The Open University Statistics Group

2013 Report

1. Membership

Professor K.J. McConway (Head of Group)
Professor F. Critchley
Dr F. Elfadaly
Dr D.G. Enki (to May)
Dr A.E. Faria
Professor C.P. Farrington
Professor P.H. Garthwaite
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.K. Vines
Dr H.J. Whitaker
Mrs J.E. Williams

Long-term visitors

Dr K.A. Anaya (London School of Hygiene and Tropical Medicine, University of Bath) [Visiting Research Fellow]
Dr D.G. Enki (University of Plymouth) [Visiting Research Fellow]
Professor J.C. Gower [Visiting & Emeritus Professor]
Dr S. Unkel (Justus-Liebig-Universität Giessen, Ludwig-Maximilians-Universität München, both Germany) [Visiting Research Fellow]

Full–time research students

Mr T.G. Gebru [from March]
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 A. Gustar [joint with Music]
Mr R. Harris

Visitor

Professor P. Marriott (University of Waterloo, Canada)

2. Introduction

Here is an overview of the considerable and successful activity in Statistics at the Open University in 2013, in both research and teaching.

On the research side, in addition to all our usual activities and excellent outputs described in Section 4 onwards of this report, 2013 was a terrific year for grant success. Several members of the group received major grants, totalling over £1.2 million (see Section 10 for details). Inter alia, these and existing grants mean that there will be 6 postdoctoral researchers working with us next year.

Another major success, for which hearty congratulations are also due, was Paddy Farrington's award of the Royal Statistical Society's prestigious Bradford Hill Medal, "awarded every three years to a fellow of the Society for outstanding or influential contributions to the development, application or exposition of medical statistics".

In the course of the year, members of the Group published 21 papers in refereed journals (see Section 5). Pre-publication papers can be found in our technical report series, at http://statistics.open.ac.uk/technical_reports

In teaching, production proceeded at frenetic pace on our new first level 30-point module M140 *Introducing Statistics.* This has been produced in double-quick time by a large Module Team headed by Paul Garthwaite and Karen Vines. Its first presentation started in October 2013, attracting well over 1100 students. It will also have a second presentation each year, the first of these beginning in February 2014. This brings us to a grand total of six modules in Statistics (details of which are given in Section 3 below), offered in a total of seven presentations per year (all six starting in October, M140 additionally starting in February). All told, just short of 3000 students are enrolled on the Statistics modules presented for study from October 2013.

At the end of the year, the Group was, however, reduced in number by one by the retirement of our friend and colleague Gillian Iossif. Gillian was a member of the group/department for more than 31 years during which she contributed enormously to our teaching. Gillian was a steadfast member and leader of numerous Course/Module Teams both in the production and presentation of courses (sometimes, almost single-handedly!). Her presence will be sorely missed and we wish her all the very best for a long and happy retirement.

Earlier in the year, postdoc Doyo Enki left us to take up a permanent position in the Plymouth University Peninsula Schools of Medicine and Dentistry.

The Statistical Advisory Service continues to run successfully and has been joined by a Statistical Consultancy Service. The two support research in other departments across the Open University, and are run by Alvaro Faria, with assistance from Paul Garthwaite. Two web pages explain the services: http://statistics.open.ac.uk/advisory_service, http://statistics.open.ac.uk/stats_consult

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

3. Teaching

Presentation

By the end of the year, the Group was presenting a record six modules of 30 CATS points each. These, with numbers of students starting the October 2013 presentation of each course given in square brackets, are:

- [M140](#) *Introducing Statistics* (Chairs: Rachel Hilliam, Karen Vines) [1177]

- [M248](#) *Analysing Data* (Chair: Faria) [483]

- [M249](#) *Practical Modern Statistics* (Chairs: Farrington, Queen) [307]

- [M343](#) *Applications of Probability* (Chair: Iossif) [404]

- [M346](#) *Linear Statistical Modelling* (Chairs: McConway, Whitaker) [366]

- [M347](#) *Mathematical Statistics* (Chairs: McConway, Jones) [252].

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 Sciences*.

Production

As mentioned above, this year saw a huge effort on the (almost complete!) production of [M140](#) *Introducing Statistics*. This module forms part of the Open University's new first level curriculum for Mathematics-based degrees and some joint honours degrees. Paul Garthwaite and Karen Vines were production team co-chairs; other team members included Paddy Farrington, Chris Jones, Kevin McConway, Angela Noufaily, Heather Whitaker and Jane Williams.

Paddy Farrington and Gillian Iossif completed 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. The Mathematics and Statistics degree was awarded to 144 students in 2013. 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 114 students in 2013.

4. Research interests

The Department is home to three Research Groups:

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

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

Medical Statistics: Whitaker (co-ordinator), Enki, Farrington, Garthwaite, Hilliam, 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, earth sciences and ecology.

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

5. Publications

5.1 Publications in refereed academic journals

Anacleto, O., **Queen, C.M.** & Albers, C.J. (2013) Forecasting multivariate road traffic flows using Bayesian dynamic graphical models, splines and other traffic variables. *Australian and New Zealand Journal of Statistics*, **55**, 69-86.

Anacleto, O., **Queen, C.M.** & Albers, C.J. (2013) Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. *Applied Statistics, **62**, 251-270.

Elfadaly, F.G. & **Garthwaite, P.H.** (2013). Eliciting Dirichlet and Connor-Mosimann prior distributions for multinomial models. *Test*, **22**, 628-646.

Enki, D.G., Noufaily, A., Garthwaite, P.H., Andrews, N.J., Charlett, A**.,** Lane, C. & **Farrington, C.P. (2013)** Automated biosurveillance data from England and Wales, 1991-2011. *Emerging Infectious Diseases*, **19**, 35-42.

Enki, D.G., Trendafilov, N.T. & Jolliffe, I. T. (2013) A clustering approach to interpretable principal components. *Journal of Applied Statistics*, **40**, 583-599.

Fardet, L., Nazareth, I., **Whitaker, H.J.** & Petersen, I. (2013) Severe neuropsychiatric outcomes following discontinuation of long term glucocorticoid therapy: a cohort study. *Journal of Clinical Psychiatry, **74**, e281-6.

Farrington, C.P., Unkel S. & Anaya-Izquierdo, K. (2013) Estimation of basic reproduction numbers: individual heterogeneity and robustness to perturbation of the contact function. *Biostatistics* **14**, 528-540.

Farrington C.P., Whitaker H.J., Unkel S. & Pebody R. (2013) Correlated infections: quantifying individual heterogeneity in the spread of infectious diseases. *American Journal of Epidemiology*, **177**, 474-486.

Garthwaite, P.H., Al-Awadhi, S.A., **Elfadaly, F.G.** & Jenkinson, D.J. (2013). Prior distribution elicitation for generalized linear and piecewise-linear models. *Journal of Applied Statistics*, **40**, 59-75.

Hanf M., Quantin C., **Farrington P.,** Benzenine E., Hocine N.M., Velten M., Tubert-Bitter P. & Escolano S. (2013) Validation of the French national health insurance information system as a tool in vaccine safety assessment: Application to febrile convulsions after paediatric measles/mumps/rubella immunization. *Vaccine, **31**, 5856-5862.

Hua, W., Sun, G., Dodd, C.N., Romio, S.A., **Whitaker, H.J.,** Izurieta, H.S., Black, S., Sturkenboom, M.C.J.M., Davis, R.L., Deceuninck, G. & Andrews, N.J. (2013) A simulation study to compare three self-controlled case series approaches: Correction for violation of assumption and explanation of bias.

Pharmacoepidemiology and Drug Safety, **22**, 819-825.

Jones, M.C. (2013) Perlman and Wellner's circular and transformed circular copulas are particular beta and t copulas. *Symmetry*, **5**, 81-85.

Kafatos, G., Andrews, N., ***McConway K.J.*** & ***Farrington P.*** (2013) Regression models for censored serological data. *Journal of Medical Microbiology*, **62**, 93-100.

Kato, S. & ***Jones, M.C.*** (2013) An extended family of circular distributions related to wrapped Cauchy distributions via Brownian motion. *Bernoulli*, **19**, 154-171.

Noufaily, A., Enki, D.G., Farrington, P., Garthwaite, P., Andrews, N. & Charlett, A. (2013) An improved algorithm for outbreak detection in multiple surveillance systems. *Statistics in Medicine*, **32**, 1206-1222.

Noufaily, A. & ***Jones, M.C.*** (2013) On maximisation of the likelihood for the generalized gamma distribution. *Computational Statistics*, **28**, 505-517.

Noufaily, A. & ***Jones, M.C.*** (2013) Parametric quantile regression based on the generalized gamma distribution. *Applied Statistics*, **62**, 723-740*.*

Quantin, C., Benzenine, E., Velten, M., Huet, F., ***Farrington, C.P.*** & Tubert-Bitter, P. (2013) Self-controlled case series and misclassification bias induced by case selection from administrative hospital databases: Application to febrile convulsions in pediatric vaccine pharmacoepidemiology. *American Journal of Epidemiology*, **178**, 1731-1739.

Trendafilov, N.T., Unkel, S. & Krzanowski, W. (2013) Exploratory factor and principal component analyzes: some new aspects. *Statistics and Computing*, **23**, 209--220.

Unkel, S. & ***Trendafilov, N.T.*** (2013) Zig-zag exploratory factor analysis with more variables than observations. *Computational Statistics*, **28**, 107-125.

Wijlaars, L.P.M.M., Nazareth, I., ***Whitaker, H.J.,*** Evans, S.J.W. & Petersen, I. (2013) Suicide-related events in young people following prescription of SSRIs and other antidepressants: A self-controlled case series analysis. *BMJ Open*, **3**, e003247.

5.2 Refereed conference proceedings

Adachi, K. & ***Trendafilov, N.T.*** (2013) A direct cardinality constraint approach to sparse principal components. In *Proceedings of the 27th Symposium of Japanese Society of Computational Statistics*, pp. 165-168.

Adachi, K. & ***Trendafilov, N.T.*** (2013) Perfect simple structure factor analysis with alternating least squares algorithm. In *Proceedings of the 41st Annual Meeting of the Behaviormetric Society of Japan*, pp. 200-203.

Adachi, K. & ***Trendafilov, N.T.*** (2013) Sparse factor analysis for identifying optimal perfect simple structure. In *Proceedings of the Joint Meeting of the IASC Satellite Conference for the 59th ISI WSC and the 8th Asian Regional Section of the IASC*, Seoul, South Korea, pp. 285-290.

Adachi, K. & ***Trendafilov, N.T.*** (2013) Sparse orthogonal factor analysis. In *Advances in Latent Variables*, Eds: E. Brentari & M. Carpita; Milan, Italy, 6 pages.

Anaya-Izquierdo, K., **Critchley, F.,** Marriott, P & Vos, P. (2013) Computational information geometry in statistics: foundations. In *GSI 2013 Proceedings*, Eds: F. Nielsen & F. Barbaresco, pp. 311-318.

Anaya-Izquierdo, K., **Critchley, F.,** Marriott, P & Vos, P. (2013) Computational information geometry in statistics: mixture modelling. In *GSI 2013 Proceedings*, Eds: F. Nielsen & F. Barbaresco, pp. 319-326.

Pereira, I.C. **Faria, A.E.** & Da Silva, M.B. (2013) Influence of feed rate and threaded length in thread forming and tapping operations. In: *Proceedings of the World Congress on Engineering 2013*, Eds: S.I. Ao, L. Gelman, D.W.L. Hukins, A. Hunter & A.M. Korsunsky, pp. 1781-1784.

Trendafilov, N.T. & Adachi, K. (2013) Sparse vs. simple structure loadings. In *Advances in Latent Variables,* Eds: E. Brentari & M. Carpita; Milan, Italy, 6 pages.

5.3 Outreach and other publications

Farrington P. (2013) Censoring on outcome is not valid in self-controlled case series studies. Letter to the Editor. *Journal of Clinical Epidemiology,* **66,** 1428-1429.

McConway, K. (2013) Selling the appeal of statistics. *New Scientist*, **217**, 47.

6. Seminars at The Open University

In 2013, we took a temporary break from hosting our long-standing series of one-day and longer conferences. We ran our usual **seminar programme** with invited speakers. Over the year we heard:

John Aston (University of Warwick) *Functions, covariances, and speaking European languages*

Paul Clark (University of Bristol) *Estimating structural mean models with multiple instrumental variables using the generalised method of moments*

Nicolai Meinshausen (University of Oxford) *Random intersection trees for interaction search in large, sparse datasets*

Lars Eldén (Linköping University, Sweden) *Partial least squares regression - algorithms, analysis and modern applications*

Richard Boys (University of Newcastle) *Inference for population dynamics in the Neolithic period*

Bill Browne (University of Bristol) *Stat-JR and other software developed at the multilevel modelling centre*

Helen Higgins (University of Bristol) *Taking a truly Bayesian approach: triumphs and tribulations of eliciting vets' beliefs*

Michele Gallo (University of Naples "L'Orientale", Italy) *N-way partial least squares for compositional data*

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

Another **internal research morning** was also held at which the speakers were Steffen Unkel, Angela Noufaily, Doyo Enki, Chris Jones, Paddy Farrington and Nickolay Trendafilov.

7. Conference & workshop papers and posters given

F. Critchley

Geometric Science of Information Conference GSI'13, Paris, France *Computational information geometry in statistics: foundations*

C.P. Farrington

Royal Statistical Society Conference RSS 2013, Newcastle *The self-controlled case series method: Genesis and early history*

P.H. Garthwaite

Joint Statistics Meeting, Montreal, Canada *Overview of elicitation methods and software*

T. G. Gebru

6th International Conference of the ERCIM Working Group on Computing and Statistics (ERCIM'13), London *Sparse linear discriminant analysis in high-dimensions: a review*

J.C. Gower

Royal Statistical Society one-day conference in honour of Fisher's 125th anniversary, London *Fisher at Rothamsted*

International Federation of Classification Societies 2013 Conference, Tilburg, Netherlands *Visualisation of three way arrays*

International Federation of Classification Societies 2013 Conference, Tilburg, Netherlands. **Presenter on pre-conference course on** *Biplots*

M.C. Jones

Non-Gaussian Multivariate Statistical Models and Their Applications, Banff, Canada *Alternative skew-symmetric distributions*

Leeds Annual Statistics Research Conference LASR 2013, Leeds *Families of unimodal distributions on the circle*

Royal Statistical Society Conference RSS 2013, Newcastle *Simple thoughts on simple quantile regression*

First International Conference on Statistical Distributions and Their Applications, Mount Pleasant, MI, USA *Families of unimodal distributions on the circle* [Keynote talk]

A. Noufaily

Royal Statistical Society Conference RSS 2013, Newcastle *Modelling reporting delays for infectious diseases using splines*

N.T. Trendafilov

78th Annual Meeting of the Psychometric Society IMPS 2013, Arnhem, Netherlands *Computational identification of optimal confirmatory factor analysis model using sparseness constraint*

6th International Conference of the ERCIM Working Group on Computing and Statistics (ERCIM'13), London *A new sparse PCA procedure via decomposing PCA loss function*

Y.G. Weldeselassie

34th Annual Meeting of the International Society for Clinical Biostatistics, Munich, Germany *Flexible modelling of vaccine effects in the self-controlled case series method*

8. Other seminars and talks given

F. Critchley

University of Leeds *Invariant coordinate selection … revisited*

M.C. Jones

Institute of Statistical Mathematics, Tokyo, Japan *Extending univariate families of distributions to the bivariate case*

K.J. McConway

Royal Statistical Society *Academic statistics in the UK*

Association of Statistics Lecturers in Universities/RSS Medical Section *The future of academic statistics in the UK*

A. Noufaily

Queen Mary University of London *An improved algorithm for outbreak detection in multiple surveillance systems*

University of Durham *Robust threshold estimation for outbreak detection in multiple surveillance systems*

C. Queen

University of Warwick *A graphical dynamic approach to forecasting flows in road traffic networks*

University of Bath *A graphical dynamic approach to forecasting flows in road traffic networks*

9.   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

- Editorial Board Member: *Epidemiology and Infection*

M.C. Jones

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

K.J. McConway

- Senior Editor/Statistical Advisor: *Journal of Endocrinology*
- Senior Editor/Statistical Advisor: *Journal of Molecular Endocrinology*

N.T. Trendafilov

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

10. Major grants awarded

F. Critchley

A grant of £627,874 from the EPSRC for the project *Emerging Geometries for Statistical Science: Articulating the Vision*

C.P. Farrington & H.J. Whitaker

Paddy Farrington is Principal Investigator, Heather Whitaker Co-Investigator, on the project *Software Tools and Online Resources for the Self-Controlled Case Series Method and its Extension* funded by a grant of £330,000 from the MRC

Paddy Farrington and Heather Whitaker also obtained a £76,000 share of the €10.7 million EU Innovative Medicines Initiative project *Accelerated Development of Vaccine Benefit-Risk Collaboration in Europe (ADVANCE)*, as joint partners with 20 other European centres

N.T. Trendafilov

A grant of £173,257 from the Leverhulme Trust for the project *Sparse Factor Analysis with Application to Large Data Sets*

11. Other activities

F. Critchley

- organised a session on *Computational aspects of information geometry in statistics* at the Geometric Science of Information conference GSI'13, Paris, France

- visited Instituto Superior Técnico, Lisbon, Portugal

- examined a PhD at the University of Leeds

C.P. Farrington

- was awarded the RSS Bradford Hill medal

- holds a Royal Society Wolfson Merit Award

- was appointed as a member of the RSS Honours Committee

- took part in the review panel for the *IAME Infection, Antimicrobials, Modelling, Evolution Unit*, Universités Paris-Nord et Denis Diderot, Paris, under the aegis of the AERES (*Agence d'Evaluation de la Recherche et de l'Enseignement Supérieur*)

- worked as expert adviser to the court (in Paris) in litigation over the safety of benfluorex (*Mediator*™)

- examined a PhD at the University of Nottingham

- recorded a promotional video for the OU on vaccine safety

P.H. Garthwaite

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

- examined a PhD at Brunel University

J.C. Gower

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

R.M. Hilliam

- presented a one day workshop for University of Derby doctoral students entitled *Statistics; what you need to know*

M.C. Jones

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

- visited Institute of Statistical Mathematics, Tokyo, Japan

- examined a PhD at the University of Warwick

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 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'

- interviewed on air for Inside Science (BBC Radio 4), BBC Scotland and BBC Three Counties Radio

- wrote invited comments on statistical aspects of scientific papers for the Science Media Centre

- gave talks to the "Skeptics in the Pub" in Leicester, Bedford and High Wycombe

- invited member of Roundtable meetings (arranged by HEFCE) on enhancements to Unistats

- trustee of the Southern England Consortium for Credit Accumulation and Transfer

A. Noufaily

- organised a session on *Quantile regression* at the Royal Statistical Society Conference RSS 2013, Newcastle

- recorded a promotional video for the OU on outbreak control

C.M. Queen

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