The Open University Statistics Group

2012 Report

1. Membership

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

Long-term visitors

Dr K.A. Anaya (London School of Hygiene and Tropical Medicine) [Visiting Research Fellow]
Professor J.C. Gower [Visiting & Emeritus Professor]
Dr S. Unkel (Justus-Liebig-University Giessen, Germany) [Visiting Research Fellow] (from April)

Full–time research students

Mr O. Anacleto
Mr F. Elfadaly (to June)
Ms L. Gunn [joint with Earth Sciences]
Mr O. Oyebamiji [joint with Earth Sciences]
Mr A.J. Santos
Ms S. Villers
Mr Y.G. Weldeselassie

Part–time research students

Mr N. Calleja
Mr A. Gustar [joint with Music]
Mr R. Harris (from October)

Visitors

Dr S. Dossou-Gbété (Université de Pau et des Pays de l'Adour, France)
Dr S. Kato (Institute of Statistical Mathematics, Tokyo, Japan)
Professor I. Koch (University of Adelaide, Australia)
Mr J.F. Rosco-Nieves (Universidad de Extremadura, Spain)

2. Introduction

Here is an overview of the considerable and successful activity in Statistics at the Open University in 2012, in both research and teaching.

Production started and then proceeded at pace on our new first level 30-point module M140 *Introducing Statistics.* This is being produced in double-quick time by a large Module Team headed by Paul Garthwaite and Karen Vines in order to be first presented in October 2013.

Both our new web-based module M347 *Mathematical Statistics* and the revision of M343 *Applications of Probability* saw their first, successful, presentations in 2012. Almost 2000 students took the four (out of five) Statistics modules which were presented with a start date of February 2012. This year, we also had the switchover of the presentation of all current Statistics modules from February to October starts, and a little more than 2000 students are enrolled on the (five) Statistics modules presented for study from October.

In the course of the year, members of the Group published 20 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 25th instalment to our long-standing series of Open University conferences in Statistics.

Early in the year, Steffen Unkel left us for a post at Justus-Liebig-University Giessen, Germany; we wish him well and are delighted that he continues to hold a visiting research appointment with us. Note that Dr D.G. Enki in this report and Dr D. Gragn in previous reports are one and the same person!

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://statistics.open.ac.uk/advisory_service.

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

3. Teaching

Presentation

By the end of the year, the Group was presenting five modules of 30 CATS points each, namely, M248 *Analysing Data* (Chair: Alvaro Faria), M249 *Practical Modern Statistics* (Chair: Paddy Farrington), M343 *Applications of Probability* (Chair: Gillian Iossif), M346 *Linear Statistical Modelling* (Chairs: Paul Garthwaite, Kevin McConway), M347 *Mathematical Statistics* (Chair: Kevin McConway). (M346 was the module not presented in the early part of the year.)

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); Jane Williams represented us on the latter Module Team. Kevin McConway contributed in a small way to the presentation of SDK125 *Introducing Health Science*.

Production

This year the Group began the production of M140 *Introducing Statistics* in earnest. This module will form part of the Open University's new first level curriculum for Mathematics-based degrees and some joint honours degrees. The module is primarily print-based although it will be following M347's pioneering footsteps in having an online presence as well. Paul Garthwaite and Karen Vines are production team co-chairs and authors; Paddy Farrington, Chris Jones, Kevin McConway, Heather Whitaker and Jane Williams make up the rest of the authoring team. The new module is based on our defunct module MDST242 *Statistics in Society*, but with many modifications, updates and modernisations.

In the early months of the year, Gillian Iossif and Chris Jones put the final touches to the production of the revision of M343 *Applications of Probability* and the new module M347 *Mathematical Statistics*, respectively.

Paddy Farrington and Gillian Iossif started work on a minor revision to M249 *Practical Modern Statistics*, to allow for use of a more current version of SPSS.

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 four students were awarded the Mathematics and Statistics degree in 2012. The undergraduate Diploma in Statistics, no longer available to new students, is still awarded to current students who successfully study M248, M249 and two modules from M343, M346 and M347 (or their predecessors). The Diploma was awarded to 52 students in 2012.

4. Ph.D.s awarded

Osvaldo Anacleto Junior for his thesis *Bayesian dynamic graphical models for high-dimensional flow forecasting in road traffic networks* (supervisors: *C.M. Queen* and *P.H. Garthwaite*)

Neville Calleja for his thesis *Mitigating misclassification in health surveys* (supervisors: *P. H. Garthwaite* and *C. P. Farrington*)

Fadlalla Elfadaly for his thesis *Elicitation of subjective probability distributions* (Supervisors: *P. H. Garthwaite* and R. Laney (Computing))

5. Research interests

The Department is home to three Research Groups:

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

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

Medical Statistics: Whitaker (co-ordinator), Enki, Farrington, Garthwaite, Noufaily, 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 Publications in refereed academic journals

Anacleto, O. & Louzada, F. (2012) Bootstrap confidence intervals for industrial recurrent event data. *Pesquisa Operacional*, **32**, 103-119.

Araya, Y.N., Silvertown, J., Gowing, D.J., ***McConway, K.J.,*** Linder, H.P. & Midgley G. (2012) Do niche-structured plant communities exhibit phylogenetic conservatism? A test case in an endemic clade. *Journal of Ecology*, **100,** 1434-1439.

Crawford, J.R., Cullum, C.M., ***Garthwaite, P.H.,*** Lycett, E. & Allsopp, K.J. (2012) Point and interval estimates of percentile ranks for scores on the Texas Functional Living Scale. *Clinical Neuropsychologist*, **26,** 1154-1165.

Crawford, J.R. & ***Garthwaite, P.H.*** (2012) Single case research in neuropsychology: a comparison of five forms of *t*-test for comparing a case to controls. *Cortex*, **48,** 1009-1016.

Crawford, J.R., ***Garthwaite, P.H.,*** Denham, A.K. & Chelune, G.J. (2012) Using regression equations built from summary data in the psychological assessment of the individual case: extension to multiple regression. *Psychological Assessment*, **24**, 801-814.

Crawford, J.R., ***Garthwaite, P.H.,*** Longman, J.S. & Batty, A.M. (2012) Some supplementary methods for the analysis of WAIS-IV index scores in neuropsychological assessment. *Journal of Neuropsychology*, **6,** 192-211.

Crawford, J.R, ***Garthwaite, P.H.,*** Morrice, N. & Duff, K. (2012) Some supplementary methods for the analysis of RBANS. *Psychological Assessment,* **24**, 365-374.

Dallas, N.L., Malone, P.R., Jones, A., Doggart, A.J., ***McConway, K.J.*** & Rogers, P.B. (2012) The results of real-time brachytherapy for the management of low- and intermediate-risk prostate cancer in patients with prostate volumes up to 100 mL. *BJU International*, **110**, 383-390.

Enki, D.G. & ***Trendafilov, N.T.*** (2012) Sparse principal components by semi-partition clustering. *Computational Statistics*, **27,** 605-626.

Farrington C.P., Unkel S. & ***Anaya-Izquierdo, K.*** (2012) The relative frailty variance and shared frailty models. *Journal of the Royal Statistical Society Series B*, **74**, 673–696.

Garthwaite, P.H., Critchley, F., Anaya-Izquierdo, K. & Mubwandarikwa, E. (2012) Orthogonalization of vectors with minimal adjustment. *Biometrika*, **99**, 787-798.

Jones, M.C. (2012) Relationships between distributions with certain symmetries. *Statistics and Probability Letters,* **82,** 1737-1744.

Jones, M.C. & Pewsey, A. (2012) Inverse Batschelet distributions for circular data. *Biometrics,* **68,** 183-193.

Pariente A., Fourrier-Réglat A., Ducruet T., **Farrington P.,** Béland S.G., Dartigues J.F., Moore N. & Moride, Y. (2012) Antipsychotic use and myocardial infarction in older patients with treated dementia. *Archives of Internal Medicine,* **172,** 648-653.

Tanner, J., Gould, D., Jenkins, P., **Hilliam, R.,** Mistry, N. & Walsh, S. (2012) A fresh look at preoperative body washing. *Journal of Infection Prevention*, **13**, 11-15.

Trendafilov, N.T. (2012) DINDSCAL: Direct INDSCAL. *Statistics and Computing*, **22**, 445–454.

Unkel S. & *Farrington C.P.* (2012) A new measure of time-varying association for shared frailty models with bivariate current status data. *Biostatistics*, **13**, 665–679.

Unkel, S., Farrington, C.P., Garthwaite, P.H., Robertson, C. & Andrews, N. (2012) Statistical methods for the prospective detection of infectious disease outbreaks: A review. *Journal of the Royal Statistical Society Series A*, **175,** 49-82.

Warren-Gash C., Hayward, A.C., Hemingway, H., Denaxas, S., Thomas, S.L., Timmis, A.D., **Whitaker, H.** & Smeeth, L. (2012) Influenza infection and risk of acute myocardial infarction in England and Wales: a CALIBER self-controlled case series study. *Journal of Infectious Diseases,* **206**, 1652-1659.

Worthington, J.P., Silvertown, J., Cook, L., Cameron, R., Dodd, M., Greenwood, R.M., **McConway, K.** & Skelton, P. (2012) Evolution MegaLab: a case-study in citizen science methods. *Methods in Ecology and Evolution*, **3**, 303-309.

6.2 Outreach and other publications

Farrington C.P. & *Whitaker, H.J.* (2012) Mortality and the self-controlled case series method: Letter to the Editor. *Pharmacoepidemiology and Drug Safety*, **21,** 906.

McConway, K. & Spiegelhalter, D. (2012) Score and ignore: A radio listener's guide to ignoring health stories. *Significance,* **9**(5), 45-48.

7. Seminars and conferences at The Open University

In May, the Statistics group organized its 25th approximately annual conference, this one being a two-day conference on **Statistical Methods for Infectious Diseases**. The organising committee comprised Sara Griffin & Emma Howard (secretaries in Mathematics & Statistics Department), Paddy Farrington, Angela Noufaily and Steffen Unkel. External financial support for the conference was provided by the Royal Society, the Wolfson Foundation and the Medical Research Council. Speakers at the conference were:

Karim Anaya-Izquierdo (London School of Hygiene and Tropical Medicine) *Estimation of spatial spillover effects in cluster randomized trials*

Nick Andrews (Health Protection Agency, London) *Using the indirect cohort (Broome) method to estimate vaccine effectiveness in the presence of serotype replacement*

Simon Cauchemez (Imperial College London) *Influenza, measurement errors and the interpretation of paired serology*

Azra Ghani (Imperial College London) *Spatial modeling of malaria control and elimination in Africa*

Gavin Gibson (Heriot-Watt University, Edinburgh) *Epidemic models with non-standard infection processes*

Nele Goeyvaerts (Hasselt University, Belgium) *Multivariate decay models for longitudinal censored maternal antibody levels*

Neil Hens (Hasselt University, Belgium) *Estimating infectious disease parameters in a time-heterogeneous setting: hepatitis A in Flanders, Belgium*

Michael Höhle (Robert Koch Institute, Berlin, Germany) *Statistical methods for interpreting the epidemic curve of the STEC O104:H4 outbreak in Germany, 2011*

Valerie Isham (University College London) *The effect of population structure on transmission of infection*

Mirjam Kretzschmar (National Institute for Public Health and the Environment, Bilthoven, and University Medical Centre, Utrecht, The Netherlands) *Two sides of a coin: mathematical and statistical modeling of infectious diseases*

Andrea Mann (London School of Hygiene and Tropical Medicine) *Modeling influenza incidence data using multivariate hidden Markov models*

Piero Manfredi (University of Pisa, Italy) *Hope-Simpson's progressive immunity hypothesis finally explains Zoster data well*

Philip O'Neill (University of Nottingham) *Recent developments in Bayesian inference for infectious disease models*

Tim Peto (John Radcliffe Hospital, Oxford) *Clostridium difficile transmission in Oxfordshire: comparison between network analyses and MCMC modelling approaches*
Anne Presanis (MRC Biostatistics Unit, Cambridge) *Evidence synthesis for infectious diseases*

Kate Soldan (Health Protection Agency, London) *Early evaluation of the impact of the HPV vaccination programme in England*

Steffen Unkel (Justus-Liebig-University Giessen, Germany) *Time-varying frailty models and the estimation of heterogeneities in transmission of infectious diseases*

Jacco Wallinga (National Institute for Public Health and the Environment, Bilthoven, The Netherlands) *Searching for variables that explain the variability in transmission intensity of influenza*

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

John Stevens (University of Sheffield) *Bayesian methods in health economics*

Peter Challenor (National Oceanography Centre, Southampton) *Climate, models and uncertainty*

Inge Koch (University of Adelaide, Australia) *Classification for high-dimensional low sample size data*

Peter Green (University of Bristol and University of Technology, Sydney, Australia) *Emission tomography and Bayesian inverse problems*

Piotr Fryzlewicz (London School of Economics) *Segmenting low- and high-dimensional time series via binary segmentation and its modern variants*

Simplice Dossou-Gbété (Université de Pau et des Pays de l'Adour, France) *Association rules mining approach for binary classification when response's distribution is imbalanced: application to the prediction of in-hospital mortality in Senegal and Mali*

Philip Dawid (University of Cambridge) *Proper local scoring rules*

We also ran our ninth **Annual Research Students' Research Day**, where speakers were Osvaldo Anacleto, Oluwole Oyebamiji, Alexandre Santos, Sofia Villers and Yonas Weldeselassie.

An **internal research morning** was also held at which the speakers were Frank Critchley, Doyo Enki, Paul Garthwaite and Angela Noufaily.

8. Conference & workshop papers and posters given

O. Anacleto

Workshop on Structure and Uncertainty: Modelling, Inference and Computation in Complex Stochastic Systems, Bristol *Dynamic chain graph models* [Poster]

Eighth World Congress in Probability and Statistics, Istanbul, Turkey *Dynamic chain graph models for multivariate time series forecasting*

ISBA [International Society for Bayesian Analysis] 2012 World Meeting, Kyoto, Japan *Dynamic graphical models for real-time multivariate road traffic flow forecasting* [Poster]

Stochastiktage, Mainz, Germany G*raphical model representations of multivariate time series for road traffic flow forecasting*

F. Critchley

Fields Institute Workshop: Graphical Models: Mathematics, Statistics and Computer Science, Toronto, Canada *Computational information geometry and graphical models*

ICMS [International Centre for Mathematical Sciences] Workshop: Advances in Markov Chain Monte Carlo, Edinburgh *Orthogonal foliations: geometry of MCMC for hierarchical models*

D.G. Enki

2012 ISDS (International Society for Disease Surveillance) conference, San Diego, USA *Statistical models for biosurveillance of multiple organisms*

33rd Annual Meeting of the International Society for Clinical Biostatistics, Bergen, Norway *Statistical models for biosurveillance: an empirical investigation*

C.P. Farrington

RSS General Applications Section half-day event, London *Outbreak detection: can the HPA system be improved*?

33rd Annual Meeting of the International Society for Clinical Biostatistics, Bergen, Norway *Estimation of the basic reproduction number for infectious diseases with age-varying individual heterogeneity in contact rates*

33rd Annual Meeting of the International Society for Clinical Biostatistics, Bergen, Norway *Classification biases in case series analysis and application to febrile convulsions*

P.H. Garthwaite

Fifth European Public Health Conference, Malta *Using expert knowledge and single patient studies* [Plenary talk]

ERMITAGE (Enhancing Robustness and Model Integration for the Assessment of Global Environmental Change) Workshop, Berlin *Emulation using partial least squares*

J.C. Gower

RSS, British Society for the History of Mathematics and British Society for the History of Science 50th anniversary commemoration of R.A. Fisher, London *The Rothamsted years*

K.J. McConway

Higher Education Academy Social Sciences Teaching and Learning Summit, Coventry *Public engagement: how one academic statistician got in on the act*

A. Noufaily

33rd Annual Meeting of the International Society for Clinical Biostatistics, Bergen, Norway *An improved algorithm for outbreak detection in multiple surveillance systems*

2012 ISDS (International Society for Disease Surveillance) conference, San Diego, USA An improved algorithm for outbreak detection in multiple surveillance systems [Poster]

A.J. Santos

35th Research Students' Conference in Probability and Statistics, Southampton *Dynamic Bayesian Smooth Transition Autoregressive (DBSTAR) models*

N.T. Trendafilov

5th International Conference of the ERCIM Working Group on Computing and Statistics (ERCIM'12), Oviedo, Spain *From simple structure to sparse components: a comparative introduction*

Y.G. Weldeselassie

35th Research Students' Conference in Probability and Statistics, Southampton *Application of spline based self controlled case series method to vaccine studies of large datasets*

33rd Annual Meeting of the International Society for Clinical Biostatistics, Bergen, Norway *Self-controlled case series method with smooth age effect*

University of Leeds, Work in Progress Meeting *Self-controlled case series method with smooth age effect*

H.J. Whitaker

Primary Care Databases Symposium, Wellcome Collection, London *The self-controlled case series method*

9. Other seminars and talks given

O. Anacleto

Universidade Federal de São Paulo, Brazil *Statistical models for road traffic flow forecasting*

Universidade de São Paulo, Brazil *Dynamic graphical models for multivariate time series*

P.H. Garthwaite

University of Adelaide, Australia *The analysis of single patient studies*

University of Adelaide, Australia *Workshop: Elicitation of expert opinion as a prior distribution*

K.J. McConway

National Audit Office, London *Statistics in the media*

Royal Berkshire Hospital, Reading *Statistics: doing the right thing, and then telling others what you did*

A. Noufaily

Universitè de Pau et des Pays de l'Adour, France *Regréssion quantile paramétrique basée sur la loi gamma generalizée*

10. Editorial roles

F. Critchley

- Associate Editor: *Journal de la Société Francaise de Statistique*
- Editorial Board Member: *Studies in Classification, Data Analysis, and Knowledge Organisation* (book series, Springer)

C.P. Farrington

- Associate Editor: *Biostatistics* (to May)
- Editorial Board Member: *Epidemiology and Infection*

M.C. Jones

- Associate Editor: *Annals of the Institute of Statistical Mathematics*
- Associate Editor: *Communications in Statistics*
- Associate Editor: *Journal of the Korean Statistical Society*
- Associate Editor: *Statistical Methodology*

N.T. Trendafilov

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

11. Major grants awarded

P.H. Garthwaite

A grant of £480,000 from the MRC/NIHR Methodology Research Panel for the project *Development of Statistical Methods for the Analysis of Single Patient Data*. Co-investigators: J.R. Crawford (University of Aberdeen) and M.C. Jones

12. Other activities

F. Critchley

- co-organised meeting on Graphical Models: Mathematics, Statistics and Computer Science at the Fields Institute, Toronto, Canada

- visited University of Waterloo, Canada

C.P. Farrington

- holds a Royal Society Wolfson Merit Award

- examined two PhDs, one at Freiburg University, Germany, the other at Hasselt University, Belgium

- appointed as expert adviser to the court (in Paris) in litigation over the safety of benfluorex (*Mediator*^{TM})

P.H. Garthwaite

- external examiner for undergraduate statistics courses at the University of Aberystwyth

- member of the College of Experts for the NIHR Evaluation, Trials and Studies Coordinating Centre

- visited University of Adelaide, Australia

J.C. Gower

- member of the Sir Ronald Aylmer Fisher Memorial Committee of Great Britain

M.C. Jones

- member and vice-chair of the RSS Research Section committee

K.J. McConway

- RSS Vice-president for Academic Affairs and an elected member of the RSS Council

- member of the Council for the Mathematical Sciences

- external examiner for undergraduate and postgraduate statistics and optimisation courses at the University of Birmingham

- external examiner for undergraduate statistics courses at Queen Mary, University of London

- member, Institutional Review Panel for the Arab Open University

- member of the South Central Regional Advisory Committee for the Research for Patient Benefit Programme of the NHS National Institute of Health Research

- member, Natural Environment Research Council Peer Review College

- academic consultant on three more series of the BBC Radio4/OU series 'More or Less'

- gave a talk to the "Skeptics in the Pub" in Milton Keynes

A. Noufaily

- organiser, half-day meeting on Surveillance Systems for Outbreak Detection of Infectious Diseases on behalf of the General Applications Section, RSS, London

C.M. Queen

- external examiner for undergraduate statistics and related modules at Birkbeck College, University of London

- examined a PhD at the University of Southampton

N.T. Trendafilov

- organised a special session on Sparse Dimension Reduction at the 5th International Conference of the ERCIM Working Group on Computing and Statistics (ERCIM'12), Oviedo, Spain