

The Open University Statistics Group

2014 Report

1. Membership

Professor K.J. McConway (Head of Group to March)
Dr C.M. Queen (Head of Group from April)
Miss C. Calvert (from September)
Professor F. Critchley
Dr F. Elfadaly
Dr A.E. Faria
Professor C.P. Farrington
Dr S. Fontanella (from February)
Professor P.H. Garthwaite
Dr R.M. Hilliam
Professor M.C. Jones
Dr A. Noufaily
Dr R. Sabolova (from October)
Dr N.T. Trendafilov
Dr G. Van Bever (from October)
Dr S.K. Vines
Dr Y.G. Weldeselassie (from May)
Dr H.J. Whitaker
Mrs J.E. Williams (to August)

Long-term associates

Dr K.A. Anaya (University of Bath) [Visiting Research Fellow, to July]
Dr D.G. Enki (University of Plymouth) [Visiting Research Fellow, to July]
Professor J.C. Gower [Emeritus Professor]
Dr S. Unkel (Justus-Liebig-Universität Giessen, Ludwig-Maximilians-Universität München, Universitätsmedizin Göttingen, all Germany) [Visiting Research Fellow]

Full-time research students

Mr T.G. Gebru
Ms L. Gunn [joint with Earth Sciences, to July]
Mr O. Oyebamiji [joint with Earth Sciences]
Mr M.Z. Rahman Shabuz (from October)
Mr A.J. Santos (to September)
Ms S. Villers (to September)
Mr Y.G. Weldeselassie (to April)

Part-time research students

Mr A. Gustar [joint with Music, to September]
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 2014, in both research and teaching.

On the research side, all our usual activities and excellent outputs are described in Section 4 onwards of this report. There are now no fewer than 6 postdoctoral researchers working in the Group (Elfadaly, Fontanella, Noufaily, Sabolova, Van Bever, Weldeselassie). Our efforts formed a substantial part of the Open University's submission to UoA 10, Mathematical Sciences, of the REF, the outcome of which appears to have been satisfactory.

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

The excellence of our new first level 30-point module [M140](#) *Introducing Statistics* (first presented from October 2013) was recognised by the Open University through one of its coveted Teaching Awards. The award, for Excellence in Promoting Equality, was made to the Module Team of Paul Garthwaite and Karen Vines (chairs), Paddy Farrington, Rachel Hilliam, Chris Jones, Kevin McConway, Heather Whitaker and Jane Williams, and Curriculum Manager Gloria Baldi.

All told, just short of 3500 students were enrolled on the Statistics modules presented for study from dates in 2014 (see Section 3).

We are delighted to be able to report the promotions of Catriona Queen to Senior Lecturer and Nickolay Trendafilov to Reader, and congratulate them on their success.

Our Staff Tutor colleague Jane Williams retired in August after very many years of service at the Open University. Over the years, Jane held various positions both regionally and centrally, several of them senior, as well as being a highly valued member of the Group and of numerous Course/Module Teams in Statistics. Best wishes to Jane for a long and happy retirement. Staff Tutor numbers in Statistics were shortly thereafter replenished (to two) by the appointment of Carol Calvert.

The Statistical Advisory and Consultancy Services continue to run successfully, supporting research in other departments across the Open University. These are co-ordinated by Alvaro Faria, assisted by 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

For the first time, the Group presented six modules of 30 CATS points each with one of them (M140) being presented twice in the year. These, with numbers of students starting the presentations of each course given in square brackets, are:

- [M140](#) *Introducing Statistics* (Chairs: Vines, Hilliam) [February 756; October 1063]
- [M248](#) *Analysing Data* (Chair: Faria) [399]
- [M249](#) *Practical Modern Statistics* (Chair: Queen) [249, yes, really]
- [M343](#) *Applications of Probability* (Chair: Trendafilov) [403]
- [M346](#) *Linear Statistical Modelling* (Chair: Whitaker) [357]
- [M347](#) *Mathematical Statistics* (Chair: Jones) [242].

The first level mathematics modules, [MU123](#) *Discovering Mathematics* and the recently finished MST121 *Using Mathematics*, contain 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

Final touches were put to the new first level module [M140](#) *Introducing Statistics*.

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 109 students in 2014. 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 67 students in 2013. A further 44 students were awarded other joint degrees that include a substantial element of statistics.

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), 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, robust statistics, time series forecasting, influence analysis and statistics in the media 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. PhDs awarded

Leanne Gunn for her thesis

The duration of volcanic eruptions: empirical probabilistic forecasting models based on historic eruption data

(supervisors: S. Blake, H Rymer (both Environment, Earth & Ecosystems), **M.C. Jones** and P. Einarsson (U. Iceland))

Andrew Gustar for his thesis

The demographics of musical works: a statistical approach to historical musicology

(supervisors: D. Rowland (Music) and **K.J. McConway**)

Alexandre Santos for his thesis

Dynamic Bayesian Smooth Transition Autoregressive (DBSTAR) models for non-stationary nonlinear time series

(supervisors: **A.E. Faria**, **C.M. Queen** and R. Souza (Pontifícia Universidade Católica do Rio de Janeiro))

Sofia Villers for her thesis

Measuring orienteers' performance: a method based on lower rank approximation of asymmetrical matrices

(supervisors: **K. Vines** and **K.J. McConway**)

Yonas Weldeselassie for his thesis

Smooth risk functions for self-controlled case series method

(supervisors: **H.J. Whitaker** and **C.P. Farrington**)

6. Publications

6.1 Publications in refereed journals

Albers, C. J. & **Gower, J.C.** (2014) Canonical analysis: ranks, ratios and fits. *Journal of Classification*, **31**, 2-27.

Albers, C. J. & **Gower, J.C.** (2014) A contribution to the visualisation of three-way tables. *Journal of Multivariate Analysis*, **132**, 1-8.

Alkahtani, S., Cherrill, J., Millward, C., Grayson, K., **Hilliam, R.**, Sammons, H. & Choonara, I. (2014) Access to medicines by child refugees in the East Midlands region of England: a cross-sectional study. *BMJ Open*, **4**, article e006421.

Anaya-Izquierdo, K., **Critchley, F.** & Marriott, P. (2014) When are first-order asymptotics adequate? A diagnostic. *Stat*, **3**, 17–22.

Connolly-Andersen, A.M., Hammargren, E., **Whitaker, H.**, Eliasson, M., Holmgren, L., Klingström, J. & Ahlm, C. (2014) Increased risk of acute myocardial infarction and stroke during hemorrhagic fever with renal syndrome: a self-controlled case series study. *Circulation*, **129**, 1295-1302.

Critchley, F. & Marriott, P. (2014) Computational information geometry in statistics: theory and practice. *Entropy*, **16**, 2454-2471.

Elfadaly, F.G. & El Gayar, S.M. (2014). Two-term Edgeworth expansions for the classes of U- and V-statistics. *Communications in Statistics – Theory and Methods*, **43**, 4679-4706.

Enki, D.G., **Noufaily, A.** & **Farrington, C.P.** (2014). A time-varying shared frailty model with application to infectious diseases. *Annals of Applied Statistics*, **8**, 430-447.

Fakis, A., **Hilliam, R.**, Stoneley, H. & Townend, M. (2014) Quantitative analysis of qualitative information from interviews: a systematic literature review. *Journal of Mixed Methods Research*, **8**, 139-161.

Fakis, A., **Hilliam, R.**, Townend, M., Stoneley, H. & Robinson, G. (2014) Mothers' and grandmothers' perceptions relating to causality, treatment and support for families of a child with a diagnosis of attention deficit hyperactivity disorder: applying Enosis, an alternative mixed methods approach. *International Journal of Multiple Research Approaches*, **8**, 148-165.

Ghebremichael-Weldeselassie, Y., Whitaker, H.J. & **Farrington, C.P.** (2014) Self-controlled case series with smooth age effect. *Statistics in Medicine*, **33**, 639-649.

Gower, J., Groenen, P.F.J., Van de Velden, M. & **Vines, K.** (2014) Better perceptual maps: introducing explanatory icons to facilitate interpretation. *Food Quality and Preference*, **36**, 61-69.

Gower, J.C., Le Roux, N.J. & Gardner-Lubbe, S. (2014) The canonical analysis of distance. *Journal of Classification*, **31**, 107-128.

Gunn, L.S., Blake, S., **Jones, M.C.** & Rymer, H. (2014) Forecasting the duration of future flank eruptions: an empirical probabilistic model. *Bulletin of Volcanology*, **76**, Article 780.

Holden, P.B., Edwards, N.R., **Garthwaite, P.H.**, Fraedrich, F., Lunkeit, F., Kirk, E., Labriet, M., Kanudia, A. & Babonneau, F. (2014). PLASIM-ENTSem v1.0: a spatio-temporal emulator of future climate change for impacts assessment. *Geoscientific Model Development*, **7**, 433-451.

Jones, M.C. (2014) Generating distributions by transformation of scale. *Statistica Sinica*, **24**, 749-771.

LeRoux, N.J., Gardner-Lubbe, S. & **Gower, J.C.** (2014) The analysis of distance of grouped data with categorical variables: categorical canonical variate analysis. *Journal of Multivariate Analysis*, **132**, 9-24.

Menendez, P., Fan, Y., **Garthwaite, P.H.** & Sisson, S.A. (2014). Simultaneous adjustment of bias and coverage probabilities for confidence intervals. *Computational Statistics and Data Analysis*, **70**, 35-44.

Trendafilov, N.T. (2014) From simple structure to sparse components: a review. *Computational Statistics*, **29**, 431-454.

Unkel, S., **Farrington, C.P., Whitaker, H.J.** & Pebody, R. (2014) Time varying frailty models and the estimation of heterogeneities in transmission of infectious diseases. *Journal of the Royal Statistical Society Series C*, **63**, 141-158.

6.2 Book chapters and refereed conference proceedings

Adachi, K. & **Trendafilov, N.T.** (2014) Penalty-free sparse PCA. In *Proceedings of COMPSTAT 2014: 21st International Conference on Computational Statistics*, Eds: M. Gilli, G. Gonzalez-Rodriguez & A. Nieto-Reyes, pp. 197-203.

Adachi, K. & **Trendafilov, N.T.** (2014) Factor analysis for clustering variables with EM algorithm. In *Proceedings of the 28th Annual Meeting of the Japanese Society of Computational Statistics*, pp. 159-162.

Fontanella S., Trendafilov, N.T. & Adachi, K. (2014) Sparse exploratory factor analysis. In *Proceedings of COMPSTAT 2014: 21st International Conference on Computational Statistics*, Eds: M. Gilli, G. Gonzalez-Rodriguez & A. Nieto-Reyes, pp. 281-288.

Gower, J.C. (2014) Some history of algebraic forms and data analysis. In *Visualization and Verbalization of Data*, Eds: J. Blasius & M. Greenacre; Chapman & Hall/CRC, pp. 17-29.

Gower, J.C. (2014) Retirement is a beginning. In *Psychology and statistics: Ready for the future?* (Willem Heiser Festschrift) Eds: M. de Rooij & C.M. van Putten; Institute of Psychology, Leiden University, pp. 103-108.

Trendafilov, N.T. & Adachi, K. (2014) Sparse orthogonal factor analysis. In *Advances in Latent Variables: Methods, Models and Applications*. Eds: M. Carpita, E. Brentari & E.M. Qannari; Springer, pp. 227-239.

Trendafilov, N.T. & Gallo, M. (2014) Sparse PCA for compositional data. In *Proceedings of the 47th Scientific Meeting of the Italian Statistical Society*, Eds: S. Cabras, T. Di Battista & W. Racugno.

6.3 Others

Trendafilov, N., Kleinstaubler, M. and Zou, H. (2014) Editorial: sparse matrices in data analysis. *Computational Statistics*, **29**, 403-405.

7. Seminars at The Open University

We ran our usual **seminar programme** with invited speakers. Over the year we heard:

Luigi Ippoliti (Gabriele d'Annunzio University of Chieti-Pescara, Italy) *The offset normal shape distribution for dynamic shape analysis*

Krzysztof Latuszynski (University of Warwick) *The Bernoulli factory, extensions and applications*

Jordan Stoyanov (University of Newcastle) *Intriguing problems from combinatorics, algebra, analysis & probability*

Marc Hallin (Université Libre de Bruxelles, Belgium, and Princeton University, USA) *The double sin of the skew-normal, skew-symmetric families of distributions and Fisher information*

Susan Holmes (Stanford University, USA) *Incorporating multiple tables, trees and networks to analyze the microbiome*

Rajen Shah (University of Cambridge) *Min-wise hashing for large-scale regression*

Antonio Gasparrini (London School of Hygiene and Tropical Medicine) *Modelling exposure-lag-response associations*

Stephen Buckland (University of St Andrews) *Model-based distance sampling methods*

Ian Vernon (University of Durham) *Confronting a galaxy formation computer model with conflicting observed data sets: a necessarily subjective Bayesian analysis*

We also ran our eleventh **Annual Research Students' Research Day**, where speakers were Oluwole Oyebamiji, Ross Harris, Sofia Villers and Tsegay Gebru.

And we held an **internal research morning** at which the speakers were Germain Van Bever, Radka Sabolova, Paul Garthwaite, Sara Fontanella and Chris Jones.

8. Conference & workshop papers and posters given

F. Critchley

NORDSTAT 2014, the 25th Nordic Conference in Mathematical Statistics, Turku, Finland *Invariant coordinate selection*

One Day Conference on Geometry and Statistics, Bath *When are first order asymptotics adequate? ... a diagnostic*

F. Elfadaly

ISBA 2014, Twelfth World Meeting of the International Society for Bayesian Statistics, Cancun, Mexico *On eliciting some prior distributions for probabilities of multinomial models* [poster]

A.E. Faria

34th International Symposium on Forecasting, Rotterdam, Netherlands *Dynamic Bayesian smooth transition autoregressive models for non-linear non-stationary processes*

C.P. Farrington

Royal Statistical Society Medical Section Meeting on the Self-Controlled Case Series Method, London *Control without controls: genesis, theory, controversy*

S. Fontanella

COMPSTAT 2014: 21st International Conference on Computational Statistics, Geneva, Switzerland *Sparse exploratory factor analysis*

P.H. Garthwaite

2nd Subjective Bayesian Meeting, Durham *Quantifying expert opinion as a probability distribution*

Royal Statistical Society International Conference RSS 2014, Sheffield *Analysis for single patient studies: examining a patient's profile of scores*

ISBA 2014, Twelfth World Meeting of the International Society for Bayesian Statistics, Cancun, Mexico *Analysis of single patient data and probability matching priors* [poster]

T.G. Gebru

COMPSTAT 2014: 21st International Conference on Computational Statistics, Geneva, Switzerland *Function constrained sparse linear discriminant analysis*

R. Hilliam & J.E. Williams

HEA STEM Annual Learning and Teaching Conference, Edinburgh *Bringing statistics to life for a diverse population* [poster]

M.C. Jones

Advances in Directional Statistics (ADISTA14) International Workshop, Brussels, Belgium *Families of unimodal distributions on the circle*

Mathematical Methods in Reliability Joint Research Group Meeting on Probability Distributions and Statistical Models for Reliability and Survival Analysis, Brunel University, Uxbridge *Distributions on R^+ : shape parameters and transformations*

7th Trondheim Symposium in Statistics, Selbu, Norway *Some parametric families of distributions: connections, comparisons and extensions*

7th International Conference of the ERCIM Working Group on Computational and Methodological Statistics (ERCIM 2014), Pisa, Italy *Families of unimodal distributions on the circle*

K.J. McConway

Statistics in Journalism Practice and Education Conference, Sheffield *Statistics and the media: a statistician's view*

Rustat Conference on Risk and the Use and Misuse of Statistics, Cambridge *Improving the understanding of risk and the better use of statistics*

Royal Statistical Society International Conference RSS 2014, Sheffield *Working with the media*

A. Noufaily

2nd General Assembly Meeting of Accelerated Development of Vaccine Benefit-Risk Collaboration in Europe (ADVANCE), Barcelona, Spain *Risk assessment: progress so far and next steps*

Workshop on Mathematical Challenges for Long Epidemic Time Series, Warwick *Detection of infectious disease outbreaks from laboratory data with reporting delays*

N.T. Trendafilov

7th International Conference of the ERCIM Working Group on Computational and Methodological Statistics (ERCIM 2014), Pisa, Italy *Common principal components estimation: a dynamical system approach*

Y.G. Weldeselassie

35th Annual Meeting of the International Society for Clinical Biostatistics, Vienna, Austria *Non-parametric self-controlled case series method*

9. Other seminars and talks given

F. Critchley

Université Libre de Bruxelles, Belgium *Computational information geometry: an introduction and overview*

University of Cambridge *Fisher symmetry and symmetric component analysis*

C.P. Farrington

Roslin Institute, Edinburgh *Correlated infections – and what they can tell us about disease transmission*

Heriot-Watt University, Edinburgh *Estimation of individual heterogeneity in infections transmission from multivariate serological survey data*

P.H. Garthwaite

University of Bristol *Analysis for single patient studies*

M.C. Jones

University of Bath *Some parametric families of distributions; what they are and how they compare*

C.M. Queen

University of Bristol *Dynamic chain graphs for high-dimensional time series: an application to real-time traffic flow forecasting*

G. Van Bever

Université Libre de Bruxelles, Belgium *Advances in invariant coordinate analysis*

Y.G. Weldeselassie

University College London *The self-controlled case series (SCCS) method with smooth risk functions*

10. Editorial roles

F. Critchley

- Associate Editor: *Journal de la Société Française 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* (to October)

M.C. Jones

- Associate Editor: *Annals of the Institute of Statistical Mathematics*
- Associate Editor: *Journal of the Korean Statistical Society*
- Associate Editor: *Statistical Methodology*
- Associate Editor: *Statistics and Probability Letters* (from October)

K.J. McConway

- Senior Editor/Statistical Advisor: *Endocrine-Related Cancer*
- 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*

11. Major grants awarded

F. Critchley

ICMS Workshop grant, Edinburgh, for *Computational information geometry for image and signal processing*: £20,500; co-organiser with CTJ Dodson (Manchester) and Frank Nielsen (Ecole Polytechnique, Paris); workshop to be held 21-25 September 2015.

12. Other activities

F. Critchley

- examined a PhD at the University of Jyväskylä, Finland

C.P. Farrington

- holds a Royal Society Wolfson Merit Award
- is a member of the RSS Honours Committee
- examined a PhD at the University of Strathclyde

P.H. Garthwaite

- joined the RSS Research Section committee
- external examiner for undergraduate statistics courses at the University of Aberystwyth
- examined a PhD at Queen Mary University of London

J.C. Gower

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

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 (to March)
- member, Natural Environment Research Council Peer Review College
- academic consultant on two more series of the BBC Radio4/OU series 'More or Less'
- wrote invited comments on statistical aspects of scientific papers for the Science Media Centre
- STEM Ambassador (working with schools and young people to promote STEM subjects), including giving talks at two schools in Wellingborough
- judge for the 'SET for Britain' parliamentary poster competition for career-young researchers
- trustee of the Southern England Consortium for Credit Accumulation and Transfer

C.M. Queen

- external examiner for undergraduate statistics and related modules at Birkbeck College, University of London
- examined a PhD at the University of Warwick

R. Sabolova

- visited the University of Turku, Finland

N.T. Trendafilov

- visited the University of Naples "L'Orientale", Italy, Université Catholique de Louvain, Louvain-la-Neuve, Belgium and the University of Osaka, Japan