Fundação para a Ciência e a Tecnologia

Evaluation Reseach Unit 2013

Document Date: Mon Feb 03 08:45:20 WET 2014

Locked Date: Mon Dec 23 10:43:00 WET 2013

1. IDENTIFICATION OF THE R&D UNIT

1.0 Reference

6

1.1 Name of the R&D Unit

Centre of Statistics and its Applications

1.2 Acronym

CEAUL

1.3 Coordinator

Maria Antonia Conceição Abrantes Amaral Turkman

1.4 Multidisciplinary/interdisciplinary R&D Unit

No

Scientific areas

Mathematics**(Main Scientific Area)**

1.5 Profile of the R&D Unit

70% Basic research

30% Applied research and/or Experimental development

1.6 Keywords

Statistical Inference

Statistical Modelling and Data Analysis

Biostatistics

Statistical Computing

1.7 Link to the R&D Unit's page on the Internet

http://www.ceaul.fc.ul.pt

1.8 Unit registration options

Keep

2. INSTITUTIONS AND THEIR ROLES

2.1 Main Host Institution

Fundação da Faculdade de Ciências (FFC/FC/UL)

2.2 Other(s) Host Institution(s)

2.3 Participating Institution(s)

Institution Name Host Institution

2.4 Institutional commitment

FCUL (Faculdade de Ciências da Universidade de Lisboa) created in 1911, is the participating institution for all its research units (formerly Instituição de Acolhimento). FCUL supplies space, general computing infrastructure, communications and, through its foundation (FFCUL - Fundação da FCUL), administrative and financial R&D (research and development) management.

Research units are associated to (one or several) FCUL departments, which are in charge of space allocation and may support interface with FCUL and/or FFCUL (see below) operational services. FCUL is also using space within the University of Lisbon Institute for Interdisciplinary Research, located at the Campo Grande university campus and at walking distance from FCUL, for some of its R&D units in mathematics and in physics.

FCUL associated R&D units have natural access to all FCUL labs and teaching areas to held meetings, seminars and workshops, especially during non-academic periods. R&D units also benefit from IT services, enabling most of them to locate their internet sites within the University servers. Legal, contractual and IP support is also directly provided by FCUL when requested by units.

Administrative and financial management are provided by FFCUL - the Main Host Institution - a non-profit organization recently reclassified as a public institution created in 1993 as an initiative of the FCUL to promote and facilitate scientific research, technology and development services to society, and to organize qualified human resources training, offer consulting expertise and promote knowledge dissemination.

FFCUL acts as the legal front institution of multiple research units, from different scientific fields, carrying out research work and managing financial and administratively R&D (research and development) projects, with more than 400 ongoing projects. Many of these R&D activities are developed together with international teams and are funded both at national and European levels.

FFCUL provides the financial stability, the accounting support and the specialized human resources needed to comply with the overall project applications and financing programs and all different expenses reports required throughout project development. FFCUL human resources include an 12 person team in the financial department and 14 project managers.

CEAUL is a Unit I&D, close to Department of Statistics and Operations Research (DEIO) of Faculty of Sciences of the University of Lisbon, located at the 4th floor of the C6 building, which offers physical facilities to CEAUL- infrastructures, installations, budget management and contribution of a technician - to the prosecution of CEAUL's activities. CEAUL contributes largely, with its annual budget to DEIO's library by regularly acquiring the most recent and important books in every domain of Statistics. This Library with more than 4000 specialized books is open to the scientific community and is very sought for its actuality and excellence.

3. R&D UNIT DESCRIPTION AND ACHIEVEMENTS

3.1 Description of the R&D Unit

This I&D Unit plays a central role in Portugal as a Research Centre in Probability, Statistics and its Applications,

developing multidisciplinary activities that aim to carry out research in domains as diverse as Health Sciences and Biology, Financial Risk and hazards posed by nature in Geophysics and Environment. CEAUL is a member of CNM -Adhering Organization and Committee for Mathematics/ International Mathematical Union (IMU) and is an associated member of CIM - Centro Internacional de Matemática/ERCOM - European Research Centres on Mathematics. CEAUL keeps an updated library, open to the general scientific community.

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, Environmental Statistics, Statistics in Genomics, Statistics in Medicine, Longitudinal models and Survival Analysis, Time Series Analysis, Categorical Data Analysis and Missing Data, Sampling, Principal Component Analysis, Classification Methods, Temporal and Spatial Statistics, Design of Experiments and History and Teaching of Statistics.

During the period 2008-2012, the main objective of CEAUL was to contribute to the development of new theoretical framework and methodological issues addressed to applications to real problems, making efforts towards the correct use of Statistics in Society and Academics. That was attained through different inter-linked activities, among which we mention the organization and participation in conferences; organization of advanced courses and seminars open to the general scientific community; dissemination of results through the scientific publication both in National and International peer review journals; supervision of Ph.D. and M.Sc students; collaboration with researchers in other scientific areas and services to community.

Also an important objective of this research unit was to maintain activities in order to promote Statistics in Society and Schools. Towards that, CEAUL has strong links and cooperates with the National Institute of Statistics (INE), Portuguese Statistical Society (SPE), Portuguese Mathematical Society (SPM), Center of International Mathematics (CIM), Universities and Polytechnics, among other institutions. In particular it is important to refer to (i) the editorial support given by senior members of CEAUL to REVSTAT - Statistical Journal, an International Journal in the area of Probability and Statistics in English, edited by INE; (ii) the production of the Portuguese Statistical Glossary and the edition of the SPE newsletter, both as part of the joint activities with SPE. In the teaching of Statistics at non-university level, some collaborators of the center have a prominent role in ALEA project (http://alea-estp.ine.pt/, jointly with INE), and their activity is recognized as a landmark in improving the teaching of Statistics at secondary school level. Also, several members were involved in the revision of basic and secondary school books of Mathematics (Probability and Statistics) and in producing documentation for teachers.

In order to achieve its main objectives, members of CEAUL were grouped, according to their research interests, in three main research lines (GI), led by senior researchers who defined the program of activities in order to give coherence to the group, but keeping always in mind the fulfillment of the objectives and main goals of the research unit. The GI's were always encouraged to interact between them in their activities, whenever that was appealing, mainly in the organization of advanced courses, conferences and seminars.

The group leaders, together with the elect Coordinator and Vice-Coordinator, were the core of the scientific panel whose responsibility was to monitor and inflect the main goals of the research unit, and to rule the investment policies. The research groups were

• GI 1: Order Statistics, Extremes and Applications
• GI 2: Probability, Modelling and Data Analysis
• GI 3: Statistical Modelling in Environmental and Life Sciences

The hot research topic of GI1 group is in Extreme Value Theory and the main topics of research in 2008 and 2012 were on fields of univariate, multivariate, multidimensional and spatial extremes, with special emphasis on their applications to life sciences, environment, risk, insurance, finance, outliers and biometry. Topics in the areas of statistical quality control, design of experiments and forensic statistics were of related interest. The team GI2 had a wide range of interests and the research program involved work on parametric and semi-parametric inference, distribution theory, dynamical systems, sampling, optimal design, and applications in Life Sciences, sample augmentation and meta-analysis issues, History of Statistics and Teaching of Statistics. Team GI3 divided their activities in methodological and applied research. From a methodological point of view the team worked on a variety of ideas reflecting their main motivating topics, namely Bayesian statistics, spatio-temporal models and categorical data analysis. The spectrum of statistical applications was large, particularly in the fields of environmental and health risks, ranging from the modeling and analysis of forest fire data in Portugal to the application of state-space models to estimate the total abundance of certain species of sharks and black scabbard fish, passing through applications to road safety, genetics, epidemiology and health.

The Internal Regulation of CEAUL provides a Scientific Committee, which approves in a plenary meeting the general guidelines of Scientific and Financial management. During 2008-2012, part of the budget was used for common expenses (Periodicals and Books for the Library, Common Software and Informatics Assistance as well as maintenance of WebPage) and the remaining was shared among the GI's, according to the number of PhDs in each GI. Although collaborators were not sponsored by FCT, some of their activities were financed by the Center, particularly if they were Ph.D. students, under the recommendation of the GI leaders. There was a scientific as well as a financial autonomy of the GIs.

3.2 Major achievements

(i) Dissemination of research: CEAUL as a whole has been increasing the scientific production and this was one of the main achievements during 2008-2012. His members made 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. A complete list of publications is available in http://www.ceaul.fc.ul.pt/files/20082012_PUBLICACOES_3.pdf

(ii) Interaction of research: There has been a direct evidence of interaction between the theoretical and applied components of the work. Moreover, there has been a considerable effort on theoretical issues reinforced with involvement in applications to real world. During 2008-2012 members of CEAUL actively collaborated with researchers from several research institutions in different fields, resulting in international publications either in periodicals from the area of Statistics or in periodicals from specialized areas in other fields. The motivating problems of collaboration were on diverse areas such as: health, forest fires, ecology, fisheries, chemical production, radioactivity, quality control, insurance, finance, and education.

(iii) Reinforcement: There has been a new reinforcement in some branches, as for instance: PORT methods in Extreme Values, Internet Auction Markets, History of Probability and Statistics in Portugal in the early XXth century, Teaching of Statistics with Documentation for Teachers, in collaboration with INE's Alea Project, Quantification of Regional Risk of Large Wildfire Sizes, Study of the Evolution of the Human Resources of the Faculty of Sciences of Lisbon, Theoretical Issues in Applications with Genome Analysis, Statistical Models to Genetics and Epidemiological Data, Optimal Screening Methods in Gene Profile Classification, Missing Data and Categorical Data Analysis.

(iv) Projects: Integration of the research teams of international projects, such as the DEEPFISHMAN project, regarding the Black scabbard fish abundance estimate in the Portuguese coastal waters; 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. The project on the data support centre of MalariaGEN - Genomic Epidemiology Network of Malaria; the Spanish Biostatistics Net named BIOSTATNET; the project - Bayesian analysis of complex stochastic systems: methodology and applications - approved in the framework of bilateral agreement FCT(Portugal)/CAPES(Brazil).

(v) Software development: Members of the centre have been involved in software development. The flubase package developed by Nunes B, Natario I and Carvalho L and The R package bild - a package for Binary Longitudinal Data, developed by Gonçalves, H., Cabral, S. and Azzalini, A , and Package ACD: Categorical data analysis with complete or missing responses, developed by Paulino and collaborators, are already made available to the R community.

(vi) Internationalization: International recognition in the areas of Statistics of Extremes, Environmental and Financial Statistics, Bayesian Statistics, Time Series and Biostatistics was also a main achievement during the past five years. This recognition is well manifested particularly through publications with international researchers, invitations received to give talks in conferences and seminars in well recognized international institutions, members as associated editors of international periodicals.

(vii) Advanced Training: Besides the natural involvement in the current Ph.D. programs available in the variety of Institutions that CEAUL members belong to, CEAUL organized a large number of advanced courses on emerging areas of Statistics for researchers and advanced students from all over the country, as well as courses on more general areas of Statistics, for researchers and Ph.D. students who want to learn more about some specific subject. These courses attracted many researchers all over the country and increased the recognition of CEAUL as a leader of Statistics in the Portuguese scene.

(viii) Consultancy and Service to Community: The Coordinator of CEAUL acted as consultant for the project led by the consortium Siemens Joint Commission International (JC) towards the development of a National Health Care Evaluation System in Portugal (SINAS - Sistema Nacional de Avaliação em Saúde), in response to a challenge set by the Portuguese Health Regulatory Authority. This was also an opportunity for three MSc students to initiate their research activity while taking part of this project. Ana Luisa Papoila gave consultancy support to the Research Unit of Centro Hospitalar de Lisboa Central (CHLC) as co-head of the Epidemiology and Statistics Consulting Core. Also CEAUL was the motor for the signature during 2009, through the Faculty of Sciences, of a protocol of cooperation with EXIGO consultores, which aims to provide consultancy, scientific and technical services, production and edition of scientific information in the area of health. As a result of this protocol, members of CEAUL, helped in the planning of a trial in order to identify if a certain medical product should be put in the market or not, collaborated in meta-analyses studies and were involved in a research project aimed to assess the financing system of the public hospitals belonging to the Portuguese National Health Service. The report of this project was used by the Portuguese Court of Auditors in its final report audit.

(ix) Outreached activities: Important for the scientific community in general was the release of the new version of English-Portuguese Statistical Glossary through the sites of Portuguese Statistical Society and Brazilian Statistical Association (http://glossario.spestatistica.pt/). The group responsible to produce this glossary had 6 members, three from each country. Two of the Portuguese members were from CEAUL, namely Daniel Paulino as the principal coordinator and Dinis Pestana; the responsible for the edition of the glossary was Paulo Soares, also a member of CEAUL

Also REVSTAT - Statistical Journal (http://www.ine.pt/revstat/inicio.html), whose editor is Ivette Gomes and co-editor Antónia Turkman, gained international recognition by being included in SCOPUS and ISI web of knowledge.

Lucília Carvalho was the Leader of the Working Group of the National Statistical Institute (INE) on Labor Force Survey - Implementation of Computer Assisted Telephone Interviewing /CATI.

Key Publications

Araújo Santos, P. and Fraga Alves, M.I. (2012). "A new class of independence tests for interval forecasts evaluation". Computational Statistics and Data Analysis 56:11, 3366-3380

Buishand, A., de Haan, L. and C. Zhou (2008).On spatial extremes; with application to a rainfall problem. Annals of Applied Statistics, 2, 624-642.

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

392-413.

Canto e Castro, L. & S. Dias (2011). Generalized Pickands' estimators for the tail index parameter and max-semistability.Extremes 14:4, 429-449

de Zea Bermudez, P. and Kotz, S. (2010). Parameter estimation of the generalized Pareto distribution - Part I. Journal Statistical Planning and Inference, 140, 6, 1353-1373 and Part II, 1374-1388

Nunes, B., Natário, I. and Carvalho, L. (2011).Time series methods for obtaining excess mortality attributable to influenza epidemics. Journal of Statistical Methods in Medical Research.;20(4):331-45.

Poleto F., Singer J. and Paulino C.D. (2010).Missing data mechanisms and their implications on the analysis of categorical data. Statistics and Computing 21, 31-43

Silva, G.L., Dean, C.B, Niyonsenga, T. and Vanasse, A. (2008).Hierarchical Bayesian spatiotemporal analysis of revascularization odds using smoothing splines. Statistics in Medicine, 27, 2381-2401

Turkman, K.F., Amaral Turkman, M.A. and J.M. Pereira(2010) Asymptotic models and inference for extremes of spatio-temporal data, Extremes 13, 375-397

Pereira, C., Telles Antunes, M., Pestana, D., Mendonça, M.C., and Santos, J.C. (2012). An Unidentified Skeletal Assemblage from a Post-1755 Mass Grave of Lisbon: Dental Morphology and Population Affinity. Bulletin International Association of Paleodontology, 6(1):12-26.

3.3 **External Advisory Committee Reports**

R&D Unit	File
Centre of Statistics and its Applications	2011.pdf
Centre of Statistics and its Applications	2013.pdf
Centre of Statistics and its Applications	2010.pdf
Centre of Statistics and its Applications	2012.pdf

3.4 **Composition of the External Advisory Committee**

R&D Unit	Member Name	Institution
Centre of Statistics and its Applications	Jürg Hüsler	Institute of Mathematical Statistics and Actuarial Science, University of Bern, Switzerland
Centre of Statistics and its Applications	Wolfgang Urfer	Department of Statistics, University of Dortmund, Germany
Centre of Statistics and its Applications	António Gabriel da Silva St'Aubyn	Instituto Superior de Agronomia, Universidade de Lisboa, Portugal

3.5 **Brief description of the output indicators 2008/2012 of the research team of the new unit that support the vision and objectives of the strategic program**

This item is only compulsory for new R&D units. However since we set, as an important objective, the collaboration with researchers from other areas, namely health and environment, we give some examples of the 56 collaborative works published in periodicals from other scientific domains, by members of the present team, during 2008-2012.

Achcar, J. A., Martinez, E. Z., Ruffino-Netto, A., Paulino, C. D., Soares, P.(2008), A statistical model investigating the prevalence of tuberculosis in New York City using counting processes with two change-points, EPIDEMIOLOGY AND INFECTION, Vol.136, pp. 1599-1605.

Barreto-Hernandez, E., Gama-Carvalho, M., Sousa, L.,(2011) Pre-processing optimization of RNA immunoprecipitation microarray data, JOURNAL OF COMPUTATIONAL BIOLOGY (2011), Vol. 18, pp. 1319-1328.

Bruno-Soares, A.M., Cadima, J. & Matos, T.J.S., (2010) Predicting degradability parameters of diets for ruminants using regressions on chemical parameters. Journal of the Science of Food and Agriculture, 90, 949-955

Costa V, Lourenço HM, Figueiredo I, Carvalho L, Lopes H, Farias I, Pires L, Afonso C, Vieira AR, Nunes ML, Gordo LS,, Mercury, cadmium and lead in black scabbardfish (Aphanopus carbo Lowe, 1839) from mainland Portugal and the Azores and Madeira archipelagos, SCIENTIA MARINA (2009), Vol.73, pp. 77-88.

de Zea Bermudez, P, Mendes-Lopes, J, Pereira, J M C, Turkman, K F, Vasconcelos, M J P, Spatial and temporal extremes of wildfires sizes in Portugal (1984-2004), INTERNATIONAL JOURNAL OF WILDLAND FIRE (2009), Vol.18, pp. 983-991.

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 (2009), 9: 125.

Gonçalves, L., A. Subtil, M.R. Oliveira, V.E. Rosário, P. Lee, M.F. Shaio (2012). Bayesian latent class models in malaria diagnosis, PLoS ONE, 7, Nº7, e40633.

Pinto, A. E., Pires, A., Silva, G. L., Bicho, C., André, S. and Soares, J. (2011), Ploidy and S-phase fraction as predictive markers of response to radiotherapy in cervical cancer, PATHOLOGY RESEARCH AND PRACTICE, Vol.207, pp. 623-627.

Sepúlveda, N., Paulino, C.D. and Carneiro, J (2010). Estimation of T-cell repertoire diversity and clonal size distribution by Poisson abundance models. Journal of Immunological Methods, 353: 124-137

Silva, L., Turkman, A., Paulino, C.D., Bayesian statistical evaluation of peak area measurements in gamma spectrometry, APPLIED RADIATION AND ISOTOPES (2010), Vol. 68, pp. 1397-1402.

Soares, M.O. and Canto e Castro, L. (2012). Continuous time simulation and discretized models for cost-effectiveness analysis. Pharmacoeconomics 30:12, 1101-1117.

4. FUNDING 2008/2012

4.1

Description	2008	2009	2010	2011	2012	TOTAL (€)
Pluriannual Programme/Strategic project	168075	181500	176915	12271	175956	714717
FCT-funded projects	11880	19751	71904	12312	30971	146818
European Commission-funded projects	0	2143	6649	4897	2481	16170
Other international projects	0	0	0	0	0	0
Other national projects	7784	16614	11754	16332	18126	70610
National industry projects	0	0	0	0	0	0
International industry projects	0	0	0	0	0	0
TOTAL (€)	187739	220008	267222	45812	227534	948315

5. GENERAL INDICATORS 2008/2012

5.1

Description	2008	2009	2010	2011	2012	TOTAL ()
No. of researchers	97	100	102	103	104	506
No. of integrated researchers	45	50	53	53	56	257
No. of technicians and administrative staff	2	2	2	2	2	10
PhD theses under the supervision of integrated members	7	5	3	7	6	28
Publications in international peer reviewed journals	40	57	46	42	46	231
Books and chapters of international distribution	2	3	3	10	3	21
Models	0	0	0	0	0	0
Patents	0	0	0	0	0	0
Prototypes	0	0	0	0	0	0
Industry research contracts	0	0	0	0	0	0
Research contracts with national or international bodies	3	3	4	3	4	17

Software, computer code and algorithms	0	1	1	0	2	4
Books in portuguese	2	3	3	3	4	15
International Conference Proceedings	8	11	16	26	25	86
National Conference Proceedings	12	22	0	6	1	41
Encyclopedia entries	1	0	0	5	1	7
Organisation of Conferences, Workshops and Seminars	20	8	29	14	17	88
Organisation of Advanced Courses	3	4	2	2	2	13
Other National Projects	3	4	5	6	4	22

5.2 Overall description of indicators and research outputs/Highlights

The total number of researchers in the centre includes the Ph.D. students and collaborators with Ph.D, who are integrated members in other research centres, but with whom joint research is carried. There has been, as well, a policy to take as collaborators of the Centre previously integrated members, who, for some reason, stopped being scientifically active. Another policy of CEAUL has been to integrate the Ph.D. students as members, as soon as they finish their Ph.D., provided they have permission from the High Education Institution where they belong to.
Of the 28 PhD theses supervised, 19 were from students coming from the Ph.D. programme in Statistics and Operations Research of the University of Lisbon. Ph.D. students were mainly teachers coming from High Education Institutions.
There has been a steady flow of publications in international peer reviewed journals. The drop of the number of this type of publications in 2011 was compensated by an increase of publications, the same year, in chapters of books. Of the 227 publications in peer review journals, 160 are in the ISI web of knowledge. Publications in the area of Statistics account for 60% of the total. The other 40% were published in periodicals from other scientific areas (85% in the general areas of health, biology, ecology and genetics and 15% in the environmental sciences).
The research contracts referred to are (i) the European funded project and (ii) national projects funded by FCT with a member of CEAUL as Principal Investigator and FFCUL being the main Institution. The numbers indicated correspond to active projects in the correspondent year. Members of this R&D unit participated individually in other projects which were not contemplated here because FFCUL was not the main Institution. They are however referred to in the respective group outputs, section 9.4.19.
There are very few conferences in the area of Statistics which publish regularly the respective proceedings. As such, the number of publications in conference proceedings does not reflect the participation in conferences. Indeed participation in conferences and workshops was very high as it can be judged by the achievements reported by the group leaders in section 9.3.2.
At national level, the Portuguese Statistical Society organizes a national conference every year. Till 2009 a selection of refereed papers were published the following year in a book of proceedings. However, since 2009 the proceedings started being published by Springer, but the issue from 2009 and 2010 conferences only appeared in 2013. This justifies the fact that the number of papers in national proceedings dropped so drastically. Books published in Portuguese were scholarly editions.
Software and computer codes correspond to the publication in CRAN of three packages for R software and a computer code in R.
The full list of the activities referred to as conferences, workshops, seminars and advanced courses is in http://www.ceaul.fc.ul.pt/cs.html. They are relative to those organized directly by CEAUL. Participation as members of scientific committees or otherwise are not considered here.
In other national projects we refer to projects which were the result of contracts with external bodies, namely as consultancy, services to community, requested specialized courses, etc. See in section 4.1 the corresponding funding.

6. SCIENTIFIC COMPONENT - STRATEGIC PROGRAMME 2015/2020

6.1 Abstract in Portuguese for publication

O Centro de Estatística e Aplicações (CEAUL) está sediado no Departamento de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa, sendo esta a Instituição de Acolhimento, disponibilizando instalações e infra-estruturas necessárias à prossecução das suas actividades. A gestão orçamental está a cargo da Fundação da Faculdade de Ciências da Universidade de Lisboa (FFCUL) na qualidade de instituição principal.

Criado em 1975 por J. Tiago de Oliveira, o CEAUL desempenha um papel central em Portugal como o único centro dedicado exclusivamente à investigação metodológica nas grandes áreas de Probabilidade, Estatística e Processos Estocásticos e suas aplicações. É um centro com actividades multidisciplinares, onde se alia investigação fundamental à investigação aplicada em domínios tão diversos como Ciências da Vida e da Saúde, Ciências Naturais e do

Ambiente e Ciências Sociais (particularmente em Risco Financeiro). A política seguida no CEAUL tem sido a de desenvolver fortes ligações com outras áreas e domínios científicos. Neste sentido tem realizado frequentemente seminários, cursos, encontros e participado na divulgação da Estatística.

O CEAUL é membro da Comissão Nacional de Matemática e membro associado do Centro Internacional de Matemática. Tem ligações científicas com o Instituto Nacional de Estatística através da participação de alguns dos seus membros, na qualidade de editor-chefe, co-editor e editores associados, da revista internacional Revstat-Statistical Journal, indexada em bases de dados internacionais. Está ligado internacionalmente a várias Universidades prestigiadas no domínio da Estatística, através de colaboração científica e participação dos seus membros em redes e projectos de investigação.

Num sentido muito geral a Estatística é o estudo da incerteza presente em dados que resultam de fenómenos onde o mecanismo aleatório desempenha um papel especial. Os estatísticos são cientistas que desenvolvem instrumentos para medir este tipo de incerteza e quantificar os riscos associados à possível tomada de decisões em contextos de incerteza. Consequentemente a Estatística desempenha um papel primordial em problemas em que o ingrediente principal é a variabilidade inerente que conduz à incerteza associada. Como tal a Estatística não pode aparecer divorciada das outras ciências e do mundo real. As teorias e estudos dentro do domínio da Estatística devem ser dirigidos para o desenvolvimento de procedimentos cuja implementação contribua para o avanço das Ciências e tomada de decisões em face da incerteza.

Sendo uma disciplina relativamente jovem, a Estatística conheceu os seus maiores desenvolvimentos a partir de meados do século XX. No entanto, com o advento dos computadores e a grande capacidade actual de recolha e armazenamento de dados, não só o paradigma original da Estatística tem vindo a mudar rapidamente, como também o apoio do estatístico é cada vez mais necessário para dar uma resposta adequada a problemas suscitados em uma grande variedade de actividades de investigação e desenvolvimento científico. Em consequência, a estratégia desta unidade em Estatística e Aplicações é a de acompanhar os novos avanços no domínio científico da Estatística e continuar a fazer investigação fundamental nas suas áreas de excelência, desenvolvendo novas metodologias e técnicas de modo a poder transferir e estimular o uso dos conhecimentos mais actualizados na colaboração com investigadores de outros campos científicos.

O programa estratégico do CEAUL para 2015-2020 pauta-se, como consequência deste raciocínio e dos principais objectivos do CEAUL, pelas seguintes linhas:

1. Desenvolver investigação fundamental em diferentes áreas da Estatística, com o objectivo de expandir as fronteiras do conhecimento, adequando actualizados métodos e técnicas estatísticas às ciências experimentais e observacionais, na colaboração com investigadores de outros domínios e áreas científicas.
2. Integrar-se dinamicamente em projectos nacionais e internacionais de natureza interdisciplinar.
3. Participar activamente na sociedade, dando apoio científico no quadro de projectos relevantes ou na forma de cursos de actualização, na Indústria, Comércio, Serviços e Administração.
4. Promover actividades de formação avançada, através de programas doutorais e oferta de cursos de especialização em áreas emergentes da Estatística.

De modo a implementar adequadamente os seus objectivos, esta unidade de investigação está dividida em dois grupos de investigação e três linhas temáticas, transversais a esses grupos. Investigação fundamental é feita ao nível dos grupos em diversos temas de Estatística, tais como, mas não exclusivamente, Estatística de Extremos, Técnicas de Reamostragem, Teoria do Risco, Processos Estocásticos, Séries Temporais, Análise Espaço-Temporal, Estatística Bayesiana, Estatística Multivariada, Análise de Sobrevivência, Análise de Dados Incompletos, Gestão e Controlo da Qualidade. As três linhas temáticas escolhidas, Estatística Ambiental, Bioestatística e Risco em Extremos Financeiros, exprimem a experiência em aplicações e conhecimento fundamental acumulados pela unidade, nomeadamente nessas áreas. O objectivo destas linhas temáticas é dar apoio estatístico a investigadores de outras áreas científicas, quer a título individual, quer através da participação em projectos nacionais e internacionais, e fazer investigação aplicada através de trabalhos de consultoria e participação em projectos ligados à Indústria, Serviços e Administração Pública.

Para prosseguir os objectivos e implementar a estratégia proposta, o Centro de Estatística e Aplicações conta com 38 membros integrados, provenientes de diversas Universidades e Institutos Politécnicos de Norte a Sul do País, da Madeira e Açores, 34 membros colaboradores pertencentes a outras unidades de investigação das áreas de Matemática e de Ciências do Ambiente, da Vida e da Saúde, e 17 estudantes de doutoramento.

6.2 Abstract in English for evaluation

The Centre of Statistics and its Applications (CEAUL) is a research and development (R&D) unit connected to the Department of Statistics and Operations Research (DEIO) at the Faculty of Sciences of the University of Lisbon (FCUL). FCUL offers physical facilities to CEAUL, general computing infrastructure and communications, needed for the prosecution of its activities. The Foundation of FCUL (FFCUL), the main host institution, supplies administrative and financial R&D management.

Established in 1975 by Prof. J. Tiago de Oliveira, CEAUL plays a central role in Portugal as the only R&D unit dedicated exclusively to fundamental research in the areas of Probability, Statistics and Stochastic Processes. Besides fundamental research, CEAUL is also dedicated to interdisciplinary activities in domains as diverse as Life, Health,

Natural and Environmental Sciences and Financial Risk.

CEAUL is a member of CNM-Adhering Organization and Committee for Mathematics/ International Mathematical Union and is an associated member of CIM-International Centre of Mathematics. It has straight connections with the Portuguese National Institute of Statistics (INE), through the editorial support given by senior members of CEAUL to Revstat-Statistical Journal, an international journal in the area of Probability and Statistics, edited by INE and indexed by the ISI Web of Knowledge. At the international level CEAUL is connected to several prestigious universities through the scientific collaboration and participation of its members in networks and research activities.

In a very general sense Statistics may be viewed as the study of the uncertainty present in data arising from any phenomena and situations where chance mechanisms have a key action. Such study aims at describing the characteristics of interest in each practical situation and providing guidance for the making of decisions in contexts of unpredictability. Given the unavoidable randomness inherent to the vast majority of real world phenomena, statistical thinking today infuses so many scientific fields that it plays a crucial backdrop to the advance of the Science in general. Statistical theory and methodology cannot exist divorced from the needs of other sciences felt by the corresponding researchers and users.

Statistics, as a scientific discipline, is relatively young with its foundations and main developments dating back to the XX century. However, with the advances of technology and consequent capacity for acquisition and data warehouse, the paradigm of Statistics has been changing in a fast speed and statistical support is more and more required to give adequate answers to problems posed in a great variety of research activities. Consequently the strategy of this research unit is to keep pace with the recent advances in the scientific domain of Statistics, to continue fundamental research in its areas of excellence, developing new methodologies and tools with the goal of communicating and using the most up-to-date knowledge, while collaborating with researchers from other scientific areas.

As a consequence, the strategic program of CEAUL for 2015-2020 aims to achieve the following objectives:

1. Develop fundamental research on diverse domains of Statistics, with the aim of extending the frontiers of statistical knowledge and provide statistical tools for experimental sciences, while collaborating with researchers from other domains and scientific areas.
2. Participate in national and international projects of an interdisciplinary nature.
3. Participate actively in Society by giving support to Industry, Commerce, Services and Business in the form of consultancy or specialized courses.
4. Develop advanced training activities through the involvement in Ph.D. programs of High Education Institutions and the organization of advanced courses in emerging areas of Statistics.

In order to attain the strategic objectives, the R&D unit adopted an organizational framework based on two groups, namely (1) Order Statistics, Extremes, Modelling and Simulation and (2) Statistical Models and Methods for Complex Systems, and three thematic lines horizontal to the groups.

Fundamental research will be carried out within the research groups, in a diversified range of themes including but not restricted to, Statistics of Extremes, Computational Statistics and Simulation, Resampling Methodologies, Quality Control, Bayesian Statistics, Biostatistics, Environmental Statistics, Statistics in Genomics Immunology and Epidemiology, Statistics in Medicine, Longitudinal Models and Survival Analysis, Time Series Analysis, Incomplete Data Analysis, Multivariate Analysis, Temporal and Spatial Statistics, Design of Experiments, History and Teaching of Statistics. Although the two groups work on different themes and areas of Statistics, they share similar interests in terms of areas of application. The three thematic lines, Environmental Statistics, Biostatistics and Risk in Financial Extremes, reflect the scientific domains and areas where the research carried by the unit has more impact. The objective of these thematic lines is to give statistical support to researchers coming from other fields of science. This collaborative work is done either at an individual level, or through the participation in national and international projects. Knowledge transfer, in terms of consultancy work, participation in projects coming from the Industry, Services and Public Administration, is also a strategic objective of the thematic lines.

To carry out the objectives set and to implement the proposed strategy, CEAUL counts with (i) thirty eight integrated members, coming from a range of Universities and Polytechnic Institutes from North to South of Portugal, Madeira and Azores, (ii) thirty four collaborators who are integrated members in other research units from the area of Mathematics, Environmental and Health Sciences, and seventeen PhD students.

6.3 Strategy and vision of the unit and future management

In a very general sense, Statistics is the study of the uncertainty present in data arising from any phenomena where chance mechanism plays a special role. Statisticians are the scientists who develop tools to measure and quantify risks associated with possible decisions. As a consequence Statistics plays a greater role in topics that have variability, giving rise to uncertainty, as an essential ingredient. Statisticians are not only theoreticians but above all they are practitioners. Statisticians cannot exist alone divorced from other sciences and from the real world. Their theories and studies should be directed towards the development of statistical procedures whose implementation should improve decision-making under uncertainty. Hence statisticians should closely collaborate with the scientists who encounter this type of situations in their field of research or study. The benefit can be two fold. Most commonly statisticians help researchers from other fields using known existing tools to solve their problems. However, very often, the problems statisticians encounter in other fields are very challenging in the sense that they raise questions for which the statistician has not yet direct or known answer. These problems feed advanced research in Statistics and consequently they are essential for devising new and more sophisticated statistical methods and tools, leading to fundamental research.

Quoting the October 19th 2013 edition of the Economist entitled How Science Goes Wrong, - Scientists' grasp of Statistics has not kept pace with the development of complex techniques for crunching data. Some scientists use inappropriate techniques because those are the ones they feel comfortable with; others latch on to new ones without understanding their subtleties. Some just rely on the methods built into their software, even if they do not understand them - Thus, the main vision and strategy of this unit in Statistics and Applications is to follow closely new advances in the area as well as to develop further techniques in diverse fields of Statistics and communicate them to scientists with the objective of helping them in improving their scientific findings. Fundamental research, on the main areas of knowledge, has a very important role to play in this vision and strategy, in order to keep up with new demands and needs coming from other scientific fields. Just to give an example: today, dealing with Massive Data is one of the key issues of statistical research, as a response to the needs of many scientific areas; techniques developed for small or moderate-sized datasets simply do not translate to modern massive data sets which are a direct consequence of the new technologies in monitoring and gathering data.

CEAUL is a very well established research group in Statistics with a very good record of accumulated knowledge on several areas of Statistics, such as Statistics of Extremes, Statistical Quality Control, Time Series, Spatial Statistics, Bayesian Statistics, Environmental Statistics and Biostatistics, to be able to implement this vision and strategy.
In order to implement this vision and strategy the unit is organized in two complementary research groups, classified according to statistical methodologies on which fundamental research will be carried out, and three thematic strands according to different fields of applications. The dominant topics of group 1 are on extreme value theory, modelling and assessment of risk, resampling methodologies and computational statistics; those of group 2 are on Bayesian Statistics, spatio-temporal analysis, incomplete data analysis and epidemiological processes. Thematic strands are transversal to the research groups and are based on the expertise of the R&D unit on the diverse field of applications, namely on Environmental Statistics, Biostatistics and Risk in Financial Extremes. Fundamental research and methodological support to the thematic strands are developed inside and across the research groups. In general lines, the objective of the thematic strands is to give statistical support to researchers in other scientific fields of research. This will be done on three different levels:
Level 1 - individual collaboration through the application of statistical methods to solve problems presented by individual researchers coming from other fields;
Level 2 - collaboration in joint national and/or international projects with other scientific units seeking statistical solutions; such level of collaboration will be extended over longer periods of time and will need full cooperation of the research groups involved, making it more demanding than level 1 type collaboration;
Level3 - Knowledge transfer of economic and social nature, having a direct impact on society and economy at regional or national level.

Each research group as well as each thematic strand will be coordinated by a senior researcher. The coordinators of research groups are responsible for defining the scientific policies and research strategies of the group. The coordinators of the thematic strand are responsible in seeking collaborations, defining working strategies, organizing the activities and communicating with the group leaders, guaranteeing that projects are accomplished timely.
The success of the implementation of the vision and strategy of the unit depends heavily on the ability of the coordinators of the research and thematic groups to communicate and coordinate their activities in an optimal manner. This is the reason why they were chosen among the most senior members of the R&D unit and who have a long experience in the coordination of research activities.
The unit will be run by a Director (the unit coordinator) presiding a scientific committee composed by the coordinators and two representatives of the research and thematic groups. This scientific committee is responsible for steering, verifying the activities of the unit, aided by an external advisory board. The administrative part of the unit will be run by the Director, a Vice-Director appointed by the scientific committee and a Secretary. On a regular basis the groups and thematic strands will organize meetings to discuss scientific progresses achieved and there will be, each year, a general meeting of the Centre with all members, to discuss the scientific, as well as the administrative aspects of the unit, with the objective of being accountable to its members.

6.4 Laboratory intensity level of the unit

Low or absent

Laboratory intensity level Justification

Due to the nature of research we do, the level of laboratory intensity is low, although computation and simulation are a very important part of the research to be developed. We have to deal with large data sets and to work with very demanding software and hence, not only each member should have its own up-to-date personal computer, but also as a group we may need workstations or access to main frame computers, depending on the nature of problems at hand. However, free access to R programming and availability of vast amount of R packages reduces the need for acquiring expensive software.

6.5 General objectives

The general objectives can be resumed in four different, but interconnected categories:

1. Fundamental research on statistical methods and models with the objective of extending the frontiers of statistical knowledge and provide statistical tools for experimental sciences.

2. Collaborative research with scientists coming from other fields with the objective of applying statistical methods, models and tools needed in improving their findings in their respective fields of science.

3. Active participation in Society by giving support to Industry, Commerce, Services and Business in the form of consultancy or specialized courses.

4. Participation in the activities of High Education Institutions by the transfer of scientific knowledge through Ph.D. programs and advanced courses.

The detailed description of the general objectives is as follows:

1. Fundamental Research

Fundamental research activities of this unit can be classified into short to medium and long-term interests.

Short to medium term research activities correspond to fields of statistics, probability and stochastic processes in which this unit has been actively making contributions, namely in Statistics of Extremes, Computational Statistics and Simulation, Resampling Methodologies, Quality Control, Bayesian Statistics, Biostatistics, Environmental Statistics, Statistics in Genomics, Immunology and Epidemiology, Statistics in Medicine, Longitudinal Models and Survival Analysis, Time Series Analysis, Categorical Data Analysis and Missing Data, Sampling, Principal Component Analysis, Classification Methods, Temporal and Spatial Statistics, Design of Experiments and History and Teaching of Statistics. Detailed description of identified problems within these general fields will be given in more detail as objectives of the two groups of research.

Long-term fundamental research activities will correspond to recent and, yet to come, trends in Statistical Sciences and problems that are identified as important resulting from the collaborative research that the unit will carry with other scientists. Traditionally the paradigm of statistics has been to extract information from hard to come by, expensive and consequently small data sets. This paradigm, with the changes in instrumentation and technology in general, has been changing rapidly towards methods to deal with massive or big data. Together with these changes, statistical methods and models, as well as interface between statistics and computation has been advancing in a very fast pace. The new paradigm in Statistics can be described as Large-Scale Statistical Science. The effect of these can be seen in many fields from Genomics, Environmental Sciences, Economy and Health. We can give many examples how needs in other Sciences influence the new trends in fundamental research in Statistics. Just to mention a few, we see application of Statistics in Neuroimaging, which needs fundamental research on Functional Data Analysis. Genomic Applications and Statistical Issues, such as Millimeter-long DNA Molecules, Next Generation Sequencing, need specific statistical methodologies to deal, for instance, with multiple testing procedures for hierarchically dependent hypothesis and take conclusions. In Environmental Sciences, there is the need to analyse data obtained at very high resolutions. Corresponding models have complex dependent structures needing very sophisticated computational needs. In Finances, data appear in continuous time, which needs non-linear continuous time series models. The research unit intends to keep up with such fast changes in other sciences and consequent need for new statistical methods. The objective will be to keep up with these changes to bring the new advances into collaborative research and contribute in fundamental advances.

2. Collaborative research

Collaborative research will be basically in the areas where there is a very good accumulated knowledge and experience, namely in Environmental problems, Health and Biological applications. However the unit is open to collaborative research in other scientific areas. For example, although the unit has a lot of accumulated knowledge on Extreme Value Statistics and Modeling of Rare Events fundamental in the areas of Finance, there has not been much collaborative work on these areas. The unit is signaling its willingness to open up to new collaborative research by having a new thematic strand on Risk in Financial Extremes with this objective in mind. The short-medium and long-term objectives of different collaborative research are described in the respective thematic strands.

3. Participation in Society

This research unit has good potential in technology and knowledge transfer to Industry, Services and Business through actions such has consultancy and specialized courses. However this potential so far has not been fully put in use although there have been some very interesting results in the area as described before in the achievements. The unit has always chosen, as priority, fundamental and collaborative research. However in coming years the unit intends to invest more in this objective. This is also facilitated by the fact that there is a general recognition among the Industry, Services and Business that statistical methods can improve their performances. Recent consultancy work carried out for EDP (Portuguese Electricity Board) in producing risk maps for disruption in power lines, is one good example. We intend to make this objective as one of the top priorities of the unit. The unit chooses the Economics of Health one of the priority areas of consultancy.

4 Participation in activities of High Education Institutions

CEAUL is one of the driving forces of the Ph.D. program in Statistics and Operations Research of the University of Lisbon. However this transfer of knowledge is being diversified by collaborating in other Ph.D. programs such as the Ph.D. program in Earth Sciences, of FCUL, BEGC - Biology and Ecology of Global Change, University of Aveiro, BioSys - PhD Program in Biological Systems, Functional & Integrative Genomics, from FCUL and also collaborating with the research units of CEMAT (IST) and CEMAPRE (ISEG) with the objective of creating a new Ph.D. program in Statistics. This will be given in more detail in the section opportunities for advanced training. This R&D unit will continue offering advanced courses on emerging areas of Statistics open for researchers and advanced students from all over the country, as well as courses on more general areas of Statistics, for researchers and Ph.D. students who want to learn more about some specific subject.

6.6 Implementation

Taking into account the diversity of scientific areas and subareas involved in the research programs proposed, it was considered that the most efficient way to attain the strategic objectives of the R&D unit would be to adopt an organizational framework based on two groups and three thematic lines horizontal to the groups.

Fundamental research on the areas of expertise and interest of this R&D unit, will be carried out within the research groups, namely (i) Order Statistics, Extremes, Modelling and Simulation (RG-6-70) and (ii) Statistical Models and Methods for Complex Systems(RG-6-74). Each group has a list of collaborators and PhD students who are supervised by the group members. Collaborators are in general integrated members of other research units in Mathematics or other scientific domains, with whom a strong scientific collaboration already exists. The connection between integrated members and collaborators is done at individual level. PhD students are supposed to meet on a regular basis with their supervisors who are responsible for their research activities. Although the two groups work on different themes and areas of Statistics, they share similar interests in terms of areas of application. They also have research points of contact at a methodological level, such as statistics of extremes, planning of experiments, simulation and computation, multivariate analysis, to mention a few. These points of contact are particularly important for the development of the activities within the thematic strands. The three thematic lines, Environmental Statistics (TL-6-42), Biostatistics (TL-6-1194) and Risk in Financial Extremes (TL-6-1195), reflect, simultaneously, the scientific domains and areas where the research of the unit has more impact, and the interest and previous experience of the members of this R&D unit in what concerns collaboration with researchers from other fields. Collaboration with other research units, participation in national and international interdisciplinary projects, consultancy and general activities will be done within the thematic lines. The thematic lines on Environmental Statistics and Biostatistics are horizontal to both research groups and are a result of a well-established experience in collaborative work with researchers from other fields of science. The thematic line on Extremes in Financial Risk is a recent addition and collaboration will be more on methodological aspects. It will be basically within the competences of the research group on extremes (RG-6-70), although it is expected that members of the research group RG-06-74, who work on max stable processes and nonlinear time series, will be involved as well. Hence, this thematic line has different general objectives from the other two.

Some of the research themes which are fundamental for the prosecution of the activities of the thematic strand on Environmental Statistics are: extreme value statistics in quantification of risk of rare events (RG-6-70); max-stable processes to deal with spatial extremes(RG-6-70 and RG-6-74); spatio-temporal and Bayesian hierarchical modelling (RG-6-74), to deal with the evolution of events in space and time; statistical screening methodologies (RG-6-74) to deal with the setting of environmental standards and taking conclusions from causal relationships; state-space models (RG-6-74) in managing fish stocks. These activities are inserted within projects and contracts with other national research units and /or inserted into international projects of multidisciplinary nature. The coordinator of the thematic strand has the responsibility to organize the activities, to communicate with the project partners and to find the members of the research groups who will carry out the project work. There are ongoing collaborations which will continue, such as joint work with the Centre of Forest Research and Institute D. Luiz, the collaboration within the network - Against Wildfires, a Multidisciplinary Approach - , with partners from Spain and other Mediterranean countries; collaboration with the IPMA in fisheries and other future collaborations that will result from actions already taken, or to be taken towards joint ventures.

This R&D unit has already a lot of experience in collaborative work in the area of Biostatistics; indeed around 35% of the publications of the R&D unit during 2008-2012 were a result of collaborative work in this area. General research themes which are fundamental for the prosecution of the activities of the thematic strand on Biostatistics are: spatio-temporal models (RG-6-74, RG-6-70); categorical data analysis (RG-6-74); survival and longitudinal models (RG-6-74); extreme events related to censored data (RG-6-70); incomplete data analysis (RG-6-74); non-parametric Bayesian analysis (RG-6-74); generalized linear mixed models (RG-6-74); epidemic models and processes (RG-6-74); dynamical analysis of population growth models (RG-6-70); experimental design (RG-6-74, RG-6-70), multivariate analysis (RG-6-74, RG-6-70), forensic statistics (RG-6-70). There are quite a number of ongoing projects which will continue as activities of this thematic strand and they are all described in section 10 reserved for the thematic strands. This thematic strand will be responsible for several of the activities on epidemiology and health contributing towards regional strategy. At the international level it is important to refer to, among others, the involvement in (i) the INLA network on survival and longitudinal joint models, with members of the statistics group of the Centre of Mathematics at the University of Minho and the INLA team coordinated by Harvard Rue from the Norwegian University of Science and Technology; (ii) PERSSILAA FP7 project -PERsonalised ICT Supported Service for Independent Living and Active Ageing. Again the coordinator of the thematic strand on Biostatistics has the responsibility to organize the activities, to communicate with the project partners and to find the members of the research groups who will carry out the project work.

The proposal of a thematic strand on Risk in Financial Extremes was a natural consequence of the recognition that in fields such as in insurance and finance, the stochastic modelling of extremes is of major importance. The group RG-06-70 is internationally recognized in the area of Extremes and has contributed with fundamental research on themes which have strong impact on insurance and finance. The focus of this thematic strand is aimed to build a bridge between Extreme Values and Applications in Finance, Economics and Insurance, under the EVT framework. At the national level the partners come from CEMAPRE - the Centre for Applied Mathematics and Economics and Centre of Mathematics and Applications (UNL); international partners are from Drexel University, Katholieke Universiteit Leuven, Erasmus University Rotterdam, Universidad Complutense de Madrid. The coordinator of the thematic strand will be responsible for organizing the activities and communicate with the national and international partners.

Typically problems addressed within the thematic strands will be published in periodicals related to the fields of application. The research produced by the groups is typically of a fundamental nature and will be published in periodicals in the area of Probability, Statistics, Stochastic Processes and related fields. Problems brought to the thematic strands by the collaborative work with researchers from other fields, may as well give rise to publications in the area of Statistics, like it was the case during the period 2008-2012. On the other hand, outcomes of statistical analyses performed by members of the unit within the scope of joint works with external entities, may be deciding factors to a possible social or economic impact of novel operative procedures, treatment forms or protocols, intervention means, etc. for the field at issue, which in itself is an outstanding performance indicator for the overall team.

The expected indicators were calculated assuming that conditions to carry out research (financing, lecturing, etc) stay stationary during the coming 6 years. We also assume a very conservative hypothesis that the 38 integrated members, who compose the present team, will continue publishing at a similar pace as before. The same rationale applies to other indicators like books, chapter of books and proceedings, although with conference proceedings this reasoning is not so linear. Indeed, in the area of Statistics there are few conferences which publish regularly proceedings and those conferences are not always the most appropriate for the adequate dissemination of the results of the team.

Relatively to research contracts our estimate includes projects where integrated members will eventually participate, even if the contracts signed are not with the host institution. As it is referred to along the description of the thematic lines, we have already some ongoing projects which will carry beyond 2015. Contracts with Industry are likely to happen particularly in the area of quality control with The Portuguese Institute of Quality, Auto Europa, Corticeira Amorim; we also expect contracts with services, consultancy companies, etc. We include all these in the contracts with Industry.
The estimates for the indicators relatively to research contracts and contracts with Industry are equally spread along the years, as an overall general indicator. Research contracts are likely to carry over on average for three years and hence a research contract counts for the indicator as many times as its duration in years.
Relatively to PhD students, of the present 17 students, 3 are expected to finish during 2014. The other 14 will finish the latest by 2017. On average we expect to recruit 3 PhD students each year and we expect to hire 3 PhD students to work specifically on the objectives of the thematic lines (this is explained in section 12).

6.7 Contributions for the regional strategy

The CEAUL is located in the Region of Lisboa e Vale do Tejo (RLVT). The strategy of this region is not yet public. Therefore, we took as reference the analysis published in September 2013 by the CCRLVT (Coordinating Commission of the Region of Lisbon and Tagus Valley), entitled - Plano de Ação Regional de Lisboa 2014-2020, Diagnóstico Prospetivo -. The SWOT analysis published in this document was considered in CEAUL's strategic planning,

Regarding the Intelligent Growth and considering that the region has growth potential in the Health and Marine sectors, the CEAUL will namely contribute to reinforce the expertise of the region in these areas by means of the following actions.

Epidemiology: Several cancer research centres of excellence are located in RLVT, and statistical methodologies in cancer epidemiology are important to predict the incidence of cancer. This is a difficult problem associated with uncertainty, due to time variability of the observations. CEAUL's members do fundamental research and collaborative work in this particular area. There is an ongoing project of collaboration between CEAUL and Portuguese Southern Cancer Registry (ROR Sul), that started two years ago and will continue due to its high importance for the southern Portugal.

Health: Quality of services in the health care sector is a major concern of the managers and health professionals. CEAUL is cooperating with the Commission on Quality and Patient Safety Centre of Centro Hospitalar de Lisboa Central, EPE(CHLC), to evaluate the user satisfaction of CHLC patients, in a systematic approach that is integrated in CHLC's Continuous Improvement Program. CEAUL will have the responsibility to construct the questionnaires and analysis of surveys of customer satisfaction, to be applied in several public hospital wards chosen by CHLC.

Marine: CEAUL has a close collaboration with the Portuguese Institute of Marine and Atmospheric Sciences (IPMA). I. Figueiredo (new integrated member and a researcher at that Institute) has a long-term collaboration with Lucilia Carvalho, namely studying the black scabbard fish (DEEPFISHMANFP7 Project). In 2015-2020 this collaboration will continue with studies on the modelling of the dynamics of the black scabbard fish in NE Atlantic. A PROMAR project to evaluate the stock levels of the ray fishery, and to study of technical measures for its management is also planned.

The actions proposed will also contribute to overcome major challenges in RLVT, namely the increase of cooperation between companies and research centres, and the increase in the efficiency of technology and knowledge transfer from the academia to the economy.

6.8 Opportunities for advanced training

Members of CEAUL come from different Universities and participate in PhD programs in Mathematics/Statistics of their respective Universities. CEAUL has been one of the driving forces of the Ph.D. program in Statistics and Operations Research of the University of Lisbon (DEIO/FCUL). The recent merge of this University with the Technical University of Lisbon (UTL), created an opportunity for the proposal of a Ph.D. program in Statistics and Applied Mathematics by joining DEIO's PhD program with the existing PhD programs in the different departments of Mathematics of UTL, namely DM/IST, DM/ISEG and ISA. Some members of CEAUL belong to the Portuguese Open University which offers, as well, a PhD program in Mathematics with specialization in areas of Statistics.

Members of CEAUL also collaborate in Ph.D. programs financed by FCT in other scientific areas, such as the Ph.D. program in Earth Sciences (FCUL), in Biology and Ecology of Global Change (University of Aveiro), in Biological Systems, Functional & Integrative Genomics (FCUL). This is an opportunity for members of CEAUL who are involved in those programs to offer PhD themes and act also as co-supervisors.

Another opportunity is with the PhD program of The Chronic Diseases Research Center (CEDOC) from Nova Medical School in Health Sciences (NMSHS). A protocol of collaboration between CEAUL and CEDOC allows CEAUL members to act as co-supervisors of PhD thesis when substantial advice is offered. This medical school offers also a PhD program in Biostatistics and CEAUL members from that school have the opportunity of supervising theses on that area.

Recruitment for Academy has been drastically reduced in Portugal and Industry and Services are not yet fully aware of the benefits of hiring highly qualified personnel with a Ph.D. in Statistics. The lack of interest within the students from Portugal to follow a Ph.D. program, has been observed in recent years with a consequent drastic reduction of PhD candidates in Statistics, since prospects of finding a job after a PhD, particularly in academic career, are feeble. However there are opportunities from abroad. We mention the ISA/UL protocol with Cape Verde and Angola towards advanced studies in Statistics in which members of CEAUL are collaborating. Also we are setting up collaboration with researchers from Brazilian federal universities towards joint research and advanced training/supervision of students in postgraduate studies. The creation of the Iberian Region of the International Biometric Society, in which members of CEAUL are deeply involved, will also open an opportunity for joint advanced training with Spanish biometricians. Advertisement of the different PhD programs is also a priority so that we capture international students.

CEAUL will keep also its tradition of organizing advanced courses on emerging areas of Statistics for researchers and advanced students from all over the country.

6.9 Internationalization

Internationalization is a fundamental activity for CEAUL team. It is intended that researchers participate in International Seminars, Workshops, Conferences, Short Courses and other international meetings. The goal is to keep up-to-date with scientific advances, present the work to international peers, and interact with expert researchers for scientific collaboration. Organization of international meetings in Portugal will also be one of the goals. As a result, it is expected a good portfolio of collaborative publications in peer-reviewed International Journals and other kind of international disseminationof the main results.

We mention some of formal and informal networks and initiatives that members of CEAUL are connected with and are relevant for the strategic plan.
• Strengthening International Cooperation of the KLIMATEXT Research Team. CZ.1.07./2.3.00/20.0086. PI: Jan Picek (members of group RG-6-70).
• European Network for Model-driven Investigation of Environmental Processes. Informal network led by A Gelfand (members of group RG-6-74).
• Network CEAUL/ ELSA-Brasil in Rio de Janeiro (at Fundação Oswaldo Cruz) (members of group RG-6-74).
• BIOSTATNET an interdisciplinary Biostatistics network (members of group RG-6-74).
• Network: Against Wildfires, a Multidisciplinary Approach (members of group RG-6-74).
• INLA project on Survival and longitudinal joint models: a network CEAUL/CMAT/NUST, coordinated by Havard Rue (members of group RG-6-74).
• PERsonalised ICT Supported Service for Independent Living and Active Ageing. Europeanproject (FP7-ICT-2013-10); KF Turkman, MA Turkman, M Antunes (started 10/2013)
• Joint work with the International Council for the Exploration of the Sea and the Institut Français de Recherche pour l'Exploitation de la Mer to model the dynamics of the black scabbard fish in NE Atlantic. (L Carvalho, I Figueiredo)
• Mesoscale Weather Extremes: Theory, Spatial Modeling and Prediction; with M.Schlather as coordinator. Participation of A Ferreira and L de Haan, starting 2014.

International collaborators include among many others: J A Achcar, A Al-Hassan, TA Alonzo, C Anderson, A Azzalini, J Beirlant, A Buishand, J Cai, MS Carvalho, D Chor, CB Dean, H Drees, J Einmahl, E Faerstein, W González-Manteiga, S Hammoudeh, J Á Jiménez, IT Jolliffe, Z Kabluchko, R Leadbetter, V. Leiva, M McAleer, E Z Martinez, S Mejza, G Molenberghs, P Mueller, T Niyonsenga, L Peng, T Pérez Amaral, F Poleto, A Ruffino-Netto, M Schlather, J Singer, M Swash, A Vanasse, CG Vries, E Ursu, C. Zhou.

Some members of this R&D unit are or have recently been members of editorial boards of international periodicals, such as Extremes, Journal of Statistical Planning and Inference, Journal of Statistical Theory and Practice, Revstat Statistical Journal, Chaotic Modeling and Simulation, Chilean Journal of Statistics. Ivette Gomes is a member of the ISI Nominations Committee 2013-2015.

Knowledge transfer

This R&D unit has accumulated knowledge on issues that can be used in benefit of Institutions, Industry, Hospitals and Services.

The applied research undertaken is mainly oriented to support research activities in other scientific domains. This is done through the involvement in research networks, Institutions, Hospitals, etc. Social and economic benefits are expected as a byproduct of our collaboration.The structure and objectives of the thematic strands on Environmental Statistics and Biostatistics are partly directed towards this goal.

We refer to some of the initiatives towards knowledge transfer:

A) Some wildfires go out of control causing large damages. Therefore, one is often concerned to forecast the onset of a possible extreme wildfire and its subsequent evolution both in space and time. We propose to construct an on-line alarm system to predict the event of extreme fires. Such alarm system, once implemented can be a valuable tool in the organization of fire fighting capacity and allocation of resources.

B) Estimation of fish stocks, its variation in space and time, is an important decision support tool for predicting and harnessing fish stocks in a sustainable manner. This R&D unit will continue collaboration with IPMA on these issues.

C) In seeking to control pollution, it is necessary to impose requirements on the levels of pollution in an effort to limit its effects, by achieving an optimal balance between a relevant measure of benefit and an appropriate assessment of cost in the resulting regulatory process. This objective requires probabilistic assessment of uncertainty and statistical methods to infer on the causal effect of pollutants on population and environment. Our proposed contribution is to assess the existing environmental standards on water pollution and help setting standards on new class of pollutants.

D) Patient satisfaction evaluation originates important indicators to the implementation of strategies for continuous quality improvement of health care. Evaluation of satisfaction as an indicator of quality of health services is part of a protocol CEAUL/ Commission on Quality and Patient Safety Centre of Centro Hospitalar de Lisboa Central.

E) Sharing knowledge for the joint development of technical-scientific services to the health sector is part of the protocol between Exigo/CEAUL via FFCUL. Activities programmed are (i) lectures, short exercises and workshops tailored to the needs of individual clients. For non-statisticians the aim is to increase knowledge and understanding of statistics to enable participants to communicate with statisticians through a common language and deal more effectively with statistical issues. (ii) Initiate the establishment of a Bayesian approach to Health Technology Assessment (HTA) as a comprehensive bridging methodology between HTA to Multiple criteria decision making while providing a powerful framework for determining optimal decision behaviour in face of uncertainty.

Ethical issues

This unit will follow the general ethical guidelines applicable to all sciences and in particular will follow closely the American Statistical Association (ASA)'s Ethical Guidelines for Statistical Practice. We quote ASA on this issue:

- Statistical tools and methods, as with many other technologies, can be employed either for social good or evil. The professionalism encouraged by these guidelines is predicated on their use in socially responsible pursuits by morally responsible societies, governments, and employers. Where the end purpose of a statistical application is itself morally reprehensible, statistical professionalism ceases to have ethical worth -.

The complete set of guidelines can be found in http://www.amstat.org/about/ethicalguidelines.cfm

Since we collaborate with other scientists from Engineers to Medical Doctors, whose results and research findings have direct consequences on the well-being of the Society, we give extreme care in statistical analysis to make sure that conclusions taken from planned experiments and consequent data analysis will reflect the true state of the nature.

7. EXPECTED INDICATORS OF THE STRATEGIC PROGRAMME 2015/2020

7.1

Description	2015	2016	2017	2018	2019	2020	TOTAL ()
Publications in peer-review international journals	31	31	31	31	31	31	186
Patents and performing patents	0	0	0	0	0	0	0
Books and chapters of international distribution	3	3	3	3	3	3	18
PhD theses under the supervision of integrated members	4	5	5	3	3	6	26
Conference proceedings	12	12	12	12	12	12	72
New materials, devices, products and processes, software, computer codes and algorithms	1	1	1	1	1	1	6
Books, including single-authored works (including scholarly editions of oral or written texts and translations with introduction and commentary)	2	2	2	2	2	2	12

Edited special issues of journals, with substantial research input on the part of the researcher	0	1	0	1	0	1	3
Chapters in books, including contributions to conference proceedings, essays in collections	15	15	15	15	15	15	90
Creative writing (to the extent that it embodies research)	0	0	0	0	0	0	0
Dictionary entries (forming part of investigation)	0	0	0	0	0	0	0
Encyclopedia entries (to the extent that they embody research)	0	0	0	0	0	0	0
Audio/visual and electronic/digital materials	0	0	0	0	0	0	0
Other categories, including web-based resources; video and audio recordings (to the extent that they embody research)	0	0	0	0	0	0	0
Performances and exhibitions to the extent that they embody research	0	0	0	0	0	0	0
Industrial research contracts	2	2	2	2	2	2	12
Research contracts with national or international bodies	7	7	7	7	7	7	42

8. PROPOSED RESEARCH TEAM

8.1 Criteria adopted by the R&D unit for the definition of integrated member, if different from FCT reference table

The criteria adopted by the R&D unit for the definition of integrated member was the one proposed by FCT for the area of Mathematics, namely four outputs during the last five years period, (refereed papers in international periodicals, refereed papers in proceedings of conferences, software published in The Comprehensive R Archive Network, chapters of books). Two membersdid not fulfill in total these criteria, but the justification for including them is as follows:
Ana Ferreira: one of her publications is in an encyclopedia entry; however, her knowledge and capacity is important for the centre. She is known to be an excellent researcher with international recognition; apart from being a co-author of a book on Extreme Value Theory: An Introduction (417 pp.) from Springer, one of her papers is to be published in Bernoulli in 2013 on The generalized Pareto process; with a view towards application and simulation (http://arxiv.org/abs/1203.2551) and another on Spatial aggregation and high quantile estimation applied to extreme precipitation (http://homes.stat.unipd.it/mgri/SIS2010/Program/13-SSXIII_SPS/899-1555-1-RV.pdf) is under review. She was invited by Martin Schlatter to participate on the Project: MODEXTREME - Modelling vegetation response to extreme events, from the EU FP7 program.
Salomé Cabral: one of the publications is from 2013, but it was published on line during 2012; hence we accepted her publications for the period 2009-2013. She is also an important element for the team because of her knowledge on count longitudinal data, which is one of the methodologies important for the thematic strand on Biostatistics. She has inclusive another R package, already in CRAN entitled cold: A Package for Count Longitudinal Data, Version 1.0-3. URL: http://CRAN.R-project.org/package=cold .

8.2 List of Integrated Members / 10 nuclear CVs

Name	Nuclear CV
Maria Antonia Conceição Abrantes Amaral Turkman (Coordinator)	Yes
Amílcar Manuel do Rosário Oliveira	No
Ana Luisa Trigoso Papoila da Silva	No
Ana Maria Santos Ferreira Gorjão Henriques	No
Carina Soares da Silva Fortes	No
Carlos Daniel Mimoso Paulino	Yes
Clara Maria Henrique Cordeiro	No
Cristiana Maria Palmela Pereira	No
Dinis Duarte Ferreira Pestana	Yes
Fernanda Otilia de Sousa Figueiredo	No
Giovani Loiola da Silva	No
Ivone Maria Ribeiro Figueiredo da Silva Rosa (CO-ORIENTADORA)	No
Jorge Filipe Campinos Landerset Cadima	No

José Leonel Linhares da Rocha	No
João Paulo Oliveira Martins	No
Kamil Feridun Turkman	Yes
Laurentius Franciscus Maria de Haan	Yes
Ligia Carla Pinto Henriques Jorge Rodrigues	Yes
Lisete Maria Ribeiro de Sousa	No
Luisa da Conceicao dos Santos do Canto e Castro de Loura	Yes
Luzia Augusta Pires Gonçalves	Yes
MARIA ISABEL FRAGA ALVES	Yes
MARIA IVETTE LEAL DE CARVALHO GOMES	Yes
Manjunath, Bangalore Govinda Raju	No
Maria Lucília Salema e Carvalho	No
Maria Manuela Costa Neves Figueiredo	No
Maria Salomé Esteves Cabral	No
Maria de Fátima Almeida Brilhante	No
Marília Cristina de Sousa Antunes	No
Miguel Martins Felgueiras	No
Patricia Cortes Zea Bermudez	No
Paula Cristina Martins dos Reis	No
Paulo Jose Araujo dos Santos	No
Sandra Cristina de Faria Ramos	No
Sandra Maria Freitas Mendonça	No
Sandra Maria da Silva Figueiredo Aleixo	No
Teresa Paula Costa Azinheira Oliveira	No
Valeska Lima Andreozzi	No

9. PROPOSED RESEARCH GROUPS

Reference Name		Principal Investigator
RG-6-70	Order Statistics, Extremes, Modelling and Simulation	MARIA IVETTE LEAL DE CARVALHO GOMES
RG-6-74	Statistical Models and Methods for Complex Systems	Maria Antonia Conceição Abrantes Amaral Turkman

(RG-6-70) Order Statistics, Extremes, Modelling and Simulation

9.1. IDENTIFICATION OF THE RESEARCH GROUP

9.1.1 Reference of the research group

RG-6-70

9.1.2 Name of the Research Group in portuguese

Estatísticas Ordinais, Extremos Modelação e Simulação

9.1.3 Name of the Research Group in English

Order Statistics, Extremes, Modelling and Simulation

9.1.4 Keyword(s)

Extreme Value Theory

Modelling and Risk

Computational Statistics

Resampling Methodologies

9.1.5 Existed in 2008/2012

Yes

9.1.6 Participating Institution(s) to which the Research Group belongs

Fundação da Faculdade de Ciências (FFC/FC/UL)

9.2. RESEARCHERS IN THE GROUP

9.2.1 List of Integrated Members / 3 nuclear CVs

Name	Principal Investigator	Nuclear CV
MARIA IVETTE LEAL DE CARVALHO GOMES	Yes	Yes
MARIA ISABEL FRAGA ALVES	No	Yes
Dinis Duarte Ferreira Pestana	No	Yes
Amílcar Manuel do Rosário Oliveira	No	No
Ana Maria Santos Ferreira Gorjão Henriques	No	No
Clara Maria Henrique Cordeiro	No	No
Cristiana Maria Palmela Pereira	No	No
Fernanda Otilia de Sousa Figueiredo	No	No
José Leonel Linhares da Rocha	No	No
João Paulo Oliveira Martins	No	No
Ligia Carla Pinto Henriques Jorge Rodrigues	No	No
Luisa da Conceicao dos Santos do Canto e Castro de Loura	No	No
Manjunath, Bangalore Govinda Raju	No	No
Maria Manuela Costa Neves Figueiredo	No	No
Maria de Fátima Almeida Brilhante	No	No
Miguel Martins Felgueiras	No	No
Paula Cristina Martins dos Reis	No	No
Paulo Jose Araujo dos Santos	No	No
Sandra Maria Freitas Mendonça	No	No
Sandra Maria da Silva Figueiredo Aleixo	No	No
Teresa Paula Costa Azinheira Oliveira	No	No

9.2.2 List of current PhD students

NAME

César Duarte Alves da Rocha

Délia Canha Gouveia

Helena Alexandra Couceiro Feio de Almeida Penalva

José Antonio Seijas Macias

Manuel Ana Mário Alberto

Paulo Jorge Sampaio Morais

Paulo José Figueira Semblano

9.2.3 List of other researchers of the Research Group

NAME

Maria Marise Simões de Almeida

Regina Maria Baltazar Bispo

Frederico Almeida Gião Gonçalves Caeiro

Miguel Brás de Carvalho

Sandra Cristina Pires Dias

Helena Maria Simões Ferreira

Marta Susana Ribeiro Ferreira

Adelaide Maria de Sousa Figueiredo

Helena Maria Iglésias Pereira

Maria Joao Teixeira Martins

Pedro Rui Mazeda Gil

Maria Cristina Souto Miranda

Dora Susana Raposo Prata Gomes

Cláudia Margarida Pedrosa Neves

Pedro Duarte Leal Gomes Pestana

Rui Filipe Vargas de Sousa Santos

fernando jose araujo correia da ponte sequeira

LUÍS MIGUEL DOS SANTOS REIS SOARES DE ALMEIDA

Maria da Graca Santos Temido Neves Mendes

Björn Jochen Vandewalle

Rita Maria César e Sá Fernandes de Vasconcelos

9.3. RESEARCH GROUP DESCRIPTION AND ACHIEVEMENTS FOR 2008/2012

9.3.1 Description of the Research Group

This Research Group was composed by two teams, on Order Statistics, Extremes and Applications and on Probability, Modelling and Data Analysis, working in different but related topics of Statistics, with two senior researchers as leaders, M. Ivette Gomes and Dinis Pestana, respectively. The main research interests of the group have been in the fields of Univariate, Multivariate, Multidimensional and Spatial Extreme Value Theory (EVT), but developments in the areas of Statistical Quality Control (SQC), Reliability, Design of Experiments, Robustness and Outliers, Forensic Statistics, Chaotic Modelling, Meta-Analysis and Biometry have also been dealt with. Among the specific topics of research undertaken by the members of this Research Group, we mention:

1) In the field of EVT, the development of diverse estimation procedures for parameters of extreme events and extremes in population growth dynamics have been worked by M.F. Brilhante, M.I. Fraga Alves, A. Ferreira, M.I. Gomes, L. de Haan, S. Mendonça, M.C. Miranda, C. Neves and D. Pestana. Penultimate versus ultimate EVT models in Reliability were in charge of M.L. Canto e Castro, M.I. Gomes and P. Reis. New multivariate and spatial models in EVT were introduced and studied, essentially by A. Ferreira, L. de Haan, C. Neves, M.M. Neves and collaborators. M.I. Fraga Alves, M.I. Gomes, L. Henriques-Rodrigues, B.G. Manjunath and C. Neves performed a deeper study of the second and third order structure in statistics of extremes. Adaptive choice of threshold, with the use of heuristic and resampling-based methodologies has been advanced, essentially with the participation of M.F. Brilhante, C. Cordeiro,

F. Figueiredo, M.I. Fraga Alves, M.I. Gomes, M.M. Neves and D. Pestana. M.I. Gomes and M.M. Neves have developed adequate methodology for the analysis of extreme events related to censored data.

2) In the area of SQC, special emphasis was given to the role of the quantile function in the building of robust control charts, with research performed by F. Figueiredo, M.I. Gomes, A. Oliveira and T. Oliveira. Different topics in the areas of Experimental Design, Multivariate Data Analysis and Long-Term Experiences in Mathematics E-Learning were in charge of A. Oliveira and T. Oliveira. The multivariate Statis approach was used by F. Figueiredo to deal with a possible restructuration of privatized firms. M. Figueiredo-Neves and F. Rosado performed new developments on outlier selection.

3) Forecast Evaluation was a topic essentially in charge of P. Araújo Santos, C. Cordeiro, M.I. Fraga Alves and M.M. Neves.

4) S. Aleixo, M.F. Brilhante, M. Felgueiras, J.P. Martins, D. Pestana and F. Sequeira dealt with research on mixtures of beta densities in meta-analysis and extensions of the beta family. The use of statistical tools in Biomedicine research was in charge of C. Pereira, D. Pestana and L. Soares de Almeida. J. Cadima and collaborators have performed applications in animal and forestry data. S. Aleixo, M.F. Brilhante, M.I. Gomes, D. Pestana, P. Pestana, and L. Rocha did research in Population dynamics, and random repair models. S. Cabral and H. Gonçalves devveloped the bild package, on analysis of longitudinal data. Further developments in distance sampling were in charge of T.A. Marques and his collaborators.

5) Participation in the ALEA project, invited conferences at secondary schools and university colleges, and at Gulbenkian Foundation in charge of M.E. Graça Martins and D. Pestana. Tutorials for training scientific translators were in charge of D. Pestana and C. Ventura. Studies on Probability and Statistics publications at the beginning of the XXth Century were in charge of R. Santos and D. Pestana.

6) All new methodologies were applied not only to simulated data but also to real data essentially in the areas of Environment, Finance, Forensics, Insurance Life Sciences, Medicine and Risk Assessment.

9.3.2 Main achievements

Among the achievements, we refer the 17 articles published as chapters of international books, on topics as diverse as Tail Behaviour, Testing Super Heavy Tails, Population Growth, Completed Joint Regression Analysis, Synchronizability and the Topology of Networks, Bootstrap and Imputation Techniques, Determinism and Randomness, Dynamical Analysis, Logistic Regression, Fractility of Asymptotic Models, Data Mining and Quality in Service and Industry and Comparison of Higher-Education Models, among others. We further mention the 5 articles, on Extreme Value Distributions, Statistical Process Control, Acceptance Sampling, Distance Sampling and p-values Combination, invited to International Encyclopedia of Statistical Science, edited by M. Lovric, 1 article on Extreme Values in Reliability, invited to the Encyclopedia of Quantitative Risk Assessment and Analysis, edited by E. Melnick and B. Everitt, and 1 article on Conventional Distance Sampling, invited to The Encyclopedia Sponsored by Statistics and Probability Societies.

It is also worth mentioning the edition of a volume of Revstat (6:1), on Statistics of Extremes and Related Fields, of two volumes of Discussiones Mathematicae Probability and Statistics (29:2 and 30:1) and the edition of 5 books related to the series of the Workshop on Statistics, Mathematics and Computation (3rd, 4th, 5th and 6th) and to the Workshop on Risk and Extreme Values in Insurance and Finance, a two-day workshop aimed to build a bridge between researchers and practitioners, held in the framework of the 100-year anniversary of the Faculty of Sciences of the University of Lisbon in 2011, a unique event which brought together in Portugal the three authors of the book Modelling Extremal Events for Insurance and Finance, P. Embrechts (Zurich, Switzerland), C. Klüppelberg (Munchen, Germany), and T. Mikosch (Copenhagen, Denmark).

Among the international publications in peer-review journals, we have counted a total number of 125 published articles in the period 2008-2012 (21 + 33 + 22 + 24 + 25), in journals like Acoustics Today, Advances in Applied Probability, Annals of Applied Statistics, Annals of Statistics, Biometrics, Communications in Statistics-Simulation and Computation & Theory and Methods, Ecology, Economic letters, Endangered Species Research, Extremes, Computatinal Statistics and Data Analysis, GEM, Hidrological Processes, ICES Journal of Marine Sciences, International J. of Bifurcation and Chaos, International J. Dermatology, JABES, J. American Statistical Association, J. of Applied Ecology, Journal of Applied Statistics, J. Forensic and Legal Medicine, J. of Mammalogy, J. Multivariate Analysis, J. of Nonlinear Mathematical Physics, Journal of Ornithology, J. of the Royal Statistical Society, J. Statistical Computation and Simulation, J. Statististical Planning and Inference, Journal of Statistical Software, Life and Marine Sciences, Metrika, Pharmacoeconomics, Portugaliae Mathematica, Spatial Statistics, Revstat, Significance, Skin Cancer, Statistical Neerlandica, Statistics and Probability Letters, Structural Change and Economic Dynamics, Test, The American Journal of Gastroenterology, The Annals of Probability and Wildlife Biology in Practice.

Apart from more than 100 published abstracts and extended abstracts, we have further counted 61 articles in international proceedings, 26 articles in national proceedings and 15 articles in national journals. It is also worth

mentioning the edition of 6 international books, the authorship of 8 national books, more than 100 invited conferences and the scientific involvement in 44 workshops and conferences in this period. There were 18 PHD students who got their degree in this period, under the supervision of members of the Research Group.

9.4. RESEARCH GROUP OUTPUT 2008/2012

9.4.1 Publications in peer reviewed journals and/or other publications

Araújo Santos, P. and Fraga Alves, M.I. (2012). "A new class of independence tests for interval forecasts evaluation". Computational Statistics and Data Analysis 56:11, 3366-3380.

Beirlant, J., Figueiredo, F., Gomes, M.I. & Vandewalle, B. (2008). Improved reduced-bias tail index and quantile estimators. J. Statist. Plann. and Inference 138: 6, 1851-1870.

Beirlant, J., Caeiro, F. and Gomes, M.I. (2012). An overview and open research topics in the field of statistics of univariate extremes. Revstat 10:1, 1-31.

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.

Canto e Castro, L. & S. Dias (2011). Generalized Pickands' estimators for the tail index parameter and max-semistability. Extremes 14:4, 429-449.

Ferreira A., de Haan L. and Zhou C. (2012). Exceedance probability of the integral of a stochastic process. Journal of Multivariate Analysis 10:1, 241-257.

Fraga Alves, M.I., Gomes, M.I., de Haan, L. & Neves, C. (2009). Mixed Moment Estimators and Location Invariant Alternatives. Extremes 12, 149-185.

Gomes, M. I., de Haan, L. and Henriques Rodrigues, L. (2008). Tail index estimation for heavy-tailed models: accommodation of bias in the weighted log-excesses. J. Royal Statist. Soc. B, 70:1, 31-52.

Gomes, M.I., Figueiredo, F. and Neves, M. (2012). "Adaptive estimation of heavy right tails: the bootstrap methodology in action". Extremes15, 463-489.

Soares de Almeida, L., Requena, L., Rütten, A., Kutzner, H., Garbe, C., Pestana, D., and Marques Gomes, M. (2008). "Desmoplasic Malignant Melanoma: A Clinico-Pathologic Analysis of 113 Cases", American J. Dermatopathology 30, 207-215.

9.4.2 Completed PhD theses

Ferreira, Marta (2008). "Extremos em Séries Temporais Max-Autoregressivas". Local: Universidade de Lisboa, Supervisor: L. Canto e Castro.

Prata Gomes, Dora (2008). "Métodos Computacionais na Estimação Pontual e Intervalar do Índice Extremal". Local: Universidade Nova de Lisboa, Supervisors: M. M. Neves e J. Tiago Mexia.

Henriques-Rodrigues, Lígia (2009). "Estimação de Viés Reduzido em Estatística de Extremos".Local: Universidade de Lisboa, Supervisor: M. Ivette Gomes.

Martins, João Paulo (2009). "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.

Caneco, Acilina do Nascimento (2010). "Sincronização Caótica de Sistemas Dinâmicos não Lineares Acopolados", Doutoramento em Matemática, Local: Departamento de Matemática da Escola de Ciências e Tecnologia da Universidade de Évora (supervisors: J. L. Rocha and Maria Clara Grácio, UE).

Rocha, Maria Luísa (2010). "Modelos e Evidência Empírica na Limitação de Riscos Financeiros", Local: Universidade dos Açores (co-supervisors: António de Menezes, Dinis Pestana e Maria de Fátima Brilhante)

Araújo-Santos, Paulo (2011). "Excesses, Durations and Forecasting Value-at-Risk". Doutoramento em Estatística e Investigação Operacional, Ramo de Probabilidadee e Estatística, Local: Universidade de Lisboa, Supervisor: M. Isabel Fraga Alves

Cordeiro, Clara Maria Henrique (2011). "Métodos de Reamostragem e Previsão em Séries Temporais". Doutoramento em Matemática, Ramo de Estatística, Local: Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Supervisor: M.Manuela Neves.

Branco, Fernando José (2011). "Investigação Experimental: Potência Estatística dos Testes de Aleatorização na Comparação de dois Grupos Independentes". Doutoramento em Matemática, Planeamento de Experiências, Local: Universidade Aberta, 2011, Supervisor: Teresa Oliveira

Reis, Paula (2012). "Modelos Limite para a Fiabilidade de Sistemas de Grande Dimensão". Doutoramento em Estatística e Investigação Operacional, Ramos de Probabilidade e Estatística, Local: Universidade de Lisboa, Supervisor: Luísa Canto e Castro e Caldeira Duarte

9.4.3 Patents and Prototypes or other research outputs

9.4.4 Books and book chapters of international circulation

Aleixo, S., Rocha, J. L., and Pestana, D. (2011). Probabilistic Methods in Dynamical Analysis: Population Growths Associated to Models Beta (p,q) with Allee Effect. In Peixoto, M.M., Pinto, A.A., Rand, D.A.J. (Eds.), Dynamics, Games and Science, in honour of Mauricio Peixoto and David Rand, vol II, Ch. 5, 79-95, Springer Verlag, New York, ISBN 978-3-642-14787-6, e-ISBN 978-3-642-14788-3.

Caneco, A., S. Fernandes, C. Grácio and J. L. Rocha (2011). Networks Synchronizability, Local Dynamics and Some Graph Invariants. In Peixoto, M.M., Pinto, A.A. and Rand, D.A. (Eds), Dynamics, Games and Science, Springer Proceedings in Mathematics 1, Springer-Verlag Berlin Heidelberg, Vol. 1, Chap. 14, 221-238, ISBN 978-3-642-11456-4.

Cordeiro, C., Machás, A. and Neves, M. (2010). A Case Study of a Customer Satisfaction Problem: Bootstrap and Imputation Techniques. In Esposito Vinzi, V., Chin, W.W., Henseler, J., Wang, H. (Eds.), Handbook of Partial Least Squares Concepts, Methods and Applications in Marketing and Related Fields. Springer Handbooks of Computational Statistics. 279-288, ISBN 978-3-540-32827-8.

Fraga Alves, M.I., Neves, C. and Cormann, U. (2010). Heavy and Super-Heavy Tail Analysis. In Ch2 Section 2.7, pgs 75-101, of Laws of Small Numbers: Extremes and Rare Events, by Michael Falk, Jürg Hüsler and Rolf-Dieter Reiss. Third Edition, Springer-Basel, December 2010, XVI, 509 pages, ISBN: 978-3-0348-0008-2.

Gomes, M.I., Pestana, D., Henriques-Rodrigues, L. & Viseu, C. (2008). Tail Behaviour: an Empirical Study. In Arnold, B.C., Balakrishnan, N., Sarabia, J.M. & Mínguez, R. (eds.), Advances in Mathematical and Statistical Modelling, Chapter 14, 195-207, Birkhauser, Boston, ISBN 978-0-8176- 4625-7.

Neves, C. & Fraga Alves, M.I. (2008). Ratio of Maximum to the Sum for Testing Super Heavy Tails. In Arnold, B.C., Balakrishnan, N., Sarabia, J.M. and Mínguez, R. (eds.), Advances in Mathematical and Statistical Modeling, Chapter 13, 181-194, Birkhauser, Boston, ISBN 978-0-8176- 4625-7.

Oliveira, T.A., Oliveira, A. and Pérez-Bonilla, A. (2012). Data Mining and Quality in Service Industry: Review and some Applications. In J. Faurlin, A. Juan, S. Grasman & M. Fry (eds.), Decision Making under Uncertainty in the Service Industry, CRC/Taylor & Francis, 83-107, ISBN 9781439867341.

Pestana, D., Aleixo, S., and Rocha, J. L. (2011). "Regular variation, paretian distributions, and the interplay of light and heavy tails in the fractality of asymptotic models". In C.H. Skiadas, I. Dimotikalis & C. Skiadas (Eds.), Chaos Theory: Modeling, Simulation and Applications, World Scientific Books, 309-316, ISBN: 978-981-4350-33-4.

Seco, A., Felgueiras, M., Fdez-Riverola, F. and Pereira, A. (2011). Elder Care Alert Management-Decision Support by a Logistic Regression Model. Advances in Intelligent and Soft Computing Vol. 90. Trends in Practical Applications of Agents and Multiagent Systems, 9-16. Springer, Berlin, ISBN: 978-3-642-19874-8.

Trenholm, S., Juan, A.A., Simosa, J., Oliveira, A. and Oliveira, T.A. (2012). Long-Term Experiences in Mathematics E-Learning in Europe and the US: A Comparative Study Regarding Differing Higher-Education Models. Invited Chapter Book in IGI Global Publisher -Math-E-Learning: Teaching Mathematics Online: Emergent Technologies and Methodologies, 238-258, ISSN 1698-580X.

9.4.5 Conference proceedings

Aleixo, S., Rocha, J. L., and Pestana, D. (2008). Populational Growth Models in the Light of Symbolic Dynamics. In Luzar-Stiffler, V., Dobric, V.H. and Bekic, Z. (eds.), Proceedings of the ITI 2008, 30th International Conference on Information Technology Interfaces, 311-316.

Brilhante, M. F., Mendonça, S., Pestana, D., and Sequeira, F. (2010). Using Products and Powers of Products to Test Uniformity. In Luzar-Stiffler, V., Jarec, I. and Bekic, Z. (eds.), Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces, IEEE CFP10498-PRT, 509-514.

Cordeiro, C. and Neves, M. (2008). Bootstrap and exponencial smoothing working together in forecasting time series. In Paula Brito (Ed.), Proceedings in Computational Statistics (COMPSTAT 2008), 891-899, Physica-Verlag

Felgueiras, M., Martins, J. and Santos, R. (2012). Pseudo-convex Mixtures. In T. E. Simos, G. Psihoyios, Ch. Tsitouras and Z. Anastassi (eds.) Numerical Analysis and Applied Mathematics ICNAAM 2012, American Institute of Physics, AIP Conf. Proc. 1479, 1125-1128, 2012 (doi: 10.1063/1.4756346).

Ferreira, A. (2010), Spatial Extremes and High Quantile Estimation of Aggregated Rainfall Proceedings of the 45th Scientific Meeting of the Italian Statistical Society, Padua 16-18 June, 2010, p.1-9.

Fraga Alves, M.I., de Haan, L. & C. Neves (2011). Right Endpoint of a Distribution in Gumbel Domain of Attraction - Statistical Inference. Bullettin International Statistical Institute LXIV, Int. Statistical Inst.: Proc. 58th World Statistical Congress, 2011, Dublin (Session STS040) 2994-2995).

Gomes, D., Nunes, C. and Canto e Castro, L. (2008) - 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. 2008. Holland.

Gomes, M.I., Caeiro, F. and Henriques-Rodrigues, L. (2012). PORT-PPWM extreme value index estimation. In Colubi, A., Fokianos, K., Kontoghiorghes, E.J. and González-Rodríguez, G. (eds.). Proceedings of COMPSTAT 2012, The International Statistical Institute/International Association for Statistical Computing, 259-270.

Neves, M., Gomes, M.I., Figueiredo, F. and Prata-Gomes, D. (2012). Modelling extreme events: sample fraction adaptive choice in parameter estimation. In T. E. Simos, G. Psihoyios, Ch. Tsitouras and Z. Anastassi (eds.), Numerical

Analysis and Applied Mathematics ICNAAM 2012, AIP Conference Proceedings 1479, American Institute of Physics, 1110-1113, ISBN 978-0-7354-1089.

Oliveira, T.A. and Oliveira, A. (2012). "Ineffectiveness of the FIM in selecting optimal BIB designs for testing block effects". In Colubi, A., Fokianos, K., Kontoghiorghes, E.J. and González-Rodríguez, G. (eds.). Proceedings of COMPSTAT 2012, 701-712, The International Statistical Institute/International Association for Statistical Computing.

9.4.6 New materials, devices, products and processes

9.4.7 Software, computer code and algorithms

Gonçalves, M.H., Cabral, M.S., and Azzalini, A. (2010). bild: a package for Binary Longitudinal Data, Version 1.0. URL: http://CRAN.R-project.org/package=bild

Stefan Wilhelm and B. G. Manjunath, tmvtnorm: A Package for the Truncated Multivariate Normal Distribution, The R Journal 2(1) (2010), 25-29 (a basis for packages under development for parametrical and semi-parametrical analysis of extreme events)

9.4.8 Books, including single-authored works (including scholarly editions of oral or written texts and translations with introduction and commentary)

Brilhante, M. F., Pestana, D. D., Rocha, J., Rocha, M.L., e Velosa, S. (2011). Inferência Estatística Sobre a Localização Usando a Escala, XVIII+289 p., INE edistions, Lisboa, ISBN: 978-989-25-0149-9.

Cabral, M.S. e Gonçalves, M.H. (2011). Análise de Dados Longitudinais. Sociedade Portuguesa de Estatística, ISBN: 978-972-8890-24-7.

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

Figueiredo, F., Figueiredo A., Ramos, A. e Teles, P. (2009). Estatística Descritiva e Probabilidades-Problemas Resolvidos e Propostos com Aplicações em R, Escolar Editora, 2ª edição (409 pgs.), ISBN: 9789725922491.

Gomes, M.I., Figueiredo, F. & Barão, M.I. (2010). Controlo Estatístico da Qualidade. Edições I.N.E. (xii+305 pages), 2ª edition, revised, ISBN: 978-972-8890-23-0, 2010.

Pereira, C. (2012). Medicina Dentária Forense. Lisboa, Lidel, ISBN: 978-972-757-753-8.

Pestana, D.D. e Velosa, S. (2008). Introdução à Probabilidade e à Estatística, 3ª ed. revista e aumentada, 1164 pp., Calouste Gulbenkian Edition, ISBN: 978-972-31-1248-1.

Pestana, D.D. e Velosa, S. (2010). Introdução à Probabilidade e à Estatística, 4ª ed. revista, 1164 pp., Calouste Gulbenkian Edition, ISBN: 9789723111507.

9.4.9 Edited special issues of journals, with substantial research input on the part of the researcher

Beirlant, J., Fraga Alves, I. and Leadbetter, R. (2008) Special issue on "Statistics of Extremes and Related Fields, REVSTAT - Statistical Journal, Volume 6, Number 1, ISSN: 1645-6726.

Oliveira, T.A., Mejza, S. (2008): Third Workshop on Statistics, Mathematics and Computation, First Portuguese-Polish Workshop on Biometry - Book of Abstracts, Universidade Aberta Editions, ISBN: 978-972-674-541-9.

Rosado, F. (2008-2012). Edição dos Boletins Primavera e Outono da Sociedade Portuguesa de Estatística (SPE), Edições SPE, ISSN: 1646-5903.

Agra Coelho, C. and Pestana, D. (2009). Guest editors of the special issue of Discussiones Mathematicae Probability and Statistics 29:2, ISSN 1509-9423 (print version), ISSN 2084-0381 (electronic version).

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

Fraga Alves, M.I., Gomes, M.I., de Haan, L. and C. Neves (2011). "Risk and Extreme Values in Insurance and Finance": Book of Abstracts. CEAUL Editions, ISBN: 978-989-8203-73-1 (x + 120 pages).

Gomes, M.I., Pestana, D. & Silva, P. (2008). Bull. Internat. Statist. Inst. LXII (Electronic publication), ISI editions.

Agra Coelho, C. and Pestana, D. (2010). Guest editors of the special issue of Discussiones Mathematicae Probability and Statistics 30:1, ISSN 1509-9423 (print version), ISSN 2084-0381 (electronic version).

Oliveira, T.A., Castela, G. Oliveira, A., Grilo, L. and V. Vairinhos (2011). "Book of Abstracts: 5th Workshop on Statistics, Mathematics and Computation". ISBN: 978-972-9473-54-8. Instituto Politécnico de Tomar Editions.

Oliveira, T.A., Mejza, S., Ferreira, D., Oliveira, A. and Grilo, L. (2012). "Book of Abstracts -6th Workshop on Statistics, Mathematics and Computation & 3rd Portuguese-Polish Workshop on Biometry", July 3-4, 2012, Covilhã, Portugal. ISBN: 978-972-9473-62-3, Depósito Legal: 345933/12, Published by Instituto Politécnico de Tomar, 112 pages.

9.4.10 Chapters in books, including contributions to conference proceedings, essays in collections

Felgueiras, M. (2008). Considerações sobre a Distribuição Pareto. In Hill, M.M. et al. (eds.), Estatística: da Teoria à Prática (Actas XV Congresso Anual da SPE), Edições S.P.E.,193-202, 2008.

Ferreira, M. e Canto e Castro, L. (2008). Distribuição estacionária e comportamento extremal de um processo RARMAXp. In Hill, M.M., Ferreira, M.A., Dias, J.G., Salgueiro, M.F., Carvalho, H., Vicente, P. E Braumann, C. (eds.), Estatística: da Teoria à Prática (Actas XV Congresso Anual da SPE), 203-214.

Figueiredo, F. e Gomes, M.I. (2008). Monitorização de um processo de produção de rolhas - comparação da performance das cartas de controlo tradicionais com a de cartas robustas. In Hill, M.M. et al. (eds.), Estatística: da Teoria à Prática, Edições S.P.E., 245-258

Henriques-Rodrigues, L. (2008). Estimação do índice de cauda em modelos de caudas pesadas acomodação do vies nos excessos acima de um "threshold" elevado. In Hill, M.M. et al. (eds.), Estatística: da Teoria à Prática (Actas XV Congresso Anual da SPE), 45-69 (Prémio SPE), Edição

Martins, J. e Mendonça, S. (2008). Transformações estabilizadoras da variância. In Hill, M.M. et al. (eds.), Estatística da Teoria à Prática, Actas do XV Congresso Anual da Sociedade Portuguesa de Estatística, Edições SPE, Lisboa, 343-350

. Monteiro, S., Oliveira, T.A. (2008). Estruturas dos Planos Factoriais Fraccionados de Base Dois. In Hill, M.M. et al. (eds.), Estatística da Teoria à Prática, Actas do XV Congresso Anual da SPE, Edições SPE, 2008, pp. 375-389

. Figueiredo, A., Figueiredo, F. e Monteiro, N.P. (2009). Partilha de rendas nos bancos privatizados em Portugal. Em Irene Oliveira et al. (eds.), Estatística-Arte de Explicar o Acaso, Actas do XVI Congresso Anual da SPE, Edições S.P.E, 275-286.

Gomes, M.I. e Pestana, D. (2009). Caudas leves em desporto: estimação de parâmetros úteis. In Oliveira, I., M.M. et al. (eds.), Estatística: Arte de Explicar o Acaso, Actas do XVI Congresso Anual da SPE , Edições S.P.E., 307-318

Neves, M. e Santos, J.A. (2009). Modelação e estimação não paramétrica e semiparamétrica'. in Estatística. Arte de Explicar o Acaso (Irene Oliveira, et al. edts) Actas do XVI Congresso Anual da Sociedade Portuguesa de Estatística, SPE, 59-79.

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. In G. Cunha e J. Varanda (eds.), Estatística e Qualidade na Saúde: Novos Rumos, Novos Desafios, Instituto Nacional de Estatística, Lisboa, 74-90

9.4.11 **Creative writing (to the extent that it embodies research)**

9.4.12 **Encyclopedia entries (to the extent that they embody research)**

Ferreira, A. (2008) Extreme Values in Reliability in Encyclopedia of Quantitative Risk Assessment and Analysis, Melnick, E., and Everitt, B. (eds). John Wiley & Sons, Ltd, Chichester, UK, pp 686-691, Online ISBN: 9780470061596.

Fraga Alves, M.I. & C. Neves (2011). Extreme Value Distributions. In Lovric, M. (ed.), International Encyclopedia of Statistical Science, Springer-Verlag, 493-496, ISBN 978-3-642-04897-5.

. Gomes, M.I. (2011). Statistical Process Control. In Lovric, M. (ed.), International Encyclopedia of Statistical Science, Springer-Verlag, Part 19, 1459-1463, ISBN 978-3-642-04897-5.

Gomes, M.I. (2011). Acceptance Sampling. In Lovric, M. (ed.), International Encyclopedia of Statistical Science, Springer-Verlag, Part 1, 5-7, ISBN 978-3-642-04897-5.

Marques, T. A., Buckland, S. T., Borchers, D. L., Rexstad, E. & Thomas, L. (2011), Distance Sampling, In Lovric, M. (ed.), International Encyclopedia of Statistical Science, Springer-Verlag, 398-400, ISBN 978-3-642-04897-5.

Marques, T.A., Buckland, S.T., Borchers, D.L., Rexstad, E. and Thomas, L. (2012). Conventional Distance Sampling (version 1). StatProb: The Encyclopedia Sponsored by Statistics and Probability Societies. Freely available at: http://statprob.com/encyclopedia/ConventionalDistanceSampling.html

Pestana, D. (2011). Combining p-values. In Lovric, M. (ed.), International Encyclopedia of Statistical Science, Springer-Verlag, Part 1, 1145-1147, ISBN 978-3-642-04897-5.

9.4.13 **Audio/visual and electronic/digital materials**

9.4.14 **Other categories, including web-based resources; video and audio recordings (to the extent that they embody research)**

9.4.15 **Performances and exhibitions to the extent that they embody research**

9.4.16 **Other research outputs**

Araújo Santos, P., Soares Lopes, M.V. e Vieira de São João, R.M. (2012). A Estatística no Instituto Politécnico de Santarém. Boletim da Sociedade Portuguesa de Estatística, Primavera, Edições SPE, 57-58.

. Brilhante, M. F., Pestana, D., e Rocha, M. L. (2011). Betices. Boletim da Sociedade Portuguesa de Matemática, Número Especial (Actas do Encontro Nacional da SPM 2010), 238-246.

Felgueiras, M. (2011). Mistura de gaussianas e distribuição t de Student em Finanças. Boletim da Sociedade Portuguesa de Matemática, Número Especial (Actas do Encontro Nacional da SPM 2010), 210-214.

Ferreira A. e de Haan L. (2010). Processos Max-Estáveis: Uma caracterização simples e exemplos. Boletim da Sociedade Portuguesa de Estatística, Outono de 2010, 17-22.

Gomes, M.I. (2009). Algumas reflexões avulsas sobre a Sociedade Portuguesa de Estatística. Boletim da Sociedade Portuguesa de Estatística, Primavera de 2009, 50-52.

Henriques-Rodrigues, L., Grilo, L.M., Carvalho, F. e Pitacas, I. (2012). Instituto Politécnico de Tomar. Boletim Sociedade Portuguesa de Estatística, Primavera, 17-23.

Oliveira, T.A. (2011). "Planos em blocos incompletes equilibrados (BIB e PBIB designs): na fronteira entre a Estatística e a Matemática". Boletim da Sociedade Portuguesa de Matemática, Número Especial (Actas do Encontro Nacional da SPM 2010), 238-246.

Pestana, D. (2010). Devemos Acreditar em Análises Clínicas? Salutis Scientia 2, 4-14.

. Sampaio, T., Zwolinski, J e Neves M. (2010). A Análise de Dados Espaciais em duas áreas das Ciências Biológicas''+. Boletim Sociedade Portuguesa Estatística, Outono de 2010, 29-38.

Santos, J.A. & M.M. Neves (2011). A semiparametric model of the logistic regression. Boletim da Sociedade Portuguesa de Matemática, Número Especial (Actas do Encontro Nacional da SPM 2010), 197-203.

9.4.17 Organisation of scientific dissemination activities

"III Workshop on Statistics, Mathematics and Computation" (jointly with the "I Portuguese-Polish Workshop? on Biometry"), Lisboa, 21-22 July, 2008.

Statistics for Extremes. Invited Session to EVA 2009, Fort-Collins, Colorado, USA, July 22-26, 2009.

Statistics of Extremes in Today's World. Invited Session to the SIS 2010, "The 45th Scientific Meeting of the Italian Statistical Society, University of Padua, June 6-18, 2010.

Extremes and Risks. Invited Session to IWAP 2010, International Workshop on Applied Probability, Madrid, Spain, July 5-8, 2010.

Thematic sessions on "Probability and Statistics". Encontro Nacional da Sociedade Portuguesa de Matemática, July 8-10, 2010, Instituto Politécnico de Leiria.

5th Scientific Meeting ISLA: Data Mining and Business Intelligence, Methods and Applications, Santarém, Portugal, November 11-13, 2010.

Workshop on Risk and Extreme Values in Insurance and Finance, Lisbon, Portugal, June 6-7, 2011.

Statistics of Extremes in Today's World. Invited Session to ISI 2011, 58th Meeting of the International Statistical Institute, Dublin, Ireland, August 21-26, 2011.

XIX Congresso Anual da Sociedade Portuguesa de Estatística, Nazaré, Portugal, September 28-October 1, 2011.

Statistics of Extremes and Risk Assessment. Invited Session to ERCIM 2012, ERCIM WG on Computing and Statistics, Oviedo, 1-3 December, 2012.

9.4.18 Research contracts with national or international entities

Without any formal contract, and with the main objective introducing in the industry the recent developed robust methods in Statistical Process Control (SPC), we have kept on for a long time contacts with The Portuguese Institute of Quality ("Instituto Português da Qualidade")

Without any formal contract, and with the main objective introducing in the industry the recent developed robust methods in Statistical Process Control (SPC), we have kept on for a long time contacts with Auto-Europa

A research contract with CIM (Companhia Industrial da Matola, Mozambique), led to the supervision of a M.Sc. thesis on the Application of Experimental Design and Statistical Control in the Production of Cookies

We further recently began contacts with Corticeira Amorim, already formalized, in order to adequately control a cork stopper's process production. The title of the project is "Delineamento de Plano de Amostragem por Variáveis para Deteção do Composto 2,4,6-TCA em Rolhas de Cortiça", AMORIM & IRMÃOS, S.A. The project is running until the end of 2013, with a support of 17,500 EUR.

9.4.19 Projects funded in national and international competitive calls

CZ.1.07./2.3.00/20.0086, "STRENGTHENING INTERNATIONAL COOPERATION OF THE KLIMATEXT RESEARH

TEAM"

FCT/ PTDC/MAT/112770/2009: "EXTREMOS NO ESPAÇO"

FCT/PTDC/MAT/101736/2008 : "EXTREMA, Statistics of Extremes in Today's World"

Consorcium RISCOS - International Working Group on the Reform of the State Laboratories, Period: 2009-2012.

ESF Forward Look on "Mathematics and Industry"

Percurso Científico do Património Museológico da Academia de Ciências de Lisboa, referência HC/0074/2009

FCT Base (RG-MATH-LVT-Lisboa-6-948).

FCT/ PTDC/MAT/64924/2006: "EXES - SPATIAL EXTREMES"

FCT/PPCDT/ MAT/58876/2004: "ERSE - EXTREMES, RISK, SAFETY and the ENVIRONMENT"

PEst-OE/MAT/UI0006/2011. Strategic Project - UI 6 - 2011-2012.

9.5. ORGANISATIONAL STRUCTURE AND OBJECTIVES OF THE RESEARCH GROUP 2015/2020

9.5.1 Structure of the Research Group

This Research Group is going to be a reasonable large but homogeneous group, working in different but related topics of Statistics, within the general field of Order Statistics and Applications. M. Ivette Gomes, an expert in the areas of Extreme Value Theory (EVT), Statistical Quality Control (SQC), Computational Statistics, Exploratory Data Analysis and Resampling Methodologies, is going to be the leader of the Research Group. There will be a co-leader, Dinis Pestana, an expert in the areas of Modelling, Functional Methods in Probability, Stable and Extreme Models, Population Dynamics, Sampling, Risk Processes, History of Probability and Statistics and Meta-Analysis.
1) In the field of EVT, Order Statistics and Resampling Methodologies it is expected the involvement of 14 integrated researchers. There will be the collaboration of 11 researchers, who have had strong links with CEAUL. The four senior researchers, L. Canto e Castro, M.I. Fraga Alves, M.I. Gomes and M.M. Neves, will respectively co-ordinate the themes Penultimate Models, Extremes and Risk in Finantial Data (in strong connection with the thematic strand on "Risk in Financial Extremes"), Extremes and Risk in Environmental Data (in connection with the thematic strand on "Enviromental Statistics") and Applications of Extremes to Life Sciences and Biostatistics (now related to the thematic strand on "Biostatistics").
2) Research on Risk Assessment and Risk Analysis, Statistical Modelling and Computational Statistics, with applications in areas as diverse as Health Sciences, Medicine, Agriculture, Education Sciences and Biology wil be developed by different members of this Research Group, among whom we mention P. Araújo Santos, M.I. Fraga Alves, M.I.Gomes, A. Oliveira and T. Oliveira. As applications in Medicine, Dentistry and Forensic Medicine, essentially with the participation of D. Pestana and C. Pereira, together with the collaboration of M. Almeida and a Ph.D. student, it is worth mentioning theoretical issues of meta-analysis. Dealing with generalized and random p-values is one of the main research subjects of D. Pestana and F. Brilhante. Further developments will be investigated, with the collaboration of L. Soares de Almeida, F. Sequeira and R. Santos, and namely meta-analysis of results on desmoplastic melanoma is under way. R. Santos, one of the current collaborators of this Researcg Group, jointly with M. Felgueiras and J. P. Martins are specializing in extensions of Dorfman's theory on composite testing, one of the main goals being the analysis of toxic portfolios. But this problem has persistent consequences in health sciences, and will ultimately be one of the issues in the cooperation with the new research unit in epidemics and environmental health problems.
3) Fractal Issues in Population Modeling are also under the objectives of this team, with the main purpose of developing population dynamics models. Among the integrated members working in the topic, we mention J.L. Rocha, S. Aleixo, D. Pestana, M.I. Gomes, and F. Brilhante. L. Rocha and S. Aleixo will further deal with topics like dynamical analysis of population growth models, synchronization in and characterization of chaotic growth models with strong and weak Allee effect, adjustment of growth models to real data, bifurcations structure in population dynamics and definition and characterization of ordinal weighted correlation coefficients. D. Pestana, F. Brilhante and S. Mendonça will develop Scale Analysis Issues, using the independence of scale and location estimators. In the area of History of Science, D. Pestana is developing research in Probability and Statistics in Portugal during the first Republic, with the collaboration of R. Santos. Developments in the areas of Experimental Design, Multivariate Analysis and SQC will be essentially in charge of F. Figueiredo, M.I. Gomes, A. Oliveira and T. Oliveira, with the collaboration of two Ph.D. students and with A. Figueiredo, as a collaborator.

9.5.2 Objectives of the Research Group

The main objectives of the team for 2015-2020 are related to the development of new methodologies in the fields of Univariate, Multivariate, Multidimensional and Spatial Extremes, with special emphasis on their applications to Life Sciences, Environment, Risk, Insurance, Finance, Outliers and Biometry. Topics in the areas of Statistical Quality Control, Design of Experiments, Robustness, Population Dynamics and Scale Analysis will also be of related interest. We further mention E-Learning, Applications in Medicine, Dentistry and Forensic Medicine and History of Science.
1) In the field of Extreme Value Theory and Order Statistics, we intend to go on with the practical and theoretical

development of PORT methodologies for the EVI, VaR and other Parameters of Extreme Events estimation, not only for independent, identically distributed frameworks, but also for dependent and spatial set-ups. Adaptive Choice of Thresholds and Bias-Corretion are two additional topics of high relevance in the area, to be revisited and reformulated, as well as the use of Parametric Methodologies and the R Language for the reliable estimation of parameters of extreme or even rare events. The Bootstrap Methodology and other Resampling-based Methods will also be put in action. Attention will also be paid to semi-parametric testing and to the extremal index prediction.

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 univariate and multivariate control charts, as well as control charts for batch processes. 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. Long-Term Experiences in Mathematics E-Learning as well as Fractal Issues in Population Modeling are also under the objectives of this team.

3) On the other hand, stuttering fractal samples will be used to models for artificial intelligence and random repair. Related topics to be developed are dynamical analysis of population growth models, synchronization in chaotic growth models, modelling and characterization of growth models with strong and weak Allee effect, among others. Research in Probability and Statistics in Portugal during the first Republic will be developed. Scale analysis using the independence of scale and location estimators will also be a topic to be dealt with. Forensic Sciences, mainly Forensic Dentistry, will also be under investigation. Forensic Dentistry projects are now required to develop sophisticated statistical skills. Using careful planning of experiments and sophisticated statistical analysis, we are confident that we shall publish more reliable and accurate mathematical models for different forensic applications, contributing for instance for the improvement of the protocol of age estimation. Other important fields are human identification and characterization of past population such as paleopathology and paleodemography.

3) All the new developed methodologies, most of them devised for the analysis of spatial data, will be studied computationally and applied not only to simulated data but also to real data essentially in the areas of Environment, Finance, Insurance, Life Sciences, Medicine and Risk. We further intend to develop R packages related to those new methodological procedures.

(RG-6-74) Statistical Models and Methods for Complex Systems

9.1. IDENTIFICATION OF THE RESEARCH GROUP

9.1.1 Reference of the research group

RG-6-74

9.1.2 Name of the Research Group in portuguese

Modelos e metodos estatisticos para sistemas complexos

9.1.3 Name of the Research Group in English

Statistical Models and Methods for Complex Systems

9.1.4 Keyword(s)

Bayesian Statistics

Spatio-temporal Analysis

Incomplete and Censored Data Analysis

Longitudinal Data Analysis

9.1.5 Existed in 2008/2012

Yes

9.1.6 Participating Institution(s) to which the Research Group belongs

Fundação da Faculdade de Ciências (FFC/FC/UL)

9.2.1 List of Integrated Members / 3 nuclear CVs

Name	Principal Investigator	Nuclear CV
Maria Antonia Conceição Abrantes Amaral Turkman	Yes	Yes
Kamil Feridun Turkman	No	Yes
Carlos Daniel Mimoso Paulino	No	Yes
Ana Luisa Trigoso Papoila da Silva	No	No
Carina Soares da Silva Fortes	No	No
Giovani Loiola da Silva	No	No
Ivone Maria Ribeiro Figueiredo da Silva Rosa (CO-ORIENTADORA)	No	No
Jorge Filipe Campinos Landerset Cadima	No	No
Laurentius Franciscus Maria de Haan	No	No
Lisete Maria Ribeiro de Sousa	No	No
Luzia Augusta Pires Gonçalves	No	No
Maria Lucília Salema e Carvalho	No	No
Maria Salomé Esteves Cabral	No	No
Patricia Cortes Zea Bermudez	No	No
Marília Cristina de Sousa Antunes	No	No
Sandra Cristina de Faria Ramos	No	No
Valeska Lima Andreozzi	No	No

9.2.2 List of current PhD students

NAME

Soraia Alexandra Gonçalves Pereira

Paula Cristina Sequeira Pereira

Clandio Timm Marques

Maria João Fernandes Pereira Polidoro

Ana Patrícia Subtil da Graça Freitas Garcia

Carlos José Brás Geraldes

Ana Sofia Monteiro de Araújo Soares

Ricardo Miguel Vieira de São João

Maria Inês Tomás de Oliveira Pascoal de Sousa Dias

Fernando José do Nascimento Sebastião

9.2.3 List of other researchers of the Research Group

NAME

José Miguel Oliveira Cardoso Pereira

Maria da Conceicao Rodrigues Ribeiro

Cristina Maria Tristão Simões Rocha

Vanda Calhau Fernandes Inácio

Manuel Gonzalez Scotto

Nuno Henriques dos Santos de Sepulveda

Carlos do Carmo de Portugal e Castro da Camara

Paulo José de Jesus Soares

Maria Helena Neves de Queirós Gonçalves

Isabel Cristina Maciel Natário

Luis Filipe Meira Machado

Cristina Maria Sousa Catita

Ines Pereira Silva Cunha de Sousa

9.3. RESEARCH GROUP DESCRIPTION AND ACHIEVEMENTS FOR 2008/2012

9.3.1 Description of the Research Group

This research group was previously named Statistical Modelling in Environmental and Life Sciences. The main objective was to make contributions on specific probability and statistical issues in modelling environmental and biological data and to create a group with expertise in analyzing data coming from these fields. This last goal was achieved through the intensive collaboration with researchers from those fields. Examples were in
• Fisheries - Institute of Marine Sciences;
• Forest fires - Forest Fire Research Centre;
• Immunology and Genetics - Gulbenkian Institute of Science, University College London;
• Biology, Medicine, Epidemiology and Health - University of São Paulo, Fio Cruz-Brasil, Hasselt University, University of Texas, Portuguese Oncology Institute and Research Unit of Centro Hospitalar Lisboa Central and Institute of Hygiene & Tropical Medicine;
• Analysis of microarray data with researchers from diverse groups on molecular biology;
We also organized 5 workshops, 52 seminars and 12 advanced courses on themes related to the research interests of the group.
Among the specific topics of research we mention work on:

A) Time Series and Stochastic Processes
• Maxima of stationary processes.
• Exceedance probability of stochastic process.
• Non-linear Time Series Analysis.
• Periodic autoregressive model identification using genetic algorithms.
• Spatio-temporal non-homogeneous marked point processes with emphasis in modelling forest fires.
• Time series methods and hidden Markov models for influenza epidemics.
• Non-homogeneous Poisson models with a change-point.

B) Spatio-temporal analysis and methods: Hierarchical Bayesian models and simulation based inference techniques for spatio-temporal processes were studied and applied in several fields:
• Quantification of regional risk of large wildfire sizes;
• Joint modelling of probability of ignition and fire size in Australia;
• Analysis of fire frequency data using non-homogeneous Markov and discrete Weibull models;
• Over-dispersed longitudinal spatially correlated data with application to the analysis of regional variation for revascularization odds of patients hospitalized for acute coronary syndrome.
• Stomach cancer incidence in Southern Portugal

C) Missing observations and categorical data analysis
• Sensitivity analysis for incomplete continuous data;
• Missing data mechanisms and inferential implications of over-parameterization on analysis of categorical data;
• Comparison of diagnostic tests with missing data;

D) Bayesian latent class models: studies to measure deprivation among Portuguese households and in diagnosis of malaria; tools were explored to validate the assumption of conditional independence in these models.

E) Survival analysis and Longitudinal data:
• joint modelling of repeated measured and survival data with application to AIDS;
• Survival data analysis in prospective study of pain intensity in cancer patients and to cystic fibrosis cohort from Brazil;
• Hierarchical survival models with frailty in the identification of factors responsible for increasing the hazard of HIV related hospital mortality in Portugal.

F) Bayesian non-parametric methods:
• Estimation of the ROC surface;
• Screening methods in supervised classification;

G) Statistics in Genetics

• Robust statistical methods for the analysis of microarray data;
• Algorithms in the identification of differentially expressed genes and the analysis of gene expression profiles;
• Poisson mixture models in the estimation of T-cell repertoire diversity and clonal size distribution;
• Bayesian analysis of allelic penetrance models for complex binary traits.

H) Statistical Epidemiology and Public Health: among many works we mention:
• Counting processes with two change-points for the study of prevalence of tuberculosis;
• Assessment of the predictive value of diagnostic tests;
• Bayesian approach to cytogenetic dosimetry;
• Modelling over week patterns of alcohol consumption.

9.3.2 Main achievements

General Productivity: the group published 96 papers in international periodicals, 25 articles in conference proceedings and contributed with communications in 127 international conferences, of which 29 as invited speaker. At national level we mention 15 publications in conference proceedings and 7 national books. Ten PhD students got their degree in this period. CD Paulino' student was awarded with two prizes for the best PhD thesis in the area of Statistics and Mathematics in Brazil in 2011/2012.

We refer to the highlights of members of the present team in each of the previous themes.
A) Time Series & Stochastic Processes: KF Turkman(KFT), P Zea Bermudez(PB), L Carvalho(LC), L de Haan(LH), CD Paulino (CDP).
KFT published a chapter of a book on the asymptotic joint distribution of maxima of stationary processes and their discrete version; LH's work on exceedance probability of a stochastic process was published in the J. Multiv. Analysis. Studies on Non-linear Time Series were undertaken by PB and KFT and resulted on a book to be published by Springer. LC's work on influenza epidemics was published in J of Stat Meth Med Res and on Stat in Med (2013). She co-authored the package flubase. CDP's work on Non-homogeneous Poisson process applied to ozone peaks in Mexico City was published in Environ.&Ecol.Stats. (EES).
B) Spatio-temporal analysis: KFT, PB, MA Turkman(MAT), G Silva(GS) and AL Papoila (ALP)
KFT, MAT and PB´s work on spatial extremes and applications to wildfires resulted on 3 papers in Int. J. of Wildland Fire, EES and Extremes. Analysis of Australia fires by MAT and KFT was published in EES. GS's work on spatially correlated binomial data was published in Stat in Med. MAT and ALP worked on disease mapping; their paper on stomach cancer incidence was accepted in Biometrical J.
C) Missing observations and categorical data analysis: CDP's works on this topic were published in Stats&Computing, Test, Int Stat Review and J Appl Stat. He co-authored the R package ACD.
D) Latent class models:CDP and L Gonçalves (LG).
The work relative to the subtopics mentioned were published, respectively, in JAppl Stats, PLoSONE and Stats&Prob Letters.
E) Survival analysis and Longitudinal data:GS, V Andreozzi (VA) and ALP.
VA's work on hierarchical survival models was published on BMC Health Serv Res; books on survival analysis were published by VA and ALP. ALP's work on current status data using GAMS was published in the Int J Appl Maths&Stats. GS and VA supervised a PhD student on joint modeling of survival and longitudinal data (2013).
F) Bayesian non-parametric methods:MAT, S Ramos(SR), M Antunes (MA) and CDP.
The work on ROC surface by MAT and PhD student V Inacio was published in Biometrical J. SR, MAT and MA worked on screening methods on supervised classification. The work published in Comp Stats& Data Anal (CSDA) was generalized using Bayesian non-parametric methods and was part of SR's PhD thesis. CDP's work on Bayesian semi-parametric approach for the differential analysis of sequence counts data was published in JRSSC (2013).
G) Statistics in Genetics: L Sousa (LS), MAT, CDP, MA, SR, C Silva-Fortes (CSF).
LS and MA's work on an algorithm to identify differentially expressed genes was published in Biometrical J. CSF, MAT and LS developed and implemented a new plot for the analysis of gene expression profiles, published in BMC Bioinf.; SR, MA and MAT proposed a new Bayesian method for the classification of gene expression data published in CSDA. CDP's work on immunology and genetics was also published in CSDA.LS collaboration with research groups on molecular biology resulted on 11 papers;
H)Statistical Epidemiology and Public Health: ALP, LG, VA, MA, GL, CDP, MAT, CSF.
The topics mentioned in the description of this activity resulted in papers 4 published by CDP,MAT, CDP and VA, respectively.
Collaboration of ALP in health sciences resulted in 11 papers, LG in tropical medicine in 4 papers and VA in epidemiology in 7 papers.

9.4. RESEARCH GROUP OUTPUT 2008/2012

9.4.1 Publications in peer reviewed journals and/or other publications

Amaral-Turkman, MA, Turkman, KF, Le Page, Y and Pereira, JM (2011). "Hierarchical space-time models for fire ignition and percentage of land burned by wildfires". Environmental and Ecological Statistics, Volume 18, Issue 4, Page 601-617

Antunes, Marília and Sousa, Lisete. (2008) Bayesian Classification and Non-Bayesian Label Estimation via EM Algorithm to Identify Differentially Expressed Genes: a Comparative Study. Biometrical Journal 50, 5, 824-836

Gonçalves, L., A. Subtil, M.R. Oliveira, V.E. Rosário, P. Lee, M.F. Shaio (2012). "Bayesian latent class models in malariadiagnosis", PLoS ONE, 7, Nº7, e40633

Machado, C., Paulino, C.D. and Nunes, F. (2009). Deprivation Analysis based on Bayesian Latent Class Models. J. Applied Statistics, 36/8, 871-891

Sepúlveda, N., Paulino, C.D. and Penha- Gonçalves, C. (2009).Bayesian Analysis of Allelic Penetrance Models for Complex Binary Traits.Computational Statistics and Data Analysis 53, 1271-1283

Mendes, J., de Zea Bermudez, P., Pereira, J., Turkman, K.F. and Vasconcelos, V M. (2010). Spatial extremes of wildfire sizes: Bayesian hierarchical models for extremes. Environmental and Ecological Statistics, 17, 1, 1-28

Ramos, S, Amaral Turkman, MA, Antunes, M. (2010).Bayesian classification for bivariate normal gene expression. Computational Statistics and Data Analysis 54, 2012-2020

Inácio, V., Amaral Turkman, M .A., Nakas, C.T., Alonzo, T.A. (2011) "Nonparametric Bayesian estimation of the three-way receiver operating characteristic surface", Biometrical Journal, 53, 1011-1024.

Fortes, C.S., Amaral Turkman, M.A. and Sousa, L. (2012). Arrow Plot: a new graphical tool for selecting up and down regulated genes and genes differentially expressed on sample subgroups". BMC Bioinformatics 13:147 (10.1186/1471-2105-13-147)

Ursu, E. and Turkman, K.F. (2012). Periodic autoregressive model identification using genetic algorithms".J. of Time Series Analysis, 33(3) 398-405

9.4.2 Completed PhD theses

Correia, Elisete Maria Rodrigues, (2008). O Papel de Distribuicoes Condicionalmente Especificadas em Estatistica Bayesiana, Local: Universidade de Lisboa, Supervised by: MariaAntonia Amaral Turkman

Ramos, S. (2010) Metodos Estatisticos de Screening em Classificao Supervisionada, Ph.D. thesis, Local: Universidade de Lisboa, December 2010. Supervised by M A Amaral Turkman and Marília Antunes

Barreto Hernandez, Emiliano (2011) Statistical Methodologies for the Analysis and Normalization of RIP-CHIP Data, October 2011. PhD Thesis in Biostatistics and Bioinformatics, Local: DEIO - FCUL, Universidade de Lisboa, supervised by Lisete Sousa

Poleto, F. (2011). Analise de Dados Categorizados com Omissao em Variaveis Explicativas e Respostas. PhD Thesis, Department of Statistics, Local: University of Sao Paulo, April 2011. Supervised by Julio Singer and C. D. Paulino. Thesis awarded with Prize Casimiro Montenegro Filho, edicao 2011, da Secretaria de Assuntos Estrategicos da Presidencia da Republica do Brasil and prize CAPES 2012.

Reboucas, Silvia Pedro, (2011) Metodologias de Classificacao Supervisionada para Analise de Dados de Microarrays, June 2011. PhD Thesis in Probability and Statistics, Local: DEIO - FCUL, Universidade de Lisboa, supervised by Lisete Sousa and Ana Pires

Vicente, Leonel Silva (2011) Modelos Bayesianos Hierarquicos no Planeamento de Recursos Humanos, November 2011, PhD Thesis in Probability and Statistics,Local: DEIO - FCUL,Universidade de Lisboa, Supervised by Kamil Feridun Turkman

Inacio, Vanda Calhau, (2012). Semiparametric and Nonparametric Modeling of Diagnostic Data, Local: Universidade de Lisboa, Supervisor: Maria Antonia Amaral Turkman

Nunes, Baltazar. (2012). Statistical methods for modeling and nowcasting the impacts of influenza epidemics. PhD Thesis in Probability and Statistics, Local: DEIO - FCUL, Universidade de Lisboa, Supervised by Lucilia Carvalho and Isabel Natario.

Dias, Sara Marques Simões, (2012). Modelling HIV/AIDS length of stay in Portugal, Local: ISEGI, Universidade Nova de Lisboa, supervised by Rosario Oliveira and Valeska Andreozzi

Silva Fortes, Carina (2012) Aplicacao da Metodologia ROC na Analise de Dados de Microarrays. Local: Universidade de Lisboa. Supervisors: Maria Antonia Turkman e Lisete Ribeiro de Sousa

9.4.3 Patents and Prototypes or other research outputs

9.4.4 Books and book chapters of international circulation

Napoleao Patrícia , Mafalda Selas, Claudia Freixo, Catarina Ramos, Antonia Turkman, Valeska Andreozzi, Miguel Mota Carmo, Ana Maria Viegas-Crespo, Rui Cruz Ferreira, Teresa Pinheiro (2011) THE ROLE OF INFLAMMATORY BIOMARKERS IN THE ASSESSMENT OF CORONARY ARTERY DISEASE, In Coronary Angiography - Advances in Noninvasive Imaging Approach for Evaluation of Coronary Artery Disease. Edited by Prof. Baskot Branislav. Coronary Angiography. Chapter 16, 281-314.ISBN978-953-307-675-1

Sepulveda, N. & Carneiro, J. (2011). Repertoire dynamics of peripheral regulatory and effector T cells competing for antigen presenting cells. In Mathematical Models and Immune Cell Biology (C. Molina-Paris and G. Lythe, eds.), p.

275-304, Springer Verlag, London.ISBN 978-1-4419-7724-3

Turkman, K.F (2012). Discrete-Continuous Time Extremes of Stationary Processes, p 565-582. Handbook of statistics-vol 30, 2012.Time Series-Methods and Applications. Ed. C.R. Rao and Subba Rao, Elsevier-North Holland. ISBN: 978-0-444-53858-1

9.4.5 Conference proceedings

Ramos, S., Amaral Turkman, M.A. Antunes, M. (2008). Optimal Screening Methods in Gene Expression Profiles Classification. In Proceedings in Computational Statistics 2008, edited by Paula Brito, Verlag Heidelberg, vol II, part III, 137-144

Poleto F., Molenberghs G., Paulino C.D. and Singer J. (2010). Sensitivity analysis for incomplete continuous data: In Proceedings of the 25th International Workshop on Statistical Modelling, Glasgow, July 5-9 (A. Bowman, ed.), 445-448

Goncalves, L., Subtil, A., Oliveira, M.R. (2011) Sampling Strategies and Bayesian Latent Class Models for Estimating Sensitivity and Specificity?. In Proceedings of XIII Spanish Biometry Conference and III Ibero-American Biometry Meeting, Barcelona, September 7-9, 2011 (Electronic ed., Number 101, 4 pages).

Olivença, D., Zea Bermudez, P. , Figueiredo, J., Belo, S., Gonçalves, L. (2011) Count models - an application to the study of the Schistosomiasis?. In Proceedings of XIII Spanish Biometry Conference and III Ibero-American Biometry Meeting, Barcelona, September 7-9, 2011 (Electronic ed., Number 110, 4 pages)

Paulino, C. D. Silva, G. L. and Branco, M. (2011). A Bayesian approach to cytogenetic dosimetry?. In Proceedings of XIII Spanish Biometry Conference and III Ibero-American Biometry Meeting, Barcelona, September 7-9, 2011 (Electronic ed., Oral commun., Number 96, 6 pages)

Sousa, L. (2009) 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

Sepulveda, N., Guindani, M., Paulino, C. D. and Mueller, P. (2011). T-cell receptor diversity estimation through Poisson mixture models?. In Proceedings of XIII Spanish Biometry Conference and III Ibero-American Biometry Meeting, Barcelona. September 7-9, 2011 (Electronic ed., Oral commun., Number 98, 4 pages)

Silva, G. L. and Dias, M. I. (2011). Spatio-temporal analysis of forest fires in Portugal?. In Proceedings of XIII Spanish Biometry Conference and III Ibero-American Biometry Meeting, Barcelona, September 7-9, 2011 (Electronic ed., Oral commun., Number 98, 4 pages)

Gonçalves, L. A. Subtil, N. Brites, M.R. Oliveira, A.M. Alho, M. Landum , J. Meireles , L.M. Madeira de Carvalho, S. Belo (2012). ?Bayesian Latent Class Models in Veterinary and Human Epidemiology?, 46th Scientific Meeting of the Italian Statistical Society, 8 pages.

Papoila, A.L., Ribeiro,C., Sao Joao, R., Geraldes, C., Amaral-Turkman, M.A., Miranda, A. (2012). ?Stomach cancer incidence in Southern Portugal 1998-2006: a Spatio-temporal analysis?. Proceedings of the VI International Workshop on Spatio-Temporal Modelling (METMA), Guimar?es, Portugal, 12-14 September 2012. (4 pages-edited in CD-room)

9.4.6 New materials, devices, products and processes

9.4.7 Software, computer code and algorithms

Nunes, B.; Natario, I.; Carvalho, L. (2009). R package flubase.
http://cran.r-project.org/web/packages/flubase/index.html

Silva-Fortes, C., Amaral Turkman, M.A. and Sousa, L. (2012) Arrow plot: a new graphical tool for selecting up and down regulated genes and genes differentially expressed on sample subgroups. R code for implementation of Algorithms 1 and 2. http://www.biomedcentral.com/1471-2105/13/147/additional

Poleto, F., Singer, J., Paulino, C.D., Correa, F. and Jelihovschi, E. (2012). Package ACD: Categorical data analysis with complete or missing responses. Version 1.5. CRAN (Comprehensive R Archive Network).

9.4.8 Books, including single-authored works (including scholarly editions of oral or written texts and translations with introduction and commentary)

Carvalho, M. L. & Natário, I. (2008). Análise de Dados Espaciais. Sociedade Portuguesa de Estatística. (ISBN: 978-972-8890-18-6)

Antunes, M. (2010) CRM e Prospecção de dados. Edições CEAUL.ISBN: 978-989-8203-45-8 (144 pgs).ISBN: 978-989-8203

Rocha, C. e Papoila, A.L. (2009). Análise de Sobrevivência. Edições SPE, Lisboa.ISBN: 978-972-8890-22-3

Carvalho, MS; Andreozzi, VL; Codeço, CT; Campos, DP; Barbosa, MTS; Shimakura, SE. (2011). "Análise de Sobrevivência: Teoria e aplicações em saúde". Editora Fiocruz, Rio de Janeiro, Brazil. ISBN: 978-85-7541-216-9. 2ª edição (revista e ampliada).

Murteira, B. e Antunes, M. (2012). Probabilidades e Estatística, Volume I. Escolar Editora, Lisboa, 501 p.

ISBN:978-972-592-355-9.

Murteira, B. e Antunes, M. (2012). Probabilidades e Estatística, Volume II. Escolar Editora, Lisboa, 612 p. ISBN:978-972-592-359-7.

Natário, I. (2012). Probabilidades e Estatística. Arquivo Escolar.

9.4.9 Edited special issues of journals, with substantial research input on the part of the researcher

Sousa, L., Andreozzi, V., Antunes, M. and Gonçalves, L. (2011) Special issue on Statistical Modelling: Challenges in Health. REVSTAT - Statistical Journal, Volume 9, No. 1, March 2011. ISSN 0873-4275

9.4.10 Chapters in books, including contributions to conference proceedings, essays in collections

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

Silva, G.L., Oliveira, M. and Borges, J. (2009).Modelling and analysis of forest fire data in Portugal. In Invited Paper Meeting (Statistics in the Omics Era), Proceedings of the SIS 2009 Statistical Conference - Statistical Methods for the analysis of large data-sets, pg. 211-214, Chieti-Pescara, Italy

Pereira, P.S., Turkman, K.F. (2011) Spatial Point Processes Applied to the Study of Forest Fires In Portugal, In Proceedings of the Spatial Data Methods for Environmental and Ecological Processes - 2nd Edition, Foggia-Baiadelle Zagare, Italy, (B.Cafarelli, ed.), 443-446.

Ribeiro, C., Amaral Turkman, M.A. and Cardoso, J.C. (2011) Bayesian Hierarchical Models: An analysis of Portugal Road Accident Data, In Proceedings of the Spatial Data Methods for Environmental and Ecological Processes - 2nd Edition, Italy, Foggia-Baiadelle Zagare September, 01-02, (B.Cafarelli, ed.), 101-104.

Machado, C., Paulino, C.D. e Nunes, F. (2008). Analise multidimensional da pobreza em Portugal com base em modelos bayesianos de classes latentes. Em Estatistica da Teoria a Prática (Actas do XV Congresso Anual da Sociedade Portuguesa de Estatistica), 277-287, SPE

Poleto, F., Singer, J. e Paulino, C.D. (2008). Analise log-linear em dados categorizados estratificados com respostas omissas. In Estatistica da Teoria a Pratica,Hill, M.M. et al. (eds.) (Actas do XV Congresso Anual da Sociedade Portuguesa de Estatistica), 437-450, SPE.

Sepulveda, N., Paulino, C.D. e Carneiro, J. (2008). Diversidade de linfocitos T Reguladores e Efectores. Em Estatistica da Teoria a Prática (Actas do XV Congresso Anual da Sociedade Portuguesa de Estatistica), Hill, M.M. et al. (eds.) 513-523, SPE

Silva-Fortes,C., Hernandez,E.B., Sousa,L., Amaral Turkman,M.A. and Gama-Carvalho,M. (2008). Análise por microarrays de complexos ribonucleoproteicos: normalização e selecção de genes com associação diferencial. In Estatistica, da Teoria a Prática: Actas do XV Congresso Anual da SPE (Hill et al. Eds.), Ediçoes SPE, p. 537-547

Castel-Branco, Marta, Antunes, Marilia, Cara d'Anjo, Andrea e Bernardo, Fernando (2009). Factores de risco associados a prevalencia de Salmonella em explorações de frangos destinados a produção de carne em Portugal.In Oliveira, I., M.M. et al. (eds.), Estatistica: Arte de Explicar o Acaso Actas do XVI Congresso da Sociedade Portuguesa de Estatistica, Ediçoes SPE, 187-194.

Teles, J.; Gonçalves, L. (2008) Power of normality tests under a mixture of Gaussian Distributions: a simulation Study. In Proceedings of 23rd International Workshop on Statistical Modelling (P. Eilers, Editor), pp.410-413

9.4.11 Creative writing (to the extent that it embodies research)

9.4.12 Encyclopedia entries (to the extent that they embody research)

9.4.13 Audio/visual and electronic/digital materials

9.4.14 Other categories, including web-based resources; video and audio recordings (to the extent that they embody research)

9.4.15 Performances and exhibitions to the extent that they embody research

9.4.16 Other research outputs

• Figueiredo, I. F., Carvalho, L., & Natario, I. (2010). DEEPFISHMAN. A FP7 Project: Management and Monitoring of Deep-sea Fisheries and Stocks. Case Study 3 C Report. Black scabbardfish in IXa. Unpublished manuscript.

• Government/Organization contract research: L Carvalho was the leader of the Working Group of the National Statistical Institute (INE) on Labour Force Survey - Implementation of Computer Assisted Telephone Interviewing/CATI)

• Under the protocol between the FFCUL and Exigo Consultores, P Bermudez, M Antunes, V Andreozzi, KF Turkman and MA Turkman, were involved in a research project aimed to assess the financing system of the public hospitals belonging to the Portuguese National Health Service. The report of this project was used by the Portuguese Court of Auditors in its final report audit available
athttp://www.tcontas.pt/pt/actos/rel_auditoria/2011/2s/audit-dgtc-rel030-2011-2s.shtm

• Antonia Turkman acts as a consultant to the project led by the consortium Siemens ? Joint Commission International (JC) towards the development of a National Health Care Evaluation System in Portugal (SINAS) aimed at a quality assessment and improvement across the health care system.

• A. Papoila gaveconsultancy support to the Research Unit of Centro Hospitalar de Lisboa Central (CHLC) as co-head of the Epidemiology and Statistics Consulting Core.

9.4.17 Organisation of scientific dissemination activities

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

StaM2010 - Statistical Modelling: Challenges in Health, Workshop, 9-12 May 2010, Lisbon. Organized by MA Turkman, L Sousa, V Andreozzi and M Antunes

Half-day Workshop on Spatio-temporal Modelling (1) Testing the Dependence Structure in Spatial and Spatio-temporal Models by W Gonzalez Manteiga, (2) A Kernel Indicator Variogram and its Application to Environmental Data, by R Menezes, (3) Discrete-Continuous Time Extremes of Stationary Processes, by KF Turkman. 11th January 2011, organized by G Silva.

Half-day Workshop on Biostatistics (1) Bayesian nonparametric ROC regression modeling by V Inacio and M Carvalho/ (2) Extending induced ROC methodology to the functional context by V Inacio/ (3) Detection of copy number variations using whole genome sequence data by N Sepulveda, CEAUL, 6 January 2012. Organized by G Silva.

CEAUL Courses in Applied Statistics to Environmental and Health Sciences, organized and lectured by: M Antunes, P de Zea Bermudez and V Andreozzi, 12-26 July 2012.

Short course on Bayesian Statistics via WinBUGS, organized and lectured by D Paulino, G Silva and P Soares, CEAUL, 17th-18th December 2009.

Short course on Categorical Data Analysis with Missing Observations. Organized and Lectured by D Paulino, CEAUL 22-23 January 2009.

Practical advanced course on Survival Analysis: Theory and Applications in Health. Organized by V Andreozzi and lectured by V Andreozzi and M Sa Carvalho, 28-30 July 2008.

Advanced Course on Bayesian Nonparametric Statistical Methods - Theory and Applications: Prof. Peter Mueller, (University of Texas M D Anderson Cancer Center, Houston - USA)-CEAUL, Lisbon, 11-15 January 2010. Organized by D Paulino.

Advanced Course on Bayesian Computing with INLA; Havard Rue, (Department of Mathematical Sciences - Norwegian University of Science and Technology) 8/11/2010. Organized by G Silva.

9.4.18 Research contracts with national or international entities

DEEPFISHMAN, FP7-KBBE-2008-1-4-02. Management and Monitoring of Deep-sea Fisheries and Stocks: Case study: Black scabbardfish. (CEAUL team: L. Carvalho, I. Natario and I. Figueiredo).

9.4.19 Projects funded in national and international competitive calls

Statistical Methods in Environmental and Epidemiological Processes, PTDC/MAT/118335/2010 - duration 1/1/2012-31/12/2014.Coordinator K.F. Turkman, Financed by FCT (team from CEAUL: K F Turkman, MA Turkman, P Bermudez, M Antunes, L Sousa, V Andreozzi, S Ramos, V Inacio, P Pereira)

Statistical Methods in Genetics and Environment, PTDC/MAT/64353/2006 - duration 1/1/2008-31/12/2010, Coordinator K.F. Turkman, Financed by FCT(team from CEAUL: K F Turkman, MA Turkman, P Bermudez, M Antunes, L Sousa, V Andreozzi, S Ramos, C Silva-Fortes)

Dengue in Madeira archipelago.Risk assessment for the emergence of Aedesaegypti mediated arboviroses and tools for vector control. (PTDC/SAU-EPI/115853/2009). IP: Carla Sousa. Team from CEAUL: L Gonçalves). 2012-2014

Extremes in Space : PTDC/MAT/112770/2009. - duration 1/1/2011-31/12/2013.Coordinator L. de Haan Financed by FCT. (in the group from 2012)

Ecoepidemiology of Dirofilariaspp: molecular characterization, potential vectors and transmission dynamics (PTDC/SAU-SAP/113523/2009). IP: Silvana Belo. Financed by FCT(team from CEAUL: L Gonçalves), 2011-2014.

New arboviruses isolated in Portugal. Risk assessment and public health application.PTDC/SAU-SAP/119199/2010. PI: Maria João P. F. Alves. Financed by FCT (team from CEAUL M Antunes) 2012-2014.

The relationship of circulating biomarkers of apoptosis and endothelial function with the plaque composition using VH IVUS (PIC/IC/82734/2007). PI: Teresa Pinheiro. Financed by FCT. (team from CEAUL: M A Turkman and V Andreozzi) 2009-2012

Eating habits in preschool children: a longitudinal approach to the identification of determinants and effects on body composition (PTDC/SAU-ESA/108577/2008.) PI: Carla Lopes. Financed by FCT (team from CEAUL: V. Andreozzi) 2010-2013.

Bayesian analysis of stochastic complex systems: methodological aspects and applications (Proc.º 4.4.1.00 CAPES) within the ambit of the protocol of scientific and technological cooperation FCT/CAPES (Brasil). PI: CD Paulino, (team from CEAUL: MA Turkman, KF Turkman, G Silva, V Andreozzi, M Antunes, P de Zea Bermudez, PhD students, Rui Martins and Inês Dias), 2010-2011.

Environment and Health in children day care centres (ENVIRH) (PTDC/SAU-ESA/100275/2008) PI: Nuno Neuparth (NOVA Medical School). Financedby FCT. (team from CEAUL: Ana Luisa Papoila). 01-02-2010 a 31-8-2013

9.5. ORGANISATIONAL STRUCTURE AND OBJECTIVES OF THE RESEARCH GROUP 2015/2020

9.5.1 Structure of the Research Group

Research undertaken by this group has been motivated by problems in areas of Environmental, Life and Health Sciences; It started,several years back, with three senior researchers specialized respectively on (i) Time Series and Extremes of Stochastic Processes, (ii) Bayesian Statistics and Decision,(iii) Stochastic Processes and Epidemic Models, with the goal of creating a group with expert knowledge to develop methodological and applied work, not only on these broad areas of research, but as well as on emerging areas relevant to solve problems in those scientific fields. This was achieved thanks to the advanced training program undertaken by the group, with the PhD students joining the group after getting their degree. The actual group comes from the addition of researchers on the area of Bayesian Statistics and Biostatistics in 2007. In 2012 Laurens de Haan joined the group due to our interest in the study of spatial extremes and applications in the environmental sciences. At present, with the restructure of the R&D unit, other two members of CEAUL and a collaborator joined the group: (i) S Cabral whose research interests are on Longitudinal Data Analysis and Repeated Measures, Generalized Linear Models; Analysis of Variance and Design of Experiments, (ii) J Cadima whose research interests are on Multivariate Statistics, Linear and Matrix Algebra, fundamental knowledge for the persecution of the research strategy of the group and (iii) I Figueiredo, a long term collaborator of the group, an expert researcher on modelling and management of fishing resources.

Strategically the group does fundamental research on areas with strong impact in applications. We believe that Statistics is an interdisciplinary Science and therefore we give utmost importance to collaboration with researchers from other scientific areas. As such, the group has three core teams, although members can participate and, usually do so, in more than one team:

A. A team who undertakes methodological research; all members of the group are encouraged to do methodological research, although some will contribute more on the applications. The main topics of research are: spatio-temporal analysis (KF Turkman, MA Turkman, G Silva, AL Papoila, L Carvalho, I Figueiredo); non-linear time series(KF Turkman, MA Turkman, P De Zea Bermudez); spatial extremes (KF Turkman, L de Haan), stochastic and epidemic processes (L Carvalho, I Figueiredo, V Andreozzi, M Antunes, AL Papoila, MA Turkman, L Gonçalves), Bayesian statistics (MA Turkman, CD Paulino, G Silva, KF Turkman, S Ramos, M Antunes, P de Zea Bermudez, V Andreozzi, L Gonçalves) survival analysis (AL Papoila, V Andreozzi, G Silva), generalized linear mixed models (V Andreozzi, S Cabral, G Silva, MA Turkman, AL Papoila), latent class models (CD Paulino, L Gonçalves), categorical and incomplete data analysis (CD Paulino, G Silva, S Cabral), multivariate statistics (J Cadima, S Ramos), statistical screening and supervised classification methods (M Antunes, MA Turkman, S Ramos, J Cadima, KF Turkman), Statistics in Genetics (L Sousa, C Silva-Fortes, CD Paulino, S Ramos, MA Turkman).

B. A team who interacts with researchers from environmental sciences and contributes for the activities of the thematic strand on environmental statistics;

C. A team who interacts with researchers from life and health sciences and contributes to the activities of the thematic strand on Biostatistics.

The activities of core teams on methodological work are coordinated by the group leader. The leaders of the thematic strands on environmental statistics and on Biostatistics are both from this research group, and all members are committed to contribute with their time and knowledge towards the success of the initiatives of the thematic strands, participating in the tasks which the thematic leaders assign to them.

9.5.2 Objectives of the Research Group

In consonance with the global strategy of the R&D unit, we set as objectives
1) Do fundamental research in statistical issues regarding the modeling of complex systems and processes that arises in environmental, biological and health sciences, contributing to problem solving regarding activities of the thematic strands.

2) Collaborate with researchers from other scientific areas, either through the participation in national or international projects, or at an individual level, contributing to national and international recognition of our R&D unit.

3) Collaborate with national and international institutions and partners in knowledge transfer activities, through participation in interdisciplinary projects and consultancy activities.

4) Promote advanced training programs through organization of advanced courses and workshops on emerging areas of statistics and supervision of PhD and post-doc students.

5) Seek partners and foment strategic collaboration with other research units from statistics and other fields of science, with the objective of participating in Horizon2020 projects, contributing towards the recognition of our R&D unit as a potential partner in interdisciplinary projects.

The group will continue to do methodological work in the areas of interest and which are relevant for a proper answer to problems brought by the thematic strands where the group is involved.

Here we refer to some of the proposed methodological work, which is foreseen for the prosecution of the objectives of the group.

A) It is known that one of the most important effects of global warming is the increased intensity and sizes of extreme events in Environment ranging from forest fires, rainfall to wind speeds and hurricanes. These extremal processes are by nature temporal and spatial. Max stable processes are particular important in modeling this type of events. Themes to tackle are, e.g., discrete monitoring of a process in continuous time, two dimensional max-stable processes.

B) Environmental and Biological problems show great complexities observed at different time and space resolutions often expressed in terms of thousands of parameters and very large data sets. These complex systems are easier to work through the use of Bayesian hierarchical models and simulation based inference. We propose new developments regarding (i) dynamic beta/Dirichlet regression; (ii) Bayesian hierarchical multivariate age-period-cohort model taking into account the geographical component.

C) Proposals on Statistics in Genetics: Bayesian semi-parametric approach for the differential analysis of sequence counts data; differential analysis of biological sequence count data; development of multiple testing procedures for hierarchical dependent hypothesis; exploitation of statistical methods in Next Generation Sequencing;

D) Some proposals in categorical data analysis: a product-multinomial framework for categorical data analysis with missing responses; semi-parametric Bayesian analysis of binary responses with a MNAR continuous covariate.

E) Bayesian methods: in meta-analysis; as a framework to integrate multiple criteria decision making into health technology assessment; on a nonparametric approach to statistical screening and supervised classification.

F) Novel and alternative geometric interpretation of Mahalanobis distances for multivariate datasets; exploitation of techniques in multivariate data reduction.

G) Stochastic and epidemic processes: models for risk assessment of hospital infections; state space models for flu epidemics and fish stocks;

H) Survival and Longitudinal data analysis: methodologies to analyze discrete longitudinal data with missing data; Beta regression models for proportions for longitudinal/clustered data when a time to event is also of interest. Contributions for R INLA package.

International collaboration expected: M Guindani, P Muller, J Singer, G Molenberghs, A Gelfand, H Rue, W Gonzaléz Manteiga, C Dean, M Schlatter, among others.

10. PROPOSED THEMATIC LINES

Reference	Name	Principal Investigator
TL-6-42	Environmental Statistics	Kamil Feridun Turkman
TL-6-1194	Biostatistics	Carlos Daniel Mimoso Paulino
TL-6-1195	Risk in Financial Extremes	MARIA ISABEL FRAGA ALVES

(TL-6-42) Environmental Statistics

10.1. IDENTIFICATION OF THE THEMATIC LINE

10.1.1 Reference of Thematic Line

TL-6-42

Name of the Thematic Line in Portuguese

Estat?stica Ambiental

Name of the Thematic Line in english

Environmental Statistics

Principal Investigator

Kamil Feridun Turkman

Scientific areas

Agricultural and Forestry Sciences

Marine Sciences and Technologies

Biological Sciences or Environmental Biology

Environmental Sciences

10.2. DESCRIPTION OF THE THEMATIC LINE

Description of the Thematic Line

Environmental issues are at the forefront of concern of all nations. A proper understanding of mechanisms and relationships that govern the environmental issues are vital to reconcile the economic and environmental conflicts. There are many inter-connected themes that need to be addressed by scientist as decision support tools for the decision makers. These topics include, but not restricted to, problems such as environmental monitoring and sampling, measuring levels and consequences of pollution, setting environmental standards, climatological and meteorological issues such as global warming, water resources, management of fish populations and conservation of forests. Two major objectives of such studies are to understand the mechanisms that govern these processes and predict the future realizations. Statistics has a major role to play in such areas by quantifying the effects, assessing consequences, measuring risks and interpreting the evidence.

By nature, almost all environmental problems are of complex dynamic nature over a spatial domain, observed over different time scales and spatial resolutions. The complexity of the environmental problems are augmented by great advances in environmental observation instrumentation in recent years which produce a scale of observations for which classical analytical routines are not designed to handle. As the available information increases, the need for even more complex models also increases, which in return limit the execution of simulation or inference even by the most sophisticated computing systems.

Cox and Isham (1994) broadly define three classes of mathematical models for environmental processes:

• Deterministic spatio-temporal models in which large systems of simultaneous nonlinear partial differential equations are solved numerically. These equations are chosen to represent fairly realistically the physical processes involved.

• Intermediate spatio-temporal stochastic parametric models in which a fairly large number of parameters explain the underlying physical processes.

• Empirical spatio-temporal statistical models where the state of the process is linked to experience or to data observed in previous time and spatial locations and other relevant explanatory variables.

All of these different approaches have an important role to play and their choice depends on the purpose of the study. For example, deterministic models by nature represent the dynamics of the underlying processes much better than the empirical statistical models, but are not capable of quantifying the inherent random fluctuations and their effect on the conclusions which are taken from the simulation studies. On the other hand, empirical statistical models are constructed to quantify the degree of uncertainty inherent in these models, but they are not devised to capture realistically the underlying physical processes. Recently, there has been great interest in producing models that couple the capacities of deterministic and empirical models to overcome these shortcomings, but these challenges are formidable as theyinvolvenumerical solutions for high number of nonlinear stochastic partial differential equations. At present, such idealistic models are far from being operational. Scientists who base their studies on complex deterministic models, such as global circulation models in climatology, would only accept empirical statistical models if the statisticians are capable of producing realistic models for challenging data sets, overcome the associated computational problems and give credible results. However, typically such realistic empirical models involve numerical integrations or matrix inversion operations over hundreds or thousands of unknown model parameters which are far too complex for existing computing technology used for standard statistical models. The seminal paper of Gelfand and Smith(1990) opened way to simulation based inferential methods and the resulting Markov chain Monte Carlo (MCMC) revolution, with which statisticians and other scientists are able to write down incredibly complex models in a compartmental way through hierarchical specification and perform the inference on high number of unknown model

parameters by simulation. These models and inferential methods extend to data which have complex spatio-temporal non-linear dependence structures and show clear non-Gaussian behavior. The bottleneck in these models is twofold and corresponding solutions are now hot topics of environmental statistics. One of the issues is the curse of dimension. Typically the spatio-temporal dependence structure inherent in most data are represented in the model by a latent Gaussian random field and the inferential methods depend on the inversion of the covariance matrix of this process which can run to hundred thousands of rows and columns. Present numerical methods cannot handle such matrix operations. There are two possible solutions to this problem. One is the observation that although covariance matrices are dense, inverses of these matrices are fairly sparse and numerical methods make use of this property. Second approach is to approximate latent Gaussian random fields by Gaussian Markov random fields (Lindgren et al, 2010). Typically this is done by showing that certain classes of Gaussian random fields such as the Matern class appear as solutions to certain stochastic differential equations and these solutions can be approximated by Gaussian Markov random fields using the standard finite elements method.

Other associated problem with these empirical models is that MCMC methodology depends heavily on the analytical expression of the likelihood for the model and in many problems the likelihood cannot be expressed analytically, although one can simulate with reasonable ease from these models. Likelihood free computational methods such as Approximate Bayesian Computation, have been used successfully for such difficult problems. Today these incredible methodological advances in environmental statistics are also accompanied by great achievements in associated computational advances. Programming R and many software packages such as OPENBUGS and INLA bring these methodological advances to the service of scientists.

Summing up, Statistics is trying hard to catch up with ever demanding environmental data, both on two fronts: demand for more complex models and computational problems attached to massive data sets that appear with most problems. In doing so, it is absolutely crucial that statisticians and environmental scientists should collaborate and work closely to produce proper statistical and computational methods that serve the purpose.

10.3. RESEARCH GROUPS INVOLVED IN THE THEMATIC LINE

Reference Name

RG-6-70   Order Statistics, Extremes, Modelling and Simulation

RG-6-74   Statistical Models and Methods for Complex Systems

10.4. ORGANISATIONAL STRUCTURE AND OBJECTIVES OF THE THEMATIC LINE

10.4.1 Structure of the Thematic Line

The organizational structure can be divided into two major groups: internal and external. Internally the complementary competences of the members of each research group will have to be administered to get the best productive results. The internal structure that contains these members aims to combine probabilistic modeling, inferential and computational skills of the team needed in formulating and resolving environmental problems. External organizational structure involves close collaboration with research units and institutes specialized in environmental problems. These units include, but not restricted to, Institute D. Luiz (IDL/FCUL), Portuguese Institute for the Sea and the Atmosphere (IPMA), Forest Research Centre (ISA/UL) and Centre for Environmental and Marine Studies (CESAM/university Aveiro), and Center of Chemistry and Biochemistry (CQB/FCUL).

Typically the research group on Extremes (RG-6-70) will contribute on models for extreme events related to environmental issues and consequent risk analysis. The research group on modeling complex systems (RG-6-74) will contribute on models for spatio-temporal processes and their application in environmental problems and will be responsible for inferential and computational issues that arise in analyzing related data sets. There also be close connections and collaborations with other thematic strands namely Biostatistics and Risk in Financial Extremes, since many statistical and computational issues are common to all thematic strands.

Some of key members of these external institutions are also collaborators of this research unit.

The coordinator of the thematic strand will have the responsibility of identifying problems on environmental issues by collaborating closely and on regular basis with research units on environmental themes. These problems will have to be classified and ordered according to their importance and executability with the capabilities of this research unit. Then for each identified problem the coordinator will form task forces blending necessary skills and make PERT Analysis to define the period of the project and its cost analysis. These identified problems will have priority taking into account possible projects within Horizon 2020. The execution of each task will be closely followed by the coordinator of the thematic strand.

The research unit has significant accumulated knowledge and experience in administering the internal and external organizational structures as explained above through realization of, and participation in significant number of research projects and publications on environmental issues.

10.4.2 Objectives of the Thematic Line

The objectives will be classified according to short-medium term and long term

The short-medium term objectives are well described projects with clear indication what and how to be achieved with the objective of applying for grants within the Horizon 2020 and COST programs of the EU. Some of these proposals

are at the stage of negotiation with National and International partners.

A) Applications in forestry and wild fires

Wildfires, in Portugal, are a relevant public policy issue due to the significant economic and social damage they cause. Contributions in this area will be:

1) Optimal alarm systems for prediction of large fires

We propose to construct an on-line alarm system to predict the event of extreme fires.

Such alarm system, once implemented can be a very valuable tool in the organization of fire fighting capacity and allocation of resources. The major partner of this project is IDL.

2) Fire risk maps as decision support tools

Annual probability maps at different spatial resolutions describing the fire risk across Portugal are important decision support tools in allocation of resources to fight effectively against fires during fire season. The R&D unit has already produced work on these issues jointly with Forestry Center, IDL and at present is part of a COST proposal within an International Network composed of more than 12 partners.

3) Advanced statistical methods for forest surveillance and intervention.

i) beta and Dirichlet regressions for modeling and predicting the proportion of burned area related to forest fires in wild and populated areas, (ii) Bayesian dynamic modelsfor forest surveillance and intervention

B) Applications in fisheries

Portugal identified Sea and its economical values as priority for the coming decade. Statistical methodologies have much to offer in predicting and harnessing fish stocks in a sustainable manner. This research unit has been collaborating with IPMA for many years and will continue with this collaboration through ongoing projects.

C) Applications in water quality

1. Monitoring drinking water for pollutants is an expensive procedure and statistical screening methods can be useful in devising cost-effective monitoring tools. This objective is now integrated into a major Water JPI call on emerging water contaminants with several national and international partners.

2. Setting environmental standards: European Commission intends to define new sets of pollutants in checking for water quality. Setting environmental standards for these new pollutants requires probabilistic assessment of uncertainty and statistical methods to infer on the causal effect of pollutions on the population and environment. Our proposed contribution is to help setting standards on new class of pollutants within the WJPI call, which will be coordinated by our partner in CQB.

D) Applications in meteorology and climatology:

1. Calibration: Many climate processes are simulated by models at high space/temporal resolution. Data also exist at different resolutions. Often the simulated and observed data do not match particularly at extreme observations; it is therefore desirable to calibrate simulated data based on observed data. These methods will specifically be applied to data on wind speeds.

2. Assessing risk in extreme rainfall: The objective is to devise a statistical procedure for extreme rainfall events over time and space with an underlying dependence structure described by suitable anisotropic stochastic processes. This will be a joint project with CESAM and the two research groups.

The long term objectives are based upon the application of methodological advances proposed by the two groups. Therefore, at this moment it is not possible to give a detailed description of individual objectives and proposals. Major long term efforts will be on the application of multivariate max-stable processes and Bayesian hierarchical models and simulation based inference for complex systems.

(TL-6-1194) Biostatistics

10.1. IDENTIFICATION OF THE THEMATIC LINE

10.1.1 Reference of Thematic Line

TL-6-1194

10.1.2 Name of the Thematic Line in Portuguese

Bioestatistica

10.1.3 Name of the Thematic Line in english

Biostatistics

10.1.4 Principal Investigator

Carlos Daniel Mimoso Paulino

Scientific areas

Diagnostic, Therapies and Public Health

Biomedicine

Experimental Biology

Agricultural and Forestry Sciences

10.2. DESCRIPTION OF THE THEMATIC LINE

10.2.1 Description of the Thematic Line

In general terms Biostatistics may be defined as a study field intimately associated with application of Statistical Science to a wide range of topics in the comprehensive domain of the Biosciences, involving many areas from Life and Health Sciences and Natural and Environmental Sciences. Biostatistics, being closely connected with the wider field of Biometry, encompasses issues of design and data analysis concerning experiments from such varied areas as Medicine and Biology (in their broadest sense), Pharmacy and Public Health, Agriculture, Animal Breeding and Forestry, Fishery and Ecology.

The application-multifaceted nature of Biostatistics and the traditional scientific relationship of many members of our research unit with aforementioned domains help explain the choice of Biostatistics as a thematic strand appropriate to bring together a significant number of joint scientific activities involving several other groups and institutions. This thematic line will perform research in a multidisciplinary environment of biostatisticians, physicians, medical and biological scientists, bioinformaticians and health researchers, playing a highly collaborative role in providing statistical assistance to designing experiments and analyzing data, co-authoring ensuing publications, delivering advanced training and co-supervising PhD. theses. Its overall scientific objective is contributing to be applied appropriate statistical methodology, either known or novel (to be developed within the unit at the level of the research groups) so as to achieve the goals of specific projects proposed by members of our research unit or then by external researchers to some of these members.

Taking into account:
- The diversity of scientific areas and subareas involved in these research programs proposed by or to members of the unit, and accepted by them;
- The need of taking better advantage of the knowledge and skills of its members, as well as their interrelations and common interests;
the scientific council of the research unit has decided that the most efficient way to attain its strategic objectives would be to adopt an organizational framework based upon thematic lines and work specific programs across both research groups and members integrated into the unit.

The scientific programs embraced within this thematic strand are of several types. In terms of scope, there are national and international projects, involving or not formal consortia, and a few of them have already got funding. Some of them are underway but extending over the period 2015-2020 and other are proposed to start from 2014. Regarding their origin, most refer to projects or plans proposed by other research units and groups, or other kind of public or private institutions, which are mentioned in the part of organizational structure. These programs are expected not to exhaust the joint scientific activity of the set of line members once there is short-term potential of further scientific interrelations and the previous experience has shown that the requests for statistical collaboration within our unit arise frequently.
To convey a more concrete idea of the nature and purposes of the set of currently planned scientific programs, some additional information about a few selected projects is given in terms of scientific areas/subareas embracing them and respective general objectives. Further information is given in part 10.4.
A. Diagnostics, Therapeutics and Public Health /Epidemiology
Ongoing collaborative work on international projects around epidemiology of malaria, run in African countries, which are funded by institutions (such as Wellcome Trust, Royal Society Leverhulme Trust) and integrated into a genomic epidemiology network (MalariaGEN). Their objectives are studying the malaria transmission dynamics, determining the relationship between genomic variations and immunity and identifying genetic determinants associated with severe malaria.
B. Biomedicine / Oncology
Study of investigating the prognostic value of biomarkers and other risk factors to overall survival and remission times in populations of patients suffering from different types of carcinoma. In this oncologic study it is intended to apply novel statistical models accounting for the incomplete nature of some data, to be developed by a subset of the Group RG-6-74.

C. Anthropology /Bioanthropology
Collaboration in a national project on medico-legal age estimation in non-identified cadavers by aspartic acid racemization. For this purpose it is intended to construct a regression model that will be able to provide useful guidance to forensic dentistry, especially in what concerns

identification of cadavers, namely in procedural international cooperation protocols.

D. Agriculture and Forestry /Plant Biology /Genomics

Ongoing collaboration in a national project, funded by FCT, involving another research unit, with the goal of applying statistical methodologies that allow the identification of genes differentially expressed in data concerning cork oaks coming from NGS (Next Generation Sequencing) technologies.

E. Clinical Research / Treatment Regimes

Ongoing collaboration with researchers from national and international institutions on a new evidence-based approach to optimize therapeutic decisions applied to patients with rheumatoid arthritis receiving biotechnological therapies (having already got the Pfizer Award in Clinical Research 2013).

Projects in Diagnostics, Therapeutics and Public Health:

F. Public Health and Environmental Factors

Ongoing international research project, financed by FCT, aiming to provide critical information for understanding how socio-economic-cultural inequalities and the urban morphology affect the health of the residents in several areas of the city of Praia, Cape Verde.

G. Public Health and Environmental Factors

Ongoing collaboration in a national project funded by FCT, involving several institutions, on health effects of air quality and ventilation in elderly care centers. The ultimate objective is to define and disseminate recommendations for improvement of these conditions affecting health of the users of these services.

H. Healthcare Service Improvement

European project PERSSILAA funded by FP7 involving a consortium of partners, from Netherlands, Italy, Ireland, Spain and Portugal (through a CEAUL team for statistical analysis), with varied skills merging social, medical and technological sciences with industry, academia and end user organisations. It aims to develop and validate a new service model, to screen for and prevent frailty in community dwelling older adults, integrating nutrition, physical and cognitive function.

I. Public Health and Environmental Factors

Maintaining the collaboration with Brazilian institutions (namely FIOCruz, Rio de Janeiro), statistical support will be given to solving complex problems around ELSA (Longitudinal Study of Adult Health).

10.3. RESEARCH GROUPS INVOLVED IN THE THEMATIC LINE

Reference Name

RG-6-70   Order Statistics, Extremes, Modelling and Simulation

RG-6-74   Statistical Models and Methods for Complex Systems

10.4. ORGANISATIONAL STRUCTURE AND OBJECTIVES OF THE THEMATIC LINE

10.4.1 Structure of the Thematic Line

The organisational structure presents two levels, labelled internal and external, as follows. Internal organisation aims to take advantage of the experience and complementary skills of the members involved in this strand to the modelling, inference and computation challenges posed by the several applied problems addressed to each team. External organisation has to do with the close collaboration with external researchers and units who propose joint programs and work plans to subsets of its members or then to the overall strand. In the latter case, the strand coordinator will identify, classify and sort by priority the proposals according to their scientific and practical relevance and capability of the research groups.

For a more complete idea of the nature and diversity of the external institutions that are expected to collaborate with this strand, we list some of those associated with programs already proposed to individual researchers or teams involved in this thematic line, either mentioned in 10.2 or to be referred to herein.

A. London School of Hygiene and Tropical Medicine, University of London/ Liverpool School of Tropical Disease

B. Unit of Pathologic Anatomy, Portuguese Institute of Oncology, Lisboa

C. Public Health Schools/ Faculty of Dental Medicine/ National Institute of Legal medicine

D. Center of Biodiversity, Integrative and Functional Genomics - Univ. Lisbon/ Univ. Minho

E. Rheumatology Unit - Hospital Garcia de Orta, Almada/ Gulbenkian Institute for Science/ Sanquin Research Institute, Immunopathology Dept., Amsterdam

F. Instituto de Higiene e Medicina Tropical/ Faculdade de Ciências e Tecnologia, Univ. Nova de Lisboa/ Cape Verde Health Ministry.

G. Instituto de Saúde Pública-Univ. Oporto/ Centro de Estudos de Doenças Crónicas/ Faculdade de Ciências Médicas-Univ. Nova Lisboa/ Instituto Nacional de Saúde Ricardo Jorge-Porto/ Laboratório Nacional de Engenharia Civil.

H. University of Twente/ University of Ireland-Cork/ Universidad Politecnica de Madrid

I. Fundação Oswaldo Cruz - Fiocruz, Rio de Janeiro, Brazil

Here it is a listing of further projects with mention to their purposes and scientific areas embracing them as well as external institutional partners:

Biomedicine/ Disease mechanisms
1. Detection of key genetic polymorphisms associated with multiple sclerosis through statistical analysis of the genotype-phenotype data: Gulbenkian Institute for Science/ ICBAS, Univ. Oporto
GeoSciences/ Nature Risks on Health
2. Effect of cold waves on the cardiovascular disease mortality: Centro de Estudos em Geografia e Ordenamento do Território - Univ. Coimbra

Experimental Biology/ Genetics & Genomics
3. Study on the detection of Intragenic initiation in cells from Next Generation Sequencing Data: Institute of Molecular Medicine - Faculty of Medicine, Univ. Lisbon

Immunology and Infection
4. Assessment of influence of concomitant immuno-suppressive therapies on anti-drug antibodies production in patients with Crohn's Disease: Group of studies on intestine inflamatory disease and Gastrenterology Unit of regional hospitals

Diagnostics, Therapeutics and Public Health/

Epidemiology/ Disease Mapping
5. Space-time analysis toward disease mapping: Public Health Institute, Univ. Oporto/ Faculty of Science at the University of Western Ontario, Canada/ Regional Oncologic Registry - South Portugal

Pharmacology &Toxicology

6. Epidemiology and diagnosis of renal disease and nephrotoxic effects of antiretrovirals in the HIV population (funded by FCT): Hospital Amadora-Sintra/ Faculdade de Ciências Médicas-Univ. Nova Lisboa/ Nephrology Dept., IIS-Fundacón Jiménez Díaz, Madrid

Healthcare quality assessment
7. Rehabilitation and quality of life in residential facilities for people with long-term mental illness for assessment of quality of care of these services and effectiveness of ensuing interventions on training of professionals: Mental Health Dept., Faculdade Ciências Médicas-UNL/University College, London.

10.4.2 Objectives of the Thematic Line

In the foregoing material concerning this thematic strand, several projects and scientific programs have been described with reference to their objectives and external partners. Others will be mentioned below once they also have already been proposed to individual members and teams of this research group.
Obviously, it is not expected all this list does exhaust the projects to be proposed within Horizon 2020, given the close relationship of this research unit with many national and international groups and institutions. Some of them have already expressed their interest and strong will to set up a collaborative scheme that ensures fulfilment of their purposes embodied in problem solving, scientific production and advanced training. We may refer to the specific examples of the Chronic Diseases Research Center from Faculdade de Ciências Médicas (Univ. Nova Lisboa), Comissão de Qualidade e Segurança do Doente do Centro Hospitalar de Lisboa Central and groups from Universidade de São Paulo and Universidade Federal de Lavras, Brazil.

Listing of further collaborative studies:

Animal Science and Veterinary Sciences/ Animal Health and Epidemiology
8. Impact of new treatments for immunodeficiency in infected cats: Faculty of Veterinary Sciences, Univ. Lisbon.

Biochemical Sciences/ Biochemistry

9. Statistical analysis of qualitative and quantitative data generated by a liquid chromatography-mass spectrometry technique in order to identify and quantify the metabolites characterizing the situation of infection (vine) and symbiosis (cork oak): Center of Biodiversity, Integrative and Functional Genomics - Univ. Lisbon.

Immunology and Infection
10. Optimizing the CD34 cell collection process: Imunohemotherapy Unit - Hospital Santa Maria, Lisboa.

Neurosciences, Ageing and Degenerative Diseases/ Neurophysiology
11. Finding prognostics factors and impact of new treatments for neurophysiology diseases: Clinical & Translational Physiology Unit - Molecular Medicine Institute/ Physiology Institute, Faculty of Medicine, Univ. Lisbon.

Diagnostics, Therapeutics and Public Health/

Epidemiology
12. Pursuing the ongoing work on modelling flu epidemics, towards using state space models to estimate the incidence of this disease based on data collected by syndromic surveillance systems: Direção Geral de Saúde/ Faculty of Sciences and Technology-Univ Nova Lisboa.
Healthcare Assessment
13. Comparing the effectiveness of educational outreach visits in implementation of clinical guidelines in primary care against usual strategies and conducting a cost-effectiveness analysis of this method: Portuguese National Health Institute Dr. Ricardo Jorge.

Public Health, Physical Exercise, Nutrition
14. Study of adherence to exercise, nutrition and pharmacological therapies by elderly in council of Faro towards improving of the type 2 diabetes patients' self-care (initiative within the ambit of European Innovation Partnership on Active and Healthy Ageing): Univ. Algarve/ Administração Regional de Saúde do Algarve/ Camara Municipal de Faro.

Healthcare Improvement
15. Assessment of factors associated with depression, anxiety, pain and quality of life in cancer patients during chemotherapy in order to develop and implement programs and intervention strategies aimed at people at greater risk (project awarded with a grant Terry Fox): Hospital Amadora-Sintra.

Healthcare Improvement
16. Project (funded by FCT) on organizational and informational system to improve the comunication management of healthcare associated infections in hospitals: Institute of Hygiene and Tropical Medicine.

(TL-6-1195) Risk in Financial Extremes

10.1. IDENTIFICATION OF THE THEMATIC LINE

10.1.1 Reference of Thematic Line

TL-6-1195

10.1.2 Name of the Thematic Line in Portuguese

Risco em Extremos Financeiros

10.1.3 Name of the Thematic Line in english

Risk in Financial Extremes

10.1.4 Principal Investigator

MARIA ISABEL FRAGA ALVES

10.1.5

Scientific areas

Economics

Finance

10.2.1 Description of the Thematic Line

In fields such as in insurance and finance, the stochastic modelling of extremes is of major importance. There are innumerous motivating examples: if we think for instance of notions as very large claims with catastrophic consequences for Insurance or Reinsurance companies, Value-at-Risk (VaR), probable maximal loss or other crucial measures for financial investors, and even the Pareto law for which portfolios with a small percentage representing large claims typically explaining for the total claim amount. A comprehensive textbook treatment of extremes in insurance and finance is Embrechts, Klüppelberg and Mikosch (1997), including an extensive list of references therein.

Moreover, recurrent black swans financial events are of major concern for both investors and regulators because of the extreme price changes they cause, despite their very low probability of occurrence. These events exist not only in finance, economics, insurance, but also in environment, earth sciences, among others. In economics and finance, the worst-case events have become more frequent than before, while they have kept their devastating impacts. Examples of recent financial extreme events include the Black Monday of the stock market crash that occurred on October 19, 1987, the turmoil in the bond market in February 1994, the 1997 Asian financial crisis, and the 2007/2009 global financial crisis. These crises are a major concern for regulators and owners of financial institutions because of their heavy consequences. Many economists and financial analysts have shown increasing interest in examining the behavior of financial markets, testing financial stress and managing risks during those events. Under such circumstances, significant and extreme drops in prices and returns of these assets have become more probable, with potentially damaging consequences on portfolios of individuals and institutions. These circumstances have also made risk management strategies for these high flying commodities and highly volatile stocks more challenging, particularly as the percentages of violations of confidence targets have compounded. The quantification of the potential size of losses and assessing risk levels for individual stocks and portfolios composed of them is fundamental in designing prudent risk management and portfolio strategies. Value-at-Risk (VaR) models have become an important instrument within the financial markets for quantifying and assessing market downside risks associated with financial and commodity asset price fluctuations. They determine the maximum expected loss an asset or a portfolio can generate over a certain holding period, with a pre-determined probability value. Thus, a VaR model can be used to evaluate the performance of portfolio managers by providing downside risk quantification, together with asset and portfolio returns. It can also help investors and portfolio managers to determine the most effective risk management strategy for a given situation. Moreover, quantification of the extreme losses in asset markets is important in the current market environment. Extreme value theory (EVT) provides a comprehensive theoretical forum through which statistical models describing extreme scenarios can be developed.

Three recent updated overviews on the subject are Chang, Allen and McAleer (2013), Hammoudeh and Michael McAleer (2013) and Chang, Allen, McAleer and Amaral (2013).At this stage, it is worthwhile to mention that in Tsay (2010) the whole chapter 7 is devoted to Extreme Values, Quantiles, and Value at Risk.

A special topic under semi-parametric approach to extreme values statistical inference concerns the VaR estimation. Although the classical VaR estimators based on an intermediate number of top-order statistics, such as Weissman-Hill estimators, do not enjoy the adequate linear property of quantiles, recently had appeared the corresponding PORT-VaR and quasi-PORT reduced-bias VaR estimators enjoying a linear property; they are based on a partially shifted version of a Minimum-Variance Reduced-Bias (MVRB) estimator of the Extreme Value Index (EVI), the primary parameter in Statistics of Extremes. Associated to these new methodological research, there has been in parallel practical investigation about a possible heuristic stability criterion for the choice of the convenient threshold, providing applications of the methodology to simulated data and to log-returns of financial stocks. In fact, if we think of statistical inference for VaR, the extremal quantile estimation constitutes an unavoidable problem. In this respect, heavy tailed models, i.e., models with positive EVI for which the tail decays at a polynomial rate, play a central role in financial related fields and bias reduction techniques for heavy tails have been revealed very useful for reliable estimation, and several applications have been under analysis, such as log-exchange rates or financial stocks indexes.

In Gomes and Pestana (2012) the authors deal with the semiparametric estimation of VaR for heavy tails, using bias reduction techniques. Bias-corrected estimation of high quantiles has also been handled in Matthys and Beirlant (2003).

In the field of non-life insurance it is also of major importance the most convenient modelling for large claims, from a specific portfolio of policies; in fact, a good diagnostic to model large claim sizes is determinant for a reliable estimation of interesting characteristics in insurance setup, such as ruin probabilities or reinsurance premium, for instance. For a good seminal overview of this subject under the EVT approach we refer to Beirlant and Teugels (1992).

As mentioned before, Pareto distribution is widely used in modeling large claims, which are captured by the heavy tailed nature of this model. However, it should be mentioned that in practice sometimes it is used an upper truncated version, which ensures that the moments of the distribution exist. Concerning this subject, we mention Aban, Meerschaert and Panorska (2006) and Clark (2013), which provides some background on the characteristics of the upper-truncated Pareto distribution and suggests some diagnostics, based on order statistics, to assist in selecting the upper truncation point.

REFERENCES A (see subsection 10.4.2).

10.3. RESEARCH GROUPS INVOLVED IN THE THEMATIC LINE

Reference Name

RG-6-70   Order Statistics, Extremes, Modelling and Simulation

10.4. ORGANISATIONAL STRUCTURE AND OBJECTIVES OF THE THEMATIC LINE

10.4.1 Structure of the Thematic Line

The organizational structure will reflect the natural interactions of Group 1, embracing both internal and external cooperation with national and international research groups, with the main task of collaborative publication, research, or other forms of participation of the team at the international level.

Concerning the thematic strand on Risk in Financial Extremes, it is intended to keep the international contacts in the field, pursuing a potential collaborative work with other partners: Shawkat Hammoudeh (Drexel University, US), Abdullah Al-Hassan (International Monetary Fund, US), Michael McAleer (Erasmus University Rotterdam, NL), Juan Ángel Jiménez (Universidad Complutense de Madrid, ES), Teodosio Pérez Amaral (Universidad Complutense de Madrid, ES), Jan Beirlant (KatholiekeUniversiteit Leuven, BE) among other potential collaborations in the future. At a national level, it is also a policy of this strand to join competences of its members with collaborators integrated in other FCT research units, profiting of complementary skills oriented to identify the convenient statisticalapproach to risk analysis forproblems in extreme values under financial context; these foreseen collaborations include researchers from University of Aveiro (Claudia Neves), CEMAPRE (UL), CMA (FCT,UNL). Moreover, the members of this research thematic strand have a significant historical knowledge in organizational structures through realization of thematic meetings and a rich portfolio of publications on EVT as a comprehensive theoretical tool for modelling extreme events, including financial and insurance issues, either at particular networks level or integrated in research projects.For a very brief set of recent publications and international meetings relevant for the knowhow of the team see REFERENCES B.The coordinator of this thematic strand will make the necessary efforts to make the team members regularly updated with the undergoing research projects.

REFERENCES B

Araújo Santos, P. and Fraga Alves, M. I. (2012). A new class of independence tests for interval forecasts. Computational Statistics & Data Analysis, 56, Issue 11, 3366- 3380.

Araújo Santos, P., and FragaAlves, M. I. (2013). Forecasting Value-at-Risk with a duration-based POT method (2013). Mathematics and Computers in Simulation, 94, 295-309.

Beirlant, J., Caeiro, F. and Gomes, M.I. (2012). An overview and open research topics in the field of statistics of univariate extremes. Revstat 10:1, 1-31.

Hammoudeh, S., Araújo-Santos, P., and Al-Hassan, A. (2013). Downside risk management and VaR-based optimal portfolios for precious metals, oil and stocks. North American Journal of Economics and Finance, 25, 318-334.

Figueiredo, F., Gomes, M. I.,Henriques-Rodrigues, L. and Cristina Miranda, M.C. (2012). A computational study of a quasi-PORT methodology for VaR based on second-order reduced-bias estimation. Journal of Statistical Computation and Simulation, 82, 587-602.

Caeiro, F. M., Gomes, M.I. (2012). Semi-parametric second-order reduced-bias high quantile estimation. Test, 18(2), 392-413.

Gomes, M.I., Mendonça, S.andPestana, D. (2011). Adaptive Reduced-Bias Tail Index and VaR Estimation via the Bootstrap Methodology. Communications in Statistics-Theory and Methods, 40, 2946-2968.

Figueiredo, F., Gomes, M.I. and Neves, M.M. (2011). Adaptive reduced bias estimation of financial log-returns. In FragaAlves et al. (eds.). Risk and Extreme Values in Insurance and Finance: Book of Abstracts, CEAUL editions, 43-46.

Fraga Alves, M.I., Gomes, M.I., deHaan, L. and Neves, C. (2011).Risk and Extreme Values in Insurance and Finance: Book of Abstracts. Edições CEAUL, ISBN: 978-989-8203-73-1, Depósito Legal: 325764 / 11 (x + 120 pages).

Workshop on Risk & Extreme Values in Insurance and Finance - University Lisbon, 2011, June, 6 - 7 Organized by CEAUL, Faculty of Sciences, UL.

10.4.2 Objectives of the Thematic Line

The focus of this thematic strand is aimed to build a bridge between Extreme Values and Applications in Finance, Economics and Insurance, under the EVT framework. Risk in Financial Extremes aims to investigate risk modelling in financial fields, taking advantage of recent probabilistic developments and statistical tools aimed to handle with extremes and rare events, both by the theoretical side and by applications to case studies. This thematic strand will join together specialists working on extremes and rare events, fulfilling an exchange program between expert researches, helping to create a network of knowledge, ideas and experience.

Some team members of this strand will pursue with the study of unconditional and conditional methods, such as the high quantile EVT models of DPOT (Duration-based Peak over Threshold) and quasi-PORT (Peaks Over Random Threshold), to estimate the Value-at-Risk with very small probability values for adequately long financial time series to obtain a reasonable number of violations for backtesting, which has revealed a very powerful framework to build

models for forecasting one-day ahead VaR. It is our goal to go on with comparative studies with global stock market indices, to provide evidence that the developed models can perform better than the state-of-the art for risk models and also better then the widely used RiskMetrics model in terms of unconditional coverage, clustering of violations and capital requirements.

Another strong mainstream for some team members is to obtain new probabilistic and statistical developments for the particular case of heavy tailed models, since these models play a central role in financial related fields, using bias reduction techniques aiming to attain reliable estimation, for real data in financial and insurancefields.

In the field of non-life insurance it is also of major importance the convenient modelling for large claims, from a specific portfolio of policies;somehow related to this topic, it constitutes also a research objective of some team members the study of statistical inference procedures for the upper-truncated Pareto distribution, both concerning estimation of parameters of interest suggests some diagnostics, based on order statistics, to assist in selecting the upper truncation point.

REFERENCES A

Paul Embrechts, P. Claudia Klüppelberg, Thomas Mikosch. Modelling Extremal Events for Insurance and Finance, Springer-Verlag, Berlin, 1997. 4th Printing2002, Edition for China, 2003.
Chia-Lin Chang, David Allen and Michael McAleer. Recent developments in financial economics and econometrics: An overview, The North American Journal of Economics and Finance, In Press, Corrected Proof, Available online 13 March 2013.
ShawkatHammoudeh, Michael McAleer (2013). Risk management and financial derivatives: An overview, The North American Journal of Economics and Finance, 25, 109-115.
Chia-Lin Chang, David E. Allen, Michael McAleer and Teodosio Perez Amaral (2013). Risk modelling and management: An overview. Mathematics and Computers in Simulation, 94, 159-163.
Ruey S. Tsay (2010). Analysis of Financial Time Series. 3rded. John WIiley& Sons. New Jersey.
M. Ivette Gomes, Dinis Pestana (2007). A Sturdy Reduced-Bias Extreme Quantile (VaR) Estimator. Journal of the American Statistical Association, 102, 280-292.
G. Matthys and J. Beirlant (2003). Estimating The Extreme Value Index And High Quantiles With Exponential Regression Models. StatisticaSinica 13, 853-880. • Beirlant, J. and Teugels, J.L. (1992). Modeling large claims in non-life insurance. Insurance: Mathematics and Economics,11, 17-29.
Aban, M. Meerschaert, A. Panorska (2006). Parameter Estimation for the Truncated Pareto Distribution, Journal of the American Statistical Association, Vol. 101, 270-277.
Clark, David R. (2013). A Note on the Upper-Truncated Pareto Distribution. Casualty Actuarial Society E-Forum, Winter 2013, Vol. 1, 1-22.

11. BUDGET FOR THE STRATEGIC PROGRAMME 2015/2020

11.1 The unit is a candidate for evaluation and intend to apply for funding?

Yes

11.2

Host institution's budget

Funda??o da Faculdade de Ci?ncias (FFC/FC/UL)

Description	2015	2016	2017	2018	2019	2020	TOTAL (€)
Human Resources	0	53620	53620	71660	53620	18040	250560
Missions	72000	72000	72000	72000	72000	72000	432000
Consultants	7200	7200	7200	7200	7200	7200	43200
Service Procurement and Acquisitions	21280	21280	21280	21280	21280	21280	127680
Patent Registration	0	0	0	0	0	0	0
Adaptation of Buildings and Facilities	0	0	0	0	0	0	0
Overheads	26920	42125	43025	46635	42125	30530	231360

| Equipment | 7200 | 14400 | 18000 | 14400 | 14400 | 3600 | 72000 |
| TOTAL (€) | 134600 | 210625 | 215125 | 233175 | 210625 | 152650 | 1156800 |

Overall budget

Description	2015	2016	2017	2018	2019	2020	TOTAL (€)
Human Resources	0	53620	53620	71660	53620	18040	250560
Missions	72000	72000	72000	72000	72000	72000	432000
Consultants	7200	7200	7200	7200	7200	7200	43200
Service Procurement and Acquisitions	21280	21280	21280	21280	21280	21280	127680
Patent Registration	0	0	0	0	0	0	0
Adaptation of Buildings and Facilities	0	0	0	0	0	0	0
Overheads	26920	42125	43025	46635	42125	30530	231360
Equipment	7200	14400	18000	14400	14400	3600	72000
TOTAL (€)	134600	210625	215125	233175	210625	152650	1156800

12. STRATEGIC PROGRAMME AND BUDGET RATIONALE 2015/2020

12.1 Overall budget rationale

The basic needs of this R&D are in human resources, missions, consultants, services and acquisitions. Equipment needs are basically restricted to computing facilities. The major expenses in human resources are due to post-doc and PhD students procured for the specific objectives of the R&D unit. The unit will continue receiving PhD students and post-doc who apply through the usual channels to work on themes related to the objectives of the groups. However, for the prosecution of the objectives of the thematic strands, we seek for students who will work on very specific problems. The rationale behind the human resources and equipment is given in detail below.

In order to follow recent trends in the scientific area and collaborate fully with international partners, missions are the most important part of the budget. We plan at least one international mission per person per year at an expected cost of 1800 per person per year, which includes registration, travel expenses and accommodation. We also expect to receive approximately six collaborators and members of the advisory board per year at an expected cost of 1200 euros per visit. As it has been in the past, a policy of this R&D unit is to finance PhD students to attend to conferences, workshops and advanced courses. This is usually done by supervisors, using the budget for missions allocated to them.

Due to the extensive use of computing facilities, the R&D unit needs to acquire the services of a computer technician on regular basis. At present rates this is estimated at 800 euros per month. Other services include repairs of computer equipment and purchase of nondurable materials, such as toners, printing material, USB devices, connecting cables, ink cartridges, paper, photocopies and other current expenses relevant to the development of the research. We expect that during the duration of the project, the overhead paid to the host institution will be kept at 20%.

12.2 Human Resources rationale

For the proposed strategic plan, the R&D unit will require 2 post-doc and 3 PhD students over the planned six year period according to the following table:

2015 - 0 post-doc and 0 PhD; with total estimated cost 0 euros
2016 - 1 post-doc and 3 PhD; with total estimated cost 53620 euros
2017 - 1 post-doc and 3 PhD with total estimated cost 53620 euros
2018 - 2 post-doc and 3 PhD; with total estimated cost 71660 euros
2019 - 1 post-doc and 3 PhD; with total estimated cost 53620 euros
2020 - 1post-doc and 0 PhD; with total estimated cost 18040 euros

It is expected that each post-doc will stay for three years. Hence the post-doc we hire at the beginning of on 2016 will leave at the end of 2018 and the one we hire at the beginning of 2018 will stay till the end of 2020. Each PhD student is expected to stay for a maximum of four years. At the beginning of 2016 we would hire 3 PhD students who would be expected to finish the PhD by the end of 2019 and on 2017 we would hire an extra PhD student who would leave by the end of 2020.

These costs are estimated according to the present day rates applied by FCT, namely 1495 euros monthly expenses + 100 euros yearly insurance for a post-doc student and 980 euros monthly expenses + 100 euros yearly insurance for a PhD student.

12.3 Equipment rationale

The R&D unit basically needs computing facilities (laptops, desk computers, printers, etc); we calculated that during the period of the execution of the project that each member will spend approximately 1800 euros. The purchase of this equipment will be phased according to individual needs. Computing facilities are also expected to be given to PhD students and pos-docs. So, as for missions, the budget allocated to an integrated member for equipment at his/hers request may be spent for the benefit of their students.

13. REVIEWERS PROPOSED BY THE R&D UNIT(S)

13.1

Name	Institution	Email	Scientific Areas
Richard A. Davis	Department of Statistics, Columbia University, New York, USA	rdavis@stat.columbia.edu	Mathematics
Jef (Jozef) Teugels	Katholieke Universiteit Leuven, Belgium	jef.teugels@wis.kuleuven.be	Mathematics
Leonhard Held	University of Zurich, Institute of Social and Preventive Medicine, Switzerland	leonhard.held@ifspm.uzh.ch	Mathematics