The Open University Statistics Group

2016 Report

1. Membership

Miss A. Bromley (from June)
Miss C. Calvert
Professor F. Critchley
Dr F. Elfadaly (from July)
Dr Á.E. Faria
Dr S. Fontanella (to November)
Professor P.H. Garthwaite
Dr R.M. Hilliam
Professor M.C. Jones
Professor K.J. McConway (to January)
Dr C.M. Queen
Dr R. Sabolová
Dr N.T. Trendafilov
Dr G. Van Bever (to July)
Dr S.K. Vines
Dr Y.G. Weldeselassie
Dr H.J. Whitaker

Long-term associates

Dr F. Elfadaly (Imperial College Faculty of Medicine) [Visiting Research Fellow to July]
Professor C.P. Farrington [Emeritus Professor]
Professor J.C. Gower [Emeritus Professor]
Professor K.J. McConway [Emeritus Professor from January]
Dr A. Noufaily (Warwick Medical School) [Visiting Research Fellow]

Full–time research students

Mr T.G. Gebru
Mr M.Z. Rahman Shabuz

Part–time research students

Mr R. Harris

Visitors

Dr K. Burke (University of Limerick, Ireland)
Professor P. Marriott (University of Waterloo, Canada)
Professor H. Oja (University of Turku, Finland)

2. Introduction

Here is an overview of the considerable and successful activity in Statistics at the Open University in 2016, in both research and teaching.

On the research side, all our usual activities and outputs are described in Section 4 onwards of this report. In the course of the year, members of the Group published 27 papers in refereed journals — *a record number* — along with a number of other outputs (see Section 6). Pre-publication papers can be found in our technical report series, at

http://www.mathematics.open.ac.uk/research/statistics/technical-reports

On the teaching side, just under 2900 students were enrolled on the Statistics modules presented for study from dates in 2016 (see Section 3).

Our excellent colleague Kevin McConway retired in January after just 35 years, 3 months and 17 days of service at the Open University. Kevin was an exceptionally valued and admired colleague: for his outstanding contributions over so many years to teaching; for his excellent research, practical collaboration and outreach/media activities; and for his major academic managerial and administrative service: from the statisticians' point of view, he was Mr Open University. Best wishes to Kevin for a long and happy retirement.

We were also sad to see two excellent postdocs move on to the next stages of their careers, Germain van Bever to Université Libre de Bruxelles, Belgium, and Sara Fontanella to Imperial College London. We wish them very well too, of course.

Our central academic numbers were partially replenished after recent retirements, plural, by the appointment of Dr Fadlalla Elfadaly (from Imperial, but previously an O.U. Ph.D. student and postdoc) to a lectureship. Pressure on staff tutor time due to Rachel Hilliam's other commitments was assuaged by the fixed term appointment of Miss Alison Bromley. At the very end of the year, Alison was appointed to a permanent staff tutor post. Good luck, folks!

The Statistical Advisory and Consultancy Services continue to run successfully, supporting research in other departments across the Open University. These are run by Álvaro Faria and Paul Garthwaite. Two web pages explain the services:

http://www.mathematics.open.ac.uk/research/research/statistics/advisory-service

http://www.mathematics.open.ac.uk/research/statistics/statistical-consultancy-service

Further information on the Group's research activities is available through our general web pages at http://www.mathematics.open.ac.uk/research/statistics

3. Teaching

Presentation

The Group continued to present 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: Calvert, Hilliam) [February 594; October 945]

- [M248](#) *Analysing Data* (Chair: Faria) [336]

- [M249](#) *Practical Modern Statistics* (Chairs: Vines, Whitaker) [192]

- [M343](#) *Applications of Probability* (Chair: Trendafilov) [351]

- [M346](#) *Linear Statistical Modelling* (Chair: Garthwaite, Vines) [242]

- [M347](#) *Mathematical Statistics* (Chair: Jones) [231].

The first level mathematics module, [MU123](#) *Discovering Mathematics,* also contains a little statistical material.

Production

Work continued apace on a considerable revision of [M248](#) *Analysing Data,* for first presentation in October 2017. This work is being undertaken in order to take account of the new predecessor module [M140](#), to improve parts of the material, and to update aspects of the presentation of the module.

Qualifications

Our modules are compulsory components of degrees in [Mathematics and Statistics](#), [Mathematics](#), [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 102 students in 2016. The undergraduate Diploma in Statistics, no longer available to new students, is still awarded to current students who successfully study a large proportion of our modules; the Diploma was awarded to 6 students in 2016. A further 48 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 to September), Trendafilov (co-ordinator from September) Faria, Fontanella, Garthwaite, Gower, Jones, Queen, Sabolová, Van Bever, Vines.

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 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 journals

Adachi, K. & **Trendafilov, N.T.** (2016) Sparse principal component analysis subject to prespecified cardinality of loadings. *Computational Statistics*, **31,** 1403-1427.

Calleja, N. & **Garthwaite, P.H.** (2016) Running an international survey in a small country: challenges and opportunities. *Public Health Panorama*, **2,** 292-301.

Calvert, C., Hilliam, R. & Coleman, J. (2016) Improving retention for all students, studying mathematics as part of their chosen qualification, by using a voluntary diagnostic quiz. *MSOR Connections,* **14,** 3, 28-31.

Collins, T., Laney, R., Willis, A. & **Garthwaite, P.H.** (2016) Developing and evaluating computational models of musical style. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* **30,** 16-43.

Critchley, F. & Marriott, P. (2016) Computing with Fisher geodesics and extended exponential families. *Statistics and Computing,* **26,** 325-332.

Corrêa, J.M., Neto, A.C., Teixeira Jr., L.A., Carreño, E.M. & **Faria, Á.E.** (2016) Linear combination of forecasts with numerical adjustment via MINIMAX non-linear programming. *Revista GEPROS - Gestão da Produção, Operações e Sistemas*, **11**, 79–95.

Corrêa, J.M., Neto, A.C., Teixeira Jr., L.A., Franco, E.M.C. & **Faria Jr., A.E.** (2016) Time series forecasting with the WARIMAX-GARCH method. *Neurocomputing*, **216,** 805-815.

Elfadaly, F.G., Garthwaite, P.H. & Crawford, J R. (2016) On point estimation of the abnormality of a Mahalanobis index. *Computational Statistics and Data Analysis,* **99,** 115-130.

Enki, D.G., **Garthwaite, P.H., Farrington, C.P.,** Noufaily, A., Andrews, N.J. & Charlett, A. (2016) Comparison of statistical algorithms for the detection of infectious disease outbreaks in large multiple surveillance systems. *PLOS ONE*, **11,** Issue 8.

Gallo, M.**, Trendafilov, N.T.** & Buccianti, A. (2016) Sparse PCA and investigation of multi-elements compositional repositories: theory and applications. *Environmental and Ecological Statistics*, **23,** 421-434.

Garthwaite, P.H., Fan, Y. & Sisson, S.A. (2016) Adaptive optimal scaling of Metropolis-Hastings algorithms using the Robbins-Monro process. *Communications in Statistics — Theory and Methods*, **45,** 5098-5111.

Garthwaite, P.H. & Koch, I. (2016) Evaluating the contributions of individual variables to a quadratic form. *Australian and New Zealand Journal of Statistics,* **58,** 99-119.

Ghebremichael-Weldeselassie, Y., Whitaker, H.J. & **Farrington, C.P.** (2016) Flexible modelling of vaccine effect in self-controlled case series models. *Biometrical Journal,* **58,** 607-622.

Gower, J.C. (2016) Biometrics and psychometrics: origins, commonalities, and differences. *Journal of Statistical Software,* **73,** Issue 5.

Gower, J.C., Lubbe, S. & Le Roux, N.J. (2016) Biplots: qualitative data. *Wiley Interdisciplinary Reviews: Computational Statistics*, **8**, 82-111.

Jones, M.C. (2016) On bivariate transformation of scale distributions. *Communications in Statistics — Theory and Methods,* **45,** 577-588.

Kafatos, G., Andrews, N.J., **McConway, K.J.,** Maple, P.A.C., Brown, K. & **Farrington, C.P.** (2016) Is it appropriate to used fixed assay cut-offs for estimating seroprevalence? *Epidemiology and Infection,* **144,** 887-895.

Marriott, P., **Sabolová, R., Van Bever, G.** & **Critchley, F.** (2016) The information geometry of sparse goodness-of-fit testing. *Entropy,* **18,** 421.

McConway, K. (2016) Statistics and the media: A statistician's view. *Journalism,* **17,** 49-65.

Noufaily, A., **Farrington, P., Garthwaite, P.,** Enki, D.G., Andrews, N. & Charlett, A. (2016) Detection of infectious disease outbreaks from laboratory data with reporting delays. *Journal of the American Statistical Association,* **111,** 488-499.

Petersen, I., Douglas, I. & **Whitaker, H.** (2016) Self-controlled case series methods: an alternative to standard epidemiological study designs. *British Medical Journal,* **354,** Article i4515.

Ronchetti, E. & **Sabolová R.** (2016) Saddlepoint tests for quantile regression. *Canadian Journal of Statistics*, **44,** 271-299.

Root, A.A., Wong, A.Y.S., **Ghebremichael-Weldeselassie, Y.,** Smeeth, L., Bhaskaran, K., Evans, S.J.W., Brauer, B., Wong, I.C.K., Navaratnam, V. & Douglas, I. (2016) Evaluation of the risk of cardiovascular events with clarithromycin using both propensity score and self-controlled study designs. *British Journal of Clinical Pharmacology,* **82,** 512-521.

Tanner, J., Kiernan, M., **Hilliam, R.,** Davey, S., Collins, E., Wood, T., Ball, J. & Leaper, D. (2016) Effectiveness of a care bundle to reduce surgical site infections in patients having open colorectal surgery. *Annals of the Royal College of Surgeons of England*, **98,** 270-274.

Trendafilov, N.T. & **Gebru, T.** (2016) Recipes for sparse LDA of horizontal data. *Metron,* **74,** 207-221.

Van Bever, G. (2016) Bivariate simplicial tests for randomness. *Statistics and Probability Letters*, **12,** 20-25.

Wong, A.Y.S., Root, A., Douglas, I.J., Chui, C.S., Chan, E.W., **Ghebremichael-Weldeselassie, Y.,** Siu, C.W., Smeeth L. & Wong I.C.K. (2016) Cardiovascular

outcomes associated with use of clarithromycin: population based study. *British Medical Journal,* **352,** Article h6926.

5.2 Book chapters and refereed conference proceedings

Adachi, K. & ***Trendafilov, N.T.*** (2016) Some mathematical notes on comprehensive factor analysis. In *Proceedings of COMPSTAT 2016: 22nd International Conference on Computational Statistics*, Eds: A. Colubi, A. Blanco & C. Gatu, pp. 25-35.

Albers, C.J., ***Critchley, F.*** & ***Gower, J.C.*** (2016) Group average representations in Euclidean distance cones. In *Selected Contributions in Data Analysis and Classification (Edwin Diday Festschrift),* Eds: P. Brito, P. Bertrand, G. Cucumel & F. de Carvalho; Springer, pp. 445-454.

Hickmann, T., Teixeira, Jr, L.A., ***Faria, A.,*** Rodrigues, S.B., Côrrea, J.M., Garcia, E.L. & Aracayo, L.A.S. (2016) Temperature field forecast in concrete dam with the use of ARIMA models and the finite element method. In *Annals of the XLVIII Brazilian Operational Research Symposium,* pp. 2218-2229.

Teixeira Jr., L.A., ***Faria, Á.E.,*** Pereira, R.V. de B., Souza, R.C. & Franco, E.M.C. (2016) Combinação linear wavelet SARIMA-RNA com estágios multiplos na previsão de series temporais. In: *Blucher Marine Engineering Proceedings*, Vol. 2, No. 1, pp. 295–307.

5.3 Others

Critchley, F., Marriott, P., ***Sabolová, R.*** & ***Van Bever, G.*** (2016) Discussion of "Of quantiles and expectiles: consistent scoring functions, Choquet representations and forecast rankings", by W. Ehm, T. Gneiting, A. Jordan & F. Krüger. *Journal of the Royal Statistical Society Series B,* **78,** 536-537.

McConway, K. (2016) The nature of probability. Available at http://www.metoffice.gov.uk/about-us/who/accuracy/forecasts/probability

6. Seminars at The Open University

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

Roman Belavkin (Middlesex University) *Information-theoretic view of mutation and optimal search of sequences in Hamming spaces*

Kateřina Helisová (Czech Technical University in Prague, Czech Republic) *Assessing dissimilarity of random sets through convex compact approximations, support functions and envelope tests*

Michael Betancourt (University of Warwick) *Scalable Bayesian inference with Hamiltonian Monte Carlo*

Cosetta Minelli (Imperial College London) *Statistical methods for imputation of missing variables across studies: Bayesian vs. classical approach*

Irene Petersen (Imperial College London) *Evidence beyond randomised controlled trials*

Kevin Burke (University of Limerick, Ireland) *Non-proportional hazard effects*

Robin Huang (University of Cambridge*) Detecting treatment effects in longitudinal clinical trials for Alzheimer's disease: can we find better strategies than fitting Z-composite scores?*

Michail Papathomas (University of St Andrews) *Exploring dependence between categorical variables: benefits and limitations of using variable selection within Bayesian clustering in relation to searching for interactions*

Jane Hutton (University of Warwick) *Chain event graphs, and understanding complex missing data*

We also held an **internal research morning** at which the speakers were Zillur Shabuz, Karen Vines, Chris Jones, Radka Sabolová, Germain Van Bever and Frank Critchley

7. Conference & workshop papers and posters given

C. Calvert

Higher Education Academy STEM Conference, Nottingham *Improving retention for all students, studying mathematics as part of their chosen qualification, by using a voluntary diagnostic quiz* [with R. Hilliam]

F. Elfadaly

International Society for Bayesian Analysis (ISBA) World Meeting 2016, Sardinia, Italy *Bayesian imputation of missing variables across studies in linear mixed-effects models* [Poster; with P. Garthwaite]

Á.E. Faria

XIII Brazilian Meeting on Bayesian Statistics (EBEB 2016), Belo Horizonte, Brazil *An application of dynamic Bayesian smooth transition autoregressive models to electricity loading in Brazil*

S. Fontanella

30th meeting of the Japanese Society of Computational Statistics, Kyoto, Japan *Constrained versus unconstrained exploratory approach to Bayesian item factor analysis*

COMPSTAT 2016, Oviedo, Spain *An unconstrained approach to rotational indeterminacy in Bayesian exploratory multidimensional IRT models*

P.H. Garthwaite

Workshop on Methods for Big Data, Brunel University, Uxbridge *Model coupling and the emulation of global climate change impacts on crop yields*

see also F. Elfadaly

R. Hilliam

Horizons in STEM Higher Education Conference: Making Connections and Sharing Pedagogy, Leicester *Interactive statistics for a large student cohort* [with C. Calvert]

M.C. Jones

Flexible Statistical Modeling: Past, Present and Future, Ghent, Belgium *On families of distributions with shape parameters: review and extension of a review* [Plenary]

36th Conference on Applied Statistics in Ireland, Limerick, Ireland *Log-location-scale-log-concave distributions for lifetime data*

K.J. McConway

Communicating Risks and Statistics conference, University College London *What statisticians can and can't expect in working with journalists*

R. Sabolová

International Conference on Robust Statistics (ICORS 2016), Geneva, Switzerland *Robustness, information geometry and sparse goodness-of-t testing*

Information Geometry and its Applications IV (IGAIA IV), Liblice, Czech Republic *Binary information geometry: some theory and applications* [Poster]

9th International Conference of the ERCIM Working Group on Computational and Methodological Statistics, Seville, Spain *Recovery of Fisher's linear discriminant subspace by invariant coordinate selection (ICS) methods: latest advances* [Poster]

M.Z. Shabuz

42nd Young Statisticians Meeting, University College London *Contribution of individual variables to the regression sum of squares* [Poster]

N.T. Trendafilov

4th Stochastic Modeling Techniques and Data Analysis International Conference, Valletta, Malta *Algebra of exploratory factor analysis with more variables than observations*

9th International Conference of the ERCIM Working Group on Computational and Methodological Statistics, Seville, Spain *Sparse dimension reduction of multiple data on same observations*

G. Van Bever

3rd Conference of the International Society for Non-Parametric Statistics (ISNPS), Avignon, France *Joint diagonalisation of scatter operators: Hilbertian fourth order blind identification and other applications*

Y.G. Weldeselassie

37th Annual Conference of the International Society for Clinical Biostatistics, Birmingham *Self-controlled case series with multiple event types*

32nd International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Dublin, Ireland *Testing assumptions of the self-controlled case series method* [Poster; with H.J. Whitaker]

H.J. Whitaker

37th Annual Conference of the International Society for Clinical Biostatistics, Birmingham *Testing assumptions of the self-controlled case series method*

32nd International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Dublin, Ireland *Advances in the self-controlled case series method*

see also Y.G. Weldeselassie

8. Other seminars and talks given

P.H. Garthwaite

Royal Statistical Society Highlands Local Group, Aberdeen *'Tweaking' variables to make them uncorrelated*

University of Aberdeen *Learning more from a test statistic*

K.J. McConway

Queen Mary University of London *Statistics and the media: a statistician's view*

N.T. Trendafilov

Stockholm University, Sweden *Archetypal analysis of large data*

G. Van Bever

Middlesex University, London *Hilbertian fourth order blind identification*

Aalto University, Finland *Shape depth*

University of Bath *Hilbertian fourth order blind identification*

H.J. Whitaker

University of Bristol *The self-controlled case series method*

9.   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)

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*

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*

10. Grants awarded

N.T. Trendafilov

Awarded a Santander Mobility Scholarship to the tune of £4,500 for *Sparse principal component analysis in multi-way data context*

11. Other activities

C. Calvert

- was a member of the RSS Professional Affairs Committee

- lead author (with L.J. Brown, D.A. Edward & R. Hilliam) of an Open University eSTEeM project on 'Success Against the Odds'

- co-authored two talks and two posters at the Open University eSTEeM conference

- co-authored an OU Institute of Educational Technology publication

- obtained further internal funding to evaluate retention on an introductory module

F. Critchley

- was a research visitor to Instituto Superior Técnico, Universidade de Lisboa, Portugal

- was a research visitor to Université de Toulouse, France

F. Elfadaly

- is an Honorary Lecturer in the Population Health & Gene Therapy section of the National Heart and Lung Institute, Faculty of Medicine, Imperial College London

Á.E. Faria

- examined a PhD at the University of Warwick

C.P. Farrington

- was a member of the RSS Honours Committee

P.H. Garthwaite

- is member and deputy chair of the RSS Research Section committee

- external examiner for MSc in Applied Statistics with specialisation in Actuarial Science at the University of Mauritius

- examined a PhD at the University of Leeds

R. Hilliam

- co-authored two talks and two posters at the Open University eSTEeM conference

- co-authored an OU Institute of Educational Technology publication

- is a visiting Fellow in Medical Statistics at the University of Derby

K.J. McConway

- is chair of the RSS Examinations Board

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

- academic consultant on one further series of the BBC Radio4/OU series 'More or Less', including one contribution on air

- STEM Ambassador (working with schools and young people to promote STEM subjects), including taking part in a STEM careers session at a school in Milton Keynes

- judge for the 'SET for Britain' parliamentary poster competition for career-young researchers

- wrote invited comments on statistical aspects of scientific papers and ran a training session on statistics for journalists for the Science Media Centre

C.M. Queen

- external examiner for undergraduate and postgraduate statistics modules at the University of Warwick

N.T. Trendafilov

- was a research visitor to the University of Osaka, Japan

- was a research visitor to the Department of Meteorology, Stockholm University, Sweden

S.K. Vines

- external examiner for undergraduate statistics modules at Kingston University London

- continued co-chairing (with C. Hughes) an Open University eSTEeM project on 'Sonification of Depictions of Numerical Data'

- presented a talk and contributed to a workshop at the Open University eSTEeM conference

H.J. Whitaker

- was a member of the scientific organising committee of the 37th Annual Conference of the International Society of Clinical Biostatistics conference held in Birmingham, 2016

- is a member of the International Society for Pharmacoepidemiology working group on Guidance for the Application and Reporting of Self-Controlled Study Designs in Pharmacoepidemiology