

Conference Directory

RSS International Conference 2024

THE UK'S ANNUAL SHOWCASE **FOR** STATISTICS & DATA SCIENCE

BRIGHTON 2-5 September 2024

ALL WELCOME

CONNECT • SHARE • LEARN

rss.org.uk/conference2024 #RSS2024Conf @RSSAnnualConf



RSS International Conference

A **WELCOMING**
and **INCLUSIVE**
event for all statisticians
and data scientists

A polite reminder to all:

- Be courteous and respectful
- Be professional
- Be kind to others

The RSS does not tolerate harassment of conference participants in any form.

rss.org.uk/conduct



ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

Connect • Share • Learn

Strengthen the profession

Corporate Partner ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

As a corporate partner, you will be helping your staff to develop their careers as well as demonstrating your – and their – commitment to the statistical profession. You will also enjoy exclusive discounts.

To find out more, please visit:
www.rss.org.uk/corporate
Tel: +44 207 614 3948
Email: membership@rss.org.uk



Welcome

On behalf of the Conference and events committee, it is my privilege and pleasure to welcome you to the Royal Statistical Society International Conference 2024 in Brighton. We hope the sun shines and that we can all enjoy the beautiful south coast venue as a backdrop to our exciting conference.

We have a wide variety of keynote, invited and contributed sessions as well as workshop and professional development sessions that provide a great choice of sessions to attend – there is surely something to interest and enthuse everyone; and that is before we add in poster presentations and rapid-fire talks.

The conference begins with our first Keynote from Hadley Wickham and is followed by a welcome reception on Brighton Pier. Who can resist an evening with a key person from the world of R and a trip to one of Brighton's most famous and iconic landmarks!

Over the following three days we have the opportunity to continue an exploration of use of R in data science and statistics as well as consider causal inference, through Karla Diaz Ordaz's keynote and several contributed sessions that relate causal inference to different contexts. Those of us with a passion for teaching statistics have a choice of a variety of sessions while anybody who wish to consider methods and theory will find many opportunities to do so. There are many other themes and keynotes of interest and importance in the world of statistics and data science too, including a focus on statistics in sport on Thursday.

Of course, we must not forget the ample opportunities to network and connect with new colleagues from around the world as well as meet with old friends and acquaintances. What better places to do so than at the quiz night on Tuesday evening, awards ceremony and poster reception on Wednesday evening, and at the conference dinner on Thursday

evening at the famous Grand Hotel on the sea front which has been welcoming guests since 1864 and is a stunning example of Victorian architecture.

Paul Gentry and his team have again created a wonderful chance for us to enjoy a stimulating conference in beautiful surroundings. Thank you Paul and team.



Professor Beverley Hale
Programme Chair
RSS 2024 Conference



General event information

Badges

Conference attendees are asked to always wear their badges and lanyards while in the conference venue.

The badges provided are intended to last for the duration of the conference, however if your badge gets damaged please go to the registration desk where a replacement can be printed.

Lost badges will be replaced at the discretion of the conference organisers. A replacement fee may be charged.

Cloakroom

There is a cloakroom available at the hotel reception for coats and small bags. Please avoid bringing suitcases to the venue (unless you are staying at the hotel!).

The Society accepts no responsibility for any items which may be lost or damaged.

Conduct policy

The Society operates an event conduct policy and requires attendees to always adhere to the policy while in the conference and its associated events. The full text of the policy can be found at rss.org.uk/conduct.

Conference app

The conference app is where you can view the most up to date information about the conference programme, plan your conference diary, message other attendees, submit feedback and questions.

A link with login details will be sent to all attendees in advance of conference – the email will go to the email address you used to register and will be sent from 'VenuIQ' (please check your junk/spam folders).

Daytime catering

Lunch as well as tea/coffee at the start of the day, the mid-morning break and the mid-afternoon break is provided as part of your conference fee. All refreshments will be served in the Durham Hall on the first floor.

There may be queues for refreshments, especially at lunchtime, so we do ask you to be patient and stagger your arrival at the serving points wherever possible.

Dietary requirements

For those attendees who have notified the organisers of special dietary requirements* please ask a member of catering staff for guidance when collecting your lunch.

For the conference dinner please collect a place card on arrival at the Grand Hotel.

The conference organisers have passed all dietary requirements notified to us to the venues for lunches and the conference dinner, however it is not always possible to provide dietary options for the other evening social events. If you are concerned about this, please contact the RSS conference team.

** There will be vegetarian options provided at lunchtime which do not require any special arrangements..*

Discounts for delegates

Visit Brighton have worked with local businesses to provide discounts for conference delegates. Simply show your conference badge at the time of purchase to take advantage of money-saving offers at restaurants, bars, shops and attractions. For available offers visit www.visitbrighton.com/special-offers/delegate-deals-o2311

Evening social events

There are no tickets for these events, however, please only attend if you have booked as we are only catering for the number of people booked. The Monday evening welcome reception is fully booked so please only go to the pier if you have booked a place.

Exhibition

Please make time to visit our exhibitors.

The exhibitors' stands are in the Durham Hall alongside the catering area. This area will be open at the following times:

TUESDAY 3 SEPTEMBER	8.20AM – 5.00PM
---------------------	-----------------

WEDNESDAY 4 SEPTEMBER	8.20AM – 7.45PM
-----------------------	-----------------

THURSDAY 5 SEPTEMBER	8.30AM – 3.30PM
----------------------	-----------------

Local information

Information about the local area and amenities, including maps and guides, can be found on the Visit Brighton website:

www.visitbrighton.com/plan-your-visit/visitor-information

Mobile devices

We ask that you switch your mobile phones and other hand-held devices to silent when you are in sessions.

Photography and filming sessions

The Society will have a photographer taking photographs in many sessions and at social events. The Society may use photographs taken in conference reports, publications and in future publicity materials. If you would prefer your photograph not to be used in this way, please contact Paul Gentry (Head of Conferences & Events).

Please note that the Society is filming a selection of invited sessions, the recordings of which will then be made available on the Society's YouTube channel.

Poster presentations

Posters will be on display in the Durham Hall throughout the conference. The poster reception on Wednesday evening will provide an opportunity to speak with the authors –the final list of posters being presented will be available on the conference app and on information boards in the venue.

Presentation desk

For speakers wishing to submit or amend their presentations during the conference and for poster presenters registering their poster the presentation desk is located in the Sussex Lounge next to registration.

Quiet/Prayer Room

A quiet room is available for those needing a quiet space or for prayer or meditation. Please ask the registration desk for more information. Please be respectful of others using the room.

Registration desk

The conference registration desk can be found in the Sussex Lounge on the first floor – on entering the hotel please make your way to the staircase at the rear of the hotel reception area.

The Registration desk will open at the following times:

MONDAY 2 SEPTEMBER	3.00PM – 6.30PM
TUESDAY 3 SEPTEMBER	8.00AM – 5.30PM
WEDNESDAY 4 SEPTEMBER	8.15AM – 6.15PM
THURSDAY 5 SEPTEMBER	8.30AM – 3.30PM

Session locations

Plenary sessions will be held in the Oxford Hall on the first floor. All other sessions will be held in one of the conference rooms on the ground, first and second floors of the hotel.

ALL sessions are open to all attendees registered to attend on the relevant day (unless otherwise stated).

Social media

You can follow the conference on X: RSSAnnualConf – please use the hashtag #RSS2024Conf

Stewards and staff

RSS stewards (in RSS t-shirts) and RSS staff members (identified by their badges) will be happy to assist you throughout the conference.

Timings for presentations

The following timings for presentations will normally apply:

- **Keynote talks:** 30-45 minutes plus time for questions and discussion
- **Invited talks:** 20-30 minutes plus time for questions and discussion
- **Contributed talks:** 20 minutes including time for questions and discussion
- **Rapid-fire talks:** 5 minutes with no questions or discussion.

WiFi access

Free Wi-Fi is available within the venue. In order to access the network you will need to:

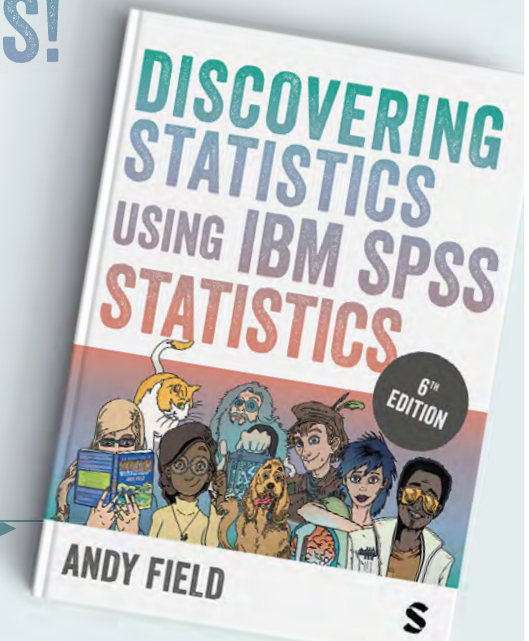
- select the Hilton Honors network
- click 'I have a promotional code'
- use VIP as the code

Please note free wi-fi may not be suitable for large downloads or streaming.

**OUT NOW: NEW 6TH EDITION OF THE
STATISTICS LIFESAVER LOVED BY
OVER 500,000 STUDENTS!**

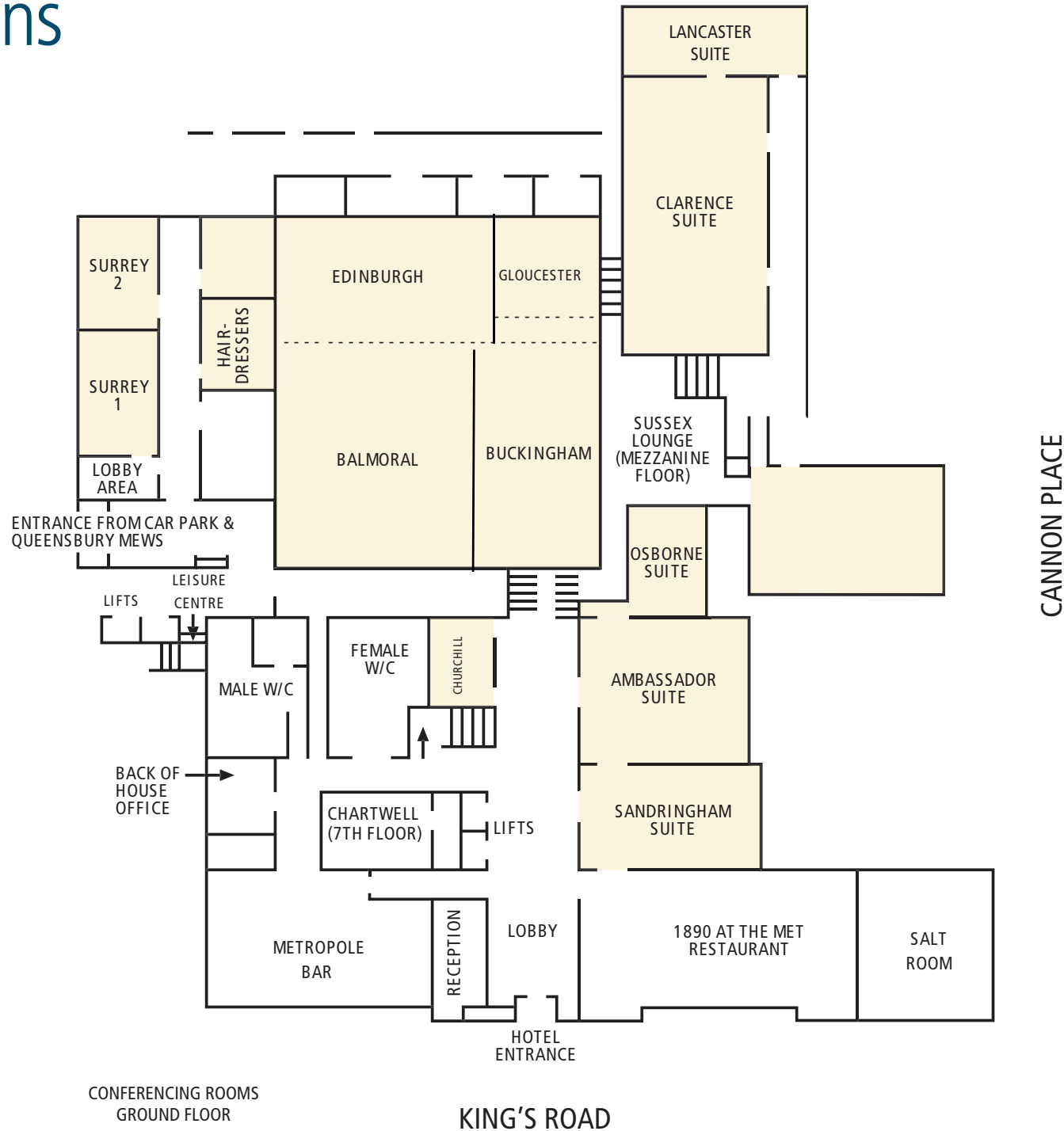


**GET YOUR COPY HERE
AND SAVE 25%
WITH CODE **RSSCON25****



Venue Plans

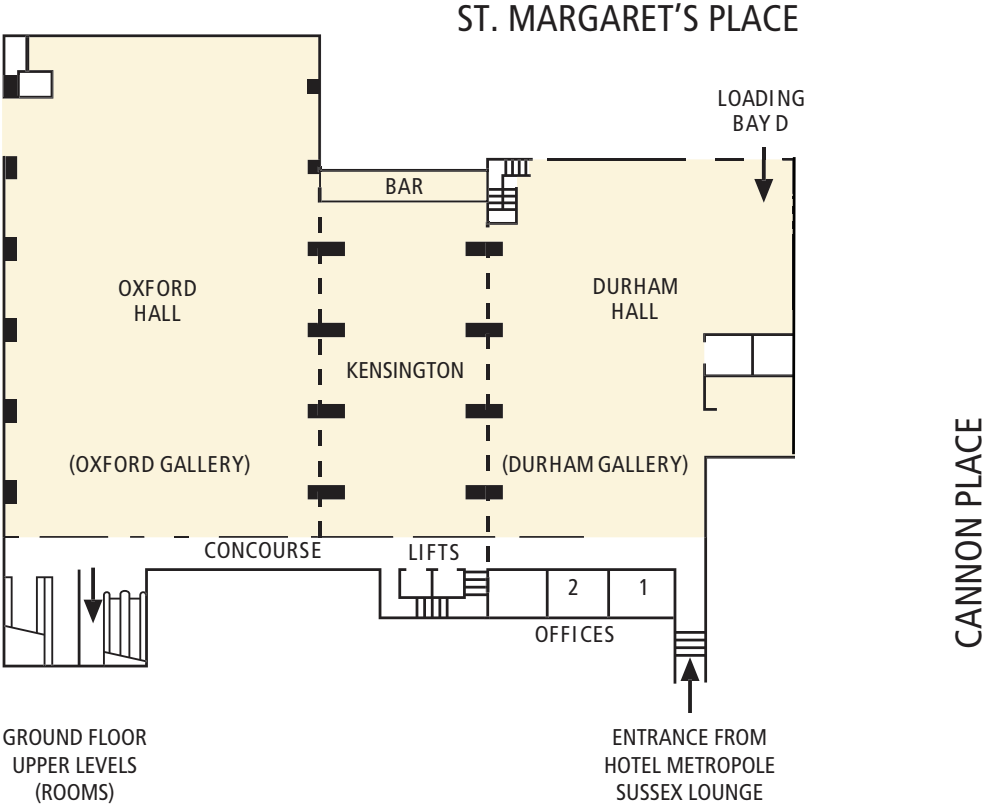
Ground floor



CONFERRING ROOMS
GROUND FLOOR

KING'S ROAD

Upper levels



Exhibitors at RSS 2023 Conference

1 Oxford University Press

At Oxford University Press, we're part of one of the greatest universities in the world, and everything we do furthers the University of Oxford's objective of excellence in research, scholarship, and education by publishing worldwide. It's why we invest back into education and research, and champion a love of learning. We are the world's largest university press with the widest global presence. Our products and services are used in over 200 countries and regions, in more than 85 languages, and we publish over 2,500 titles per year.



2 Timberlake

Timberlake is a global brand with over forty years of experience and expertise as a supplier of statistical, econometric and forecasting software packages; the delivery of quality training courses; and a consultancy service provider. We provide a total solution to our diverse range of clients across the fields of statistics, econometrics, forecasting, quantitative and qualitative research, epidemiology, finance, political and social sciences as well as data visualisation.

Our vision is to be a world-class provider of cutting-edge statistical, econometric and operational research software and training solutions.

timberlake

rss.org.uk/conference2024

#RSS2024Conf

3 & 4 Royal Statistical Society

Joining the RSS will help you expand your network in a global community of like-minded individuals and stay informed of developments in the world of statistics and data science. Speak to the friendly team at the main RSS stand to find out how membership will work for you.



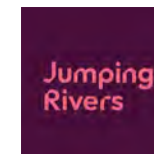
5 Chapman & Hall/ CRC Press

CRC Press is the world's leading publisher of books in Statistics through their Chapman & Hall/CRC imprint. They publish c.75 new books every year, and have an extensive backlist covering statistical theory and methods, computational statistics and data science, biostatistics and epidemiology, and much more. They are part of the Taylor & Francis Group, a world leader in statistics journals; the publishing partner for the American Statistical Association journals.



6 Jumping Rivers

Jumping Rivers is a leading UK-based R & Python consultancy known for its data science, machine learning, and statistical analysis expertise. We host and maintain Posit/RStudio products for various organisations. With a team of experienced data scientists, we assist with making the switch from legacy software to newer data science tools.



7 Cambridge University Press

We make sure academics can access high-quality, digitally interconnected materials that enhance understanding and the global impact of research. Our Cambridge University Press imprint publishes more than 380 peer-reviewed academic journals and thousands of books for research and higher education.



8 UK Data Service

Sponsors of best poster first prize

Funded by the Economic and Social Research Council, we curate the UK's largest collection of economic, population and social research data. Free at the point of service for non-commercial use and with over 52,000 registered users from 146 countries, we are firmly established as pioneers in providing training and access to data for teaching, learning and public benefit.



9 ADR UK (Administrative Data Research UK)

ADR UK (Administrative Data Research UK) is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve people's lives. By linking together data held by different parts of government, and facilitating secure access for accredited researchers to these and de-identified datasets, ADR UK is creating a sustainable body of knowledge about how our society and economy function. This approach is tailored to give decision makers the answers they need to solve important policy questions.

ADR UK is made up of four national partnerships (ADR England, ADR Northern Ireland, ADR Scotland and ADR Wales), and the Office for National Statistics. The partnership is coordinated by a team within the Economic and Social Research Council, part of UK Research and Innovation.



10 ISI – Statistical Science for a Better World

Established in 1885, the International Statistical Institute is a non-profit, non-government organisation, with individual and institutional members in over 150 countries. The ISI comprises a central core of eminent statisticians and seven Associations, each focusing on a different area of statistics. The ISI Mission is to lead, support and promote the understanding, development and good practice of statistics worldwide, by providing the core global network for statistics. Our biennial World Statistics Congresses bring together several thousands of statisticians. In addition, numerous smaller meetings, conferences and workshops take place under the auspices of the ISI and its Associations. The ISI's portfolio also includes the publication of several professional and scientific journals. ISI has a consultative status with the United Nations and a broad network around the globe. We connect to many international and national organisations, including national statistical societies, with a special focus on organisations in or supporting developing regions of the world. A majority of the world's central bureaus of statistics and many central banks are represented within the ISI network of institutional members.

<https://isi-web.org/>



11 Minitab

Sponsors of the Applications in Business, Industry & Finance stream

Visit Minitab for new solutions in statistical analysis, predictive analytics, machine learning and real-time process monitoring, and continuous improvement. We are proud for Minitab to sponsor the RSS conference stream 'Applications in Business, Industry & Finance', as organizations worldwide trust Minitab to deliver accurate, cutting-edge Solutions Analytics. From intuitive data science software to world-class training and expert services, the Minitab team is excited to help you succeed in new ways!

www.minitab.com/en-us/solutions/analytics/data-science-machine-learning/



12 Smith Institute

Sponsors of the Awards & Poster Reception

Smith Institute is an advanced mathematics, data science and AI consultancy. We help some of the biggest names in energy, transport, defence, security and consumer goods to solve their most complex challenges. Our mission is to use our skills, expertise and experience to create positive and visible impact for society, the economy and the environment.



Sponsors

MSD

Headline sponsor and sponsor of the Young Statisticians' reception

MSD - Merck Sharp & Dohme LLC – are a global health care company working to deliver innovative health solutions through medicines, vaccines, biologic therapies and animal health products. We're focused on discovering new solutions for today and the future.

<https://www.msd.com/>



Phastar

Sponsors of the Young Statisticians' Quiz Night

Phastar is a leading, award-winning, specialist Biometrics CRO that provides statistical consulting, analysis, reporting, data management and data science services to small and large biotechnology, pharmaceutical and medical device companies across the world. We leverage collaboration, technology, clinical data, and our worldwide presence to optimize decision-making for safety, quality, and efficiency. With 15 offices worldwide, our number one priority is ensuring on-time, quality work and assisting our customers in optimizing their clinical research.

<https://phastar.com>



Amelco

Sponsors of the Statistics in Sport day and the Paris 2024 prediction competition

With A Technical Reputation Established In The Banking Sector Developing Software Solutions For Pricing, Trading And Execution Platforms Amelco Is Now A Leading Provider Of Sports Betting Platforms For Global Tier 1 Operators.

<https://amelco.co.uk>



Effective Statistician

Sponsor of the prize for best rapid fire talk

Reach your potential, lead great science and serve patients.

We help you as a statistician, data scientist, programmer and other quantitative scientist to have more influence and drive innovation, grow your knowledge about the healthcare industry, and improve your personal effectiveness.

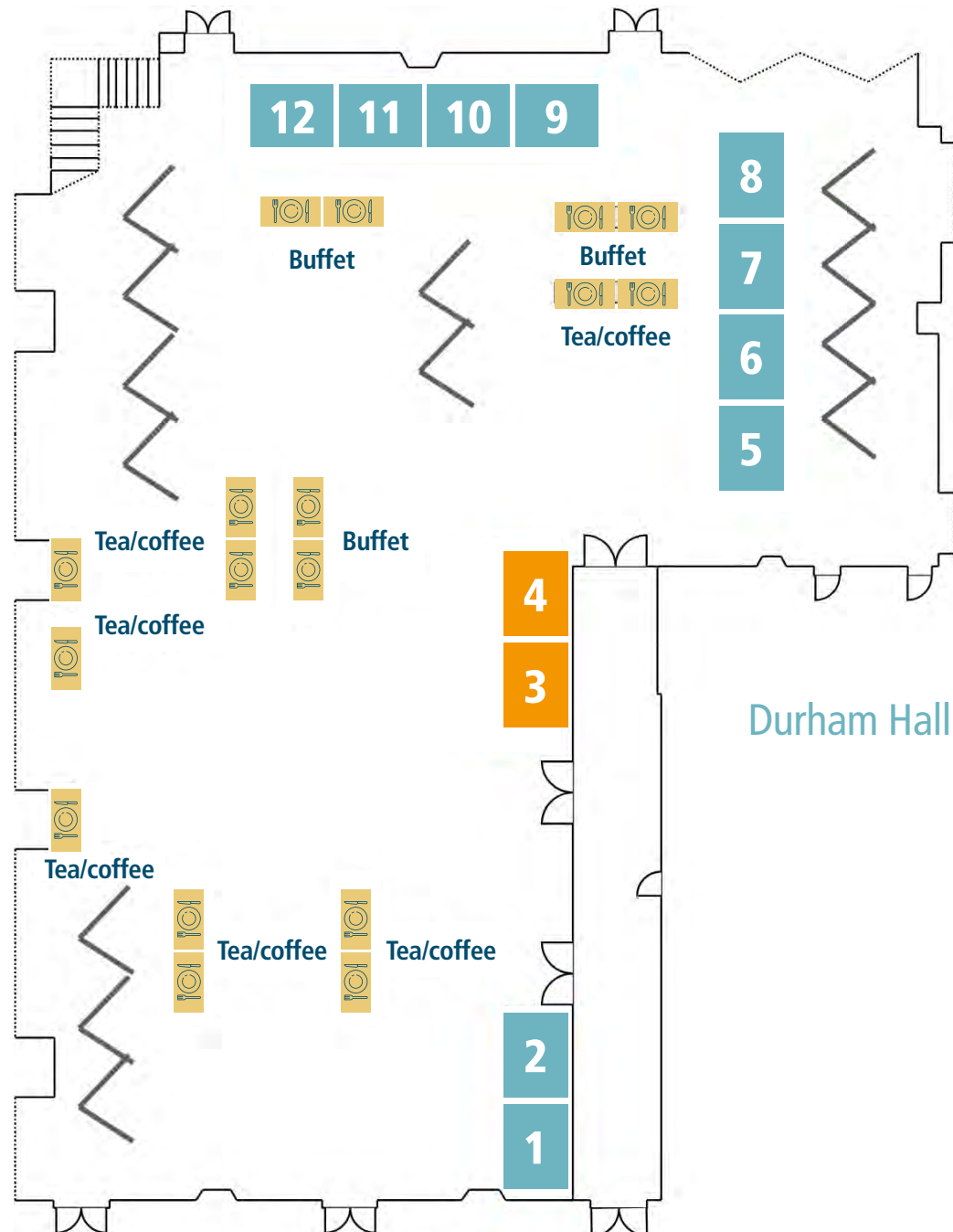
The Effective Statistician runs a weekly podcast, frequent webinars, an online conference twice a year, and other online and in-person learning opportunities. Check the website for The Effective Statistician to access many free resources to boost your career.

<https://theeffectivestatistician.com/>














Exhibitor Plan








BRIGHTON
2-5 September 2024



Full Programme

All conference sessions are organised in streams. For ease of reference each stream is represented by a colour in the programme:

	Applications in Business, Industry & Finance	
	Break	
	Communicating & Teaching Statistics	
	Data Science & AI	
	Environmental & Spatial Statistics	
	Invite only	
	Medical Statistics	
	Methods & Theory	
	Networking	
	Official Statistics & Public Policy	

	Other
	Other Applications of Statistics
	Plenary
	Professional Development
	Rapid Fire
	Social & Wellbeing Statistics
	Prize winners

The programme here is correct at the time of going to print. –

Please check the conference app and the online programme for updates to speakers and rooms.

<https://tinyurl.com/rss2024conference>

Day 1 Monday 2 September

Day 2 Tuesday 3 September

13:00 - 17:00

Professional Development

Pre-conference workshop for early career statisticians and data scientists

Kensington

Presenters:

Srshti Putcha - Young Statisticians Section

Ellie Van Vogt - Imperial College London

Dario Domingo - Durham University

Olga Mendes - Royal Statistical Society

Li-Vern Teo - Accenture

Nicola Rennie - Lancaster University

This workshop is for pre-booked attendees only.

Organised by RSS Young Statisticians Section

17:15 – 18:30

Plenary

Welcome to conference followed by Keynote 1

Oxford Hall

Hadley Wickham – Chief Scientist, Posit

19:00 - 21:00

Networking

Welcome Reception

Palm Court, Brighton Pier

08:20 – 08:50

Refreshment break

Durham Hall

08:50 – 09:50

Prize winners

RSS Prize winners: Best presentations from YSM 2024

Clarence

Hear the prize winning presentations from the Young Statisticians Meeting (YSM) 2024

Organised by Jake Emmerson for RSS Young Statisticians Section

Professional Development

How to do it all: advice from female statisticians

Kensington

This session promotes practical advice from women in statistics and data science. We discuss issues such as how to achieve a work-life balance, how to identify opportunities available when returning from a career break, how to explore flexible working arrangements and other issues that statisticians and data scientists may face in their day-to-day lives.

Organised by Joy Leahy for RSS Young Statisticians Section and RSS Celebrating Diversity Special Interest Group

Communicating & Teaching

Contributed: R Markdown, Shiny and tidymodels

Lancaster

Optimising the workflow: using R and R Markdown to generate reproducible statistical reports in the higher education sector
Iryna Schlackow - University of Oxford, UK

Sharing Results Using Shiny Dashboards: An Introduction
Erin Barker - York Health Economics Consortium, UK

Survival analysis with tidymodels
Hannah Frick - Posit, USA

Environmental & Spatial Statistics

Contributed: Environmental Modelling 1

Sandringham

Dealing with disagreeable data: How a proliferation of atmospheric datasets can hinder policy making and why data science provides an answer.

Hamish Steptoe - Met Office, UK

A copula based analysis of joint distributions of pollutants monthly maxima: an application to data from Mexico City
Eliane R. Rodrigues - Universidad Nacional Autonoma de Mexico, Mexico

Air Pollution in urban areas: a Structural Equation Model Index to identify the most critical European cities.

Elena Grimaccia - National Statistical Institute, Rome, Italy.

Day 2 Tuesday 3 September

08:50 – 09:50

Medical Statistics

Contributed: Causal Inference in medical statistics

Balmoral

Causal inference and survival analysis with time-varying exposure and interval censoring - An application to Parkinson's disease cohort
Anahita Nodehi - University of Bristol, UK

Two sample Mendelian Randomisation using an outcome from a multilevel model of disease progression
Michael A Lawton - University of Bristol, UK

Methods & Theory

Contributed: Longitudinal & Time Series

Ambassador

Bayesian Approaches to Identifiability and Estimation of Growth Mixture Models
Xingyao Xiao - University of California, Berkeley, USA.

Bias-reduction and finiteness of the GEE estimator with correlated binary responses
Anestis Touloumis - University of Brighton, UK

Flexible models for simple longitudinal data
Helen E Ogden - University of Southampton, UK

Methods & Theory

Contributed: Quantile estimation

Churchill

Distributional outcome regression via quantile functions
Rahul Ghosal - University of South Carolina, Columbia, USA.

Extreme marginal quantile treatment effect for high dimensional data
Jing Zhou - University of East Anglia, Norwich, UK

Approximate selection of all components whose high-quantiles are larger than a given value
Jun-ichiro Fukuchi - Gakushuin University, Tokyo, Japan

Official Statistics & Public Policy

Contributed: International Applications

Osborne

Quantifying uncertainty in headline administrative data-based international migration estimates
Kevin McCafferty - Office for National Statistics, London, UK

Bayesian multinomial logit models for the U.S. Voting Rights Act (VRA) determination
Joseph Kang - U.S. Census Bureau, Washington D.C., USA.

Poverty and its spatial association with the opportunity to access ART in Malawi
Zvifadzo Matsena Zingoni - University of California Los Angeles, USA

Medical Statistics

Contributed: Risk and Overall Survival in medical statistics

Buckingham

How much disease-risk is due to old age and established risk factors?
Anthony J. Webster - University of Oxford, UK

Bayesian profile regression reveals how multiple risk factors are jointly associated with the risk of antibacterial resistance.
Michail Papathomas - University of St Andrews, UK

Identifying & Assessing an Alternative Research Question for Overall Survival in the Presence of Treatment Discontinuation: Initial Results from the SOLVE Study
Kara-Louise Royle - Leeds Cancer Research UK Clinical Trials Unit, UK

Medical Statistics

Contributed: GLMMs and GAMs in medical statistics

Gloucester

Digital aspect in clinical research: reliability estimates of digital mobility endpoints in Parkinson's disease patients in Mobilise D clinical validation study
Julia Chernova - Bayer plc., Reading, UK

Economic assessment of the impact of telecare on the use of social care resources using a zero-inflated, hierarchical linear statistical model
Joe W. E. Moss - York Health Economics Consortium, UK

Measurement of routine cardiometabolic disease risk factors in primary care following Covid-19 pandemic
Frederick Ho - University of Glasgow, UK

Day 2 Tuesday 3 September

10:00 – 11:00

Plenary

Keynote 2 - From causal inference to machine learning and back: a two-way street towards better science

Oxford Hall

Karla Diaz Ordaz - University College London

Scientific inquiry often involves uncovering causal relationships and distinguishing these from correlations. Statistical causal inference aims to develop a rigorous framework for establishing causal effects, often from observational data. This involves carefully formulating the causal question of interest, specifying the estimand, and explicitly stating the assumptions under which it can be estimated using the data at hand. The estimation step often involves developing and using novel statistical and computational tools.

Machine learning methods have become established for prediction problems, but there is increasing interest in using these algorithms for causal inference. However, causal effect estimation often involves counterfactuals, and prediction tools from the machine learning literature cannot be used “out-of-the-box” for causal inference.

At the same time, there is an increasing interest in using causal reasoning when building and interpreting machine learning algorithms. Doing so can help reduce unfairness and other algorithmic biases stemming from the training data not being representative of the target population. Causality can also help with interpretability and explainability of machine learning outputs.

In this talk, I will review Orthogonal statistical learning, a framework to “de-bias” standard machine learning algorithms so they perform well for causal tasks. I will also discuss the role causal inference can play in machine learning to improve fairness and explainability of so-called “black-box” models.

This two-way street opens the way to making better use of the data and obtaining reliable answers to real-life scientific problems, while maintaining good statistical principles.

11:00 – 11:30

Refreshment break

Durham Hall

11:30 – 12:50

Communicating & Teaching

R without fear: teaching strategies to make R accessible to everyone

Lancaster

Learning R in the classroom or online? A natural experiment
Sam Dumble - Statistics for Sustainable Development, Reading, UK

More interactive, more engaging: teaching R in a lecture environment
Nicola Rennie - Lancaster University, UK

Overcoming Emotional and Attitudinal Barriers in the Teaching and Learning of R
Jenny Terry - University of Sussex, Brighton, UK

Splitting the Load – Making students comfortable with R before throwing t-tests at them
Maria D Christodoulou & Mariagrazia Zottoli, University of Oxford, UK

Organised by Oxford University Statistical Consulting:
Maria Christodoulou, Cora Mezger, Mariagrazia Zottoli

Communicating & Teaching

How and when should we teach machine learning?

Oxford Hall

There is a large interest from students in courses on Machine Learning, but the pedagogy of this material is much less established than traditional statistical courses. These talks will discuss where machine learning might be placed in comparison to traditional statistics, and whether it should be taught in the same way as traditional statistics or seen as an entirely separate topic.

How can the pedagogy of Statistics inform the teaching of Machine Learning?

Paul Fannon - Cambridge Centre for Data Driven Discovery

Receivers, creators, innovators. How have different users of Machine Learning been taught in the biological sciences at Cambridge?

Matt Castle - Head of Bioinformatics Training Institute, University of Cambridge

What mathematics is useful in understanding machine learning?

Irina Mohorianu - Turing Institute

Organised by Paul Fannon, Cambridge Centre for Data Driven Discovery



Day 2 Tuesday 3 September

11:30 – 12:50

Data Science & AI

Innovative approaches to filling data gaps on forced displacement

Osborne

Migration and forced displacement are increasingly important to understand to inform effective humanitarian response, long-term development programming, and policy. People move faster than data systems can evolve and forced displacement contexts are often hard to reach. Therefore, any edge that data science can give operations and policy decision makers by providing estimates using alternative data sources could improve the lives of many.

Machine learning and satellite imagery to detect shelters in Somalia as part of census planning
Nicci Potts - Foreign, Commonwealth and Development Office & Office for National Statistics, Newport, UK

Cost-effective high-resolution gridded population mapping and customised national sampling frame for displacement (refugees)
Sarchil H Qader - University of Southampton, UK

Nowcasting refugee population figures at UNHCR
Giulia Del Panta - UNHCR, Copenhagen

Assessing pockets of stability in multi-crisis displacement contexts
Esther Mulwa & Prithvi Hirani - IOM, London, UK

Organized & moderated by Patrick Brock (Senior Data Scientist, World Bank UNHCR Joint Data Center) for the RSS International Development section.

Environmental & Spatial Statistics

Environmental Modelling in Industry and Government

Clarence

The aim of this session is to act as a showcase for the environmental statistics research being carried out in UK Government and Industry.

The relationship between Salmonella incidence cases and England temperature, and the impact of public health interventions:
A timeseries analysis

Jennifer Israelsson - UK Health Security Agency, London, UK

Causal inference and the voluntary carbon market
Aoibheann Brady - BeZero Carbon, London, UK

Joint Probabilities of Hot and Dry Summers in the United Kingdom
Philip G Sansom - Met Office, Exeter, UK

Organised by Theresa Smith for RSS Environmental Statistics Section

Medical Statistics

Patient and public involvement (PPI) for medical statistics methodology research: why bother?

Balmoral

Patient and Public Involvement (PPI) is crucial in medical research, but guidance for incorporating PPI into research often focuses on applied studies. In this session, members of the PPI with Statistical Methods and Research Techniques (PPI-SMART) team will provide an overview of what PPI is and why it is also vital to statistical methodology research in the field of health and social care. The session will end with an opportunity for the audience to ask the panel questions, offer their own experiences, and seek advice.

Introduction to impactful PPI for statistical methodology research
Laura Gray - University of Leicester

Case study: Immortal time bias
Freya Tyrer - University of Leicester

Case study: Natural history modelling
Jonathan Broomfield - University of Leicester

Case study: Survival analysis in cancer trials
Kara-Louise Royle - University of Leeds

Case study: Estimands in clinical trials
Suzie Cro - Imperial College London

Organised by Laura Gray, Freya Tyrer, Jonathan Broomfield, Kara-Louise Royle & Suzie Cro for PPISMART

rss.org.uk/conference2024

#RSS2024Conf



Day 2 Tuesday 3 September

Methods & Theory

Network Stochastic Processes and Time Series

Buckingham

This session considers dynamic networks that occur in many fields of science, technology and medicine, as well as everyday life. Whilst simple stochastic network models are well-established in maths and computer science, the recent explosion of dynamic network data has exposed a gap in our ability to process real-life networks. This session lies at the intersection of Data Science, Theory and Methods and encapsulates a diverse range of areas ranging from applied probability, to applications and computation.

NIRVAR: Network Informed Vector Autoregression
Francesco Sanna Passino - Imperial College London, UK

Locally stationary network time series modelling
Matthew A Nunes - University of Bath, UK

Network-autoregressive models for edge processes
Gesine Reinert - University of Oxford, UK

Organised by Guy Nason, Imperial College London

Methods & Theory

Something doesn't add up? Introduction to Compositional Data Analysis

Edinburgh

In this session, we introduce compositional data analysis, outline the challenges faced when using statistical methods assuming normality, and highlight some novel and interesting developments to tackle challenges in CoDA.

This session will be followed by the session "Applications of Compositional Data Analysis".

Advances in compositional data analysis and its scientific applications
Javier Palarea-Albaladejo - University of Girona, Spain

Compositional count data modelling using the generalized-Dirichlet-multinomial method

Oliver Stoner - University of Glasgow, UK

Analysis of composition through statistical models
David Firth - University of Warwick, UK

Organised by Anastasia Frantsuzova, Oliver Stoner, Michael Papathomas, Tereza Neocleous

Official Statistics & Public Policy

Celebrating statistical excellence: Insights from award-winning teams in official and non-official statistics

Sandringham

As statisticians and data scientists both within and outside of government, continue to grapple with an unprecedented demand for data, the importance of celebrating and recognising good practice has never been more crucial.

In this session, co-chaired by Sir Robert Chote (chair, UK Statistics Authority) and Ed Humpherson (director general, Office for Statistics Regulation), you will hear from the 2024 winning teams for the Champion Award for Excellence in Official Statistics and the Trustworthiness, Quality & Value Award. And for those considering entering the awards, you'll also hear what the judging panels are looking for in terms of meeting user needs and going above and beyond.

Our 2024 Champion winners, the ONS Explore Subnational Statistics team will discuss their work in improving availability and access to local data in the UK and how they engaged with a wide variety of users to ensure demands for local data were being met.

You'll also hear from our TQV winners, the Human Fertilisation and Embryology Authority on their award-winning dashboard. Thought to be the first dashboard of its kind, bringing together data from across the UK on fertility treatments and outcomes, the team will discuss how they voluntarily applied the 'pillars' of Trustworthiness, Quality and Value that stand at the heart of the Code of Practice for Statistics, to their work.

Organised by Mags Wiley for RSS Statistical Excellence Awards Committees



Day 2 Tuesday 3 September

11:30 – 12:50

Professional Development

Learn from my mistakes: career advice from highly successful statisticians

Kensington

Join us to hear career tales from some of the most prominent statisticians today.

They didn't get to where they are today without making a few mistakes along the way. In this session, a panel of highly successful statisticians share with us their career highs, lows and mistakes made.

Let's learn from the mistakes of others. After all, we can't live long enough to make all the mistakes ourselves!

Organised by Joy Leahy for RSS Young Statisticians Section

12:50 – 13:50

Lunch

Durham Hall

rss.org.uk/conference2024

#RSS2024Conf

13:00 – 13:50

Networking

Young Statisticians Reception

Ambassador

Come and meet the RSS Young Statisticians' Section over lunch.

All welcome!

Sponsored by MSD



13:50 – 15:10

Applications in Business, Industry & Finance

FinTech, Economic Modelling, and Complexity

Churchill

This session includes talks related to the study the FinTech and application of Machine Learning and Complexity principles in the financial field and evolution of the Fintech industry. It brings together academic and industry experts to discuss their understanding of Fintech from data scientist or social scientist perspectives.

Know Your Customer: Balancing innovation and regulation for financial inclusion

Karen Elliott - University of Birmingham

Trading on online social mood: A machine learning strategy based on Twitter sentiment

Ning Wang - University of Oxford

Organised by Daniela Cialfi for RSS Economics and Finance Section

Stream sponsor Minitab The Minitab logo consists of the word "Minitab" in a bold, blue, sans-serif font, followed by a blue square icon containing a white stylized 'M' shape.

Communicating & Teaching Statistics

Using R, Shiny and associated technology in statistical education

Kensington

This session will explore cutting-edge uses of R and RStudio to enhance statistics education. We begin by very briefly overviewing the student experience of learning statistics through computers and explore some of the barriers and opportunities. We will discuss various packages based on shiny in R (learnr, gradethis and flexTeaching) to generate and mark assessments, and to generate rich but efficient student feedback.

Introduction: The student experience

Andy Field – University of Sussex, Brighton, UK

flexTeaching: serving randomly-generated statistical assignments on the web using RMarkdown

Richard D. Morey - Cardiff University, UK

Data simulation in the classroom

Lisa M DeBruine - University of Glasgow, UK

Level Up Learning: Enhancing Student Engagement Through Interactive Technology and Gamification

Danielle N Evans - University of Sussex, Brighton, UK

Shiny MarkR: Using R and Shiny to Provide Feedback on Statistics Assignments

Martina Sladekova - University of Sussex, Brighton, UK

Organised by Andy Field for RSS Teaching Statistics Section

Day 2 Tuesday 3 September

Data Science & AI

Uncertainty Quantification for Digital Twins

Edinburgh

In this session, we will explore modern methods for inference and uncertainty quantification in digital twins, using recent advances in statistics and machine learning. We will look at 3 projects involving different digital twins, looking at the recent advances that underpin their successes.

Modelling London's Air Quality Using Spatio-Temporal Gaussian Processes

Oliver Hamelijnck - University of Warwick / Alan Turing Institute

Digital twins for buildings

Fiona Bradley - University of Glasgow

Digital twins for cardiovascular disease modelling: addressing patient-specific biophysical parameter estimation and uncertainty quantification with emulation.

Dirk Husmeier - University of Glasgow

Organised by Vinny Davies on behalf of RSS Computational Statistics and Machine Learning Section

Communicating & Teaching Statistics

Developing quantitative data driven research skills in social science students outside the classroom

Lancaster

There is international demand for quantitative data skills in the workplace. If SHAPE (Social Science, Humanities and the Arts for People and the Economy) students are trained in data skills they can enter into statistical professions and data careers. This session will showcase extra-curricular quantitative data skills training initiatives for SHAPE students and will include a discussion on the impact/outcomes that these initiatives have on the students who engage with them.

Statistical Literacy: Essential skills for SHAPE students

Rhys Christopher Jones - Educational Consultant, Cardiff, UK

Developing SHAPE undergraduates' data skills: What ten years of a data fellowship programme has taught us.

Jackie Carter - University of Manchester, UK

How online quantitative data skills training can develop data skills in SHAPE students: Evidence from the UK Data Service.

Vanessa A Higgins - UK Data Service, University of Manchester, UK

Organised by Vanessa Higgins, UK Data Service

Environmental & Spatial Statistics

Statistics of natural processes and vulnerability to environmental hazards

Osborne

This session will bring together speakers from across academia and industry to talk about statistical modelling of environmental hazards, with a focus on flood risk.

Probabilistic flood modelling on a truly global scale
Rob Lamb and Fay Luxford - JBA Consulting

Statistical flood frequency estimation

Adam Griffin - UKCEH

Multi-Model data assimilation techniques for flood forecasting
Gwyneth Matthews - University of Reading

Organised by Adam Sykulski and Claudia Neves for RSS Environmental Statistics Section



Day 2 Tuesday 3 September

13:50 – 15:10

Medical Statistics

Advancements in Digital Health Technology: Statistical Opportunities, Challenges and Solutions

Buckingham

This invited session provides a unique opportunity to explore the intricacies of DHT, bringing together industry experts and other researchers to discuss the challenges and innovations that shape the future of healthcare. Attendees can expect to gain valuable insights into the latest trends, technologies, and solutions in this dynamic field.

Statistical Challenges in Decentralized Clinical Trials and Digital Health Technology Implications within Industry

Pablo Hernandez Alonso - ICON plc, & Universitat Oberta de Catalunya

Handling high frequency digital data: The role of Algorithms and Machine Learning, Reflections from a Knowledge Exchange Event
Rosemary Abbott - ICON plc & **Mia Tackney** - MRC Biostatistics Unit, Cambridge University

Digital Technologies: Sources of and Approaches to Handling Missing Data

Karen Smith - ICON plc

Organised by Pablo Hernandez Alonso, Karen Smith, Eileen Irvin and Rosemary Abbott

Methods & Theory

Inference for Functional Data

Balmoral

Functional data analysis methods for biological growth processes

Davide Pigoli - King's College London, UK

Mixture of function-on-function regression models

Julien Jacques - Université Lyon, France

On boxplots for functional data

Stanislav Nagy - Charles University, Prague, Czech Republic

Organised by Nicolas Hernandez, Queen Mary University London

Official Statistics & Public Policy

Statistics Under Pressure: how to inform decision-making at pace

Oxford Hall

The RSS's 'Statistics Under Pressure' project aims to support statisticians providing data and analyses at pace to inform time-pressured decisions. It can be challenging to judge when imperfect statistics are good enough to inform decision-making, and which trade-offs can appropriately be made. These considerations are crucial to allow data and statistics to best inform decision-making.

In this session, we will discuss a range of case studies that illustrate examples in which trade-offs had to be made to ensure that the most appropriate data and statistics could inform timely decision-making. The case studies will cover a broad range of topic areas, from Covid-19 to current government statistics.

They will highlight the considerations that had to be taken into account, and why.

The session will be chaired by RSS President-elect and chair of the Statistics Under Pressure Steering Group, John Aston. Individuals from the Statistics Under Pressure steering group and from the Government Statistical Service will take us through their experiences of the work behind the case studies.

We will hear reflections from Ed Humpherson, Head of the Office for Statistics Regulation, on how this approach intersects with the Code of Practice for Statistics. There will then be a panel discussion on the initiative and the lessons that can be learned to empower statisticians to balance trade-offs and provide the best possible statistics, data, and analyses under pressure.

Organised by Deniz Gursul for RSS Statistics Under Pressure Steering Group

rss.org.uk/conference2024

[#RSS2024Conf](https://twitter.com/RSS2024Conf)

Day 2 Tuesday 3 September

Other Applications of Statistics

Present challenges in the field of cyber-security that statistics can play a role

Gloucester

Statistical cybersecurity: a brief discussion of challenges, data structures and future directions

Francesco Sanna Passino - Imperial College London

Anomaly detection in streamed telecommunications data related to cybersecurity

Edward Austin - Lancaster University

Intelligent Cybersecure Systems, Trustworthy and Secure AI Systems

Ali Safaa Sadiq Al Shakarchi

Cyber Attacks Threat Models for IoT devices and Hybrid systems

Erisa Karafili - University of Southampton

Cybercrime: A social ecology

Alice Hutchings - Cambridge University

Organised by Ayse Ulgen, Nottingham Trent University

Professional Development

RSS Future Leaders - Inspiring Data Careers for All

Clarence

The Royal Statistical Society, with support from the Academy of Social Sciences and the Economic and Social Research Council, is recruiting a cohort of new members from underrepresented groups to serve as role models, mentors, and spokespersons for RSS and the statistics and data science professions.

Inspired by the phrase "You cannot be what you cannot see," this initiative aims to make visible the diversity of the statistics and data science profession and support and inspire a new generation of statisticians and data scientists to engage with RSS and join the profession.

For RSS Conference, our cohort of future leaders will report back on their work to develop mentoring, training and other professional support for budding statisticians and data scientists. They will discuss key achievements, lessons learned, and offer advice to conference delegates looking to rollout similar support in their own teams and workplaces.

Organised by Ricky McGowan for the Royal Statistical Society

Social & Wellbeing Statistics

Charting sustainable, healthy, and fairer futures: how Nesta draws the line from data to action

Sandingham

At Nesta, we recognize that complex societal challenges require multifaceted solutions. This session will offer a window into the work we do at the confluence of multiple disciplines.

Exploring heat pump flexibility using mixed methods research:

A collaboration between Nesta and Centre for Net Zero

Oliver Berry - Nesta, London, UK

Charting sustainable, healthy, and fairer futures: how Nesta draws the line from data to action

Rachel Wilcock - Nesta, London, UK

Charting sustainable, healthy, and fairer futures: how Nesta draws the line from data to action

Darren Hilliard - Nesta, London, UK

Organised by Federico Andreis, Nesta



Day 2 Tuesday 3 September

15:20 – 16:20

Methods & Theory

RSS Prize Winners: Kendall Lecture

Clarence

The David G Kendall prize of which this lecture is part was awarded for the recipient's contributions to, among others, empirical process theory, nonparametric inference under shape constraints, and asymptotics for multivariate models.

Mean-field asymptotics: some recent progress beyond Gaussian data
Qiyang Han - Rutgers University, USA

Professional Development

Statisticians for Society: A Journey through our Evolution and Future Opportunities

Lancaster

During this session, our speakers will delve into the journey and accomplishments of Statisticians for Society since 2017. You shall gain insights into the future operations of the scheme, exploring exciting new opportunities on the horizon and discovering how members can get involved.

Organised by Emma Lawford for Statisticians for Society

Other Applications of Statistics

Contributed: Forensics, LCA and randomisations

Osborne

Statistics in Forensic Science: Building a research relationship with impact.

Melissa A Humphries - The University of Adelaide, Australia

Intersectional social identities and mental health outcomes among minoritised ethnic youth in Aotearoa New Zealand (NZ): a latent class analysis

Shanthi Ameratunga - The University of Auckland, & Health New Zealand, and UCL Institute of Epidemiology & Health Care, London, UK

Randomization: Don't leave it to chance

Penny S Reynolds - University of Florida, Gainesville, USA

Communicating & Teaching Statistics

Contributed: Teaching Statistics

Edinburgh

Professional Ethics in Statistics Classrooms

Ayse Aysin Bilgin - Macquarie University, Australia

Students as Statistical Consultants: Ideas and Experiences

Jessica Hargreaves - University of York, UK

Putting Hypothesis Testing in its Place

Neil H Spencer - University of Hertfordshire, UK

Medical Statistics

Contributed: Infectious disease modelling

Sandringham

The complexity of pooled testing for infectious disease detection
Christopher R Bilder - University of Nebraska-Lincoln, Lincoln, USA.

Evaluating pooled testing designs in community and healthcare settings for emerging epidemics

Bethany Heath - MRC Biostatistics Unit, University of Cambridge, UK

Modelling Malaria Elimination in Nigeria: Modified SEIR-SEI Model with Intervention Effectiveness Analysis

Oluwafemi S Oyamakin - University of Ibadan, Nigeria.

Medical Statistics

Contributed: Handling bias in medical statistics

Buckingham

Systematically missing data in individual participant data meta-analysis: multiple imputation when data cannot be pooled

Robert Thiesmeier - Karolinska Institutet, Stockholm, Sweden

Bias corrections for the study weights in meta-analyses.

Stephen Walter - McMaster University, Hamilton, Canada

Scoping review of software implementations of quantitative bias analysis to informatively missing data

Rachael A Hughes - University of Bristol, UK

Day 2 Tuesday 3 September

Medical Statistics

Contributed: Clustering and phenotyping in medical statistics

Ambassador

A variational Bayes approach to patient phenotyping from large electronic health records data

Brian Buckley - University College Dublin, Ireland

Optimising Stem Cell Therapies for Corneal Damage: Statistical Insights
Patrick Parkinson - Newcastle University, UK

Uncovering multivariable effect modifiers in omics association studies through clustering analysis

Francisco J. Perez-Reche - University of Aberdeen, UK

Methods & Theory

Contributed: Modelling complex data

Balmoral

Advancements in Functional Data Analysis: Exploring Density Definitions and Gaussian Measures in $L_2[a, b]$ Space.

Adhiraj Mandal - University of Glasgow, UK

Bayesian Partially Reduced-Rank Regression

Maria F Pintado - Queen Mary University of London, UK

Spectral CLTs with long memory for large language and large multimodal models

Andrej Srakar - Institute for Economic Research, Ljubljana & University of Ljubljana, Slovenia

Official Statistics & Public Policy

Contributed: RAP, Dashboards and patient experience

Kensington

That's a RAP – working across teams to develop official statistics, dashboards and accessibility

Laura Stott, Alice Byers - Scottish Government, Edinburgh, UK

Developing the world's first fertility treatment dashboard
Elliot N Bridges - Human Fertilisation and Embryology Authority, London, UK

What are patient priorities for primary care? Understanding drivers of patient experience of primary care using the updated GP Patient Survey questionnaire

Eileen Irvin - Ipsos, London, UK

16:20 – 16:45

Refreshments

Durham Hall

16:45 – 18:45

Plenary

Keynote 3 - Discussion Meeting: Analysis of citizen science data

Oxford Hall

Paper 1: 'Efficient statistical inference methods for assessing changes in species'

Emily B Dennis, Butterfly Conservation & University of Kent, Alex Diana, University of Essex, Eleni Matechou, University of Kent, and Byron J T Morgan, University of Kent

The global decline of biodiversity, driven by habitat degradation and climate breakdown, is a significant concern. Accurate measures of change are crucial to provide reliable evidence of species' population changes. Meanwhile citizen science data have witnessed a remarkable expansion in both quantity and sources and serve as the foundation for assessing species' status. The growing data reservoir presents opportunities for novel and improved inference but often comes with computational costs: computational efficiency is paramount, especially as regular analysis updates are necessary. Building upon recent research, we present illustrations of computationally efficient methods for fitting new models, applied to three major citizen science data sets for butterflies. We extend a method for modelling abundance changes of seasonal organisms, firstly to accommodate multiple years of count data efficiently, and secondly for application to counts from a snapshot mass-participation survey. We also present a variational inference approach for fitting occupancy models efficiently to opportunistic citizen science data. The continuous growth of citizen science data offers unprecedented opportunities to enhance our understanding of how species respond to anthropogenic pressures. Efficient techniques in fitting new models are vital for accurately assessing species' status, supporting policy-making, setting measurable targets, and enabling effective conservation efforts.

Day 2 Tuesday 3 September

16:45 – 18:45

Plenary

Keynote 3 - Discussion Meeting: Analysis of citizen science data

Oxford Hall

Paper 2: 'Frequentist Prediction Sets for Species Abundance using Indirect Information'

Elizabeth Bersson and Peter D Hoff, Duke University, Durham, USA

Citizen science databases that consist of volunteer-led sampling efforts of species communities are relied on as essential sources of data in ecology. Summarising such data across counties with frequentist-valid prediction sets for each county provides an interpretable comparison across counties of varying size or composition. As citizen science data often feature unequal sampling efforts across a spatial domain, prediction sets constructed with indirect methods that share information across counties may be used to improve precision. In this article, we present a nonparametric framework to obtain precise prediction sets for a multinomial random sample based on indirect information that maintain frequentist coverage guarantees for each county. We detail a simple algorithm to obtain prediction sets for each county using indirect information where the computation time does not depend on the sample size and scales nicely with the number of species considered. The indirect information may be estimated by a proposed empirical Bayes procedure based on information from auxiliary data. Our approach makes inference for under-sampled counties more precise, while maintaining area-specific frequentist validity for each county. Our method is used to provide a useful description of avian species abundance in North Carolina, USA based on citizen science data from the eBird database.

Paper3: 'Extreme-value modelling of migratory bird arrival dates: Insights from citizen science data'

Jonathan Koh, University of Bern, Switzerland, and
Thomas Opitz, INRAE, France

Citizen science mobilises many observers and gathers huge datasets but often without strict sampling protocols, which results in observation biases due to heterogeneity in sampling effort that can lead to biased statistical inferences. We develop a spatiotemporal Bayesian hierarchical model for bias-corrected estimation of arrival dates of the first migratory bird individuals at a breeding site. Higher sampling effort could be correlated with earlier observed dates. We implement data fusion of two citizen-science datasets with fundamentally different protocols (BBS, eBird) and map posterior distributions of the latent process, which contains four spatial components with Gaussian process priors: species niche; sampling effort; position and scale parameters of annual first date of arrival. The data layer includes four response variables: counts of observed eBird locations (Poisson); presence-absence at observed eBird locations (Binomial); BBS occurrence counts (Poisson); first arrival dates (Generalized Extreme-Value). We devise a Markov Chain Monte Carlo scheme and check by simulation that the latent process components are identifiable. We apply our model to several migratory bird species in the northeastern US for 2001--2021, and find that the sampling effort significantly modulates the observed first arrival date. We exploit this relationship to effectively bias-correct predictions of the true first arrival dates.

The papers will be published in Journal of the Royal Statistical Society, Series A.

19:30 – 22:00

Networking

Young Statisticians Section Quiz Night

The Seahorse, 26 Kings Road, Brighton BN1 2LN



Advance your career in 2025 with an RSS course

Our new programme for 2025 has both virtual and face-to-face courses, all of them designed to give you as much practical experience as possible to advance your skills and career.

See below for a selection of courses on our programme, for our full course list please see our website.

- Data Visualisation - Foundation level
- Effective Analytical Leadership - Foundation level
- Interactive Dashboards & Web Apps using R & Shiny - Intermediate level
- Introduction to Bayesian Analysis using Stan - Intermediate level
- Introduction to Machine Learning in R - Intermediate level
- Bayesian Meta-analysis - Professional level
- Multilevel Modelling - Professional level
- Survival Analysis - Professional level

Explore all our courses online at rss.org.uk/public-courses or contact us on +44 (0)207 638 8998 training@rss.org.uk

Discounts are available for early bookings, group bookings and RSS members. Become an RSS member and enjoy our full range of benefits rss.org.uk/join



Day 3 Wednesday 4 September

08:20 – 08:50

Refreshment Break

Durham Hall

08:50 – 09:50

Professional Development

NIHR RSS at the RSS

Churchill

The National Institute of Health Research (NIHR) has recently launched its Research Support Service (RSS). Crucial to the delivery of this service is the professional development of its research staff, these are typically individuals who are embedded within existing academic Departments. There is capacity building as part of the NIHR RSS but it would be beneficial if various stakeholders come together to exchange ideas and explore avenues for the development of these statisticians to help them with their career progression journey and retention.

Organised by Mona Kanaan and Catherine Hewitt for RSS Medical Section

Applications in Business, Industry & Finance

Contributed: Business, Industry & Finance

Gloucester

Big Data Analytics for Effective Engine Time Series Analysis
Stela Iordache & Marios Kokmotos, Rolls-Royce, UK

Sequential Monitoring for Changes in Dynamic Semiparametric Risk Models
Xiaohan Xue - University of Bath, UK

Enhancing U.S. Treasury Yield Curve Forecasting Through Machine Learning and Forecast Combination Techniques
Jinjun Liu - HKBU, Hong Kong

Stream sponsor Minitab 

Communicating & Teaching Statistics

Contributed: Learning, Teaching and Surveying

Lancaster

Learning statistics: from anxiety to confidence
Rachel M Hilliam - The Open University, UK

Mind your Language! Teaching Statistics in the Health Sciences.
Philip M. Sedgwick - St. George's, University of London, UK

Experimental approaches in transitioning from face-to-face to push-to-web: Learnings from the Childcare and early years survey of parents
Tom Huskinson - Ipsos UK, London, UK

Data Science & AI

Contributed: Bayesian Methods

Buckingham

Bayesian Comparative Judgement: A New Approach to Enhancing Accuracy in Ranking Pairwise Comparisons.
Andy Gray - Swansea University, & Bath Spa University, UK

A Bayesian Approach to Bootstrap in Clustering techniques
Elena Ballante - University of Pavia, & IRCCS Mondino Foundation, Pavia, Italy

Harnessing Data Streams with Multivariate Hidden Markov Models for Enhanced Aircraft Engine Utilization Tracking
Christopher S Dodd - Rolls-Royce, Derby, UK

Prize Winners

RSS Prize Winners: Best presentations from RSC 2023

Clarence

Hear the prize winning presentations from the Research Students' Conference (RSC) in Probability and Statistics 2023.

Incorporating Memory into Continuous-Time Spatial Capture-Recapture Models
Clara Panchaud - University of Edinburgh, UK

Latent Position Models for Dynamic Multiview Networks
Brian Hassett - University College Dublin, & SFI Centre for Research Training in Foundations of Data Science, Dublin, Ireland.

Are you shuffling your cards well enough?
Erin G Russell - University of Bristol, UK

Organised by James Grant for RSS Young Statisticians Section

Day 3 Wednesday 4 September

Environmental & Spatial

Contributed: Climate models

Ambassador

Disentangling climate change & land use change effects on river flows: a probabilistic approach

Nicholas M Wray - University of Edinburgh, UK

Bayesian estimation of the climate sensitivity based on a simple climate model fitted to observations of ocean heat content and hemispheric temperatures

Magne Aldrin - Norwegian Computing Center, Oslo, Norway.

Illusionary Wind Energy Predictability Imperils the Energy Transition: Evidence from the Offshore Wind Energy Industry in Great Britain

Kevin F. Forbes - Energy and Environmental Data Science, Malahide, Ireland

Medical

Contributed: Topics on predictive modelling in medical statistics

Balmoral

Standard errors in case-cohort studies with small sample sizes and adjustment for prognostic covariates

Neil Wright - University of Oxford, Oxford, United Kingdom

Sample size calculations for developing a clinical prediction model: an approach to target precise individual-level predictions for decision making

Richard D Riley - University of Birmingham, UK

Understanding algorithmic fairness for clinical prediction in terms of net benefit and health equity

Jose Benitez-Aurioles - University of Manchester, UK

Methods & Theory

Contributed: Statistical theory and inference

Edinburgh

Testing for association between multi-view data and outcomes

Jing Ma - Fred Hutchinson Cancer Center, Seattle, USA.

The Impacts of Errors in Factor Levels on Second-Order Orthogonal Composite Designs

Abimibola Victoria Oladugba - University of Nigeria, Nsukka, Enugu, Nigeria

On the Hausdorff Metric and its Application in One and Two Sample Goodness of Fit Testing

Yun Jia - City, University of London, UK

Official Statistics & Public Policy

Contributed: Area based statistics

Sandringham

Modelling Income levels for Local Areas (Small Area Income Estimates) using Survey and Administrative data in England & Wales

Andrew I Zelin - ONS, Newport, UK

Using Census data to derive a new UK-wide area-based measure of deprivation

Siobhan Donnelly - HESA, Cheltenham, UK

Mapping Violence Against Women and Girls at Community Level

Rowland G Seymour - University of Birmingham, UK

Methods & Theory

Contributed: Non-parametric approaches

Osborne

Debiasing Welch's Method for Spectral Density Estimation

Lachlan C Astfalck - The University of Western Australia, Perth, Australia.

A Bayesian nonparametric hierarchical meta-regression: a bias correction approach when combining studies of varying quality and different types

Pablo E. Verde - University Hospital Düsseldorf, Germany.

Optimally-weighted Estimators of the Maximum Mean Discrepancy for Likelihood-Free Inference

Masha Naslidnyk - University College London, UK



Day 3 Wednesday 4 September

08:50 – 09:50

Professional Development

Improve Your Data Visualisations: A Practical Tutorial

Kensington

Join us for an engaging workshop designed to elevate your data visualisation skills!

Topics Covered:

- Four Key Questions: Learn the essential questions to ask before starting any visualisation project.
- Gestalt Principles: Discover how these principles can help you create clear and intuitive visualisations.
- Pre-attentive Attributes: Understand how to use these attributes to draw attention to the most important parts of your data.
- R for Data Visualisation: Gain practical skills and hands-on experience with R, one of the most powerful tools for creating data visualisations

Interactive Workbook: Participate in hands-on activities with our interactive workbook. You'll have the opportunity to apply what you learn in real-time, ensuring you leave with practical, applicable skills.

By the end of the workshop, you'll feel confident in your ability to design visualisations that communicate your data's story powerfully and effectively.

Organised and presented by Alexander Schacht, founder of The Effective Statistician, and Chantelle Cornett, an experienced R user.

rss.org.uk/conference2024

[#RSS2024Conf](https://twitter.com/RSS2024Conf)

28

10:00 – 11:00

Plenary

Keynote 4: Campion (President's Invited) Lecture -
Tim Harford

Oxford Hall

11:00 – 11:30

Refreshment break

Durham Hall

11:00 – 11:25

Networking

Data Science & Artificial Intelligence Journal launch

Clarence

Join us for the launch of the Royal Statistical Society's first new journal since 1952.

Refreshments provided

11:30 – 12:50

Applications in Business, Industry & Finance

Decision Making in Financial Markets

Churchill

An overview of the statistical methods which underpin decision-making in financial markets.

Carol Alexander – University of Sussex, UK

Random Orthogonal Matrix Simulation Targeting High-order Moments

Wei Wei - University of Bristol, UK

Generalized linear pools for combining probabilistic forecasts

Xiaochun Meng - University of Sussex, Brighton, UK

Determinants of the Value of an Option to Invest

Xi Chen - University of Sussex Business School, Brighton, UK

Dynamic Term Structure Models with Nonlinearities using Gaussian Processes

Nikolaos Karouzakis - University of Sussex Business School, Brighton, UK

Organised by Carol Alexander, University of Sussex

Communicating & Teaching Statistics

Crafting Compelling Narratives: Effectively Communicating Statistics in the Media and Beyond

Edinburgh

This session will explore the role of statistics in a post-pandemic media landscape. There will be icebreaker activities, discussions, and tips on dealing with the media and conducting public outreach, presented by the RSS statistical ambassadors. Additionally, there will be a panel discussion with John Burn-Murdoch, chief data reporter at the Financial Times, who will share insights on his approach to data journalism and what he is looking for from experts.

Chair: Professor Jennifer Visser-Rogers

Organised by Mags Wiley for RSS Statistical Ambassadors

Day 3 Wednesday 4 September

Data Science & AI

Statistical Applications of Generative AI

Balmoral

Impressive advances in generative AI have been highlighted by the wide use of generative models such as ChatGPT across a wide range of applications. Advanced methods exist for the generation of, among other things, images, text, and networks, and have been used with great potential across a variety of applications. To enhance both statistics and modern methods in AI, researchers are combining methods for generative AI with more traditional statistical methods. In this session, we will look at two case studies that demonstrate the potential for combining generative AI with statistical modelling, and discuss how we can use generative AI in combination with statistical methodology in the future.

Synthetic networks

Gesine Reinert - University of Oxford

Generating Deep Fake Left Ventricle Images for Improved Statistical Emulation

Vinny Davies & Andrew Elliott - University of Glasgow

Organised by Vinny Davies & Andrew Elliott for RSS Computational Statistics and Machine Learning Section

Data Science & AI

Data science for social good: plumbing vs wizardry

Osborne

At Nesta, our approach to impact is to use innovative methods, such as data science, to solve challenging societal issues. However, we also recognise that impact and innovation can come into tension with each other, requiring us to think critically about how we apply our methods to achieve impact. This session will shed light on how Nesta uses data science within its strategy to make progress towards its missing goals.

Data science for social good: plumbing vs wizardry

Aidan S Kelly - Nesta, London, UK

Breaking the binary of green and not-green jobs using NLP

Elizabeth Mary Gallagher & Jack Vines - Nesta, London, UK

Gaps in the market: Building a picture of the UK food environment via machine learning

Izzy Stewart & Max Hadley - Nesta, London, UK

Organised by Federico Andreis for Nesta

Environmental & Spatial

Statistical literacy and climate change – the work of the RSS Climate Change Task Force

Oxford Hall

In 2023 the RSS established a Climate Change Task Force to bring together collective RSS expertise on climate and environmental issues. The work of the Task Force has focused on promoting public understanding of statistical issues related to the climate change debate – publishing a series of responses to questions about climate change such as: how is global mean warming estimated? What does global mean warming mean locally? What is the relationship between climate change and extreme weather events?

In this session members we will explore some of the statistical issues around climate change where there is a role for the RSS in improving public understanding. We will discuss the Task Force's work to date and reflect on how statisticians can help shed light on these issues.

Organised by Andy Garrett, David Stephenson, Jonathan Everett for the RSS Climate Change Task Force



Day 3 Wednesday 4 September

11:30 – 12:50

Medical Statistics

Spatial statistical methods for global public health

Sandringham

Spatial statistics play a pivotal role in the study and control of tropical diseases, offering a powerful tool to understand their distribution, transmission dynamics, and the factors that drive their prevalence. In addition, spatial statistics is also being used to evaluate the progress of control interventions, such as vaccination campaigns, to ensure equitable immunization coverage and safeguarding vulnerable populations from preventable diseases.

This session delves into the application of cutting-edge spatial statistical methods to address pressing public health challenges that significantly impact populations in low-middle income countries, and provide pathways for more effective disease control and public health interventions.

A novel spatio-temporal modelling framework for correlated time series with heterogenous patterns with applications to malaria outbreak detection

Alejandro Rozo Posada - KU Leuven, & Hasselt University, Belgium

Geostatistical and machine learning approaches for mapping the coverage of childhood vaccination in low- and middle-income countries

Edson C Utazi - University of Southampton, UK

Modelling the impact of climate and vector control coverage interventions to changes in malaria parasite prevalence in sub-Saharan Africa: a spatio-temporal analysis

Penelope Vounatsou - Swiss and Tropical Public Health Institute, Basel, Switzerland

Organised by Emanuele Giorgi, Lancaster University

rss.org.uk/conference2024

#RSS2024Conf

30

Methods & Theory

Papers from the RSS journals' on the topic of 'causal inference'

Clarence

Selective Inference for Effect Modification Via the Lasso

Qingyuan Zhao - University of Cambridge, UK

Nonparametric identification of causal effects in clustered observational studies with differential selection

Luke J Keele - UPenn, Philadelphia, USA

Causal analysis at extreme quantiles with application to London traffic flow data

Emma J McCoy - London School of Economics and Political Science, UK

Organised by Judith Shorten & Ciara Aaron on behalf of RSS Journals

Official Statistics & Public Policy

Exploring the use of administrative data for population and migration estimates

Kensington

Would you like to learn more about our admin-based population and migration estimates (ABPEs) using our dynamic population model (DPM)? This approach draws strength from a range of data sources, such as administrative data and survey data to enhance population and migration estimates. We are continuing to develop our research into this new method which will enable us to estimate population and population change in a timely way, to better respond to user needs. Join us for an opportunity to get more 'hands-on' as we explore the details of the model.

Speakers:

Louisa Blackwell and Dominic Webber, Office for National Statistics

Organised by Catherine Bremner, Office for National Statistics

Medical Statistics

Predictive Modelling in Healthcare

Buckingham

It is a well-established premise that the utilisation of cutting-edge statistical algorithms can provide significant value through predictions and recommendations at both an individual patient level as well as at a wider population level. Such insights can empower clinicians to make efficient, effective, datadriven patient-centred decisions. This session will discuss several examples of the utilisation of advanced statistical methods in various healthcare settings, exploring the utilisation of both static and dynamic predictive modelling approaches. The challenges with such modelling will be discussed alongside routes for further exploration of the potential for embedding statistical modelling in clinical support systems with the ever-expanding collection of electronic health records.

Predicting adverse outcomes in pregnant women with pre-eclampsia

Kim Kavanagh - University of Strathclyde, Glasgow, UK

Dynamic risk prediction models and assessing performance

David McLernon - University of Aberdeen, UK

Predictive Modelling for Intensive Care Patients

Lisa McFetridge - Queen's University Belfast, UK

Organised by Lisa McFetridge, Queen's University Belfast and sponsored by the International Centre for Mathematical Sciences (ICMS)

Day 3 Wednesday 4 September

Methods & Theory

Advancements in Statistical Learning: Subgroup Selection, Transfer and Deep Learning

Ambassador

This session explores the forefront of statistical learning, covering a diverse array of advanced data science topics. From subgroup selection to the estimation of heterogeneous treatment effects, and from deep learning to transfer learning, we delve into various cutting-edge areas. Highlights include the introduction of a novel method for subgroup identification in covariate-response pairs, adaptable to classification, quantile regression, and the estimation of heterogeneous effects. Additionally, we'll explore advanced transfer learning strategies for high-dimensional linear models and robust estimation techniques within neural networks. Each presentation provides unique insights and contribute to the dynamic field of statistical learning.

Isotonic subgroup selection

Richard J. Samworth - University of Cambridge, UK

A statistical analysis of an image classification problem

Juntong Chen - University of Twente, Enschede, Netherlands

Residual Importance-Weighted Transfer Learning for High-dimensional Linear Regression

Chenlei Leng – University of Warwick, Coventry, UK

Organised by Chengchun Shi for IMS

Professional Development

Introduction to machine learning with {tidymodels} in R

Lancaster

Machine learning (ML) is an important aspect of data science that can be used to create predictions, make classifications, and uncover insights in data that can be difficult to detect. {tidymodels} is a collection of R packages that can be used for various aspects of machine learning pipelines, including sampling data, building and fitting models, and performance evaluation. {tidymodels} provides a consistent, user-friendly approach to fitting machine learning models in R. This interactive workshop will introduce some common machine learning techniques such as random forests and support vector machines, and demonstrate how to fit these models in R using {tidymodels}.

We'll also cover some of the common concepts of machine learning such as cross-validation, hyperparameter tuning and model evaluation. No previous knowledge of machine learning is required for this workshop, though familiarity with some statistical concepts such as correlation, variability, and simple linear regression may be helpful. Being reasonably comfortable with data wrangling using {dplyr} and {tidyr} would be beneficial to attendees.

Organised and presented by Nicola Rennie - Lancaster University

12:50 – 13:40

Lunch

Durham Hall

13:40 – 15:00

Applications in Business, Industry & Finance

Model Risk in Financial and Insurance Markets

Osborne

This session covers model risk in financial and insurance markets. The Federal Reserve (Federal Reserve Board of Governors, 2011) mentions explicitly that “model risk should be managed like other types of risk” and that “banks should identify the sources of model risk and assess the magnitude”.

The main focus will be on highlighting the great impact that parameter estimation risk and model specification risk can have on valuations routinely required in financial and insurance markets. It will also discuss the important role that regulators play in these markets and it will provide several important examples using real world data. A very important feature of this session is trying to reconcile statistical theory around model risk with current practice in two of the most important areas for society, finance and insurance.

Measuring Model Risk for Market Risk Measures

Emese Lazar - University of Reading, UK

Risk Management of Model Risk over the last Two Decades

Jón Daníelsson - London School of Economics, UK

Model Risk and Regulation: Reconciling Theory and Practice.

Radu S Tunaru - University of Sussex, UK

Model Risk in Pricing and Regulation of the Non-negative Equity Guarantee Clause in UK Equity Release Mortgages

Enoch N. B. Quaye - University of Bristol, UK

Organised by Radu Tunaru, University of Sussex Business School

Stream sponsor Minitab 

Day 3 Wednesday 4 September

13:40 – 15:00

Communicating & Teaching Statistics

Shut up and listen: ensuring statistical communication addresses the concerns of the audience

Oxford Hall

Assumptions about how statistics will be used and understood have to be pragmatic, but they benefit from starting with an understanding of the intended audience: who they are, what they already know, and what matters to them. Experienced communicators do a lot of this preparation behind the scenes, and the panel will discuss what they do and what is important to different members of the public.

- **Ed Humpherson** will chair the discussion with a view to how statistics can be used to answer questions important to people, as promoted by the Office for Statistics Regulation.
- **Timandra Harkness** is well known as a broadcaster, but also speaks at various industry conferences, as well as directly to people concerned about data and technology when researching her books.
- **Kevin McConway** was the first academic advisor to the More or Less programme produced by the BBC, and more recently trustee of the Science Media Centre.
- **Hannah Ritchie** is deputy editor and formerly head of research at Our World in Data, based at the the Oxford Martin School, and now an honorary fellow of the RSS.
- **David Spiegelhalter** studies the perceived trustworthiness of statistical communication, as well as putting that into practice in the media and public inquiries.

The panel will also consider topics raised by delegates around issues such as misinformation, statistical literacy, and what statisticians can do to help the public engage with more confidence and understanding in crucial public debates.

Organised by Tom King for RSS Data Ethics & Governance Section

Environmental & Spatial Statistics

Spatial Data Science Using R

Lancaster

Spatial data arise in many fields including health, ecology, environment and business. In this course, we will learn statistical methods, modelling approaches, and visualization techniques to analyse spatial data using R. We will also learn how to create interactive maps and dashboards that facilitate the communication of insights to collaborators and policymakers. We will work through several fully reproducible data science examples using real-world data such as disease risk mapping, air pollution prediction, species distribution modelling, crime mapping and real state analyses. We will cover the following topics:

- Spatial data including areal, geostatistical and point patterns.
- R packages for retrieval, manipulation and visualization of spatial data.
- Statistical methods to describe, analyze, and simulate spatial data.
- Fitting and interpreting Bayesian spatial models using the integrated nested Laplace approximation (INLA) and stochastic partial differential equation (SPDE) approaches.
- Communicating results with interactive visualizations and dashboards.

Prerequisites: It is assumed participants are familiar with R and it is recommended a working knowledge of generalized linear models. Participants should bring their laptops with R and RStudio installed.

Organiser and Presenter: Paula Moraga, KAUST, Saudi Arabia

Data Science & AI

All Statisticians are Data Scientists?

Clarence

We've looked at the actual activities that Data Scientists do in industry and what this role really means.

Are all statisticians data scientists and are all data scientists statisticians? Join us for a lively discussion on the role of statistics in data science in 2024.

Organised by Janet Bastiman, William Browne & Giles Pavey for RSS Data Science and AI Section



Day 3 Wednesday 4 September

Medical Statistics

Enhancing statistical rigour in health data research

Buckingham

Many medical studies are subject to sources of bias such as non-ignorable missing data, measurement error or misclassification, and unmeasured confounding. A quantitative bias analysis enables analysts to quantify the likely impact of such biases on their study's conclusions. Recent developments have improved the accessibility of quantitative bias analysis methods; thus enhancing the robustness and reproducibility of health-related research.

Accessible sensitivity analysis for handling missing data using multiple imputation

Suzie Cro - Imperial College London, UK

Quantitative bias analysis to unmeasured confounding or measurement error

Rolf Groenwold - LUMC, Leiden, Netherlands

Organised by Rachael Hughes, University of Bristol

Official Statistics & Public Policy

Why continuous improvement is important for population and migration statistics

Balmoral

Census 2021 for England and Wales currently provides our most detailed snapshot of the population, however there continues to be changes; people move homes, change jobs, some leave the country. As part of our continuous improvement, we have been looking at how we can move towards greater use of administrative data which should provide users with more timely estimates with a more consistent level of quality. In 2024 we will publish the recommendation to our consultation on the future of population and migration statistics in England and Wales. In this session you will hear more about our longer term plans.

Mary Gregory - Director, Population Statistics Division, Office for National Statistics

Fiona Dawe - Deputy Director, Characteristics and Inclusion Division, Office for National Statistics

Esta Clark - Head of Statistical Design, Scotland's Census Programme, National Records of Scotland

Organised by Catherine Bremner, Office for National Statistics

Other Applications of Statistics

Applications of Compositional Data Analysis

Ambassador

Compositional data analysis (CoDA) is a branch of multivariate statistics that describes and models data that has a property of a constant sum. For example, compositions of minerals in a rock as percentages, food compositions or the hours in the day spent doing an activity.

If approached by standard statistical techniques, compositional data give misleading results because of the constrained constant-sum space the data lies on.

This session, which follows from the session "Something doesn't add up? Introduction to Compositional Data Analysis", will give examples of application areas of CoDA, from environmental sciences, medicine and genetics to voting patterns.

A novel approach to mutational signature analysis of cancer sequencing data using compositional receptor models

Víctor Velasco-Pardo - University of St Andrews, UK

Two graphical displays for multi-part compositions

David Firth - University of Warwick, Coventry, UK

On Criteria for Compositional Distances for Voting Data

Zeeshan Ali Ali - University of Hull, UK

Organised by Anastasia Frantsuzova, Oliver Stoner, Michael Papathomas & Tereza Neocleous



Day 3 Wednesday 4 September

13:40 – 15:00

Professional Development

Getting your work to work

Edinburgh

Quantitative research findings are relevant only if they are computationally reproducible – rerunning the same analysis on the same dataset yields the same numerical results, figures, and inferential conclusions. Studies have found that across many research fields, only about a quarter to a third of published research is computationally reproducible. Such statistics erode confidence in science and require that we establish new processes to ensure research integrity. This session will include practical tips and advice on checking whether statistical experiments and data science projects are carried out appropriately and accurately at each stage of the data analysis. We will discuss problems that can arise during the initial pre-planning stages of an analysis to prevent cherry-picking and other questionable research practices. We will also introduce an initiative running at the School of Psychology at the University of Sussex through which researchers can check the reproducibility of their analyses before papers are submitted for publication. The workshop will be based on examples from medical and psychology research, but the content will be relevant to researchers and professionals conducting quantitative research in any field. The aim of the session is to help researchers recognise the benefits of reproducible workflows as a standard part of the research process to strengthen confidence in scientific integrity.

Organised by Reny Baykova and Teresa Lee for RSS Young Statisticians Section

Professional Development

Meta-analysis using Stata

Kensington

Meta-analysis combines results of multiple similar studies to provide an estimate of the overall effect. This overall estimate may not always be representative of a true effect. Often, studies report results that vary in magnitude and even direction of the effect, which leads to between-study heterogeneity. And sometimes the actual studies selected in a meta-analysis are not representative of the population of interest, which happens, for instance, in the presence of publication bias. Meta-analysis provides the tools to investigate and address these complications. Stata has a long history of meta-analysis methods contributed by Stata researchers. This workshop will describe the methodology for meta-analysis and cover the use of Stata's meta suite to perform meta-analysis. No prior knowledge of Stata is required. The participants of the workshop will also be provided with a temporary license to Stata 18.

Presented by **Chuck Huber** - Director, Statistical Outreach at StataCorp LLC, adjunct associate professor of biostatistics at the Texas A&M School of Public Health, and adjunct associate professor of biostatistics at the New York University School of Global Public Health.

Social & Wellbeing Statistics

Causal inference for societal impact

Sandringham

The session showcases the breadth of uses of causal inference techniques in thinking about the world and drawing inference from data. We will look at a range of methods and applications in causal inference that have some societal benefit. This includes applications in mental health, criminology and transportation as well as methods to understand fairness, leverage natural experiments and reinforcement learning.

The effect of immigration policy reform on mental health in people from minoritised ethnic groups in England: an interrupted times series analysis of longitudinal data from the UK
Annie Jeffery – University College London, UK

A/B testing in Ridesharing
Chengchun Shi – London School of Economics, UK

Ethnic Disparities in Sentencing: Warranted or Unwarranted
Jose Pina Sanchez – University of Leeds, UK

TBC
Joshua Loftus - London School of Economics, UK

Organised by Sara Geneletti, London School of Economics

rss.org.uk/conference2024

[#RSS2024Conf](https://twitter.com/RSS2024Conf)

Day 3 Wednesday 4 September

15:10 – 15:50

Rapid-fire sessions

Taking place across 10 rooms.

The following talks are expected to be presented – the final listing and allocations to sessions and rooms will be available on the online programme and the conference app.

We are grateful to **The Effective Statistician** for sponsoring the prize for best rapid-fire talk – the winner will be announced at the conference dinner.



Multistage models for risk stratification and understanding
Anthony J Webster - University of Oxford, UK

XAI for DNA: A forensic Shapley value application for high-dimensional, time-series-like data

Melissa A Humphries - The University of Adelaide, Australia

Appearances can be deceiving: predicting whether a footballer will "make it"

Jessica Hargreaves - University of York, UK

Understanding COVID-19 testing behaviour in England through a sociodemographic lens: A framework for monitoring equitable uptake of interventions

Sumali Bajaj - University of Oxford, UK

MultiOmicsIntegrator – MOI: An integrated solution for omics analyses
Bianca Alexandra Pasat - Apoptosis Research Centre, Galway, Ireland; Biomedical Research Foundation of the Academy of Athens, Greece; & Science Foundation Ireland, Galway, Ireland

Using machine learning to improve the sampling and estimation strategy for the Community Innovation Survey
Jonas Klingwort - Statistics Netherlands (CBS), Heerlen, Netherlands.

Advancing Dengue Fever Forecasting: A Model-Based Approach Incorporating Climate and Environmental Covariates in Rio de Janeiro, Brazil

Xiang Chen - King Abdullah University of Science and Technology, Jeddah, Saudi Arabia.

Adaptive smoothing of graphon probability surfaces using hybrid of node and link communities.

Arthur Verdeyme - EPFL, Lausanne, Switzerland

Understanding the replication crisis by combining statistical and diagnostic concepts to estimate outcome probabilities directly from observed proportions

Huw Llewelyn - Aberystwyth University, UK

Evaluating Randomness Assumption: A Novel Graph Theoretic Approach for Linear and Circular Data

Shriya Gehlot - Indian Institute of Management Ahmedabad, India.

Evaluating the Accuracy of the Cauchy Combination Test Distribution for Count Data under an Arbitrary Correlation Structure

Huda Alsulami - Queen Mary University of London, UK & King Abdulaziz University, Jeddah, Saudi Arabia.

Sample size estimation for external validation of a clinical prediction model when the outcomes are count or zero inflated count data.

Murthy N Mittinty - Flinders University, Adelaide, Australia

Joint Statistical Model using Longitudinal and Survival Data for identification of Biomarkers with strong prognostic capabilities

Margaret Moyinoluwa Igbalajobi - University of Bristol, UK

Optimal combinations of mean square error and directional forecast accuracy for model selection

Robert Mauritius Kunst - Institute for Advanced Studies, Vienna, Austria

Approximate Bayesian Computation for Factor Copula Models
Hanwen Xuan - University of New South Wales, Sydney, Australia.

Measuring the Information Value of Education and Likert Scales with Realist Ontology and Entropy

Wendy Olsen - University of Manchester, UK

Handling Model Uncertainty in the EM Algorithm by Model Averaging
Guangyuan Cui - City University of Hong Kong, China

Enhancing the identification and profiling of patients with heart failure in a population-based cohort through linkage of primary and secondary care data

Ellie Iob - PHARMO Institute, London, UK

Bayesian Additive Mixed Model in migration studies: Imputing migration flows time series

Andrea Lisette Aparicio Castro - University College of London, & University of Manchester, UK

A Bayesian Lasso for Tail Index Regression

Myung Won Lee - University of Edinburgh, UK

An Age–Period–Cohort Model in a Dirichlet Framework: A Coherent Causes of Death Estimation

Andrea Nigri - University of Foggia, Italy

The Diversity Dividend: Quantifying the long-term business opportunity of inclusive advertising

Annie Quinn - Choreograph, London, UK

Statistics of Armed Conflicts in Nigeria and the Rising Number of Women and Children in Refugee Shelters and IDP Camps

Polycarp E. Chigbu - University of Nigeria, Nsukka, Nigeria

What is the kurtosis of a network? An introduction to graph cumulants
Gecia Bravo-Hermesdorff & Lee M Gunderson - University College London, UK

Combining Country Indicators and Individual Variables to Predict Soil-Transmitted Helminth Infections Among Migrant Populations: A Case Study from Naples, Italy

Jana C Purkiss - Centre for Health Informatics, Computation, and Statistics (CHICAS), Lancaster University Medical School, UK

Robust Likelihood inferences for crossover designs for binary data
Tsung-Shan Tsou - Graduate Institute of Statistics, National Central University, Taiwan

Generalized Linear Mixed Model Inference for two-stage Survey Design using Laplace Approximation

Abiodun Kanmi Olakiigbe - Nigerian Institute of Medical Research Yaba, Lagos State, & Federal University of Agriculture Abeokuta, Ogun state, Nigeria

Derivative Pricing With Strategic Competition For Liquidity
Puru Gupta - University of Warwick, Coventry, UK

Day 3 Wednesday 4 September

15:10 – 15:50

Models for unbounded n -valued zero-inflated count data

Ayobami Fadilat Gboyega - Redeemer's University, Ede, Lecturer, Nigeria

Evaluating the effects of policy in econometrics and epidemiology by assuming parallel trends

Colin P Birch - Animal and Plant Health Agency, Weybridge, UK

A Test For Discriminating Between Members of The Odd Weibull-G Family of Distributions

Boikanyo Makubate - Botswana International University of Science and Technology, Palapye, Botswana

Empowerment Models for Improving Human Immunodeficiency Virus Patients' Adherence to Antiretroviral Drugs in Lagos, Nigeria

Kazeem Adewale Osulale - Nigerian Institute of Medical Research, Lagos, Nigeria

Impact of Brexit and COVID-19 on the Credit Risks of UK Companies

Kirill Romanyuk - HSE University, St Petersburg, Russian Federation

Estimation of Proportion of Students who had Cheated in Examination Using Poisson Distribution

Ayobami Fadilat Gboyega - Redeemer's University, Ede, Osun State, Nigeria.

Relative Range: a tool for outlier detection

Hana Sulieman - American University of Sharjah, UAE

Insightful Vision: Spatial Statistical Modelling in Neurodegenerative Research

Hannah J Mitchell - Queen's University Belfast, UK

A simulation study to compare using full population cohorts versus sampled comparator groups to evaluate the education outcomes of children with rare diseases

Joachim Tan - UCL Great Ormond Street Institute of Child Health, London, UK

Bayesian design for mathematical models of fruit growth based on misspecified prior information

Nushrath Najimuddin - Queensland University of Technology, Brisbane, Australia

Disentangling the effects of fine-scale mobility on leptospirosis infection using telemetry data

Pablo Ruiz Cuenca - Lancaster University, UK

Students' Difficulties with the Discipline-Oriented Teaching of R

Valentin Vancak - The Hebrew University of Jerusalem & Holon Institute of Technology, Israel

Physical Activity and Nutrition Behaviour in Female Adolescent Hockey Players within the competitive season.

Tilly J Spurr - University of Chichester, Chichester, UK

Integrating features based on correlations can maximize the predictive ability of a statistical model

Sergio Sabroso-Lasa - Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Multiversal Methods: Theory and Applications

Giulio G Cantone - University of Sussex, Brighton, UK

Sledgehammers and nuts: A modern case for classical machine learning techniques

Jonny Davis - ADSP, London, UK

Enhancing postgraduate health sciences education: teaching statistics through simulations

Nicola Orsini - Karolinska Institutet, Stockholm, Sweden

Identifying long COVID: How patterns in electronic health records can reveal cases of poorly-recorded health conditions

Karen Jeffrey - University of Edinburgh, UK

Berry-Esseen bounds for drift parameter estimation of discretely observed fractional Vasicek-type process

Fares Alazemi - Kuwait University, Kuwait

A Novel Flexible Generalized Lindley Distribution: Simulations and its Application in Medical Data Analysis

Samuel Adewale Aderoju - Kwara State University, Nigeria

Is there a spatial element to 'Don't know' and 'Prefer not to say' survey item nonresponses?

Dominick F Sutton - University of Glasgow, UK

An Experimental Design for Anytime-Valid Causal Inference on Multi-Armed Bandits

Biyonka Liang - Harvard University, Cambridge, USA.

Multi-state modelling of intermittently-observed data: a new Bayesian model and software package "msmbayes"

Christopher H Jackson - MRC Biostatistics Unit, University of Cambridge, UK

Binomial Confidence Intervals for Rare Events: Importance of Defining Margin of Error Relative to Magnitude of Proportion

Owen McGrath - University of Limerick, Ireland

Changes in features and diversity of first and last names among livebirths between 1920-2004 in a large US population voter registry

Joseph Lam - University College London, UK

Innovative use of NLP / LLM for creating official statistics from prison and probation case notes

Yuanyi Zhu & Caroline Tudor - Ministry of Justice, London, UK

Modelling temporary emigration in open-population capture-recapture data using the latent multinomial model

Katarina Skopalova - University of Glasgow, UK

Organ-Specific Biological Aging and Cognitive Changes: Implications for Neurodegenerative Disease Risk

Sujin Kang - Imperial College London, UK

Combining Radiomics and Machine Learning to assess PD-L1 expression from Magnetic Resonance Imaging acquisitions of lung cancer patients

Agnese Robustelli Test - University of Pavia, Italy.

Estimating causal effect of count outcome using zero-inflated models to account for noncompliance in randomised clinical trials

Bee-Choo Tai - National University of Singapore, Singapore

Nowcasting & forecasting hospital admissions due to infectious diseases at UKHSA

Jack Kennedy - UKHSA, London, UK

Day 3 Wednesday 4 September

Data Science via OpenSAFELY: 58 million patients' GP records analysed securely through open code

Colm D Andrews - University of Oxford, UK

Osmolality as a strong predictor of COVID-19 mortality or hospitalization and its possible links to other biomarkers

Ayşe Ulgen - Nottingham Trent University, UK, & Sirin Cetin - Amasya University, Turkey

Could US states have lowered COVID-19 infections? Insights from Functional Data Analysis.

Francesco Tripoli - École Polytechnique Fédérale de Lausanne, Switzerland

School Inspections in England: A Large Language Analysis of Trends in Ofsted Reports

Loki F Dunn - Imperial College London, UK

Generalised Fourier Autoregressive Model for Economic and Climatic Cyclical-Seasonal Time Series Datasets

Abass Ishola Taiwo - Olabisi Onabanjo University, Ago-Iwoye, Nigeria

Modelling Hepatitis C virus infection and treatment impact through serological surveillance data

Conor Egan - MRC BSU University of Cambridge, Cambridge, UK

Using online student focus groups in the development of new educational resources

Gian Carlo Diluvi - University of British Columbia, Vancouver, Canada.

Comparison of methods for multistate models of the natural history of cancer

Eddymurphy U. Akwivu - Amsterdam University Medical Centers, Netherlands.

Powerful Partial Conjunction Hypothesis Testing via Conditioning

Biyonka Liang - Harvard University, Cambridge, USA.

Data Science for Health Equity: building a world where data improves everyone's health

Brieuc Lehmann - Data Science for Health Equity, & University College London, UK

Augmented Spline Regression for Advanced Data Analysis: Generalized Additive Models and Functional Gradient Boosting with Geometrically Designed (GeD) Splines.

Emilio Luis Sáenz Guillén - Bayes Business School – City University of London, UK

A deep learning approach to modelling joint environmental extremes

Jordan Richards - University of Edinburgh, UK

Using linked data to understand mortality patterns in UK armed forces veterans

David Mais & Michael Archbold - Office for National Statistics, Newport, UK

Celestial Symmetry: Statistical Insights into the Callanish Standing Stones

Rosalind Sampson - Plus-Project Partnership, Edinburgh, UK

Modelling Composition Response Data with Application to Clot Composition Observed for Acute Ischemic Stroke (AIS) Patients

Malak Almutairi - University of Galway, Ireland

Real-Time Dengue Nowcasting in Brazil Using Google Trends and Official Epidemiological Data

Yang Xiao - King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

Statistical considerations for assessing the incremental value of additional diagnostic screening tests in paired study designs.

Nicholas B Larson - Mayo Clinic, Rochester, USA

Variance estimators for mixed-effect proportional hazards models fitted to complex samples.

Bradley A Drayton - The University of Auckland, New Zealand

How are published clinical prediction model studies assessing and reporting calibration performance at external validation?

Mia Mann - Keele University, Staffordshire, UK

Spatial variation of socioeconomic variables in England: a wavelet approach

Duncan Cook - University of Cambridge, UK

A Metabolomics Analysis of Hypertensive Disorders of Pregnancy

Zoe Winston - United States Military Academy, West Point, USA.

Adversarial Variational Autoencoders to extend and improve generative model

Loc Nguyen - Vietnam

Navigating Data Science in 2024/2025: Trends and Transformations

Amieroh Abrahams - Jumping Rivers, Newcastle, UK

Standardized Assessment of Metabolic Bariatric Surgery Outcomes: Swiss-Finnish BARIatric Metabolic Outcome Score (SF-BARI Score)

Saija A.M. Hurme & Sofia Grönroos - University of Turku & Turku University Hospital, Finland

Transforming Folk Singing Survey Data into Interactive Dashboards using R and Quarto

Grace Accad - University of Sheffield, UK



Day 3 Wednesday 4 September

15:50 – 16:15

Refreshment Break

Durham Hall

16:15 – 17:15

Plenary

Keynote 5 - Significance Lecture: Erica Thompson

Oxford Hall

17:15 – 18:00

Plenary

RSS Awards Ceremony

Oxford Hall

Each year the Society awards medals and prizes to those who have made outstanding contributions to the development of statistics and data science.

During the ceremony the Society President Andrew Garrett will welcome this year's recipients to receive their awards.

The list of recipients can be viewed here:

<https://rss.org.uk/news-publication/news-publications/2024/general-news/announcing-our-honours-recipients-for-2024/>

rss.org.uk/conference2024

#RSS2024Conf

38

18:00 – 19:45

Networking

Awards & Poster reception

Durham Hall

Celebrate the success of the award winners and take the opportunity to view this year's poster presentations and chat to the presenters.

At the time of going to print the following posters were expected to be presented – please refer to the online programme and the conference app for the latest list.

Sponsored by Smith Institute



Insightful Vision: Spatial Statistical Modelling in Neurodegenerative Research

Hannah J Mitchell - Queen's University Belfast, UK

A simulation study to compare using full population cohorts versus sampled comparator groups to evaluate the education outcomes of children with rare diseases

Joachim Tan - UCL Great Ormond Street Institute of Child Health, London, UK

Bayesian design for mathematical models of fruit growth based on misspecified prior information

Nushrath Najimuddin - Queensland University of Technology, Brisbane, Australia

Disentangling the effects of fine-scale mobility on leptospirosis infection using telemetry data

Pablo Ruiz Cuenca - Lancaster University, UK

Students' Difficulties with the Discipline-Oriented Teaching of R

Valentin Vancak - The Hebrew University of Jerusalem, & Holon Institute of Technology, Israel

Physical Activity and Nutrition Behaviour in Female Adolescent Hockey Players within the competitive season.

Tilly J Spurr - University of Chichester, UK

Integrating features based on correlations can maximize the predictive ability of a statistical model

Sergio Sabroso-Lasa - Spanish National Cancer Research Centre (CNIO), Madrid, Spain.

Multiversal Methods: Theory and Applications

Giulio G Cantone - University of Sussex, Brighton, UK

Definition, Identification, and Estimation of the Direct and Indirect Number Needed to Treat

Valentin Vancak - Holon Institute of Technology, Israel

Using address matching to link Energy Performance Certificate data to the English Longitudinal Study of Ageing

Shaun Scholes - University College London, UK

Causal attribution of smoking and BMI to disease

Anthony J Webster - University of Oxford, UK

Kumaraswamy-Normal (Kw-N) Distribution Approach in Process Monitoring: An Application to Relative Humidity of Environmental Science

Saheed Abiodun Afolabi - King Fahd University of Petroleum and Minerals, Dammam, Saudi Arabia

Making the 'world's first' fertility dashboard accessible

Abigail Ng - Human Fertilisation and Embryology Authority, London, UK

Predicting the outcome of a basketball match based on the score difference

Jessica Hargreaves - University of York, UK

Development of a reporting guideline for Scientific Machine Learning based studies.

Evangelos Danopoulos - University of Cambridge, UK

Don't Stop 'til You Get Enough - Sequential Hypothesis Testing Stop Criterion with "Enhanced Precision Is The Goal"

Eyal A Kazin - Zimmer Biomet, London, UK

A data-driven approach to monitor and predict the careers of students at Italian universities

Andrea Nigri - University of Foggia, Italy

Predictive performance of power posteriors

Yann McLatchie - University College London, UK

Advancing Dengue Fever Forecasting: A Model-Based Approach Incorporating Climate and Environmental Covariates in Rio de Janeiro, Brazil

Xiang Chen - King Abdullah University of Science and Technology, Jeddah, Saudi Arabia.

Using online student focus groups in the development of new educational resources

Gian Carlo Diluvi - University of British Columbia, Vancouver, Canada.

Comparison of methods for multistate models of the natural history of cancer

Eddymurphy U. Akwiwu - Amsterdam University Medical Centers, Netherlands

Powerful Partial Conjunction Hypothesis Testing via Conditioning

Biyonka Liang - Harvard University, Cambridge, USA.

Data Science for Health Equity: building a world where data improves everyone's health

Brieuc Lehmann - Data Science for Health Equity, & University College London, UK

Augmented Spline Regression for Advanced Data Analysis: Generalized Additive Models and Functional Gradient Boosting with Geometrically Designed (GeD) Splines.

Emilio Luis Sáenz Guillén - Bayes Business School, City University of London, UK

A deep learning approach to modelling joint environmental extremes

Jordan Richards - University of Edinburgh, UK

Using linked data to understand mortality patterns in UK armed forces veterans

David Mais & Michael Archbold - Office for National Statistics, Newport, UK

Celestial Symmetry: Statistical Insights into the Callanish Standing Stones

Rosalind Sampson - Plus-Project Partnership, Edinburgh, UK

Modelling Composition Response Data with Application to Clot Composition Observed for Acute Ischemic Stroke (AIS) Patients

Malak Almutairi - University of Galway, Ireland

Real-Time Dengue Nowcasting in Brazil Using Google Trends and Official Epidemiological Data

Yang Xiao - King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Statistical considerations for assessing the incremental value of additional diagnostic screening tests in paired study designs.

Nicholas B Larson - Mayo Clinic, Rochester, USA

Variance estimators for mixed-effect proportional hazards models fitted to complex samples.

Bradley A Drayton - The University of Auckland, New Zealand

How are published clinical prediction model studies assessing and reporting calibration performance at external validation?

Mia Mann - Keele University, Staffordshire, UK

Spatial variation of socioeconomic variables in England: a wavelet approach

Duncan Cook - University of Cambridge, UK

A Metabolomics Analysis of Hypertensive Disorders of Pregnancy

Zoe Winston - United States Military Academy, West Point, USA.

Exploring patterns of progression of eGFR in ethnically diverse population with Type II Diabetes using Group-Based Trajectory Models (GBTM).

Thamer S I Alobaid - King's College London, UK & Ministry of Health, Kuwait.

Least squares estimation for non-ergodic weighted fractional Ornstein-Uhlenbeck process of general parameters

Abdulaziz Alsenafi - Kuwait University, Kuwait

A Monte Carlo Simulation Study of the Type-I-Error and Power of ANOVA F-Test, Friedman Test and Randomization Test for Randomized Complete Block Design

Abimibola Victoria Oladugba - University of Nigeria, Nsukka, Nigeria

Addressing Criminal Recidivism: Using Regression Models for Effective Public Policy

Thomaz M Bonfim - UTFPR - Universidade Tecnológica Federal Do Paraná, Apucarana, Brazil.

Classification models applied in agriculture 4.0

João M Barros - UTFPR- Universidade Tecnológica Federal Do Paraná, Brazil.

Nonparametric Predictive Inference with Multivariate Data using copulas

Taghreed Almasoud - Durham University, UK

Structural breaks detection for dependent extremes

Xutian Yao - King's College London, UK

Bayesian Sample Size Determination for Longitudinal Intervention Studies with Linear and Log-linear Growth

Ulrich Lösenner - Utrecht University, Netherlands.

Scalable high-dimensional non-parametric density estimation, with Bayesian applications

Robert L Grant - BayesCamp Ltd, & Kingston University, UK

A Bayesian propagation algebraic algorithm to accommodate non-normal time series data: An environmental case of prediction

Ali Gargoum - UAE University, Al Ain, UAE

Rethinking Research Impact: Combining Effect Sizes and Economic Significance in Finance

Barry Quinn - Queen's Business School, Belfast, UK

MIIPW: An R Package for GEE with Missing Data Integration using a Combination of Mean Score and Inverse Probability Weighted Approaches and Multiple Imputation

Bhriku Kumar Rajbongshi - Indian Institute of Technology (Indian School of Mines), Dhanbad, India

Measuring SMS response rates from the PainPath study

Noor Daud - Keele University, UK

Tail index estimation at the interface between block maxima and POT

Chang Xu - King's College London, UK

Independent participant data meta analysis to evaluate the implementation of a parenting intervention program at scale

Francesca Little - University of Cape Town, South Africa.

Revolutionizing Health: Impact of Electronic Medical Records (EMR) on Health information management and data collection efficiency to improve health systems

Moses K Mwebaza - Jinja Regional Referral Hospital, Jinja, Uganda

Bayesian outcome-guided Clustering Identifies Clinically Predictive Subgroups of Alcoholic Hepatitis Patients

Kazeem Adedayo Adeleke - University of the West of England, Bristol, UK

Testing the equality of dependent intraclass correlation coefficients with skewed data using non-parametric bootstrap

Josep L Carrasco - University of Barcelona, Spain

Factors Affecting Infant and Child Mortality in Nigeria: A Comparison of Ridge and Lasso Regression Models

Adedotun O Omisore - Osun State Polytechnic, Iree, & Federal Polytechnic, Ilaro, Ogun State, Nigeria

Evaluating level of agreement among teams classifying acoustic survey data of mixed fish aggregations

Ingunn F Tvete - The Norwegian Computing Center, Oslo, Senior researcher, Norway

Comparing time series and machine learning for prediction of the Saudi stock market

Kamel Alanazi - University of Strathclyde, Glasgow, UK

Modelling of income profiles

Friederike Schmal - University of Münster, Germany

Impact of COVID-19 on NHS Scotland Waiting Times

Shannon Brolly - University of Glasgow, & Phastar, Glasgow, UK

Pregnancy outcomes during the COVID-19 pandemic: insights from eLIXIR, Born in South-London

Florence Tydeman - King's College London, UK

The Double Burden of Malnutrition: An Emerging Occurrence in Low and Middle-Income Countries

Phillips Edomwonyi Obasohan - Niger State Polytechnics, Zungeru, Nigeria

Bayesian Network Analysis of Domestic Water and Energy Use in Bristol

Fiorella P La Matta Romero - University of The West of England, Bristol, UK

Machine learning techniques for prediction using observational data

Morgan Crook - Phastar, Glasgow, UK

Understanding models and data in health protection

Helen E Clough - University of Liverpool, UK

The Role of Frailty in Shaping Social Contact Patterns in Belgium, 2022-2023

Neilshan Loedy - Hasselt University, Diepenbeek, Belgium.

Study designs for clinical trials: a case study within the tobacco industry

Alexandra H Hunt - Phastar, Sheffield, UK, & Philip Morris International, Neuchatel, Switzerland

Causal machine learning for generalising heterogeneous treatment effects from a randomized trial to a target population

Vanessa Rodriguez - University College London, & HDRUK, London, UK

Graphon estimation beyond binary edges

Charles Dufour - EPFL, Lausanne, Switzerland

Passive data for remote monitoring in psychosis and schizophrenia: a systematic review

Sian Bladon - University of Manchester, UK

Statistical Shape Analysis for Facial Deformity

Xiangyu Wu - University of Leeds, UK

A role for low cost air quality monitors.

Henry Simon - Open University, Milton Keynes, UK

Exploring Statistical Mishaps

Elizabeth Y Yuu - Hasso Plattner Institute, Potsdam, Germany.

A new tool to analyze the utility of predictive models: Clinical efficacy curve

Luis Mariano Esteban - Polytechnic University School of La Almunia, University of Zaragoza, Spain

Multi-state modelling of mood and disability and their impact on mortality in post-stroke patients.

Elijah Dawson - Lancaster University, UK

Correcting for Data Distribution Shift when Developing Clinical Prediction Models

Haya Elayan - University of Manchester, UK

Linking Third Sector Business Data for a better understanding of the Civil Society

Hannah R Goode - Office for National Statistics, Titchfield, UK

Bayesian Modelling for Biodiversity

Kabiru Abubakari - Queen Mary University of London, & Kew Gardens, London, UK

A novel intuitive approach to teaching linear mixed models to researchers

Vicki Hodgson - University of Cambridge, UK

Multiple Imputation of General Practice Data: Estimating Scotland level Disease Prevalence

Julia Romero & Debbie Cannon - Public Health Scotland, Edinburgh, UK

A Content Analysis of Data Science Degree Programmes in the UK

Khristin Fabian - Edinburgh Napier University, UK

Prediction of brain amyloid positivity from blood-based and digital biomarkers to accelerate Alzheimer's disease clinical research

Víctor Velasco-Pardo - University of St Andrews, UK

A comparison of generalised linear models to determine the academic success of university students from traditionally disadvantaged backgrounds

Louise Kelly - University of Strathclyde, Glasgow, UK

The Royal College of Ophthalmologists' National Ophthalmology Database Study: Risk Factors for post-cataract surgery rhegmatogenous retinal detachment

Paul HJ Donachie - The Royal College of Ophthalmologists' National Ophthalmology Audit, London, & Gloucestershire Hospitals NHS Foundation Trust, Gloucester, UK

Forecasting emergency theatre attendance: A vine copula-based approach for time series with seasonality.

Michael J Gulley - University of Plymouth, UK

Living in a fantasy: Statistical models in psychology are not equipped to deal with real-world data

Martina Sladekova - University of Sussex, Brighton, UK

A Review of Browne's Method and the Just-About-Right Approach to Sample Size Estimation.

Scholastica Chinenye Obodo - University of the West of England, Bristol, UK

Identifying metabolic biomarkers associated with risk of pancreatic cancer: a case-subcohort study

Christiana Kartsonaki - University of Oxford, UK

Application of statistical and AI techniques to build models for cancer diagnosis

Rocío Aznar-Gimeno - Instituto Tecnológico de Aragón, Zaragoza, Spain.

Mixed effects modelling of UK Biobank primary care data to assess the longitudinal impact of major depressive disorder on glycaemic control in type 2 diabetes

Alexandra C Gillett - King's College London, UK

Utilising Metaphors and Analogies to Improve Sport and Exercise Science Student's Use of Descriptive Statistics: A Pilot

Christopher Vine - University of Chichester, UK

Multivariate Geometric Anisotropic Cox Processes

James S Martin - Imperial College London, UK

Estimating the complier average causal effect (CACE) for randomised therapy trials – is a binary compliance definition appropriate?

Teresa Lee - UCL, London, UK

A comparison of semi-parametric and parametric survival methods on student dropouts from higher education.

David Young - University of Strathclyde, Glasgow, UK

Statistical considerations in the design of a soil monitoring scheme
Jacqueline M. Potts - Biomathematics and Statistics Scotland (BioSS), Aberdeen, UK

Statistics Under Pressure: informing decision-making at pace

Deniz Gursul – Royal Statistical Society, London, UK

Analysis of Factors Influencing Treatment Decision Making in Resectable Non-Small Cell Lung Cancer (US) using Logistic Regression and Automated Machine Learning

Alice Wang - AstraZeneca, Gaithersburg, USA, & Phastar, London, UK

Adversarial Variational Autoencoders to extend and improve generative model

Loc Nguyen - Vietnam

NPI-B-RP of preliminary tests in multiple-sample problems

Norah Alshahrani - University of Durham, UK

Change-point detection control charts for monitoring Weibull scale parameter

Longcheen Huwang - Institute of Statistics, Hsinchu, Taiwan

Quantitative Analysis of Tissue Volumes in Perthes Disease Using Contrast-Enhanced MRI

Hannah Kane - University College Dublin, Dublin, Ireland.

We are grateful to the UK Data Service for sponsoring the first prize of £300 for the best poster.



ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

INTRODUCING:

RSS Data Science and Artificial Intelligence



RSS: Data Science and Artificial Intelligence

Location: RSS conference, Clarence Suite

Date: Wednesday 4th September

Time: 11.00 am



RSS: Data Science and Artificial Intelligence – the first addition to the RSS family of world class statistics journals since 1952 – is now open for submissions and is seeking high quality papers from across the breadth of data science, including AI, statistics, machine learning, econometrics, bioinformatics, engineering, computational social sciences, and beyond.

Explore the journal's website now and discover why *RSS: Data Science and Artificial Intelligence*, published by Oxford University Press, is the ideal platform for showcasing your research.

Full details can be found at: academic.oup.com/rssdat/pages/why-publish
or contact journal@rss.org.uk

Day 4 Thursday 5 September

08:30 – 09:00

Refreshment break

Durham Hall

09:00 – 10:00

Professional Development

Secrets for securing research funding

Kensington

Securing funding for your research work is a crucial skill for researchers.

Writing applications is just the only tip of the iceberg – many soft skills and actions are needed (such as identifying suitable funding opportunities, assessing your likelihood of success) before you start writing to be efficient and effective in funding procurement.

This session will set out key considerations for Early Career Researchers (ECRs) in relation to securing funding for your research and helping you to develop and implement your own research funding strategy.

Research Collaboration with Industry
Jake Ansell - The University of Edinburgh, UK

Secrets for Securing Research Funding
Peter WF Smith - University of Southampton, UK

Why and how to get research funding? - sharing experience from Engineering and Physical Sciences and Medical and Applied Health Research
Yinghui Wei - University of Plymouth, UK

Organised by Zexun Chen for RSS Young Statisticians Section

Communicating & Teaching

Statistical storytelling: Significance prize winners

Clarence

A special session in which the three finalists of the Statistical Excellence Award for Early Career Writing 2024 competition will deliver presentations based on their articles.

Join us for refreshments before the start of the session.

Organised by Joy Leahy and Anna Britten for the RSS Young Statisticians Section and Significance magazine

Data Science & AI

Contributed: Time series and neural networks

Osborne

Start asking your data “Why?” - A Gentle Introduction To Causal Inference
Eyal A Kazin - Zimmer Biomet, London, UK

Advancing tourism demand forecasting: A practical high-frequency framework
Mitchell L Horrocks - Bond Business School, Bond University, Gold Coast, Australia.

Environmental & Spatial Statistics

Contributed: Environmental modelling 2

Edinburgh

Modeling Data-Level Spatial Dependence in Multivariate Generalized Extreme Value Distributions with Copulas Based on Gaussian Markov Random Fields
Brynjólfur Gauti Guðrúnar Jónsson - University of Iceland, Reykjavik, Iceland

Exploring hidden dynamics: statistical inference in periodically inhomogeneous hidden Markov models for animal behaviour
Jan-Ole Koslik - Bielefeld University, Bielefeld, Germany.

Simulation-based evaluation of treatment strategies to fight parasites in salmon farming with respect to resistance development
Ragnar B Huseby - Norwegian Computing Center, Oslo, Norway.

Medical Statistics

Contributed: Time-to-event models in medical statistics

Balmoral

A simulation study to compare machine learning algorithms and Cox proportional hazards model for survival analysis of stroke patients.
Josline Adhiambo Otieno - Umeå University, Sweden.

An Evaluation of Predictive Value and Associated Uncertainty of Kinetic Information from Dynamic PET Imaging in Breast Cancer Patients
Fengyun Gu - North China Electric Power University, Beijing, China, & **Finbarr O'Sullivan** - University College Cork, Ireland

Application of statistical models on analyzing adverse event data in cancer clinical trial
Dung-Tsa Chen - Moffitt Cancer Center, Tampa, USA.

rss.org.uk/conference2024

[#RSS2024Conf](https://twitter.com/RSS2024Conf)

Day 4 Thursday 5 September

Medical Statistics

Contributed: Disease specific predictions

Sandringham

Prediction of ADHD in unselected schoolchildren using a large-scale machine learning model based on imbalance problems

Chao Xu - University of Glasgow, UK

Development and validation of a dynamic prediction model for dementia in Parkinson's disease: an individual participant data meta-analysis

Yan Li - University of Aberdeen, UK

Situational awareness and forecasting during the COVID-19 pandemic in Norway

Solveig Engebretsen - Norwegian Computing Center, Oslo, Norway

Methods & Theory

Contributed: Estimation techniques

Buckingham

Bias-reducing adjustments for generalised additive models

Oliver Kemp - University of Warwick, Coventry, UK

Theory of Posterior Concentration for Generalized Bayesian Additive Regression Trees

Enakshi Saha - Harvard University, Boston, USA

Robust and Efficient Estimation in Ordinal Response Models using the Density Power Divergence

Arijit Pyne - Indian Statistical Institute, Kolkata, India.

Official Statistics & Public Policy

Contributed: Methodological approaches

Ambassador

Predicting COVID-19 hospitalizations: fostering decision-making in future pandemics

Joep Burger - Statistics Netherlands, Heerlen, Netherlands.

A methodological framework to monitor changes in proportions of Causes of Death time series.

Andrea Nigri - University of Foggia, Italy

Understanding Transport Mode Switching Behaviour: A Bayesian Approach Using British Social Attitudes Survey Data

Dilum Dissanayake - University of Birmingham, UK

Other Applications of Statistics

Contributed: Statistical applications to sport

Lancaster

The Split: Analysing Contest Design in the Scottish Premier League

Jessica K Hargreaves - University of York, UK

Designing an Inclusive Scoring System for Powerlifting

Jean C. Peyen - Queen's University Belfast, UK, &

Ruheyan Nuermaimaiti - University of Leeds, UK

Modelling the Number of Yellow Cards in English Premier League Football Matches

India E Richmond - University of York, UK

Part of the Statistics in Sport day, kindly sponsored by Amelco



Social Statistics & Wellbeing Statistics

Contributed: Health and public safety

Churchill

Estimating Prevalence of Post-war Health Disorders using Multiple Systems Data

Prajamitra Bhuyan - Indian Institute of Management Calcutta, Kolkata, India.

Lighting up evidence on lighting and public safety.

Paul R Marchant - Leeds, UK

Bayesian spatial disaggregation modeling for the detection of disease clusters

Hanan Alahmadi - King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.



Day 4 Thursday 5 September

10:10 – 11:10

Plenary

Keynote 6: Assessing competitive balance in the English Premier League and other adventures in sports statistics

Oxford Hall

Nial Friel – University College Dublin

Data arising from sporting competitions provide an abundance of interesting statistical challenges, from both a modelling and an inferential perspective. This talk will explore some of these challenges, including the question of how to cluster and rank teams (or individuals) based on pairwise comparison data. That is, when the data consists of the outcome when pairs of teams (or individuals) play against one another. A particular example will illustrate a modelling framework that can be used to assess how competitive the English Premier League has been over the past 40 years.

Part of the Statistics in Sport day, kindly sponsored by Amelco



11:10 – 11:40

Refreshment break

Durham Hall

11:40 – 13:00

Applications in Business, Industry & Finance

Statistical approaches to quality improvement in the service industry

Gloucester

Service industries are ripe with operational data. Much of this data can be utilised to provide real insights into asset management and how to improve business innovation and success. The session will consist of three talks drawing on experience of using statistics to advantage in terms of business function and lasting impact in different parts of the service sector. There will then be a focused discussion to draw up a Pareto chart of engagement with different parts of the service sector and identify areas with least activity.

In this session we will explore case studies and make recommendations for helping all sectors grasp the opportunity for quality improvement.

Dental data analytics

Shirley Coleman - Newcastle University, UK

Six Sigma quality improvement in regulation

Jonathan Smyth-Renshaw

Organised by Shirley Coleman for RSS Quality Improvement Section

Stream sponsor Minitab 

Communicating & Teaching Statistics

Hidden Figures in Statistics

Oxford Hall

Statistical techniques and theory are often taught either without any historical context, or with a narrow context using a handful of the same stories. Portraying statistics as a tool developed by real people to solve real problems could support student learning and improve their understanding of the applications of statistics to present day problems. However, when historical and social contexts are overlooked, and stories lack diversity, students from historically underserved communities may be discouraged from pursuing the discipline further.

As well as telling our students about Fisher and Bayes, we should aim to diversify and tell stories of the Black, Chinese, Women, LGBT+, or Disabled statisticians who also changed the world. These stories are inevitably harder to find, so we've collected a few of our favourites into this storytelling session, combined with some guidance on how you could fit these into your teaching. Please join us for an enjoyable hour of stories while raising and celebrating diversity in statistics.

Organised by Lucy Teece for RSS Celebrating Diversity Special Interest Group

Day 4 Thursday 5 September

Data Science & AI

How should the statistical profession adopt AI?

Balmoral

Traditionally, statisticians and data scientists receive formal training in statistical distribution, hypothesis testing, modelling and other statistical principles and analyse data using tried and trusted statistical methods. With the advancements of big data and artificial intelligence (AI) techniques, there is a growing emphasis on data-driven decision-making and applications in data analysis, instead of the formalised statistical-driven approach.

Are we ready to include AI-enhanced techniques within the statisticians toolbox? This session, jointly organised by the Data Ethics and Governance Section (DEG) and the Young Statisticians Section (YSS), will include discussions on the implications, potential dangers, new applications, technical and ethical considerations, and other aspects of AI to the future of the statistical profession.

Organised by Peter Elias and Teresa Lee for RSS Data Ethics & Governance and Young Statisticians Section

Environmental & Spatial Statistics

New advances in point processes for environmental problems

Churchill

A spatial point process is a stochastic model for a spatial point pattern, i.e., a finite collection of points specifying the locations of some events observed within a given spatial region. Over the last few years, statistical spatial point pattern analysis has observed a surge in applications in many areas of science.

Similarly, many innovations have been made in terms of statistics models, methods and software implementations. This session will showcase new contributions to the statistical analysis of point processes with a view toward environmental applications.

Statistical models and deep learning methods for the analysis, prediction and monitoring of space-time point pattern data
Jorge Mateu - University Jaume I, Castellon, Spain

A spatial statistical overview on landslide hazard across Japan via a marked point process equipped with the Barrier model.
Erin Bryce - University of Glasgow, UK

Point processes in practice: the role of spatial resolution and integration scheme.

Sara Martino - Norwegian University of Science and Technology, Trondheim, Norway

Organised by Daniela Castro-Camilo for RSS Environmental Statistics Section

Medical Statistics

Statistics and data science applications in health analytics consultancy

Edinburgh

Health analytics/informatics is one of the fastest growing sectors currently, and statistics and data science have made crucial contributions to this sector that are rarely acknowledged among statistical and data science communities. Newer techniques and communication methods are being used to support commercial functions in healthcare and life sciences, such as real world evidence (RWE) and health economics and outcomes research (HEOR), while the consultancy setup exposes statisticians and data scientists to a wide range of analyses and commercial needs. This is in contrast to the traditional medical statistics focus on clinical trials and drug safety studies.

This panel session with statistical experts in consultancies will discuss key statistical and data science topics in relation to these commercial functions, best practice for communicating analyses and practical considerations for statistical/data science consulting. These topics include high dimensional methods for population characterisation and risk prediction, and longitudinal methods for modelling and forecasting disease progression.

We plan to raise awareness of and garner interest in this sector among statisticians and data scientists, and to demonstrate how statistics and data science can be used to improve population health.

Mei Sum Chan - LCP Health Analytics

Maria Christodoulou - Oxford University Statistical Consulting
Soroosh Afyouni - Optima Partners

Organised by Mei Sum Chan, LCP Health Analytics



Day 4 Thursday 5 September

11:40 – 13:00

Methods & Theory

Survival models with multiple linear predictors

Osborne

This session introduces MPR survival models, their properties, and how they are used to analyse survival data. These wholly parametric models use a separate linear predictor for each of the basic parameters in a standard survival model. For example, the Weibull has two basic parameters – scale and shape - and so the corresponding MPR Weibull will have two linear predictors in the covariates.

Typically, these models have non proportional hazards, are more flexible than standard survival models, as they can handle increasing, waning and crossing hazards data.

After an introduction, the following talks will illustrate the use of MPR models in a variety of situations:

Inference in survival models

Gilbert MacKenzie - formerly, The University of Limerick, Ireland

Automating variable selection in distributional regression

Kevin Burke - University of Limerick, Ireland

Multiple-predictor regression survival models with a cure rate mixture

Defen Peng - University of British Columbia, Vancouver, Canada.

Piecewise Exponential Distributional Regression Model for Survival Analysis

Jack Moore - University of Limerick, Ireland.

Organised by Gilbert Mackenzie and Kevin Burke

rss.org.uk/conference2024

[#RSS2024Conf](https://twitter.com/RSS2024Conf)

46

Social & Wellbeing Statistics

Survey Data Collection Methods Collaboration (Survey Futures): Research Programme, Objectives and Early Findings

Sandringham

The UK Economic and Social Research Council (ESRC) has commissioned a £3.3M initiative that aims to deliver a step-change in survey research to ensure that it will remain possible in the UK to carry out high quality social surveys of the kinds required by the public and academic sectors. The initiative, known as Survey Futures, involves an extensive research programme, a training and capacity-building stream, workshops, conferences and other events. The need for Survey Futures arises both from challenges currently facing the survey community and from new opportunities. Paradoxically, both some of the challenges and some of the opportunities were caused or accelerated by the covid-19 pandemic. The pandemic decimated interviewing field forces but also forced the rapid adoption of new approaches to survey data collection. In this session, we will outline the objectives of Survey Futures, will present the methodological research programme, and will discuss some early findings.

Survey Futures: Addressing current challenges to assure the future of survey data collection in the UK

Peter Lynn - University of Essex, UK, & **Olga Maslovskaya** - University of Southampton, UK

Live Video Interviewing: Empirical Evidence of Opportunities and Challenges Across Seven Major UK Social Surveys

Gabriele Durrant - University of Southampton, UK, & **Tom Huskinson** - IPSOS, London, UK

Targeted survey procedures in UK self-completion surveys

Viktor Sladka - University of Essex, UK

The role of face-to-face interviewers in a post-pandemic UK

Sierra Mesplie-Escott - NatCen, London, UK

Session chair and discussant: Peter WF Smith

Organised by Olga Maslovskaya, University of Southampton

Professional Development

Reproducible data reports with Quarto

Lancaster

Quarto is an open-source scientific and technical publishing system. Quarto allows you to weave together narrative text and code to produce elegantly formatted output and tell a story. When the underlying data inevitably changes, a click of a button will update the document like magic – automatically updating plots, tables and text. This style of automated reporting makes data analysis more efficient and reproducible. Quarto supports dozens of static and dynamic output formats including HTML, PDF, MS Word, slides, dashboards, scientific articles and more.

By the end of the workshop you will be able to:

- use Markdown formatting to customise a document
- embed R code into a Quarto report
- dynamically embed plots and tables

The session will be highly interactive. Please, bring a laptop with internet access with you.

Some basic knowledge of R would be helpful - but not necessary.

Organised and presented by Myles Mitchell, Jumping Rivers

Day 4 Thursday 5 September

Data Science & AI

Data Hazards - communicating worst-case scenarios of data science projects

Kensington

In our Data Hazards workshop, participants will be given the opportunity to discuss the ethical decisions and implications associated with working on data science projects. To guide this discussion, we will provide an introduction to our Data Hazards labelling system (<https://datahazards.com/>), which gives researchers a shared vocabulary for ethical concerns in data science projects.

This workshop will bring together data scientists from diverse backgrounds to give their input on a variety of projects. We will give a brief example of a data science project working with healthcare data and medical scoring systems within the UK's National Health Service, and split off into small groups to discuss what potential data hazards might apply to such a project and how we could proactively seek to mitigate these hazards. Our hope is that participants can then go on and use the Data Hazards resources in their own work and facilitate their own discussions within various research groups.

Official Statistics & Public Policy

Official statistics for the public as well as for public policy

Clarence

Official statistics should be seen as an indispensable element in the information system of a democratic society, serving the government, businesses, and the public with data about the economic, demographic, social and environmental situation. This is according to the United Nations' fundamental principles for official statistics.

In this session, we would like you to think big – about the contribution that official statistics in the UK could and should make to our society and to reflect on where we are now, for example how far do you think UK official statistics are holding up in an increasingly complex data ecosystem and at a time when facts are questioned and re-interpreted?

There are two key pieces of work that may inform this conversation – the recently completed independent review of the UK Statistics Authority, and the RSS campaign for more public statistics, where the focus is on highlighting the questions to be answered in the public interest and then drawing on relevant data from other official and non-official sources, as well as official statistics.

We would like to build on these two sets of ideas and explore what more needs to be done so that official statistics in the UK can create the most value for the people of this country and then to touch on what the RSS might do (through the National Statistics Advisory Group NSAG and others in the society) in helping increase the public good.

We will be exploring all of this in a panel session. Do come along and take part.

Chair: Paul Allin, RSS Honorary officer for National Statistics

Panel members:

Denise Lievesley - author of the 2024 Independent Review of the UK Statistics Authority

Ed Humpherson - Director General, Office for Statistics Regulation

Anna Powell-Smith - Centre for Public Data

Camille Harrison - RSS Official Statistics Section and the Animal and Plant Health Agency (APHA)

Ken Roy - former Defra Head of Profession for Statistics

Session organised by Giuliana Battisti, Paul Allin and Jonathan Everett for RSS National Statistics Advisory Group

Other Applications of Statistics

Data driven perspectives on football

Buckingham

Understanding and application within women's football: Opportunities and challenges

Alice Harkness-Armstrong - University of Essex, UK

Applications of data and data science in elite football.

Laurie Shaw - City Football Group, Manchester, UK

Managing Expectations - A statistical view on football's managerial merry-go-round

Matt Penn - SENT Entertainment, London, UK

Organised by Bev Hale, University of Chichester

Part of the Statistics in Sport day, kindly sponsored by Amelco



Day 4 Thursday 5 September

13:00 – 14:00

Lunch

Durham Hall

13:15 – 13:45

Other

Royal Statistical Society Annual General Meeting

Ambassador

The AGM is a crucial session in the Society's calendar year, and an important reflection of the last year. It is a key opportunity for members to review RSS activities, finances, and governance arrangements.

RSS President, Andrew Garrett, will chair this meeting, offering his insight and leading in making pivotal decisions for the year ahead. Fellows will be asked to either approve or amend:

- The 2023 annual report and financial statements
- The re-appointment of the Society's auditors
- The minutes of the 2023 AGM

Non-members are very welcome to attend the event; however, only RSS fellows are eligible to vote on these matters.

Lunch served in room from 1pm

rss.org.uk/conference2024

#RSS2024Conf

48

14:00 – 15:00

Communicating & Teaching Statistics

Contributed: Tools for Interactivity and Auto-generation

Edinburgh

Auto-Generating Full Solutions to Mathematical Statistics Online Assessment Questions Using R

Thomas J Farrar - Cape Peninsula University of Technology, Cape Town, South Africa

Unlocking Intuition: When Escape Rooms Meet Statistics

Luke Middleton - University of Brighton, UK

An interactive application to conduct dose-response analysis using log-logistic model

Mintu Nath - University of Aberdeen, UK

Data Science & AI

Contributed: Real World and LLMs

Lancaster

A Comparative Analysis of Traditional Text Mining and Machine Learning for Extracting Information from Unstructured Text Data
Deepak Kovaichelvan - Rolls-Royce plc, Derby, UK

The Viability of Large Language Models for Conjoint and MaxDiff Analysis in Market Research

Chris Moore - Ipsos, London, UK

Garbage in garbage out: is your data worth the paper it is not written on?

Anne Rosemary Tate - University of Sussex, Brighton, UK

Environmental & Spatial Statistics

Contributed: Spatial statistics

Buckingham

Using geospatial disaggregation methods for small area population estimation in England and Wales

Matthew Plummer - Office for National Statistics, Newport, UK

Spectral estimation for spatial point processes and random fields
Jake P Grainger - EPFL, Lausanne, Switzerland

Conditional autoregressive models fused with random forests to improve small-area spatial prediction

Vinny Davies - University of Glasgow, UK

Medical Statistics

Contributed: Trials and NHST

Osborne

Seamless and adaptive designs in diagnostic accuracy studies: a review of the literature

Jordan L Oakley - Newcastle University, UK

A systematic comparison of methods for estimating heterogeneous treatment effects in large-scale randomized trials

Richard Dorsett - University of Westminster, London, UK

Day 4 Thursday 5 September

Medical Statistics

Contributed: Mortality, vaccines and fleeting moments

Balmoral

Modelling Cause-Specific Mortality Probabilities: An Application of Multinomial Logistic Regression

Daniel Cernin - University of Southampton, UK

Indirect treatment comparisons in vaccine: a case study comparing immunogenicity between 15- (PCV15) and 20-valent (PCV20) pneumococcal conjugate vaccines (PCV)

Shahrul Mt-Isa - MSD, Zurich, Switzerland.

Capturing Fleeting Moments: Compound Moves in Reversible Jump MCMC for Changepoint Detection in Single Molecule Imaging

Emily M Gribbin - Queen's University Belfast, UK

Methods & Theory

Contributed: Statistical modelling

Churchill

On point processes for modeling repairable systems: from better-than-minimal to minimal and worse-than-minimal repairs

Amy F Langston - Rhodes University, Makhandla, South Africa & University of the Free State, Bloemfontein, South Africa.

Hybrid classifiers for pairwise Markov models

Kristi Kuljus - University of Tartu, Estonia

Time-series imputation using low-rank matrix approximations

Owen D Jones - Cardiff University, UK

Other Applications of Statistics

Contributed: Modelling Risks

Gloucester

Reducing traffic injuries in day or/and dark for Vision Zero: A Bayesian spatial analysis approach

Jane Law - University of Waterloo, Canada

Comparing predictive microbiology models in food safety.

Ian Hunt

Biomathematics and Statistics Scotland, Edinburgh, UK

Prize winners

RSS Prize Winners: Best Presentations from YSM 2023

Clarence

Hear the prize winning presentations from the Young Statisticians Meeting 2023.

Calculating avoidable deaths for lung cancer survival disparities between black and white patients in the US

Zayn Rajan

A comparison of methods for external validation of the Kidney Donor Risk Index in the UK transplant population in the presence of semi-competing events

Stephanie Riley

Methods for estimating hidden populations

Albert Nyarko-Agyei

Organised by Jake Emmerson for RSS Young Statisticians Section

Social & Wellbeing Statistics

Contributed: Vulnerability, bias and growth

Sandringham

Exploring modern slavery and human trafficking vulnerability factors in the UK

Harry Bowles - The Alan Turing Institute, London, UK

Equity in Grant Funding: a multifaceted exploration of bias in peer review

Yesim Kakalic - The Alan Turing Institute, London, UK

Inclusive growth - how can data at local authority level support policymakers in meeting this objective?

Tej Nathwani - Higher Education Statistics Agency (now part of Jisc), Cheltenham, UK

Professional Development

Getting your work published and maximising its impact

Kensington

Are you looking to get your work published? Or do you want to raise the profile of your published work?

The session, jointly organised with Oxford University Press, will provide valuable advice for writing journal and magazine articles, give an overview of article review processes, and explore available self-promotional tools to raise the visibility of your work or research

How to get published in Significance magazine

Anna Britten – Royal Statistical Society, London, UK

Academic Journal Publishing: Submission, Publication, and Beyond

Martha Bailes - Oxford University Press, UK

Organised by Zexun Chen for RSS Young Statisticians Section

Day 4 Thursday 5 September

15:00 – 15:20

Refreshment break

Durham Hall

15:20 – 16:40

Applications in Business, Industry & Finance

Current challenges for Statistics and Data Science in Business & Industry

Sandringham

In this session we will hear from statisticians working in different sectors of business/industry about the current issues facing statistics and data science in the sectors in which they work. This will be followed by a audience discussion exploring current challenges and future directions for statistics and data science in business/industry.

David Smallbone - KPMG

Jennifer Visser-Rogers

Ron Bates - Rolls-Royce

Organised by Neil Spencer for RSS Business & Industrial Section

Stream sponsor Minitab 

rss.org.uk/conference2024

#RSS2024Conf

50

Communicating & Teaching Statistics

Schools, outreach and the digital world: how can we deliver statistical and data literacy for all?

Oxford Hall

The work of the RSS Education Policy Advisory Group (EPAG) seeks to improve the quality of the teaching, learning and assessment of statistics and data. It aims to influence education policy to ensure the curriculum is relevant and engaging, to support teachers to teach statistics and data science, and to support people who are not in formal education to be statistically literate – to ensure they have the skills to interrogate and navigate the world around them.

We believe that it is crucial to encourage curiosity whilst at school, reach out to those who might not be seeing the wonders of statistics and engage with those who have left formal education. As such, our panellists will share their insights on outreach (from an RSS William Guy Lecturer), on the challenges and opportunities around teaching statistics in the school curriculum (from a current teacher), and on engaging with the general population to inspire them about statistics and data (from a skills advisor at the Alan Turing Institute).

We will also consider how to improve statistical literacy in the general population, highlighting the key importance of statistical and data literacy to relevant real-world examples such as social media, AI, and media reporting, and the necessity of these skills to successfully understand and flourish in the world around us. The session will end with a panel and audience discussion on improving statistics and data science education, to ensure that no one gets left behind.

Panel members will include:

Thomas Gyllenship – current teacher

Matthew Forshaw - Newcastle University & The Alan Turing Institute.

Organised by Deniz Gursul for RSS Education Policy Advisory Group

Methods & Theory

Bayesian methods for spatio-temporal data that don't take ages

Edinburgh

In the analysis of modern datasets, where detailed information is increasingly collected, statistical methods for spatio-temporal data take on heightened significance. Whether applied to environmental monitoring, behavioural sciences, or other domains, spatio-temporal statistical techniques play a pivotal role in extracting meaningful insights and making accurate predictions. The four talks present new cutting-edge methods for analyzing and predicting spatio-temporal data, that build upon a wealth of Bayesian approaches, ranging from dynamic latent-space models and dependent Dirichlet processes to Hawkes processes and Gaussian processes.

Despite their flexibility and appealing interpretation, Bayesian methods often encounter non-trivial computational challenges when applied to large datasets. This session will address the computational aspect of the proposed models and offer innovative ways to speed up inference and increase their applicability inside and outside of the Bayesian community.

The session provides a comprehensive overview of the new trends in this exciting area.

Spatio-temporal stick-breaking process

Clara Grazian - University of Sydney, Australia

Cox-Hawkes: doubly stochastic spatiotemporal poisson process

Xenia Miscouridou - University of Cyprus, Nicosia, Cyprus, & Imperial College London, UK

Scalable dynamic network models for enriched demand planning

Emanuele Aliverti - University of Padova, Italy

Gaussian process regression on an incomplete and large spatio-temporal grid

Sahoko Ishida - University of Oxford, UK

Organised by Beatrice Franzolini & Francesca Panero for j-ISBA, the junior section of the International Society for Bayesian Analysis (ISBA)

Day 4 Thursday 5 September

Official Statistics & Public Policy

Creating Ambitious Population Statistics

Ambassador

In preparation of the 2023 National Statistician Recommendation, the Methods, Data and Quality Directorate, at the Office for National Statistics has been developing a range of new methods to enable the transformation of population statistics and maximise the use of alternative data sources. In this session, methodologists involved in this work will present the new approaches they have developed to meet the opportunities and challenges offered by non-survey data.

Using Bayesian Modelling to derive yearly probabilistic population and migration estimates from administrative data
Aidan Metcalfe - Office for National Statistics, UK

Estimating administrative data coverage to reduce bias in population stocks
Amy Large and **Eleanor Law** - Office for National Statistics, UK

Redefining migrations – using new sources of data for better migration statistics
Mark Edwards - Office for National Statistics, UK

Confident in our uncertainty estimates
Mark Edwards - Office for National Statistics, UK

Developing quality indicators
Karina Williams - Office for National Statistics, UK

From local authority level statistics to lower layer super output areas, how do we produce coherent annual statistics
Matt Plummer - Office for National Statistics, UK

Discussant: **Mark Fransham** - Oxford University

Chair: **Ana Basiri**

Organised by Camille Szmargd Harrison for RSS Official Statistics Section

Other Applications of Statistics

Paris Olympics 2024 Predictive Analysis Competition Winners

Clarence

To celebrate Paris Olympics 2024, the RSS Statistics in Sport Section ran an open competition to predict the Olympic Medals Table (before the Opening Ceremony) for Paris 2024. The entry with the highest correlation with the actual Medals Table, along with two entries chosen by the judging panel based on the methodology used, will present their work in this session. This will be followed by a panel discussion with the competition winners and Milt Mavrakakis from Amelco, the sponsors of the competition.

Mohamed Hammeda

Ian Simpson

Fernando Zapeda

Organised by Jessica Hargreaves for RSS Statistics in Sport Section

Part of the Statistics in Sport day, kindly sponsored by Amelco



Professional Development

Scale it with SQL

Lancaster

Structured Query Language (SQL) is the universal language of managing and querying the information stored in Relational Database Management Systems (RDBMS). SQL provides a standardised way to communicate with databases, enabling us to retrieve, manipulate, and analyse information stored in a structured database.

Join us in this workshop to explore the power of SQL in unifying data management and data querying tasks from the programming environment of your choice and integrating the analytical pipeline in your reproducible framework.

In the first section, the workshop will briefly introduce the concept of RDBMS, the SQL language framework emphasising DDL aspects like data types and important operators. It will also provide guidelines to install the necessary libraries in R and Python to harness functionalities, and instructions to use SQL functionalities in other applications (Microsoft Excel, Microsoft Access and SAS).

The second section of the workshop will dive into the DQL functionalities of SQL and explore multiple examples of querying the data. It will include a host of data wrangling issues including selecting columns, filtering rows, summarising data, searching, ordering and multiple types of aggregation of data.

In the final section of the workshop, we will get an overview of both DQL and DML functionalities of SQL where we will create an RDBMS system using SQLite and interrogate the SQLite database to query, modify and update data from R and Python environments.

The session will be delivered by citing bite-sized examples primarily from two programming environments: R and Python. However, we will also demonstrate the implementation of SQL in other common and popular applications (like Microsoft Excel, Microsoft Access, and SAS), therefore, all participants, regardless of their programming background, can follow the session. All instructions to install necessary programs and tutorial materials will be available online before the session.

Participants can follow and run scripts on their laptops during the workshop session.

Organised and presented by Mintu Nath, University of Aberdeen

Day 4 Thursday 5 September

15:20 – 16:40

Environmental & Spatial Statistics

Geopolitical Analyses of Disaster Impacts: Bias, Modelling and Decision Making

Balmoral

The frequency and severity of climate-related disasters like floods and wildfires are on the rise due to climate change, as reported by the Intergovernmental Panel on Climate Change (IPCC). But when it comes to understanding whether the disaster impacts, such as loss of life, damaged buildings, or agricultural losses, are also worsening, things get complicated. The ideal analysis would require data from almost every country, spanning several decades, covering a wide range of hazards like hurricanes, earthquakes, and storms. Unfortunately, bias can affect this data at many levels. For instance, some governments may downplay impacts for political reasons, while some may exaggerate them to receive more international aid.

We need accurate and representative hazard and impact data to train disaster risk models. In recent years, we've seen a surge in open-access data, open-source programming languages and significant improvements in statistical and machine learning models. This includes facilitated access to disaster footprint and population count data, endless Q&A forums and blogs on data science with Python and R, and professional-quality packages such as caret with hundreds of statistical and machine learning models at our fingertips. These advancements make it much easier to create accurate and useful models for disaster risk. Furthermore, advanced technologies, such as using anonymised mobile phone data to infer population displacement, have also become part of the state-of-the-art in disaster risk modelling.

In the immediate aftermath of a disaster, when there's no primary data available (such as on-the ground surveys of casualties or displacements), are predictions from disaster risk models influencing aid resource allocation? If not, what challenges are preventing us from taking a data-informed approach, and how do we ensure clear communication about the uncertainty, biases in data, and limitations of the models used? Lastly, in the realm of disaster risk modelling, how can we make sure that no one is left behind?

The session consists of six expert panellists, who will give short presentations, followed by a participative round-table discussion where the audience will have the opportunity to learn more about disasters and their impacts.

Faye Wyatt - UK Met Office

Justin Ginnetti - International Federation of Red Cross and Red Crescent Societies

Maria-Teresa Espinosa - Internal Displacement Monitoring Centre

Max Anderson Loake - University of Oxford

Sabine Loos - University of Michigan

Andrew Schroeder - Direct Relief

Organised by Hamish Patten, International Federation of Red Cross and Red Crescent Societies

Social & Wellbeing Statistics

An exploration of election predictors with reference to specific elections

Buckingham

According to Forbes, 2024 is set to be the biggest election year in history. Given this enormously significant confluence of events we will discuss election statistics: those numbers that track the progress and outcome of candidates and predictions that are made.

John Fry – University of Hull, UK

Lawrence McKay – University of Southampton, UK

Lucas Geese – University of East Anglia, UK

Timothy Martyn Hill

Organised and chaired by Timothy Martyn Hill

rss.org.uk/conference2024

[#RSS2024Conf](https://twitter.com/RSS2024Conf)



Day 4 Thursday 5 September

Medical Statistics

Now and then the relation between the 'medical' statistician and the clinical researcher

Kensington

The session will explore how far we have come since the paper written by the late Sir David Cox on the 50th anniversary of the Italian Statistical Society regarding the relationship between the statistician and the "substantive researcher worker". It will explore to what extent some of the issues highlighted in the paper still persist 35 years on and what progress has been made since then. It will also look at this relationship going forward in light of existing initiatives.

Organised by Mona Kanaan, Jason Oke, Tim Morris, Nathan Green, Christiana Kartsonaki and Rute Vieira for RSS Medical Section

16:50 – 17:50

Plenary

Keynote 7: 2024 Barnett Lecture: How to survey gibbons: An interdisciplinary story for the digital age

Oxford Hall

David Borchers - University of St Andrews

Nineteen of the twenty species of gibbon in the world are endangered or critically endangered, so getting reliable estimates of their distribution and abundance is important for conservation. But surveying gibbons is difficult because they live in thick forest and inaccessible habitat. Surveying is usually done by people standing in the forest for periods listening for gibbon calls, but this is very labour-intensive and full of potential for human error and subjectivity.

Digital acoustic recorders promise to be a much more cost-effective, objective and reliable means of surveying, but estimating animal distribution and abundance from nothing more than a set of .wav files presents a whole new set of challenges for statisticians.

This talk is the story of how a team that includes biologists, electronics engineers, computer scientists and statisticians have addressed, and are addressing these challenges. It is a story with some successes and some failures, that includes an acoustic balloon, an acoustic drone, and innovations in acoustic detector design, machine learning methods for acoustic object identification, and statistical methodology.

17:50 – 18:00

Plenary

Close of conference

Oxford Hall

19:30 – 23:59

Networking

Conference dinner

Empress Suite, Grand Hotel

WELCOME TO THE 2024 EDITION OF THE RDA WINTER SCHOOL
STOCKHOLM, SWEDEN

Red Door Analytics

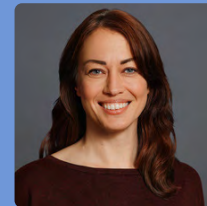
COURSE 1: 18-20 NOVEMBER

Statistical methods for population-based cancer survival analysis

This 3-day course will provide an overview of principles, methods, and applications for studies of cancer patient survival based on data collected by population-based cancer registries.



Professor Paul Dickman



Dr Sandra Eloranta

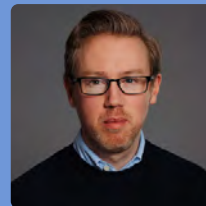


Dr Caroline Dietrich

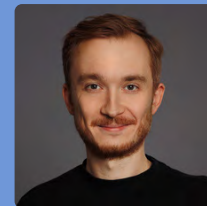
COURSE 2: 21-22 NOVEMBER

Joint modelling of longitudinal and survival data

This 2-day course will provide an overview of principles, methods, and applications for jointly modelling a longitudinal outcome, such as repeatedly measured blood pressure, together with a time-to-event outcome.



Dr Michael Crowther



Dr Alessandro Gasparini



Dr Sara Ekberg

OPEN FOR REGISTRATION NOW!
reddooranalytics.se

Thank you to those responsible for organising the conference:

Members of the RSS Conferences & Events Programme Board

Honorary Officer/Chair:

Scott Heald (Public Health Scotland)

Vice-Chair (2024 Programme Chair):

Bev Hale (University of Chichester/ Statistics in Sport Section/ RSS Council)

Karthik Bharath (University of Nottingham/ Research Section)

David Beckett (Office for National Statistics)

David Hoyle (Dunn Humby/ Manchester Local Group)

Rachael Hughes (University of Bristol)

Mona Kanaan (University of York/ Medical Section/ RSS Council)

Altea Lorenzo-Arribas (BioSS/ History of Statistics Section/ Celebrating Diversity SIG/ RSS Council)

Antonia Marsden (University of Manchester/ Manchester Local Group)

Lisa McFetridge (Queen's University Belfast)

Ciaran McInerney (University of Sheffield)

David McLernon (Institute of Applied Health Sciences Aberdeen)

Robin Nicolson (GSK/ Highlands Local Group)

Ben Swallow (University of St Andrews/ Environmental Statistics Section)

Joy Leahy (National Centre for Pharmacoeconomics/ Young Statisticians Section)

Andy Field (University of Sussex/ Teaching Statistics Section)

Kate Land (Inalytics)

Jon Simmons (Home Office)

RSS Staff and stewards

Paul Gentry (Head of Conferences & Events)

Janina Sasin (Conference Officer)

Mawreen Chapman (presentations coordination)

Joanna Ratajczak (registration team)

Maryjoy Kellman (registration team)

Teresa Espero (web & filming)

Alfie Jenkins (presentations desk)

Holly O'Brien (registration and general coordination)

Hannah De Mowbray (conference app & reporting)

Mags Wiley (reporting & social media)

Judith Shorten (Discussion Meeting)

Stuart McKendrick and fellow staff at the Royal Statistical Society.

Also, thanks to volunteers for their help with stewarding:

Sara Hilditch **Dario Domingo**

Jen Visser-Rogers **Jack Kennedy**

Katie Fisher **Lucy Teece**

Teresa Lee **Yoojung Che**

Zak Varty **Luke Middleton**

Production

Richard Gray for photography
www.rugfoot.com

Phil Sutton at Cavendish Design for design and production
www.cavendishdesign.co.uk

Andrew Mills and his team at Glenton Media for AV planning and support

Paul Creaser and team at VenulQ for the conference app

Venues

Mayesha Jalil and colleagues at the Hilton Metropole.

Sarah Shenston and the team at Brighton Palace Pier.

Charlotte Oatway and colleagues at the Grand Hotel.

Magdalen Smith, Cheryl Brunwin and colleagues at First Choice Conferences & Events for venue sourcing, site visit coordination, and for the accommodation booking service.

Darren Johnson at Visit Brighton Convention Bureau for venue sourcing and local information.

RSS International Conference 2025

THE UK'S ANNUAL SHOWCASE
FOR STATISTICS & DATA SCIENCE

EDINBURGH 1-4 September 2025

The RSS conference brings together the statistics and data science community from across the UK, Europe and around the world to share knowledge and learn about latest developments.

With attendees from over 30 countries and all sectors using statistics and data, this is a must-attend event for professionals, researchers, students and everyone interested in the impactful application of statistics and data science.

As usual, the conference programme will feature top keynote speakers, invited topic sessions, professional development workshops, contributed and rapid-fire talks, and poster presentations, as well as many opportunities for networking.

rss.org.uk/training-events/conference-2025

CONNECT • SHARE • LEARN

SUBMISSIONS
FOR INVITED
SESSIONS
NOW OPEN





Royal Statistical Society

12 Errol Street, London EC1Y 8LX

020 7638 8998

conference@rss.org.uk

rss.org.uk

CONNECT • SHARE • LEARN

rss.org.uk/conference2024 #RSS2024Conf @RSSAnnualConf

