STOR-i
excellence with impact

Lancaster University

STOR-i Workshop 2017, 12th-13th January

Abstracts

Day 1

EUSTACE: Latent Gaussian process models for weather and climate reconstruction
Professor Finn Lindgren, School of Mathematics, University of Edinburgh

The EUSTACE project will give publicly available daily estimates of surface air temperature since 1850 across the globe for the first time by combining surface and satellite data using novel statistical techniques. To this end, a spatio-temporal multiscale statistical Gaussian random field model is constructed, using connections between SPDEs and Markov random fields to obtain sparse matrices for the practical computations. The extreme size of the problem necessitates the use of iterative solvers, making use of the multiscale structure of the model to design an effective preconditioner.

Bandit approaches for border patrol
James Grant, STOR-i PhD student

We consider the challenge of patrolling a border using one or more camera-equipped drones. The problem is particularly interesting in the case where there is insufficient resource (drones) to cover all parts of the border simultaneously, and when little is known about the generating distribution of border crossings. In such a setting, one must aim to learn the optimal allocation of resource, while still endeavouring to maintain a high detection rate during the learning period.

This talk will split its focus between the mathematical formulation of the problem, as a variant of the (Combinatorial) Multi Armed Bandit problem, and discussion of Upper Confidence Bound techniques which we are developing to deal with the problem's particular features.

OR Saves Lives!
Professor Paul Harper, School of Mathematics, Cardiff University, Wales

Healthcare systems are stochastic in nature; that is they typically operate in an environment of uncertainty and variability, both at scale and within highly complex and connected networks. Furthermore, many healthcare services are under significant pressure to deliver more with less. OR methods can help healthcare providers move towards optimally configured services. This is much safer than experimenting with changes to the system for real and seeing what happens. Literally it can help save lives, for example in one major hospital our research completely redesigned the care for stroke patients, which resulted in a reduction in mortality rates by 60%. In another hospital emergency department, our work helped to save the Health Board £1.6m per year through optimised capacity planning. In this talk I will provide an overview of some current and recent OR healthcare modelling projects at Cardiff, highlighting different OR methodologies and demonstrating impact.

Paul is Professor of Operational Research and Deputy Head of the School of Mathematics, Cardiff University. He is also Director of the Health Modelling Centre Cymru (hmc2), a pan-Wales centre for modelling in healthcare, Director of Impact & Engagement for Mathematics, and a Fellow of the Learned Society of Wales. His research interests are primarily in mathematical/OR modelling and stochastic methods applied to healthcare systems, and he has been an investigator on in excess of £5.5m of funding from various research councils and direct from the health service. Paul

is an Editor for the journal Health Systems (Palgrave Macmillan) and in 2015 the Cardiff healthcare modelling team were awarded the Times Higher Education prize for 'Outstanding Contribution to Innovation and Technology'.

www.profpaulharper.com

Statistics in the world of catastrophe modelling
Ye Liu, JBA Risk Management, STOR-i PhD Alumni

The catastrophe modelling is a relatively young industry where experts from a wide range of disciplines join forces to estimate the potential loss due to catastrophic events. Statistical models are an integral part of this data-intensive process, providing methods for efficient simulation of complicated hazard scenarios and estimation of key model parameters.

This talk will showcase specific examples of how various statistical models are used at different stages of a full-scale flood catastrophe model, including the analysis of hurricane intensities in the North Atlantic Ocean, simulation of widespread flood events, and the rapid estimation of insurance losses caused by catastrophe events.
All of the examples are based on past or ongoing collaboration projects between JBA Risk Management and STOR-i, Lancaster University.

Estimation of Extreme Quantiles for Functions of Dependent Random Variables
Professor Qiwei Yao, Department of Statistics, London School of Economics

Motivated by a concrete risk management problem in financial industry, we propose a new method for estimating the extreme quantiles for a function of several dependent random variables. In contrast to the conventional approach based on extreme value theory, we do not impose the condition that the tail of the underlying distribution admits an approximate parametric form, and, furthermore, our estimation makes use of the full observed data. The proposed method is semiparametric as no parametric forms are assumed on all the marginal distributions. But we select appropriate bivariate copulas to model the joint dependence structure by taking the advantage of the recent development in constructing large dimensional vine copulas. Consequently a sample quantile resulted from a large bootstrap sample drawn from the fitted joint distribution is taken as the estimates for the extreme quantile. This estimator is proved to be consistent as long as the quantile to be estimated is not too extreme. The reliable and robust performance of the proposed method is further illustrated by simulation.

Information Engineering and Mining Data for Industry
Dr Steve King, Rolls Royce

We now take for granted Internet and mobile phone services that offer us instant access to information such as news and travel services updates. Many consumers are now making purchases via the internet for clothing, white goods, financial services and even groceries. Major retailers and shopping sites, such as Amazon, are able to construct a profile of our shopping preferences based on our history of transactions. This enables them to target advertisements and offers that are likely to be of interest to us. Credit card companies also use similar methods to identify unusual spending on our accounts to detect potential fraudulent activities. Understandably, Insurance companies are very keen to ensure their premiums will more than cover the cost of insurance claims from their client base, and will therefore want to get a reliable assessment of potential risk from various population groups. Each of these examples involves the collection and analysis of data and likely to include the process of what some might define as turning data into knowledge. Alternative descriptions include Data Mining and Exploratory Data Analysis.

This presentation provides an overview of some of the mathematical and statistical techniques that are used to derive useful information from data, but with the focus more on industrial applications. Hopefully, this will provide the audience with some insight of Data Mining and areas of where it can be used.

STOR-i
excellence with impact

Lancaster University

Day 2

On the optimal assignment of ground vehicles to ground tasks at an airport
Professor Giovanni Andreatta, Matematica Pura ed Applicata, Padova University

Modern airport management is facing, among other things, with increasing congestion of the apron area where several operations take place: aircraft taxiing, passengers and crew embarking and disembarking, refuelling, cleaning and many others. All of these ground-handling operations involve aircraft and a number of Ground Service/Support Equipment (GSE): follow-me cars, steps, tow-tractors, buses, baggage dollies etc. GSEs have to move throughout the apron, contributing to congestion and leading to possible increase in turnaround times, delays and accident rates. The efficiency and safety of ground handling operations is an important topic in airport management and has been the subject of several studies and of commercial tools concerned with aircraft-to-gate assignment, taxi planning, staff scheduling, passenger buses allocation and movements, GSE allocation etc.
The European Commission, within the Seventh Framework Programme, funded the project "Integrated Airport Apron Safety Fleet Management – AAS". Two demonstration airports were involved in the project. GSEs were equipped with On Board Units (OBU) gathering real-time information on position, current activity, state and other operational parameters, to be exploited by a decision support system for intelligent ground handling.
The presentation will focus on:
i) the proposed optimization approach to the problem of determining an optimal assignment of GSEs to ground-side operations,
ii) a discussion on how to trade efficiency of the suggested solutions with robustness against disrupting events.

Towards the online configuration of optimization algorithms
Dr Michael Epitropakis, STOR-i Associated, Management School, Lancaster University

The problem of the "online" configuration of algorithms is one of the long-standing grand challenges across different scientific fields. Adapting the behaviour (configuration) of an algorithm with an automatic, principled, and online way is an essential step for creating efficient and widely applicable real-world problem solvers that are not biased on human intuition, and/or prone to over-generalization.

Modern high-performance optimization algorithms involve a large number of design choices and parameter settings that have to be carefully tuned to efficiently solve a problem at hand. Traditional approaches involve trial-and-error design guided by expertise intuition, resulting in a limited exploration of the available choices that is time-consuming, biased and/or prone to over-generalization. Automatic algorithm configuration methodologies have emerged to efficiently address these issues/challenges. This problem has been identified across different scientific fields, such as Machine Learning, Artificial Intelligence, and Operations Research, with different names: offline/online parameter tuning, automatic algorithm configuration or design, hyper-parameter tuning, hyper-heuristics, meta-optimisation, parameter control, reactive search.

This talk will briefly summarize recent advances on automatic configuration of optimization algorithms in an online manner and discuss some of their benefits. Applications of such frameworks on challenging problems that arise in the fields of Operational Research and Software Engineering will be presented.

Best Practice Analytics
Detlef Nauck, BT, Industry Speaker

All large organisations are facing the challenge of integrating sophisticated analytics and the management of big data into their daily operations. Corporations have long realised the value of analytics for making evidence-based decisions, but they are finding it difficult to keep up with the evolving fields of business intelligence, data mining, predictive analytics, big data etc. that all somehow seem to be culminating into what is called Data Science today. The challenges of corporate analytics are not about which analytical tools to buy, which big data environment to deploy or which analytical methods produce the best predictive models. Corporations are struggling more with the

'Hows' of analytics instead of the 'Whats'. In this talk I will look at how BT's corporate research program addresses the challenges that Data Science poses to the business and give an outlook on how corporate analytics may continue to evolve and how academic research can contribute.

Biography: Detlef Nauck is Chief Research Scientist for Data Science with BT's Research and Innovation Division located at Adastral Park, Ipswich, UK. He is leading a group of international scientists working on research into Data Science covering areas like process analytics, network analytics, customer analytics, business modelling and autonomics for networks, processes and IT systems. Detlef focuses on solving business problems that require complex analysis of large and diverse data sources and the design of intelligent systems. His research programme focusses on making advances in practical machine learning methods and data analysis techniques and downstreaming them into the business. The aim is to enable networks, systems and processes to reach a level of self-awareness that allows them to automatically predict and prevent failures and so improve customer experience. Detlef has published over 100 papers and holds eight patents. He is a Visiting Professor at Bournemouth University and a Private Docent at the Otto-von-Guericke University of Magdeburg, Germany. Detlef holds an MSc (1990) and a PhD (1994) in Computer Science both from the University of Braunschweig, Germany. He also holds a Habilitation (post-doctoral degree) in Computer Science from the Otto-von-Guericke University of Magdeburg, Germany (2000).

Minimum Density Separators for Clustering

Katie Yates, STOR-i PhD Student

In the clustering problem, we seek to group data such that points assigned to the same group (cluster) are more similar to each other than points assigned to different clusters. We aim to partition dense clusters using low density separators. For computational tractability, it is necessary to restrict attention to hyperplane separators. An inherent limitation, therefore, is the inability to correctly identify clusters that are not linearly separable. We propose to overcome this restriction by mapping the input space into a kernel defined feature space where a linear separator corresponds to a non-linear separator of the original data. Through an experimental analysis, we show that our approach is capable of producing high quality partitions across various benchmark datasets.