INFLUENCE DIAGRAMS FOR REPRESENTING UNCERTAINTY IN CLIMATE-RELATED PROPOSITIONS

JIM HALL[1], CRAIG TWYMAN[1] and ALISON KAY[2]

[1]*Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, UK*
E-mail: jim.hall@bristol.ac.uk
[2]*Centre for Ecology and Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK*

Abstract. In order to respond to policy questions about the potential impacts of climate change it is usually necessary to assemble large quantities of evidence from a variety of sources. Influence diagrams provide a formal mechanism for structuring this evidence and representing its relationship with the climate-related question of interest. When populated with probabilistic measures of belief an influence diagram provides a graphical representation of uncertainty, which can help to synthesize complex and contentious arguments into a relatively simple, yet evidence-based, graphical output.

Following unusually damaging floods in October–November 2000 the UK government commissioned research with a view to establishing the extent to which the floods were a manifestation of hydrological climate change. By way of example application, influence diagrams have been used to represent the evidential reasoning and uncertainties in responding to this question. Three alternative approaches to the mathematization of uncertainty in influence diagrams are demonstrated and compared. In situations of information scarcity and imprecise expert judgements, methods based on interval probabilities have proved to be attractive. Interval probabilities can, it is argued, represent ambiguity and ignorance in a more satisfactory manner than the conventional Bayesian alternative. The analysis provides a quantified commentary on the uncertainties in the conclusion that the events of October–November 2000 were extreme, but cannot in themselves be attributed to climate change.

1. Introduction

Analysis of the impacts of climate change makes use of knowledge from diverse disciplines. This knowledge appears in a range of formats, from measurements and model predictions to expert judgements expressed in linguistic terms. Evidence may appear as time series of historic data or ensemble predictions from climate models (which lend themselves to probabilistic analysis) but also as linguistic reasoning about causal relationships, analogues, patterns or fragments of partially relevant data. One way of bringing these various sources of knowledge together is in an influence diagram, which represents the relationships and interactions between a series of propositions or processes. These relationships may be described in exact algebraic terms but may also represent less precise understanding about the mechanisms of influence. Van Lenthe et al. (1997) demonstrate how influence diagrams can be used to structure complex climate-related policy decisions, though they confine their analysis to the situation where it is possible to elicit precise probability distributions for all of the required variables in the influence diagram.

Climatic Change (2005) 69: 343–365

Kuikka and Varis (1997) used Bayesian belief networks to structure reasoning about climate variables in the climate–water system in Finland.

Influence diagrams can communicate knowledge in a formal but accessible manner. The process of constructing, populating and testing an influence diagram generates valuable insights. It provides a mechanism for decomposing complex problems or systems into recognisable sub-systems that are often easier to reason about than the system as a whole. An influence diagram can help to externalise expert judgement and hence facilitates dialogue between experts and other decision stakeholders. Whilst the influence diagrams discussed in this paper have been populated with probability estimates from experts, the approach is equally applicable to joint probability distributions estimated from data. In general the populated diagram represents a joint probability distribution over the variables or propositions in the diagram. Influence diagrams have in recent years become widely used as a structure for Bayesian updating of probabilities in multivariate statistical problems (Gilks, 1996).

The objective of this paper is to illustrate how influence diagrams can be used to analyse uncertainty in a climate-related proposition whose analysis involves making use of evidence from diverse sources, including expert judgement, in a variety of formats. The work is related to the well-known problem of attributing observed climate to anthropogenic forcing (Hegerl et al., 1997; Risbey et al., 2000; Stott et al., 2001; Risbey and Kandlikar, 2002), but deals with a specific weather event in a situation where the limited available evidence is heterogeneous in format and subject to varying degrees of epistemic uncertainty. In particular we seek to illustrate how simple methods based on interval probabilities can provide a more expressive mathematical vocabulary of uncertainty than the conventional alternative. The analysis is based upon a study commissioned by the UK Government that sought to establish the extent to which very severe floods in the UK in October–November 2000 were attributable to climate change. The paper begins with a brief review of the necessary concepts of influence diagrams. Three alternative approaches to handling uncertainty in influence diagrams are introduced and discussed. In Section 3 each of these three methods is applied to the UK floods example. Conclusions are drawn at two levels, first in relation to UK floods example and second in relation to the extent to which the three alternative methods are appropriate for handling uncertainty in influence diagrams applied to climate problems.

2. Influence Diagrams

An influence diagram (Howard and Matheson, 1981; Shachter, 1988; Pearl, 1988; Oliver and Smith, 1990; Gammerman, 1995) is a directed acyclic graph of nodes and links. An example of an influence diagram, which originates from one of the early publications on Bayesian belief networks (Spiegelhalter, 1986) and has now become something of an archetype, is illustrated in Figure 1. It shows the

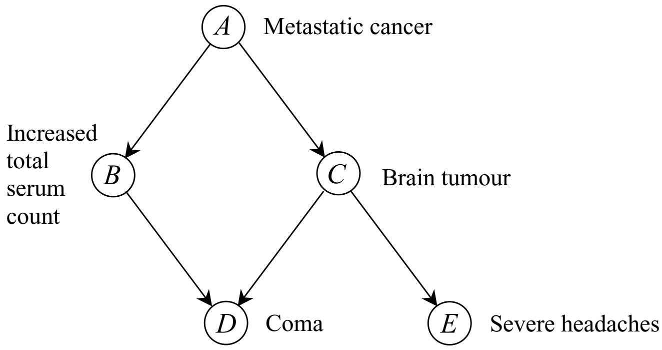

Figure 1. Typical influence diagram (after Henrion, 1988).

relationship between observable symptoms and a medical diagnosis (metastatic cancer or brain tumour). Each node in the diagram represents a proposition and each link represents a potential influence. In this paper we restrict ourselves to the situation in which the proposition can take one of only two states, say A and its negation $\neg A$. $P(A)$ is a probability that represents the degree of belief in A. The links signify the existence of direct influences on a proposition by its immediate predecessors (its *parents*). A conditional probability represents the strength of this influence. Absence of a unidirectional connection between two nodes by one or more links indicates absence of influence.

In the network in Figure 1, A is the parent of B and C, which are the parents of D. B and C are referred to as the *children* of A. C is also the parent of E. If there is only one node in the graph that does not have a parent, then the network is a hierarchy. If, furthermore, there are no closed loops in the graph, then the network is a tree.

Proposition A is referred to as the *source* proposition, and propositions D and E are referred to as the *sink* propositions. *Predictive* or *causal* inference involves reasoning from evidence about source propositions down through the network in the direction of links to sink propositions. *Diagnostic* inference involves reasoning in the reverse direction, from observations of manifestations to infer probabilities of possible causes (Henrion, 1988). In this paper we are concerned with this latter, diagnostic mode of reasoning, i.e. to what extent can it be believed that the October–November 2000 floods were a manifestation of climate change, given the available measured and modelled evidence.

It is assumed that all influence is captured in conditional probability relations between a child proposition and its parents. In other words, the child is conditionally independent of all other propositions in the network apart from its parents. If this is the case then, for example for proposition E, (see Figure 1)

$$P(E) = P(E|C)P(C) + P(E|\neg C)P(\neg C).$$

The relationship between C and E is a feature of the structure of the inference problem. For example C may be a *necessary* condition for E, in which case $P(E|\neg C) = 0$ and $P(E|C) \leq 1$, or C may be a *sufficient* condition for E, in which case $P(E|C) = 1$ and $P(E|\neg C) \leq 1$. Probabilities can be propagated in either direction in the network, provided appropriate conditional probabilities are available, so

$$P(C) = P(C|E)\,P(E) + P(C|\neg E)\,P(\neg E).$$

For more than one parent/child the following expansion, referred to as the Bayes conditionalization formula, is required, so for example in Figure 1:

$$\begin{aligned} P(D) = {} & P(D|B \wedge C)\,P(B \wedge C) + P(D|B \wedge \neg C)\,P(B \wedge \neg C) \\ & + P(D|\neg B \wedge C)\,P(\neg B \wedge C) + P(D|\neg B \wedge \neg C)\,P(\neg B \wedge \neg C) \end{aligned} \tag{1}$$

where the symbol $\wedge$ represents the conjunction of two propositions i.e. the *and* operator. The disjunction (i.e. the *or* operator) is written $\vee$. The probabilities of the conjunctions $P(B \wedge C)$, $P(B \wedge \neg C)$, $P(\neg B \wedge C)$, $P(\neg B \wedge \neg C)$, depend on the relationship between $P(B)$ and $P(C)$. If it is assumed that B and C are independent then

$$P(B \wedge C) = P(B) \cdot P(C)$$

and so on. The rationale for the independence assumption is that any dependency relationship will be represented by links to a common predecessor elsewhere in the network. The assumption of conditional independence is not made in Interval Probability Theory of Cui and Blockley (1990), which is discussed below. First we address Bayesian belief networks, where the probability $P(A)$ of a proposition A is a point value on the range [0, 1]. We then extend to consider two interval probability approaches, in which the uncertainty in $P(A)$ is represented as an interval on [0, 1].

The extension of classical point probabilities to interval probabilities holds the potential for bridging the gap in the debate about whether climate change predictions should be expressed in probabilistic terms. On the one hand it is argued that probabilities are essential to make rational decisions under conditions of uncertainty, and if probabilities are not provided then decisions will be made with implied assessments of relative likelihoods that depart, perhaps significantly, from experts' best estimates (Schneider, 2001, 2002; Pittock et al., 2001). On the other hand, it is argued that aspects of climate uncertainty, in particular uncertainties in emissions scenarios, do not lend themselves to quantification and there is no justifiable basis for constructing probability distributions (Grüber and Nakicenovic, 2001). A further argument against probabilistic quantification of uncertainty as it is currently conducted or proposed (for example Allen, 1999; Wigley and Raper, 2001), is that the uncertainties in climate model predictions due to incompleteness in the model representation of relevant processes may be under-estimated: ensemble predictions and inter-model comparisons capture only a fraction of the total model uncertainty and successful reproduction of historic climate provides only partially relevant evidence of predictive accuracy under significantly changed conditions (Young et al., 1996; Stocker and Schmittner, 1997; Shackley et al., 1998; Allen et al., 1998).

Imprecise probabilities (Walley, 1991), of which the interval probabilities presented in this paper are a special case, are able to represent *ambiguity*: the situation in which an agent is unable to distinguish (express preferences) between alternative probabilities or risks. This is achieved by the use of lower and upper bounds on an unknown probability. In the extreme case, imprecise probabilities can be used to represent complete ignorance by using vacuous probability bounds, without the customary recourse of the probabilistic analyst in the face of indeterminacy, which is to adopt a uniform probability distribution or some other arguably 'non-informative' prior. Keynes (1921) showed that incautious adoption of the uniform distribution can lead to contradictions. A uniform distribution (or indeed any other unique distribution) represents a precise statement about the relative likelihoods of different states, which in cases of legitimate indeterminacy will overstate the available knowledge.

It has been known, at least since the experiments of Ellsberg (1961), that people express preferences and behave in ways that show aversion to ambiguity. Ambiguity aversion cannot be accounted for in the conventional theory of choice under uncertainty (Savage, 1954) whilst a coherent explanation is provided by the generalisation to imprecise probabilities (Gilboa, 1987; Gilboa and Schmeidler, 1989; Schmeidler, 1989). These observations are particularly relevant in the context of climate problems because it can be argued that concepts such as the precautionary principle are motivated by the presence of ambiguity (see for example: Henry and Henry, 2002; Chevé and Congar, 2003).

2.1. BAYESIAN BELIEF NETWORKS

When the nodes and links in an influence diagram are populated with Bayesian probabilities, the influence diagram is known as a Bayesian belief network. Bayesian belief networks (Pearl, 1988; Oliver and Smith, 1990; Gammerman, 1995; Jensen, 1996) are the most widely used method for quantified uncertainty handling in influence diagrams, and several commercial software packages are available. In general Bayesian networks deal with the case in which a node can take a number of different states. Moreover, Bayesian belief networks can be structured as multiply connected directed acyclic graphs and support combinations of predictive and diagnostic inference in a general manner. Here we restrict ourselves to the case of diagnostic inference in tree structures in which each proposition has only two states: 'true' or 'false'.

The approach is illustrated with an example node from the influence diagram introduced in the following section. The node 'individual flow gauge readings were unusually high in 2000' (label H) has two children 'individual flow gauge readings rank highly among data series' (label E_1) and 'individual flow readings registered in 2000 were highly unlikely' (label E_2). Suppose that the following probabilities are set: $P(E_1) = 0.735$, $P(E_2) = 0.245$, so that $P(E_1 \wedge E_2) = 0.180$, $P(E_1 \wedge \neg E_2) = 0.555$, $P(\neg E_1 \cap E_2) = 0.065$, $P(\neg E_1 \wedge \neg E_2) = 0.200$. The conditional probabilities are set: $P(H|E_1 \wedge E_2) = 0.94$, $P(H|E_1 \wedge \neg E_2) = 0.80$, $P(H|\neg E_1 \wedge E_2) = 0.49$, $P(H|\neg E_1 \wedge \neg E_2) = 0.35$. From Equation (1), $P(H) = 0.715$.

2.2. INTERVAL PROBABILITIES

Two further approaches to quantified uncertainty handling in inference networks are now introduced, which both address the inevitable uncertainty in a probabilistic estimate $P(A)$, by dealing with it as an interval probability. Thus, whilst being founded on the axioms of probability theory, this interval probability approach allows support for a proposition to be separated from support for the negation of the proposition. If A is a proposition, an interval number is used as a probability measure, so that

$$P(A) = [S_{\mathrm{n}}(A),\ S_{\mathrm{p}}(A)]$$

where $S_{\mathrm{n}}(A)$ is the lower bound on the probability $P(A)$, or *necessary* support for A, and $S_{\mathrm{p}}(A)$ is the upper bound on the probability $P(A)$, or *possible* support for A. The negation

$$P(\neg A) = [1 - S_{\mathrm{p}}(A),\ 1 - S_{\mathrm{n}}(A)].$$

If, as here, an interval probability is interpreted as a measure of belief, then $S_n(A)$ represents the extent to which it is certainly believed that A is true, $1 - S_{\mathrm{p}}(A) = S_{\mathrm{n}}(\neg A)$ represents the extent to which it is certainly believed that A is false, and the quantity $S_{\mathrm{p}}(A) - S_{\mathrm{n}}(A)$ represents the extent of uncertainty of belief in the truth of A. Three extreme cases illustrate the meaning of this interval measure of belief:

$$P(A) = [0, 0] \text{ represents a belief that } A \text{ is certainly false,}$$

$$P(A) = [1, 1] \text{ represents a belief that } A \text{ is certainly true, and}$$

$$P(A) = [0, 1] \text{ represents a belief that } A \text{ is unknown.}$$

The Bayes conditionalization formula is used to propagate these interval probabilities through the network structure, however, in the case of interval probabilities interval analysis has to be applied to the calculation. Two separate implementations of the approach have been tested in the current analysis, which are described below. Both methods are based on local computations of inferred probabilities at each node, an approach also proposed by Fertig and Breeze (1990). Better constrained solutions can be obtained by means of more recent developments in the field of credal networks, though at considerable computational expense (Cozman, 2000).

2.2.1. *Support Logic Programming (SLP)*

Support Logic Programming (Baldwin 1986a, b) is implemented in the Fuzzy Relational Inference Language (FRIL) (Baldwin et al., 1995), which enables interval probabilities (referred to in FRIL as 'support pairs') to be associated with a set of logical propositions held in a database. It is possible to define relations between these propositions, in the form of conditional probability statements, and then query the database to generate inferences.

Consider the example introduced in Section 2.1 in which the following support pairs are now adopted $P(E_1) = [0.53, 0.94]$, $P(E_2) = [0.19, 0.30]$ (where E_1 and E_2 are as defined in Section 2.1). Write $\theta_1 = P(E_1 \wedge E_2)$, $\theta_2 = P(E_1 \wedge \neg E_2)$, $\theta_3 = P(\neg E_1 \wedge E_2)$ and $\theta_4 = P(\neg E_1 \wedge \neg E_2)$. Interval arithmetic is used to calculate the probabilities of compound propositions, assuming independence, so, for example, for θ_2

$$S_{\mathrm{n}}(E_1 \wedge \neg E_2) = 0.53 \times (1 - 0.30) = 0.371$$

and

$$S_{\mathrm{p}}(E_1 \wedge \neg E_2) = 0.94 \times (1 - 0.19) = 0.761$$

Similarly $\theta_1 = [0.101, 0.282]$, $\theta_3 = [0.011, 0.141]$ and $\theta_4 = [0.042, 0.381]$. The conditional probabilities are assigned as follows: $P(H|E_1 \wedge E_2) = [0.94, 0.94]$, $P(H|E_1 \wedge \neg E_2) = [0.80, 0.80]$, $P(H|\neg E_1 \wedge E_2) = [0.49, 0.49]$, $P(H|\neg E_1 \wedge \neg E_2) = [0.00, 0.70]$. The inference is computed according to the following interval version of the Bayes conditionalization formula (Baldwin et al., 1995):

$$S_{\mathrm{n}}(H) = \min_{\theta_1+\theta_2+\theta_3+\theta_4=1} [0.94\theta_1 + 0.8\theta_2 + 0.49\theta_3 + 0.0\theta_4]$$

where $\theta_1, \ldots, \theta_4$ are constrained so that $0.101 \leq \theta_1 \leq 0.282$, $0.371 \leq \theta_2 \leq 0.761$, $0.011 \leq \theta_3 \leq 0.141$, $0.042 \leq \theta_4 \leq 0.381$. The result of this minimization is $S_n(H) = 0.381$. The upper bound is found using the upper conditional probabilities and maximising: $S_{\mathrm{p}}(H) = 0.466$.

2.2.2. *Interval Probability Theory (IPT)*

Cui and Blockley (1990) developed previous work on interval representations by introducing the parameter ρ, which represents the degree of dependence between propositions E_1 and E_2:

$$\rho = \frac{P(E_1 \wedge E_2)}{\min(P(E_1), P(E_2))}.$$

Thus $\rho = 1$ indicates maximal dependence, whilst if E_1 and A_2 are independent

$$\rho = \max(P(E_1), P(E_2))$$

so that

$$P(E_1 \wedge E_2) = P(E_1) \cdot P(E_2).$$

The minimum value of ρ is given by

$$\rho = \max\left[\frac{P(E_1) + P(E_2) - 1}{\min(P(E_1), P(E_2))}, 0\right]$$

where $\rho = 0$ indicates that E_1 and $E_2$are mutually exclusive.

If ρ is defined as an interval $[\rho_l, \rho_u]$ then

$$S_\mathrm{n}(E_1 \wedge E_2) = \rho_l \cdot \min(S_\mathrm{n}(E1),\ S_\mathrm{n}(E2)) \tag{2}$$

$$S_\mathrm{p}(E_1 \wedge E_2) = \rho_u \cdot \min(S_\mathrm{p}(E_1),\ S_\mathrm{p}(E_2)) \tag{3}$$

$$S_\mathrm{n}(E_1 \vee E_2) = S_\mathrm{n}(E_1) + S_\mathrm{n}(E_2) - \rho_l \cdot \min(S_\mathrm{n}(E_1),\ S_\mathrm{n}(E_2)) \tag{4}$$

$$S_\mathrm{p}(E_1 \vee E_2) = S_\mathrm{p}(E_1) + S_\mathrm{p}(E_2) - \rho_\mathrm{u} \cdot \min(S_\mathrm{p}(E_1),\ S_\mathrm{p}(E_2)). \tag{5}$$

The dependency parameter ρ is an additional item of information, which is elicited in order to address explicitly the dependency between propositions. It is a convenient means of exploring different dependence relationships when the exact nature of dependence is uncertain. Consider the example introduced in Section 2.1, but in this case we recognise that there will be some dependence between flow ranks and flow likelihoods, represented by a dependency of $\rho = [0.67, 0.96]$. Therefore from Equation (2)

$$S_\mathrm{n}(E_1 \wedge E_2) = 0.67 \times \min(0.53,\ 0.19) = 0.127$$

and from Equation (3)

$$S_\mathrm{p}(E_1 \wedge E_2) = 0.96 \times \min(0.94,\ 0.30) = 0.288.$$

From Equation (4)

$$S_\mathrm{n}(E_1 \vee E_2) = 0.53 + 0.19 - 0.127 = 0.593$$

$$\text{i.e. } S_\mathrm{p}(\neg E_1 \vee \neg E_2) = 1 - 0.593 = 0.407$$

and from Equation (5)

$$S_\mathrm{p}(E_1 \vee E_2) = 0.94 + 0.30 - 0.288 = 0.952$$

$$\text{i.e. } S_\mathrm{n}(\neg E_1 \vee \neg E_2) = 1 - 0.952 = 0.048.$$

The inference is again computed according to the interval version of the Bayes conditionalization formula:

$$S_n(H) = \min_{\theta_1+\theta_2+\theta_3+\theta_4=1} [0.94\,\theta_1 + 0.8\,\theta_2 + 0.49\,\theta_3 + 0.0\,\theta_4]$$

where in this case $\theta_1, \ldots, \theta_4$ are constrained so that $0.127 \leq \theta_1 \leq 0.288$, $0.048 \leq \theta_4 \leq 0.407$, $0.53 \leq \theta_1 + \theta_2 \leq 0.94$, $0.19 \leq \theta_1 + \theta_3 \leq 0.30$, $\theta_2 \geq 0$, $\theta_3 \geq 0$. The result of this minimization is $S_n(H) = 0.473$. The upper bound is found using the upper conditional probabilities and maximising: $S_p(H) = 0.834$.

The three alternative inference mechanisms introduced above each provide practical methods for propagating uncertainty in influence diagrams. Their application is illustrated in the following example.

3. Application: The October–November 2000 Floods in the UK

In the autumn and winter of the year 2000 devastating flooding occurred throughout Britain, causing financial losses estimated at £1.4 billion (Penning-Rowsell and Chatterton, 2002). Floods occur nearly every year somewhere in the UK, but the 2000 episodes were noteworthy because vast areas of the country suffered together, in some cases repeatedly and for long durations. For most of the UK, heavy rainfall began in the middle of September and carried on patchily until late December. Three particularly extreme multiple day downpours occurred during this time, though the weather was more extraordinary for its longevity and breadth. Given other unusually severe floods in preceding years speculation was widespread as to whether the floods were an impact of climate change. John Prescott, the Deputy Prime Minister, summed up the mood (BBC, 2000):

> *This was a wake-up call that struck home. When people see and experience these ferocious storms, long summer droughts, torrential rains – more extreme and more frequent – they know something is wrong and that climate change now affects them.*

Before the flooding was even over, the UK Government had commissioned the Centre for Ecology & Hydrology (CEH) and the Met Office to assess the severity of the flood events and the potential link between the flooding and climate change. The research brought together 11 experts in hydrology, meteorology, climate change and statistics. It involved assembling and analysing diverse data from the flood events and then placing it in the context of previous observed flooding and evidence of hydrological changes that are predicted in analysis of climate change. Seeking to address the link between a specific extreme event and climate change is highly problematic from a scientific point of view and perhaps many climate modellers or hydrologists would refuse such a challenge. However, climate modellers in particular will be familiar with the problem of having to respond to policy questions with analytical tools whose shortcomings are all too familiar. The purpose of this paper is not to comment on whether the question posed about the 2000 floods was a useful one to address. Rather it is to demonstrate how difficult policy related questions (and the question used as an example in this paper is by no means unique) can be supported with influence diagrams.

The research following the October–November 2000 floods is detailed in a technical report (CEH, 2001) and briefly summarised here. The study addressed rainfall data as well as river flow data. The latter provides a more direct measure of the severity of flooding but changes in flood regime can be caused by changes to the catchment (such as alteration to floodplain storage, urban development and changes in agricultural practices) as well as climate change. Therefore, whilst the link between measured rainfall and catchment flood responses can be complicated (in particular due to the 'memory' effect whereby rain falling on an already wet catchment is more likely to produce flooding than on a dry catchment), rainfall data can provide an important indicator of change.

Rainfall data were obtained from the network of approximately 2000 rain gauges. Radar provided a second source of rainfall information. River flows were obtained from 11 river flow gauges, some of which had only short records. The analysis also addressed sea surface temperatures (SSTs), air pressure at mean sea-level (PMSL), groundwater levels, soil moisture levels and computer model simulations. Analyses ranged from simple accumulation totals to statistical searches for trends. Whilst the hydro-meteorological analysis involved statistical analysis of many continuous variables, these statistical insights were then related to logical propositions that may or may not be true *i.e.* had two discrete outcomes. In all, 24 separate items of evidence were identified that had some bearing on the following proposition:

SOURCE PROPOSITION: *The October/November 2000 flood events were a manifestation of hydrological climate change.*

This source proposition led to two child propositions:

PROPOSITION B_1: *The October/November 2000 flooding and rainfall were unusual in the historical context*
PROPOSITION B_2: *Hydrological climate change is occurring in the form of increasing frequency and/or magnitude of unusual flooding and rainfall events, particularly when considering longer durations and wide spatial coverage.*

Through study of documentation and discussions with experts each of the 24 items of evidence listed in Table I was associated with a proposition (Figure 2), which was connected, usually through a series of intermediate propositions, to one of the two propositions, B_1 and B_2 stated above.

To populate an influence diagram of the type shown in Figure 2 required the following quantified estimates:

TABLE I
Summary of evidence

Climate models: The Hadley Centre Regional Climate Model (RCM) nested in its General Circulation Model (GCM) predicted return periods (RPs) of rainfall extremes of the type experienced in late-2000 (over long durations), and the 1860 predictions were compared to the 2000 predictions (p. 25).

England & Wales series: Combined river flows for five major rivers (Thames, Severn, Trent, Dee and Wharfe) across England and Wales were analysed for trends in instantaneous annual maxima (AM), high flow days and longer duration annual maxima (pp. 88, 80, 25).

Extended duration AM: 21 sites (same sites as for 'high flow days') were analysed for trends in N-day maxima (average of daily mean flows) over durations of 3–60 days (pp. 88, 86).

Gauge assessment: 1620 rain gauges recorded rainfall amounts for the whole period (September–December). From these records, return periods were calculated according to FEH methodology (Institute of Hydrology, 1999) for the three main events (multiple-day bursts of intense rainfall (pp. 58, 59).

Groundwater: A qualitative assessment notes that 2000 flooding extended the range of recorded river flows and groundwater levels. December runoff totals were without recorded precedent (for any month) in a number of spring-fed rivers including the Kennet, Dorset, Stour and Itchen (p. 23).

Group flow gauges: Combined N-day maxima (average of daily mean river flows) for five major rivers (same data as 'England & Wales series') were compiled over durations of 10–90 days. The top 10 ranking events in recorded history are listed for each duration (p. 25).

High flow series: 21 sites (same sites as for 'extended duration AM') were analysed for trends in the number of days in each year where river flow was above the 97 percentile (p. 87, 86).

Historical extremes: A qualitative assessment notes that, when viewed in a historical context, the extreme rainfall amounts and river flows of late 2000 are rare, but do not appear to be inconsistent with recorded variability at individual locations (p. 25).

Individual flow likelihood: A qualitative assessment notes that for most English gauging stations commissioned in the last 35 years the 2000 floods resulted in unprecedented flows over the 30 and 60-day time spans (p. 24). However, the most outstanding individual events in the autumn of 2000 do not compare with the hydrological extremes locally registered during the most damaging fluvial floods of the twentieth century (p. 25). In addition, AM return periods for 9 targeted flow gauges ranged up to about 100 years. There is no suggestion that the year 2000 AM were exceptionally rare events at these sites, in the sense of being far out of line with the flood flow magnitudes expected to occur roughly once or twice every hundred years or so (p. 68).

Individual flow ranks: N-day maxima (average of daily mean flows) over durations of 10–90 days were recorded separately for four individual gauges, each on a major river. The top 8 ranking events in recorded history are listed for each duration for each gauge (p. 24).

Instantaneous AM: 30 sites were analysed for trends in instantaneous river flow annual maxima (pp. 85, 86).

Long-term rainfall: 13 sites with long records (89–147 years) were analysed for trends in rainfall amounts over durations of 1–60 days (p. 81, 82).

Long-term records 2: In the 'extended duration AM' analysis, two longest river flow records (Thames and Dee) were singled out. No positive trends were present for these records over durations of 3–60 days (p. 88).

(*Continued on next page*)

TABLE I
(*Continued*)

National rain RPs: The Tabony method was used to find return periods for 1 and 2-month duration rainfall totals for localities across the entire nation (pp. 55–57). *National rankings:* National average rainfall amounts were tallied for any non-overlapping 2, 3 and 4-month periods. The top 10 ranking episodes in recorded history were listed for each duration (p. 18).
1970 split test: Return periods were calculated for 11 rain gauges with long records over durations of 1–60 days. The return periods were calculated for two different situations: once for the whole record and once using only pre-1970 data. For a given intensity, return periods are expected to be less for the whole record than for the pre-1970 data (p. 73).
POT data: Five Peak Over Threshold (POT) series for major UK rivers (Ouse, Trent, Severn, Dee and Thames) were analysed for trends in the frequency and magnitude of very large (POT1) and medium-sized (POT3) floods (p. 87).
PMSL: Return periods for pressure at mean sea-level (PMSL) values over the UK were calculated separately for October and November 2000 (pp. 49–50).
Radar assessment: In expert in weather radar observation at the UK Met Office, made a qualitative assessment of radar patterns during the three main events (multiple-day bursts of intense rainfall) (pp. 41–42).
Regional rainfall: Rain gauge amounts from 10 regions comprising England and Wales were converted to return periods via the Tabony method for 2-month and 4-month durations (p. 19).
River flow ranges: Historical records for 8 major ‘representative rivers’ were examined to see if extreme daily flow ranges were extended in 2000 (pp. 21–22).
Short-term records: 15 sites with short records (28–39 years) were analysed for trends in rainfall amounts over durations of 1–60 days (pp. 83–84).
SMDs: A qualitative assessment noted that the eradication of soil moisture deficits (SMDs) and the accompanying recovery of groundwater levels has no recent equivalent in magnitude and rapidity (pp. 16–17).
SSTs: The standard deviation from the mean for sea-surface temperatures (SSTs) was calculated for the north Atlantic and the mid-north Atlantic (pp. 49–50).

Note: Page numbers refer to CEH (2001).

1. a probabilistic estimate of belief in each sink proposition;
2. conditional probability estimates describing the strength of relationship between propositions joined by a link; and
3. for Interval Probability Theory only, pair-wise dependency estimates at each node in the diagram.

Whilst the technical analysis involved quantified methods these did not produce results in a format that could be directly input into the influence diagram. The final stage of constructing a belief measure for each source proposition was therefore based on expert judgement in a workshop session involving seven experts (four from CEH and three from the Met Office). In conducting this type of expert elicitation it is important to, as far as possible, guard against the biases that have been widely observed in the estimation of probabilities from individuals or groups (Merkhoffer, 1987; Bell et al., 1988; Cooke, 1991; Ferrell, 1994). The literature

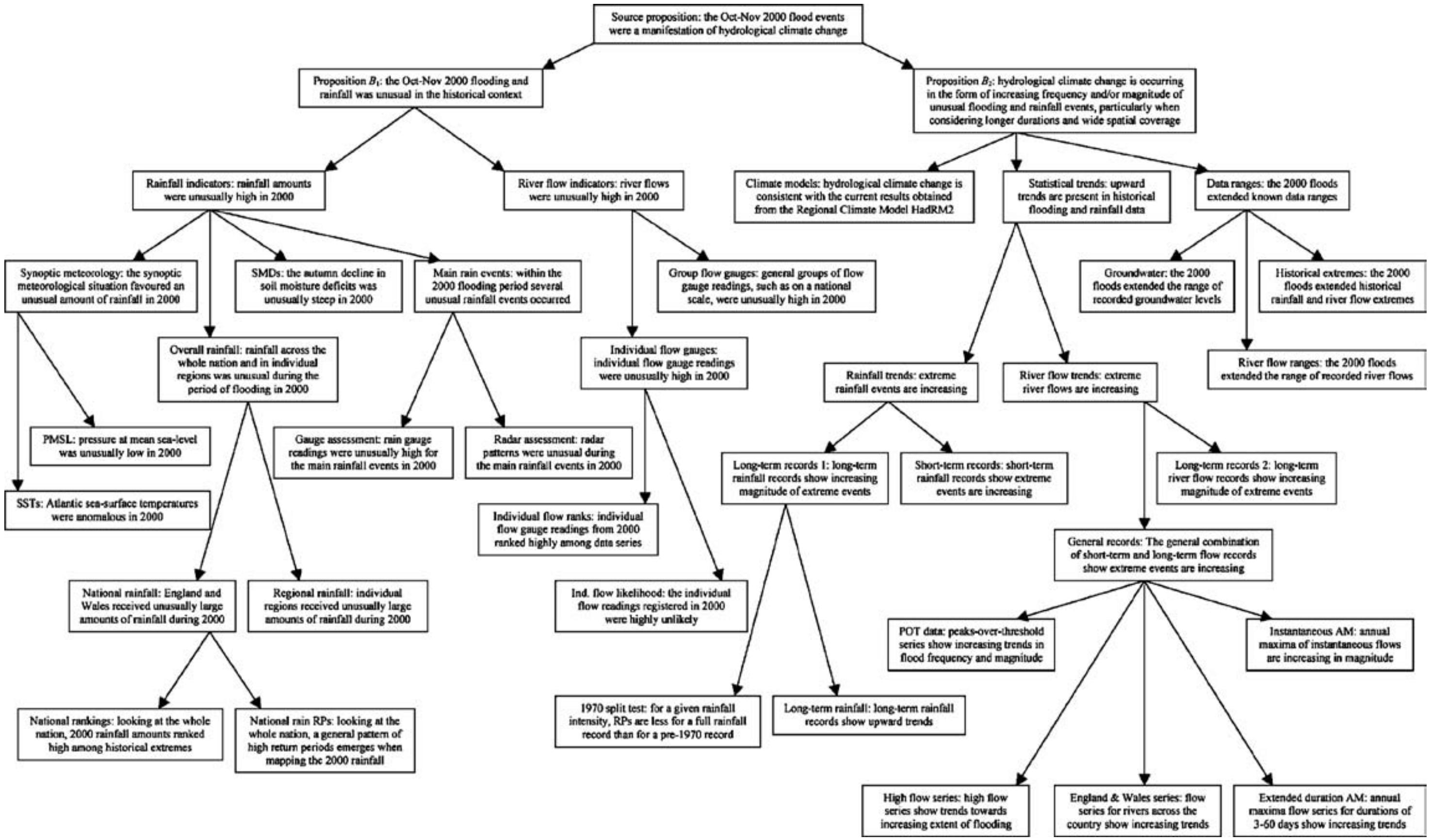

Figure 2. Influence diagram for analysis of October–November 2000 floods.

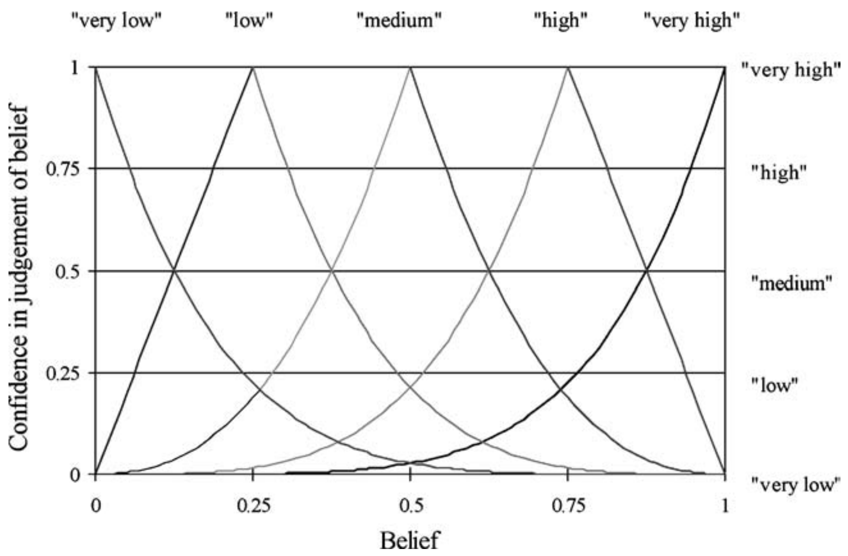

Figure 3. Curves to map from linguistic assessment of belief and confidence to interval probabilities.

on elicitation and combination of interval estimates is less developed but subjects once trained have been shown to cope effectively with the measurement of evidential support and make consistent assessments (Wallsten et al., 1983; Curley and Golden, 1994). Genest and Zidek (1986) and Sentz and Ferson (2002) provide guidance on combination of evidence from multiple sources. For each judgement the seven experts conferred in order to reach a consensus view on the weight of evidence and degree of uncertainty. Whilst for the majority of judgements experts were in broad agreement, for some of the propositions the supporting science or evidence was contested, resulting in disagreement about the interval representation of belief. In these cases an interval reflecting the range of conflicting estimates was assigned.

A method, analogous to fuzzy membership functions, was used to map from linguistic estimates of strength of belief in a proposition and confidence in that assessment to interval point probabilities (Figure 3). Two linguistic judgements are required to extract an interval from Figure 3: an assessment of belief and an assessment of confidence in that belief. The horizontal axis in Figure 3 is a probability scale, so an interval probability is an interval of real numbers on this scale. Each of the five degree of belief (‘very low’, ‘low’, ‘medium’, ‘high’ and ‘very high’) has an associated function (drawn as a curve in Figure 3) reaching a maximum value of unity on the vertical axis at the point probability that reflects the given degree of belief held with ‘very high’ confidence. The vertical axis in Figure 3 is sub-divided into five confidence levels (‘very low’, ‘low’, ‘medium’, ‘high’ and ‘very high’), which relate to the credibility or pedigree of the evidence. The interval measure of belief associated with a given proposition is obtained at the intersection of the relevant confidence level with the relevant belief curve.

The measures of confidence were interpreted subjectively, based on scientific experience. As an example, consider a hypothesis of increases in flood frequency: An upwards trend in 50 years of quality-controlled river flow data that was significant at the 99% level would be "high" belief in the hypothesis with "high" confidence. The same trend in 20 years of data would be "medium" or "low" belief with "high" confidence. A significant trend in long but poor-quality data would be "high" belief with "low" confidence.

In order to compare the different influence diagrams, the same interval probabilities (unconditional and conditional) were input into the SLP and IPT calculations. These probability intervals were converted to point probabilities input into the Bayesian belief network by setting $P(A) = (S_n(A) + S_p(A))/2$. The belief estimates for all of the source propositions are listed in Table II.

In situations where influence diagrams are used to represent phenomena for which there is a large quantity of statistical data, the conditional probabilities required in the network can be estimated from data. This is not the case in the analysis of climate-related problems, where data are scarce and only of partial relevance. Under these circumstances conditional probabilities have to be estimated by experts, based on their knowledge of the relationships between propositions. Of particular significance are the conditional probabilities associated with the links at the top of the diagram. Proposition B_2 is the most influential diagnostic of the source proposition H: the 2000 flood events cannot be a manifestation of climate change if hydrological climate change is not occurring. Proposition B_1 is of some influence, since a highly unusual event strengthens the case for climate change, but is not a necessary condition, because the 2000 flood events were potentially part of a changing climate regardless of their size. This causal reasoning is reflected in the conditional probability assignments: $P(H|B_1 \wedge B_2) = 1.0$, $P(H|B_1 \wedge \neg B_2) = 0.15$, $P(H|\neg B_1 \wedge B_2) = 0.9$, $P(H|\neg B_1 \wedge \neg B_2) = 0.0$.

3.1. RESULTS

The results from the influence diagram analysis of the proposition 'The October/November 2000 flood events were a manifestation of hydrological climate change' are listed in Table III. Figure 4 illustrates the proposition hierarchy implemented using Interval Probability Theory. The modelling tool that was adopted for the analysis (Davis and Hall, 2003) displays interval probabilities as graphical 'Italian flag' icons in which the left hand green proportion represents the belief in a proposition, the right hand red proportion represents belief in its negation and the white represents the uncertainty.

The results from the three different mathematical inference mechanisms demonstrate considerable uncertainty in the proposition that the October–November 2000

TABLE II
Linguistic assessments and numerical belief measures of sink propositions

Item of evidence (shorthand)	Support	Confidence	Interval probability	Point probability
Climate models	Good	Medium	[0.63, 0.88]	0.76
England & Wales series	Poor	Low	[0.06, 0.47]	0.27
Extended duration AM	Good	Medium	[0.63, 0.88]	0.76
Gauge assessment	Good	High	[0.70, 0.81]	0.76
Groundwater	Good	Medium	[0.63, 0.88]	0.76
Group flow gauges	Very good	High	[0.95, 1.00]	0.98
High flow series	Good	Medium	[0.63, 0.88]	0.76
Historical extremes	Poor	Medium	[0.13, 0.37]	0.25
Ind. flow likelihood	Poor	High	[0.19, 0.30]	0.25
Individual flow ranks	Good	Low	[0.53, 0.94]	0.74
Instantaneous AM	Medium	Low	[0.29, 0.71]	0.50
Long-term rainfall	Medium	Low	[0.29, 0.71]	0.50
Long-term records 2	Poor	Low	[0.06, 0.47]	0.27
National rain RPs	Very good	Medium	[0.88, 1.00]	0.94
National rankings	Very good	High	[0.95, 1.00]	0.98
1970 split test	Medium	Low	[0.29, 0.71]	0.50
POT data	Poor	Low	[0.06, 0.47]	0.27
PMSL	Good	Medium	[0.63, 0.88]	0.76
Radar assessment	Poor	Low	[0.06, 0.47]	0.27
Regional rainfall	Very good	High	[0.95, 1.00]	0.98
River flow ranges	Good	Medium	[0.63, 0.88]	0.76
Short-term records	Medium	High	[0.45, 0.55]	0.50
SMDs	Very good	High	[0.95, 1.00]	0.98
SSTs	Good	Medium	[0.63, 0.88]	0.76

floods were a manifestation of hydrological climate change. The proposition receives a support of 0.58 from the Bayesian belief network, [0.21, 0.86] from SLP and [0.37, 0.72] from IPT. The probability measures for proposition B_1 indicate that the 2000 flooding and rainfall were highly unusual in the historical context. The existence of hydrological climate change is much less certain, with nearly as much evidence against Proposition B_2 as for it. Support for the proposition that 'Upward trends are present in historical flooding and rainfall data' (shorthand: 'statistical trends') is 0.49 from the Bayesian belief network, [0.10, 0.68] from SLP and [0.23, 0.65] from IPT. The three results demonstrate more belief against the proposition than in favour of it, and SLP and IPT indicate major uncertainty. Lack of strong statistical trends has a dominant effect on all propositions that are derived from it. Short-term rainfall and river flow records show some modest positive trends,

TABLE III
Derived belief measures for non-sink propositions

	Bayes	SLP	IPT
Source proposition	0.58	[0.21, 0.86]	[0.37, 0.72]
Proposition B_1	0.89	[0.78, 0.94]	[0.82, 0.94]
Proposition B_2	0.52	[0.13, 0.82]	[0.28, 0.68]
Data ranges	0.64	[0.45, 0.75]	[0.48, 0.76]
General records	0.51	[0.24, 0.80]	[0.32, 0.66]
Individual flow gauges	0.71	[0.47, 0.83]	[0.47, 0.83]
Long-term records 1	0.52	[0.21, 0.72]	[0.27, 0.74]
Main rain events	0.74	[0.52, 0.81]	[0.57, 0.81]
National rainfall	0.94	[0.89, 0.98]	[0.89, 0.98]
Overall rainfall	0.95	[0.92, 0.99]	[0.92, 0.99]
Rainfall indicators	0.95	[0.86, 0.98]	[0.90, 0.97]
Rainfall trends	0.53	[0.24, 0.65]	[0.28, 0.67]
River flow indicators	0.85	[0.75, 0.90]	[0.75, 0.91]
River flow trends	0.72	[0.19, 0.84]	[0.28, 0.79]
Statistical trends	0.49	[0.10, 0.68]	[0.23, 0.65]
Synoptic meteorology	0.75	[0.55, 0.85]	[0.60, 0.85]

but the trend in longer records, which carries more weight, is much less clear. In particular, no trends emerge in the ‘England & Wales series’ of five major rivers over a variety of durations.

The output from the influence diagram analysis leads to the following conclusions:

1. The 2000 flooding and rainfall was extreme, and it is highly likely that the 2000 events were ‘unusual’ in the historical context.
2. The existence of long-term hydrological climate change in the UK, at least regarding 2000-type events with long duration and wide spatial extent, is at best uncertain.
3. Evidence in support of hydrological climate change provided by climate model scenarios and the, in some respects, unprecedented nature of the 2000 events, is counteracted, if not outweighed, by an overall lack of clear statistical trends.
4. Given the uncertainty surrounding Proposition B_2, it is impossible to determine whether or not the October/November flooding and rainfall was a manifestation of climate change.

The influence diagram in Figure 4 provides a readily assimilated picture of the sources of uncertainty and conflict in the evidence, illustrating that the

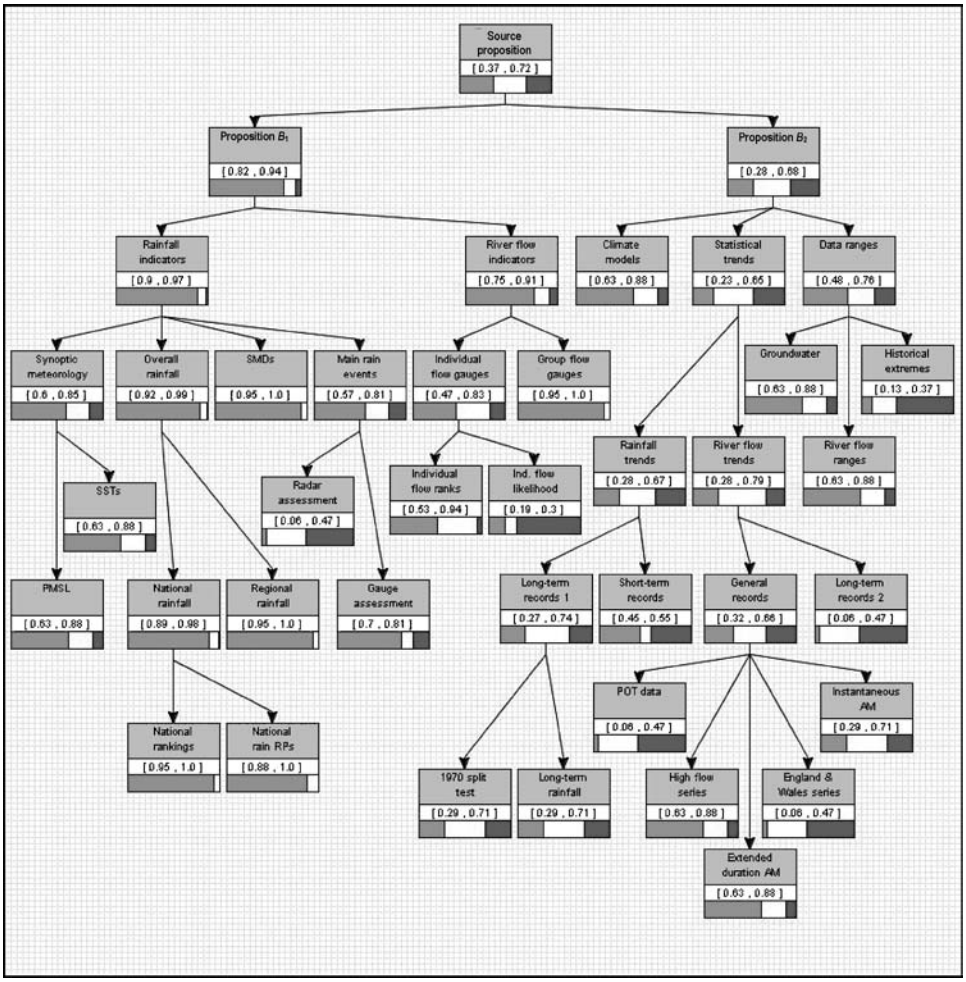

Figure 4. Interval probability influence diagram (three-shaded bars are a graphical representation of the interval probability numbers). See Figure 2 for definition of Source proposition and Propositions B_1 and B_2.

non-committal conclusions are not due to scientists sitting on the fence, but due to insufficient evidence in key arguments that relate the available data to the climate hypothesis. Thus, whilst the results are not conclusive they do provide very useful insights.

There is inevitably a degree of judgement involved in how propositions in an influence diagram, in particular the higher level more abstract propositions, are defined and related. In principle this should not result in different inferences for the parent proposition, as the conditional probabilities connecting nodes in the diagram should reflect the definition of those nodes. In practice, an alternative structure to the one presented in Figure 2 with conditional probabilities elicited from the same experts might be expected to yield slightly different results, because of the variability in expert judgements and biases that depend upon how propositions are stated. One would, nonetheless, expect the same key paths of influence and areas of uncertainty to emerge.

3.2. COMPARISON OF ALTERNATIVE INFLUENCE DIAGRAMS

There are two substantial differences between the three approximate reasoning mechanisms implemented:

1. Conventional Bayesian belief networks use point probabilities whereas SLP and IPT use interval probabilities.
2. IPT uses a dependency measure whereas Bayesian belief networks and SLP assume conditional independence between predecessor propositions.

The extension from point probabilities to interval probabilities is attractive in being able to express ambiguity in probabilistic belief statements (Henkind and Harrison, 1988; Shafer and Pearl, 1990; Krause and Clark, 1993). As discussed previously, this is particularly attractive in situations where evidence is scarce so experts wish to express legitimate indecision, perhaps verging on total ignorance, in their probabilistic estimates. For communication purposes, the simple graphical representation of interval probabilities in Figure 4 gives a convenient overview of areas of belief and uncertainty in a complex problem, which would be hard to communicate as succinctly in linguistic or indeed numerical terms. If point probabilities are entered in SLP it will generate the same results as the corresponding Bayesian belief network.

The implementation of influence diagrams is assisted by the availability of off-the-shelf computer packages, of which there are several for Bayesian belief networks and an increasing number for interval versions. Whilst convenient in some respects users should be wary of the theoretical commitments they are making in adopting a particular approach, which this paper has endeavoured to highlight. The case studied in this paper, in which each proposition could take one of two states ('true' or 'false'), was straightforward to implement in each of the tools adopted, but readers should beware that populating belief networks in which variables can take multiple states can be very time-consuming and, if based on expert elicitation rather than on estimation of probabilities from data, may exceed the ability of experts to make sufficient and consistent judgements.

The introduction of a dependency measure in IPT could be regarded as a violation of the principles of uncertainty propagation in graphical structures, where absence of a link connecting two nodes indicates conditional independence (the Markov condition). However, as the example studied here indicates, in a diagnostic mode of reasoning it is hard to reflect dependencies between different items of evidence that are ultimately related to common phenomena (hydrology and, potentially, climate change). Modelling all the underlying sources of the dependencies will often quickly become unwieldy and may be recursively complex (Ferson et al., 2004). The dependency measure avoids inappropriate independence assumptions, though it does add to the number of probability measures that have to be estimated from data or elicited from experts. It is a convenient means of exploring different dependence relationships when the exact nature of dependence is uncertain. Apart from the dependence measure, IPT is based on the same interval calculations as SLP, so if independence is assumed in IPT it generates the same probability inferences as SLP.

4. Conclusions

The use of influence diagrams for evidential reasoning applied to a problematic climate-related proposition has been demonstrated. The influence diagrams have provided a structured commentary on the conclusion that the events of October–November 2000 were extreme, but cannot in themselves be attributed to climate change. Three alternative inference mechanisms have been tested on the same influence diagram structure. Support Logic Programming and Interval Probability Theory both deal with interval bounds on an unknown probability measure and are attractive in being able to represent in a straightforward way legitimate imprecision in our ability to estimate probabilities. This is particularly useful in situations where evidence is scarce or ambiguous. Interval Probability Theory has the added attraction of being able to represent dependency relationships between evidence that are not implied by the network structure.

Influence diagrams can help to synthesize complex and contentious arguments into a relatively simple, yet evidence-based, graphical output. The graphical structure can formalise expert reasoning, facilitating dialogue between experts, policy makers and other decision stakeholders. In the case of the October–November 2000 floods in the UK, influence diagrams have demonstrated sources of uncertainty and conflict in the available evidence. The analysis has demonstrated how the reluctance of scientists to commit themselves to conclusions about the floods was due to insufficient evidence in the pivotal arguments that related available data via expert reasoning to the hypothesis that the events were a manifestation of climate change. Thus the process of constructing, populating and testing an influence diagram has generated valuable insights.

Acknowledgements

Dr Hall's research is funded by a Royal Academy of Engineering post-doctoral research fellowship. The cooperation of Rob Lamb is gratefully acknowledged.

References

Allen, M.: 1999, 'Do-it-yourself climate prediction', *Nature* **401**, 642.

Allen, M., Raper, S. and Mitchell, J.: 2001, 'Uncertainty in the IPCC's third assessment report', *Science* **293**, 430–433.

Baldwin, J. F.: 1986a, 'Support logic programming', *Int. J. Intelligent Syst.* **1**, 73–204.

Baldwin, J. F.: 1986b, 'Evidential support logic programming', *Fuzzy Sets Syst.* **24**, 1–26.

Baldwin, J. F., Martin, T. P., and Pilsworth, B. W.: 1995, *FRIL – Fuzzy and Evidential Reasoning in Artificial Intelligence*, Research Studies Press, Taunton, MA.

BBC: 2000, 'UK floods a climate alarm call', BBC News, 21 November 2000: http:///news.bbc.co.uk/1/hi/uk_politics/1033458.htm (viewed 26/09/02).

Bell, D. E., Raiffa, H., and Tversky, A. (eds.): 1988, *Decision Making: Descriptive, Normative and Prescriptive Interactions*, Cambridge University Press, Cambridge.

CEH: 2001, 'To what degree can the October/November 2000 flood events be attributed to climate change?', *DEFRA FD2304 Technical Report*, Centre for Ecology and Hydrology, Wallingford.

Chevé, M. and Congar, R.: 2003, 'Managing environmental risks under scientific uncertainty and controversy', Paper presented at the EARE 2003 conference, Bilbao, Spain.

Cooke, R. M.: 1991, *Experts in Uncertainty*, Oxford University Press, Oxford.

Cozman, F. G.: 2000, 'Credal networks', *Artif. Intelligence J.* **120**, 199–233.

Curley, S. P. and Golden, J. I., 'Using belief functions to represent degrees of belief', *Organ. Behav. Hum. Decis. Process.* **58**, 271–303.

Davis, J. P. and Hall, J. W.: 2003, 'A software supported process for assembling evidence and handling uncertainty in decision-making', *Decis. Support Syst.* **35**(3), 415–433.

Ellsberg, D.: 1961, 'Risk, ambiguity and Savage's axioms', *Q. J. Econ.* **75**, 643–669.

Ferrell, W. R.: 1994, 'Discrete subjective probabilities and decision analysis: Elicitation, calibration and combination', in Wright, G. and Ayton, P. (eds.), *Subjective Probability*, Wiley, Chichester.

Ferson, S., Nelsen, R., Hajagos, J., Berleant, D., Zhang, J., Tucker, W. T., Ginzburg, L., and Oberkampf, W. L.: 2004, 'Dependence in probabilistic modeling, Dempster-Shafer theory and probability bounds analysis', *Report SAND2004-xxxx*, Sandia National Laboratories, Albuquerque, New Mexico.

Fertig, K. W. and Breese, J. S.: 1990, 'Interval influence diagrams', in Henrion, M., Shacter, R. D., Kanal, L. N. and Lemmer, J. F. (eds.), *Uncertainty in Artificial Intelligence, 5*, North-Holland, Amsterdam, 149–161.

Gammerman, A. (ed.): 1995, *Probabilistic Reasoning and Bayesian Belief Networks*, Alfred Waller, Henley.

Genest, C. and Zidek, J. V.:1986, 'Combining probability distributions: A critique and annotated bibliography', *Stat. Sci.* **1**(1), 114–148.

Gilboa, I.: 1987, 'Expected utility with purely subjective non-additive probabilities', *J. Math. Econ.* **16**, 65–88.

Gilboa, I. and Schmeidler, D.: 1989, 'Maxmin expected utility with non-unique prior', *J. Math. Econ.* **18**, 141–153.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (eds.): 1996, *Markov chain Monte Carlo in practice*, Chapman & Hall, London.

Grüber, A. and Nakicenovic, N.: 2001, 'Identifying dangers in uncertain climate', *Nature* **412**, 15.

Hall, J. W., Blockley, D. I., and Davis, J. P.:1998, 'Uncertain inference using interval probability theory', *Int. J. Approx. Reason.* **19**(3–4), 247–264.

Hegerl, G. C., Hasselmann, K., Cubasch, U., Mitchell, J. F. B., Roeckner, E., Voss, R. and Waszkewitz, J.: 1997, 'Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change', *Clim. Dyn.* **13**(9), 613–634.

Henkind, S. J. and Harrison, M. C.: 1988, 'An analysis of four uncertainty calculi', *Trans. Syst. Man Cybern. IEEE* **18**(5), 700–714.

Henrion, M.: 1988, 'Propagating uncertainty in Bayesian Networks by Probabilistic Logic Sampling', in Lemmer, J. F. and Kanal, L. N. (eds.), *Uncertainty in Artificial Intelligence 2*, Elsevier, Amsterdam.

Henry, C. and Henry, M.: 2002, 'Formalization and application of the precautionary principle', *Discussion paper 2002009*, Institut de Recherches Economiques et Sociales (IRES), Université Catholique de Louvain.

Howard, R. A. and Matheson, J.: 1981, 'Influence Diagrams', in *The Principles and Applications of Decision Analysis*, Vol. II, Strategic Decisions Group, Menlo Park, CA.

Institute of Hydrology: 1999, *Flood Estimation Handbook*, Institute of Hydrology, Wallingford.

Jensen, F. V.: 1996, *An Introduction to Bayesian Networks*, Springer-Verlag, New York.

Keynes, J. M.: 1921, *A Treatise on Probability*, MacMillan, London.

Krause, P. J. and Clark, D. A.: 1993, *Representing Uncertain Knowledge: An Artificial Intelligence Approach*, Intellect Books, Oxford.

Kuikka, S. and Varis, O.: 1997, 'Uncertainties of climatic change impacts in Finnish watersheds: A Bayesian network analysis of expert knowledge', *Boreal Environ. Res.* **2**(1), 109–128.

Merkhoffer, M. W.: 1987, 'Quantifying judgmental uncertainty: Methodology, experiences and insights', *IEEE Trans. Syst. Man Cybern.* **17**, 741–752.

Oliver, R. M. and Smith, J. Q. (eds.): 1990, *Influence Diagrams, Belief Nets and Decision Analysis*, Wiley, New York.

Pearl, J.: 1988, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo.

Penning-Rowsell, E. C. and Chatterton, J. B.:2002, 'Autumn 2000 floods in England and Wales assessment of national economic and financial losses', Middlesex University Flood Hazard Research Centre.

Pittock, B. A., Jones, R. N. and Mitchell, C. D.: 2001, 'Probabilities will help us plan for climate change', *Nature* **413**, 249.

Risbey, J. S., Kandlikar, M., and Karoly, D. J.: 2000, 'A protocol to articulate and quantify uncertainties in climate change detection and attribution', *Clim. Res.* **16**, 61–78.

Risbey, J. and Kandlikar, M.: 2002, 'Expert assessment of uncertainties in detection and attribution of climate change', *Bull. Am. Meteorol. Soc.* **83**, 1317–1326.

Savage, L. J.: 1954, *The Foundations of Statistics*, Wiley, New York.

Schmeidler, D.: 1989, 'Subjective probability and expected utility without additivity', *Econometrica* **45**, 571–587.

Schneider, S. H.: 2001, 'What is dangerous climate change?', *Nature* **411**, 17–19.

Schneider, S. H.: 2002, 'Can we estimate the likelihood of climatic scenarios at 2100?', *Clim. Change* **52**, 441–451.

Sentz, K. and Ferson, S.: 2002, 'Combination of evidence in Dempster-Shafer theory', *Report SAND2002-0835*, Sandia National Laboratories, Albuquerque, New Mexico.

Shachter, R. D.: 1988, 'Probabilistic inference and influence diagrams', *Oper. Res.* **36**(4), 589–604.

Shackley, S., Young, P., Parkinson, S., and Wynne, B.: 1998, 'Uncertainty, complexity and concepts of good science in climate change modelling: Are GCMs the best tools?', *Clim. Change* **38**(2), 59–205.

Shafer, G.: 1976, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.

Shafer, G. and Pearl, J.: 1990, *Readings in Uncertain Reasoning*, Morgan Kaufmann, San Mateo.

Spiegelhalter, D. J.: 1986, 'A Statistical View of Uncertainty in Expert Systems', in Kanal, L. N. and Lemmer, J. F. (eds.), *Uncertainty in Artificial Intelligence*, North-Holland, New York, pp. 17–48.

Stocker, T. F. and Schmittner, A.: 1997, 'Influence of CO_2 emission rates on the stability of the thermohaline circulation', *Nature* **388**, 862–865.

Stott, P. A., Tett, S. F. B., Jones, G. S., Allen, M. R., Ingram, W. J., and Mitchell, J. F. B.: 2001, 'Attribution of twentieth century temperature change to natural and anthropogenic causes', *Clim. Dyn.* **17**(1), 1–21.

van Lenthe, J., Hendrickx, L., Biesiot, W., and Vlek, C.: 1997, 'A decision-analytic approach to the integrated assessment of climate change', *Risk Decis. Policy* **2**(3), 213–234.

Walley, P.: 1991, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, London.

Wallsten, T. S., Forsyth, B. H. and Budescu, D. V.: 1983, 'Stability and coherence of health experts', upper and lower subjective probabilities about dose-response functions', *Organ. Behav. Hum. Perform.* **31**, 277–302.

Wigley, T. M. L. and Raper, S. C. B.: 2001, 'Interpretation of high projections for global-mean warming', *Science* **293**, 451–454.

Young, P., Parkinson, S., and Lees, M.: 1996, 'Simplicity out of complexity in environmental modelling: Occam's razor revisited', *J. Appl. Stat.* **23**(2–3), 165–210.

Received 5 June 2003; in revised form 3 June 2004