

Annex 1 Case Studies

Science Case Studies:

Changepoints and online clustering:

Changepoint detection and online clustering have emerged as a key research theme within the CDT. Our work on these topics has been motivated by discussions and co-funded PhD projects with numerous partners including BT, Shell, DSTL, and Tesco. Our engagement has highlighted that the STOR toolbox is not fit for industrial purpose, e.g. for monitoring telecoms networks or oil field production. Important progress has therefore been required on key problems of mutual academic/industrial interest. Our work with a community of STOR-i students has enabled us to address many of the fundamental questions in this area. A key feature of this work has been the incorporation of novel optimization methods within the analysis framework. This has enabled us to develop substantial computational savings leading, for example, to faster detection of multiple changes. Outputs to date have appeared in highly-regarded international journals such as Bayesian Analysis; Journal of Computational and Graphical Statistics; Statistics and Computing; Journal of Machine Learning Research and IEEE Transactions on Pattern Analysis and Machine Intelligence. Lancaster's work in this area has also helped enable the UK to emerge as one of the leading international players in this area of growing importance, with STOR-i affiliated students regularly winning international student competition prizes. STOR-i affiliated staff were also invited to organise a month-long international research programme at the Newton Institute (<http://www.newton.ac.uk/event/ICP>).

One student working in this area is David Hofmeyr, he completed his PhD in 2015 and was subsequently awarded a STOR-i impact fellowship and used EPSRC IAA funds to develop his work with Kognitio. His research focused on the development of optimal projections for clustering and semi-supervised classification. David's research has led to journal papers in the two top machine learning journals: Journal of Machine Learning Research and IEEE Transactions on Pattern Analysis and Machine Intelligence, and papers in Statistics and Computing and 3 peer reviewed conference proceedings. Two other papers are currently under review. David was also awarded the best paper prize from the Conference of the Pattern Recognition Society of South Africa, and the Kingsman prize (awarded annually at LU to the top Management Science PhD). He is now a senior lecturer at U. Stellenbosch, South Africa.

Extreme Values:

Extremes values is a long-term area of strength at LU, with a tradition of research at the interface between methodology and substantive applications. It produced two 4* REF Impact Studies in REF2013. The CDT has taken this work to a new level of user-led scientific development with Shell, the Met Office and JBA all co-funding repeated studentships, 9 in total over both CDTs, and becoming fully integrated into the research of the group. Specifically Shell's Statistics team, consisting of 3-4 staff working on extremes, visit 2 days a month to work with us on developing new methods for their oil-rig risk assessment. JBA have employed 3 of our PhD students and actively collaborate with us through a KTP. We were highly commended on our joint work with JBA on spatial flood risk assessment for the UK Government's National Flood Resilience Review (2016). This used methods developed through the CDT. The depth and importance of our joint work with the Met Office is set out below. This major development of exciting research opportunity created by the CDT has helped LU attract Wadsworth, a rising star in statistics (RSS Research Prize 2015 and EPSRC Fellowship to be held 2016-8).

Our industrial partnerships have highlighted that existing methods are not fit for use in some major industrial applications where the structure and scale of data and the nature of the systems concerned introduce novel challenges. Outputs from the group, over the new CDT, span the top probability, statistical theory, applied statistics and domain journals, including: Advances in Applied Probability, JRSS B, Biometrika, Extremes, Journal of Multivariate Analysis, Technometrics, Annals of Applied Statistics, Applied Statistics, Dynamics and Statistics of the Climate System and Ocean Engineering. LU's work in this area has also helped re-affirm the UK as one of the leading international players in this area of growing importance.

One student working in this area is Hugo Winter, who graduated from STOR-i in 2015. His PhD was with the Met Office and developed statistical models for spatial and temporal heatwaves using extreme value methods. The quality of his research is evidenced by the award of the LU Faculty of Science and Technology prize for top PhD student. He has papers in JRSS Series C, Annals of Applied Statistics, Extremes, and Dynamics and Statistics of the Climate System. The methods developed by Hugo are already having an impact. They are being used in the Met Office's current work on European temperature records to inform the UK Government of the present and future risk of prolonged periods of extreme heat. He is now a Research Engineer for EDF Energy, where he is in charge of their extreme event analysis work. This will in turn have a critical impact on the flood defence design for their new nuclear power stations. He is already the industrial co-funder of a STOR-i student.

User Case Studies:

Our two user studies cover our novel partnership model and an example where the work of a STOR-i student has made substantial impact.

Industrial Partnership Model:

Historically the amount of industrially engaged, co-funded STOR research in the UK has been very limited. Arguably, therefore, one of the greatest influences of STOR-i has been to affect a culture change within the user community, developing a pioneering model of co-created, fundamental STOR research at scale motivated by, and feeding back into, areas of industrial priority. What has been particularly heartening has been to see the support of business across a range of sectors including Defence, Education, Energy, Environmental Services, FMCG, Oil & Gas and Telecoms.

Our success dates back to the very inception of the CDT in 2009/10. The proposal was originally developed with the committed support and input of a core group of organisations who sought to improve the STOR people pipeline. Since then, through persistent collaboration and substantial time-investment, we have developed a rich eco-system of industrial partners ranging from SMEs to multinationals. Currently they number 25 companies contributing to the new CDT, with 60 contributing since the Centre's initial launch. Key to the partnership has been building relationships that have depth as well as a focus on delivering quality collaborations. This is achieved via an integrated approach: including supervisors regularly visiting to provide training, attending our conference, student research placements and critically providing ample opportunities for extensive networking with all the CDT's students.

We are, in particular, grateful to our long term partners including BT, Shell and ATASS. With them we are jointly developing and delivering on long-term research plans based on extreme values (5 PhD students funded), non-stationary time series (6 PhD students funded), and sports modelling and betting markets (3 PhD students funded). These partners take the engagement with us to a different level. Shell's statistics team visit for 2 days a month; BT are heavily engaged in other long-term research initiatives (such as StatScale) and helped support the involvement of PhD students within its broad activities; ATASS fund all our students to attend a training event they host at their Exeter headquarters giving the students experience of their unique business working environment. We are particularly pleased that in the new CDT we have developed a new major OR focussed partnership with Rolls Royce, who already co-fund 3 studentships.

The scale and diversity of activity we have undertaken is, to the best of our knowledge, internationally unprecedented. This view supported by John Birge (Chicago), former President of INFORMS, who writes that "STOR-i ...[is] . . . the most innovative and comprehensive collaboration between industry and academia that I have seen in any doctoral program. . . the broadest and most industry integrated OR curricula anywhere in the world." It has also been heartening to see other UK and overseas CDTs attempt to replicate many of our best practices.

BT:

BT has been a core partner of STOR-i since 2010. To date, they have sponsored 5 PhD students within the STOR-i programme. Key to the BT-STOR-i partnership has been a shared focus on the development of novel and efficient STOR methods to address key operational challenges. These include workforce planning, network performance and fault detection. Methods developed have already been used to demonstrate proof of concept on key operational issues, and have helped highlight new insights and opportunities for the organisation.

The work of STOR-i student, Emma Ross provides an insight into the user benefits. Emma graduated from STOR-i in 2016. Her PhD developed multi-skilled workforce planning models for large-scale service industries in collaboration with BT. Emma's work has provided a solution to automate the allocation of BT engineers which takes account of their full range of skills. Her algorithms are now used in BT's planning systems, demonstrating clear benefits to the company. Emma has extended this work to incorporate the accumulation of incomplete work across a planning horizon. The associated paper *Cross-Trained Workforce Planning for Service Industries: the Effects of Temporal Demand Flexibility* is currently under review with the *Journal of the Operational Research Society*. She initiated collaboration with the Norwegian School of Economics by using STOR-i's Research Fund for an extended visit. This resulted in a paper submitted to the *European Journal of Operational Research* which develops higher-level planning models for the recommendation of training actions which mitigate future demand uncertainty.

Student Case Studies:

Our two student case studies illustrate both how our CDT's training enables students from different disciplines to transfer over to be highly successful in Statistics and Operational Research (areas of severe national shortage), and also how the breadth of our skills training has developed a very rounded and mature researcher with excellent technical skills.

Tim Park:

Tim exemplifies the core case for STOR-i's original funding to increase the pool of STOR postgraduates. Tim graduated from STOR-i in 2014 from our first cohort. He came to STOR-i from a physics degree and excelled, quickly picking up new STOR knowledge and skills and weaving these with his physics background. His PhD with Unilever was on *Wavelet Methods for Multivariate Non-Stationary Time Series*. The work led to a paper in *IEEE Transactions on Signal Processing*. He also won an American Statistical Association award for its student paper competition in statistical learning and data mining at the Joint Statistics Meeting (2014). The research has had substantial impact on Unilever. Details of this are commercially sensitive and not in the public domain at this time. Tim now works for Shell Global Solutions where he applies statistical methods to a wide range of problems in the oil and gas industry. Tim is now a funder/industrial supervisor for a STOR-i PhD student.

Kaylea Haynes:

Kaylea's PhD was awarded in 2016. Kaylea's work focused on computationally efficient methods for searching for changes in statistical structure within massive data streams, funded by DSTL. Two articles have already been accepted in leading journals (Statistics and Computing, Journal of Computational and Graphical Statistics) with a third popular science article appearing in Significance. A fourth article will result from collaboration with Prof. Axel Munk (U. Gottingen), a world-leading authority in nonparametric statistics. The impact of this work is reinforced by Kaylea's strong commitment to reproducible research, making her methods available in the form of open-source R packages, including substantial work on the Statistics group's flagship R packages (changepoint). Recognition for Kaylea's work is already strong, as evidenced by her winning the highly competitive international student paper award competition to present at JSM (the leading statistics conference in USA).

Kaylea's natural drive and enthusiasm extend beyond traditional PhD-level academic boundaries. She has also been an excellent ambassador for STOR-i and the wider university, independently identifying and establishing links with new organisations (e.g., Kognitio). The quality of her all round performance as a PhD student is illustrated through the award of the LU Faculty of Science prize for top PhD student. She is now working in sports analysis for a new start up company.