Environmental Modelling & Software 70 (2015) 97–112

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Prediction under uncertainty as a boundary problem: A general formulation using Iterative Closed Question Modelling

Joseph H.A. Guillaume [a, b, *], Matti Kummu [b], Timo A. Räsänen [b], Anthony J. Jakeman [a]

[a] National Centre for Groundwater Research and Training & Integrated Catchment Assessment and Management Centre, The Fenner School of Environment and Society, The Australian National University, Building 48A, Linnaeus Way, Canberra, ACT 0200, Australia
[b] Water & Development Research Group (WDRG), Aalto University, Tietotie 1E, 02150 Espoo, Finland

ARTICLE INFO

Article history:
Received 3 February 2014
Received in revised form
24 July 2014
Accepted 10 April 2015
Available online 15 May 2015

Keywords:
Uncertainty analysis
Iterative Closed Question Modelling
Boundary critique
Process philosophy

ABSTRACT

Making predictions about environmental systems is a challenge due to the high level of uncertainty involved. In this paper we give a general formulation of prediction under uncertainty as a boundary problem. This leads to development of a methodology for making predictions under uncertainty, named Iterative Closed Question Modelling (ICQM). ICQM involves iteratively devising questions and testing the certainty of their answers by creating complete model scenarios (complete taken to include structure, parameters and inputs for each scenario instance). The model scenarios are categorised in terms of which answer they support, and whether they are plausible or not. Using a simple two-parameter flow duration curve model, the paper demonstrates the application of ICQM using eight alternative uncertainty analysis techniques. ICQM provides a useful and generic approach to making predictions under uncertainty, helps to understand how existing techniques address the boundary problem differently and promotes the development of new techniques.

1. Introduction

Computational models are widely used for predicting consequences of changes to environmental systems. Our understanding of these systems is however unavoidably uncertain, as environmental models cannot be validated (Oreskes et al., 1994). As modellers, we must therefore live with the fact that all models are wrong (see also Sterman, 2002), and that we are limited to making predictions *under uncertainty*.

The study of uncertainty analysis has made significant advances by treating uncertainty in predictions as a quantification problem. Estimation of parameters and their uncertainty is posed as a statistical problem (e.g. Yeh, 1986; Schoups and Vrugt, 2010), which can be extended to inputs (e.g. Vrugt et al., 2009) and model structures (e.g. Hoeting et al., 1999; Thyer et al., 2009). The field of uncertainty quantification has developed efficient techniques for then determining the effect of input uncertainties on response metrics of interest (Eldred et al., 2011).

While quantitative techniques are essential in the computation of most environmental models, they often can be seen as ultimately resulting in a qualitative answer or a discrete decision based on some sharp (but possibly uncertain) boundary. This is particular apparent in the phrasing of questions asked of a model (Guillaume and Jakeman, 2012). In quantitative modelling, there is a tendency to ask for numbers, such as "what is the concentration of this pollutant?" or "what is the probability of a flood occurring?", in which case uncertainty is approached with a similar intent to measurement techniques by expressing confidence intervals that define the boundary of what is considered likely. Additionally, such questions can often be expressed with more explicit boundaries, usually corresponding to a less precise but more practical question such as: "Is the concentration below the acceptable limit?"; "Which flood management option has the greatest expected utility: A, B, or C?"; or "Will the flood occur with odds greater than 5 times in 100 years?". Uncertainty analysis could therefore be seen and treated as a *boundary problem*.

This paper shows how existing uncertainty analysis techniques can be placed in this broader context of a boundary problem. A boundary problem involves analysis and critique of boundaries and how they interact (Midgley, 2000). A problem formulation is given that defines prediction in terms of *normative* and *epistemic*

* Corresponding author. Water & Development Research Group (WDRG), Aalto University, Tietotie 1E, 02150 Espoo, Finland. Tel.: +358 504072906.
E-mail address: joseph.guillaume@aalto.fi (J.H.A. Guillaume).

http://dx.doi.org/10.1016/j.envsoft.2015.04.004
1364-8152/

boundaries, respectively corresponding to a question of interest to the model user and agreed assumptions about what model scenarios are considered plausible. Prediction is seen as a process that uses model scenarios to ascertain which of a closed question's predetermined answers are plausible. Prediction under uncertainty is therefore simply a general case where many models are considered to be plausible. A methodology, named Iterative Closed Question Modelling (ICQM), is presented on the basis of this problem formulation, and it is illustrated using a broad range of existing and emerging uncertainty analysis techniques.

Placing prediction under uncertainty in this broader boundary problem context has three key advantages. Firstly, it provides an understanding of how existing techniques relate to each other. Secondly, it promotes the development of new techniques that achieve the same end goal through different means, perhaps supported by alternative assumptions. Thirdly, and perhaps most importantly, the resulting ICQM methodology is generic enough that it can act as a first port of call, providing a structure that modellers can invoke in an initial approach to make predictions under uncertainty. It provides high level guidance while leaving sufficient flexibility for the methodology to be broadly applicable and able to be adapted to different contexts, including different types of uncertainty.

We present a general formulation of the boundary problem approach to uncertainty, but it has also been implicitly advocated by other authors. For example, Section 4 presents a number of major uncertainty analysis techniques that can be expressed in terms of boundaries, most obviously set membership, which identifies feasible sets of parameters (Walter and Piet-Lahanier, 1990; Norton, 1996; Keesman et al., 2013). Beven (2006) appealed to the notion of boundaries in arguing for an approach where “models that do not fall within the multiple prior limits of acceptability should be rejected.” Schwanitz (2013) argues for testing whether models satisfy stylised facts about their expected behaviour. Haasnoot et al. (2014) use a set of closed questions to test model performance.

More generally, our approach inherits from the systems thinking literature on boundary judgements and boundary critique. Churchman (1970) argued for broadening the boundaries of an analysis in order to challenge our “cherished assumptions”. Ulrich (1983) introduced a set of questions to structure the critique of boundaries. Midgley (2000) introduced a form of process philosophy that avoids a number of philosophical problems about how knowledge is generated by focussing on boundary judgements as the process of bringing knowledge into being.

This work therefore contributes to the harmonisation of existing quantitative uncertainty analysis techniques and the ideas of boundary critique and process philosophy. The work is presented in this paper as follows: Section 2 introduces a problem formulation of prediction under uncertainty as a boundary problem, and hence derives the ICQM methodology; Section 3 synthesises a typology of sources and levels of uncertainty which helps to understand the variety of alternative techniques to implement ICQM; Section 4 lists a selection of specific techniques to implement ICQM; Section 5 presents a demonstration problem to provide a more practical example of how ICQM is implemented; Section 6 presents the results from the demonstration problem; Section 7 is the discussion section and it reflects on the effectiveness of ICQM as an initial approach to elucidating and characterising uncertainty; Section 8 forms the conclusions.

2. Prediction under uncertainty: problem formulation and ICQM methodology

This section sets the context of prediction under uncertainty as the identification and exploration of model scenarios (Section 2.1). It then presents a problem formulation of prediction as a boundary problem (Section 2.2), allowing the expression of uncertainty analysis within that boundary problem (Section 2.3). We emphasise the need for explicit iteration in solving this boundary problem (Section 2.4). The problem formulation can be operationalised nearly directly as the proposed ICQM methodology (Section 2.5), which is applied later in this paper.

2.1. Context: identification and exploration of model scenarios

We start with the fundamental assumption that we are working in a modelling context. In particular, the focus of the analysis is the identification and exploration of *model scenarios*. We define a model scenario as an explicit representation of knowledge including a model structure, parameter values and input values. The model scenario might be represented in different forms, including conceptual diagrams, stories or computational models. It may also have different interpretations. It may for example be intended as a representation of reality, a representation of a stakeholder's view of a system, or may even be entirely fictional with no relation to reality or stakeholder if the modeller so desired.

The identification of model scenarios involves exploring the space of all possible model scenarios, which we will call *model scenario space*. This space is represented in a set diagram in Fig. 1a and b. Any point in this space (e.g. those shown as stars in Fig. 1) is a complete model scenario instance, consisting of model structure, parameter and input values. We denote a model scenario with the symbol y and model scenario space with the symbol Y, such that:

$$y \in Y \qquad (1)$$

(literally, model scenarios are elements of the model scenario space)

2.2. Prediction as a boundary problem

At the most general level, prediction involves the definition of two boundaries in model scenario space:

1. A *Normative Boundary*, which refers to a closed question of interest. A closed question has pre-determined answers, such as the categorical options A, B, C etc. or the even simpler yes/no/do not know. These answers can be used to partition the model scenario space into subsets of models with different conclusions. The boundaries defined by the closed questions “Will objectives be met?” and “Which solution is best?” respectively partition the space into subsets where objectives are and are not met (Fig. 1a and c) and into subsets where different options are best (Fig. 1b and d). The adjective ‘normative’ means it is related to or derived from a standard or norm, i.e. this boundary has meaning because a model user believes it *ought* to have meaning rather than through any intrinsic meaning of its own.
2. An *Epistemic boundary*, which refers to agreed assumptions about the system. The model scenario space can be partitioned into subsets that do and do not meet the agreed assumptions and therefore are and are not considered plausible (dashed line in Fig. 1a and b. The adjective ‘epistemic’ means that the boundary is related to knowledge, hypothesised, elicited or otherwise.

In this formulation, a prediction is seen as the use of an epistemic boundary to draw a conclusion regarding a normative boundary, or equivalently, the use of agreed assumptions to draw a conclusion about the answer to a closed question. In a modelling context, this is traditionally achieved by identifying a model scenario that is considered plausible (i.e. satisfies agreed assumptions