International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Providing scientific certainty in predictive decision support: the role of closed questions

Joseph H.A. Guillaume[1], Anthony J. Jakeman[1]

[1]*Integrated Catchment Assessment and Management (iCAM) Centre and National Centre for Groundwater Research and Training, The Fenner School of Environment and Society, The Australian National University (joseph.guillaume@anu.edu.au)*

Abstract: Uncertainty in scientific predictions of outcome of decisions is currently of substantial interest, with a common perception, at least in groundwater modelling, that some decision makers have unrealistic expectations of accuracy. A feeling shared by some practitioners is that with a proper stakeholder engagement process, uncertainty is no longer such a difficult issue. This paper argues that the key is precisely specifying purpose and expectations, and the definition of closed questions provides a useful approach to do so. A closed question, also known as a fixed-choice or pre-coded question, has a finite set of answers, e.g. yes/no or A, B, C. If an answer is possible, it will always be certain, but science may need to say "we do not know". The approach is conceptually simple, though not necessarily trivial in practice, often requiring iteration between interest groups and modellers. Closed questions provide an interface between scientists and stakeholders which has a number of advantages, illustrated using examples from the published literature. Defining closed questions demands that proper attention be paid to problem formulation, which has long been recognised as important in the treatment of messy problems so pervasive in environmental modelling. The stakeholder engagement process manages the uncertainty in the definition of the problem, freeing the scientific process to concentrate on uncertainty in the model. Closed questions specify the purpose and required precision of the prediction, allowing the use of simpler models and techniques. This helps address the challenges of defining context-dependent criteria for acceptable model quality and determining appropriate model complexity. By addressing values in a proper process involving both stakeholders and scientists, the role of the scientist in providing objective predictions is separated from their potentially more subjective contributions to the remainder of the process. The common use of scientific hypothesis testing shows that probabilistic uncertainty can also be represented, allowing stakeholders to explicitly accept risk. As a result, a well-run, engaging process using closed questions allows science to promptly answer stakeholder questions with certainty, or recognize that an answer is not possible with the allocated resources, and therefore that the problem must be treated as trans-scientific.

***Keywords**:* uncertainty; stakeholder engagement; closed questions; first order logic; falsification.

1 INTRODUCTION

Increasingly science is expected to support decisions by providing urgent answers to complex, uncertain questions. Typical complaints are that science takes too long, or provides unreliable answers that turn out to contradict stakeholders' experiences. The resulting stakeholder disappointment has even led to the comment by Voss [2011] that "groundwater modelling has become a self-supporting industry of fantastical promises that cannot be kept". One response has suggested that expectations need to be lowered. The issue is that models can seemingly predict anything, to arbitrary precision. Calculating uncertainty bounds around point estimates is proposed as an obvious solution – except that, ironically, it implies that we can also quantify the uncertainty in our estimates – *what we do not know* – to arbitrary precision. In both cases, the worry is that stakeholders will consider the model to be unreliable, destroying its utility and credibility. In other words, the aim is to provide certainty to the user (albeit conditional on current understanding). While this is obviously impossible if we are to state the value of a variable to arbitrary precision, fortunately decision makers can usually make do with far vaguer statements. The key is to specify expectations as precisely as possible. The concept is clearly not new:

"Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise." Tukey [1962]

Stakeholders must necessarily work together to define the right question, and delineate how approximate the answer can be, and still be useful. Scientists must define how vague the question can be, and still be studied. Both require certainty – of expectations for a given question, and of reliability of the answer (contingent on current understanding). This paper proposes that the use of closed questions provides an intuitive, natural language approach that satisfies these requirements. By better defining the interaction between stakeholders and scientists, it provides a useful tool for the integration of modelling and stakeholder participation, particularly for ill-structured problems, as are common in environmental management.

Rather than discussing details of methods, this paper focuses on the method-independent benefits, illustrated by simple examples. The concept of a closed question is introduced (Section 2) as it applies to integrated modelling and stakeholders. A poorly defined question can cause difficulties for scientists and result in disappointment for stakeholders. Closed questions emphasise the importance of this issue (2.1), though they are not a solution. Once defined, they do help reduce the difficulty stakeholders have in specifying in scientific terms what makes a model's performance fit for purpose (2.2) and assessing whether it is sufficiently simple or complex (2.3). Discussion of closed questions forces the use of black and white terms, rather than the shades of grey inherent in the way model outputs and probabilities are used. While this is not always appropriate, in the context of prediction, it emphasises the importance of defining the limits of science's capabilities – of being able to say "we don't know" (2.4). If the shades of grey are really necessary, the risks involved must be explicitly accepted within the closed question (2.5).

2 COMMUNICATION OF PREDICTIONS

Any interaction between stakeholders and science should have a clear purpose. Predictions are logical assertions. They make claims about what is true. The required truth claim should therefore be clear before modelling starts. This sets stakeholders expectations, and allows scientists to choose the best available method to evaluate the required truth claim. Of course this can be an iterative process where a purpose is refined in response to the scientist's conclusion that methods cannot provide a useful outcome.

If we have a truth claim such as "this plan will meet its objectives", the easiest way to turn it into a question is to ask - "will this plan meet its objectives?" A question such as "Option A is better than Option B" can however also be phrased as "which option is better – A or B?" The natural way of dealing with predictions as logical assertions is to formulate them as questions with a finite number of *a priori* stated answers, that is, closed questions. Closed-ended questions are well-established in the design of surveys and questionnaires where they are also known as fixed-choice or pre-coded questions (Schuman and Presser [1979]). This is in opposition to the vague, open questions "how does this plan perform?" and "how good is option A?" This simple observation motivates our main argument - for predictions intended to support decisions, science questions must be formulated as unambiguously-phrased closed questions.

It is obvious that yes/no answers are equivalent to testing whether the assertion is true. Where the closed question specifies a choice between options, it translates to an optimisation problem. Uncertainty is, however, omnipresent. We must ask - is there a clear answer, or is there too much uncertainty? And so, instead of yes/no, the answer may be "we cannot give a definite answer", and we might have to say that option B *could be* equally as good as option A. In the latter case, we therefore always deal with optimisation under uncertainty. If it turns out that the question cannot be conclusively answered with current knowledge, science may still be able to contribute. The question can be modified by accepting risk (see Section 2.5), or by allocating additional resources to allow science to reduce the uncertainty. In this way, uncertainty is inherently addressed in the process of defining the right question, and the production of a certain yet approximate answer.

Closed questions are proposed exclusively as a replacement for asking for predictions of arbitrary accuracy. Dialogue and lateral thinking are better encouraged with open-ended questions, or less structured discussions. Closed questions are more suitable in a setting where stakeholders seek confirmation, not in an exploratory context. Often numerical models are used to create predictions when qualitative conceptual models would suffice. If a closed question cannot be formulated, such as in an exploration exercise, a first step should therefore be to check whether a numerical prediction is actually necessary. We therefore expect the approach to be applicable in most cases where predictions are used.

2.1 Uncertainty in defining the closed question

Defining a closed question is not a trivial exercise. The importance of uncertainty in defining the question to be asked of a model is commonly recognised, but frequently only paid lip service. It is also known as Type III error in statistics, described by Kimball [1957] as "giving the right answer to the wrong problem", context location of uncertainty in Walker et al. [2003], scoping and framing as in Guillaume et al [2010], or problem formulation in an optimisation context or decision theoretic context as in Guillaume & Pierce [2011].

Refsgaard et al. [2006] provided an invaluable example where they asked five consultants to model, "which parts of this area are most vulnerable to pollution and need to be protected?" While this is a closed question (choosing between specified parts), the result was that "the differences between the five estimates are striking and clearly do not provide a sound basis for deciding anything about which areas should be protected." Refsgaard et al. use this example to argue for the need to consider and evaluate multiple conceptual models. However, they also remark that the consultants "were known to have different views and preferences on which methodologies are most suitable for assessing vulnerability". It can also be concluded from their observations that this closed question is not sufficiently specified, and that value judgements about the meaning of "vulnerability" have a significant impact.

Uncertainty in definition of the problem, and uncertainty in modelling or the scientific process are often confounded, especially in wicked or messy problems (as defined by Rittel and Webber [1973]), with the result that value judgements that are best defined by stakeholders are left to the responsibility of the scientist. It is no surprise that stakeholders would end up rejecting a model or decision support system in these circumstances. In the operations research and soft systems literature, the issue is addressed through the use of problem structuring methods, as described by Rosenhead [1996] and Rosenhead & Mingers [2001]. Guillaume & Pierce [2011] suggest that some of these methods do not go far enough, as they fail to translate the resulting stakeholder view to a model question.

Defining closed questions forces modellers and stakeholders to go through a distinct phase of problem formulation. It impels stakeholders to identify why a particular assertion is necessary, and how it will be used. The role of the scientist in providing objective predictions is separated from their role in helping to define the question which typically requires a very different approach.

While the problem formulation phase is distinct, it should be seen as iterative, and never final. Stakeholder needs may evolve, modellers may discover that additional input is required, that the questions are intractable within the allocated resources or vice-versa. Answering a closed question may in fact lead to a change in the question, contributing to the process of retrieving elements to structure a problem, as described by Simon [1973]. In the meantime, scientists have clear requirements about what their model, or other work, needs to achieve. In the scientific context, it is better to have clear requirements that change than to have vague requirements or none at all.

2.2 Defining criteria for acceptable model performance

Refsgaard & Henriksen [2004] recognise that stakeholders will ultimately decide whether a model's performance is acceptable:

"The key responsibilities of the water resources manager are to specify the objectives and define the acceptance limits of accuracy performance criteria for the model application. Furthermore, it is the manager's responsibility to define requirements for code verification and model validation."

These responsibilities seemingly involve a high level of technical expertise in modelling – and some are even open research issues in the modelling literature. Water resource managers would be well justified in resisting this extra burden. Closed questions provide a simple approach to partly overcome this issue. It is generally accepted that the precision at which a result is interpreted should be no greater than its accuracy. Closed questions implicitly specify the precision required in order to be able to answer the question. It follows that the accuracy must be at least as good as this required precision. This only partially addresses the issue because the modeller still needs an accepted method for determining accuracy – or equivalently, uncertainty. The key value judgements are however addressed and closed questions can help determine whether the complexity of the system is sufficiently represented (Section 2.3). We leave the rest to future work.

By defining the required accuracy for a specific prediction, closed questions can also make seemingly unanswerable questions feasible. Reichert and Borsuk [2005] give an example where the reduction of phosphorous loading due to a management option is assessed. The probability distributions overlap such that it is seems possible that the management option could have no effect. However, they point out that "if the major management objective is to achieve a phosphorus load reduction to the lake in any particular year, then it is important to correctly calculate the distributions of *differences across alternatives*" (our emphasis). The distribution of differences shows that the management option will always achieve a reduction. Directly asking "is there a reduction in phosphorus loading in this scenario?" provided (conditional) certainty, while answering the question indirectly by first asking "what is the phosphorus loading in this scenario?" only resulted in "we cannot tell".

Closed questions about model performance can also add value for a manager's understanding of the crucial strengths and limitations of a model. For example, stakeholders can start by asking "under what conditions are the model's predictions able to differentiate between specified options?" and "under what conditions" is the model definitely incapable of such differentiation?" In some cases these questions might require more closure. For instance one might pose "under what climate conditions (and/or land use change or demographic change or international policy changes) is the model definitely not able to differentiate between options?" Rarely does one see such delineation of conditions reported from a modelling exercise.

2.3 How complex does a model need to be?

How simple or complex a model should be is ongoing matter of debate in the modelling community, and one which influences interaction with stakeholders due to its effect on run-times, model development times and hence the time scales of stakeholder engagement processes. Both sides in the simplicity-complexity debate seem to agree on the quote attributed to Einstein, that "a model should be as simple as possible but no simpler". The point of contention is what constitutes "as simple as possible". The need for a purpose is implied, and closed questions provide a clear answer in the case of prediction. The model must be sufficiently complex so as to capture the necessary processes to answer the closed question, and verify that it is answered robustly – that no uncertainties have been missed. From this perspective, simple models may be argued to be insufficient, but the burden of proof is on complex models or techniques to show that they make an additional contribution.

The ability to capture additional processes does not necessarily require more parameters, though complex models often do. Jakeman and Hornberger [1993] present a spatially lumped rainfall-runoff model in terms of transfer functions. While distributed models describe flow spatially and temporally, "in physical terms, the transfer function representation corresponds to a flexible configuration of linear storages connected in parallel and/or series paths for the transit of excess rainfall to the stream." A much simpler model can capture the same processes by using emergent properties of flow. It does not differentiate between these processes, so it has a restricted domain of applicability – it is only appropriate for answering some questions. It can answer the question "will there be enough stream-flow for irrigation?" If there is enough historical data, it may be able to answer how flow varies with climate. It cannot deal with change of land use as this would alter the emergent properties of flow. It cannot answer the question "Will there be enough stream-flow for irrigation after the catchment is re-forested?" The precisely defined aim identified using the closed question allows the modeller to determine whether this simple model is sufficient.

2.4 Defining the limits of science using closed questions

By making the required assertion clear, closed questions clarify what is required of science, and what will be handled by other means. This is crucial with so-called trans-scientific questions. Weinberg [1972] defines trans-scientific questions as those "which can be asked of science and yet which cannot be answered by science". This includes the case where decisions are made on the basis of incomplete data, such as in engineering, and where values influence the question, as in public policy. In that context, a scientist has a responsibility "to make clear where science ends and trans-science begins".

Due to the nature of scientific uncertainty, the scientist may not be able to give a definite answer, particularly with limited resources. Consider the question: if the general population is exposed to more than 170 millirems of radiation per year, from a sample of 1000, will the number of cases of leukaemia and cancer increase by more than 5? Weinberg [1972] and Holcomb [1970] observe that at the time, science could not give a definite answer – maybe yes, maybe no. Experimentation was difficult (and on humans, forbidden), and assumptions were difficult to test. However, Weinberg's analysis [1972] was that "neither side was willing to say that the question was simply unresolvable". Honesty keeps expectations of science realistic, and directs the end-user to use methods that can handle indeterminate facts. It may be that this provides a strong argument for allocating additional resources to the scientific challenge. On the other hand, when a clear answer can be provided, this allows the end-user to claim scientific uncertainty (barring surprises in new data).

It is hoped that the process of asking and answering a closed question shares the advantages assigned by Weinberg [1972] to legal adversary procedures, in "forcing scientists to be more honest, to say where science ends and trans-science begins, as well as to help weigh the ethical issues which underlie whatever choices the society makes between technological alternatives".

2.5 Accepting risk

The desire to express a probabilistic result or degree of confidence is a common objection to the certainty embodied in closed questions. The widespread adoption of hypothesis testing in science shows how this supposed obstacle is easily overcome.

Hypothesis testing replaces certainty on a case-by-case basis for certainty in the acceptance of risk. In the theory of statistical inference formulated by Neyman and Pearson [1933], certainty is provided about the long-run frequency of errors: "Without hoping to know whether each separate hypothesis is true or false, we may search for rules to govern our behaviour with regard to them, in following which we insure that, in the long run of experience, we shall not be too often wrong".

As described by Gigerenzer et al. [1990], consider a plant that produces widgets. During testing, if the size of the widgets is too small, the plant must be stopped. The question is asked "Should the plant be stopped?" With the normal variation in widget size, it is impossible to give a definite answer, so instead the question asked is "We accept that the plant will be stopped in error up to one day a month. Accepting this risk, should the plant be stopped?" Neyman and Pearson [1933] do not advocate a standard 95% confidence level, and do not jump to conclusions by interpreting 95% confidence as certainty. They explicitly accept the risk of type I and II errors, of false acceptance and false rejection of the hypothesis. In their words: "in some cases it will be more important to avoid the first, in others the second. ... The use of these statistical tools in any given case, in determining just how the balance should be struck, must be left to the investigator".

Long-run frequency of errors is admittedly not suitable for many environmental decisions and problems. Anderson [1998] however recommends the presentation of probabilities using a frequency format, as is intuitively done with long-run errors. Cognitive errors occur due to the activation of different probability-related concepts. The frequency format allows people to intuitively use reasoning compatible with the formal rules of probability. Anderson gives an example using open questions:

Probability format: "What is the probability that this watershed, for which you have planned a forest ecosystem network, will retain all 14 species of passerines 40 years from now?"

Frequency format: "Picture 100 watersheds like this one. For each, you have planned a similar forest ecosystem network. How many will retain all 14 species of passerines 40 years from now?"

With similar wording to the widget example, we can easily formulate a closed question: "We accept that in 5 out of 100 watersheds like this one, we would lose a species. With this exception, will all species be retained 40 years from now?"

As Lehmann [1993] describes, the Bayesian view and the statistician R.A. Fisher take a contrasting evidentiary view, stating probability rather than using an acceptance/rejection procedure. The scientists 'gains a better understanding' by using the degree of confidence as a 'means of learning'. This is clearly a different purpose to prediction, and when it comes to making decisions, even Fisher himself claimed that a research worker is interested in answering the (closed!) questions: "ought I take notice of that?" and "Is this particular hypothesis overthrown, and if so at what level of significance, by this particular body of observations?".

The use of standard significance levels is now common practice in science, but Neyman and Pearson's [1933] behavioural approach seems more appropriate in the case of environmental prediction. Whereas decisions in scientific hypothesis testing typically have intangible consequences on accumulated knowledge, decisions resulting from a prediction can have significant direct impact on society at large. A 95% confidence level on river levels would imply that society accepts that flooding occurs on 5 out of 100 days. It is essential that stakeholders explicitly and consciously accept risk. Scientists must not implicitly accept risk on their behalf.

3 CONCLUSIONS

Closed questions provide an intuitive, natural language means of structuring interaction regarding predictions between science and stakeholders. They precisely define the purpose of scientific input with a number of advantages. They are easily translated to the mathematical concepts of testing of assertions and optimisation under uncertainty. They demand that proper attention be paid to problem formulation. They help modellers define acceptable model performance, and can help simplify problems. They acknowledge that it is acceptable for science not to know, and demand that risk be explicitly accepted by stakeholders. The result is that uncertainty in predictions is handled intuitively, and the expectations of stakeholders are clear as to what aspect of a prediction can be interpreted with certainty. This paper has not discussed how stakeholder groups should formulate closed questions, or how science should provide answers – there remain substantial issues in practice, but at a methodological level, the concept is so simple, that even a sceptic can give it a chance. Next time you need a prediction, consider: will it help to phrase it as a closed question?

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of the National Centre for Groundwater Research and Training (NCGRT), and stimulating discussions at Modflow & More 2011, and with a number of people including Alexey Voinov, Rachel Blakers and John Doherty.

REFERENCES

Anderson, J.L. Embracing uncertainty: The interface of Bayesian statistics and cognitive psychology. Conservation Ecology [online] 2(1): 2. Available from the Internet. URL: http://www.consecol.org/vol2/iss1/art2/, 1998.

Gigerenzer, G., Z. Swijtink and L. Daston (1990). The empire of chance: How probability changed science and everyday life, Cambridge Univ Pr.

Guillaume J.H.A., Pierce S.A. Decision problem formulation for Groundwater Allocation Planning, 2011 GSA Annual Meeting in Minneapolis, 9–12 October 2011

Guillaume, J.H.A., S.A. Pierce and A.J. Jakeman. Managing uncertainty in determining sustainable aquifer yield. Groundwater 2010. Canberra, Australia, 2010

Holcomb, R.W. Radiation Risk: A Scientific Problem? Science 167(3919): 853-855, 1970, DOI: 10.1126/science.167.3919.853

Jakeman, A.J. and G.M. Hornberger. How much complexity is warranted in a rainfall-runoff model? Water Resour. Res. 29(8): 2637-2649, 1993, DOI: 10.1029/93wr00877.

Kimball, A.W. Errors of the Third Kind in Statistical Consulting. Journal of the American Statistical Association 52(278): 133-142, 1957

Lehmann, E.L. The Fisher, Neyman-Pearson Theories of Testing Hypotheses: One Theory or Two? Journal of the American Statistical Association 88(424): 1242-1249, 1993

Neyman, J. and E.S. Pearson. On the Problem of the Most Efficient Tests of Statistical Hypotheses. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character 231: 289-337, 1933

Refsgaard, J. C. and H. J. Henriksen. "Modelling guidelines--terminology and guiding principles." Advances in Water Resources 27(1): 71-82, 2004

Refsgaard, J. C., van der Sluijs, J. P. et al. "A framework for dealing with uncertainty due to model structure error." Advances in Water Resources 29(11): 1586-1597, 2006

Reichert, P. and M.E. Borsuk. Does high forecast uncertainty preclude effective decision support? Environmental Modelling & Software 20(8): 991-1001, 2005, DOI: 10.1016/j.envsoft.2004.10.005

Rittel, H.W.J. and M.M. Webber. Dilemmas in a general theory of planning, Policy Sciences 4(2): 155-169, 1973, DOI: 10.1007/bf01405730

Rosenhead, J. What's the Problem? An Introduction to Problem Structuring Methods. Interfaces 26(6): 117-131, 1996.

Rosenhead, J. and J. Minger. Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, 2001.

Simon, H.A. The structure of ill structured problems, Artificial Intelligence, 4(3-4): 181-201, 1973, DOI: 10.1016/0004-3702(73)90011-8

Schuman, H. and S. Presser. The Open and Closed Question. American Sociological Review 44(5): 692-712, 1979, from http://www.jstor.org/stable/2094521

Tukey, J.W. The future of data analysis. Annals of Mathematical Statistics 33 (1), 1962

Voss, C. Editor’s message: Groundwater modeling fantasies: part 2, down to earth. Hydrogeology Journal: 1-4, 2011, DOI: 10.1007/s10040-011-0790-6

Walker, W.E., P. Harremoës, J. Rotmans, J.P. van der Sluijs, M.B.A. van Asselt, P. Janssen and M.P.K. von Krauss. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment 4(1): 5 – 17, 2003 DOI: 10.1076/iaij.4.1.5.16466

Weinberg, A. M. "Science and trans-science." Minerva 10(2): 209-222, 1972