The Open University Statistics Group

2011 Report

1. Membership

Professor P.H. Garthwaite (Head of Group)
Dr K. Anaya (to July)
Professor F. Critchley
Dr A.E. Faria
Professor C.P. Farrington
Dr D. Gragn (from March)
Dr R.M. Hilliam (from November)
Dr G.E. Iossif
Professor M.C. Jones
Professor K.J. McConway
Dr A. Noufaily (from March)
Dr C.M. Queen
Dr N.T. Trendafilov
Dr S. Unkel
Dr S.K. Vines
Dr H.J. Whitaker
Mrs J.E. Williams

Long-term visitors

Dr K. Anaya [Visiting Research Fellow] (from July)
Professor J.C. Gower [Visiting & Emeritus Professor]

Full-time research students

Mr O. Anacleto
Mr T.E. Collins [joint with Computing] (to September)
Mr F. Elfadaly
Mr D. Gragn (to February)
Ms L. Gunn [joint with Earth Sciences] (from October)
Ms A. Noufaily (to February)
Mr O. Oyebamiji [joint with Earth Sciences] (from October)
Mr A.J. Santos
Ms S. Villers (from October)
Mr Y.G. Weldeselassie

Part-time research students

Mr N. Calleja
Mr G. Kafatos (to June)
Mr A. Owen (to January)

Support staff

Mrs S. Frain, Group Secretary (to March)
Ms E. Howard (from March)
**Visitors**

Dr N. Al-Kandari (University of Kuwait)
Dr R. Kuhnert (Robert Koch Institute, Berlin, Germany)
Professor P.K. Marriott (University of Waterloo, Canada)
Ms C. Monteiro (Pontificia Universidade Católica do Rio de Janeiro, Brazil)
Professor H. Oja (University of Tampere, Finland)
Dr M. van de Velden (Erasmus University, Rotterdam, Netherlands)

### 2. Introduction

2011 was another extremely busy and successful year for Statistics at the Open University in both research and teaching.

It was the final year of production for both our new web-based module M347 *Mathematical Statistics* and the revision of M343 *Applications of Probability*. The Group's four modules in presentation attracted a total of over 1800 students in 2011.

In the course of the year, members of the Group published a book and many papers in refereed journals (see Section 6). Pre-publication papers can be found in our technical report series, at http://statistics.open.ac.uk/technical_reports. Members of the Group gave many presentations and took part in numerous international conferences. We also added the 24th instalment to our long-standing series of Open University conferences in Statistics. PhD students graduated in record numbers.

Congratulations to Paddy Farrington for his achievement in obtaining a highly prestigious *Royal Society Wolfson Merit Award*. Paddy also had a high profile position in the Royal Statistical Society (RSS) in 2011.

We are very pleased to welcome Rachel Hilliam to the group, as Staff Tutor in the OU's East Midlands region, part-time at first, changing to full-time early in 2012. In the other direction, Karim Anaya has left us for a lecturing post at the London School of Hygiene and Tropical Medicine; we wish him well and are delighted that he still works with us in a visiting capacity.

The Statistical Advisory Service, which supports research in other departments across the Open University, continues to run successfully. It is led by Alvaro Faria, with assistance from Paul Garthwaite. A website explains the service: http://www.mathematics.open.ac.uk/advisory.

Further information on the Group is available through our web pages at http://statistics.open.ac.uk

### 3. Teaching

**Presentation**

The Group presented four modules of 30 CATS points each, namely, M248 *Analysing Data* (Chair: Alvaro Faria), M249 *Practical Modern Statistics* (Chair: Karen Vines), M343 *Applications of Probability* (Chair: Gillian Iossif), and M346 *Linear Statistical Modelling* (Chair: Paul Garthwaite).
Each of the first level mathematics modules, MU123 Discovering Mathematics and MST121 Using Mathematics, contains statistical material, MST121 substantially so (roughly one quarter of the module); Nickolay Trendafilov was involved with the latter. Kevin McConway contributed in a small way to the presentation of SDK125 Introducing Health Science.

Production

The Group continued production of M347 Mathematical Statistics. As well as enhancing our third level provision with a new module that gives “the theory behind the methods”, this module is notable for leading the way within the Department of Mathematics and Statistics in terms of electronic presentation: it will be presented entirely on-line. Catriona Queen and Chris Jones were production team co-chairs and lead authors; Paddy Farrington made several considerable contributions; Heather Whitaker produced the screencasts; and Robert Hasson (Applied Mathematics Group) had a very important software development role, facilitating the interface between the authors’ Latex drafting and the University’s Virtual Learning Environment.

Gillian Iossif led, or perhaps more accurately did, the revision of M343 Applications of Probability. This very successful module needed a certain amount of improvement and updating. Gillian single-handedly authored the entire revision. In 2011, she had help on some aspects of production from Paddy Farrington.

Gillian also revised the M248 Analysing Data computer books for use with a newer version of Minitab.

The Group also made initial preparations for its next production, M140 Introduction to Statistics. This is a first-year statistics module which will have a leading role in the brave new world of revised qualifications.

All the modules mentioned in this subsection are 30 CATS points.

Qualifications

Our modules are compulsory components of degrees in Mathematics and Statistics, Computing & IT and Statistics, Mathematics and its Learning and Economics and Mathematical Sciences, as well being specified options in certain other qualifications. Fifty seven students were awarded the Mathematics and Statistics degree in 2011. The undergraduate Diploma in Statistics is awarded to students who successfully study M248, M249, M343 and M346 (or their predecessors). The Diploma was awarded to 69 students in 2011.

4. Ph.D.s awarded

Tom Collins for his thesis Improved methods for pattern discovery in music, with applications in automated stylistic composition (supervisors: R.Laney (Computing), A. Willis (Computing) and P.H. Garthwaite)

Doyo Gragn for his thesis Interpretable and fast dimension reduction (supervisors: N.T. Trendafilov and F. Critchley)

George Kafatos for his thesis Statistical analysis of quantitative seroepidemiological data (supervisors: C.P. Farrington and K.J. McConway)
Angela Noufaily for her thesis *Parametric quantile regression based on the generalised gamma distribution* (supervisors: M.C. Jones and H.J. Whitaker)

Alun Owen for his thesis *Development of Bayesian dynamic forecasting models of football match outcomes* (supervisors: S.K. Vines and K.J. McConway)

5. Research interests

The Department is home to three Research Groups:

*Multivariate Statistics*: Critchley (co-ordinator), Anaya, Faria, Garthwaite, Gower, Jones, Queen, Trendafilov, Unkel, Vines, Constantine (CSIRO, retired), Cook (University of Minnesota), Marriott (University of Waterloo), Taylor (University of Hertfordshire).

*Bayesian Statistics*: Queen (co-ordinator), Faria, Garthwaite, McConway, Vines, Whitaker.

*Medical Statistics*: Whitaker (co-ordinator), Farrington, Garthwaite, Gragn, Noufaily, Unkel, Vines, Weldeselassie.

Clearly, these three general groupings cover a huge variety of research projects in Statistics. Research also continues, of course, on topics not directly covered by these Research Groups, of which single case studies, distribution theory, kernel smoothing, robust statistics, time series forecasting and influence analysis are but a small selection. Strong ongoing collaborations continue in areas such as medicine, psychology, traffic modelling and ecology.

For details, see the publications, talks and other items to follow.

6. Publications

6.1 Book


6.2 Publications in refereed academic journals


6.3 Refereed conference proceedings and letter to the editor


6.4 Outreach and other publications


7. Seminars and conferences at The Open University

In May, the Statistics group organized its 24th approximately annual one-day conference, this one being on *Visualisation and Presentation in Statistics*. The organising committee comprised Sarah Frain, Sara Griffin (secretary in Mathematics & Statistics Department), John Gower and Steffen Unkel. Along with an introduction by John Gower, speakers at the conference were:

John Aldrich (University of Southampton) *Graphs before graphics – some history*

Rosemary Bailey (Queen Mary University of London) *Bad statistics*

Martin Bland (University of York) *Reporting clinical trials with confidence*

Michael Blastland (freelance journalist) *Statistics: a game for two players*
Tony Hirst (Communication and Systems Department, The Open University) Visualisations for the rest of us - How to create rich interactive visualisations without any of the pain

Jill Leyland (Royal Statistical Society) The lack of confidence in UK official statistics - are communication problems responsible? - And what can be done?

Kevin McConway (The Open University) Statistics in the media: publicity, entertainment and wallpaper

David Spiegelhalter (University of Cambridge and MRC Biostatistics Unit, Cambridge) Visualising risk and uncertainty: the power of movement

Michel van de Velden (Erasmus University Rotterdam, Netherlands) Perceptual maps: the good, the bad and the ugly

The group also ran its usual seminar programme with invited speakers. Over the year we heard:

Oliver Johnson (University of Bristol) Efficient entropy-based detection of change-points in streaming data

Tony Lawrance (University of Warwick) Three topics in non-linear time series analysis and modelling

Jian Qing Shi (University of Newcastle) Penalized Gaussian process regression and classification for high-dimensional data

Jane Hutton (University of Warwick) Modelling longitudinal data with bounded responses, with missing data

Sujit Sahu (University of Southampton) High resolution Bayesian space-time modelling for ozone concentration levels

Irini Moustaki (London School of Economics and Political Science) Full and limited information estimation methods for latent variable models with longitudinal categorical data

Philip Scarf (University of Salford) A statistical excursion in the isochronic hills

Jim Smith (University of Warwick) Chain event graphs: a new graphical representation of a discrete model

Andrew B. Lawson & Ana Corberán-Vallet (Medical University of South Carolina, Charleston, USA) A multivariate Bayesian model-based approach for on-line spatio-temporal disease surveillance

We also ran our eighth Annual Research Students’ Research Day, where speakers were Osvaldo Anacleto, Fadlalla Elfadaly, Alexandre Santos and Yonas Weldeselassie.
8. Conference & workshop papers and posters given

O. Anacleto

34th Research Students’ Conference in Probability and Statistics, Cambridge
*Graphical dynamic models for road traffic networks*

34th Research Students’ Conference in Probability and Statistics, Cambridge
*Modelling the variability of road traffic flows* [Poster]

ERCIM’11, 4th International Conference of the ERCIM Working Group on
Computing & Statistics, London *Graphical model representations of multivariate
time series for road traffic flow forecasting*

UseR! 2011, Coventry *Forecasting multivariate time series using the DLM package: An application to road traffic networks* [Poster; awarded best application poster prize]

K. Anaya

All Models are Wrong ..., Groningen, Netherlands *Computational information geometry: geometry of model choice* [Poster]

F. Critchley

Fifth Scientific Conference on Modelling and Forecasting of Socio-Economic
Phenomena, Zakopane, Poland *Invariant coordinate selection: an overview and some recent developments* [Plenary talk]

WOGAS3, Third Workshop on Geometric and Algebraic Statistics, Coventry *Computational information geometry for statistical science: theory and practice*

ICORS 2011, International Conference on Robust Statistics, Valladolid, Spain *Invariant coordinate selection revisited: Fisher symmetry and symmetric component analysis*

ERCIM’11, 4th International Conference of the ERCIM Working Group on
Computing & Statistics, London *Invariant coordinate selection revisited: a symmetry group perspective*

F. Elfadaly

34th Research Students’ Conference in Probability and Statistics, Cambridge *On eliciting some prior distributions for the probabilities of multinomial models*

Hierarchical Models and Markov Chain Monte Carlo, Conference in Honour of Professor Adrian F.M. Smith, Crete, Greece *On Eliciting Logistic Normal Priors for Multinomial Models*

C.P. Farrington

NIHR Workshop, Manchester *Methodologic development of syndromic and laboratory statistical surveillance systems in England*

University of Bremen Institute for Prevention Research and Social Medicine workshop on pharmacoepidemiology, Bremen, Germany *The self-controlled case series method. I: Principles and methodology; II: Variations, applications, design*
RIVM Workshop on Human Papillomavirus, Bilthoven, Netherlands Correlated infections: quantifying heterogeneity in transmission and identifying transmission routes from multivariate serological survey data

**P.H. Garthwaite**

47th Annual Statistics Conference at Gregynog, Newtown The analysis of single case studies

**J.C. Gower**

CARME 2011, Correspondence Analysis and Related Methods, Rennes, France

_The analysis of single case studies_

**D. Gragn**

ISCB 2011, 32nd Annual Conference of the International Society for Clinical Biostatistics, Ottawa, Canada Evaluation of statistical algorithms for the detection of infectious disease outbreaks

**M.C. Jones**

Statistics 2011 Canada/IMST 2011-FIM XX, Montreal, Canada Some parametric families of distributions, and their interconnections [Plenary talk]

International Conference on Advances in Probability and Statistics – Theory and Applications: a Celebration of N. Balakrishnan’s 30 Years of Contributions to Statistics, Sha Tin, Hong Kong Some parametric families of distributions and their interconnections [Plenary talk]

**K.J. McConway**

Visualisation and Presentation in Statistics, The Open University _Statistics in the media: publicity, entertainment and wallpaper_

**A. Noufaily**

ISCB 2011, 32nd Annual Conference of the International Society for Clinical Biostatistics, Ottawa, Canada Robust threshold estimation in regression-based statistical surveillance systems

**A.J. Santos**

34th Research Students' Conference in Probability and Statistics, Cambridge Dynamic Bayesian smooth transition autoregressive (DBSTAR) models

31st Annual International Symposium on Forecasting (ISF), Prague, Czech Republic Dynamic Bayesian smooth transition autoregressive (DBSTAR) models

7th Workshop on Bayesian Inference in Stochastic Process (BISP), Madrid, Spain Dynamic Bayesian smooth transition autoregressive (DBSTAR) models
**N.T. Trendafilov**

Symposium on Life Sciences and Statistics, Osaka, Japan *Exploratory factor analysis: modern aspects* [Plenary talk]

ASMDA’11, Applied Stochastic Models and Data Analysis, Rome, Italy *sBarse: sparse biplots component analysis*

**S. Unkel**

ISCB 2011, 32nd Annual Conference of the International Society for Clinical Biostatistics, Ottawa, Canada *The relative frailty variance and shared frailty models*

57th Biometric Colloquium of the International Biometric Society, German Region, Zurich, Switzerland *The relative frailty variance and shared frailty models*

**Y.G. Weldeselassie**

34th Research Students' Conference in Probability and Statistics, Cambridge *Smooth risk functions for self-controlled case series models*

**9. Other seminars and talks given**

**O. Anacleto**

University of Leeds *Graphical dynamic models for road traffic networks*

**F. Critchley**

University College London *Computational information geometry for statistical science: Theory and practice*

University of Bristol *Computational information geometry for statistical science*

**A.E. Faria**

University of Southampton *Dynamic spatial Bayesian models for the interface between computer codes*

**C.P. Farrington**

International Society for Pharmacoepidemiology webinar *The self-controlled case series method in pharmacoepidemiology*

Université de Pau et des Pays de l’Adour, Pau, France *Modèles de fragilité et données de survie à statut actuel*

Biostatistics Network, University College London *The self-controlled case series method: recent developments*

**P.H. Garthwaite**

University of New South Wales, Sydney, Australia *The analysis of single case studies*
J.C. Gower
Polytechnic University of Barcelona, Spain Applications of quadratic minimisation problems in statistics
Polytechnic University of Barcelona, Spain Fisher, Rothamsted 1919 – 1932

K.J. McConway
Royal Statistical Society, Sheffield Local Group Statistics in the media: the good, the bad and the ugly
Max Planck Institute for Human Development, Berlin Statistics as reported in the British mass media: publicity, entertainment and wallpaper

A. Noufaily
Royal Statistical Society, South West Local Group, Plymouth Parametric quantile regression based on the generalized gamma distribution

N.T. Trendafilov
University of Bologna, Italy Exploratory factor analysis of data matrices with more variables than observations

S. Unkel
University of Durham On assessing time dependence of association in bivariate current status data
University of Essex On assessing time dependence of association in bivariate current status data
University of Zurich, Switzerland On assessing time dependence of association in bivariate current status data

H.J. Whitaker
London School of Hygiene and Tropical Medicine The self-controlled case series method: general overview and what next?

10. Editorial roles

F. Critchley
- Associate Editor: Journal de la Société Francaise de Statistique

C.P. Farrington
- Associate Editor: Biostatistics
- Editorial Board Member: Epidemiology and Infection

J.C. Gower
- Associate Editor: The Mathematical Scientist
M.C. Jones

- Associate Editor: *Annals of the Institute of Statistical Mathematics*
- Associate Editor: *Communications in Statistics*
- Associate Editor: *Journal of Statistical Planning and Inference*
- Associate Editor: *Journal of the Korean Statistical Society*
- Associate Editor: *Statistica Sinica* (to August)
- Associate Editor: *Statistical Methodology*

N.T. Trendafilov

- Associate Editor: *Computational Statistics*
- Guest Editor: Special Issue on Sparse Methods in Multivariate Analysis, *Computational Statistics*

11. Grants awarded

F. Critchley

A grant of £4,000 from *Centre for Research in Statistical Methodology (CRiSM), University of Warwick* to put on WOGAS3, Third Workshop on Geometric and Algebraic Statistics

A grant of US$500 from *The Minerva Foundation* to participate in ICORS2011, International Conference on Robust Statistics, Valladolid, Spain

(with H. Massam (York University, Canada), F. Matús (Academy of Sciences, Czech Republic) and J. Smith (University of Warwick)) A grant of CAN$9,900 from *The Fields Institute, Toronto* to organise the Workshop on Graphical Models: Mathematics, Statistics and Computer Science

C.P. Farrington

Received a *Royal Society Wolfson Merit Award*, grant of £56,000 over five years

A.J. Santos

A grant of US$660 from the *International Institute of Forecasters* to participate in the 31st Annual International Symposium on Forecasting (ISF), Prague, Czech Republic

12. Other activities

F. Critchley

- co-organised WOGAS3, Third Workshop on Geometric and Algebraic Statistics, University of Warwick, and a session therein

- international rapporteur for PhD at École Polytechnique Fédérale de Lausanne, Switzerland

- visited Instituto Superior Técnico, Lisbon, Portugal
**C.P. Farrington**
- RSS Vice-President for Academic Affairs
- member of the RSS Council and Council representative on RSS Research Section committee. Elected from Council to serve on the Executive Committee
- external examiner for the MSc in Statistical Science at the University of Lancaster
- external assessor for the Periodic Quality Review of the University of Lancaster’s Department of Mathematics and Statistics
- examined two PhDs, one at Hasselt University, Belgium, the other for the Science Faculty of the Open University
- is a co-investigator on a French grant approved in 2010, from the Agence Nationale pour la Recherche
- is a co-investigator (as is Heather Whitaker) on two Canadian research grants approved in 2010, from the Ministry of Health and Long-Term Care, and from the Canadian Institutes of Health Research

**P.H. Garthwaite**
- external examiner for undergraduate statistics courses at the University of Aberystwyth
- examined a PhD for the University of Durham
- member of the College of Experts for the NIHR Evaluation, Trials and Studies Coordinating Centre

**J.C. Gower**
- member of the Sir Ronald Aylmer Fisher Memorial Committee of Great Britain
- foreign member of the Advisory Board of IOPS (Interuniversitaire Onderzoekschool voor Psychometrie en Sociometrie)

**M.C. Jones**
- member and vice-chair of the RSS Research Section committee

**K.J. McConway**
- wrote a major report for the RSS Council on the state of academic statistics in the UK
- spent long visits at the Statistical Laboratory, University of Cambridge (including contributing to the website at understandinguncertainty.org) and the Centre for Adaptive Behaviour and Cognition at the Max Planck Institute for Human Development, Berlin, Germany
- examined a PhD for the Economics Department of the Open University
- member of the South Central Regional Funding Committee for the Research for Patient Benefit Programme of the NHS National Institute of Health Research
• academic consultant on three more series of the BBC Radio4/OU series ‘More or Less’ (including contributions on air)

C.M. Queen

• examined a PhD for the University of Warwick

N.T. Trendafilov

• organised session at ASMDA’11, Applied Stochastic Models and Data Analysis, Rome, Italy

• visited University of Osaka, Japan

H.J. Whitaker

• examined a PhD at the London School of Hygiene and Tropical Medicine

• gave lecture on course ‘Practical Pharmacoepidemiology’ at the London School of Hygiene and Tropical Medicine

• is a co-investigator (as is Paddy Farrington) on two Canadian research grants approved in 2010, from the Ministry of Health and Long-Term Care, and from the Canadian Institutes of Health Research