Structured expert judgment using the Classical Method

When complex decisions must be made while data is unavailable, structured expert judgment can be used to combine uncertainty distributions resulting from experts' assessments. Roger Cooke, the Chauncey Starr Senior Fellow at Resources for the Future, and Emeritus Professor at Delft University of Technology, created the well-known Classical Model to quantify uncertainty using expert opinion. Together with Dr Tina Nane, from Delft University of Technology, and Dr Anca Hanea, from the University of Melbourne, Professor Cooke presents 'Decision Making Under Uncertainty: Introduction to Structured Expert Judgment', the only freely available online course on structured expert judgment.

Image by U.S. Government / Rawpixel.com

Expert opinions are frequently sought when complex decisions must be made in situations where appropriate information cannot be acquired from existing data and models. Experts are asked to quantify their uncertainty over quantities of interest that inform the decision-making process. Furthermore, the experts are unlikely to be in complete agreement with one another. In such situations, expert judgement can be employed to quantify the uncertainty that ensues and to aggregate expert opinion.

Roger Cooke, the Chauncey Starr Senior Fellow at Resources for the Future, and Emeritus Professor at Delft University of Technology, created the Classical Model, also known as Cooke's method, for quantifying uncertainty when using expert opinion. Throughout the three decades since its formulation, the Classical Model has been used to perform structured expert judgment in a diverse range of applications, including climate change, disaster management,

epidemiology, public and global health, ecology, aviation, nuclear safety, environment, and ecology. In addition, together with Dr Tina Nane, also from Delft University of Technology, and Dr Anca Hanea, from the University of Melbourne, Professor Cooke presents the only available online module on structured expert judgment.

Expert judgment can range from asking an individual expert for their best guess, to following a formal, structured approach to systematically obtain and combine probabilistic judgments. This synthesis of opinions is called expert elicitation. The validation of expert judgments is challenging since they are only called for when other data are unavailable. Measuring their accuracy is, therefore, an arduous task.

STRUCTURED EXPERT JUDGMENT
Structured expert judgment aggregates experts' uncertainty distributions. Cooke explains that structured expert judgment methods are intended to

'quantify uncertainty, not to remove it from the decision process'. If expert data is to be recognised as scientific data, it should be subjected to the same quality controls as any other kind of data. He proposed a class of methods, known as structured expert judgment, that satisfy four principles required for any method described as 'scientific'. These are scrutability/accountability, neutrality, fairness, and empirical quality control. Cooke's Classical Model is arguably the most rigorous method for quantifying uncertainty by using expert opinion.

THE CLASSICAL MODEL
The naming of the 'Classical Model' highlights the method's association with classical statistics. The Classical Model uses objective performance measures to validate expert opinion. The experts assess uncertain target questions together with a set of calibration questions. The calibration questions are from the experts' field of knowledge, have observed true values and often involve data from official reports that

have not yet been made public. The experts are scored on their performance in assessing the calibration questions. Validation is achieved by assessing the statistical accuracy of an expert's assessments together with how much information they provide. These two quantitative measures of performance are used to calculate performance-based weights.

The Classical Model combines experts' distributions using these performance-based weights, optimising the performance of the combined expert, or 'Decision Maker'. Several interesting mathematical issues arise from optimising this performance-based aggregation. In non-technical terms, Cooke describes how the performance measure must reward both the experts' statistical accuracy and informativeness, while discouraging the experts from stating judgments that differ from their true opinions. The performance of the Decision Maker can be evaluated in the same way as that of the experts, using the same performance measures. Performance-optimised Decision Makers correspond to virtual experts and can be adopted by the real-life decision maker. Cross validation is also applied whereby subsets of calibration variables are used to form weights and predict the excluded calibration variables.

APPLICATIONS
Numerous studies have been conducted using the Classical Model. Highlights among these include its application to nuclear safety in the 1990s in research being carried out by the European Union and the United States Nuclear Regulatory Commission. During the prolonged volcanic eruption on the island of Montserrat in the West Indies, from 1995 to 2018, the Classical Model was a key decision-support procedure. Harvard University and the Kuwait government used Cooke's method in their 2004–2005 study of fine particulates, pollution in the form of tiny particles or droplets suspended in the air. It was also used in an investigation into foodborne diseases for the World Health Organization in 2011–2013.

ILLNESSES TRANSMITTED BY FOOD AND WATER
More recently, the Centers for Disease Control and Prevention and the University

of Florida, together with Roger Cooke, Tina Nane, and Willy Aspinall, performed a large, structured expert judgment study using Cooke's Classical Model in their work to control and prevent illnesses transmitted through food and water in the United States. The expert elicitation took place at a two-day workshop in May 2017 and involved 48 experts from various professional and scientific backgrounds. Estimates were obtained for the proportion of 33 pathogens including bacteria, such as Salmonella and Legionella, and viruses, such as norovirus and hepatitis A, attributed to each of five major transmission pathways (foodborne, waterborne, person-to-person, animal contact, and environmental), and six associated sub-pathways.

The researchers commented on how the method made it possible for the estimates

Structured expert judgment methods are intended to 'quantify uncertainty, not to remove it from the decision process'.

to be informed by multiple data sources, such as outbreak surveillance data, studies of sporadic illnesses, case reports, and the experts' professional knowledge. They also pointed out that using calibration questions to weigh expert responses, a unique feature of the Classical Model, 'introduces mathematical rigor not found with other elicitation methods'. The findings provide an understanding of the multiple transmission pathways for the identified pathogens and support the targeting of resources and prioritisation of public health interventions, as well as informing policy.

CLIMATE CHANGE
Climate change is riddled with uncertainty, and Cooke observes that both the scientific community and the general population make errors when reasoning under uncertainty, and fail to convey it accurately. Faced with multiple uncertain quantities, most people will identify what they consider the most likely outcome for each quantity and then reason as if those values were certain. Uncertainty is 'taken into account'

after the fact by adding qualifiers such as 'highly confident', 'most likely', and 'virtually certain'. This may be satisfactory when deciding whether to take an umbrella to work, but not when deciding how society should deal with climate issues impacting life as we know it. The uncritical aggregation of high confidences sets and baits the 'confidence trap': thinking that high confidence in each of several statements confers high confidence in all statements jointly. Consider: you may be highly confident that a six will not come up on the first throw of a dice, and on the second, and third and fourth. However, the probability is about one half that a six will come up on one of the four throws. Reasoning under uncertainty must obey the laws of probability, even if the probabilities are subjective. Communicating

uncertainty to the lay public is difficult, but if the communicators themselves don't understand uncertainty, then it is well-nigh impossible. This, in Cooke's opinion, is a major challenge of dealing with climate change. We must decide before the facts are in. That means deciding under uncertainty – which we do very badly.

In a review of 49 professionally contracted studies, Cooke highlighted the challenges involved in ensuring that the use of expert subjective probabilities is scientific. This evaluation revealed pervasive overconfidence among experts. It gave insight into how the role of domain expertise and experience can affect statistical accuracy and informativeness. Moreover, the review demonstrated the need for cross validation, to gauge how well performance on calibration variables predicts performance on the variables of interest.

Researchers at the University of Bristol, UK, Princeton University and Rutgers

University in the US, and Resources for the Future, have completed a significant structured expert judgement study into climate change. They employed the Classical Model to investigate the contribution made by the dynamic effects of ice sheets to the global mean sea-level rise.

Forecasting the imminent rise in sea level is challenging. Even so, the quantification of future sea-level rise uncertainties, particularly upper-end estimates, are urgently required to inform adaptation strategies. Expert elicitation took place at two separate, two-day workshops held in the US and UK in 2018 and involved 22 experts. The format and questions were identical, so that the findings could be combined using the Classical Method. The research team found that by 2100, sea-level rise could exceed 2m, more than twice the upper value put forward by the United Nations Intergovernmental Panel on Climate Change in the Fifth Assessment Report. Moreover, this would have profound consequences for humanity with a potential land loss of 1.79 million km², including areas of food production, and up to 187 million people displaced.

The article detailing this study, published in the *Proceedings of the National Academy of Sciences of the United States of America* (*PNAS*), has received a great deal of attention. It has been mentioned in 331 news stories from 263 outlets, and as the 70th most-discussed paper in 2019, it sits in the top 5% of all research outputs scored by altmetric.com.

ONLINE MODULE
The Classical Model is also the focus of a Massive Open Online Course (MOOC): 'Decision Making Under Uncertainty: Introduction to Structured Expert Judgment', presented by Roger Cooke, Tina Nane, and Anca Hanea. This course is arranged into six parts, combining theory and applications in an interactive and engaging manner.

Learners are introduced to structured expert judgment methods, particularly the Classical Model, and advised as to why and when they should be applied. They learn how to account for uncertainty assessments and biases within a complex decision-making

Image by U.S. Government / Rawpixel.com

The Centers for Disease Control and Prevention applied the Classical Model in their work to control food-borne illnesses, including Salmonella.

> ## Using calibration to weight expert responses, a unique feature of the Classical Model, 'introduces mathematical rigor not found with other elicitation methods'.

setting. There is also the opportunity to analyse expert data with EXCALIBUR or Anduryl, software packages implementing the Classical Model, and obtain answers to questions of interest. Optional modules include exploring dependence elicitation and eliciting probabilities, applying structured expert-judgment methods to real-world scenarios, and using a different method, the IDEA Protocol module. This is the only available module on expert judgment. In its three runs to date, it has attracted more than 7,000 participants from 121 countries.

A follow-up online course is dedicated to applying structured expert judgment. Learners have the opportunity to perform their own study, under close guidance from Tina Nane and Roger Cooke. The learners will design the study and gather, assess the performance of,

and combine, expert opinion. They will also carry out the analysis and report the findings of their study.

EVALUATION OF THE CLASSICAL MODEL
Cooke has collected expert data from many studies over the years to evaluate the Classical Model and compare it with other possible weighting schemes, including equal weighting and quantile aggregation. The Classical Model outperformed the other aggregation methods considered in the analysis. The findings demonstrate the superiority of the Classical Model, both in terms of in-sample and out-of-sample validation and in terms of point forecast accuracy. Moreover, when compared with other aggregation methods, the performance-based combination of experts generated by the Classical Model is more statistically accurate and more informative.

Behind the Research

Dr Tina Nane

Professor Roger Cooke

E: g.f.nane@tudelft.nl **W:** www.tinanane.com **E:** cooke@rff.org **W:** rogermcooke.net

Research Objectives

Tina Nane has designed the Massive Open Online Course (MOOC) 'Decision making under uncertainty: an introduction to structured expert judgment', based on Roger Cooke's Classical Model.

Bio

Tina Nane is an Assistant Professor of Applied Probability at Delft University of Technology. She received her PhD in Statistics in 2013 and a cum laude MSc degree in Risk and Environmental Modelling in 2008 from Delft University of Technology. Dr Nane's current research primarily focuses on uncertainty quantification and analysis, both data-driven and by employing expert opinion. Dr Nane has designed and is the main lecturer of the Massive Open Online Course (MOOC) 'Decision Making Under

Uncertainty: Introduction to Structured Expert Judgment'.
Roger Cooke received his bachelor's degree (1968, magna cum laude) and PhD (1975) from Yale University in mathematics and philosophy. From 1975–2005, he worked in the Netherlands as Assistant Professor in Logic and Philosophy of Science at the University of Amsterdam, later as Professor of Applied Decision Theory at the Delft University of Technology. In 2005, he moved back to the USA as senior fellow at Resources

for the Future. From 2006 to 2008, he supervised the development of non-parametric continuous-discrete Bayesian Belief Nets for the Dutch Ministry of Transport. In 2008, he was elected fellow of the Society for Risk Analysis. In 2010, Cooke was named lead author in the 5th assessment of the Intergovernmental Panel on Climate Change for the chapter on Risk and Uncertainty. In 2011, he received the Lifetime Distinguished Achievement Award from the Society for Risk Analysis.

References

Decision Making Under Uncertainty: Introduction to Structured Expert Judgment. Online Course (MOOC). online-learning.tudelft.nl/courses/decision-making-under-uncertainty-introduction-to-structured-expert-judgment [Accessed 23rd September 2021]

Beshearse, E, Bruce, BB, Nane, GF, Cooke, RM, Aspinall, W, Hald, T, et al, (2021) Attribution of Illnesses Transmitted by Food and Water to Comprehensive Transmission Pathways Using Structured Expert Judgment, United States. *Emerging Infectious Diseases*, 27(1), 182–195. doi.org/10.3201/eid2701.200316

Hanea, AM, Nane, GF, (2021) An In-Depth Perspective on the Classical Model. In: Hanea, AM, Nane, GF, Bedford, T, French, S, (eds) Expert Judgement in Risk and Decision Analysis. International Series in Operations Research & Management Science, 293, 225–256. Springer, Cham. doi.org/10.1007/978-3-030-46474-5_10

Cooke, R, (2020) The Science of Forecasting: The Special Sauce. Resources, [online]. www.resources.org/common-resources/science-forecasting-special-sauce/ [Accessed 23rd September 2021]

Cooke, RM, Marti, D, Mazzuchi, T, (2020) Expert forecasting with and without uncertainty quantification and weighting: what do the data say? *International Journal of Forecasting*, 37, 378–387. doi.org/10.1016/j.ijforecast.2020.06.007

Bamber, J, Oppenheimer, M, Kopp, R, Aspinall, W, Cooke, R, (2019) Ice sheet contributions to future sea-level rise from structured expert judgment. *PNAS*, 116(23), 11195–11200. doi.org/10.1073/pnas.1817205116

Cooke, R, (2019) The Iceman Cometh. Resources [online]. www.resources.org/common-resources/iceman-cometh [Accessed 23rd September 2021]

Cooke, R, (2017) Respectability Looming for Expert Judgment. Resources [online]. www.resources.org/common-resources/respectability-looming-for-expert-judgment [Accessed 23rd September 2021]

Cooke, RM, (2015) Messaging climate change uncertainty. *Nature Climate Change*, 5, 8–10. www.nature.com/nclimate/journal/v5/n1/full/nclimate2466.html

Personal response

What inspired you to include performance-based weights in the Classical Model?

❝ Expert subjective probabilities began appearing in technical risk analyses of nuclear power plants in the 1970s. The rigorous reporting in these early studies exposed very wide differences in experts' judgments and teed up issues of validating and synthesising expert judgment. Any new measurement device, such as Galileo's telescope, is first 'calibrated' by applying it to things we know before employing it to measure things we don't know. Expert judgment constitutes a new measuring device. Applying this simple idea led to treating experts as statistical hypotheses and validating these hypotheses against calibration variables from their field to which the true values were known. ❞

TUDelft