

John Hilary Maindonald – Curriculum Vitae

(focusing primarily on work since 1998)

Email: john at [statsresearch.co.nz](mailto:john@statsresearch.co.nz)

<http://www.maths.anu.edu.au/~johnm>

Birth year: 1937.

Home Address: 40 Futuna Close, Karori, Wellington 6012, New Zealand.

Phone +64 4 4769607

Research Interests and Experience:

My interests and experience have extended over many different areas of statistical application; see the selected list of papers that appears below. Particular interests have been:

Statistical perspectives on work that comes broadly under the name “data mining”.

Systems and software for statistical computation, and especially the R system (I am the senior author of a major book that uses the R system to demonstrate modern approaches to data analysis and graphics);

Implications of new developments in statistical methodology for statistical practice;

Statistical issues in experimental design and research planning more generally;

Bioinformation Science, especially Design and Analysis of Microarray experiments.

Other areas of involvement have included: public health, mortality projection, population genetics, disinfestation statistics, and numerical matrix computation.

Professional activity

From 2004, to 2008 I was on the Program Committee for the Australian Data Mining Conferences (AusDM04, AusDM05, ... AusDM0).

I was a member of the Program Committee for the 2008 Australian Statistical Conference in Melbourne, and organized the highly successful 3-day R workshop that preceded this conference.

I have been examiner for seven PhD theses and two Masters theses, for universities in Australia and New Zealand.

Consulting experience in the past 18 years has included:

Collaborative work with researchers from many different areas of application, including horticulture, disinfestation research, entomology, public health, demography, economic history and genomics. I have published numerous papers jointly with other scientists, in Australia, Hawaii, Canada, New Caledonia and N. Z. This has tested my ability, building on prior knowledge, to get quickly to grips with pertinent background science including, most recently, genomics, population genetics and related areas;

Commercial consulting, with the New Zealand Apple and Pear Board, ENZAFRUIT, the New Zealand Police, and Copyright Agency Ltd.

Analysis of expression array data, and provision of advice, for members of the Molecular Genetics and Evolution Group in the Research School of Biological Sciences at Australian National University.

Teaching experience since 1998 has included:

During my employment with the ANU Statistical Consulting Unit over 1998-2003, I taught a wide range of **short courses that were primarily aimed at ANU research students** and staff, but that also had substantial attendances from Government and Industry.

The initial version of my book with John Braun was taken, largely, from lecture notes from these courses.

In 2004 I gave several lectures, and ran weekly two-hour laboratories that used the R system for statistical and data mining exercises, in the Math3346 data mining honours course. The laboratories were repeated in each of the years 2005 - 2010. Over 2006–2010, I was Course Supervisor for this course, contributing around 20 hours of lectures in 20 in each of these years. Laboratory exercises and notes, initially prepared for use in 2004 and submitted for use by APAC institutions, were updated for use in 2005 and 2006. These are available for more general unrestricted use for educational purposes.

In 2002, in collaboration with Biolateral Ltd, I ran successful courses on analysis of microarray data.

I have fronted numerous workshops on the R system, starting in 2003: Queensland Branch of the Statistical Society of Australia (SSAI) meeting; 2005: University of Wollongong, Victorian Branch of the SSAI; 2006: Swinburne University of Technology, Australian Government Antarctic Division; 2007: UQ Institute of Molecular Biology, Lund University Centre for Mathematical Statistics, Queensland Branch of the SSAI, Canberra Branch of the SSAI, New South Wales Department of Primary Industry (1 course also in 2008), and CSIRO Plant Industry; 2008: Queensland Treasury Office of Economic and Statistical Research, Australian National University Extreme Values preparatory workshop, Charles Sturt University School of Animal and Veterinary Sciences, CSIRO Sustainable Ecosystems; 2009 and 2010: Charles Darwin University; 2009-2015: ACSPRI (Australian Council for Social and Political Research Incorporated, 10 5-day courses); 2010: NSW Dept of Environment and Climate Change; 2012: SSAI (Statistical Society of Australia Incorporated); 2012: Defence Science and Technology Organisation; 2013 and 2014: Genomics Discovery Unit, ANU (1-week courses); 2014: Civil Air Services Australia: 2 1-week courses.

Present position

Visiting Research Fellow (retired), Centre for Mathematics and Its Applications, ANU.

Previous Positions

19-24 July 2009: Member of Faculty and Keynote Speaker -- University of Warwick Summer School on Model Validation (Mathematics in Biomedical Engineering).

9 May 2007 – 16 June 2007: Visiting professor, Lund Institute of Technology, Sweden. (Gave PhD course that was funded by the STINT program).

2001-2005: Research Fellow, Centre for Bioinformation Science, ANU

1998-2001: Statistical Consultant/Research Scientist (ANU Statistical Consulting Unit)

1996-1997: Statistical Consultant/Research Scientist (Statistics Department and Centre for Clinical Epidemiology and Biostatistics, University of Newcastle)

1992-1996: Biometrician/Research Scientist, Horticulture & Food Research Institute of NZ Ltd

1978-1992: Officer in Charge of Mt Albert Substation. Applied Mathematics Division, NZ Dept of Scientific and Industrial Research.

1971- 1978: Biometrical Consultant, Faculty of Science, Victoria University, Wellington, NZ.

Academic Awards

1958: Senior Scholar in Mathematics, University of New Zealand (Auckland)

Academic Qualifications:

- 1958 B.Sc., University of N. Z. (Auckland).
 1959 M.Sc. with first class honours in mathematics, University of N. Z. (Auckland)
 1962 Diploma in Teaching, Auckland Teachers' College
 2004 Ph.D. (by submission of published work) The Australian National University

Selected Joint Papers

- Boot, H. M. and **Maindonald, J. H.** 2008. New estimates of age- and sex- specific earnings and the male-female earnings gap in the British cotton industry, 1833-1906. *Economic History Review*, 61: 380-408.
- Daley, D.J.; **Maindonald, J.H.** 1989. A unified view of models describing the avoidance of Superparasitism. *IMA Journal of Mathematics Applied in Medicine and Biology*. 6, 161 - 178.
- Driscoll D. A., Smith A. L., Blight S. & Maindonald J. (2012) Reptile responses to fire and the risk of post-disturbance sampling bias. *Biodiversity and Conservation* 21: 1607-1625.
- Fazey, I., Salisbury, J., Lindenmayer, D.B., Douglas, R. and **Maindonald, J.H.** 2004: Can conservation learn from medicine about bridging the gap between scientists and practitioners? *Environmental Conservation* 31: 190–198.
- Grasso, L.G., **Maindonald, J.**, Rudd, S., Hayward, D.C., Saint, R., Miller, D.J. and Ball, E.E. 2008. Microarray analysis identifies candidate genes for key roles in coral development. *BMC Genomics* 9:540. doi:10.1186/1471-2164-9-54
- Hales, S., de Wet, N., **Maindonald, J.** and Woodward, A. 2002. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. *The Lancet*. Published online August 6, 2002.
- Harker, F. R.; **Maindonald, J.H.** 1994. Ripening of nectarine fruit. *Plant Physiology* 106: 165-171.
- McMahon CR, Buscot M-J, Wiggins NL, Collier N, **Maindonald JH.** 2011 A Two-Phase Model for Smoothly Joining Disparate Growth Phases in the Macropodid *Thylogale billardierii*. *PLoS ONE* 6(10): e24934. doi:10.1371/journal.pone.0024934
- Maindonald, J.H.**; Waddell B.C.; Birtles D.B. 1991. Response to Methyl Bromide Fumigation of Codling Moth (Lepidoptera: Tortricidae) on Cherries. *Journal of Economic Entomology* 85: 1222-1230.
- Maindonald, J.H.**; Markwick N.P. 1986. The avoidance of superparasitism in four species of parasitic wasp — mathematical models and experimental results. *Researches on Population Ecology* 28: 1-16.
- Maindonald, J.H.** and Finch G.R. 1986. Apple Transport in Wooden Bins. *N. Z. Journal of Technology* 2: 171-177.
- Maindonald, J.H.** and Burden, C. J. 2005. Selection bias in plots of microarray or other data that have been sampled from a high-dimensional space. In R. May and A. J. Roberts, eds., *Proceedings of 12th Computational Techniques and Applications Conference CTAC-2004*, volume 46, pp. C59–C74. <http://anziamj.austms.org.au/v46/CTAC2004/Main> [March 15, 2005]
- Maindonald, J.** and Richardson, A.E. 2004. This passionate study: a dialogue with Florence Nightingale. *Journal of Statistics Education* 12, Number 1. Published online at <http://www.amstat.org/publications/jse/v12n1/maindonald.html>

Maindonald, J.H.; Waddell, B.C.; Petry, R.J. 2001. Apple cultivar effects on codling moth (Lepidoptera: Tortricidae) egg mortality following fumigation with methyl bromide. *Postharvest Biology and Technology* 22: 99-110.

Encyclopedia Articles and Book Chapter

Maindonald, J.H. 2005. Algorithm. *Encyclopedia of Biostatistics, 2nd edn, 149-154.*

Lawrence, C.J. and **Maindonald, J.H.** 2005. Computer Languages and Programs. *Encyclopedia of Biostatistics, 2nd edn, 1077-1081.*

Maindonald, J.H. and Stott, H.J. 2005. Confidentiality and Computers. *Encyclopedia of Biostatistics, 2nd edn, 1109-1113.*

Maindonald, J.H. 2005. Matrix Computations. *Encyclopedia of Biostatistics, 2nd edn: 3050-3058.* Wiley, Chichester U. K.

Maindonald, J., Pittelkow, Y. and Wilson, S. 2003. Some considerations for the design of microarray experiments. In *Science and Statistics*, ed. D.R. Goldstein, pp.367-390. Institute of mathematical Statistics, Lecture Notes series, volume 40.

Selected Papers

Maindonald, J. H. 2011. Technology-enabled advance in the worlds of statistics, machine learning and data mining. *New Zealand Science Review* 68: 113-116.

Maindonald, J., 2006. Data mining methodological weaknesses and suggested fixes. In P. J. Christen; P. J. Kennedy; J. Li; S. J. Simoff; and G. J. Williams, eds., *Fifth Australasian Data Mining Conference (AusDM2006)*, volume 61 of CRPIT, pp. 9–16. ACS, Sydney, Australia.

Maindonald, J. H. 2005. Data, science and new computing technology. *New Zealand Journal of Science* 62: 126-128.

Maindonald, J.H. 2003. The Role of Models in Predictive Validation (Statistics for Budding Data Miners). Invited Paper, ISI Meeting, Berlin, 2003.

Maindonald, J.H. 1990. Avoidance-modified Urn Models and Parasitic Wasps. *The Mathematical Scientist* 15: 128-136.

Maindonald, J.H. 1992. Statistical Design, Analysis and Presentation Issues. *N. Z. Journal of Agricultural Research* 35: 121-141.

Books

Maindonald, J.H. and Braun, W.J. 2010. *Data Analysis and Graphics Using R. An Example-Based Approach.* 3rd edition, Cambridge University Press, June 2010. [see the web page <http://wwwmaths.anu.edu.au/~johnm/r-book.html>].

Maindonald, J.H. 1984. *Statistical Computation*, 370+xviii pp., Wiley. (Translated into Russian, 1988).

Book Reviews (2006 and later)

(ISR= International Statistical Review; JSS= Journal of Statistical Software)

Paul S. P. Cowpertwait, Andrew V. Metcalfe. *Introductory Time Series with R.* ISR (2010) 78:452-453.

Bertrand Clarke, Ernest Fokoué, Hao Helen Zhang. *Principles and Theory for Data Mining and Machine Learning.* ISR (2011) 79: 119–120.

- Ismael Vaccaro, Eric Alden Smith, Shankar Aswani (Editors). *Environmental Social Sciences: Methods and Research Design*. ISR (2011) 79: 496.
- Kevin B. Korb, Ann E. Nicholson. *Bayesian Artificial Intelligence, Second Edition*. ISR (2011) 79:497.
- François Husson, Sébastien Lê, Jérôme Pagès. *Exploratory Multivariate Analysis by Example Using R*. ISR (2011) 79:498.
- Robert K. Yin. *Qualitative Research from Start to Finish*. ISR (2011) 79:499-500.
- Heping Zhang, Burton H. Singer. *Recursive Partitioning and Applications, Second Edition*. ISR (2011) 79:500-501.
- Christopher M. Bishop, *Pattern Recognition and Machine Learning*. JSS 17, Book Review 5, 2007.
- Dianne Cook; Deborah F Swayne: *Interactive and Dynamic Graphics for Data Analysis: With R and GGobi*. ISR (2008), 76: 436–463
- Simon N. Wood, *Generalized Additive Models: An Introduction with R*. JSS16, Book Review 3 2006.
- Alan J Izenman, *Modern Multivariate Statistical Techniques: Regression, Classification and Manifold Learning*. JSS 29, Book Review 12, 2009.
- Richard A. Berk, *Statistical Learning from a Regression Perspective*. JSS 29, Book Review 12, 2009.
- David R Anderson, *Model Based Inference in the Life Sciences: A Primer on Evidence*, ISR (2009), 77: 479-480
- Jonathan D. Cryer, Kung-Sik Chan, *Time Series Analysis With Applications in R, Second Edition*. ISR (2009), 77: 300-301.
- Julien Claude, “Morphometrics with R”. ISR (2009), 77: 302-303
- Yosef Cohen, Jeremiah Y. Cohen: “Statistics and Data with R: An Applied Approach Through Examples”. ISR (2010), 78, 1, 141-142.
- Jane Horgan. “Probability with R: An Introduction with Computer Science Applications”. ISR (2010), 78, 1, 146.
- Paul S. P. Cowpertwait, Andrew V. Metcalfe. “Introductory Time Series with R”. ISR (2010), 78, 1, 452-453.
- Bertrand Clarke, Ernest Fokoue, Hao Helen Zhang. “Principles and Theory for Data Mining and Machine Learning”. ISR (2011), 79: 119-120.
- Robert A. Muenchen. *R for SAS and SPSS Users, 2nd edn*. ISR (2012), 80: 200–201.
- Graham Williams. *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*. ISR (2012) , 80: 199–200.
- Yu-Kang Tu, Mark S. Gilthorpe. *Statistical Thinking in Epidemiology*. ISR (2014), 81: 154.
- Noel A. Card. *Applied Meta-Analysis for Social Science Research*. ISR (2014), 81:165-166.
- Adelchi Azzalini, Bruno Scarpa. *Data Analysis and Data Mining: An Introduction*. ISR (2014), 81: 170-171.
- Giuseppe Modica, Laura Poggiolini. *A First Course in Probability and Markov Chains*. ISR (2014), 82: 148
- Yuelin Li, Jonathan Baron. *Behavioral Research Data Analysis with R*. ISR (2014), 82: 147–148

Andrzej Galecki, Tomasz Burzykowski: Linear Mixed-Effects Models Using R: A Step-by-Step Approach. *ISR* (2014), 82: 315–317

Xihong Lin et al, editors: Past, Present, and Future of Statistical Science. *ISR* ((2015), 83: 332-334.

R Packages

The following R packages are available from the Comprehensive R Archive Network (CRAN):

DAAG: Data and Functions to accompany Data Analysis and Graphics Using R (with John Braun; CUP, 3rd edition, 2010)

gamclass: Functions and Data for a Course on Modern Regression and Classification.

DAAGbio: Data and Functions supplementary to DAAG, intended for use in courses that focus on expression array and other applications in molecular biology.

hddplot: Use known groups in high-dimensional data to derive scores for low-dimensional plots that are unbiased with respect to selection effects.

hwde: Models and tests for departure from Hardy-Weinberg equilibrium.