaso, SPE, 591-602.

[ON2.6] Soares de Almeida, L, Pestana, D, e Marques Gomes, M. (2009). "Evolução da actividade do Laboratório de Histopatologia Cutânea do Serviço de Dermatologia do Hospital de Santa Maria — uma comparação de três lustros", em G. Cunha e J. Varanda (eds.), Estatística e Qualidade na Saúde: Novos Rumos, Novos Desafios, Instituto Nacional de Estatística, Lisboa, 74-90.

[ON2.7] Sousa, R. (2009). O problema de Dorfman revisitado — análises quantitativas, em G. Cunha e J. Varanda (eds.), Estatística e Qualidade na Saúde: Novos Rumos, Novos Desafios, Instituto Nacional de Estatística, Lisboa, 140-144.

[ON2.8] Trigueirão, E., Gonçalves, H., Cabral, M.S. (2009). Análise Estatística de dados longitudinais discretos numa experiência sobre a influência de lípidos no adenocarcinoma mamário. Estatística e Qualidade na Saúde: Novos Rumos, Novos Desafios, Instituto Nacional de Estatística, Lisboa, 22-29, 2009.

[ON2.9] Trigueirão, E., Gonçalves, H., Cabral, M.S. O modelo linear generalizado misto na análise de dados longitudinais discretos. Notas e Comunicações do CEAUL, 19/2009.

##### ABSTRACTS AND EXTENDED ABSTRACTS

[ON2.10] Aleixo, S. M., Diamantino, M. F., e Pestana, D. (2009). A Família GLE. XVI Congresso Anual da SPE — Programa e Resumos, 251-252, SPE

[ON2.11] Bispo, R. Bernardino, J., Marques, T, Pestana, D. (2009). Discriminação entre modelos paramétricos para tempos de permanência de cadáveres de aves em parques eólicos. XVI Congresso Anual da SPE — Programa e Resumos, 35-37, SPE

[ON2.12] Brilhante, M. F., Malva, M., Mendonça, S., Pestana, D., Sequeira, F., e Velosa, S. (2009). Uniformidade. XVI Congresso Anual da SPE — Programa e Resumos, 345-346, SPE

[ON2.13] Brilhante, M. F., Pestana, D., Sequeira, F. (2009). Combinação de valores de prova e valores de prova aleatórios. XVI Congresso Anual da SPE — Programa e Resumos, 256-258, SPE

[ON2.14] Brilhante, M. F., Pestana, D., Sequeira, F. (2009). Funções de variáveis aleatórias com suporte em (0,1) com aplicações em meta-análise. XVI Congresso Anual da SPE — Programa e Resumos, 186-187, SPE

[ON2.15] Cadima, J., Calheiros, F. L., e Preto, I. (2009). Análise em Componentes Principais: Analisando os Dados ou a Escolha de Variáveis. XVI Congresso Anual da SPE — Programa e Resumos, 45-46, SPE

[ON2.16] Carrasquinha, E., Gonçalves H., Cabral, M.S. (2009). O modelo linear generalizado misto na análise de dados longitudinais discretos. XVII Congresso Anual da Sociedade Portuguesa de Estatística. Programa e Resumos 49-50, SPE editions.

[ON2.17] Felgueiras, M. (2009). Misturas com parâmetro de escala Pareto. XVI Congresso Anual da SPE — Programa e Resumos, 132-133, SPE

[ON2.18] Gonçalves, H., Trigueirão, E., Cabral, M.S. (2009). O modelo marginalizado na análise de um estudo longitudinal sobre a influência de lípidos no desenvolvimento do cancro da mama. XVI Jornadas de Classificação e Análise de Dados (JOCLAD 2009) – Programa e Resumos, 179-181.

[ON2.19] Marques, T. A., and Thomas, L. (2009). Spatially explicit mark recapture methods for estimating minke whale "boing" density. XVI Congresso Anual da SPE — Programa e Resumos, 245-246, SPE

[ON2.20] Martins, J. P. (2009) "Dependência entre a média e variância amostrais". XVI Congresso Anual da SPE — Programa e Resumos, p. 183, SPE

[ON2.21] Pestana, D., Vasconcelos, R., Velosa, S. (2009). Viés na publicação e sínteses meta-analíticas. XVI Congresso Anual da SPE — Programa e Resumos, 361-362, SPE

[ON2.22] Sousa, R. (2009) Análises conjuntas quantitativas, XVI Congresso Anual da SPE — Programa e Resumos, p. 310-311, SPE

[S2.1] Aleixo, S., Rocha, J. L., and Pestana, D. (2009). Dynamical Analysis in Probabilistic Models: Random Beta(p,2) Cantor Sets, Notas e Comunicações do CEAUL.

[S2.2] Brilhante, M. F., Mendonça, S., Pestana, D., (2009). How normal is the "normal" law, and how central is the "central" limit theorem? Notas e Comunicações do CEAUL.

[S2.3] Brilhante, M. F., Malva, M., Mendonça, S., Pestana, D., Sequeira, and Velosa, S. (2009). Uniformity, Notas e Comunicações do CEAUL.

[S2.4] Brilhante, M. F., Mendonça, S., Pestana, D., and Sequeira (2009). Using Products and Powers of Products to Test Uniformity, Notas e Comunicações do CEAUL.

[S2.5] Brilhante, M. F., Pestana, D., and Sequeira, F. (2009). Combining p-values and Random p-values, Notas e Comunicações do CEAUL.

[S2.6] Cabral, M.S. (2009). Comentário ao Modelo de Valor Acrescentado. In M. Eugénia Ferrão (eds), Modelo de Valor Acrescentado: algumas evidências a partir do projecto 3EM, com discussão, Parte V, 81-83. ISBN 978-989-654-023-4.

[S2.7] Gomes, M. I., Mendonça S., and Pestana, D. (2009). The Bootstrap Methodology and Adaptive Reduced-bias Tail Index and Value-at-Risk Estimation. Notas e Comunicações do CEAUL, 7/2009.

[S2.8] Gomes, M. I., Mendonça S., and Pestana, D., and Henriques-Rodrigues, L. (2009). Athletic Events and Statistics of Extremes: Estimation of Useful Parameters. Notas e Comunicações do CEAUL, 6/2009.

[S2.9] Gomes, M. I., Pestana, D., and Henriques-Rodrigues, L. (2009). Athletic Events and Statistics of Extremes: Estimation of Useful Parameters. Notas e Comunicações do CEAUL, 6/2009.

### Master and Ph.D. thesis completed

Ph. D. Thesis:

[PhD2.1] João Paulo Martins: "Feira dos Momentos — Planeamento Experimental e Investigação de Localização e Escala em Populações não Gaussianas".

Local: Universidade de Lisboa

Supervisor: Dinis Pestana and Sandra Mendonça.

[PhD2.2] Miguel Felgueiras: "Modelação Estatística com Misturas e Pseudo-Misturas".

Local: Universidade Nova de Lisboa

Supervisors: Dinis Pestana.

[PhD2.3] Fernando Sequeira: "Meta Análise — Harmonização de Testes Usando os Valores de Prova".

Local: Universidade de Lisboa

Supervisors: Dinis Pestana and Sílvio Velosa.

M. Sc. Thesis:

[MSc2.1] João Curado Silva: "Cartões de Crédito Particulares".

Local: Universidade de Lisboa

Supervisors: Dinis Pestana and Ana Cristina Moita

[MSc2.2] Eunice Isabel Ganhão Carrasquinha Trigueirão: "Análise de dados longitudinais discretos: uma aplicação ao estudo da influência de lípidos no adenocarcinoma mamário".

Local: Universidade de Lisboa

Supervisor: M. Salomé Cabral

Cláudi [MSc2.3] Rossana Velosa Pereira, "Uma introdução às filas de espera".

Local: Universidade da Madeira

Supervisor: Sandra Mendonça

### Organization of conferences

5. Organization of conferences (2000 ca.)

D. Pestana was a member of the scientific and organizing committees of the International Workshop in Honour of Professor J. Tiago Mexia, Universidade Nova de Lisboa.

M. H. Gonçalves, M.H., was member of the scientific committee and of the local organizing committee XVI Jornadas de Classificação e Análise de Dados (JOCCLAD 2009, <http://fct.ualg.pt/joclad2009>)

Short course on Statistical Analysis with R. Took place in ISA and J Cadima lectured part of the course

### Internationalization

D. Pestana has been an invited speaker at the Klein group Workshop on Didactics of Mathematics as a Mathematical Discipline.

T. A. Marques maintained his strong bonds with S. Buckland's St. Andrews team working in distance sampling.

D. Pestana is a Portuguese member of CENE, a joint committee contributing to the multilingual ISI dictionary of Statistics (CENE provided Portuguese translation for more than 95% of the entries) and standardization of Portuguese-Brazilian glossary of Statistics. Invited Klein Workshop

J. Cadima, S. Mendonça, D. Pestana and R. Sousa acted as reviewers for JABES, Computational Statistics and Data Analysis, J. Applied Statistics, J. Korean Statistical Society, Mathematical Reviews, Revstat, and Acquaculture. J. Cadima is associate editor of JABES, D. Pestana is associate editor of Revstat. D. Pestana acted as guest co-editor of the 2009 and 2010 issues of *Discussiones Mathematicae — Probability and Statistics*.

D. Pestana is an invited speaker at the LinStat, Tomar ("Meta-Analytical Issues in Linear Models", with J. P. Martins).

### Future Research

#### Objectives

• Meta analysis and applications in Medicine; uniformity issues related to combined p-values and random p-values, data transformations and sample augmentation. Reserach team: F. Sequeira, F. Brilhante, S. Mendonça, D. Pestana, L. Soares de Almeida, R. Vasconcelos, M. Malva, J. P. Martins, M. L.

Rocha.

- Structure of population matrices and Principal Components Analysis. Research team: J. L. Cadima and co-workers from ISA.
- History of Science — Research in Probability and Statistics in Portugal during the first Republic. Research team: R. Santos, D. Pestana.
- Fractal issues in population models. Research team: J. L. Rocha, S. Aleixo, D. Pestana, R. Santos.
- Scale analysis using the independence of scale and location estimators. Research team: D. Pestana, M. Marques dos Santos, F. Brilhante, S. Mendonça, M. Malva.
- Analysis of binary longitudinal data, in view of medical applications. Research team: S. Cabral, H. Gonçalves.
- Further work on errors in distance sampling models, and modeling bird death rates in eolic plants. Research team: T. A. Marques, R. Bispo.
- Teaching of Statistics: Documentation for teachers, and collaboration with INE's Alea Project. Research team: E. Graça Martins, D. Pestana, S. Velosa
- Extensions of Dorfman's theory on analysis of combined blood samples: specificity and sensibility issues in quantitative tests; mixture models. Research team: D. Pestana, M. Felgueiras, R. Sousa, R. Vasconcelos.
- Simulation studies of the Rényi's rarefaction, and applications to sampling theory. Research team: S. Aleixo, F. Diamantino, D. Pestana.
- Scientific translation in the area of Probability and Statistics, and development of an e-learning platform for professional translators for TRADULÍNGUAS. Research team: C. Ventura, S. Velosa, D. Pestana.

**Funding, source, dates**

All the funding received during 2009 and reported in the section 1 of this group description report was transferred to 2010

<b>Group Description</b>	
<b>Title of Research Group:</b>	(RG-LVT-6-951) Statistical Modelling in Environmental and Life Sciences
<b>Principal Investigator:</b>	Maria Antonia Conceição Abrantes Amaral Turkman
<b>Main Scientific Domain:</b>	Matemática
<b>Group Host Institution:</b>	Faculdade de Ciências - Universidade de Lisboa
<b>Funding, source, dates</b>	
<b>Funding, source, dates</b>	
<p>ALL THE REFERRED AMOUNTS INCLUDE OVERHEADS TO PAY FFCUL</p> <p>1. FCT Base for Research Group Title: (RG-MATH-LVT-Lisboa-6-951) - Statistical Modelling in Environmental and Life Sciences</p> <p>Funding:</p> <p>corresponding to 2009:</p> <p>date: 3 April 2009/ 38.169,25 Eur</p> <p>date: 7 October 2009/ 16.102,74 Eur</p> <p>2. Projects: The following Project has Principal Investigator from the Group:</p> <p>(RG-MATH-LVT-Lisboa-6-951) - Statistical Modelling in Environmental and Life Sciences</p> <p>FCT/ PTDC/MAT/64353/2006</p> <p>MEGA – STATISTICAL METHODS IN GENETIC AND ENVIRONMENT</p> <p>Funding Entity: PTDC - FCT</p> <p>Total Award Ammount: 60 000 Eur</p> <p>Prime Contractor: Universidade de Lisboa - Fundação da F.C.U.L.</p> <p>Period Covered: from September of 2007 to September of 2010 (extension asked till 31 December 2010)</p> <p>Coordinator: Prof. K.F.Turkman</p>	
<b>Objectives &amp; Achievements</b>	
<b>Objectives</b>	
<p>The main objective of this research group is to make contributions on specific probability and statistical issues in modelling environmental and biological data, as well as, to create a group with expertise in analyzing data coming from these fields. To achieve this goal we seek collaboration with researchers from the environmental and biological sciences, not only with the objective to gather data to apply the methodology developed, but as well to help them in solving their own problems. Another objective is to continue the Ph.D. programme on the main research themes, namely Hierarchical Bayesian Modelling, Spatial-temporal Modelling, Statistics in Genetics and Life Sciences. Together with applied research we have as an objective to continue developing fundamental research in the areas of interest, namely on: spatial extremes ; hierarchical bayesian spatiotemporal analysis; stochastic modelling of dynamic systems; bayesian latent class models; non-linear time series ; missing data mechanisms; bayesian classification methods; predictive quality of diagnostic tests; survival analysis and longitudinal models; bayesian analysis of allelic penetrance models for complex binary traits; ecological studies; modelling infectious diseases; syndromic surveillance; microarray data analysis.</p> <p>For the year 2009 we set as objectives:</p> <ol style="list-style-type: none"> <li>1. To prepare the 2nd edition of the books Bayesian Statistics (D. Paulino and MA Amaral Turkman); Survival analysis: theory and applications to health (V. Andreozzi); Finish the book on "Survival Analysis" (C. Rocha and AL Papoila); continue working on the books Non-linear time series (Feridun Turkman); Statistical modelling: applications to health" (V. Andreozzi, M. Antunes and collaborators from FIOCRUZ)</li> <li>2. To organize regularly seminars and advanced courses on themes related to the areas of interest of the group, which may also interest the community in general.</li> <li>3. To revise and publish several submitted papers on: Estimation on Generalized Pareto distribution; Missing Categorical Data; Supervised classification; Spatio-temporal extremes; Bayesian Latent class models; non-homogeneous Poisson models; Mercury, cadmium and lead bioaccumulation in black scabbard fish; predictors of mortality in HIV patients; influenza epidemics; Identification of plant genes; estimation of T-cell repertoire diversity; Bayesian analysis of allelic penetrance models; Identification of plant genes involved on the initial contact between ectomycorrhizal symbionts.</li> <li>4. To continue research and submit articles on, among others, hierarchical space-time models for high dimensioned fire data; excess of mortality attributable to influenza epidemics, Bayesian approach to peak area measurements in Gamma spectrometry; Pre-processing Optimization of RNA Immunoprecipitation Microarray Data; Recurrent wheezing after Respiratory Syncytial Virus bronchiolitis: a longitudinal study; Bayesian Approach to Gene Expression Profile Classification.</li> <li>5. To continue the collaboration with researchers from the environmental and biological sciences.</li> <li>6. To participate in International and national conferences.</li> <li>7. To publish the R package : Flubase Baseline of mortality free of influenza epidemics.</li> <li>8. To work on a package on Bayesian Approach to Gene Expression Profile Classification.</li> </ol>	



## Main Achievements

Reporting to the above objectives set for 2009 the main achievements were:

1. The book on "Análise de Sobrevivência" [B1] was published. V Andreozzi and M Antunes were in FIOCRUZ to work on the book "Modelação Estatística: Aplicações à saúde"; the first 5 chapters of the book on "Non-linear time series were completed.
2. 1 National Conference, 1 International workshop, 6 seminars and 4 advanced courses were organized by the members of the group. Also it was initiated a cycle of closed workshops, with only members of the group, to present and discuss matters of scientific nature.
3. Revision of the work on generalized Pareto distribution [P11-P12] was finalized. In these two papers, several methods of estimation are discussed and compared. The implication of missing data mechanisms on the analysis of categorical data is discussed on [P16]. Spatial extremes were dealt in [P10, P13, P20] with application to large fires in Portugal. [P8] corresponds to an innovative application of Bayesian latent class models to measure deprivation among Portuguese households. Non-homogeneous Poisson models considering four different types of rate function were applied to ozone measurements in Mexico City [P1]. Total mercury (THg), cadmium (Cd) and lead (Pb) concentrations were investigated in muscle, liver and gonad tissue of black scabbardfish [P4]. Using the Cox proportional model with frailty [P5] several factors were identified as responsible for increasing the hazard of HIV related hospital mortality in Portugal. In [P18] estimation of TCR diversity is considered. Definitions of recessiveness are discussed under the allelic penetrance models in [P19] and Bayesian methods are applied to analyze two data sets. Collaboration with researchers from plant biology on the statistical analysis of microarray data resulted on the publication [P17].
4. The work on Bayesian spatio-temporal methods to model jointly the probability of ignition and fire sizes in Australia, based on annual observations of several explanatory variables [C1] was submitted for publication. Results from determinations of peak areas for a radioactive source containing several radionuclides were analyzed using Bayesian methods [A5]. The work, already accepted for publication, shaded some doubts about the adequacy of the Poisson model assumption for the counting process at some gamma-ray energies and the need to proceed to a further study where more comprehensive models can be contemplated. Estimation of diseases using imperfect data was done through a state space models approach. As a result the article [A3] was accepted for publication. Time series methods were applied in order to evaluate the excess of mortality attributable to influenza epidemics and the work [A5] was accepted for publication. It was proposed a non-parametric methodology based on ranks for enriched gene selection applied to RIP-chip (RNA immunoprecipitation on chip) data analysis, namely Affymetrix HG U133 plus 2.0 Gene Chip [N13]. A new Bayesian classification methodology based entirely on marker gene pairs expression monitoring by DNA microarrays was proposed. The results were accepted for publication [A6].
5. Collaborative work with researchers from (i) health sciences resulted on the international publications P[2], [P3], [P7], [P9], [P14], [P15], [A1], national publication [N2] and several communications at international conferences (9) and national conferences (10), (ii) veterinary sciences resulted in 2 national publications, (iii) forest fires resulted in several already reported, (iii) other sciences resulted in international publications [P4], [P6], [P17], [A5].
6. Members of the group were invited speakers at international conferences, namely [C4], [C5], [C6], [C7], [C8], [C22], [C24]. Other physical participation in international conferences by the members of the group corresponds to [C1], [C10], [C11], [C18], [C19], [C21].
7. A software package for R named flubase, was published. This function estimates the baseline of mortality free of influenza epidemics, and the respective excess deaths, for more than one time series (age groups, gender, regions, etc) using one of the four methods available.
8. The package on Bayesian classification is being developed and part of the functions to be made available were published together with the article [A6] as a requirement from the publisher.

## Group Productivity

### Publications in peer review Journals

- [P1] Achcar J, Rodrigues E, Paulino CD and Soares P, 2009. Non-homogeneous Poisson models with a change-point: an application to ozone peaks in Mexico City. *Environmental and Ecological Statistics* (DOI: 10.1007/s10651-009-0114-3).
- [P2] Borrego LM, Arroz MJ, Videira P, Martins C, Guimarães H, Nunes G, Papoila AL, Trindade H, 2009. Regulatory cells, cytokine pattern and clinical risk factors for asthma in infants and young children with recurrent wheeze. *Clinical Experimental Allergy*, 39: 1160-1169
- [P3] Chalfun G, Mello RR, Dutra MVP, Andreozzi V, Silva KS, 2009. Risk Factor for Respiratory Morbidity at 12 to 36 months in Very Low Birth Weight Premature Infants Previously admitted to a Public Neonatal Intensive Care Unit. *Reports in Public Health* 25(6):1399-1408
- [P4] Costa V, Lourenço HM, Figueiredo I, Carvalho L, Lopes H, Farias I, Pires L, Afonso C, Vieira AR, Nunes ML and Gordo LS, 2009 Mercury, cadmium and lead in black scabbardfish (*Aphanopus carbo* Lowe, 1839) from mainland Portugal and the Azores and Madeira archipelagos. *Scientia Marina* 73S2, 77-88, doi: 10.3989/scimar.2009.73s2077
- [P5] Dias S, ANDREOZZI V, Martins MO, Torgal J, 2009. Predictors of mortality in HIV-associated hospitalizations in Portugal: a hierarchical survival model. *BMC Health Services Research*, 23:9-125
- [P6] Gaspar H, Moiteiro C, Amaral Turkman A, Coutinho J and Carnide V, 2009. Influence of soil fertility on dye flavonoids production in weld (*Reseda luteola* L.) accessions from Portugal. *L. Sep. Sci* 32, 1-7
- [P7] Gonçalves-Pereira M, Carmo I, Alves da Silva J, Papoila AL, Mateos R, Zarit SH (2009). Caregiving experiences and knowledge about dementia in Portuguese clinical outpatient settings. *International Psychogeriatrics* (in Press).
- [P8] Machado C, Paulino CD and Nunes F, 2009. Deprivation Analysis based on Bayesian Latent Class Models. *J. Applied Statistics*, 36/8, 871-891.
- [P9] Martins P, J Rosado-Pinto, M Teixeira, N Neuparth, O Silva, H Tavares, JL Spencer, D Mascarenhas, AL Papoila, N Khaltaev, I Annesi-Maesano, 2009. Underreport and underdiagnosis of chronic respiratory diseases in an african country. *Allergy* 64 1061-7.
- [P10] Mendes J, P de Zea Bermudez, J Pereira and KF Turkman, 2009. Spatial extremes of wildfire sizes: Bayesian hierarchical models for extremes. *Environmental and Ecological Statistics*, DOI: 10.1007/s10651-008-0099-3.
- [P11] P de Zea Bermudez and S Kotz, 2009. Parameter estimation of the Generalized Pareto Distribution - Part I. *Journal of Statistical Planning and Inference* (in press).
- [P12] P de Zea Bermudez and S Kotz, 2009. Parameter estimation of the Generalized Pareto Distribution - Part II. *Journal of Statistical Planning and Inference* (in press).
- [P13] P de Zea Bermudez, J Mendes, J Pereira, K F Turkman & MJ Vasconcelos, 2009. Spatial and temporal extremes of wildfire sizes in Portugal (1984-2004). *International Journal of Wildland Fire*, 18, 985-993.
- [P14] Pinto S, Amaral Turkman MA, Pinto A, M Swash & M Carvalho, 2009. Predicting respiratory insufficiency in amyotrophic lateral sclerosis: The role of phrenic nerve studies. *Clinical Neurophysiology* 120 941-946
- [P15] Pinto AE, André S, Silva GL, Vieira S, Santos AC, Dias S & Soares J, 2009. BCL-6 oncoprotein in breast cancer: loss of expression in disease progression. *Pathobiology*, 76, 235-42. (DOI: 10.1159/000228899).

- [P16] Poletto F, Singer J & Paulino CD ,2009. Missing data mechanisms and their implications on the analysis of categorical data. *Statistics and Computing* (DOI: 10.1007/s11222-009-9143-x).
- [P17] Sebastiana M, Figueiredo A, Acioli-Santos B, Sousa L, Pessoa F, Baldé A. & Pais MS, 2009. Identification of plant genes involved on the initial contact between ectomycorrhizal symbionts (*Castanea sativa* - European chestnut and *Pisolithus tinctorius*). *EJSB*, 45, 275–282
- [P18] Sepúlveda N, Paulino CD & Carneiro J ,2009. Estimation of T-cell repertoire diversity and clonal size distribution by Poisson abundance models. *Journal of Immunological Methods* (DOI:10.1016/j.jim.2009.11.009).
- [P19] Sepúlveda N, Paulino CD & Penha- Gonçalves C ,2009. Bayesian Analysis of Allelic Penetrance Models for Complex Binary Traits. *Computational Statistics and Data Analysis* 53, 1271-1283.
- [P20] Turkman KF, Amaral Turkman MA & JM Pereira, 2009 Asymptotic models and inference for extremes of spatio-temporal data. *Extremes* 13. <http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s10687-009-0092-8>

#### Other publications International

##### Accepted

- [A1] Diogo L, Papoila AL (2009). Is it possible to characterize objectively Sensitive Skin?. *Skin Res Technol*.
- [A2] Malta, F.M., Medeiros-Filho, J.E.M., Azevedo, R.S., Gonçalves, L., da Silva, L.C., Carilho, F.J., Pinho, J.R.R. (2009) Sequencing of E2 and NS5A regions of HCV genotype 3a in Brazilian patients with chronic hepatitis. (in press, *Mem Inst Oswaldo Cruz*)
- [A3] Natário I, Carvalho ML, Gomes MG (2009). Addressing the problem of lack of representativeness on syndromic schemes. Accepted in *Discussiones Mathematicae – Probability and Statistics*
- [A4] Nunes B, Natário I, Carvalho ML (2009). Time series methods for obtaining excess mortality attributable to influenza epidemics. Accepted in *Statistical Methods in Medical Research*
- [A5] Silva, L., Amaral Turkman, MA. and Paulino, C.D. (2009) Bayesian Statistical Evaluation of Peak Area Measurements in GAMMA SPECTOMETRY. Accepted for publication, December 2009 in *Applied Radiation and Isotopes*
- [A6] Ramos, S, Turkman, MA and Antunes, M. Bayesian classification for bivariate gene expression. Accepted for publication in *Computational Statistics and Data Analysis*

##### PROCEEDINGS INTERNACIONAL

- [C1] Amaral Turkman, M.A. and Turkman, KF. (2009) Hierarchical space-time models for fire ignition and percentage of land burned by wildfires) *Proceedings of the 24th International Workshop on Statistical Modelling*. James G. Booth, Editor, 49-52.
- [C2] Gonçalves Pereira, M; Papoila, A., Xavier, M. (2009). The FAPS study on severe mental illness: views on the implementation of psychoeducational family work in Portuguese settings. In *Proceedings of the 17th European Congress of Psychiatry, Lisbon, 2009* (Abstract: Symposium 44-03 on CD-Rom edited by European Psychiatry, Elsevier Masson SAS).
- [C3] Gonçalves Pereira, M; Alves da Silva, J; Carmo, I; Papoila, A. et al. (2009). Preliminary data from FAMIDEM survey: can we assume who is at risk regarding informal caregiving in dementia ? In *Proceedings of the 17th European Congress of Psychiatry, Lisbon, 2009* (Abstract: Poster 03-101 on CD-Rom edited by European Psychiatry, Elsevier Masson SAS).
- [C4] Sousa, L. "Proteomics: predicting proteins structure". In *Book of Short Papers: Statistical Methods for the Analysis of Large Data-Sets (Pescara - Italy, 23-25 Sept 2009)*, p. 207-210. Invited talk
- [C5] Silva, G.L. (2009). Invited paper: Modeling and analysis of forest fire data in Portugal. *SIS 2009 Statistical Conference - Statistical Methods for the analysis of large data-sets, Chieti-Pescara - Italy, 23-25th September: 2009*. 211-214

##### Abstracts International Conferences

The following were invited speakers at the 4th Workshop on Statistics, Mathematics and Computation, 9-10 November, Lisbon

- [C6] Antunes, M (2009) EMB: a combination of EM-algorithm and Bayesian classification model for gene identification, *Book of Abstracts*, 23
- [C7] Papoila, AL (2009) Prediction in Medicine: Prognosis and Diagnosis, *Book of Abstracts* 15
- [C8] Paulino, CD (2009) A fully Bayesian parametric approach for cytogenetic dosimetry, *Book of Abstracts*, 17

##### Others

- [C9] Amaral Turkman, MA and Turkman, KF (2009) Hierarchical modeling strategies for extremal data, *SETA 2009*, Lisbon (<http://seta.ceaul.fc.ul.pt>); invited talk
- [C10] Amaral Turkman MA & Turkman KF, (2009) Hierarchical space-time models for fire ignition and percentage of land burned by wildfires. *24th International workshop in Statistical Modelling*, Cornell University, 18-23 July 2009
- [C11] Branco, M. (2009). Contributed paper: A new parametric Bayesian approach for cytogenetic dosimetry (joint with C.D. Paulino & G. Silva). *XI Escola de Modelos de Regressão*, Recife, Brasil, Março 2009.
- [C12] Borrego L, Arroz M, Videira P, Martins C, Guimarães H, Marques G, Papoila A, Trindade H, editors. Regulatory cells, cytokine pattern and clinical risk factors for asthma in infants and young children with recurrent wheeze. *28th Congress of the European-Academy-of-Allergyand-Clinical-Immunology*; 2009 Jun 06-10; Warsaw, POLAND: Wiley-Blackwell Publishing, Inc
- [C13] Borrego LM, Arroz MJ, Videira P, Martins C, Guimarães H, Nunes G, Papoila AL, Trindade H (2009). Regulatory cells, cytokine pattern and clinical risk factors for asthma in infants and young children with recurrent wheeze. *XXVII Congress of the E.A.A.C.I.*, Warsaw, June 6-10, 2009
- [C14] Coelho, S.,A.L.Papoila, B. Rodrigues, H. Martins, A. Borges, S. Lopes e K. Soto (2009). Plasmatic and Urinary Neutrophil Gelatinase-Associated Lipocalin (NGAL) and Serum Cystatin C (CysC) as biomarkers of Acute Kidney Injury. Preliminary results *XXXIX Congreso Nacional de la Sociedad Española de Nefrología*, Pamplona 3-6 Outubro de 2009
- [C15] Gonçalves Pereira, M; Papoila, A., Xavier, M. (2009). The FAPS study on severe mental illness: views on the implementation of psychoeducational family work in Portuguese settings. In *Proceedings of the 17th European Congress of Psychiatry, Lisbon, 2009* (Abstract: Symposium 44-03 on CD-Rom edited by European Psychiatry, Elsevier Masson SAS).
- [C16] Gonçalves Pereira, M; Papoila, A., Xavier, M. (2009). The FAPS study on severe mental illness: views on the implementation of psychoeducational family work in Portuguese settings. *17th European Congress of Psychiatry, Lisbon, 2009*.
- [C17] Karina Soto., H. Martins, AL Papoila, S. Lopes, Cuna, L. and B. Rodrigues (2009). Performances of cystatin C and creatinine as markers of acute kidney injury and prerenal azotemia, in emergency room: A preliminary study. *XV Congreso de la SLANH, VI Congreso Iberoamericano de Nefrología*, México, 2009

- [C18] Natário I. Estimating disease outbreaks based on self-reporting. RSS2009, International Conference of the Royal Statistical Society 2009. Edinburgh, United Kingdom. 2009. Pg.243 of the Abstracts booklet
- [C19] Nunes B, Natário I e Carvalho L. "Modelling mortality baseline free of influenza epidemics" Eastern Mediterranean Region of the International Biometric Society (EMR-IBS) will be held at the Military Museum, İstanbul, Turkey, on May 10 – 14 2009
- [C21] Silva, G.L. (2009). Invited paper: Bayesian overdispersed models with smoothing splines for spatial age-specific data. II Workshop of the ERCIM Working Group on Computing and Statistics, Limassol - Cyprus, 29-31st October: 2009.
- [C22] Sepúlveda, N. (2009). Biodiversity estimation: quantifying the immune repertoire in the body's cellular ecosystem. 16th European Young Statistician Meeting, Bucharest, Romania, August 2009. Invited talk
- [C23] Soto, K., S. Coelho, A. Papoila, S. Lopes, B. Rodrigues, P. Frade and P. Devaradjan (2009). Performances of Cystatin C and Creatinine as Markers of Acute Kidney Injury in the Emergency Room Setting. ASN Renal Week, October 27-November 1 2009, San Diego, California, USA
- [C24] Turkman KF & Amaral Turkman MA. Modeling strategies for spatial extremes; A case study in extreme fire sizes. Spatial extremes and applications. Lausanne, July 13-17, 2009. Invited talk

#### Other publications National

##### Books

- [B1] Rocha, C. and Papoila, A.L.(2009) Análise de Sobrevivência. Edições SPE, ISBN: 978-972-8890-22-3

##### Periodicals

- [N1] Andreozzi, V e Amaral Turkman, MA (2009) Inferência bayesiana no R através do WinBUGS. Boletim SPE, 88-94, Outono
- [N2] Napoleão P, Selas M, Toste A, Amaral Turkman MA, Andreozzi V, Viegas-Crespo AM, Pinheiro T, Ferreira RC (2009). Serial Changes of Oxidized Low-density Lipoprotein Associated with Culprit Vessel in ST-Elevation Myocardial Infarction – A Promising Marker? Revista Portuguesa de Cardiologia 28 303-308
- [N3] Rosado-Ganhão, S. (2009), Fractals in Architecture in Revista Artitextos. Nº8. Centro Editorial da Faculdade de Arquitectura, 261-271
- [N4] Henriques, C., Rosado-Ganhão, S., (2009) Estatística na Análise do Território. in Boletim da SPE, 82-85, Spring
- [N5] Sepúlveda, N. (2009). Surfando nos mares da Biologia: peripécias de um jovem estatístico aspirante a biólogo. Boletim SPE, 75-81, Spring

##### Proceedings

- [N6] Pedro, S., Pires, A. and Sousa, L. "Aplicação de redes neuronais e árvores de classificação na análise de dados de microarrays". Book of Abstracts XVI JOCLAD, 150-153.
- [N7] Silva-Fortes, C., Turkman, MA and Sousa, L. (2009). Avaliação da capacidade preditiva de testes de diagnóstico de cancro pancreático. In Estatística e Qualidade na Saúde: novos rumos novos desafios (Cunha, G. and Varanda, J. Eds. ) 161-169.
- The following are In Oliveira, I., M.M. et al. (eds.), Estatística: Arte de Explicar o Acaso, Edições SPE, 2009:
- [N8] Castel-Branco, M., Antunes, M., Cara D'Anjo, A. and Bernardo, F. . Factores de risco associados à prevalência de Salmonella em explorações de frangos destinados à produção de carne em Portugal. 187-194
- [N9] Correia, E. e Amaral Turkman, MA (2009) Inferência Bayesiana na Razão de Médias com Precisoões Desconhecidas., 195-204
- [N10] Dias, I.S., Silva, G.L. and Louro, C. (2009). Análise bayesiana do efeito da poluição atmosférica na admissão hospitalar de crianças com problemas respiratórios no Barreiro. 217-224.
- [N11] Paulino, C. D. (2009). Dados categorizados com omissão num quadro Produto de Multinomiais: uma panorâmica de métodos de análise. Invited paper 81-106, SPE.
- [N12] Sepúlveda, N. (2009). Abordagem por Penetrâncias Alélicas: uma nova modelação estatística para a análise de fenótipos binários complexos. Prémio SPE in Estatística: Arte de explicar o acaso (Actas do XVI Congresso Anual da Sociedade Portuguesa de Estatística), 109-129, SPE.

##### Others

- [N13] Barreto-Hernandez, E., Gama-Carvalho, M. and Sousa, L. (2009). Pre-processing Optimization of RNA Immunoprecipitation Microarray Data. CEAUL Notas e Comunicações 12/2009
- [N14] Pedro, S., Pires, A. and Sousa, L. (2009). Metodologias de Classificação Supervisionada para Análise de Dados de Microarrays. CEAUL Notas e Comunicações 26/2009.

##### Abstracts

- Dumba, B., Belo, S., Gonçalves, L., Grácio, M.A.A. Helminthoses, género e actividades recreativas da população infantil em Angola (Huambo e Bié). XI Congresso Ibérico de Parasitologia, Lisboa, 15-18 Setembro 2009. Resumo: Acta Parasitológica Portuguesa, 16 (1/2): 250-251.
- In Programa e Resumos, SPE
- Abreu, A. and Rocha, C. (2009) Um modelo de cura paramétrico com covariáveis, 33-34
- Silva-Fortes, C., Amaral Turkman, MA, Lança, L. Silva, R. and Marques, G. (2009) Uma aplicação das curvas ROC MRMCM em Radiologia, 72
- Oliveira, R., Gonçalves, L, Pacheco, A. And Valadas, R. (2009) Era uma vez um médico, um engenheiro e um estatístico, 83-84
- Nunes, B. Natário, N. and Carvalho, L. (2009) Modelação de uma linha de base para a mortalidade sem o efeito de epidemias de gripe, 125-126
- Teles, J. and Amaral Turkman, MA (2009) Some remarks about Gibbs variable selection performance, 141-142
- Correia, C. Sepúlveda, N. and Paulino, C.D. (2009) Mapeamento bayesiano de BTL com a comparação de métodos de cálculo de factores de Bayes, 173-174
- Ribeiro, C. Amaral Turkman, MA and Cardoso, J. (2009) Modelação Estatística em segurança rodoviária, 175-176
- Sepúlveda, N. Sousa, V. Paulino, CD and Carneiro, J. (2009) Quantificação e comparação de repertórios de linfócitos T através do modelo Poisson-lognormal, 177-178
- Vicente, L. and Turkman, KF (2009) Modelos bayesianos hierárquicos no planeamento de recursos humanos, 179-180
- Barreto-Hernandez, E., Gama-Carvalho, M. and Sousa, L. (2009) Statistical Methodologies for RIP-Chip data analysis, 214-215
- Ramos, S. Amaral Turkman, MA and Antunes, M. (2009) Construção semi-paramétrica da região de screening óptima em classificação supervisionada, 216
- Antunes, M., Castel-Branco, M. Inácio, P and Cardo, M. (2009) Colheita de dados e especificações técnicas do estudo de base para a monitorização de

Salmonella e Campylobacter em carne de aves – do planeamento à implementação, 242

Geraldes, C., Papoila, AL, Rocha, C. and Xufre, P. (2009) Modelos lineares generalizados, modelos aditivos generalizados e redes neuronais: estudo comparativo na área da Medicina, 273-275

São João, R., Papoila, AL, Ligeiro, D. and Trindade, H. (2009) Frequências alélicas e haplotípicas do HLA no registo Português de doadores de medula óssea, 281-283

Oliveira, H., Natário, I and Oliveira, MM (2009) Um estudo epidemiológico do cancro da mama na região do Alentejo, 287-288

Polidoro, M.J., Magalhães, F. and Amaral Turkman, MA (2009) O problema da adequação de modelos numa perspectiva bayesiana, 290

Natário, I., Oloveira, MM, Mateus, S and Pires, C. (2009) Um estudo epidemiológico de doenças vasculares cerebrais na região do Alentejo, 293-294

Rosado-Ganhão (2009) Modelação e análise estatística do uso do solo na cidade de Maputo, 299

Silva, A.R., Papoila, AL, Sancho, R. and Trindade, H. (2009) Transplante da córnea: factores de prognóstico para a evolução do enxerto, 304-305

Subtil, A., Oliveira, R. and Gonçalves, L (2009) A validação da hipótese de independência condicional no modelo de classes latentes: um estudo de simulação, 315-316

Alves, M., Rocha, C. and Paixão, MT (2009) Estudo do tempo de evolução da infecção pelo VIH-1 até ao estadio de SIDA em Portugal, 320-321

Zea Bermudez, P. and Mendes, Z. (2009) Diferenças regionais nas medições efectuadas durante uma campanha Nacional de avaliação do risco cardiovascular. Programa e Resumos, SPE, p.322

Correia, E. and Amaral Turkman, MA (2009) Inferência Bayesiana na razão de proporções, 323

Castel-Branco, M., Antunes, M. Inácio, P and Cardo, M. (2009) Prevalência de Salmonella e Campylobacter em carcaças de frango para consumo humano – a realidade portuguesa, 324

#### Master and Ph.D. thesis completed

Nuno Sepúlveda - PhD. in Biomedical Sciences, Instituto Ciências Biomédicas Abel Salazar, University of Oporto (Thesis: How is the T-cell repertoire shaped?), October 2009.

César Correia - MSc. in Statistics, Instituto Superior Técnico, UTL (Thesis: Bayesian Mapping for Complex Binary Traits), November 2009 -supervised by C.D. Paulino.

Susana Luísa Augusto Esteves, MSc. in Biostatistics 2009 'Utilização de Metodologia Bayesiana e Métodos de Comparação Indirecta em Meta-Análise: Uma Aplicação ao Tratamento de Mieloma Múltiplo', Faculdade de Ciências da Universidade de Lisboa, Supervised by Maria Antónia Amaral Turkman

Liliana Sofia Mendonça Cardoso, MSc. in Biostatistics 2009 'Classificação de Genes em Hibridação Genómica Comparativa de Estirpes de Streptococcus pneumoniae' Faculdade de Ciências da Universidade de Lisboa, Supervised by: Marília Antunes

Marta Susana Ferreira Leitão Alves, MSc. in Biostatistics 2009 'Análise de Sobrevida Aplicada ao Estudo da Infecção VIH/Sida em Portugal' Faculdade de Ciências da Universidade de Lisboa, Supervised by: Cristina Rocha

Sérgio Santos, MSc in Bioinformatics 2009; "Desenvolvimento de Sistema on-line para a Predição da Topologia de Proteínas Transmembranares: Algoritmo BETH". Faculdade de Ciências da Universidade de Lisboa, January 2009. Supervised by Lisete Sousa.

#### Patents/protopypes

Baltazar Nunes, Isabel Natário and Lucilia Carvalho developed a software package for R package

flubase, <http://cran.r-project.org/web/packages/flubase/>.

flubase: Baseline of mortality free of influenza epidemics

This function estimates the baseline of mortality free of influenza epidemics, and the respective excess deaths, for more than one time series (age groups, gender, regions, etc) using one of the four methods available.

The description of the package is available at <http://cran.r-project.org/web/packages/flubase/flubase.pdf>

#### Organization of conferences

Workshops and Conferences

KF Turkman and MA Turkman organized, jointly with L de Haan SETA 2009 – Spatial Extremes, Theory and Applications, Lisbon, 6-8/04/09

Isabel Natário was a member of the organizing committee of the XVII Congresso Anual da Sociedade Portuguesa de Estatística, 30 September - 3 October, Sesimbra. [www.fct.unl.pt/spe2009](http://www.fct.unl.pt/spe2009). 273 participants.

MA Amaral Turkman was a member of the Scientific committee of the 4th Workshop on Statistics and Computation, Lisbon, 9-10 Nov 2009

M Antunes organized the 1st General Scientific Meeting of the this research group in Vimeiro 14-15 Feb 2009.

Seminars organized

Prof. Toby Prevost - Medical Statistics Kings College London - Statistics in Health and Lifestyle Research, CEAUL, 23/10/09

Seminar Bayesian Statistics in CEAUL - Modelos sobreparametrizados e suas Implicações Inferenciais: estudo em Análise de Dados Categorizados em Omissão

Dr. Frederico Poletto – Visiting student at DM-IST e CEAUL, 21/10/09

Multimodality on the Geometric Combination of Bayesian Forecasting Models

Prof. Álvaro Faria - Open University Uk. CEAUL 7/10/09

Estudo Longitudinal da Saúde do Adulto: Organização do primeiro estudo de coorte sobre determinantes de doenças cardiovasculares e diabetes no Brasil

Profª Dóra Chor - Fundação Oswaldo Cruz (ELSA)-Rio de Janeiro. CEAUL, 12/05/09

#### NetWork Approaches to Disease Surveillance

Prof. Ben Reis - Harvard Medical School and Boston Childrens Hospital. CEAUL 16/03/09

#### Statistical Facial Identification

Prof. Nick Fieller - University of Sheffield. CEAUL 5/03/09

#### COURSES

##### Introdução à Estatística Bayesiana via WinBUGS

- CD Paulino, G Silva e P Soares, IST-UTL/CEAUL
- 9h-12h 30 e 14h-17h 30
- 17-18/09
- <http://www.math.ist.utl.pt/~psoares/IEBW/>

##### Max-Stable Processes, Theory and Practice

- Prof. Mathieu Ribatet - EPFL - Switzerland
- 9h 30-12h / 14h-16h
- 14/10/09

##### Statistical Thinking for Non-Statisticians in the Pharmaceutical Industry

- Prof. Richard Kay - Independent Statistical Consultant
- 4-5/06/09
- <http://coursekay.ceaul.fc.ul.pt>

##### MiniCurso Análise de Dados Categorizados com Informação Omissa

- CD Paulino - IST/CEAUL e F Poeto - Universidade São Paulo
- FCUL 9h:12.30h / 14h-17.30h
- 22-23/01/09

#### Industry contract research

The group was the motor for the signature during 2009, through the Faculty of Sciences, of a protocol of cooperation with EXIGO consultores. EXIGO started its activity in 2001 and the aim is to provide consultancy, scientific and technical services, production and edition of scientific information in the area of health. As a result of this protocol, some members of this research group, helped in the planning of a trial in order to identify if a certain medical product should be put in the market or not.

#### Internationalization

MA Amaral Turkman was a visiting professor at the Institute of Mathematics of EPFL, Switzerland from 1/06 to 18/07 2009 where she collaborated on work related to the Bayesian Inference for max-stable processes. She was also invited to organize a session entitled "Statistics in the Omics Era", on the Conference "Statistical Methods for the analysis of large data-sets", which took place in Pescara, Italy, from 23-25/09/09".

KF Turkman was an invited speaker at Workshop on Spatial Extremes and Applications, EPFL Lausanne, 13-17/07/09. By invitation of Anthony Davison, he spent from 1/06 to 18/07 2009, as visiting professor at the Institute of Mathematics and CIB of EPFL. He worked with M. Ribatet and A. Davison on max-stable models and started investigation on Bayesian Inference for those models. Following this collaboration M Ribatet came to CEAUL in October. An article on "Models and Likelihood-based inference for Spatial Threshold Exceedances" is under preparation. KF Turkman was also invited by the Royal Statistical Society to contribute to the discussion of the read paper "Geostatistical inference under preferential sampling" PJ. Diggle et al. His contribution was accepted for publication.

L Carvalho and I Natario participate in 7º Programa-quadro DEEPFISHMAN, FP7 - KBBE- 2008-1-4-02. Management and Monitoring of Deep-sea Fisheries and Stocks: Case study: Black scabbardfish.

CD Paulino collaborates with F Poeto and J Singer from University of São Paulo. He is co-supervising the Ph.D. thesis of F Poeto. Their work on missing data mechanisms was published on "Statistics and Computing". Jointly with members of the Institute of Mathematics and Statistics of the UFRJ, Paulino submitted a project, which was approved, to CAPES/FCT for exchange of Ph.D. students and researchers. Together with P Soares he collaborated with Brazilian researchers Achcar and Rodrigues and published a joint paper.

V Andreozzi collaborates regularly with researchers from ENSP Sergio Arouca, Brasil on problems related to public health. As a result several papers were published. Jointly with M Antunes and Brazilian researchers they are preparing a book on Regression Models Applications in Health.

L. Gonçalves is a member of the projects

"Maternal Deaths and Severe Maternal Morbidity in Maputo City and Province" (SAREC - Swedish Agency for Research Cooperation)

and

"Support to Integrated Malaria Control in the Chókwè Region - Moçambique" (EU)

#### Future Research

##### Objectives

The group will continue working towards the main objectives (i) provide relevant methodological contributions to probability and statistical issues in modelling environmental and biological data, (ii) continue the Ph.D. program on the main research themes, namely Hierarchical Bayesian Modelling, Spatial-temporal Modelling, Statistics in Genetics and Life Sciences (iii) seek interdisciplinary collaboration (iv) organize seminars, workshops and advanced courses of general interest in the areas of expertise of the group.

During 2010 we have as objectives:

1. Prepare the 2nd edition of the textbook "Bayesian Statistics"; continue working on the book "Non-linear time series analysis"; start working on the

book "Survival analysis: theory and applications to health".

2. Revise the papers which were already submitted towards their final publication and . finish ongoing work towards their submission.
3. Start the collaboration with UFRJ team of the project "Análise bayesiana de sistemas estocásticos complexos: aspectos metodológicos e aplicações" approved in the framework of bilateral agreement FCT(Portugal)/CAPES(Brazil).
4. Continue the collaboration in the Portuguese Statistical Glossary, organized by SPE
5. Continue the consultancy role in DEEPFISHMAN project, regarding the Black scabbardfish abundance estimate in the Portuguese coastal waters and establishing management policies regarding its exploitation.
6. Continue to give statistical advice on project, applying GLMs and spatial methods for detecting and modeling accident patterns, aiming to establish intervention actions.
7. Continue the collaboration in the projects "Maternal Deaths and Severe Maternal Morbidity in Maputo City and Province" and "Support to Integrated Malaria Control in the Chókwè Region – Moçambique
8. Start the scientific collaboration with EXIGO consultants
9. Submit the Ph.D. dissertations of Emiliano Barreto, Sandra Ramos and Leonel Vicente
10. Participate actively in national and international conferences of the speciality
11. Organize seminars in the areas of interest, open to all the scientific community, and continue with the organization of an annual meeting of the group to discuss ongoing work.
12. Prepare a webpage specific for the group.
13. Work on the following topics
  - Bayesian inference on max-stable processes; the use of composite likelihood and application of Approximate Bayesian Computation methods.
  - Implementation of fully Bayesian hierarchical model for spatial Pareto process.
  - Spatial point processes, preferential sampling and their use in statistical modeling of wild fires.
  - Bayesian hierarchical modeling of airborne pollution data; application to Portuguese data.
  - Model Bayesian assessment; non-parametric methods to assess model adequacy
  - Bayesian hierarchical models for road safety data.
  - Bayesian non-parametric methods in supervised classification
  - Bayesian hierarchical models applied to the study of human resources
  - Bayesian hierarchical model to hospital length of stay: Comprehensive and accurate information about hospital length of stay (LOS) have important implications in the strategic planning and deployment of financial, human, and physical resources for health care providers. The objective is to propose a Bayesian hierarchical model to LOS and implement it using free available WinBugs and R software and, thus, provides a useful tool for health-environment investigations and health practice surveillance.
  - Collaborative work on the applications of statistical methods in tropical medicine, giving a particular attention to the latent class model in the diagnostic testing of some tropical diseases.
  - Further study about the empirical power of some normality tests for detecting location contaminated Gaussian mixtures
  - Further developments to the sample size determination using several methods and algorithms implemented in the R program
  - Diagnostic tests: development of scores to evaluate the predictive power of diagnostic tests. This objective will be achieved by studying issues such as, measures that adequately summarize positive and negative predictive values, measures that accommodate costs of wrong decisions; diagnostic tests based on more than one variable and link functions to deal with asymmetry. ROC methodology for bilateral diagnostic tests.
  - Syndromic surveillance: Finding a state-space model connected to the SIR model for epidemics with better fitting to the flu epidemic in Portugal, as well as the development and application of non-homogeneous Markov chains to flu epidemic surveillance.
  - Hidden Markov Models: the application of hidden Markov models to the flu epidemics in Portugal with the aim of doing short term projections.
  - State-space models: the application of state-space models to estimate the total abundance of certain species of sharks and black scabbardfish.
  - Study in a greater detail spatial point processes, sophisticated MCMC and sampling-resampling techniques, in order to apply them to the resolution of some problems, namely incidence estimation with imperfect data.
  - Use of Generalized Additive Neural Networks for prediction and comparison of its performance with Generalized Additive Models.
  - Use of PLS Path Modeling with medical data
  - Start the study of multivariate survival analysis toward the breast cancer data problem and the study of the joint analysis of survival and longitudinal modeling.
  - Identification of differentially expressed genes: two different approaches will be considered. One approach based on the area under the ROC curve as a procedure to rank up- and down-regulated genes and another based on principal components analysis (PCA). Robust PCA followed by supervised classification methods as classification trees, neural networks and nearest neighbours will be considered.
  - Web tool for predicting the topology of transmembranes proteins based on (Bayesian Estimation of Transmembrane Proteins) BETH algorithm, at <http://beth.deio.fc.ul.pt>, will be improved according to: (i) include the newer version of the algorithm which adopts a non homogeneous Markov model;(ii) incorporation of the prediction of signal peptides.

#### Funding, source, dates

All the funding received during 2009 and reported in the section 1 of this group description report was transferred to 2010

Projects: The following Project has Principal Investigator from the Group:

(RG-MATH-LVT-Lisboa-6-951) - Statistical Modelling in Environmental and Life Sciences

FCT/ PTDC/MAT/64353/2006

MEGA – STATISTICAL METHODS IN GENETIC AND ENVIRONMENT

Funding Entity: PTDC - FCT

Total Award Ammount: 60 000 Eur

Prime Contractor: Universidade de Lisboa - Fundação da F.C.U.L.

Period Covered: from September of 2007 to September of 2010 (extension asked till 31 December 2010)

Coordinator: Prof. K.F.Turkman

<b>Group Description</b>	
<b>Title of Research Group:</b>	(RG-LVT-6-1519) Applied Multivariate Data Analysis and Modelling
<b>Principal Investigator:</b>	Maria Helena Bacelar Valente da Costa Nicolau
<b>Main Scientific Domain:</b>	Matemática
<b>Group Host Institution:</b>	Faculdade de Ciências - Universidade de Lisboa
<b>Funding, source, dates</b>	
<b>Funding, source, dates</b>	
ALL THE REFERRED AMOUNTS INCLUDE OVERHEADS TO PAY FFCUL	
FCT Base for Research Group Title: (RG-MATH-LVT-Lisboa-6-1519) - Applied Multivariate Data Analysis and Modelling	
Funding:	
corresponding to 2009:	
date: 3 April 2009/ 6.735,76Eur	
date: 7 October 2009/ 2.841,64 Eur	
<b>Objectives &amp; Achievements</b>	
<b>Objectives</b>	
<p>This research project was based on the development and utilization of statistical models of multivariate analysis for extracting and analysing knowledge from large multivariate data sets. Applications were especially concerned with Behavioural Sciences, Education and Biomedicine.</p> <p>Major aims were:</p> <ul style="list-style-type: none"> <li>- Empirical and probabilistic approach to hierarchical clustering methods of variables or of statistical data units in a tridimensional and complex (using heterogeneous variables) context based on generalized affinity coefficients. Extension to interval variables.</li> <li>- Development of empirical and probabilistic procedures for validating and interpreting the results on real and on simulated data;</li> <li>- Research in randomization tests analysis related to power and serial dependency and development of applications for Conventional MANOVA, Nonrandomized Designs and Nonrandomized Classification of Variables</li> <li>-Development of new combining models in Discrete Discriminant Analysis (DDA): a new methodological approach was proposed, based on a linear combination of the First-order Independence Model (FOIM) and the Dependence Trees Model (DTM).</li> <li>- Research concerning the Hierarchical Coupling Model (HIERM), exploring the comparison of the HIERM model with other models where the choice of the "branches" of the binary tree is made using one of the traditional similarity coefficients in Cluster Analysis in alternative to affinity coefficient.</li> <li>- Investigation and Comparison of some Biometrical Methods useful in Genotypes Stability Evaluation in Series of Experiments as well as some Experimental Design techniques.</li> <li>- Using R software on the algorithms' implementation and applications.</li> </ul>	
<b>Main Achievements</b>	
<ul style="list-style-type: none"> <li>- Accomplishment and/or development of several PhD and Master thesis, some issued from scientific and technical co-operation developed inside European projects. All these thesis present several kind of applications, being particularly connected with the main areas of application indicated as main objectives.</li> <li>- Publication and presentation of papers in national and international conferences; in this point investment was made, in the domain of exploratory and confirmatory multivariate data analysis methodology and applications, in high quality books of proceedings, as well as journals or others.</li> <li>- Organization and co-organization of national and international meetings, seminars, workshops or thematic /invited sessions in the areas related to our research, as well as participation in scientific committees.</li> </ul>	
<b>Group Productivity</b>	
<b>Publications in peer review Journals</b>	
<p>Helena Bacelar-Nicolau, Fernando Nicolau, Áurea Sousa, Leonor Bacelar-Nicolau. "Measuring similarity of complex and heterogeneous data in clustering of large data sets". Biocybernetics and Biomedical Engineering, PWN-Polish Scientific Publishers Warszawa, Vol 29, Nº 2: 9-18 - Indexed and abstracted in: Science Citation Index Expanded (SciSearch®) and Journal Citation Reports/Science Edition</p> <p>Ferro, J.M., Bacelar-Nicolau, H., Rodrigues, T., Bacelar-Nicolau, L., Canhão, P., Crassard, I., Bousser, M.G., Dutra, A.P., Massaro, A., Mackowiack-Cordiolani, MA, Leys, D., Fontes, J., Stam, J. &amp; Barinagarrementeria, F. "A Risk Score to Predict the Outcome of Patients with Cerebral Vein and Dural Sinus Thrombosis", Cerebrovascular Diseases Journal, Karger, Vol. 28: 39-44 - IF 2008: 2.534</p> <p>Ramos, M. R. Carolino, E., Oliveira, T.A., Silva, A. P. , Carvalho, R. and Bicho, M.. Haptoglobin, acid phosphatase and demographic factors: obesity risk. In Biometrical Letters, Vol. 46, nº1, 43-54.</p> <p>Valente, V., Oliveira, T.A. (2009). Hierarchical Linear Models in Education Sciences: an application. In Biometrical Letters, Vol. 46, nº1, 71-86.</p> <p>Justo, M. R. M., J. Marques, R. M., A., Sousa Ferreira, A. (2009). Caesarean Birth Vs. Vaginal Delivery. Effects upon Maternal Satisfaction, Maternal Confidence and Maternal Perception of the Newborn Behaviour. INFAD, International Journal of Developmental and Educational Psychology ,ISSN 0214-9877, XXI, 2(1), 45-57.</p>	



Áurea Sousa, Fernando Nicolau, Helena Bacelar-Nicolau, Osvaldo Silva (2008). Weighted Generalised Affinity Coefficient in Cluster Analysis of Complex Data of the Interval Type. *Biometrical Letters - Special Issue* (reference IWLBP-2008) (in press).

Sousa Ferreira, A., Lima R. (2009). Discriminant Analysis in career studying "decision/indecision": The Career Factors Inventory (CFI) as a diagnostic measure. *Spanish Journal of Psychology* (to appear) (2008 IMPACT FACTOR = 0.558)

Branco, F., Oliveira, T.A., Oliveira A.: The Impact of Distribution Shape on the Power of Randomization Tests for two independent groups: A simulation study using small balanced samples. (accepted *Biometrical Letters* 2010).

#### Other publications International

D. G. Pereira, P. C. Rodrigues, A. Oliveira, J. T. Mexia. (2009). Joint Regression Analysis and Completed Joint Regression Analysis. In *Plant Breeding* (F. Columbus Eds.) Nova Science Publishers, New York, 2009. ISBN: 978-1-61668-211-8.

Sousa Ferreira, A. (2009). A Comparative Study on Discrete Discriminant Analysis through a Hierarchical Coupling Approach. In *Classification as a Tool for Research, Studies in Classification, Data Analysis, and Knowledge Organization*; Hermann Locarek-Junge, Claus Weihs (Eds.), Springer, Berlin-Heidelberg-New York (to appear).

Oliveira, T.A.(2009)."BIB Designs with Repeated Blocks: Review and Perspectives". Invited Session: Frontiers in Statistics.10th Islamic Countries Conference on Statistical Sciences (ICCS-X), Islamic Countries Society of Statistical Sciences (ISOSS), in American University in Cairo, Egypt, 20-23 de Dezembro de 2009 (to appear).

Oliveira, A., Oliveira, T.A. (2009). "Recursos Computacionais no ensino e na Aprendizagem da Estatística em ambiente online". Proceedings of XIII Encontro Iberoamericano de Educação Superior a Distância (in CD).Editor: Universidade Aberta. ISBN 978-972-674-689-8.

Bacelar-Nicolau L, Craveiro I, Pereira Miguel J, Saporta G, "HIA of Social Welfare Payment for Inclusion Policy: Tools to Improve Health Gaps, Promote Equity" HIA09 On the Move- 10th Conference on Health Impact Assessment (Rotterdam, Holanda)

Bacelar-Nicolau L, Craveiro I, Pereira Miguel J, Saporta G, "Health Impact Assessment of a Social Welfare Payment for Inclusion Policy: A Tool to Improve Health Gaps and Promote Equity?" 2nd Joint European Public Health Conference 2009 (Lozd, Poland)

Helena Bacelar-Nicolau, Fernando Nicolau, Áurea Sousa, Leonor Bacelar-Nicolau (2009), Hierarchical Clustering models of Complex Data Described by Heterogeneous Variables in a Three-way Context. Book of Abstract 11th IFCS International Conference 2009 (IFCS 2009), Technische Universität Dresden Eds,

Oliveira, T.A.(2009). "Estatística e Planeamento de Experiências: Desafios do Ensino Online." Invited Talk: Workshop "Matemática a Distância no Contexto Ibérico", Universidade Aberta-UOC-UNED, Lisbon, 24-26th June, 2009.

Oliveira, A. (2009). Estatística Computacional com R em ambiente online. Invited Talk: Workshop "Matemática a Distância no Contexto Ibérico", Universidade Aberta-UOC-UNED, Lisbon, 24-26th June, 2009.

Regra, C., Oliveira, T.A.(2009). "Análise Estatística do Processo de Reconhecimento, Validação e Certificação de Competências: O impacto da formação no CNO dos CTT ". Workshop Matemática a Distância no Contexto Ibérico, Universidade Aberta-UOC-UNED, Lisbon, 24-26th June, 2009.

Valente, V., Oliveira, T.A.(2009). "Modelação do sucesso escolar nas diferentes áreas científicas: Uma aplicação do MLH usando os recursos computacionais SPSS e MLwiN". Workshop Matemática a Distância no Contexto Ibérico, Universidade Aberta-UOC-UNED, Lisbon, 24-26th June, 2009.

Oliveira T., Oliveira A..(2009). Upper Contour Method of a Joint Regression Analysis. The R User Conference 2009, Book of Abstracts, 148. [http://www.agrocampus-ouest.fr/math/useR-2009/abstracts/book\\_of\\_abstracts](http://www.agrocampus-ouest.fr/math/useR-2009/abstracts/book_of_abstracts)

Oliveira, T.A., Oliveira, A., Mejza, S.. Some Biometrical Methods useful in Genotypes Stability Evaluation in Séries of Experiments. EMR-IBS, Military Museum in Istanbul, 10-14 de Maio de 2009. Poster presentation.

Serpa, M., S. Caldeira, C. Gomes, A. Sousa & O. Silva (2009). Students' Perspectives on Themselves and on School Work – Contributing to the Construction of Evaluation Scales. Paper presented at the 11th International Conference on Education – ATINER, May 9-11, in Atenas, Greece.

Serpa, M., S. Caldeira, C. Gomes, A. Sousa & O. Silva (2009). Estudo preliminar de escalas de avaliação do funcionamento do aluno (Preliminary Study of Scales to Evaluate Student Performance) Paper presented at the X Congresso Galego-Português de Psicopedagogia, September 9-11, in Braga, Portugal.

Sousa Ferreira, A., (2009). A Comparative Study on Discrete Discriminant Analysis through a Hierarchical Coupling Approach. Book of Abstract (IFCS 2009), Technische Universität Dresden Eds, p. 134.

Marques, A.; Sousa Ferreira, A.; Cardoso, M. (2009). Discrete Discriminant Analysis: the performance of combining models. Book of Abstract (IFCS 2009), Technische Universität Dresden Eds, p. 95.

Serpa, M., S. Caldeira, C. Gomes, A. Sousa & O. Silva (2009). Evaluative Scales of the Students Functioning: Two Instruments to Support Teaching. Proceedings of ICERI2009 Conference. November 16-18, in Madrid, Spain. (ISBN:978-84-613-2955-7).

Helena Bacelar-Nicolau,(2009). Researching & Teaching Hierarchical Clustering Models. Workshop on Statistics, Mathematics and Computation (WSMC 2009) November 9-10, 2009, Lisbon - Fundação Calouste Gulbenkian, Portugal (invited talk). In Program and Abstracts, pp11.

Isabel Pinto Doria, Ana Sousa Ferreira, Otília Dias, Helena Bacelar-Nicolau, Georges Le Calvé,(2009). Comparison of Multivariate Analysis Methodologies in a Palliative Care Setting. Workshop on Statistics, Mathematics and Computation (WSMC 2009) November 9-10, 2009, Lisbon - Fundação Calouste Gulbenkian, Portugal In Program and Abstracts, pp51.

Oliveira, T.A., Oliveira, A.(2009): Fourth Workshop on Statistics, Mathematics and Computation, Book of Abstracts, Universidade Aberta

#### Other publications National

Fernando Cascais, Helena Bacelar-Nicolau, José Vidal Oliveira (2009). "Sondagens e Inquéritos de Opinião – Diagnóstico e sugestões de medidas a adoptar", Ed. ERC (Entidade Reguladora para a Comunicação Social), 78 pag.

Vedes, A. M., Narciso, I. Sousa Ferreira, A. (2009). Um Estudo Exploratório em Casais Portugueses sobre a Relação entre Percepções de Satisfação Conjugal, Satisfação com a Intimidade e a Comunicação e Conflitos, Proximidade e Vinculação ao Cônjuge. *Psychologica, Revista da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra*, 54 (in press).

Freire, I., Veiga Simão, A. M., Caetano, A. P., Sousa Ferreira, A. Cardoso, A. & Gouveia, S. (2009). Prevenção da violência entre pares/ promoção da

convivência nas escolas – um projecto colaborativo, Revista Interações. Escola Superior de Educação de Santarém (in press).

Sousa Ferreira, A., Dias, O. (2008). Os cuidados Paliativos em Contexto Domiciliário pelo Olhar do Estatístico. Revista de Filosofia (in press).

Gonçalves, B. Fagulha, T. & Sousa Ferreira, A. (2007). Depressão, depressão crónica e sintomatologia depressiva nas mulheres de meia-idade: relação com o nível sócio-cultural e a zona de residência [Depression, chronic depression and depressive symptomatology in women half-age: relationship with the socio-cultural level and the area of residence]. Revista Portuguesa de Psicologia (in press).

Fagulha, T., Gonçalves, B. & Sousa Ferreira, A. (2008). Olhares sobre a menopausa: prevalência e intensidade da sintomatologia menopáusica em função do nível educacional e zona de residência. Revista Portuguesa de Psicologia (in press).

Helena Bacelar-Nicolau, Áurea Sousa, Leonor Bacelar-Nicolau, M. Silvério Marques (2008). Do univariado ao multivariado: a escala de elementos tangíveis, suas relações com outras escalas e mais além. Revista Portuguesa de Filosofia (in press).

Monteiro, S., Oliveira, T.A.. (2009). Construção de fracções nos planos factoriais de base dois. Livro de Actas do XVI Congresso Anual da SPE, Edições SPE, 2009, pp. 457-464.

Marques, A., Sousa Ferreira, A., Cardoso, M.. Uma proposta de combinação de modelos em Análise Discriminante. Estatística - Arte de Explicar o Acaso", In Oliveira, I. et al. Eds., Ciência Estatística, Edições S.P.E., p. 393-403.

Sousa, A.; Nicolau, F.; Bacelar-Nicolau, H. (2009). Análise Classificatória de Dados "Three-Way": uma aplicação a dados reais. In: Oliveira, I. e tal. Eds. Actas do XVI Congresso Anual da Sociedade Portuguesa de Estatística, 2008, Edições S.P.E., 673-682.

Pinto, A., Lobo, V., Bação, F., Bacelar-Nicolau, H., "A auto-percepção do estado de saúde e as diferenças sócio-económicas na utilização dos serviços de saúde, ao nível da NUTS II" (submitted).

Oliveira, A. (2009). Análise Conjunta de Regressões Completada na Condução Dinâmica de Planos de Melhoramento. Workshop em Estatística pela Jubilação do Professor João Tiago Mexia, FCT-Universidade Nova de Lisboa, 4 e 5 de Junho 2009. (Apresentação Oral).

A. Oliveira, J.T. Mexia (2009). Análise Conjunta de Regressões: uma extensão da técnica. XVII Congresso da Sociedade Portuguesa de Estatística, 30 de Setembro a 3 de Outubro 2009, Sesimbra, pág 285.

Marta Guedes, Tiago Ferreira, Ana Bivar, Alberto Cardoso, Isabel Doria.(2009) Variação do Ângulo Irido-Corneano Após Implante de ICL. Comunicação apresentada no 52º Congresso Português de Oftalmologia, Vilamoura, 5 a 7 de Dezembro de 2009.

Fernando Cascais, Helena Bacelar-Nicolau, José Vidal Oliveira (2009). "Sondagens e Inquéritos de Opinião. Diagnóstico e Sugestões de Medidas a Adoptar", Painel convidado – "Sondagens e Jornalismo. Práticas e boas práticas", IIIª Conferência Anual da Entidade Reguladora para a Comunicação Social (ERC), "A Comunicação Social num Contexto de Crise e de Mudança de Paradigma", 20-21 de Outubro de 2009, Lisboa, Fundação Calouste Gulbenkian.

Doria, I.; Sousa, A.; Bacelar-Nicolau, H.; Calve, G. (2009). Comparação de variáveis modais em análise multivariada. Em Programa e Resumos do XVII Congresso Anual da Sociedade Portuguesa de Estatística, 269-270.

Sousa Ferreira, A., Cardoso, M. (2009). Comparação de Resultados em Análise Discriminante Discreta. In: Book of Abstracts of XVII Congresso Anual da Sociedade Portuguesa de Estatística (SPE), SPE Eds., 47, Sesimbra, Outubro 2009.

Dias, O., Pinto Doria, I., Sousa Ferreira, A. (2009). A Análise Multivariada e as Competências Específicas de um Licenciado em Engenharia Civil. In: Book of Abstracts of XVII Congresso Anual da Sociedade Portuguesa de Estatística (SPE), SPE Eds., 327, Sesimbra, Outubro 2009.

Marques, A.; Sousa Ferreira, A.; Cardoso, M. (2009). Resultados de uma Escala de Sugestionabilidade: Classificação em grupos demográficos. Book of Abstract, XVI Jornadas de Classificação e Análise de Dados (JOCLAD 2009), Universidade do Algarve, p. 69.

Oswaldo Silva, Helena Bacelar-Nicolau, Fernando Nicolau, "Avaliação e comparação de partições em Análise Classificatória Hierárquica", XVI Congresso da SPE, SPE 2009, Sesimbra, 2009, Edições S.P.E.

Oswaldo Silva, Áurea Sousa, Margarida Serpa, Helena Bacelar-Nicolau, "Análise Classificatória Hierárquica no Contexto da Identificação de Aspectos do Funcionamento do Aluno", XVI Jornadas de Classificação e Análise de Dados, Livro de Resumos das JOCLAD 2009, Universidade do Algarve, 2009, CLAD Ed., 76-78.

Áurea Sousa, Licínio Tomás, Oswaldo Silva, Helena Bacelar-Nicolau, "Análise da Satisfação com os Serviços de Leitura Presencial por Métodos da Estatística Multivariada", XVI Jornadas de Classificação e Análise de Dados, Livro de Resumos das JOCLAD 2009, Universidade do Algarve, 2009, CLAD Ed., 79-81.

Helena Bacelar-Nicolau, Fernando Nicolau, Áurea Sousa, Leonor Bacelar-Nicolau, "Can we combine binary and interval variables to cluster complex statistical data units? Yes, we can!", XVI Jornadas de Classificação e Análise de Dados, Livro de Resumos das JOCLAD 2009, Universidade do Algarve, 2009, CLAD Ed., 82-84.

Palmira Caseiro, Fernando Nicolau, Helena Bacelar-Nicolau, "Estatísticas da Educação nas Escolas de Angola: Análise Preliminar com Métodos da Estatística Multivariada", XVI Jornadas de Classificação e Análise de Dados, Livro de Resumos das JOCLAD 2009, Universidade do Algarve, 2009, CLAD Ed., 142-145.

Isabel Pinto Dória, Otilia Dias, Ana Sousa Ferreira, Helena Bacelar-Nicolau, Georges Le Calvé, "O Perfil de Competências num Licenciado em Engenharia Civil numa Instituição do Ensino Superior: Uma abordagem pela Análise Multivariada", XVI Jornadas de Classificação e Análise de Dados, Livro de Resumos das JOCLAD 2009, Universidade do Algarve, 2009, CLAD Ed., 175-179.

Oliveira, T.A. (2009). Abordagem geométrica e estruturas dos múltiplos do PBIE (7,7,3,3,1). Resumo nas Actas do Congresso da SPE, Sesimbra, Portugal, organized by FCT – Universidade Nova de Lisboa, de 30 de Setembro a 3 de Outubro de 2009.

Roçadas, C., Oliveira, T.A. (2009). Modelos de Cadeias de Markov: Aplicação ao Sector Bancário. Workshop em Estatística pela Jubilação do Professor João Tiago Mexia, FCT - Universidade Nova de Lisboa, 4 e 5 de Junho 2009.

#### Master and Ph.D. thesis completed

Doria, Isabel Pinto. "Representações Euclidianas de Dados: Uma Abordagem para Variáveis Heterogéneas". Doutoramento em Biomatemática.

Supervisors: H. Bacelar-Nicolau, Universidade de Lisboa, G. Le Calvé, Universidade de Rennes II, França.

Oliveira, T.A. supervised the PhD thesis of Cláudia Vanessa Rosa Leitão de Macedo Roçadas. PhD in Matemática-Modelação Estatística, "Populações Emparelhadas com Reclassificação Periódica; Aplicação a uma carteira de clientes", in Universidade Aberta, co-supervised by Professor João Tiago Mexia, FCT- Universidade Nova de Lisboa.

Oliveira, T.A. supervised Paula Cecília dos Santos Leitão Caetano Alves in "Modelos de Planos em Blocos Incompletos: Revisão e Perspectivas", Master Thesis on Statistics and Optimization, FCT, New University of Lisbon, (approved in 2009, May 29th).

Oliveira, T.A. supervised Carlos Manuel Fernandes Regra in "Análise Estatística do Processo de Reconhecimento, Validação e Certificação de Competências (Em particular no CNO dos CTT)", Master Thesis on Statistics Mathematics and Computation - Computational Statistics, Universidade Aberta, (approved in 2010, february 5).

#### Organization of conferences

XVI Jornadas da Associação Portuguesa de Classificação e Análise de Dados, JOCLAD 2009, Abril 2 – 4, 2009, Universidade do Algarve, Faro. H. Bacelar-Nicolau was member of the Scientific Committee and Organizer of an Invited Session on "Análise Classificatória"

Oliveira, T.A. "Planos de Blocos Incompletos Equilibrados com Repetições: Revisão e Perspectivas". Seminar in Évora University, 2009, May 28th.

IV WSMC- IV Workshop in Statistics, Mathematics and Computation, Universidade Aberta, Fundação Calouste Gulbenkian, Lisbon, November 2009, 9-10. Oliveira, T.A. was Chair and Member of Scientific and Organizing Committees. Oliveira, A. was member of Scientific and Organizing Committees.

Oliveira, T.A. was Chair in Invited Session on the "Workshop em Estatística pela Jubilação do Professor Doutor João Tiago Mexia", 2009, June 4- 5.

Jornadas Anuais da Associação Portuguesa de Classificação e Análise de Dados (JOCLAD 2009), was held at Faro in 2-4 Abril 2008. A. Sousa Ferreira was member of the Scientific Committee. <http://www.fct.ualg.pt/joclad2009/JOCLAD2009/Apresentacao.html>

#### Internationalization

- European Inter-University Cooperation Programme (SOCRATES-ERASMUS), 2006-2009. Includes the Annual Intensive Seminar on "Mathematical and Computational Models in the Psychological Sciences" as well as students and teachers mobility

H. Bacelar-Nicolau is the Portuguese Responsible for this Project, as Representing Professor of the University of Lisbon.

Partner Universities are: Glasgow (England) (IP coordinator); Graz (Austria); Tartu (Estonia); Bremen (Germany); Oldenburg (Germany); Leuven (Belgium); Padova (Italy) ; Debrecen (Hungary); University of Lisbon (Portugal); Complutense Madrid (Spain); Tubingen (Germany); Amsterdam (The Netherlands).

#### Future Research

##### Objectives

No objectives were introduced in this group research, since the group will disappear as such. Members of this research group will integrate the other research groups.

##### Funding, source, dates

All the funding received during 2009 and reported in the section 1 of this group description report was transferred to 2010