
# The R User Conference 2011

## General Information

THE UNIVERSITY OF
WARWICK

# Welcome!

Dear useRs,

Welcome to the R user conference, useR! 2011, hosted by the Department of Statistics, University of Warwick. Following six successful meetings, the useR! conference is focused on: R as the 'lingua franca' of data analysis and statistical computing; providing a platform for R users to discuss and exchange ideas on how R can be used for statistical computations, data analysis, visualization and exciting applications in various fields; and giving an overview of the new features of the rapidly evolving R project. The program consists of invited talks discussing new R developments and interesting applications of R, and user-contributed presentations reflecting the wide range of fields in which R is used to analyze data.

The following pages provide you with information regarding the R User Conference, useR! 2011. We hope that you find the conference enjoyable and stimulating, and also hope that you enjoy your time here at the University of Warwick.

With kind regards from the useR! 2011 Organizing Committee.

| | |
|---|---|
| John Aston | Julia Brettschneider |
| David Firth | Ashley Ford |
| Ioannis Kosmidis | Tom Nichols (Chair) |
| Jennifer Rogers | Elke Thönnes |
| Heather Turner | |

MANGOSOLUTIONS
data analysis that delivers

Mango Solutions is delighted to sponsor UseR! 2011. We are the largest provider of R Services in the world today; our portfolio includes Application Building, Consultancy, R Training, R Validation and Imaging Services. Headquartered in the UK, we have offices in both China and Switzerland and a US office opening later this year.

Generally we work within 5 main areas where there is a particularly good match for our abilities and experience – these are the Pharmaceutical, Finance, Energy, Sensory and CRM Industries. Our client list includes F. Hoffmann-La Roche, Lloyd's of London, BP, Kraft, Novartis, GlaxoSmithKline, Aviva, Google and HSBC to name but a few.

## Consultancy

We provide a range of consulting options from ongoing consultative engagements to smaller one-off projects or tasks. Statistical support can be provided on an hourly, daily or annual basis.


Examples of our consultancy work include:

- Analysis of a customer's data utilizing complex analysis algorithms
- Statistical programming - the creation of new and innovative algorithms using R Programming
- Database modelling and optimizing for customers wishing to profile their data
- Business consultancy, assistance and advice for organizations looking at their workflows, strategies and software choices
- Code maintenance and updating of legacy code
- Validation services for analysis taking place in regulated environments

## Training

Our public courses are scheduled throughout the year, worldwide. Additionally we run private courses for customers as diverse as Morgan Stanley and Novartis. We can provide training on specific software products or more varied training on statistics and general data analysis. Our R Training courses include but are not limited to Introduction to R, Advanced R, R Graphics, R for Pharmacometrics, and R for Financial Data Analysis.

## Contact Us

Mango Solutions
Unit 2, Greenways Business Park
Chippenham,
Wiltshire SN15 1BN
Tel: +44 (0)1249 767 700
Email: info@mango-solutions.com
www.mango-solutions.com

# Programme Overview

**Monday 15th August**

| | |
|---|---|
| 18:00 - 19:30 | Dinner[1] |
| 19:30 - 21:00 | Opening Mixer, sponsored by CRiSM |

**Tuesday 16th August**

| | |
|---|---|
| 08:45 - 09:00 | Introductory Remarks |
| 09:00 - 09:45 | Invited Talk - Brian Ripley |
| 09:50 - 10:50 | Kaleidoscope I |
| 10:50 - 11:15 | Coffee Break, sponsored by GCHQ |
| 11:15 - 12:35 | Focus I |
| 12:35 - 14:00 | Lunch |
| 14:00 - 14:45 | Invited Talk - Ulrike Grömping |
| 14:45 - 15:30 | Invited Talk - Jonathan Rougier |
| 15:30 - 16:00 | Coffee Break, sponsored by Google |
| 16:00 - 17:00 | Focus II |
| 17:05 - 18:05 | Lightning Talks |
| 18:30 - 20:00 | Dinner[1] |
| 20:00 - 23:00 | Poster Session, sponsored by Revolution Analytics |

**Wednesday 17th August**

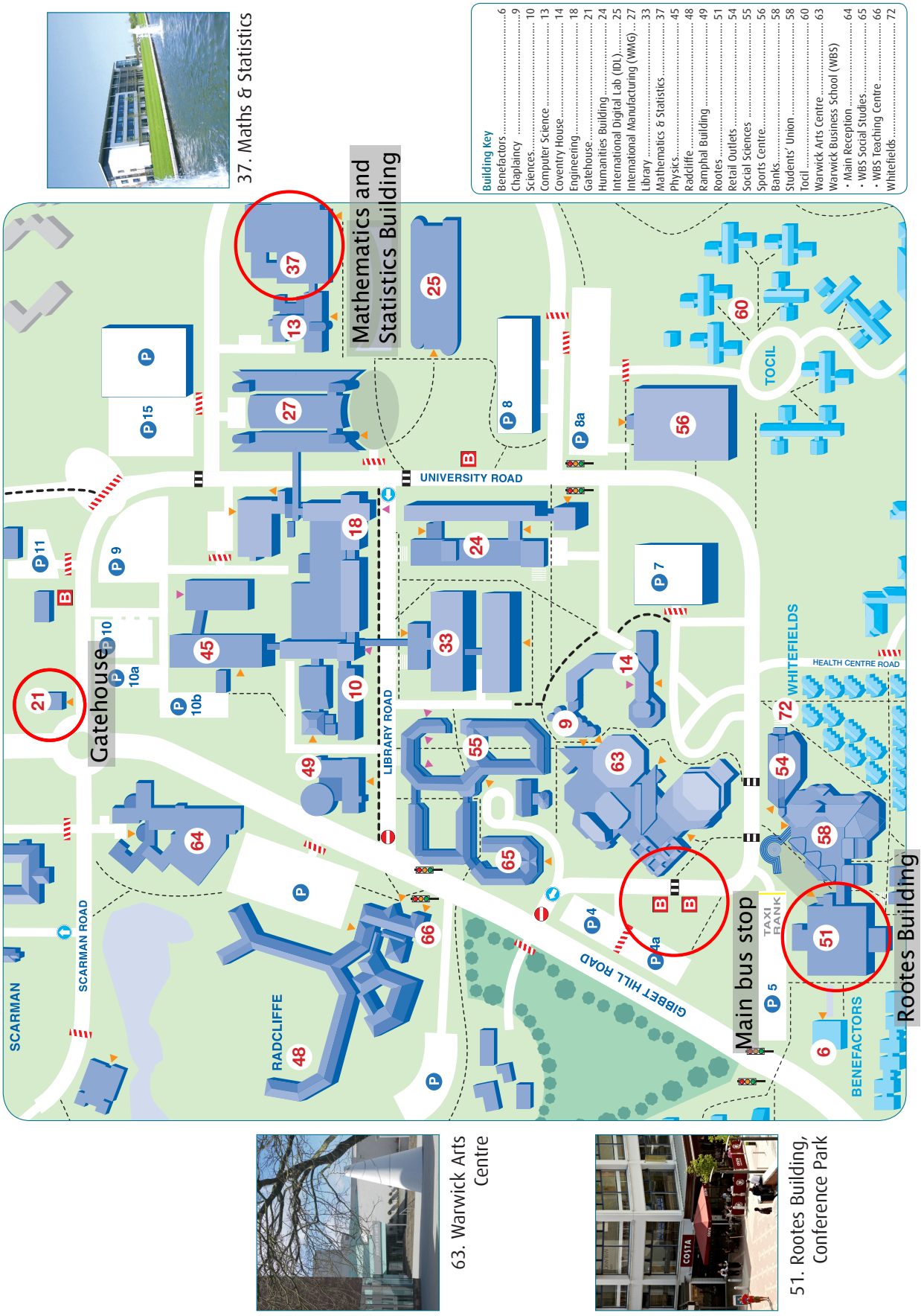
| | |
|---|---|
| 09:00 - 09:45 | Invited Talk - Lee Edlefsen |
| 09:50 - 10:50 | Kaleidoscope II |
| 10:50 - 11:15 | Coffee Break, sponsored by Mango Solutions |
| 11:15 - 12:35 | Focus III |
| 12:35 - 14:00 | Lunch |
| 14:00 - 14:45 | Invited Talk - Adrian Bowman |
| 14:45 - 15:30 | Invited Talk - Brandon Whitcher |
| 15:30 - 16:00 | Coffee Break, sponsored by Department of Statistics, University of Warwick |
| 16:00 - 17:00 | Focus IV |
| 17:05 - 18:05 | Focus V |
| 19:30 - 23:00 | Conference Dinner[2], sponsored by RStudio |

**Thursday 18th August**

| | |
|---|---|
| 09:00 - 09:45 | Invited Talk - Wolfgang Huber |
| 09:50 - 10:50 | Focus VI |
| 10:50 - 11:15 | Coffee Break, sponsored by Centre for Scientific Computing, University of Warwick |
| 11:15 - 12:35 | Focus VII |
| 12:35 - 14:00 | Lunch |
| 14:00 - 15:00 | Kaleidoscope III |
| 15:05 - 15:50 | Invited Talk - Simon Urbanek |
| 15:50 - 16:00 | Closing Remarks |

[1] note that dinner will only be provided for attendees staying in on-campus accommodation
[2] admission to the Conference Dinner will be via ticket only

The central campus of the University of Warwick.



# Travel information

## Venue

The conference takes place in the

Mathematics & Statistics Building,
University of Warwick,
Coventry, CV4 7AL,
UK

The conference venue is building no. **37** on the campus map (upper right corner) and it is also known as the "Zeeman Building".

## Note:

The University is **not** in the town of Warwick (which is about 8 miles away), and it is **not** the same as Coventry University (which is a different university, located in the centre of Coventry). This is important when telling taxi drivers where you want to be!

## Travelling to the University of Warwick

Information on getting to the University of Warwick from Coventry, as well as from other directions locally and further afield, can be found at `http://www2.warwick.ac.uk/about/visiting/directions/` (and via links provided from the navigation bar of that page).

## Parking

Complimentary car parking is available for conference attendees in the allocated car parks on campus (7, 8a and 15). Car park 7 is nearest to the Rootes Building and car parks 8a and 15 are nearest to the Mathematics & Statistics building. When you enter the car park, take the token from the machine at the entrance, which will need to be exchanged for your car park exit code at Rootes Reception. The campus map displays the location of the aforementioned car parks. Disabled parking spaces are available close to the entrance of main buildings. The useR! help desk can provide further information regarding car parking arrangements.

# General information

## Accommodation

Check in for accommodation is at Rootes Reception, located in the Rootes Building (building no. 51 on the campus map).

Rooms will be available for check in after 15:00, and all rooms must be vacated by 9:30 on your day of departure. Note that storage facilities for your luggage are available prior to check in and after you have checked out on your final day. Please ask at Rootes Reception for this. Rootes Reception will close at 23:00 each day. The staff at the Gatehouse (building no. 21 on the campus map) can provide local information to participants that arrive after this time. If you plan to arrive after this time it is advised that you contact the conference organisers before your arrival so that special arrangements can be made.

## Conference registration

Registration for the conference will take place on Sunday in the Rootes Building and on Monday to Thursday in the main atrium of the Mathematics & Statistics Building (no. 37 on the map). The registration desk will be available at the following times:

- Sunday, 14th August 2011, 17:00 – 18:00
- Monday, 15th August 2011, 08:00 – 18:00
- Tuesday, 16th August 2011, 08:00 – 18:00
- Wednesday, 17th August 2011, 08:30 – 18:00
- Thursday, 18th August 2011, 08:30 – 17:00

During the above times, the staff at the registration desk will be happy to provide any assistance and information to participants.

## Internet access

If you have a WiFi enabled laptop you may access basic internet services[3] by connecting to the "conferences" wireless network and starting up a web browser. You will need to enter a unique username and password which will be given to you at registration. For participants staying on-campus, cabled internet connection is available in every bedroom. Participants will need to supply their own Ethernet cables for this. There will be a small number of Ethernet cables available at the registration desk on request. If necessary Ethernet cables can be purchased from Costcutter shop (in building 54 on the map).

The use of any campus network connection is subject to the JANET Acceptable Use Policy. Please see `http://www.ja.net/company/policies/janet-aup.html`.

[3] but note that sending email via SMTP is not allowed

## Meals

During the conference, coffee will be provided in the main atrium of the Mathematics & Statistics Building and in the Mathematics Workroom, A0.05. Lunches will be served in Rootes Restaurant, located on the first floor of the Rootes Building (building no. 51 on the map). The Opening Mixer (Monday 19:30 - 21:00) will be held in *The Bar* on the first floor of the Rootes Building, and the Poster Reception (Tuesday 20:00 - 23:00) will be in the Panorama Suite on the second floor of the Rootes Building. Beer, wine, non-alcoholic drinks and light snacks will be served at both.

For those staying in campus accommodation, breakfast and dinner are included and will also be served in Rootes Restaurant. Breakfast is served from 07:30 - 09:30 each morning. Dinner will be served at 18:00 on Monday 15th August and at 18:30 on Tuesday 16th August. The conference dinner will take place on Wednesday 17th August at 19.30 and will be held in the Panorama Suite of the Rootes Building.

For participants who have not registered for dinners, some dining options include the *Dirty Duck* on the second floor of the Students Union building (building no. 58 on the campus map), and the Varsity pub, which is on Gibbet Hill Road, a 5 minute walk away from the Mathematics & Statistics Building.

Other options for refreshments and light food options include the *Café Bar* in the Warwick Arts Centre (building no. 63 on the campus map, open 09:00 - 21:00 from Monday to Saturday, and 15:00 - 19:30 on Sunday), and *Costa Coffee* on the ground floor of the Rootes Building (building no. 51 on the campus map, open 08:00 - 16:00 from Monday to Friday, and 10:00 - 16:00 on Saturday and Sunday). Additionally, *The Bar* serves draught beers, a good selection of bottled beers, wines, spirits, soft drinks and a variety of teas and coffees.

## Messages

The telephone number for colleagues or family to leave an urgent message for you during office hours is +44(0)24 765 74812. For emergency messages outside these times, please call the main University Switchboard on +44(0)24 765 23523.

## Smoking Policy

Smoking is forbidden anywhere in University buildings, including on-campus accommodation and bar areas. Furthermore, smoking is not permitted outside buildings if the smoke is likely to drift through open doors or windows, or inconvenience people entering or leaving the building. A smoking area is provided around the corner from the main entrance to the Zeeman Building.

# Program Committee and Conference Sponsors

## Program Committee

| | |
|---|---|
| Ramón Diaz-Uriarte | John Fox |
| Romain François | Robert Gramacy |
| Paul Hewson | Torsten Hothorn |
| Kate Mullen | Brian Peterson |
| Thomas Petzoldt | Anthony Rossini |
| Barry Rowlingson | Carolin Strobl |
| Stephan Theussl | Heather Turner (Chair) |
| Hadley Wickham | Achim Zeileis |

## Conference Sponsors

Support from our sponsors is gratefully acknowledged.

### Conference sponsors

Centre for Scientific Computing, University of Warwick
CRiSM, University of Warwick
Department of Statistics, University of Warwick
GCHQ
Google
Mango Solutions
Open Analytics
Revolution Analytics
RStudio
TIBCO

### Principal sponsors of the bursary scheme and/or scholarships

CRiSM, University of Warwick
Revolution Analytics
RStudio

### Major contributors to the bursary scheme

GCHQ
Google
John Wiley & Sons
Mango Solutions
Open Analytics
Springer
Taylor and Francis
TIBCO

## Exhibitors

GCHQ, Mango Solutions, Open Analytics, Revolution Analytics and RStudio will be exhibiting in the main atrium. The publishers Cambridge University Press, John Wiley & Sons, Springer, and Taylor and Francis will be exhibiting in the adjacent Mathematics Workroom.

# Medical and emergency information

## Medical assistance

The University Health Centre is open Monday-Friday 09:00 - 13:00 and 14:00 - 17:00. Visitors in need of emergency assistance should contact Security on internal extension 22222.

## Emergency services and fire procedures

For help in an emergency dial 22222 from any internal telephone and your call will be directed appropriately. Visitors are asked to familiarise themselves with the University's fire procedures which are displayed in each room.

*On discovering a fire in a building:*
Raise the alarm by breaking the glass in the nearest *Break Glass* point.

*On hearing the fire alarm:*
Stop what you are doing.
Leave by the nearest Fire Exit.
Walk calmly, do not run.
Do not stop to collect personal belongings.
Make your way to the nearest evacuation point, standing well clear of the building.
Do not re-enter the building until told to do so by the Fire Service or University Security Staff.

## Reporting incidents

In the event of an accident or other incident occurring on University premises, please report immediately to Rootes Reception (Ext. 22280) who will then take appropriate action. Please dial 22222 from an internal phone, or +44(0)24 765 22222 externally to be in direct contact with our 24–hour security staff.

# Tutorials

## Monday 15th August

Participants are welcome to arrive at 08:00 for registration. The morning tutorials will be held from 09:00 - 12:30, with a break for coffee at 10:30 - 11:00. Lunch (provided for all tutorial participants) is scheduled for 12:30 - 14:00. The afternoon tutorials will be held from 14:00 - 17:30, with a break for coffee at 15:30 - 16:00.

### Morning Tutorials

- **Douglas Bates:** Fitting and Evaluating Mixed Models using lme4
- **Roger Bivand:** Handling and Analyzing Spatio-temporal Data in R
- **Marine Cadoret and Sébastien Lê:** Analysing Categorical Data in R
- **Søren Højsgaard:** Graphical Models and Bayesian Networks with R
- **G. Jay Kerns:** Introductory Probability and Statistics using R
- **Max Kuhn:** Predictive Modeling with R and the caret Package
- **Fausto Molinari, Enrico Branca and Rocco Claudio Cannizzaro:** R-Adamant: Applied Financial Analysis and Risk Management
- **Antony Unwin:** Graphical Data Analysis
- **Brandon Whitcher, Jörg Polzehl, Karsten Tabelow:** Medical Image Analysis for MRI

### Afternoon Tutorials

- **Stephen Eglen:** Emacs Speaks Statistics
- **Andrea Foulkes:** High-dimensional Data Methods with R
- **Frank E. Harrell, Jr:** Regression Modeling Strategies using the R package rms
- **Martin Morgan:** Bioconductor for the Analysis of High-throughput Genomic Data
- **Paul Murrell:** Introduction to Grid Graphics
- **Giovanni Petris:** State Space Models in R
- **Karline Soetaert and Thomas Petzoldt:** Simulating Differential Equation Models in R

## Opening Mixer

All conference attendees are welcome to join us for drinks at *The Bar* on the first floor of the Rootes Building, 19:30 - 21:00. Beer, wine, non-alcoholic drinks and light snacks will be available for all attendees.

# Conference Timetable

## Tuesday 16th August

| | | |
|---|---|---|
| **08:45** | **Introductory Remarks, MS.01/MS.02** | |
| **09:00** | **Invited Talk, MS.01/MS.02, Chair: Heather Turner** | |
| | Brian Ripley | The R Development Process |
| **09:50** | **Kaleidoscope Ia, MS.03, Chair: Dieter Menne** | |
| | Claudia Beleites | Spectroscopic Data in R and Validation of Soft Classifiers: Classifying Cells and Tissues by Raman Spectroscopy |
| | Jonathan Rosenblatt | Revisiting Multi-Subject Random Effects in fMRI |
| | Zoe Hoare | Putting the R into Randomisation |
| | **Kaleidoscope Ib, MS.01, Chair: Simon Urbanek** | |
| | Teh Amouh | Bringing the power of complex data analysis methods into R |
| | Markus Gesmann | Using the Google Visualisation API with R |
| | Matthew S. Shotwell | Experimenting with a `tty` connection for R |
| | **Kaleidoscope Ic, MS.02, Chair: Achim Zeileis** | |
| | David Smith | The R Ecosystem |
| | E. James Harner | Rc2: R collaboration in the cloud |
| | J.J. Allaire | RStudio: Integrated Development Environment for R |
| **10:50** | **Coffee, Main atrium and Mathematics Workroom, A0.05** | |
| **11:15** | **Portfolio Management, B3.02, Chair: Patrick Burns** | |
| | Jagrata Minardi | R in the Practice of Risk Management Today |
| | Rocco Claudio Cannizzaro | Stress Testing with R-Adamant |
| | Jovan Njegic | Handling Multiple Time Scales in Hedge Fund Index Products |
| | **Bioinformatics and High-Throughput Data, B3.03, Chair: Hervé Pagès** | |
| | Thierry Onkelinx | AFLP: generating objective and repeatable genetic data |
| | Przemysław Biecek | The nz.seq package for handling genetic sequences in the Netezza Performance Server |
| | Stefanie Tauber | Classification of Coverage Patterns |
| | Norma Coffey | Finite Mixture Model Clustering of SNP Data |
| | **High Performance Computing, MS.03, Chair: Stefan Theussl** | |
| | Willem Ligtenberg | GPU computing and R |
| | Pragneshkumar B. Patel | Deploying and Benchmarking R on a Large Shared Memory System |
| | Kazutaka Doi | The CUtil package which enables GPU computation in R |
| | Manuel Quesada | OBANSoft: integrated software for Bayesian statistics and high performance computing with R |
| | **Reporting Technologies and Workflows, MS.01, Chair: Martin Mächler** | |
| | Nathan Uyttendaele | R2wd: writing Word Documents from R |
| | Dmitry Bliznyk | FRAD - Fast Results Analysis and Display |
| | Andreas Leha | The Emacs Org-mode: Reproducible Research and Beyond |
| | Peter Baker | Efficient data analysis workflow in R |
| | **Teaching, MS.02, Chair: Jay G. Kerns** | |
| | Ian Holliday | Teaching Statistics to Psychology Students using Reproducible Computing package RC and supporting Peer Review Framework |
| | Kristian Hovde Liland | Teaching applied statistics to students in natural sciences using the R Commander |
| | Achim Zeileis | Automatic generation of exams in R |
| **12:35** | **Lunch, Rootes Restaurant** | |

Kaleidoscope I

Focus I

| | | |
|---|---|---|
| **14:00** | **Invited Talk, MS.01/MS.02, Chair: David Firth** | |
| | Ulrike Grömping | Design of Experiments in R |
| **14:45** | **Invited Talk, MS.01/MS.02, Chair: David Firth** | |
| | Jonathan Rougier | Nomograms for visualising relationships between three variables |
| **15:30** | **Coffee, Main atrium and Mathematics Workroom, A0.05** | |
| **16:00** | **Modelling Systems and Networks, B3.02, Chair: Jonathan Rougier** | |
| | Rachel Oxlade | An S4 Object structure for emulation - the approximation of complex functions |
| | Robert J. B. Goudie | The structmcmc package: Structural inference of Bayesian networks using MCMC |
| | Christophe Dutang | Computation of generalized Nash equilibria |
| | **Computational Physics and Chemometrics, B3.03, Chair: Claudia Beleites** | |
| | Irina Roslyakova | Segmented regression in thermo-physics modeling |
| | Katarina Domijan | Sparse Bayesian kernel projections for classification of near-infrared spectroscopy data |
| | Jinhie Skarda | Recovering Signals and Information From Radio Frequencies Using R (A high school student's experience) |
| | **Visualisation, MS.04, Chair: Antony Unwin** | |
| | Andrej Blejec | animatoR: dynamic graphics in R |
| | Richard M. Heiberger | Graphical Syntax for Structables and their Mosaic Plots |
| | Alexander Pilhöfer | RMB: Visualising categorical data with Relative Multiple Barcharts |
| | **Official and Social Statistics, MS.03, Chair: Mark van der Loo** | |
| | Neil Diamond | ObsSensitivity: An R package for power analysis for sensitivity analyses of Observational Studies |
| | Jason Bryer | Visualizing Multilevel Propensity Score Analysis |
| | **Dimensionality Reduction and Variable Selection, MS.01, Chair: Matthias Schmid** | |
| | Marie Chavent | ClustOfVar: an R package for the clustering of variables |
| | Jürg Schelldorfer | Variable Screening and Parameter Estimation for High-Dimensional Generalized Linear Mixed Models Using l1-Penalization |
| | Benjamin Hofner | gamboostLSS: boosting generalized additive models for location, scale and shape |
| | **Business Management, MS.02, Chair: Enrico Branca** | |
| | Marlene S. Marchena | SCperf: An inventory management package for R |
| | Pairach Piboonrungroj | Using R to test transaction cost measurement for supply chain relationship: A structural equation model |
| | Fabrizio Ortolani | Integrating R and Excel for automatic business forecasting |
| **17:05** | **Lightning Talks** | |
| **18:30** | **Dinner (Rootes Restaurant, for participants staying in campus accommodation)** | |
| **20:00** | **Poster Session, Panorama Suite, Rootes Building** | |

Focus II

# Lightning Talks

## Community and Communication, MS.02, Chair: Ashley Ford

**Stefan Theussl:** R-Forge: Collaborative Software Development in the R Community

**Barry Rowlingson:** Why R-help Must Die! My StackOverfloweth!

**George Zhang:** China R user conference

**Heather Turner:** The R Journal

**Tal Galili:** Blogging and R - present and future

**Timothee Carayol:** R in the cloud with RGoogleDocs and RApache

**Markus Schmidberger:** Get your R application onto a powerful and fully-configured Cloud Computing environment in less than 5 minutes.

**Ian Holliday:** Reproducible Computing package RC: applications in psychology education and research

**Eirini Koutoumanou:** Teaching R to Non Package Literate Users

**Randall Pruim:** Teaching Statistics using the mosaic Package

## Statistics and Programming, MS.01, Chair: Elke Thönnes

**G. Jay Kerns:** Org-mode and reproducible research

**Michael K. Stanko:** Use R for Innovative Benchmarking with DEA (Data Envelopment Analysis)

**Toby Dylan Hocking:** Fast, named capture regular expressions in R2.14

**John C. Nash:** Developments in optimization tools for R

**Christophe Dutang:** A Unified Approach to fit probability distributions

**Marlene Müller:** R scripts for nonparametric function estimation

**John Davey:** Genetic markers for all: analysing RAD Sequencing data with R

**Véronique Storme:** Association study on metabolomic data

**Roger Dean:** Time Series analysis of Music Perception

**Branimir Dolicki:** Time Series Modeling of Daily Web Traffic

## Package Showcase, MS.03, Chair: Jennifer Rogers

**Jan Graffelman:** Biplot calibration with the calibrate package

**Claudia Beleites:** hyperSpec: working with spectroscopic data

**James Foadi:** cRy: statistical applications in macromolecular crystallography

**Marlene Silva Marchena:** SCperf: an inventory management package for R

**Emilio López:** Six Sigma is possible with R

**Jonathan Clayden:** Medical image processing with TractoR

**Federico Comoglio:** The Rknots package, when R meets topology

**Richard A. Bilonick:** Using merror 2.0 to Analyze Measurement Error and Determine Calibration Curves

**Martin Mächler:** Arbitrary Accurate Computation with R: The Rmpfr Package

**Matthew Shotwell:** Reading SAS7BDAT data files in R, without SAS!

# Regular[1] Posters

### Statistics/Visualization

**1.** Toby Dylan Hocking: Adding direct labels to plots

**2.** Cláudio Souza Filho: iRegression: a regression library to symbolic interval-valued variables

**3.** Carlos E. Rodríguez: Mixtures of Unimodal Distributions

### HPC/Programming/Interfaces

**4.** Frank Kramer: Accelerating Simulations in R using Automatically Generated GPGPU-Code

**5.** Tobias Verbeke: The State of StatET

**6.** Tobias Verbeke: The R Service Bus

### Teaching

**7.** Thomas Roth: Teaching Measurement Systems Analysis to Engineers Using R

**8.** Jason Waddell: The cards Package

### Finance/Actuarial Statistics

**9.** Viren Patel and Gao Yu: Using R in Insurance: Examples from Lloyd's

**10.** Yu-Min Yen: Solving Norm Constrained Portfolio Optimizations via Coordinate-Wise Descent Algorithms

### Ecology/Environmental Statistics

**11.** Alejandro Alija: Using R for air quality data analysis: A tool for designing improved large-scale air pollution prevention programs

**12.** Elizabeth Elliott: Invasions by polymorphic species

**13.** Naim Matasci: Using R to Empower a New Plant Biology

### Bioinformatics/Biochemistry

**14.** Federico Comoglio: Rknots, an R package for the topological analysis of knotted biological polymers

**15.** James Foadi: A specialised software for statistical applications in macromolecular crystallography

**16.** Pierre Gestraud: EMA - A R package for Easy Microarray Data Analysis

**17.** Neetika Nath: Classification of Enzymes via Machine Learning Approaches

**18.** Mike Smith: BeadDataPackR: A Compression Tool For Raw Illumina Microarray Data

**19.** Daniel Taliun: GWAtoolbox: An R Package for Time Efficient Quality Control of Multiple GWAS Data Files

### Medical Statistics

**20.** Anette Luther Christensen: Is the seasonal variation in hospitalisations rates of atrial fibrillation induced strokes in Denmark and New Zealand dynamic?

**21.** Kelvin Lam: Predictive Tool Development : A Time-to-event study of Diabetes Complications in Ontario

### Official/Social Statistics

**22.** Anthony Damico: Analyzing the American Community Survey with R

**23.** Sandra D. Griffith: Nonparametric estimation of a heaping mechanism for precise and heaped self-report data

### Operational Research

**24.** Osvaldo Anacleto-Junior: Forecasting multivariate time series using the DLM package: An application to road traffic networks

**25.** Sarah Bolt: Who's in the Waiting Room? Modelling Multivariate Time Series of Counts of Patients to Hospital Emergency Departments

**26.** Katrine Damgaard: Algorithm for defining hospital stays

[1]There will in addition be "Late-breaking" posters not listed here.

# Wednesday 17th August

| | | |
|---|---|---|
| **09:00** | **Invited Talk, MS.01/MS.02, Chair: Ioannis Kosmidis** | |
| | Lee E. Edlefsen | Scalable Data Analysis in R |
| **09:50** | **Kaleidoscope IIa, MS.03, Chair: Michael Rutter** | |
| **Kaleidoscope II** | Tal Galili | Using R to quantify the buildup in extent of free exploration in mice |
| | Rebecca Killick | Changepoint analysis with the changepoint package in R |
| | Karen R. Ryberg | Clustering patterns in streamflow to produce regionally or anthropogenically similar groups |
| | **Panel Discussion I, MS.01, Chair: Derek Norton** | |
| | | Starting & Building a Local R User Group |
| | **Kaleidoscope IIb, MS.02, Chair: Max Kuhn** | |
| | Loren Collingwood | RTextTools |
| | Jason Waddell | The Role of R in Lab Automation |
| | Louis Bajuk-Yorgan | Using R data functions with TIBCO Spotfire |
| **10:50** | **Coffee, Main atrium and Mathematics Workroom, A0.05** | |
| **11:15** | **Spatio-Temporal Statistics, B3.02, Chair: Julian Stander** | |
| **Focus III** | Khandoker Shuvo Bakar | Spatio-Temporal Bayesian Modelling using R |
| | Jason Lessels | Applying geospatial techniques to temporal data |
| | Nikolaus Umlauf | Structured Additive Regression Models: An R Interface to BayesX |
| | **Molecular and Cell Biology, B3.03, Chair: Andrea Foulkes** | |
| | Sandra Barragán | The R package isocir for Isotonic Inference for Circular Data. Applications to Problems Encountered in Cell Biology |
| | Juan José Fernández-Durán | CircNNTSR: An R Package for the Statistical Analysis of Circular Data based on Nonnegative Trigonometric Sums |
| | Matthew Nunes | Summary statistics selection for ABC inference in R |
| | Maarten van Iterson | Power and minimal sample size for multivariate analysis of microarrays |
| | **Mixed Effect Models, MS.03, Chair: Douglas Bates** | |
| | Ulrich Halekoh | Kenward-Roger modification of the F-statistic for some linear mixed models fitted with lmer |
| | Marco Geraci | lqmm: Estimating Quantile Regression Models for Independent and Hierarchical Data with R |
| | Benoit Liquet | lcmm: an R package for estimation of latent class mixed models and joint latent class models |
| | Kenneth Knoblauch | Mixed-effects Maximum Likelihood Difference Scaling |
| | **Programming, MS.01, Chair: Uwe Ligges** | |
| | Ray Brownrigg | Tricks and Traps for Young Players |
| | Friedrich Schuster | Software design patterns in R |
| | Patrick Burns | Random input testing with R |
| | Tobias Verbeke | An Open Source Visual R Debugger in StatET |
| | **Data Mining Applications, MS.02, Chair: Przemysław Biecek** | |
| | Stephan Stahlschmidt | Predicting the offender's age |
| | Daniel Chapsky | Leveraging Online Social Network Data and External Data Sources to Predict Personality |
| | Douglas Galagate | Using R to Model Click-Stream Data to Understand Users' Path To Conversion |
| **12:35** | **Lunch, Rootes Restaurant** | |
| **14:00** | **Invited Talk, MS.01/MS.02, Chair: John Aston** | |
| | Adrian Bowman | Modelling Three-dimensional Surfaces in R |
| **14:45** | **Invited Talk, MS.01/MS.02, Chair: John Aston** | |
| | Brandon Whitcher | Quantitative Medical Image Analysis |
| **15:30** | **Coffee, Main atrium and Mathematics Workroom, A0.05** | |

| | | |
|---|---|---|
| **16:00** | **Development of R, B3.02, Chair: John C. Nash** | |
| Focus IV | Michael A. Rutter | Packaging R for Ubuntu: Recent Changes and Future Opportunities |
| | Andrew R. Runnalls | Interpreter Internals: Unearthing Buried Treasure with CXXR |
| | GSoC 2011 Admins | R's Participation in the Google Summer of Code 2011 |
| | **Geospatial Techniques, B3.03, Chair: Roger Bivand** | |
| | Binbin Lu | Converting a spatial network to a graph in R |
| | Rainer M Krug | Spatial modelling with the R–GRASS Interface |
| | Daniel Nüst | sos4R - Accessing SensorWeb Data from R |
| | **Genomics and Bioinformatics, MS.03, Chair: Ramón Diaz-Uriarte** | |
| | Sebastian Gibb | MALDIquant: Quantitative Analysis of MALDI-TOF Proteomics Data |
| | Karl Kugler | QuACN: Analysis of Complex Biological Networks using R |
| | Marion Ouedraogo | Investigate clusters of co-expressed and co-located genes at a genomic scale using CoCoMap |
| | **Regression Modelling, MS.01, Chair: Cristiano Varin** | |
| | Bettina Grün | Beta Regression: Shaken, Stirred, Mixed, and Partitioned |
| | Rune Haubo B. Christensen | Regression Models for Ordinal Data: Introducing R-package ordinal |
| | Giuseppe Bruno | Multiple choice models: why not the same answer? A comparison among LIMDEP, R, SAS and Stata |
| | **R in the Business World, MS.02, Chair: David Smith** | |
| | Derek McCrae Norton | Odysseus vs. Ajax: How to build an R presence in a corporate SAS environment |
| | Ian Cook | A Validation/Qualification Solution for R |
| | Enrico Branca | R as a statistical tool for human factor engineering |
| **17:05** | **Hydrology and Soil Science, B3.02, Chair: Thomas Petzoldt** | |
| Focus V | Wayne Jones | GWSDAT (GroundWater Spatiotemporal Data Analysis Tool) |
| | Fabio Veronesi | IntR – Interactive GUI for R |
| | Pierre Roudier | Visualisation and modelling of soil data using the aqp package |
| | **Biostatistical Modelling, B3.03, Chair: Holger Hoefling** | |
| | Matthias Schmid | survAUC: Estimators of Prediction Accuracy for Time-to-Event Data |
| | Annamaria Guolo | Higher-order likelihood inference in meta-analysis using R |
| | Cristiano Varin | Gaussian copula regression using R |
| | **Psychometrics, MS.03, Chair: Yves Rosseel** | |
| | Florian Wickelmaier | Multinomial Processing Tree Models in R |
| | Basil Abou El-Komboz | Detecting Invariance in Psychometric Models with the psychotree Package |
| | Shiu-Lien Wu | Investigating multidimensional unfolding models using R2WinBUGS |
| | **Multivariate Data, MS.01, Chair: Peter Dalgaard** | |
| | John Fox | Tests for Multivariate Linear Models with the car Package |
| | Julie Josse | missMDA: a package to handle missing values in and with multivariate exploratory data analysis methods |
| | António Pedro Duarte Silva | MAINT.DATA: Modeling and Analysing Interval Data in R |
| | **Interfaces, MS.02, Chair: Matthew Shotwell** | |
| | Xavier de Pedro Puente | Web 2.0 for R scripts and workflows: Tiki and PluginR |
| | David Nicolaides | Browser Based Applications Supported by R in Pipeline Pilot |
| | Sheri Gilley | A new task-based GUI for R |
| **19:30** | **Conference Dinner, Panorama Suite, Rootes Building (ticket required)** | |

# Thursday 18th August

| | | |
|---|---|---|
| **09:00** | **Invited Talk, MS.01/MS.02, Chair: Julia Brettschneider** | |
| | Wolfgang Huber | Genomes and phenotypes |
| **09:50** | **Financial Models, B3.02, Chair: Giovanni Petris** | |
| | Helgi Tomasson | Computational aspects of continuous-time-arma (CARMA) models: The ctarma package |
| | Bernhard Spangl | robKalman - An R package for robust Kalman filtering revisited |
| | Peter Ruckdeschel | (Robust) Online Filtering in Regime Switching Models and Application to Investment Strategies for Asset Allocation |
| | **Ecology and Ecological Modelling, B3.03, Chair: Karline Soetaert** | |
| | Christian Kampichler | Using R for the Analysis of Bird Demography on a Europe-wide Scale |
| | Johan P. Dahlgren | Using OpenBUGS and lmer to study variation in plant demographic rates over several spatial and temporal scales |
| | John C. Nash | An effort to improve nonlinear modeling practice |
| | **Generalized Linear Models, MS.03, Chair: Kenneth Knoblauch** | |
| | Ioannis Kosmidis | brglm: Bias reduction in generalized linear models |
| | Przemysław Biecek | Large Scale, Massively Parallel Logistic Regression in R with the Netezza Analytics Package |
| | Merete K. Hansen | The binomTools package: Performing model diagnostics on binomial regression models |
| | **Reporting Data, MS.01, Chair: Martyn Plummer** | |
| | Sina Rüeger | uniPlot – A package to uniform and customize R graphics |
| | Alexander Kowarik | sparkTable: Generating Graphical Tables for Websites and Documents with R |
| | Isaac Subirana | compareGroups package, updated and improved |
| | **Process Optimization, MS.02, Chair: Tobias Verbeke** | |
| | Emilio López | Six Sigma Quality Using R: Tools and Training |
| | Thomas Roth | Process Performance and Capability Statistics for Non-Normal Distributions in R |
| | Nikolaus Rudak | R-Package JOP: Optimization of Multiple Responses |
| **10:50** | **Coffee, Main atrium and Mathematics Workroom A0.05** | |
| **11:15** | **Inference, B3.02, Chair: Peter Ruckdeschel** | |
| | Henry Deng | Density Estimation Packages in R |
| | Henrike Weinert | The benchden Package: Benchmark Densities for Nonparametric Density Estimation |
| | Taylor Arnold | Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions |
| | Patrick Rubin-Delanchy | An algorithm for the computation of the power of Monte Carlo tests with guaranteed precision |
| | **Population Genetics and Genetics Association Studies, B3.03, Chair: Martin Morgan** | |
| | Benjamin French | Simple haplotype analyses in R |
| | Jing Hua Zhao | Mixed models of large pedigrees in genetic association studies |
| | Jan Graffelman | Graphical tools for assessing Hardy-Weinberg equilibrium for bi-allelic genetic markers |
| | **Neuroscience, MS.03, Chair: Brandon Whitcher** | |
| | Heather Turner | Detecting Drug Effects in the Brain |
| | Karsten Tabelow | Statistical Parametric Maps for Functional MRI Experiments in R: The Package fmri |
| | Marijke Welvaert | neuRosim an R package for simulation of fMRI magnitude data with realistic noise |
| | **Data Management, MS.01, Chair: Barry Rowlingson** | |
| | Susan Ranney | It's a Boy! An Analysis of Tens of Millions of Birth Records Using R |
| | Joanne Demmler | Challenges of working with a large database of routinely collected health data: Combining SQL and R |
| | John Bryant | Demographic: Classes and Methods for Data about Populations |
| | Mark van der Loo | Correcting data violating linear restrictions using the deducorrect and editrules packages |
| | **Interactive Graphics in R, MS.02, Chair: Paul Murrell** | |
| | Eleni-Anthippi Chatzimichali | iWebPlots: Introducing a new R package for the creation of interactive web-based scatter plots |
| | Ian Hansel | Rocessing: Interactive Visualizations in R |
| | Richard Cotton | Easy Interactive ggplots |
| | Adrian Waddell | RnavGraph and the tk canvas widget |

Focus VI

Focus VII

| | | |
|---|---|---|
| **12:35** | **Lunch, Rootes Restaurant** | |
| **14:00** | **Kaleidoscope IIIa, MS.03, Chair: Adrian Bowman** | |
| **Kaleidoscope III** | Thomas Petzoldt | Using R for systems understanding - a dynamic approach |
| | David L. Miller | Using multidimensional scaling with Duchon splines for reliable finite area smoothing |
| | Alastair Sanderson | Studying galaxies in the nearby Universe, using R and ggplot2 |
| | **Panel Discussion II, MS.01, Chair: Louis Bajuk-Yorgan** | |
| | | Challenges Bringing R into Commercial Environments |
| | **Kaleidoscope IIIb, MS.02, Chair: Frank Harrell** | |
| | Olaf Mersmann | microbenchmark: A package to accurately benchmark R expressions |
| | Paul Murrell | Vector Image Processing |
| **15:05** | **Invited Talk, MS.01/MS.02, Chair: Tom Nichols** | |
| | Simon Urbanek | R Graphics: Supercharged - Recent Advances in Visualization and Analysis of Large Data in R |
| **15:50** | **Closing Remarks, MS.01/MS.02** | |

# Notes

Printed abstracts will not be provided at the conference. The abstract booklet is available to view at http://www.warwick.ac.uk/statsdept/user-2011/abstract_booklet.pdf.

Please note that there are three different formats for contributed talks:

**useR! Kaleidoscope:** These sessions will give a broad overview of the many different applications of R and should appeal to a wide audience.

**useR! Focus Sessions:** These sessions will focus on topics of special interest and may be more technical.

**useR! Lightning Talks:** New for useR! 2011, these sessions, with oral presentations of 5 minutes, provide a platform for participants to speak on any R-related topic and should particularly appeal to R newbies.

## Information for Speakers, Session Chairs and Poster Presenters

- Presenters of Kaleidoscope and Focus contributed talks, please note that your talk is scheduled for 17 minutes, followed by 3 minutes discussion.
- Presenters of Lightning talks, please note that your talk is scheduled for 5 minutes, with 1 minute question/transition time. A variation of the *pecha kucha* and *ignite* formats will be used, in which you must provide 15 slides to accompany your talk and each slide will be shown for 20 seconds. Slides must be provided in PDF format and sent to useR-2011@R-project.org by 17:00 GMT, Friday 12 August, 2011.
- Each useR! Invited Lecture will last 40 minutes, with 5 minutes at the end of the lecture reserved for questions.
- Technical details: All lecture rooms are equipped with an LCD projector and a computer or laptop that is connected to the internet. Unless speakers require their own laptop for software demonstrations, speakers are expected to ensure that a PDF of their presentation slides is on the computer provided, before the start of their session. For presenters of Lightning Talks, this is achieved by sending the slides in advance as described above; for presenters of other talks, please see the conference assistant in the relevant lecture room to arrange transfer of your slides. Conference assistants will be available in the lecture rooms at least 10 minutes before the program starts each day and at least 10 minutes before the program recommences after each coffee or lunch break. Please arrive well before the start of your session and introduce yourself to the Session Chair. During your talk, please look out for the countdown cards that the Chair will use to signal that your time is coming to a close.
- Session Chairs: Please check the News Board in the main atrium for any changes to your session. Please arrive well before the start of your session to meet the speakers and ensure that they have a PDF of their slides on the room computer. Countdown cards will be provided to help you keep the speakers to time. If we are aware that a speaker is missing up to 2 hours prior to the relevant session, then the other talks will be brought forwards and a notice placed on the News Board, otherwise can we please request that Chairs keep to the timetable, and in the event of a missing speaker, advise the audience to go to an alternative talk.
- For those preparing posters, the poster boards can accommodate posters of size A0 (Portrait) or A1 (Landscape). Poster presenters should arrive at least 15 minutes prior to the start of the poster session to put up their poster (presenters can access the venue from 19:00). Regular posters will have allocated boards as numbered in the schedule; late-breaking posters can be put up on any unnumbered board. Presenters should stay in the vicinity of their poster for the first hour of the session. Please take down your poster at the end of the session; if you are unable to stay until the end (23:00) remaining posters will be removed and can be retrieved from the registration desk the next day.