Institution: University of Nottingham

Unit of Assessment: UOA 10 – Mathematical Sciences

Title of case study: Incorporating expert knowledge in complex industrial and policy applications

1. Summary of the impact

Techniques developed at The University of Nottingham (UoN) have enabled organisations to deal with uncertainty in complex industrial and policy problems that rely on the elicitation of expert opinion and knowledge. The statistical toolkit produced for use in complex decision-making processes has been deployed in a wide range of applications. It has been particularly useful in asset management planning in organisations such as the London Underground, government approaches to evidence-based policy, and the Met Office UK Climate Projection tool (UKCP09), which is used by hundreds of organisations across the UK such as environment agencies, city and county councils, water companies and tourist boards.

2. Underpinning research

Professor Tony O'Hagan's (UoN 1990-1999, Department of Mathematics and School of Mathematical Sciences) research at Nottingham between 1993 and 1998 showed how to incorporate expert knowledge into statistical analyses. Classical statistics primarily deals with situations in which there are large amounts of data, so that information in the data swamps any modelling choices made. O'Hagan's work focussed on the other extreme, where data are limited. He developed a toolkit of methods for incorporating expert beliefs in statistical analyses, enabling accurate uncertainty quantification in complex problems. The research comprises two related threads: directly using and eliciting expert judgements; and using expert knowledge that has been coded into a complex computer simulator.

Eliciting and using expert judgements

Elicitation is the process of translating someone's beliefs about an uncertain quantity into a probability distribution, and is a key aspect of many uncertainty quantification exercises, particularly risk assessments. While the incorporation of expert beliefs in a prior distribution is a fundamental part of Bayesian statistics, the problem of how to elicit these beliefs received relatively little attention before the 1990s. Expert elicitation is now widely used; for example, it has been conducted and used by at least six American federal agencies and international organisations [B6].

O'Hagan was a leading proponent of the use of subjective priors and argued strongly that the subjective interpretation of probability should be embraced. This was described in his influential textbook [A1] and in two papers [A2, A3]. He examined in detail the elicitation process, and determined which questions lead to the most accurate representation of an expert's beliefs, particularly when eliciting complex quantities such as variances and distributions.

This research was motivated by O'Hagan's work with UK water companies on developing asset management plans. The ABLE project [A2, A3, A4], with £800K of funding from 8 water companies and the DoE (now DEFRA), arose from this work, and led to a methodology and software which is used to estimate the need for (and the cost of) maintenance, renewal and replacement of assets (sewers and water pipes in the case of the water companies).

Statistical Analysis of Complex Computer Simulators

Simulators are now widely adopted in science, engineering, industry and public policy to encode and combine expert knowledge of mechanisms and processes in order to make predictions. The work of O'Hagan facilitated this by showing how to combine complex mechanistic models with statistical models in order to quantify uncertainties.

O'Hagan (with PhD students Haylock, Kennedy and Oakley) published a series of research papers between 1996 and 2002 showing how to quantify the different types of uncertainty present when using complex simulators [A5, A6, A7, A8]. The basis of these papers, and the key methodological

development, is the use of Gaussian process *emulators* (a term coined by O'Hagan) as meta-models of simulators. For simulators that are computationally expensive to evaluate, the ability to draw inferences is constrained as the simulator can only be evaluated a limited number of times, making standard uncertainty quantification approaches impossible to apply. O'Hagan showed [A6, A7] how to build a statistical model of the simulator, called an emulator, which is cheaper to run and is trained using an ensemble of simulator evaluations. Once the emulator is available, tasks such as uncertainty analysis [A8], sensitivity analysis [A5], and calibration [A7] could be undertaken, which would otherwise have been impossible.

Research on Gaussian process emulators continues within the School, applications including to sea surge and wave levels (Dr Richard Wilkinson, UoN 2009 to date) and carbon capture and storage (Professor Andrew Cliffe, UoN 2005 to date, and Wilkinson).

3. References to the research

The three publications that best indicate the quality of the research are indicated *

[A1]* O'Hagan, A. (1994). Kendall's Advanced Theory of Statistics Volume 2B, Bayesian Inference. Edward Arnold, London. ISBN: 0-340-80752-0 (available on request)

[A2] O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications. Journal of the Royal Statistical Society: Series D (The Statistician), 47, 21-35. (with discussion, pp 55–68). DOI: 10.1111/1467-9884.00114

[A3] Garthwaite, P. H. and O'Hagan, A. (2000). Quantifying expert opinion in the UK water industry: an experimental study. Journal of the Royal Statistical Society: Series D (The Statistician), 49, 455-477. DOI: 10.1111/1467-9884.00246

[A4]* O'Hagan, A. (1997). The ABLE story: Bayesian asset management in the water industry. In The Practice of Bayesian Analysis, S. French and J. Q. Smith (eds.). Arnold, 173-198. ISBN-10: 0340662409; ISBN-13: 978-0340662403 (available on request)

[A5]* O'Hagan, A., Kennedy, M. C. and Oakley, J. E. (1999). Uncertainty analysis and other inference tools for complex computer codes (with discussion). In Bayesian Statistics 6, J. M. Bernardo *et al.* (eds.). Oxford University Press, 503–524. ISBN-10: 0198504853; ISBN-13: 978-0-19-850485-6 (available on request)

[A6] Kennedy, M. and O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. Biometrika 87, 1–13. DOI: 10.1093/biomet/87.1.1
Funded by [A9] and first released as Nottingham Statistics Group Technical Report 1998-09.

[A7] Kennedy, M. C. and O'Hagan, A. (2001). Bayesian calibration of computer models (with discussion). Journal of the Royal Statistical Society B 63, 425–464. DOI: 10.1111/1467-9868.00294
Funded by [A9] and first released as Nottingham Statistics Group Technical Report 1998-10.

[A8] Oakley, J. E. and O'Hagan, A. (2002). Bayesian inference for the uncertainty distribution of computer model outputs. Biometrika 89, 769-784. DOI: 10.1093/biomet/89.4.769
First released as Nottingham Statistics Group Technical Report 1998-11.

Grants:

[A9] Bayesian uncertainty analysis and calibration of complex computer models, PI O'Hagan, EPSRC Grant GR/K54557/01, 1 October 1995 – 30 September 1998, £114,666.

4. Details of the impact

Asset Management

For companies with large infrastructure such as UK water companies, the development of an asset management plan is an expensive and time-consuming task. O'Hagan described how to combine a few high-quality studies with many lower-level expert judgements, thus realising great reductions in

the number, and hence costs, of detailed studies required [A4]. Convincing the water companies to change their approach based in classical statistics for an approach using subjective Bayesian methods (O'Hagan's ABLE methodology) represented a major success for Bayesian statistics at this time (mid to late 1990s).

From 2007 to 2011, Metronet Rail and then London Underground ran the Engineering Strategy for Economic and Efficient Management (ESTEEM) project to optimise and manage the investment in their engineering assets over a 100-year period. O'Hagan was employed as the statistical consultant for the project, and worked with ESTEEM to develop a new approach, based on ABLE [A4], for asset management on the railways by estimating asset degradation, costs, risks and their probabilities for each maintainable item. Uncertainty was carefully characterized throughout the rail system, so that credible ranges could be placed on estimates for individual assets through to the company-wide asset management plan. The project won the prestigious international Institute of Engineering and Technology award for best asset management innovation in 2010 [B1] and an independent audit [B2] has concluded that ESTEEM represented a "world best practice" approach to asset management for long-lived assets.

Following a re-organisation in September 2009, aspects of the project continued such as the Least Whole Life Cost (WLC) predictions, and these have led to significant benefits, as noted by John Darbyshire, project manager for ESTEEM at both Metronet and Transport for London [B3]:

"The ESTEEM Civil Assets least WLC predictions substantiated a basis for long term investment in the asset base and justified inclusion of preventative maintenance in a new performance contract for maintaining those assets. A particular example of the benefits reaped from this project is in the waterproofing of structures. Prior to ESTEEM, this was thought to be too expensive to justify. However, ESTEEM predictions anticipated a 20% savings in maintenance costs over a 30-year period, a saving of £5m p.a. The water-proofing was thus implemented at the end of 2009 for all concrete and masonry structures and continues to this day."

Darbyshire goes on to note that: *"A further example is the information systems used in London Underground stations. ESTEEM has become a critical operational system used to maintain and update these systems, and was fully implemented by summer 2011."*

Evidence-Based Policy

In 2009, the Food and Environment Research Agency (FERA) conducted a consultation into the formation of a new body to manage livestock diseases in the UK [B4]. The impact assessment for the consultation required estimates of the cost and frequency of various exotic infectious disease outbreaks, such as bluetongue and avian influenza. The government's timetable for the publication of a draft bill meant that there was no time for a large-scale project to assess these costs, so instead expert elicitation techniques were used to capture government veterinarians' and economists' knowledge and uncertainty about these quantities. Because the estimates from the experts were highly uncertain, it was decided that this uncertainty should be accounted for in order to make an informed decision about the value of the new body.

The FERA consultation [B4, B5] used the methodology developed by O'Hagan [A2, A3] to elicit 90% credible intervals for each of the unknown values (i.e. an estimate of uncertainty). They used O'Hagan's approach for fitting parametric distributions to the experts' opinions [A4], and followed his advice for avoiding under-estimation and for suggested feedback to the experts to ensure their satisfaction with the final results [A3]. The consultation formed a key part of the Draft Animal Health Bill [B5] presented to parliament by the Secretary of State for Environment, Food and Rural Affairs in January 2010. Although the bill never passed into law due to the change of government in June 2010, it demonstrated the increasing acceptance of the need for the quantification of uncertainty when using expert opinion, and the use of elicitation methodology to quantify this uncertainty, in government policy [B6].

Emulators

One example of the use of O'Hagan's fundamental research on emulators and the analysis of

simulators is UKCP09, an £11M DEFRA and DECC funded project, developed at the Met Office for predicting climate change and its effects in the UK. The climate projections in UKCP09 are based upon simulators such as the HadCM3 simulator, which contains over one million lines of code and takes several days to run at a single combination of parameter values. A key aspect of UKCP09 is that users are provided with the uncertainties associated with future climate information, something that would not be possible without the use of emulators.

Emulators are used by the Met Office as statistical surrogates of HadCM3 to assess the uncertainty associated with the simulator prediction [B7], because it would be too costly to run the simulator enough times to quantify the uncertainty in any predictions. The statistical methodology used in UKCP09 is based on methodology developed directly from O'Hagan's research at UoN. For example, UKCP09 uses the emulator methodology described in Rougier [B8], which draws extensively upon O'Hagan's work [A1, A8].

UKCP09 is used by a wide range of organisations, government departments, and programmes (http://ukclimateprojections.defra.gov.uk/23081). For example, Hampshire County Council (HCC)'s Emergency Planning Unit used UKCP09 to undertake comprehensive mapping of Hampshire [B9]. This enabled them, in conjunction with other tools, to anticipate future emergency planning service demands and determine if existing provisions were adequate (e.g. in heatwave events the frequency of events was used to estimate the number of deaths). Meanwhile, South West Tourism (SWT) used UKCP09 to explore the potential impact of climate change to their region [B10], results of which were included in a report that was used to inform regional tourism strategy, which notes e.g. that *"The UKCP09 regional projections and the results from this study also show that extreme weather events are likely to increase for the SW, so the likelihood of erratic and extreme weather also needs to be investigated and considered."* (Page 37 of report to be found at [B10].)

5. Sources to corroborate the impact

[B1] The Institution of Engineering and Technology Awards past winners: http://tv.theiet.org/channels/news/10135.cfm (copy also filed 6 September 2013, together with original application form)

[B2] AMEM Assessment of ESTEEM. A report for LUL Nominee BCV Limited, Trading as Metronet Rail BCV from Asset Management Consulting Limited (AMCL), May 2009. (copy on file)

[B3] Stations Whole Life Cost Manager, Transport for London (2006-2011). (copy of letter on file)

[B4] Gosling, J. P., Hart, A., Mouat, D. C., Sabirovic, M., Scanlan, S. and Simmons, A. (2012) Quantifying Experts' Uncertainty About the Future Cost of Exotic Diseases. *Risk Analysis, 32(5) 881-893.* DOI: 10.1111/j.1539-6924.2011.01704.x (copy also on file)

[B5] Annex 8, J.P. Gosling (FERA), Draft Animal Health Bill, HM Government, January 2010 www.official-documents.gov.uk/document/cm77/7784/7784.pdf (copy also on file)

[B6] Expert Elicitation Task Force, White Paper. Prepared for the U.S. Environmental Protection Agency, Washington, DC, 2011. www.epa.gov/stpc/pdfs/ee-white-paper-final.pdf (copy also on file)

[B7] Statistical methodology used by UKCP09: http://ukclimateprojections.defra.gov.uk/23253, http://ukclimateprojections.defra.gov.uk/22783 (copies also on file)

[B8] Rougier, J. (2007). Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change*, *81*(3-4), 247-264. DOI: 10.1007/s10584-006-9156-9

[B9] UKCP09 use by HCC: http://ukclimateprojections.defra.gov.uk/23089 (copy also on file)

[B10] UKCP09 use by SWT: http://ukclimateprojections.defra.gov.uk/23098 (copy also on file)