THEME ENV.2012.6.1-1

EUPORIAS

(Grant Agreement 308291)

EUPORIAS

European Provision Of Regional Impact Assessment on a

Seasonal-to-decadal timescale

Deliverable D*33.2*

Report summarising review of existing approaches for communicating confidence and uncertainty

Deliverable Title	*Report summarising review of existing approaches for communicating confidence levels*	
Brief Description	*Report summarising review of existing approaches for communicating confidence levels*	
WP number	33	
Lead Beneficiary	*Suraje Dessai, University of Leeds*	
Contributors	*Andrea Taylor, University of Leeds* *Suraje Dessai, University of Leeds* *Carlo Buontempo, Met Office* *Ghislain Dubois, TEC*	
Creation Date Version Number Version Date	24/05/13 4 29/04/14	
Deliverable Due Date Actual Delivery Date	30/04/2014	
Nature of the Deliverable	x	*R - Report*
		P - Prototype
		D - Demonstrator
		O - Other
Dissemination Level/ Audience	x	*PU - Public*
		PP - Restricted to other programme participants, including the Commission services
		RE - Restricted to a group specified by the consortium, including the Commission services
		CO - Confidential, only for members of the consortium, including the Commission services

Version	Date	Modified by	Comments
V.1.1.	24/05/13	Andrea Taylor	
V.1.2.	25/11/13	Andrea Taylor	
V.1.3.	02/04/14	Andrea Taylor	
V.1.4.	29/04/14	Andrea Taylor	

Table of Contents

List of Tables

List of Figures

1. Executive Summary

This report presents a review of the literature concerning approaches to communicating confidence and uncertainty, with the objective of informing best practice in seasonal-to-decadal climate predictions. As issues pertaining to the communication of uncertainty transcend disciplinary boundaries this review draws upon research conducted in a range of fields, including: weather, climate policy, health and medicine, environmental risk management, economics, experimental psychology, and engineering.

The review highlights a number of factors that are likely to impact on end-users' interpretation and usage of information concerning confidence and uncertainty in seasonal-to-decadal climate and climate impact predictions. These include: ambiguity aversion, trust in information providers, institutional protocol, technical expertise, level of precision, the type of visualisation tool(s) used, and systematic thought biases.

Methods of presenting uncertainty information in numeric, verbal and visual formats are discussed. Numeric presentations permit uncertainty to be formally represented as ranges and confidence limits; but those users with less experience of using statistical information may struggle to extract appropriate meaning from them. The employment of verbal descriptors and evaluative categories may enhance the ability of these users to interpret statistical uncertainty information, but providing these without accompanying numeric ranges can lead to high variability in the way in which 'uncertainty language' is interpreted.

Visualisations provide a versatile way to display uncertainty information at varying levels of complexity. For these to be developed to best effect, however, communicators must consider a) what type of information end-users want; b) how they wish to make use of it; and c) the context in which the information will be used. Various methods of visually communicating uncertainty in seasonal climate predictions already exist. However, to date there has been little systematic testing of how understandable, useful or open to misinterpretation they are. The need to rigorously test methods of communication prior to use, in particular when applied to seasonal-to-decadal climate predictions, is therefore stressed and directions for future research outlined.

2. Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOW, Section B1.1):

No.	Objective	Yes	No
1	Develop and deliver reliable and trusted impact prediction systems for a number of carefully selected case studies. These will provide working examples of end to end climate-to-impacts-decision making services operation on S2D timescales.		X
2	Assess and document key knowledge gaps and vulnerabilities of important sectors (e.g., water, energy, health, transport, agriculture, tourism), along with the needs of specific users within these sectors, through close collaboration with project stakeholders.		X

3	Develop a set of standard tools tailored to the needs of stakeholders for calibrating, downscaling, and modelling sector-specific impacts on S2D timescales.		
4	Develop techniques to map the meteorological variables from the prediction systems provided by the WMO GPCs (two of which (Met Office and MeteoFrance) are partners in the project) into variables which are directly relevant to the needs of specific stakeholders.		X
5	Develop a knowledge-sharing protocol necessary to promote the use of these technologies. This will include making uncertain information fit into the decision support systems used by stakeholders to take decisions on the S2D horizon. This objective will place Europe at the forefront of the implementation of the GFCS, through the GFCS's ambitions to develop climate services research, a climate services information system and a user interface platform.	x	
6	Assess and document the current marketability of climate services in Europe and demonstrate how climate services on S2D time horizons can be made useful to end users.		X

3. Detailed Report

3.1. Introduction

The clear and accurate communication of confidence and uncertainty in seasonal to decadal climate predictions is vital if end users are to be able to utilise these predictions in a truly informed manner. The question of how this information can best be conveyed is therefore one of high importance to both the providers and users of climate predictions. It is not however one to which a simple answer is immediately available. With end users inevitably varying in goals, information preferences, statistical understanding, and technical expertise, formats that suit one user may be perceived as overly complex (or overly simplistic) by another. While a number of methods of communicating confidence and uncertainty in climate information exist, to date relatively little research has empirically tested their efficacy specific context of seasonal to decadal climate predictions. A larger body of research does however exist with respect to the perception and communication of uncertainty in the context of a) weather; b) longer term climate projections, and c) fields outside the domain of meteorology and climatology. Hence, in examining existing approaches to communicating confidence and uncertainty this present review will not only discuss research pertaining to the communication of uncertainty in the context of climate and weather, but also findings from the domains such as economics, medicine, environmental health, policy, and engineering. Of course, as Stephens, Edwards, and Demeritt (2012) point out, best practice for communicating information about uncertainties in one domain may not necessarily apply to all others. For this reason, when discussing the findings of studies outside of the weather and climate domain, care will be taken to stress the nature of the context.

In the next section (3.2) general issues pertaining to the definition and reporting of 'confidence' and 'uncertainty' will be outlined. This will be followed by a broad discussion of various methods of presenting information about uncertainties and the 'cognitive biases' and misunderstandings that may accompany them (3.3), before we turn our attention to existing and proposed methods of visually communicating confidence uncertainty in the specific context seasonal to decadal climate predictions (3.4). Finally, we conclude by outlining directions for further research (3.5).

3.2. General issues in confidence and uncertainty

3.2.1. Defining confidence and uncertainty

One difficulty with using the terms 'uncertainty' and 'confidence' when discussing potential future states of the world is that the terms are often used differently from case to case. In certain contexts, they may be used describe the same thing, with 'high confidence' simply being taken to mean 'low uncertainty'. Alternatively, as is the case in the IPCC Fifth Assessment Report (AR5), 'confidence' may be treated as a very specific type of non-certainty, encapsulating consensus and quality of evidence rather than estimated likelihood (Mastrandrea et al., 2010). In addition to differences in formal definition, the terms may in themselves carry different connotations, with 'confidence' eliciting more positive associations and 'uncertainty' more negative ones. Hence, the choice of vocabulary is not neutral, though this is seldom discussed.

Of course uncertainty itself can be classified in different ways and attributed to different sources (see for example (Dessai & Hulme, 2004; Paté-Cornell, 1996; Spiegelhalter & Riesch, 2011). While the word is often used as something of an umbrella term, encompassing all forms and sources of non-certainty, more precise definitions are sometimes utilised. Following Knight (1921), much of the work on economic decision making and behavioural finance segregates *'decisions under risk'* from *'decisions under uncertainty'*. **Under this classification system *'risk'* concerns those situations where probabilities are well-defined, while *'uncertainty'* concerns instances where there is either a lack of explicit information or incomplete information about probabilities.** A similar distinction is drawn upon by Stirling (2007) who, in discussing risk policy, defines **risk** as a state where the probabilities of potential outcomes can be clearly defined; and **uncertainty** where probabilities cannot be assigned to potential outcomes. Stirling also identifies two further categories of incertitude: **ambiguity**, where potential outcomes are contested and/or not clearly defined; and **ignorance**, where both likelihoods and potential outcomes are unknown. Stirling notes that according to this system of classification a hazard such as flooding may fall into the category of risk if one assumes an unchanging climate; with climatology providing the means to estimate likelihood and magnitude based on historical observation. Under climate change however flooding falls into the category of uncertainty; with past observations being less representative of present likelihood.

Table 1 Stirling's four categories of "Incertitude"

Risk • Probability of events known • Outcomes and impacts of events well defined	**Ambiguity** • Probability of events known • Outcomes and impacts of events unknown or disputed
Uncertainty • Probabilities unknown or incomplete. • Outcomes and impacts of events well defined	**Ignorance** • Probabilities unknown • Outcomes and impacts of events unknown or disputed

This framework allows a distinction to be made between those instances where probabilities alone are unknown and those where there is a lack of knowledge regarding outcomes and impacts. However, it should be noted that this particular use of the term *ambiguity* differs from the way it has been used elsewhere. For example Ellsberg (1990), who formally identified the phenomenon of 'ambiguity aversion' (see subsection 3.2.1. below), uses the term 'ambiguity' to denote Knightian uncertainty (see above) or:

"... *the quality depending on the amount, type, reliability, and unanimity of information, giving rise to one's degree of confidence in an estimate of relative likelihood*" (p.657).

Indeed in much of the literature on judgement and decision making, the phrase ambiguity is used synonymously with uncertainty in to refer to any scenario where information regarding probabilities (or potential outcomes) is incomplete (see for instance (Tversky & Shafir, 1992; Van Dijk & Zeelenberg, 2003). According to this conceptualisation, seasonal climate predictions – or indeed any forecasting system that lacks perfect reliability – are ambiguous in the sense that the models from which they are derived cannot capture every single factor that may impact on the climate system. It should of course be kept in mind that these highly specific definitions of uncertainty do not always match the way in which the term is used in a more generally understood sense (i.e. to refer to instances where there is non-certainty regardless of whether likelihoods are known or not). They do however permit a distinction to be drawn between situations where probabilities are explicitly defined and those where they are not. This split is also captured by the concepts *of first order uncertainty* and *second order uncertainty*. Here, first order uncertainty is synonymous with probability, and is in keeping with Knight's conceptualisation of risk. Second order uncertainty meanwhile is consistent with Ellsberg's notion of ambiguity, encapsulating uncertainty about probabilities (see for example (Vercelli, 1999). Hence, uncertainty pertaining particular prediction or forecast may be decomposed into first order and second order components: first order being the likelihood of a particular event occurring according to model outputs (e.g. distribution of ensemble members); second order being the extent to which models actually capture reality.

This distinction is, to a degree, additionally reflected in that made between *aleatory* and *epistemic* uncertainty (e.g. Paté-Cornell, 1996). In this taxonomy, aleatory uncertainty is that which can be attributed to randomness (and thus be statistically well defined). Epistemic uncertainty meanwhile refers to instances where knowledge is incomplete (due, for example, to errors in measurement and observation, neglected variables, disagreement between researchers, model formulation). It has however been suggested that the concepts of aleatory and epistemic uncertainty are not sufficient to capture all sources of incertitude. The term *human reflexive* uncertainty is used by Dessai and Hulme (2004), to refer to the reflexive response of humans towards uncertainty information (e.g. the potential for behavioural changes made in response to climate change communication to impact on future climate states). Politi, Han, and Col (2007) meanwhile identify *perceived personal relevance* and *information complexity* as further sources of uncertainty on the part of information recipients. Here a lack of certainty regarding future states (and optimal responses) results from neither randomness nor limitations in scientific knowledge, but from recipients' ability (or lack thereof) to comprehend the information with which they are provided and meaningfully integrate it into their own decision making. Indeed end-users 'common sense' understanding of terms such as *confidence*, *uncertainty* and *ambiguity* may in themselves differ from those definitions used by scientists and information providers, rendering interpretation more complex.

In addition to being attributed to different sources, different approaches to quantifying and characterising uncertainty exist. A distinction is often drawn between Frequentist and Bayesian (also known as Subjectivist) approaches to uncertainty. Frequentist probabilities are those where a full probability distribution may be obtained from direct observation, and are held to represent objectively quantifiable properties of the world given fixed underlying parameters (though these may be subject to measurement error). With Bayesian probabilities, values of formally unknown underlying parameters may be inferred from expert belief and probabilistic estimates generated on the basis of said assumed parameters (see for instance (Campbell, 2011) for discussion). As Campbell notes, the manner in which IPCC AR4 and AR5 guidelines segregate 'confidence' and likelihood represents an attempt to separate out subjective and frequentist elements of uncertainty (see Figure 1 and Table 1 below for further details of AR5's confidence and likelihood categories).

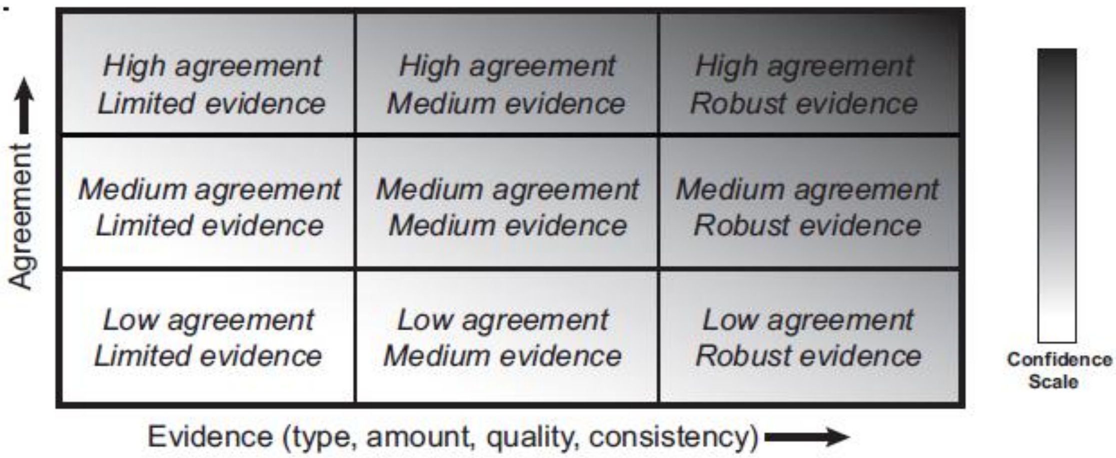

Figure 1 *"*A depiction of evidence and agreement statements and their relationship to confidence. Confidence increases towards the top-right corner as suggested by the increasing strength of shading. Generally, evidence is most robust when there are multiple, consistent independent lines of high-quality evidence". *Figure and caption reproduced from:* ***Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*** *(Mastrandrea et al., 2010)*

Table 2 Verbal likelihood scale set out in *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. *Reproduced from Mastrandrea et al. (2010)*

Term	**Likelihood of outcome**
Virtually certain	99-100% probability
Very likely	90-100% probability
Likely	66-100% probability
About as likely as not	33 to 66% probability
Unlikely	0-33% probability
Very unlikely	0-10% probability
Exceptionally unlikely	0-1% probability

Also incorporated into the AR4 and AR5 guidelines is an instruction that the level of quantitative detail provided with respect to uncertainty should appropriately reflect level of knowledge (e.g. that highly specified probability distributions should not be provided if this level of detail cannot be supported). This was based upon recommendations made by Kandlikar, Risbey, and Dessai (2005), who set out a framework for identifying the appropriate level of detail for communicators to use when conveying uncertainty. Within this framework, a full probability distribution represents the highest level of detail followed by bounded intervals, order of magnitude, expected sign or trend, ambiguous sign or trend and effective ignorance. Since being proposed these recommendations have been incorporated into the guidelines for the IPCCs Fourth and Fifth Assessment Report (AR4 and AR5).

As can be seen from this summary, uncertainty can be defined and characterised in a number of ways, some of which are more highly specified than others. For the purposes of this review however the term will used in accordance with the definition set out in the EUPORIAS glossary[1] at the time of submission.

Uncertainty *"Uncertainty means lack of precision or that the exact value for a given time is not predictable, but it does not usually imply lack of knowledge. Often, the future state of a process may not be predictable, such as a roll with dice, but the probability of finding it in a certain state may be well known (the probability of rolling a six is 1/6, and flipping tails with a coin is 1/2). In climate science, the dice may be loaded, and we may refer to uncertainties even with perfect knowledge of the odds. Uncertainties can be modelled statistically in terms of pdfs, extreme value theory and stochastic time series models."* (EUPORIAS glossary, 29/04/2014)

Hence, uncertainty covers both instances where probability distributions are well defined and those where they are not. Unlike Knight (1921) and Stirling (2007), we will not therefore classify 'decisions under uncertainty' as separate from 'decisions under risk'. Our working definition of risk will however be broadly in keeping with these conceptualisations, referring to the severity of an event weighted by its probability of occurring.

Risk "*Risk is often taken to be the product of the probability of an event and the severity of its consequences. In statistical terms, this can be expressed as Risk(Y)=Pr(X) C(Y|X), where*

[1] http://www.euporias.eu/glossary

Pr is the probability, C is the cost, X is a variable describing the magnitude of the event, and Y is a sector or region." (EUPORIAS glossary, 29/04/2014)

Like 'uncertainty' the term 'confidence' is one whose precise usage may differ depending on context. The confidence interval, for instance, is a statistical estimate of the range of values a parameter might take (e.g. a representation of dispersion). However, the word is also commonly used to refer to what might be termed 'level of belief' or the "trust you can place in a particular form of information". As mentioned above, guidelines for the IPCC AR5 make a clear distinction between confidence and likelihood. In this framework, a measure of likelihood represents an event's estimated probability of occurring (or the probability of an event occurring should a particular prior condition be fulfilled). Confidence, on the other hand, is a composite of amount of evidence (weak, moderate or robust) and level of agreement (low, medium or high). Elsewhere however this semantic distinction is not made. For instance, Han (2013), commenting on the problem of communicating uncertainty in a clinical setting, states that individual risk estimates derived from aggregate frequencies are:

"*...not "true" risk but a figurative expression of scientists' confidence based on the aggregated outcomes of individuals whose characteristics are similar."* (p. 18)

Hence, even when probabilities take the form of observed frequencies, it may be argued that subjective and objective elements cannot be entirely disentangled.

In the present report however the term confidence will be used in accordance with the definition provided by the EUPORIAS glossary (see below), which in itself is broadly in keeping with the terminology of AR5. Although statistical terms, such as *confidence level* and *confidence interval* will, of course, be used in their technically understood manner.

Confidence "*The validity of a finding based on the type, amount, quality, and consistency of evidence (e.g., mechanistic understanding, theory, data, models, expert judgment) and on the degree of agreement. Confidence is expressed qualitatively (Mastrandrea et al., 2010).*" (EUPORIAS glossary, 29/04/2014)

Key Points: Defining confidence and uncertainty

- The words uncertainty and confidence are utilised in different ways in different contexts. In the context of seasonal to decadal climate predictions it is therefore necessary to identify and address potential mismatches in definition between information providers and end-users.
- Uncertainty in predictions and projections arises from different sources (e.g. randomness, model parameters, human activities) and may be represented at different levels of detail (e.g. full probability distribution versus range versus 'direction of change'). Overly specified representations may have the potential to mislead decision makers.

3.2.2. Ambiguity aversion

Ambiguity aversion refers to a tendency, frequently observed in experimental economics, for people to respond aversely to information where probabilities and outcomes are not clearly specified. It may lead decision makers to reject options where probabilities are not precisely specified in favour of those where they are; even when second order distributions of possible likelihoods may render them equivocal. The classic illustration of this is the Ellsberg Paradox (Ellsberg, 1961): the finding that people tend to prefer an option offering a definite 1/3 chance of success to one offering an 'ambiguous' 0 – 2/3 chance of success. The presence or absence of 'missing information' has also been found to impact on judgement and decision making in situations where its actual content would not – had it been known – affect choice. Experimental studies demonstrate that where the outcome of an intermediate event is unknown (e.g. whether one has won or lost a gamble, or succeeded or failed in an endeavour) people may be a) less willing to accept further risk than they would if outcomes were known (regardless of what the known outcome is); and b) willing to pay to defer choice until the outcome of this intermediate event is known (despite making the same choice regardless of what the known outcome is)) (Tversky & Shafir, 1992). Ambiguous information about likelihoods may also be discounted entirely. For instance, in one choice experiment involving a hypothetical business decision it was found that individuals who were told that market interest in a venture was 15-30% were as likely to terminate the venture as those who received no information about likelihoods, while those who were told that interest was 15% or 30% were more likely to continue (Van Dijk & Zeelenberg, 2003).

While the economic and psychological experiments reference above may concern hypothetical or trivial scenarios, their findings have very real and non-trivial implications for communication uncertain information. Ambiguity aversion may lead information recipients – especially those with lower statistical knowledge or less understanding of forecast limitations - to a) prefer information formats that appear to contain 'less ambiguity'; or b) pay for information that reduces (or appears to reduce) it. However, failing to convey ambiguity can be both dangerously misleading (Todini et al., 2005) and detrimental to long term trust in information providers (Pidgeon & Fischhoff, 2011). **Hence, while some recipients may prefer information formats that appear to denote less ambiguity, it is important that communicators avoid creating a false perception of certainty.**

3.2.3. Institutional approach to uncertainty

Tolerance for uncertainty varies not only between individuals, but between organisations. Institutional barriers to the use of new forecasts (and thus the incorporation of new forms of uncertainty into judgement and decision making) may exist (e.g. (Ramos, Mathevet, Thielen, & Pappenberger, 2010). This is a point highlighted by Demeritt, Nobert, Cloke, and Pappenberger (2010) who, discussing uptake and usage of hydrological ensemble prediction systems, note that deviation from previously established institutional practices may lead to greater blame in the case of false alarms or false misses. Methods of integrating institutional risk appetite into the operational use of seasonal climate predictions do of course exist. For example, ROC scores can be used to determine where thresholds for action should lie given a particular organisation's tolerance for false alarms and false misses; though making use of them may be challenging for those without statistical knowledge or appropriate training. Use may be simplified and aided by the incorporation of these preferences into decision aids that provide Act/Don't Act cues (Allen & Eckel, 2012). Again however reluctance to diverge from the status quo may be a barrier to their adoption.

Institutions may also differ in the detail in uncertainty information required for decision making; with some requiring the identification of possible (or plausible) future states and others requiring a full treatment of (aleatory and epistemic) uncertainty. Writing in the field of risk analysis, Paté-Cornell (1996) sets out a six level framework for the treatment of uncertainties, going from Level 0: identification of potential hazards, to 1) identification of worst case scenario; 2) plausible upper bound (i.e. quasi-worst case); 3) best estimate of central tendency (i.e. point estimate based on mean, median or mode); 4) probabilistic risk analysis (in which all uncertainty information is represented by a single risk curve); and 5) full display of uncertainties (e.g. separate risk curves for all models/hypotheses). As she notes, the level of treatment needed may vary from case to case, with lower level representations sufficing in some instances but not in others.

3.2.4 Credibility and trust

The matter of how the presence of uncertainty influences both trust in information and trust in the providers of information is of importance to anybody wishing to communicate uncertainty in climate predictions. As previously noted, failing to adequately communicate uncertainty may, in the long term, lead to a loss of trust in information providers. When it comes to more immediate responses to uncertain information however, the broader judgement and decision making literature is however less clear. In one study examining judgement in a hypothetical medical choice scenario Longman, Turner, King, and McCaffery (2012) found that found that presenting wide likelihood ranges diminished credibility attributed to communicators (in this case clinicians). However, in another investigation – this time using a hypothetical intelligence forecast scenario – the reverse was found (Dieckmann, Mauro, & Slovic, 2010). This disparity may be the result of the different contexts used, with individuals expecting greater certainty from clinicians than intelligence reports. If this is the case one might imagine that climate predictions will be perceived as more akin to the latter than the former. Nonetheless, it is worth noting that a recent study examining trust in financial forecasting systems participants reported lower trust in the system when noise was higher than when it was lower (Goodwin, Sinan Gönül, & Önkal, 2013). However, in this case the fact that noise was high and signal weak was not explicitly conveyed.

The format in which information regarding uncertainties is presented may also influence the level of trust recipients place in it. On examining responses to an environmental hazard scenario, Gibson, Rowe, Stone, and Bruine de Bruin (2013) obtained evidence to suggest that participants presented with 'graph-with-text' displays containing uncertainty information trusted (and liked) said information more than 'graph-with-text displays' without uncertainty information. For text-only displays however a trend in the opposite direction was found. Likewise, the way in which ranges are labelled may also impact on response to the information. Goodwin et al. (2013) for instance found that labelling upper and lower confidence intervals as 'best case' and 'worst case' scenarios was associated with greater trust. This highlights the way in which seemingly inconsequential changes in format and phrasing may influence perception.

3.2.5. User expertise

In communicating uncertainty in climate predictions it should, of course, be kept in mind that one's audience may vary in statistical expertise, experience of using climate predictions, and familiarity with representations of uncertainty in a more general sense.

Research demonstrates that both mathematical ability (see (Peters, 2008) and graph literacy (e.g. Galesic & Garcia-Retamero, 2011; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012) have a pronounced impact on individuals' capacity to understand and utilise numeric and graphically presented information respectively. While it might be expected that a large proportion of those who obtain seasonal to decadal climate predictions will be mathematically adept individuals with high graph literacy, this may not universally be the case; especially where user organisations are smaller or provision more broadly available. Indeed, in the recent EUPORIAS Work Package 33 survey a notable minority of respondents indicated that they were not comfortable using measures of spread (e.g. confidence intervals, standard deviations). Hence, it is important to keep in mind that while some users will have a high degree of mathematical and technical knowledge, others may have far less experience with using statistical information, and thus struggle to interpret commonly used representations. As seasonal climate predictions become more common and reach a larger non-expert audience, this issue is likely to become more pronounced.

There are also factors not directly related to general mathematical ability that may influence the manner in which one's audience may interpret and utilise information regarding uncertainty. In a set of choice experiments featuring environmental management scenarios, Gregory et al. (2012) presented both experts (employed by the US Fish and Wildlife Service) and laypeople with information pertaining to uncertainties about the outcomes of alternative environmental management options. Information was presented in the form of either numeric ranges, evaluative categories (e.g. 'high', 'medium' or 'low' confidence in efficacy), or a combination of both. **Comparison of the responses in each condition indicated that when both quantitative and qualitative information was presented together experts tended to base their decisions on numeric ranges while laypeople tended to rely on evaluative categories.** This difference remained when mathematical ability was controlled for. As Gregory et al. notes, this suggests that presenting information in multiple forms may not necessarily lead to a convergence of interpretation between experts and non-experts, even where non-experts possess good mathematical ability.

The potential for divergence between the information preferences of experts and the understanding of information recipients is also highlighted by Bruine de Bruin and Bostrom (2013),who stress the need for experts to systematically examine the current state of their intended information recipients' knowledge prior to designing communications. Such 'mismatches' between communicators and information recipients are not, of course, restricted to the flow of information from technical providers to non-technical users. Todini et al. (2005) – discussing the communication of uncertainty in the context of flood forecasting – highlight the divergence between the preferences of researchers (for highly parameterized models) and operational flood forecast users (for representations that facilitate the recognition of clear thresholds for action).

While most users (or potential users) of seasonal to decadal climate predictions cannot be said to be laypersons in the strictest sense of the word, level of expertise and statistical knowledge is likely to vary amongst them. As will be further discussed in Section 3, semi-technical recipients of uncertainty information (i.e. non-experts who possess some statistical knowledge) do not necessarily outperform non-technical recipients in tests of comprehension (e.g. Ibrekk & Morgan, 1987). Any difficulties that non-expert recipients of information have with interpreting uncertainty information may be additionally compounded by time limitations and competing demands**. In their report on good practice in communicating uncertainties in environmental assessments Kloprogge, Sluijs, and Wardekker (2007) recommend the adoption of a 'progressive disclosure of information' strategy, whereby information regarding uncertainties is presented in custom made 'layers'; with outer layers providing a non-technical summary and inner layers becoming progressively more technically detailed.** They also stress the importance of a) ensuring that the uncertainty information is provided in a manner relevant to the intended recipient and their decision goals; b) placing information about the uncertainties most relevant to user decision making in those sections of reports most likely to be read (e.g. abstracts, summaries, conclusions, sections pertaining to key goals); and c) indicating which (potentially relevant) uncertainties were excluded from analysis. These sentiments are largely echoed in a recent report by the US Board on Population Health and Public Health Practice in their report**,** which stresses the importance of considering both the expertise of one's audience and the phase of decision making in which information will be used (*Environmental Decisions in the Face of Uncertainty*, 2013).

Conceptualising the movement of climate information from initial scientific knowledge to use in adaptation strategies as a 'process chain' Stoverinck (2011) notes that, as one moves from projection to use in adaptation, information is passed from a group of users comprised mainly of scientific experts to user groups containing a large proportion non-experts. However, in a content analysis of documents coming from various points in the 'process chain', she found that documents placed later in the chain sometimes underplayed uncertainties emphasised at earlier points in the chain, or failed to discuss uncertainties associated with projections at all. As previously noted such failures to effectively convey uncertainty threaten both organisational planning and trust in communicators.

3.2.6. Heuristics and biases

When making judgements as to how to respond to a situation decision makers may not always have the time, capacity or inclination to perform a full systematic analysis of the available information. The word heuristic is widely used in the judgement and decision making literature to refer to simple 'mental shortcuts' or 'rules of thumb' that may be used in lieu of more cognitively demanding processes in choice and judgement (e.g. Gigerenzer & Goldstein, 1996; Kahneman, 2011). The employment of these does not necessarily lead to poorer decisions. Indeed, it has been argued that correctly extracting the essential ‘gist’ of an item of information is more important to accurate interpretation than being able to recall its verbatim characteristics (e.g. Reyna, 2008). However, a number of cognitive biases have been identified that may adversely influence judgement when individuals are faced with uncertainty. Many of these have been examined within the context of psychological and economic experiments. However, their specific relevance to the communication of climate and information has been noted (e.g. Nicholls, 1999). Listed below are a set of heuristics and biases that may potentially impact upon the interpretation and usage of information regarding uncertainty in climate predictions

- Ratio bias effect

Also known as denominator neglect, the ratio bias effect is a systematic tendency to respond differently to ratios represented with larger numbers (e.g. 100 in 1000) to logically equivalent ratios represented with smaller numbers (e.g. 1 in 10) (e.g. Denes-Raj & Epstein, 1994). This bias has the potential to influence judgement of any frequency-based representation of probability that does not make use of consistent denominator information. For instance, if denominators are neglected, a ‘40 in 100’ chance of above average winter temperatures would be interpreted as more likely than a ‘4 in 10’ chance of the same. While this effect would not perhaps be anticipated amongst those recipients of climate information who possess statistical expertise, it has been observed amongst educated populations (e.g. Peters et al., 2006). This bias may be circumvented by the use of: a) consistent denominators; or b) a percentage format.

- Base rate neglect

It has frequently been observed that individuals faced with conditional probabilities may neglect base rates (e.g. Goodie & Fantino, 1996; Hoffrage & Gigerenzer, 1998). To borrow a hypothetical illustration from Nicholls (1999): if a model that is accurate 90% of the time predicts that a drought will occur in a region where there is historically a 10% chance of being in drought, then (assuming that prior probabilities remain constant) there is both a 9% chance of the model delivering a hit (90% chance of the model correctly predicting a state that occurs 10% of the time) and a 9% chance of the model producing a false positive (10% chance of the model incorrectly rejecting a state that occurs 90% of the time) (see Table 3). However, a forecaster may struggle to convince a forecast user that the actual likelihood of a drought occurring at the location is just 50%.

Table 3 Likelihood of hit, miss, false alarm, and correct non-detection for a forecasting system with 90% accuracy predicting an event (drought) with a 10% likelihood of occurring

	Drought observed	Drought not observed	*Total*
Drought predicted	**Hit** **9%** (90% chance prediction correct * 10% chance drought will occur)	**False Alarm** **9%** (10% chance that prediction is wrong * 90% chance that drought will not occur)	*18%*
Drought not predicted	**Miss** **1%** (10% chance that prediction is wrong * 10% chance that drought will occur)	**Correct non-detection** **81%** (90% chance that prediction is correct * 90% chance that drought will not occur)	*82%*
Total	*10%*	*90%*	

The forecast example used here is of course deterministic rather than probabilistic and assumes both a) an unchanging climate; and b) an unrealistic forecasting system that produces misses and false alarms at an identical 10% rate relative to observations. However, it serves to illustrate how prediction users with less statistical experience may struggle to integrate reliability and skill information with model output.

- Framing

This is a bias whereby the manner in which information is 'framed' (e.g. as a gain versus loss, or 'event occurring' versus 'not occurring') influences choice or judgement. This may lead to inconsistent responses to objectively identical information (Tversky & Kahneman, 1981). For instance, '70% chance of above average rainfall' may be judged differently than '30% chance of average or below average rainfall'. Presenting both 'probability of event' and 'probability of not-event' is recommended as a way to counter this issue (e.g. (D. Spiegelhalter, Pearson, & Short, 2011). Indeed, current IPCC AR5 guidelines explicitly instruct authors to strive to avoid framing effects (Mastrandrea et al., 2010).

- Availability

The availability heuristic refers to a tendency to base judgements regarding likelihood and magnitude on salience and the ease and 'availability' with which examples of an event come to mind (e.g. Slovic & Fischhoff, 1977; Tversky & Kahneman, 1973). For example, drought in a particular locale may be perceived as more likely if instances where this has previously occurred come easily to mind (regardless of what present meteorological indicators suggest).

- Probability weighting

This term refers to a tendency for those making judgements and decisions to subjectively weight linear increments in probability in a non-linear fashion, with sensitivity decreasing as probabilities move away from certainty (e.g. (Kahneman & Tversky, 1979). For instance, the difference between a 0 and 5% chance of an event occurring will be perceived as subjectively larger than the difference between a 50% and 55% chance of said event occurring.

- Overweighting of new information

The tendency to give new information undue ‘weight’ when making judgements has been observed amongst technical users of hydrological forecasts (Kahneman & Tversky, 1979). It would seem plausible that this bias might also be observed amongst both technical and non-technical users of other types of forecast information.

- Confirmation bias

Recipients of information may seek out or attend more to information that supports existing practices or judgements, while downplaying disconfirming evidence (see for instance (Bazerman & Moore, 2012) for discussion). Hence, information regarding prediction uncertainty may be neglected by those who favour the method of prediction being used, but focussed upon by those who do not (Vaessen, 2003); see (Kloprogge et al., 2007), for English summary). Similarly, the output of new models may be favoured when they ‘confirm’ prior beliefs, but disregarded when they do not. In the domain of seasonal to decadal forecasting, one might anticipate confirmation bias to occur when users hold strong pre-existing beliefs regarding future weather and climate states.

- Bounded rationality

As previously mentioned, reliance on simple rules of thumb is not necessarily a maladaptive response to complex situations. As Simon (1957) points out, consistently employing expected utility analyses for every decision that one makes would be impossible.

In the context of European flood forecasting it has been observed that at least some specialist users and risk manager wish to receive information in a manner that facilitates straightforward Act/Don’t Act decisions (see for instance (Demeritt et al., 2010; Todini et al., 2005). On a similar note, in action research with Australian farmers, McCown and colleagues found that users of an analytic interactive tool for simulating possible planting outcomes in different conditions tended to use it to develop ‘intuitive’ If/Then responses (McCown, 2012; McCown, Carberry, Dalgliesh, Foale, & Hochman, 2012).

On a related noted, responses to a recent survey conducted with EUPORIAS stakeholders and other interested organisations also indicates that a large proportion of respondent organisations would like to receive information in a manner that facilitates Yes/No decision making; although they may wish to receive more comprehensive data in addition to this. **The question of how uncertainty information can best be integrated into**

representations that facilitate the identification of appropriate thresholds for action (where these are desired) would thus seem to be of some importance.

Key Points: General issues in communicating confidence and uncertainty

- Ambiguity aversion means that information recipients may prefer to receive clearly defined magnitudes and likelihoods to ranges and confidence levels. However, failing to provide information about uncertainties can lead to false perceptions of certainty and be detrimental to both safety and trust in information providers.

- Organisations are likely to vary in both their tolerance for uncertainty in seasonal to decadal climate predictions and the level of detail regarding uncertainty they require. Methods of representing uncertainty that incorporate institutional thresholds for action can facilitate the use of forecast and projections.

- The information presentation preferences of end-users may not always match those of information providers. Such mismatches must be addressed.

- As experience of using both climate predictions and statistical information is likely to vary considerably amongst the end-users of seasonal to decadal climate forecasts, it is important to keep in mind firstly, that not all users may be familiar with certain statistical concepts (e.g. confidence intervals, pdfs); and secondly, that experts and non-experts users may focus on different aspects of the information being communicated (e.g. quantitative data versus qualitative evaluative categories).

- When it comes to interpreting information about uncertainty there are a number of thought biases that have the potential to lead to the neglect or misinterpretation of important information. These may affect both technical as well as non-technical users of information. When designing methods of communicating uncertainty one must take into account (and strive to mitigate) the potential for framing effects and other cognitive distortions.

3.3. Representing uncertainty

When communicating uncertainty to a recipient one is faced with the matter of how best to represent the information. As Stephens et al. (2012) note a communicator may face trade-offs between richness (i.e. level of detail and resolution), robustness (i.e. accuracy of deterministic predictions, reliability of probabilistic predictions, and the appropriate reflection of skill) and salience (i.e. comprehensibility and usability). The distinction between robustness and level of detail is stressed by Dessai, Hulme, Lempert, and Pielke (2009), who – writing with respect to longer term climate adaptation – argue that unwarranted precision may lead to poor adaptation decisions. This sentiment is echoed in the domain of health risk communication by Nelson, Hesse, and Croyle (2009), who advise against the use of unnecessary detail when reporting numeric values (e.g. reporting values to several decimal places when a whole number would be sufficient). It should however be noted that the manner in which more highly specified numeric values versus less highly specified estimates of likelihoods and outcomes are perceived by decision makers does not yet appear to have been systematically tested.

In order to provide greater detail a communicator may provide more information (or more varied representations of the same information). However, in doing so they may contend with the problem of cognitive overload. That is to say that when presented with a high volume of information people may not have the capacity to process all of it, or identify the most important features, and thus become overwhelmed or ‘overloaded’ (see Nelson et al., 2009), for review). In addition to this, information recipients – and indeed information providers – may be subject to certain cognitive biases.

The following three subsections will focus on the various ways in which uncertainty can numerically, verbally and visually represented. It is, of course, recognised that this distinction is somewhat artificial, as these are frequently combined in various ways. For instance, calibrated language may be presented with corresponding numeric ranges. Graphs and other visual representations may be appended with verbal descriptions or numeric tables. However, for the purposes of discussing the individual consideration associated with each form it was felt that this division – while artificial – is pragmatically justifiable.

3.3.1. Numeric

3.3.1.1. Likelihood format

The question of how to best present numeric information to recipients is one that has long been a point of interest in the judgement and decision making literature. Evidence strongly suggests that the manner in which numeric information is interpreted can be strongly influenced by the format in which it is presented: that is to say, whether it is presented in a percentage, ratio (also referred to as ‘frequency’) or standardised (i.e. 0 to 1) format. Slovic, Monahan, and MacGregor (2000) for instance observed that even medical professionals rated a fictitious psychiatric patient as posing a greater danger when their estimated risk of committing a violent offense was described in frequency rather than percentage format. It has been proposed that humans are innately better equipped to utilise likelihood information presented in the form of natural frequencies (e.g. “out of every 1000 cases similar to yours”) (e.g. Gigerenzer, 2003). Indeed, it does appear that presenting conditional probabilities in frequency form reduces base rate neglect (e.g. Gigerenzer & Hoffrage, 1995).

However, evidence from the field of weather forecasting has indicated that, when receiving probabilistic forecasts, members of the US public both prefer percentage based representations to frequency formats (Morss, Demuth, & Lazo, 2008) and better understand then (Joslyn & Nichols, 2009). As Stephens et al. (2012) note, communicators cannot automatically assume that a format that works well in one context will function as well in another. When contrasting forecasts of potential future weather and climate events to predictions of health outcomes, it is perhaps understandable that such differences may arise. It would seem plausible that being prompted to imagine 1000 patients like oneself may lead to a more concrete and salient mental representation than being prompted to imagine 1000 days like tomorrow. It is also worth noting that in the previously mentioned Work Package 33 user needs survey a strong majority of respondents indicated a preference for the representation *“30% chance of rain”* over the structurally identical *“3 in 10 chance of rain”* and *“.3 chance of rain*”.

3.3.1.2. Interpreting probabilistic forecasts

The extent to which those without technical or statistical expertise can effectively use information regarding forecast uncertainty is a manner that has generated considerable interest in the field of probabilistic weather forecasting. Although research conducted in this area has typically focussed on members of the public rather than operational users, this work would seem to have relevance for the communication of information regarding climate predictions to users without a statistical background.

On a positive note, recent research conducted with the US public suggests a) that those presented with probabilistic forecasts were more likely to take appropriate action than those presented with deterministic ones (Joslyn & LeClerc, 2012); and b) that people infer a degree of uncertainty into purely deterministic forecasts (Morss et al., 2008). Hence, it would appear that non-technical recipients of meteorological information can effectively utilise probabilistic information to make appropriate decisions. However, it has been noted that many people make reference class errors when presented with such forecasts (e.g. Gigerenzer, Hertwig, Van Den Broek, Fasolo, & Katsikopoulos, 2005). This is to say that a 70% chance of precipitation may be misinterpreted to mean that it will rain over 70% of a particular area, or that 70% of weather forecasters agree that it will rain tomorrow, when what is actually intended is that there is a 70% chance that rain will occur in a specific area, during a specified time period. While it has been argued that a complete, normatively correct understanding of the information presented may not always be a barrier to appropriate adaptive action (see (Handmer & Proudley, 2007), for discussion), it may lead people to make decisions they would not have otherwise made. For instance, in a hypothetical protective action task Morss, Lazo, and Demuth (2010) found that participants who believed that probability of precipitation estimates referred to 'percentage of time' it would rain had a lower threshold for taking (costly) protective action. Hence, when it comes to communicating measures of likelihood to those using said information for organisational planning and decision making, the importance of ensuring that it is made clear what said likelihoods refer to should be stressed. Again, many users and potential users of seasonal to decadal climate predictions have a high level statistical understanding, and will thus be more likely to interpret simple probabilistic information as intended. However, it is important to recognise that not all may have the same experience of working with probabilistic forecasts and prediction. It is also possible that even those with greater statistical understanding may make reference class errors when representations differ from those that they are used to.

3.3.1.3. Conveying uncertainty through numeric ranges

Probabilistic forecasts of the kind discussed above typically feature point estimates without any indication of dispersion. While such representations may be considered suitable in some contexts (e.g. weather forecasts provided for public consumption), they are unlikely to be appropriate in situations where a) dispersion is high; and b) small probabilities of extreme events of concern. They may also mislead users into believing that forecasting systems offer more certitude than they do.

One way of communicating dispersion (or spread) in estimates of likelihood and magnitude is of course through the use of numeric ranges (e.g. 30-40% chance of rain tomorrow) rather than point estimates (e.g. 35% chance of rain next week). The question of how non-technical and less statistically experienced users perceive and utilise numeric ranges versus point estimates is one that has drawn interest in various fields. In the area of health findings have been inconsistent. In one study involving a hypothetical treatment scenario Longman et al. (2012) found that participants presented with range estimates demonstrated lower understanding and higher perceived risk. **In this study participants presented with large ranges also attributed less credibility to information providers.** On the other hand, in a similar study concerning hypothetical personalised cancer risk estimates, Han et al. (2011) found that presenting 'ambiguity' in likelihood estimates as a numeric range increased worry but had no main effect on perceived risk. **In both instances 'ambiguity aversion' was cited as a reason for differences between those presented with ranges and those presented with point estimates.**

As previously stressed however, responses to information regarding uncertainty in a context such as health may differ from those in response to climate predictions. Indeed, in studies examining responses to weather forecasts, findings have been somewhat more supportive of the ability of non-technical recipients' ability to adaptively utilise uncertainty information in the form of ranges. Roulston, Bolton, Kleit, and Sears-Collins (2006) found that providing participants with information regarding a temperature forecast's error range improved performance on a hypothetical road salting task. Findings obtained by Joslyn and Savelli (2010) meanwhile suggest that the US public anticipate bias in existing probabilistic forecasts (even in instances where it is not warranted). Joslyn and Savelli thus argue that the provision of specified ranges around probabilistic forecasts may be necessary to counteract this. It should not however be assumed that the provision of numeric ranges to indicate uncertainty will necessarily lead to uniform utilitarian responses. In a hypothetical cost/loss task (where participants must indicate whether they would choose to take a costly precaution in order to reduce the impact of an even more costly potential threat), Morss et al. (2010) found that participants responded differently to scenarios featuring potential damages from frost and damages from flooding, despite the expected value and uncertainty structure remaining constant.

Key Points: Numeric representations of uncertainty

- While it is helpful to examine work conducted in other fields it should be kept in mind that the best representations of uncertainty to use in other areas (e.g. health and finance) may not be the best to use when it comes to seasonal to decadal climate predictions. Thus, recommendations to use frequencies rather than percentage representations of probability made in other domains may not be useful when it comes to communication climate predictions.

- While ambiguity aversion with respect to the presentation of range versus point estimates has been observed in certain fields, research examining the interpretation of probabilistic weather forecasts suggests that non-technical users can effectively utilise probabilistic information presented as numeric ranges – though their decision making may not be indicative of a 'normative' cost/loss trade-off.

3.3.2. Verbal

The use of verbal descriptions to convey uncertainty regarding likelihood has been suggested as one way to circumvent the problems posed by variations in probabilistic understanding and lack of precision in numeric estimates. The IPCC AR5 guidelines, for example, provide detailed instructions for the use of calibrated language to describe both likelihoods and level of confidence (amount of evidence and level of consensus) (Mastrandrea et al., 2010). Work in other domains, such as health suggests that providing verbal 'evaluative categories' can aid the comprehension and decision making when numeric information proves difficult to evaluate (e.g. Peters et al., 2009). However, as will now be discussed, caution should be taken when utilising verbal likelihoods in communication.

3.3.2.1. Verbal expressions of likelihood

One issue is, of course, that different individuals may interpret probabilistic terms such as 'likely' and 'unlikely' in markedly different ways. Budescu, Broomell, and Por (2009), for instance, found that participants presented with statements extracted from the IPCC's Fourth Assessment Report AR4 (IPCC, 2007), along with the verbal likelihood descriptors used in said report, demonstrated high variability in their estimates of probability. Although this was lower for those participant provided with accompanying numeric information that which probabilistic ranges were covered by phrases such as 'likely' or 'unlikely'. There is also evidence to suggest that people conflate verbal estimates of likelihood with representations of risk (i.e. a composite of magnitude and likelihood rather than just likelihood). Findings obtained by Patt and colleagues suggest that the interpretation of verbal statements regarding likelihoods can be influenced by a) the perceived severity of an event (Patt & Schrag, 2003); and b) whether one is being asked to act as a communicator or recipient of information (Patt & Dessai, 2005). Disparity between the intentions of communicators and the perceptions of users has also been observed with communications concerning over the counter medicines. Berry, Raynor, Knapp, and Bersellini (2004) found that, when presented with terms such as 'common' and 'rare', members of the public considerably overestimated the possibility of experiencing adverse side effects.

Keeping the above in mind, general recommendations made with respect to the communication of verbal likelihoods include:

- Using the same stem (e.g. very likely, likely, as likely as not, unlikely, very unlikely) for all terms within the scale in order to reduce (though not eliminate) variability in interpretation (Lipkus, 2007)
- Presenting numeric ranges each time verbal likelihoods are utilised, with the size of the range indicating the degree of uncertainty regarding the likelihood in question (Budescu et al., 2009)
- Establishing, prior to use, that the scale used reflects the perceptions of the intended users as far as possible (e.g. Berry et al., 2004)

3.3.2.2. The segregation of confidence and likelihood

As previously noted, current IPCC AR5 guidelines specify that (numeric and verbal) estimates of likelihood should be segregated from (verbal) summaries of 'confidence' (though high confidence should be assumed if a full probability distribution is provided). In the IPCC AR4 levels of 'confidence' were assigned numeric values, but these were later dropped on the grounds that they may this lead to confusion between 'confidence' as defined in the report (i.e. a composite of robustness of evidence and consensus amongst experts) and 'statistical confidence'. Similar schemes for rating evidence and consensus in the health domain are also in use (see Han, 2013). The U.S. Preventive Services Task Force (2008) uses a three point ‘low’, ‘medium’ and ‘high’ scale to denote “Certainty regarding net benefit”. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group segregates (qualitatively ranked) quality of evidence from strength of recommendation (Balshem et al., 2011). This separation of different sources of uncertainty is, of course, done to enable greater transparency and comprehension. However, to the author of this review’s knowledge, the manner in which these scales are interpreted by information recipients has not been systematically tested.

As Risbey and Kandlikar (2007) point out, one potential problem with segregating estimates of likelihood from estimates of strength of evidence and consensus, is that they cannot be fully separated on a conceptual level. To express ‘near certainty’ with ‘low confidence’ seems paradoxical. In response to this criticism, directed at earlier reports, guidelines for IPCC AR5 state that where low or very low confidence exists likelihood information should not be provided.

3.3.2.3. Linguistic uncertainty

Another challenge faced by those wishing to communicate information verbally is the potential for what Carey and Burgman (2008) have termed ‘linguistic uncertainty’. That is for uncertainty about the meaning of a communication to arise as a result of the ambiguity (possibility for multiple interpretations of the same word), vagueness, under-specificity or context dependence of the language used. This problem has been recognised in a number of diverse domains including weather forecasting (e.g. Handmer & Proudley, 2007), hydrology (Demeritt et al., 2010), health (Politi et al., 2007), food safety (Lofstedt, 2006), fisheries management (Hauge, Nielsen, & Korsbrekke, 2007), toxicology and engineering (Christensen, Andersen, Duijm, & Harremoës, 2003). Of particular relevance is the persistent observation that a large proportion of US and Australian recipients of probabilistic weather forecasts misunderstand what probability of precipitation estimates refer to (e.g. Hadmer & Proudley, 2007; Gigerenzer et al, 2005; Morss et al. 2008). This is to say that a “70% chance of rain tomorrow” may refer be taken to refer to a temporal (i.e. rains for 70% of the day) or spatial (i.e. rains over 70% of the area described) quality rather than a probabilistic one (i.e. it will rain on 70% of all days like tomorrow) because the ‘reference class’ has not been specified. Hence, an individual may understand what “70%” as a ratio means without correctly identifying the event it pertains to.

As Christensen et al. (2003) point out, terminology regarding uncertainty varies from field to field and it may not be possible to create a standardised set of terms. However, it is important that terms should be clearly defined within the context in which they are being utilised.

Key Points: Verbal representations of uncertainty

- Representing uncertainty verbally using calibrated language (e.g. very likely, likely, unlikely) can lead to differences in the way in which likelihoods are interpreted by information recipients. Evidence suggests that problem may be reduced by consistently including numeric ranges with each use of such language (as oppose to providing a single table of definitions).

- In presenting information about uncertainty to end users one should strive to reduce 'linguistic uncertainty' by clearly defining the terms one is using. This is especially important in the context of seasonal to decadal climate forecasts and projections where the phraseology used by information providers may have different meanings to end-users.

3.3.3. Visual and audio representations

A wealth of visualisations are available to those seeking to communicate information regarding uncertainty to recipients of varying levels of technical expertise; ranging from relatively simple bar and pie charts to elaborate interactive tools. Nelson et al. (2009) advise those seeking to visually communicate information to take into account the perceptual processes of proximity, continuation and closure (Wertheimer, 1938). *Proximity* is the propensity to *“perceive items that are close to each other in the visual field to be related in some way”* (p125); a process that can facilitate comparison if the items to be compared are placed close together in an ordered fashion (see also (Hibbard & Peters, 2003). *Continuity* refers to *“the eye’s tendency to follow lines and directions implied by separate elements of the visual field”* (p 126); with Nelson et al recommending that – when constructing tables – decimal points be aligned and alternate lines shaded. Closure, is the tendency *“for people to ‘fill in’ missing information that is not specified in a presentation to make sense of the presentation as a whole.”* (p 127); thus, the need for clear labelling that eliminates the need to ‘fill in’ is emphasised.

Again however, the question of which precise representation one should use is not one that has a simple answer.

3.3.3.1. What is being communicated?

The first question that a communicator may ask in considering how best to visually represent uncertainty regarding likelihoods or magnitudes is what the visualisation in question will be used for. Spiegelhalter et al. (2011) makes a number of recommendations for tailoring visual representations to the communication of uncertainty, including: the creation of multiple graphics for multiple users; the provision of part to whole comparisons; the avoidance of framing effects by presenting frequencies or percentages ‘with’ and ‘without’ the outcome being represented; informative labelling; the avoidance of “chart junk”; assessing the needs of the audience; and rigorously testing all visualisations prior to use.

When it comes to the representation of simple quantities, it is generally held that bar charts are useful for facilitating comparison between magnitudes; pie charts, for enabling the highlighting of a particular proportion; and line graphs for displaying trends over time (see

Nelson et al., 2009, for full review). The notion that the different representations facilitate the extraction of different information also holds true for the depiction of more complex statistical information. In their seminal research on the subject of visual communication Ibrekk and Morgan (1987) observed that, when asked to identify the mean of a binomial distribution, simple confidence intervals (on which means were explicitly marked) elicited the greatest number of correct estimates; while probability density functions and cumulative density functions tended to erroneously elicit modal and median responses respectively. When it came to making comparisons between values however, cumulative density functions were found to outperform other representations. These observations were echoed in a recent series of investigations by Edwards, Snyder, Allen, Makinson, and Hamby (2012). On presenting participants with a range of risk management scenarios where Act/Don't Act responses were required, Edwards and colleagues observed that participants were more likely to correctly identify mean values and take appropriate action when graphics rendered the correct response visually explicit (e.g. error bars depicting mean values; complementary cumulative distribution functions indicating when a threshold was crossed). These findings serve to highlight the importance of a) considering how a visual representation of uncertainty information will be utilised, and whether they render the most important and useful characteristics salient; and b) testing whether users interpret the information depicted in the manner that communicators intend.

With respect to visually conveying information regarding uncertainty in weather forecasts, Roulston and Kaplan (2009) observed that participants presented with fan charts depicting statistical confidence in temperature forecasts performed better in a decision task than those presented with point estimates. S. L. Joslyn and Nichols (2009) meanwhile found that presenting US participants with pie charts depicting chance of rain versus chance of no rain reduced reference class errors. These studies were of course conducted with members of the public rather than those using forecasts in an operational context. However, it is worth keeping in mind that not all those using climate predictions in their work will necessarily have technical and statistical understanding above that of the general public. Indeed, in the recent Work Package 33 survey it was found that representations of spread, such as error bars and fan charts were less highly favoured by those reporting lower comfort with using statistical information (although it should be kept in mind that unlike Roulston & Kaplan the survey measured preference rather than performance).

Of course, the representations discussed above suggest well defined distributions of likelihoods and magnitudes. This leaves the question of how second order uncertainty may be represented. It has been suggested that when uncertainty results from a lack of consensus or a dearth of evidence (IPCC AR5's two components of 'confidence') shading, transparency and blurring may be used to convey this (see Spielgelhalter et al, 2011 for an example). As Politi et al. (2007) note, the efficacy of using such representations to communicate uncertain information does not appear to have been widely tested. However, as will discussed further in Section 3.4, recent work by Jupp, Lowe, Stephenson and colleagues have recently explored the way in which transparency (Slingsby et al., 2009) and colour saturation (Jupp, Lowe, Coelho, & Stephenson, 2012; Lowe et al., 2013; Slingsby et al., 2009) can be used to communicate first order uncertainty (ensemble distribution) and second order uncertainty (reliability and skill) in seasonal forecasts.

The question of how resolution can be appropriately represented is also one that poses a challenge to communicators. For instance, the use of smooth contours on maps may create the impression that a prediction provides greater spatial precision that it does. However, as Stephens et al. (2012) report, the use of blocky rather than smooth contours may render the information less salient (and thus less understandable) to end users.

3.3.3.2. Familiarity

Preferences for visual representation that are already familiar have been observed amongst recipients of information. The recent WP33 user needs survey, conducted with EUPORIAS stakeholders and other interested organisations, found a consistently strong positive association between existing use of visualisations and ratings of preference. This is also echoed in the recent findings of Daron and colleagues (2014) who, on examining responses to different methods of communicating uncertainty in climate projections, found that respondents tended to prefer a more familiar bar graph representation most highly over more novel visualisations. Meanwhile, in the domain of hurricane forecasting a survey conducted by the US NWS with members of the public meanwhile indicated that respondents preferred existing representations of the 'cone of uncertainty' in hurricane forecasts (Broad, Leiserowitz, Weinkle, & Steketee, 2007). This preference for familiar representations is not limited to non-technical users of climate and weather information, Pappenberger et al. (2013) for instance, observed that when a workshop group composed predominantly of hydrologists, meteorologists and flood forecasting experts were asked to design a format for the delivery of a 10 day discharge forecast many opted for familiar formats.

Preference for familiar formats is not of course a negative thing in itself. However, greater familiarity may not always correspond with greater comprehension and usefulness. In their studies of graph interpretation, neither Ibrekk and Morgan (1987) nor Edwards et al. (2012) observed a consistent relationship between familiarity and performance. Nonetheless, if the necessary information can be accurately and coherently conveyed using a form familiar to and/or preferred by users then there would seem to be a clear case for utilising it. Again, the importance of fully testing visualisations prior to use should be stressed.

3.3.3.3. Potential for misinterpretation

As with numeric and verbal representations of uncertainty those that are visual in nature may not be interpreted in the manner a communicator intends. As previously mentioned, the results of Ibrekk and Morgan's (1987) study indicated that the 'highest point' on a probability distribution is often selected as the mean. Evaluation of the manner in which US residents interpret the cone of uncertainty has shown that many people anchor on the track line and misinterpret the cone to represent the boundaries of the area that might be affected by a hurricane (e.g. Broad et al., 2007).

The misinterpretation and misuse of graphical representations of uncertainty is not limited to non-technical and semi-technical users. Demeritt et al. (2010) for example observed instances of hydrology sector professionals stating a wish to follow single ensemble members on hydrological ensemble forecasts presented as spaghetti graphs. Also, commenting on the potential (mis)interpretation of spaghetti graphs, Dettinger (2005) notes that the structure of said diagrams may lead to misapprehensions regarding spread; with the eye being drawn to extreme upper and lower visual bounds, rather than the concentration of lines in more central areas. In cases where users are concerned with potential extremes then drawing attention to upper and lower limits may not pose a problem.

However, where overall distribution and measures of central tendency are of interest, users may struggle to extract this information from visualisations where it is not clearly depicted. The choice of whether to use minimum and maximum values or confidence intervals as upper and lower limits may also influence users' interpretations of spread and likelihood, by inducing them to anchor on visible limits.

In their analysis of climate adaptation documents Stoverinck and colleagues observed considerable inconsistency in the colours used to depict the direction predicted changes in temperature and precipitation in visualisations (Stoverinck, 2011; Stoverinck, Dubois, & Amelung). Given their use in other contexts, colours such as blue, red and green may evoke multiple pre-existing associations amongst information recipients. For instance blue may be associated with both 'water' and 'cold'; red with 'heat', 'stop' and 'danger'; green with 'verdancy' and 'go'. Hence, this variation in 'what colours mean' may lead to considerable confusion and misunderstanding. Stoverinck et al. also critique the use of maps in communicating uncertainty in climate projections; arguing that they may inhibit the integration of information and render trends difficult to interpret. The importance of avoiding the selection of counterintuitive colouring on climate maps is also emphasised by Kaye, Hartley and Hemming (2012), who also stress the importance of considering a) colour-blindness (i.e. avoiding red/green scales where possible); and b) the potential for the use of a large number of hues to render maps difficult to interpret.

Once again, the importance of fully testing methods of visual communication must be emphasised.

3.3.3.4. Context of use

The context in which the information might be used should also be considered. When presented with uncertain information, both technical and non-technical recipients are faced with the challenge of integrating it with a) other sources of information; and b) current institutional practices (e.g. Demeritt et al., 2010). Hence, information that is not presented in a manner that facilitates this may not be used (or used inappropriately). **As research with both those in the hydrology sector (e.g. Demeritt et al., 2007) and Australian agriculture (McCown et al., 2012), along with the findings of our recent user-needs study indicate, end-users may wish to have it presented in a way that facilitates straightforward 'yes'/'no' responses to situations**

The capacity of decision makers to interpret and utilise graphical representations in an operational context may also be influenced by factors such as time pressure and cognitive load. For instance, in their study of graph comprehension, Edwards et al. (2012) observed that the addition of a time pressure manipulation decreased choice accuracy. Earlier writings on the subject of visual communication have also echoed this point. Lamberti and Wallace (1987) note that graphical displays should seek to minimise cognitive load and make use of symbols that are meaningful and easily discriminable in the particular context in which they are used. While Lamberti and Wallace's focus was on displays utilised in the field of military command decisions, it is a point that would seem to generalise to any context where decision makers may be under time pressure or overburdened with information from a multitude of sources.

3.3.3.5. Using sound to represent uncertainty

While the bulk of this subsection has concerned itself with visual representations, it should also be noted the potential for sound to be utilised in communication of uncertainty is increasingly being explored (Brown & Bearman, 2012). In a recent study examining whether variations in pitch could facilitate the interpretation of uncertainty information in UKCP09 climate projections[2] it was found that those presented with sonification performed better on an interpretation task, and had faster response times than those presented with visual information alone (Bearman, 2011). This effect appears to have been strongest amongst those already familiar with UK CP09, possibly suggesting that this format could be of greatest use to more experienced users of climate information. Although given the complexity of the underlying information, it is also possible that a similar approach may facilitate the interpretation of less complex datasets amongst a broader range of users.

[2] An example of how sound was utilised in this study can be at http://vimeo.com/17029358

Key Points: Visual and audio representations of uncertainty

- In presenting information about uncertainty in seasonal to decadal predictions to end-users, information providers may have to made trade-offs between richness (detail and level of specificity), robustness (accuracy or reliability) and salience (how clear and understandable information is). It is therefore vital to understand the information requirements of end-users.
- The best visualisation to use in any given instance is likely to depend both on what being communicated and the context of use. Where an end-user has a particular threshold for action, visualisations that render the threshold salient are likely to facilitate understanding and usage.
- Preference for a particular form of uncertainty representation may not always denote better understanding of (and ability to use) it. Information providers are thus faced with the challenge of producing representations that are both acceptable to end-users and facilitate accurate interpretation.
- When seeking to provide information about uncertainty to end users the context in which the information will be used should be taken into account. Factors such as time pressure could mean that end-users need to extract the 'gist' of information quickly and without extensive deliberation. The matter of how end-users wish to use representations is also important; some may require representations that explicitly facilitate Act/Don't Act decisions.
- When designing visualisations care should be taken to reduce any confusion that may arise from choice of colour. This is especially important in the context of seasonal to decadal predictions due to strong pre-existing associations between colours and climate/weather events (e.g. red with heat and/or danger; blue with both cold and/or water). Ensuring that end-users are able to easily link colours to states is vital.
- The use of sound in the communication of uncertainty may assist interpretation amongst some users.

3.4. Current and proposed methods of visualising uncertainty in seasonal climate predictions

There are a number of ways in which information regarding uncertainty in seasonal to decadal climate predictions can be visualised. What follows will be an overview and discussion of the types of representation both currently in use amongst climate service providers and those suggested for use. Four key areas will be covered: predictions represented using maps (3.4.1), predictions represented using graphs (3.4.2), visualisations for communicating reliability and skill (3.4.3); and decision aids (3.4.4).

3.4.1 Maps

Maps are one of the most frequently utilised ways of presenting seasonal climate predictions. In a recent survey conducted with EUPORIAS stakeholders and interested organisations, maps emerged as the most highly favoured format for representing uncertainty in climate information. Given their capacity to display spatial information in a way that allows visual characteristics to be matched with real world locations, this is perhaps unsurprising. There are however a range of different way in which this format can be utilised to represent climate information at a seasonal to decadal timescale.

3.4.1.1 Terciles, quintiles and two-category.

Maps communicating seasonal climate predictions typically illustrate the extent to which predictions diverge from long term averages. One deterministic approach is to illustrate a best estimate of the extent to which a variable is expected to exceed or fall below the average (e.g. by degrees Celsius, millimetres of precipitation, wind speed mph, etc.) However, while representing predicted anomalies in this manner allows for estimated magnitude to be displayed, it does not permit uncertainty to be depicted in detail. Although a strippling effect can be used to denote regions where dispersion amongst ensemble members is low, providing an indication of how high agreement is (see for example (Kaye et al., 2012).

To illustrate exactly how the distribution of ensemble members compares to past observations two-category, tercile or quintile representations can be used. Of these terciles are perhaps the most popular. To create tercile representations, past observations are split into three categories: upper tercile (33.3%), middle tercile (33.3%) and lower tercile (33.3%). The proportion of ensemble members falling each category is then used to represent the estimated likelihood (according to the model used) of the variable falling into these upper, middle and lower categories. Likewise, a two-category representations are based on a two way split in past observations (upper 50% and lower 50%), while quintiles utilise a five way split (each category representing 20% of past observations). Maps then visually depict the proportion of ensemble members falling into a given category using gradations of colour.

The seasonal temperature maps depicted in Figures 2a - c below are taken from the Met Office website, which allows visitors to view seasonal predictions for a range of variables in the same format[3].

[3] http://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/glob-seas-prob

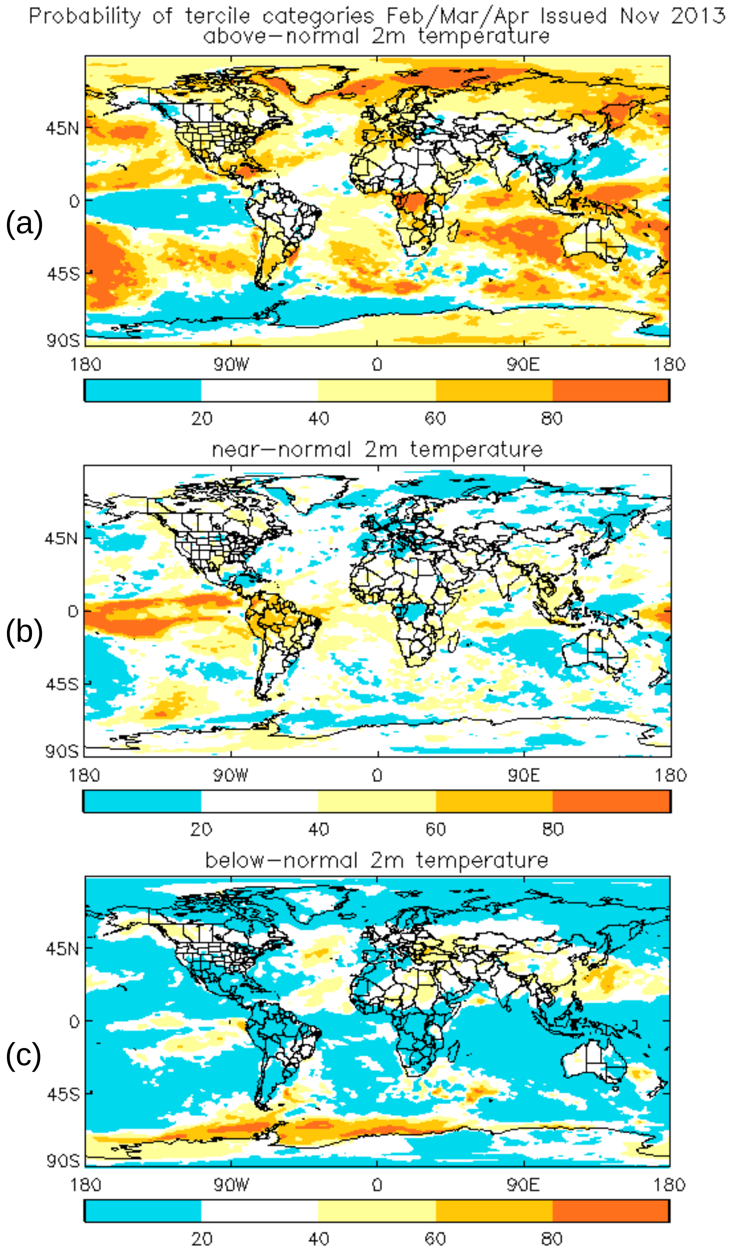

Figure 2a-c Maps illustrating predicted likelihood of 2m temperatures falling into above average (a), near average (b) and (c) below average tercile categories. *Reproduced from the Met Office website*

As is explicitly noted on the Met Office website, the visualisations represent raw data and should not be taken to constitute forecasts for a given area. However, they do illustrate one way of presenting uncertain information. The three maps represent the predicted likelihood that temperatures will fall into the upper, middle and lower tercile respectively. Here red hues are used to represent higher likelihoods, with the deepest red corresponding with the highest category of likelihood, while blues represent lower likelihoods (with deeper blues corresponding with lower likelihoods). The same information as a two-category map (e.g. predicted likelihood of temperature being above normal) and highest and lowest quintile (e.g. predicted likelihood of temperature falling into the highest or lowest 20% category). The benefit of representing terciles and quintiles in this way is that the predicted likelihood of the variable of interest falling each category can be presented. However, it also comes with the difficulty of multiple maps being required, thus potentially making comparisons more difficult. This form of representation also brings with it the question of how colours should be utilised to depict likelihoods. Using the same colour scheme for each map permits visual consistency, thus potentially facilitating comparison between maps. However, it may also confuse when the colours utilised have strong existing associations (for instance, when highly saturated reds are used to depict a high likelihood of lower than average temperatures).

Tercile information may also be displayed on a single map. One way of doing this is to display the predicted likelihood of the 'most likely tercile' for each region covered. An example of this, provided by IC3, can be in Figure 3 below. Here red, yellow and blue scales are used to depict upper, middle and lower terciles respectively. This format requires a more complex colour-scheme, but enables a consistent use of colour for different variable magnitudes. In this case blue always corresponds with 'lower than average' wind speed.

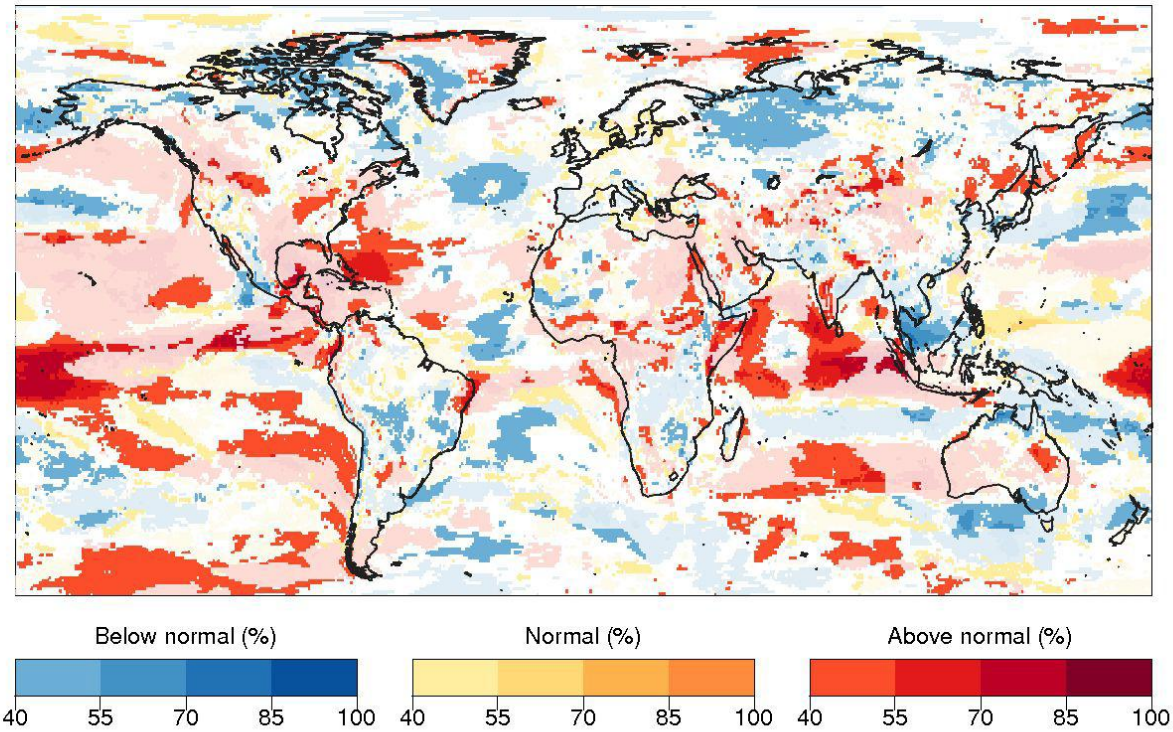

Figure 3 "Probabilistic forecast of most likely tercile for 10m wind speed (%).White colour indicates where forecasts probabilities of all 3 categories are below 40% and approximately equal. Transparent areas indicate where the observations did not match the forecast". *Visualisation and text provided by IC3.*

An extension of this visual approach is used by Lowe et al. (2012) in their visualisations for Dengue fever forecasts (Figures 4a - b). In this case however colour is determined by a set of three coordinates reflecting predicted probability of Dengue incidence rates falling into lower, middle and upper terciles, Hence, stronger shades of blue, yellow and red correspond with greater probability of incidence rate falling into lower, middle and upper tercile respectively. Saturation is also used here to display information gain relative to a reference forecast (e.g. how much the forecast differs from historical observation).

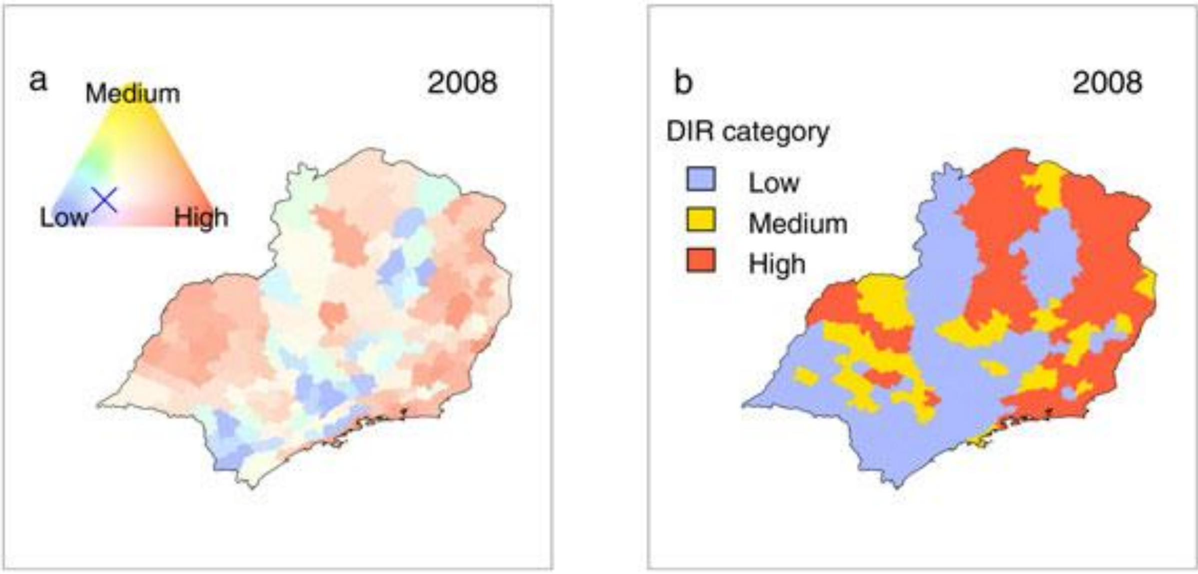

Figure 4 (a) Dengue fever incidence rate (DIR) forecast. Colour assigned to each region of the map is based on coordinates corresponding with predicted probability of DIR falling into the lower, middle or upper tercile. Saturation also denotes 'information gain' relative to a reference forecast (labelled with an X). **(b)** Observed DIR represented as Low, Medium or High tercile. *Figures reproduced from Lowe et al. (2012)*

Other ways of using a single map to present tercile information include superimpose bar graph representations of the proportion of ensemble members falling into upper, middle and lower terciles onto a map, or using glyphs where length corresponds with the predicted likelihood of a variable falling into a particular tercile category (Figure 5). By providing details about all three terciles these visualisations increase the amount of information about each region available to users. Whether this additional information is of benefit to users is likely to depend on the details of interest to decision makers (e.g. whether users wish to receive an overview of the likelihood of all states or simply information about possible extremes).

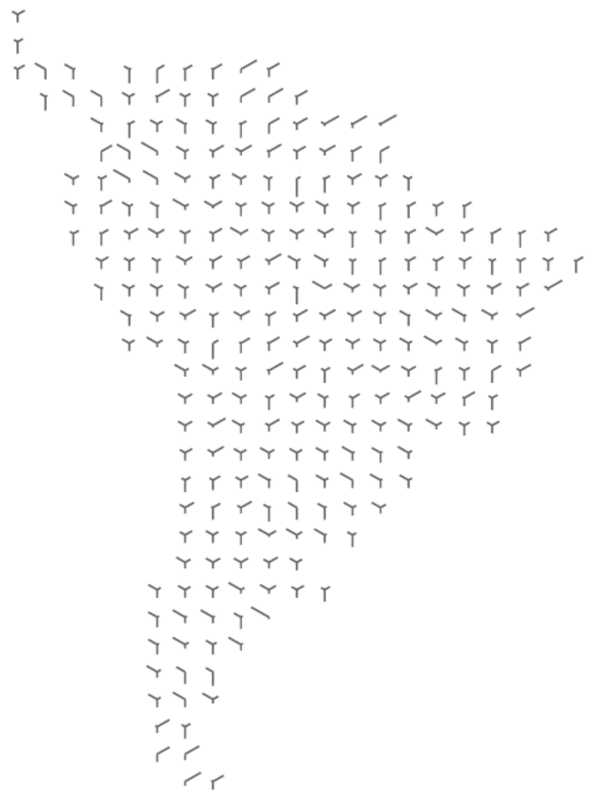

Figure 5 Tercile-based probabilistic precipitation forecast depicted using glyphs. The arm length of each glyph correspond to probability of upper, middle and lower tercile respectively (left, bottom and right arm*s).* Hence, a longer right arm corresponds with greater predicted likelihood of lower than average precipitation*. Image reproduced from Slingsby et al. (2009). http://www.gicentre.org/papers/gisruk09/climate.pdf.*

Each of the map styles discussed has benefits and limitations, and given the diversity of user needs and statistical knowledge there may not be an optimal format. More empirical research is however needed to determine which formats work best in which context

3.4.1.2 Colour scheme

As touched upon above, when constructing maps, or indeed any other form of visualisation, to represent climate information, it is important to consider both visual salience and pre-existing connotations. Colour schemes featuring red and green may be inaccessible for those with colour-blindness, while the use of a large number of hues may render it displays confusing and difficult to interpret (e.g. Kaye et al., 2012). Kaye et al. also raise the point that where saturation is used to denote differences in likelihood or magnitude, subtle variations in shade may not be obvious when areas of the same hue are far apart.

With respect to pre-existing connotations, red and blue in particular are colours that have strong associations with certain climate variables, with red being associated with heat and blue with both cold and water. Hence, a visualisation where blue corresponds with higher temperatures or lower precipitation may lead some users, especially those with less experience of using said visualisations, to misunderstand the information presented. The use of white may also be contentious, with some associating it with missing information, others with ‘middle range’, and others with ‘no signal’. This was an issue that arose in comments made in the recent Work Package 33 user needs survey. As noted in the previous section however the fact that some colours have different connotations in different disciplines likely precludes the creation of a scheme that can be perfectly intuitive for all users.

When presenting multiple maps, each representing a different measures or climate variable in the same, a question arises as to whether a single colour scheme for all. As previously noted, consistency may cause a problem when usage is counterintuitive. However, the use of a different colour scheme for each measure may confuse if many measures are being presented. Once again, this highlights the need for appropriate user testing.

3.4.1.3 Temporal dimension
While maps provide an excellent format for viewing information spatially, it is perhaps more of a challenge to incorporate a temporal dimension. One option is, of course, to present for different timeframes side-by-side or as a slide show that users may click through. Another method is to present change over time as an animation. This format has been piloted and discussed by Slingsby et al. (2009), who report a favourable response on the part of users. Although, as Slingsby et al. note, one concern with using animation in this way is that smooth transitions from predictions at one timeframe to the next may create the illusion of continuity.

3.4.2 Graphs
While maps enable spatial elements of information to be clearly and directly illustrated it is, as previously noted, a challenge to present temporal information in this manner. Likewise, maps do not offer a straightforward way to present ranges, confidence levels and other measures of dispersion (or ‘spread’). Such information may thus be more easily depicted in the form of graphs. Of the types of graph available for this purpose, some enable temporal continuity (e.g. spaghetti plots, fan graphs) while others depict timeframes discretely (e.g. probability density functions, cumulative density functions, error bars, box plots). In the recent EUPORIAS Work Package 33 user needs survey, both fan graphs and error bars received generally high favourability ratings. Although, it is worth noting that they were significantly less popular amongst those reporting lower comfort with statistics.

3.4.2.1 Spaghetti plots
Spaghetti plots are commonly used to depict the trajectory of all ensemble members over time. The examples in Figure 6 illustrate how they may be used to display predictions with low-to-moderate and high dispersion. As has previously been noted however, while they provide a clear indication of where extremes lie, they may render the overall distribution less visually salient (Dettinger, 2009).

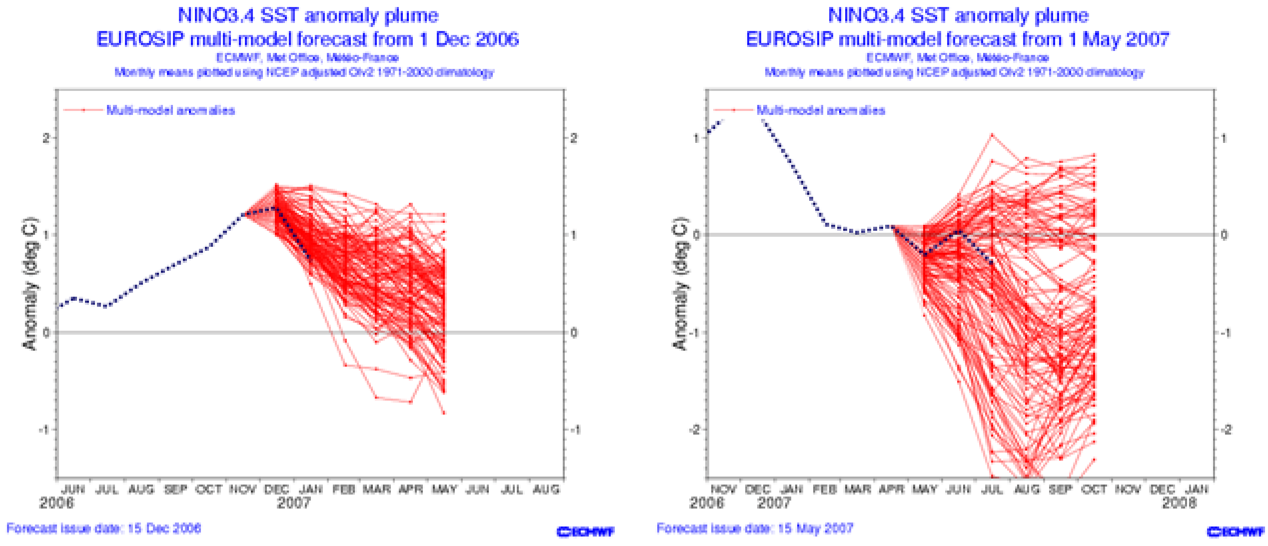

Figure 6 Spaghetti plots illustrating low to moderate and very high dispersion. *Examples provided by Jean-Pierre Ceron, Meteo France.*

3.4.2.2 Continuous confidence intervals and fan graphs

Continuous confidence intervals and fan graphs provide a way of illustrating confidence levels over time in a continuous manner. In a fan graph different levels of colour saturation are typically used to depict the range of values covered by 50, 70, and 90% confidence levels (with some also indicating a 100% limit). In some visualisations just the upper and lower bounds of a 90 or 95% confidence interval alone are shown, with a line indicating a measure of central tendency. In Figure 7 below a fan graph is used to represent the spread of a climatological reference (i.e. historical observations) for a seasonal temperature forecast.

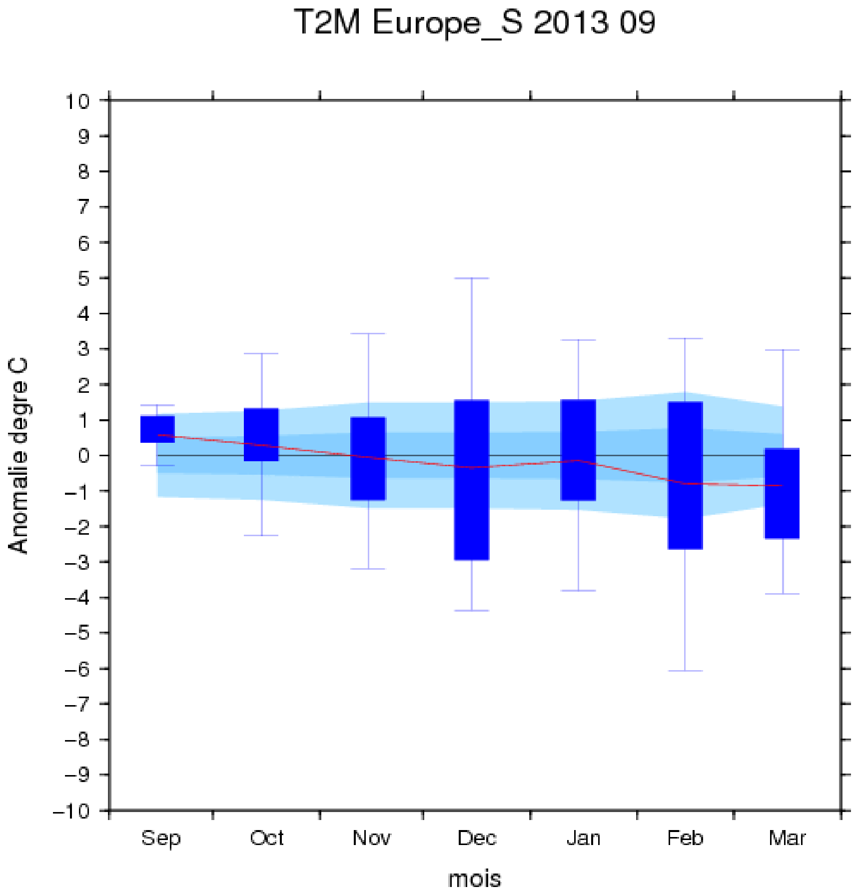

Figure 7 Example of a fan graph representing the range of a climate reference (i.e. historical observations) overlaid with boxplots representing range of a seasonal prediction. *Visualisation provided by Jean-Pierre Ceron, Meteo France.*

More complex diagrams utilising the same underlying format may also be constructed. In Figure 8 below multiple areas of concentration are depicted on the vertical axis. This highlights the fact that the distribution has multiple peaks, an element that is obscured in the standard fan graph, creating the impression of a unipolar distribution. However, this increase in complexity may bring with it a loss in salience. Once again the issue of richness versus ease of understanding has to be considered. Those with good statistical understanding for whom the recognition of multiple peaks has operational use may find this level of detail helpful. On the other hand those with lower statistical knowledge, or those simply seeking to establish upper and lower thresholds, may find that it renders visualisations difficult to interpret. Another factor to consider is, of course, the degree of detail that can be reasonably supported by the forecasting system. A very high degree of detail may lead users to overestimate the level of resolution offered by the system. Once again, it is important to consider both user expertise and context of use.

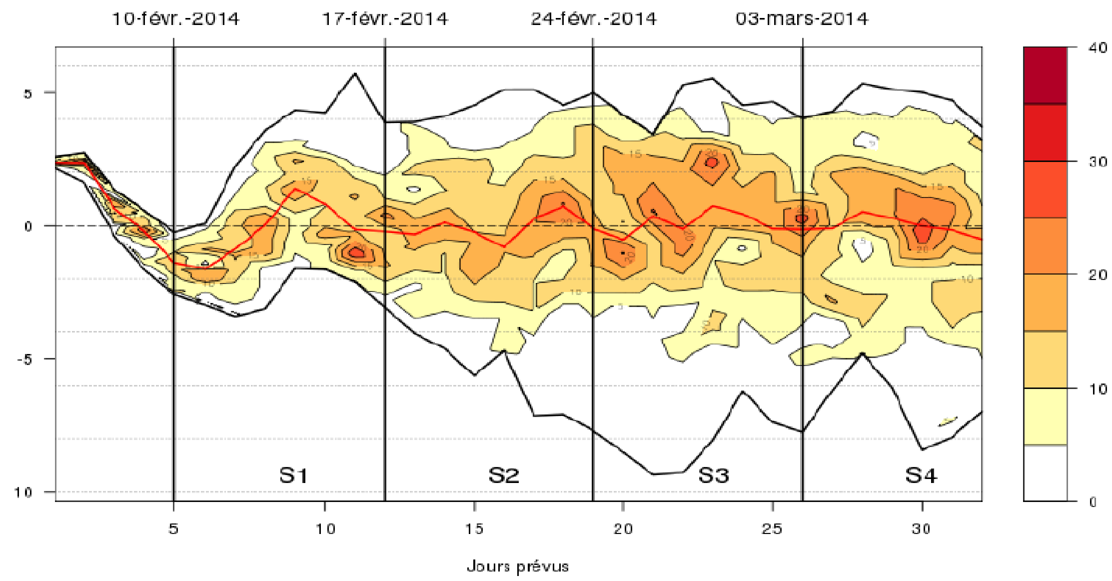

Figure 8 “Anomaly temperature forecasts issued from ECMWF VarEPS/Monthly system”
Visualisation and footnotes provided by Laurent Dubus, EDF.

Of course, while fan graphs can be used to depict confidence levels continuously, other methods may be used to present them for discrete timeframes.

3.4.2.3 Probability density functions

When depicting the distribution of a single variable for a discrete timeframe probability density functions (pdf) and cumulative density functions (cdf) offer the greatest amount of detail and can facilitate the comparison of distributions (e.g. historical observations versus current prediction) . As previously discussed however, research suggests that users with low to moderate statistical experience may have difficulty in extracting certain information, such as measures of central tendency, from these representations when they are not explicitly marked (Ibrekk & Morgan, 1989). Nonetheless, they are commonly used by those with greater technical expertise and statistical knowledge. Figure 9 illustrates how pdf’s may be used to compare climate predictions for a given period with climatology.

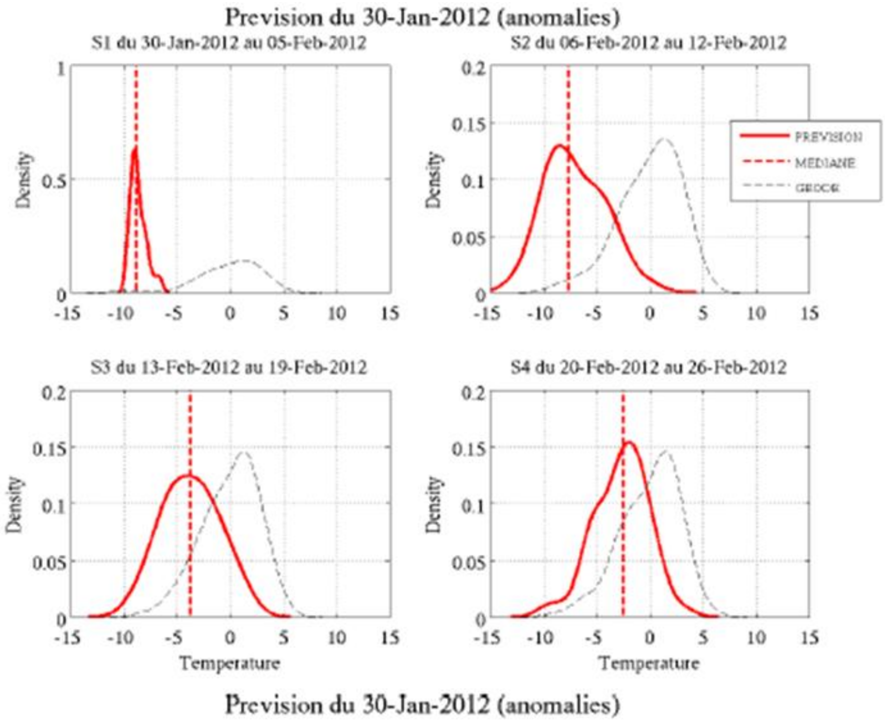

Figure 9 "Individual weeks' pdf of the 51-member forecasts (red curve), with forecast median (dash red line) and climatology (black dashed line)" *Visualisation and footnote provided by Laurent Dubus, EDF*

3.4.2.4 Error bars, boxplots and dots

Error bars are typically used to represent a single 90% or 95% confidence interval, with the mean marked as a measure of central tendency. Hence, the detail regarding the distribution is limited. Although, depending on the needs of the individual user, this level of simplicity may be desired. Boxplots contain more detail, indicating 50% and 95% (or 100%) confidence levels, with a measure of central tendency (usually the median) marked. Figure 8 above depicts a set of boxplots (representing predictions) superimposed on a fan graph (representing climatology).

Dots may also be used to display the distribution of ensemble members for a particular timeframe, with each dot representing an ensemble member. This method is used in Figure 10 and 11 below to illustrate the spread of ensemble members relative to the spread of historical (i.e. 30 year) observations. These visualisations, created as part of a suite of for UK contingency planners, use the y-axis to depict absolute measures (i.e. predicted and observed millimetres of precipitation) rather than anomaly. Figure 11 is a visually complex diagram, containing dot spreads and pdfs, both of which are superimposed on a background representing quintiles. Again, this raises the issue of the richness/simplicity trade-off. A rich and complex diagram is likely to be more useful those with greater existing statistical knowledge and graph literacy than those without. It is important to stress however that the UK Met Office provide detailed instructions (in both written and video form) as to how these visualisations can and should be used. Hence, the question of how much training users will be given with respect to utilising visualisations also needs to be asked. In contexts where more training can be provided, richer and less immediately intuitive representations may be more appropriate than in situations where less guidance is possible.

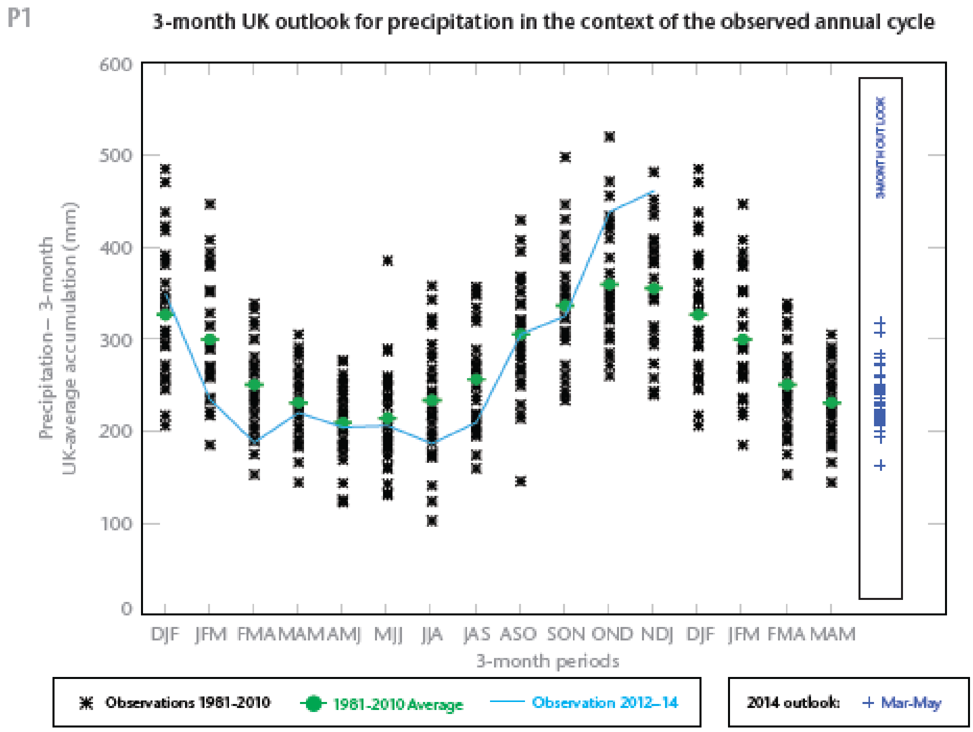

Figure 10 "3-month UK outlook for precipitation in the context of the observed annual cycle" *Visualisation for contingency planners reproduced from the Met Office website* ***(http://www.metoffice.gov.uk/publicsector/contingency-planners)***

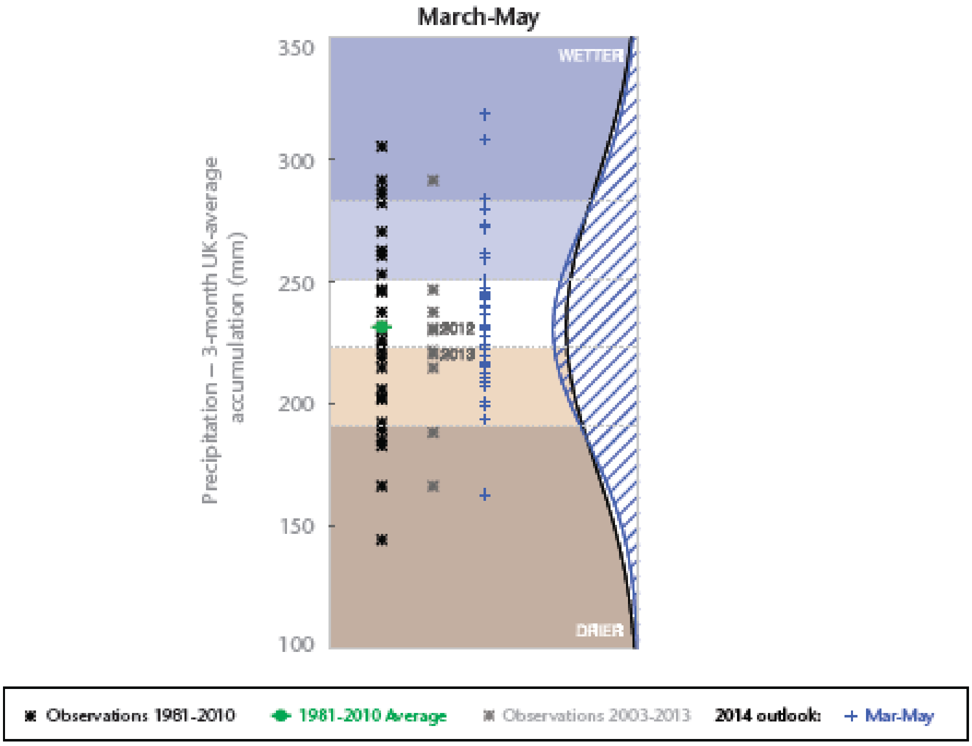

Figure 11 "3-month UK outlook for precipitation in the context of observed climatology" *Visualisation for contingency planners reproduced from the Met Office website* ***(http://www.metoffice.gov.uk/publicsector/contingency-planners)***

3.4.3 Reliability and skill

This section has thus far predominantly focussed on methods of visualising likelihoods, magnitudes and dispersion in seasonal climate predictions. However, in addition to effectively communicating ‘first order’ uncertainty derived from model output, it is also important that measures of reliability and skill are appropriately conveyed. In the recent Work Package 33 survey a substantial minority of respondents currently using seasonal predictions indicated that they did not currently receive information as to how predictions compared to observations but would like to do so. A desire to receive information about how different models compared with one another was also expressed. This indicates a pressing need for information regarding reliability and skill to be clearly communicated. However, doing this in a manner that can be easily understood by all users may pose a challenge. Numerical measures of both reliability and skill may be calculated (e.g. as skill scores). For less technical users however interpreting them may be difficult unless evaluative categories and visual aids are also provided.

3.4.3.1 Visualising skill and reliability using graphs

One way of visually representing reliability is to plot prediction against observation on a line graph (e.g. Figure 12). Here a perfectly reliable prediction would form a straight 45^{o} line. That is to say that where this perfectly reliable system indicates that there is a 60% chance of temperatures being lower than average, lower than average temperatures would occur 60% of the time that this prediction is made. The same information may also be plotted on a bar graph (Figure 13). Skill may be represented using similar formats, with the reliability of one prediction being plotted against a reference prediction.

For those lower in statistical understanding or graph literacy however such formats may prove difficult to interpret. Likewise, those more familiar with determinist forecasts (where predictions are either correct or incorrect) than probabilistic ones may also initially fail to realise that where a prediction system assigns a 60% probability to the most likely tercile occurring, the most likely tercile should not be observed 100% of the time. Again, the need to ascertain where misunderstandings may occur and address them should be emphasised.

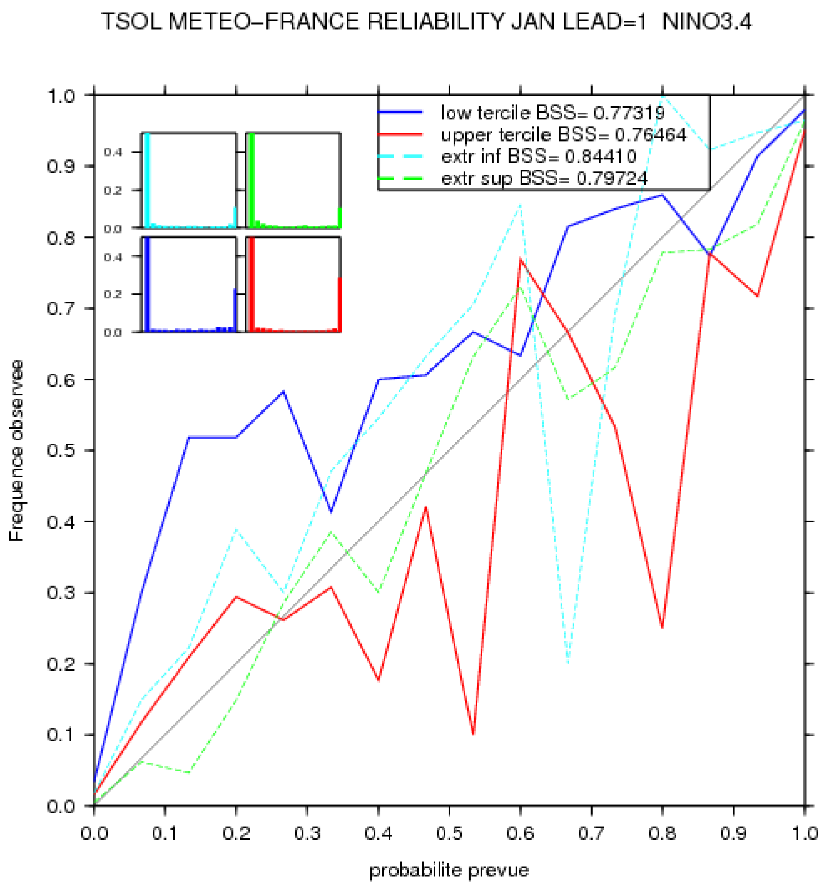

Figure 12 Reliability diagram plotting observed frequency against predicted probability. The grey line represents a hypothetical perfectly reliable forecast. *Provided by Jean-Pierre Ceron, Meteo France.*

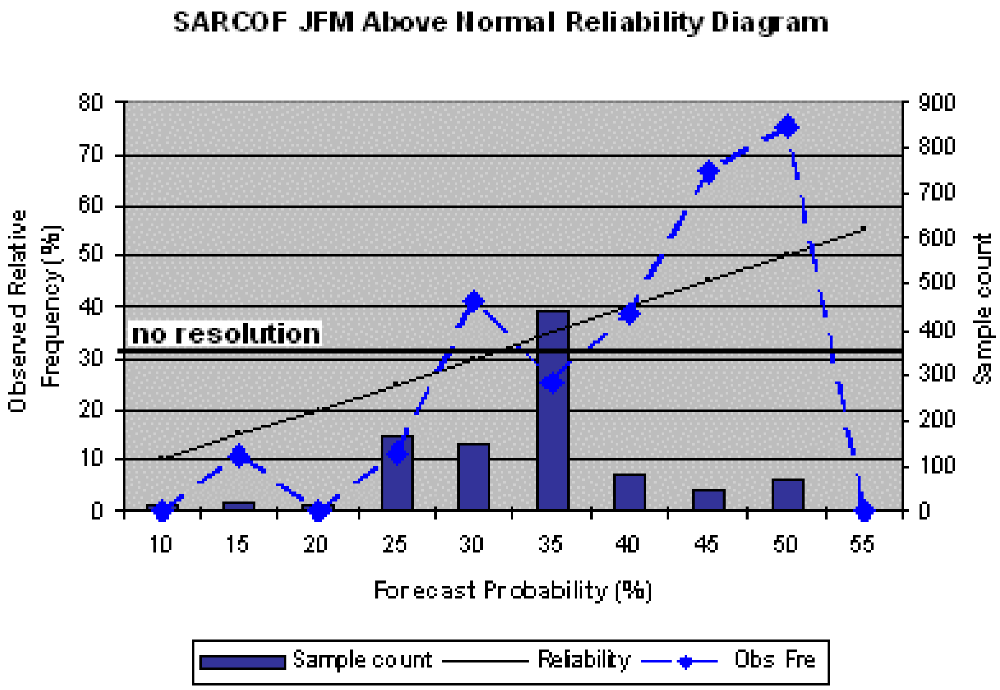

Figure 13 Reliability diagram plotting observed relative frequency against forecast probability. *Provided by Jean-Pierre Ceron, Meteo France.*

A visual comparison of predicted versus observed climate variables may be facilitated by overlaying a line (or dots) representing observations on a fan graph representing model predictions (i.e. the spread of ensemble members). A similar format may also be used to compare the performance of several different models across a particular timescale (Figure 14). Although it should be noted that in the latter case increasing visual complexity in this manner may make interpretation more difficult for less experienced users. Again, the trade-off between increasing informational content and potentially reducing salience should be considered in view of both who the users are and the amount of information required to support decision making.

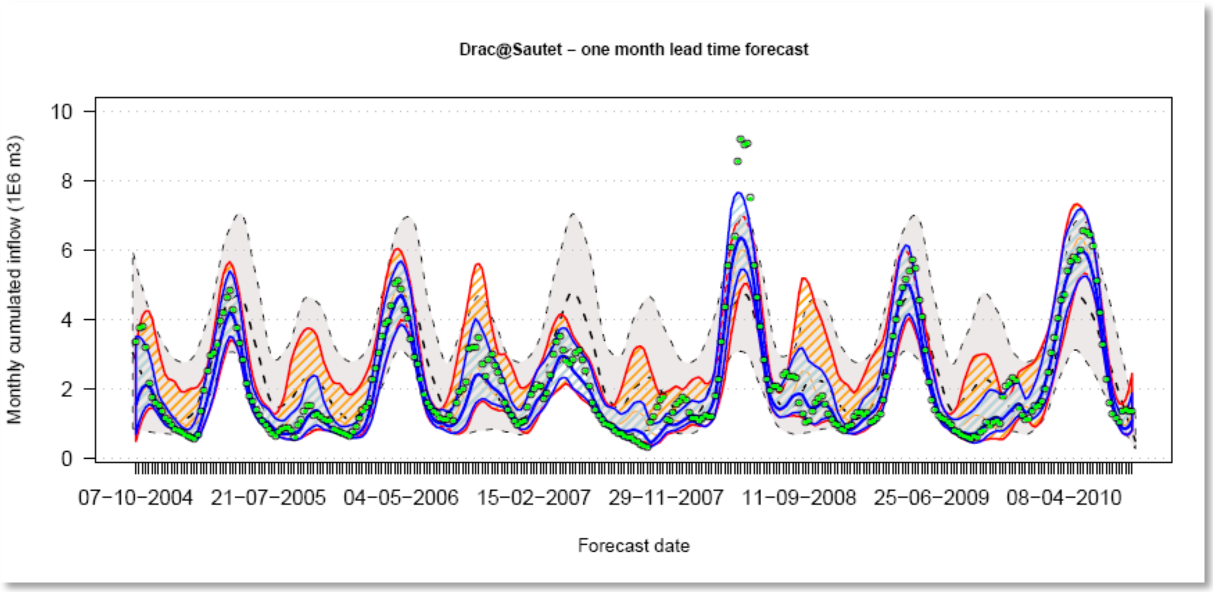

Figure 14 River Flow Monthly Forecasts. A comparison of different types of probabilistic forecasts plotted against observations (represented by green dots. *Visualisation provided by Laurent Dubus, EDF.*

3.4.3.2 Visualising reliability and skill using maps

In addition to conveying predictions themselves, maps may also be used to convey information regarding a) how well predictions have matched observed climate (accuracy in the case of deterministic forecasts reliability the case of probabilistic ones); and b) how a particular model has performed comparative to climatology or another reference forecast (skill). This may be done using colour (Figure 15), transparency (Figure 16), or presenting maps detailing ‘prediction’ and ‘observation’/‘reference’ side by side (Figure 17, see also Figure 4b). No definitive guide as to which particular format works best is however available in the current literature. Once again, more research is needed to establish which visual representations of reliability and skill work best in which context.

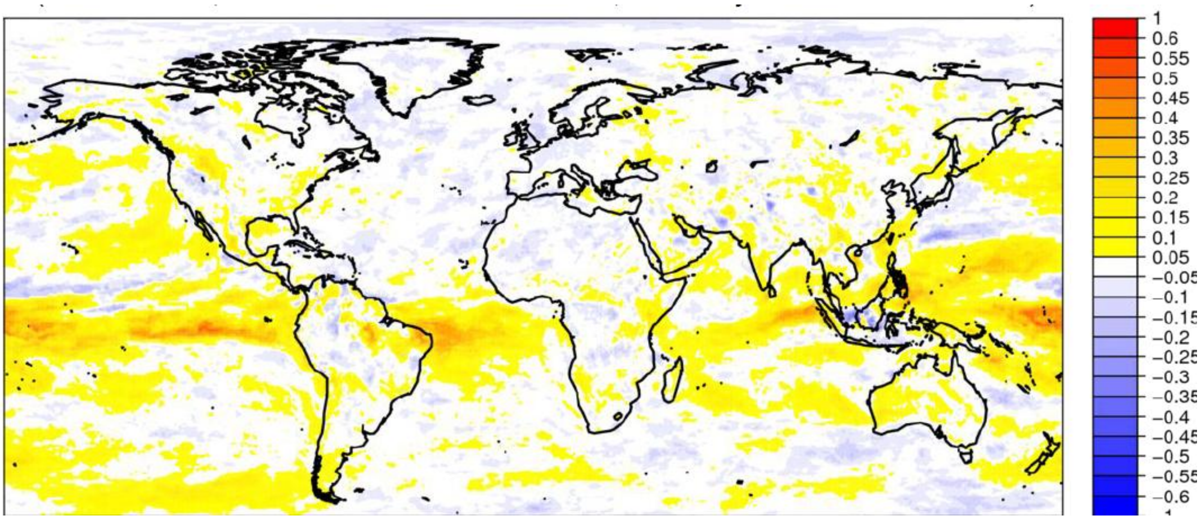

Figure 15 "Spring 10m wind resource CR probability skill score (ECMWF S4, 1 month forecast lead time, once a year from 1981-2010)" *Image and footnote provided by IC3.*

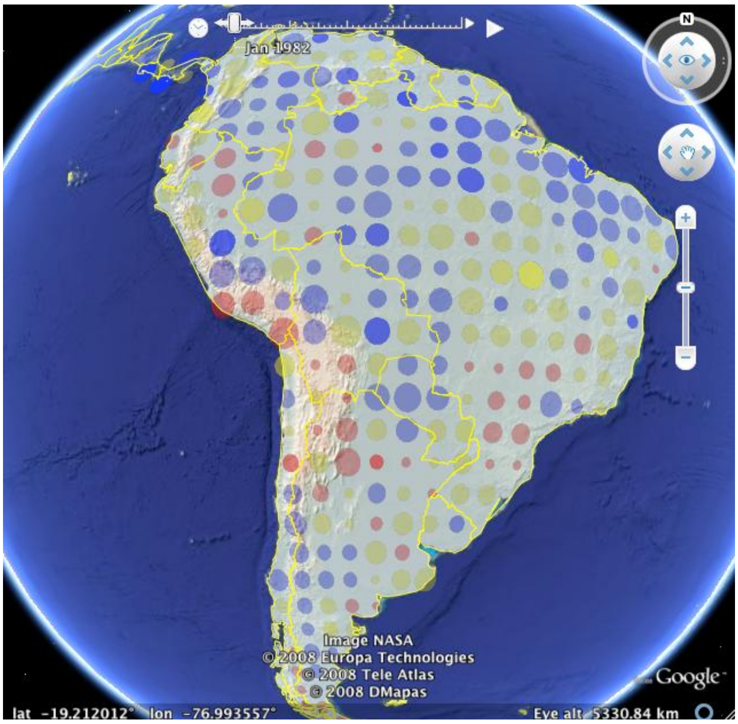

Figure 16 "Transparency used to indicate the skill (quality) of the forecast using the Brier Skill Score (BSS). High transparency indicates lower skill." *Image and footnote reproduced from Slingsby et al. (2009)*

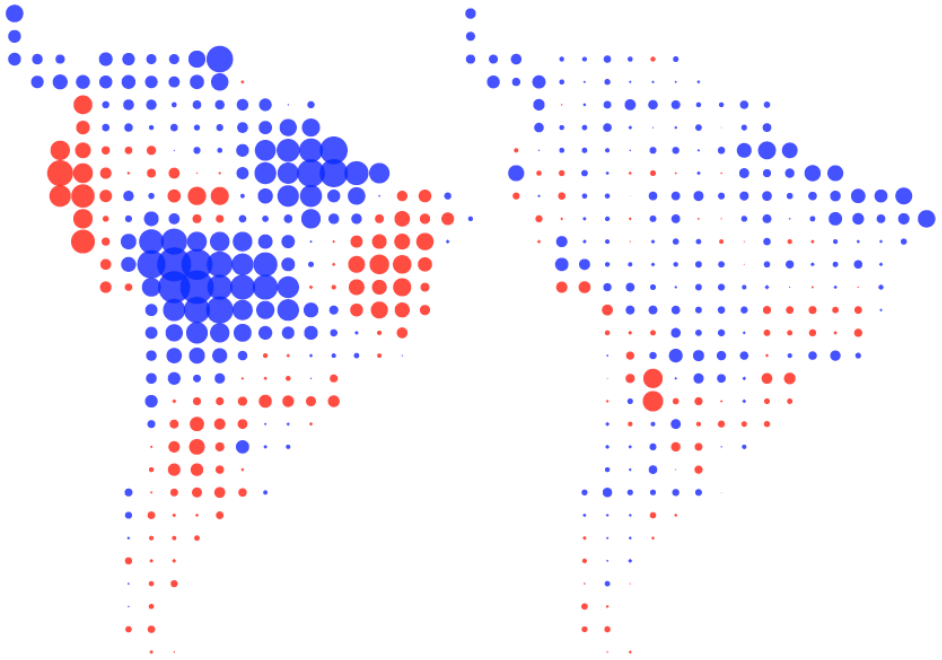

Figure 17 "Observed (left) and mean forecast (right) precipitation anomalies (red circles are negative anomalies, blue are positive – size of circle indicates magnitude of anomaly)." *Image and footnote reproduced from Slingsby et al. (2009)*

3.4.4 Decision aids and user tools

As discussed earlier in this review, some users of climate information may favour visualisations that a) explicitly indicate user-defined thresholds; and b) facilitate decision making. Received Operating Characteristic (ROC) curves provide one method of integrating organisational risk preference and tolerance for false alarms into the use of climate predictions by plotting true hits against false alarms at different levels of sensitivity (see Figure 18 for example of a temperature forecast presented in this way and Figure 19 for an explanation as to how ROC diagrams should be interpreted). Hence, as sensitivity increases the likelihood of both a true positive or false positive increases (generally in a non-linear manner). A 45^{o} line on the graph would therefore represent a prediction system for which a positive signal has an equal chance of being a true positive or a false positive. Decision makers may thus choose the point on the curve that best reflects their tolerance for false alarms and false misses. Again however, those not already familiar with such representations may struggle to utilise them effectively. Indeed, the visual similarity between ROC curves and line graphs representing reliability measures mean that some may confuse the quite different method of interpretation needed to accurately extract the meaning (i.e. on a reliability graph a 45^{o} line would indicate a perfect forecast, while on a ROC curve it indicates a forecast which returns as many false positives as it does true positives).

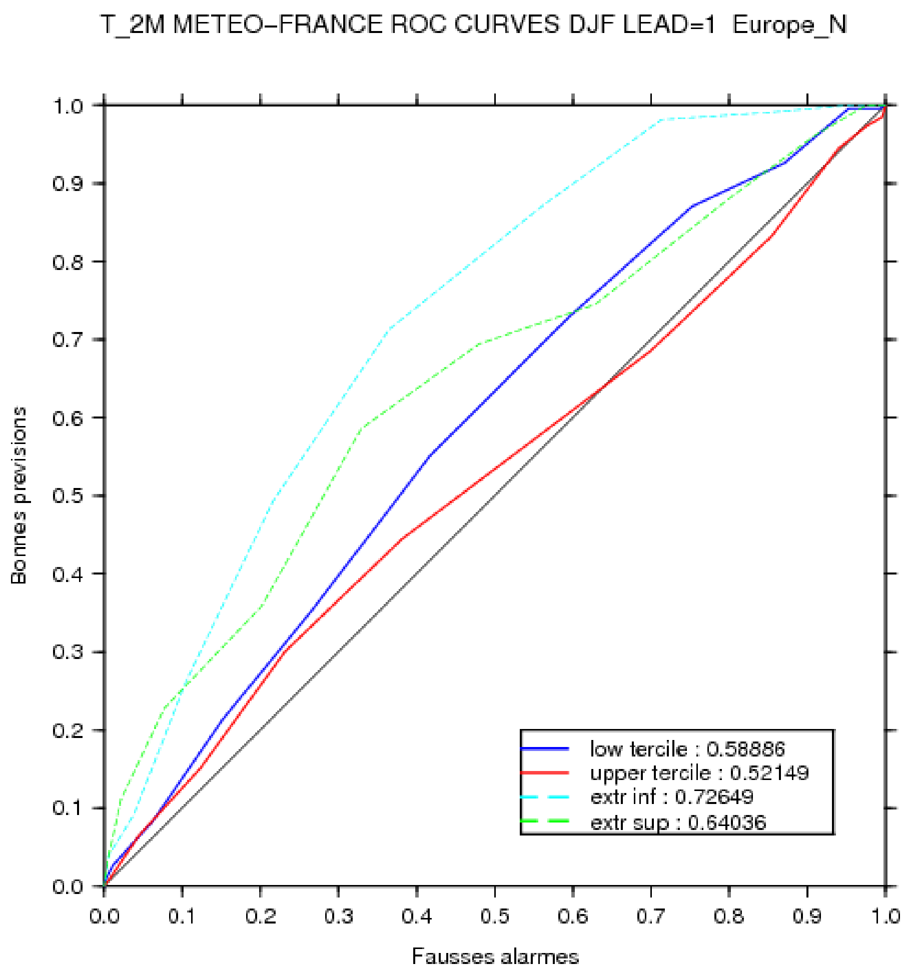

Figure 18 Example of ROC curves. Likelihood of a false alarm is plotted on the x-axis and likelihood of a true detection on the y-axis. *Visualisation provided by Jean-Pierre Ceron, Meteo France.*

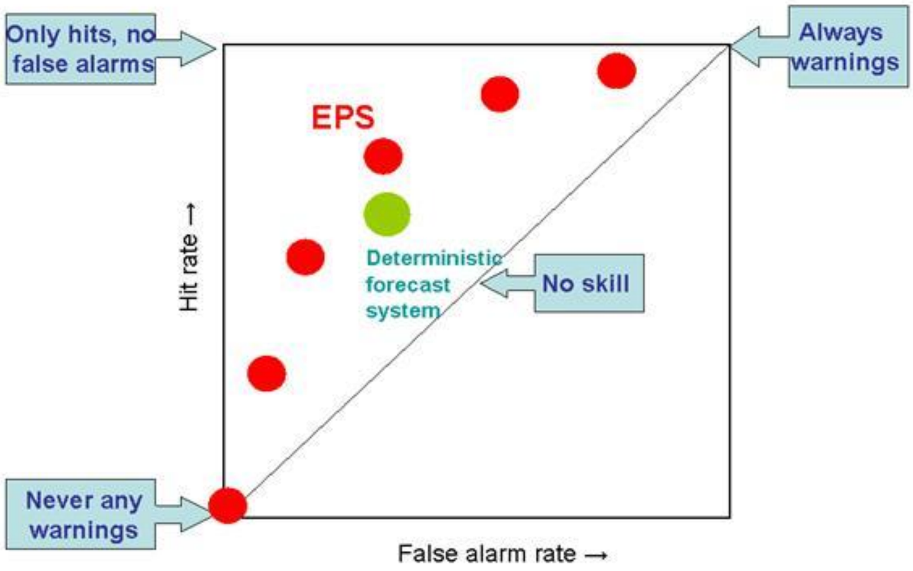

Figure 19 Illustration of how the space on a ROC diagram should be read. At the **bottom-left** no warnings are ever given (i.e. no hits or false alarms), while at the **top-right** warnings are always given regardless of whether the event will occur or not. A score in the **top-left** would denote a perfect prediction system with a hit rate of 100% and a 0% false alarm rate. Along the grey line cutting diagonally across the diagram a warning is equally likely to denote a hit or a false alarm. In this instance the diagram depicts a deterministic forecast as a single fixed point. However, as can be seen in Figure 18, curves can be added to indicate a hit-to-false alarm ratio at different levels of sensitivity. *Diagram reproduced from the European Centre for Medium-Range Weather Forecasts (ECMWF) website* ***(http://www.ecmwf.int/products/forecasts/guide/The_relative_operating_characteristics_ROC_diagram.html)***

Of course, while ROC curves may prove difficult for less statistically experienced users, this underlying information regarding the ratio of false positives to false negatives may nonetheless be integrated with user risk preferences in more user friendly tools that facilitate specific decisions. Various tools for specifying the crossing thresholds or level of preparedness required exist. Some of these utilise a 'traffic light' system to indicate threat level or need for action. Examples of such systems include the Met Office's [4] extreme weather warnings (Neal, Boyle, Grahame, Mylne, & Sharpe, 2013) and Meteo France's vigilance maps[5]. Here level of threat (represented on a green to red traffic light scale) represents a measure of risk (potential magnitude of impact weighted by likelihood).

Figure 20 below illustrates how a similar format can be used to translate probabilistic forecasts into actionable information in an operational context. For this hydrological forecast, the likelihood of river discharge exceeding a critical threshold is coded as green, yellow, orange or red, with the colours serving to provide a visual cue as to degree of preparedness/action required.

[4] http://www.metoffice.gov.uk/public/weather/warnings/#?tab=warnings&regionName=uk
[5] http://vigilance.meteofrance.com/

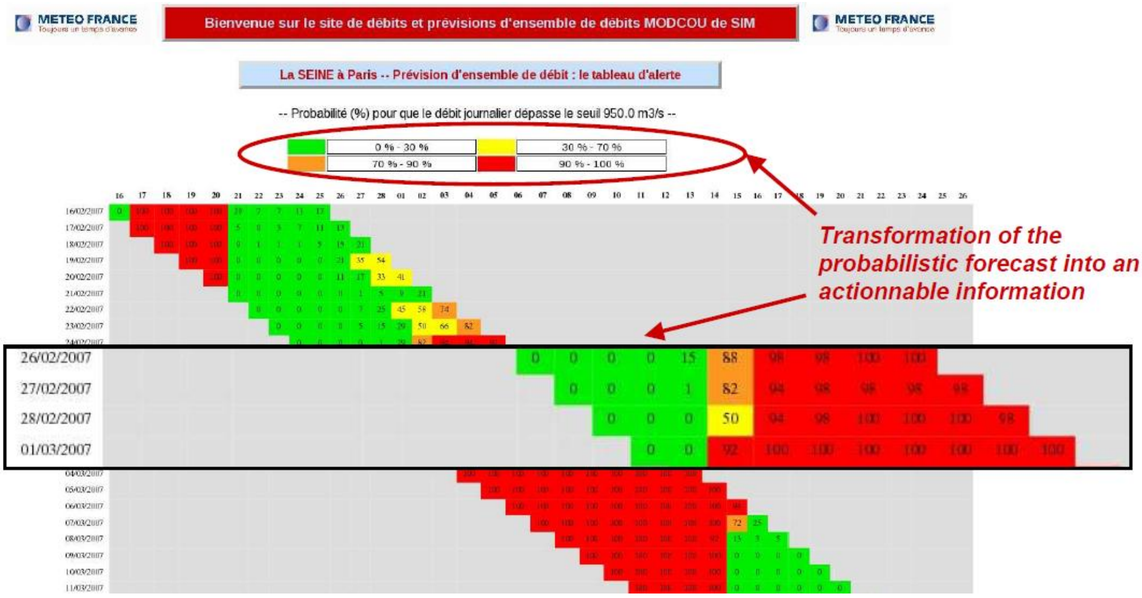

Figure 20 Traffic light coding used to indicate likelihood of river discharge crossing critical threshold. *Visualisation provided by Jean-Pierre Ceron, Meteo France.*

Recently a grid system for matching visualisations (and other information formats) to user needs has been devised by Meteo-France (see Figure 21 over page). Here visualisations may be selected depending on the particular information required. Such systems may represent a way of catering to a diverse range of users.

Analysis Grid for uncertainty presentation

Parameters		Mean Temperature	Minimal Temperature	Maximal temperature	Rainfall	Wind	Humidity	Impact parameters			
Spatial scale		Planet	Region of the globe	Continent	Country / State	Administrative region	Specific area	River Catchment	Town	Grid point	Oceanic Basin
Time Scale	**Period**	Predefined horizon	Year (continuous)	Season	Month (or severals)	Day (or severals)					
	Frequency	Annual	Seasonal	Monthly	Daily						
Models		One model (Global or Regional)	Multi-models	Nested models	Impact models						
Scenarios		One scenario	Ensemble of Scenarios	Ensemble of Simulations							
Statistics		Mean	Statistical parameters (Mediane, quantiles, Std, ...)	PDF	Differences (Model - Clim or Model1 - Model2)	Confidence interval	Significance				
Uncertainty type		Scenario uncertainty	Model uncertainty	Climate system uncertainty							
Graphics		Maps	Whisker plots	Bar charts	Plumes	Pdf (continuous)	Time serie	2D diagrams (pdf, ΔRR and ΔT)			

Figure 21 Analysis Grid. *Reproduced from Chateigner (2013)*

Key Points: Current and proposed representations of uncertainty in seasonal climate predictions

- Maps may represent probabilistic information in a number of ways (e.g. two-category maps, separate maps for different terciles/quintiles, maps depicting probability of most likely tercile, deterministic maps using strippling effects). However, the question of which formats are best understood and most useful to a range of end users requires further investigation.

- When constructing maps and other visualisations one should carefully consider one's use of colour in order to avoid a) counterintuitive representations (e.g. blue corresponding with higher temperatures or lower precipitation); and b) the use of more hues than necessary.

- Probability distributions and other spread may be communicated in a number of ways. The appropriate format to use will depend a) on whether a temporally continuous measure is desired (or appropriate); and b) whether users are concerned with potential extremes, central tendencies or more detailed distributions.

- End users may struggle to accurately interpret complex visualisations without appropriate training. Hence, the question of whether training will be provided should be considered when developing visualisations.

- More work is needed to establish how information regarding reliability and skill can best be communicated to end users.

- Decision aids providing Act/Don't Act signals and thresholds for action based on a combination of model output and user risk preference may simplify use of predictions. Although care should of course be taken to ensure that such formats do not create a false sense of certainty.

3.5 Future directions

Looking forward to future steps in Work Package 33, this review highlights both points for consideration when designing methods of communicating uncertainty in the context of seasonal to decadal climate prediction (Task 33.3) and questions to be addressed when examining their fitness for purpose (Task 33.4).

As touched upon in Section 3.2, the level of detail required by end users may vary from case to case: with some wishing for an outline of plausible scenarios and others a full statistical treatment of uncertainty. Institutions may also differ in their tolerance for uncertainty; with some rejecting new information containing high 'ambiguity' regarding likelihoods and magnitudes. In the Task 33.1 user needs survey responses indicated that most respondent organisations had at least some tolerance for uncertainty. However, reported tolerance for false alarms varied considerably (see Taylor and Dessai, 2014, for full report). Preference for information formats that facilitate Yes/No decision making also varied amongst respondents. Hence, these differences in organisational preferences need to be taken into account when devising methods of communication. The findings of the T33.1 survey also indicated that while many respondents were comfortable with using statistical measures of spread (e.g. confidence intervals and standard deviations) others were less so. Hence, the challenge of how probabilistic information can best be communicated to users with a range of statistical experience is one that should be addressed.

As has previously been stressed, the importance of testing methods of information presentation before use is paramount. Thus, the efficacy of verbal, numeric and visual methods of presentation (and combinations thereof) in communicating uncertainty information should be established; as should any systematic biases that may lead to the misinterpretation of uncertainty information. Specific questions that should be examined include:

- Are users' preferred methods of receiving uncertainty information in this domain those that they interpret most accurately?

- Do framing effects influence the interpretation of information in existing methods of presenting uncertainty information in S2D forecasts? If so, how can this be mitigated? (E.g. by presenting information in both 'event' and 'non-event' frames).

- Where uncertainty language (e.g. 'likely', 'unlikely', 'high confidence', 'low confidence') is used, is it interpreted in the manner intended by the communicator? (E.g. do users interpret the phrase 'likely' to cover the same range of likelihoods as is intended?) Does presenting calibrated language with numeric and visual representations of likelihood and spread enhance comprehension?

- How does the size of numeric ranges (displayed numerically or graphically) influence user perceptions of reliability, credibility and usefulness?

- To what extent do graph literacy and numeracy influence the comprehension of numeric and visual representations of uncertainty in seasonal to decadal climate predictions? Which visual and numeric representations are best comprehended (and perceived as most useful) by recipients possessing different levels of statistical knowledge?

- How can information regarding reliability and skill be presented in a way that best facilitates understanding? Is this information better understood when it is presented separately from the probabilistic output of the prediction model, or when both types of information are integrated (e.g. using a visual cue such as transparency to indicate skill).

3.6. Concluding remarks

The review has drawn upon research in a diverse range of fields to provide an overview of how information regarding uncertainty can be disseminated and, of course, interpreted. The words ‘confidence’ and ‘uncertainty’ can be defined in a number of ways. Hence, if using these terms when communicating with the recipients of ‘uncertainty information’, it would seem to be important to clearly state what is meant by them. The problem of ‘linguistic uncertainty’ more generally is, as has been seen, one that must be confronted when communicating with information recipients. This is especially the case if one is seeking to convey likelihoods, magnitudes, level of evidence or consensus using verbal descriptors. Explicitly calibrating this language with numeric estimates is one solution to this difficulty, though this may not be suitable if the nature of the uncertainty in question cannot be readily quantified. Hence, the thorough (and context specific) testing of any such framework prior to use is thus strongly recommended. Likewise, the rigorous testing of how numeric formats and visual representations are perceived by users is vital to ensuring that a) information is understood; and b) framing effects and other potential biases are minimised. As users (and potential users) of S2D forecasts are likely to vary in terms of technical background and experience with probabilistic information, factors such as graph literacy and numeracy should be taken into account. A range of interesting visualisations for communicating first order and second order uncertainty in seasonal climate predictions exist. However, more systematic testing is needed to determine which formats suit which type of user, and whether new forms of representation (or adjustment to existing ones) can increase ease of understanding and reduce misinterpretation.

3.7 References

Allen, M. S., & Eckel, F. A. (2012). Value from Ambiguity in Ensemble Forecasts. *Weather & Forecasting, 27*(1).

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., . . . Norris, S. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of clinical epidemiology, 64*(4), 401-406.

Bazerman, M., & Moore, D. A. (2012). Judgment in managerial decision making.

Bearman, N. (2011). *Using sound to represent uncertainty in future climate projections for the United Kingdom.* Paper presented at the Proceedings of the 17th International Conference on Auditory Display (ICAD 2011). Budapest, Hungary. CD-ROM.

Berry, D., Raynor, T., Knapp, P., & Bersellini, E. (2004). Over the counter medicines and the need for immediate action: a further evaluation of European Commission recommended wordings for communicating risk. *Patient education and counseling, 53*(2), 129-134.

Broad, K., Leiserowitz, A., Weinkle, J., & Steketee, M. (2007). Misinterpretations of the “cone of uncertainty” in Florida during the 2004 hurricane season. *Bulletin of the American Meteorological Society, 88*(5), 651-667.

Brown, E., & Bearman, N. (2012). Listening to uncertainty: Information that sings. *Significance, 9*(5), 14-17.

Budescu, D. V., Broomell, S., & Por, H.-H. (2009). Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. *Psychological science, 20*(3), 299-308.

Campbell, P. (2011). Understanding the receivers and the reception of science's uncertain messages. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369*(1956), 4891-4912.

Carey, J. M., & Burgman, M. A. (2008). Linguistic uncertainty in qualitative risk analysis and how to minimize it. *Annals of the New York Academy of Sciences, 1128*(1), 13-17.

Christensen, F. M., Andersen, O., Duijm, N. J., & Harremoës, P. (2003). Risk terminology—a platform for common understanding and better communication. *Journal of Hazardous Materials, 103*(3), 181-203.

Daron, J (2014). Visualising Climate Information: Survey Results. http://www.youtube.com/watch?v=FQQJLD-2qmI

de Bruin, W. B., & Bostrom, A. (2013). Assessing what to address in science communication. *Proceedings of the National Academy of Sciences, 110*(Supplement 3), 14062-14068.

Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J., & Ramos, M.-H. (2007). Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards, 7*(2), 115-127.

Demeritt, D., Nobert, S., Cloke, H., & Pappenberger, F. (2010). Challenges in communicating and using ensembles in operational flood forecasting. *Meteorological Applications, 17*(2), 209-222.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: when people behave against their better judgment. *Journal of personality and social*

psychology, 66(5), 819.
Dessai, S., & Hulme, M. (2004). Does climate adaptation policy need probabilities? *Climate policy, 4*(2), 107-128.
Dessai, S., Hulme, M., Lempert, R., & Pielke Jr, R. (2009). Climate prediction: a limit to adaptation. *Adapting to climate change: thresholds, values, governance*, 64-78.
Dettinger, M. D. (2005). From climate-change spaghetti to climate-change distributions for 21st Century California. *San Francisco Estuary and Watershed Science, 3*(1).
Dieckmann, N. F., Mauro, R., & Slovic, P. (2010). The effects of presenting imprecise probabilities in intelligence forecasts. *Risk analysis, 30*(6), 987-1001.
Edwards, J. A., Snyder, F. J., Allen, P. M., Makinson, K. A., & Hamby, D. M. (2012). Decision making for risk management: A comparison of graphical methods for presenting quantitative uncertainty. *Risk analysis, 32*(12), 2055-2070.
Ellsberg, D. (1990). 4. Risk, ambiguity, and the Savage axioms. *Rationality in Action: Contemporary Approaches*, 89.
Environmental Decisions in the Face of Uncertainty. (2013). The National Academies Press.
Galesic, M., & Garcia-Retamero, R. (2011). Graph Literacy A Cross-Cultural Comparison. *Medical Decision Making, 31*(3), 444-457.
Gibson, J. M., Rowe, A., Stone, E. R., & Bruine de Bruin, W. (2013). Communicating Quantitative Information About Unexploded Ordnance Risks to the Public. *Environmental science & technology, 47*(9), 4004-4013.
Gigerenzer, G. (2003). *Reckoning with risk: learning to live with uncertainty*: Penguin UK.
Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review, 103*(4), 650.
Gigerenzer, G., Hertwig, R., Van Den Broek, E., Fasolo, B., & Katsikopoulos, K. V. (2005). “A 30% chance of rain tomorrow”: How does the public understand probabilistic weather forecasts? *Risk analysis, 25*(3), 623-629.
Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological review, 102*(4), 684.
Goodie, A. S., & Fantino, E. (1996). Learning to commit or avoid the base-rate error.
Goodwin, P., Sinan Gönül, M., & Önkal, D. (2013). Antecedents and effects of trust in forecasting advice. *International Journal of Forecasting, 29*(2), 354-366.
Gregory, R., Dieckmann, N., Peters, E., Failing, L., Long, G., & Tusler, M. (2012). Deliberative disjunction: Expert and public understanding of outcome uncertainty. *Risk analysis, 32*(12), 2071-2083.
Han, P. K. (2013). Conceptual, methodological, and ethical problems in communicating uncertainty in clinical evidence. *Medical Care Research and Review, 70*(1 suppl), 14S-36S.
Han, P. K., Klein, W. M., Lehman, T., Killam, B., Massett, H., & Freedman, A. N. (2011). Communication of uncertainty regarding individualized cancer risk estimates effects and influential factors. *Medical Decision Making, 31*(2), 354-366.
Handmer, J., & Proudley, B. (2007). Communicating uncertainty via probabilities: The case of weather forecasts. *Environmental Hazards, 7*(2), 79-87.
Hauge, K. H., Nielsen, K. N., & Korsbrekke, K. (2007). Limits to transparency—exploring conceptual and operational aspects of the ICES framework for providing

precautionary fisheries management advice. *ICES Journal of Marine Science: Journal du Conseil, 64*(4), 738-743.
Hibbard, J. H., & Peters, E. (2003). Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. *Annual Review of Public Health, 24*(1), 413-433.
Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine, 73*(5), 538-540.
Ibrekk, H., & Morgan, M. G. (1987). Graphical communication of uncertain quantities to nontechnical people. *Risk analysis, 7*(4), 519-529.
Joslyn, S., & Savelli, S. (2010). Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteorological Applications, 17*(2), 180-195.
Joslyn, S. L., & LeClerc, J. E. (2012). Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of experimental psychology: applied, 18*(1), 126.
Joslyn, S. L., & Nichols, R. M. (2009). Probability or frequency? Expressing forecast uncertainty in public weather forecasts. *Meteorological Applications, 16*(3), 309-314.
Jupp, T. E., Lowe, R., Coelho, C. A., & Stephenson, D. B. (2012). On the visualization, verification and recalibration of ternary probabilistic forecasts. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370*(1962), 1100-1120.
Kahneman, D. (2011). *Thinking, fast and slow*: Macmillan.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.
Kandlikar, M., Risbey, J., & Dessai, S. (2005). Representing and communicating deep uncertainty in climate-change assessments. *Comptes Rendus Geoscience, 337*(4), 443-455.
Kaye, N., Hartley, A., & Hemming, D. (2012). Mapping the climate: guidance on appropriate techniques to map climate variables and their uncertainty. *Geoscientific Model Development, 5*(1), 245-256.
Kloprogge, P., Sluijs, J. P., & Wardekker, J. A. (2007). *Uncertainty communication: issues and good practice*: Copernicus Institute for Sustainable Development and Innovation, Utrecht University Utrecht.
Knight, F. H. (1921). Risk, uncertainty and profit. *New York: Hart, Schaffner and Marx*.
Lamberti, D., & Wallace, W. A. (1987). Presenting uncertainty in expert systems: an issue in information portrayal. *Information & Management, 13*(4), 159-169.
Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Medical Decision Making, 27*(5), 696-713.
Lofstedt, R. E. (2006). How can we make food risk communication better: where are we and where are we going? *Journal of Risk Research, 9*(8), 869-890.
Longman, T., Turner, R. M., King, M., & McCaffery, K. J. (2012). The effects of communicating uncertainty in quantitative health risk estimates. *Patient education and counseling, 89*(2), 252-259.
Lowe, R., Bailey, T. C., Stephenson, D. B., Jupp, T. E., Graham, R. J., Barcellos, C., &

Carvalho, M. S. (2013). The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil. *Statistics in medicine, 32*(5), 864-883.

Mastrandrea, M. D., Field, C. B., Stocker, T. F., Edenhofer, O., Ebi, K. L., Frame, D. J., . . . Matschoss, P. R. (2010). Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. *Intergovernmental Panel on Climate Change (IPCC)*.

McCown, R. (2012). A cognitive systems framework to inform delivery of analytic support for farmers' intuitive management under seasonal climatic variability. *Agricultural Systems, 105*(1), 7-20.

McCown, R., Carberry, P., Dalgliesh, N., Foale, M., & Hochman, Z. (2012). Farmers use intuition to reinvent analytic decision support for managing seasonal climatic variability. *Agricultural Systems, 106*(1), 33-45.

Morss, R. E., Demuth, J. L., & Lazo, J. K. (2008). Communicating uncertainty in weather forecasts: A survey of the US public. *Weather & Forecasting, 23*(5).

Morss, R. E., Lazo, J. K., & Demuth, J. L. (2010). Examining the use of weather forecasts in decision scenarios: results from a US survey with implications for uncertainty communication. *Meteorological Applications, 17*(2), 149-162.

Neal, R. A., Boyle, P., Grahame, N., Mylne, K., & Sharpe, M. (2013). Ensemble based first guess support towards a risk-based severe weather warning service. *Meteorological Applications*.

Nelson, D., Hesse, B., & Croyle, R. (2009). Making data talk: communicating public health data to the public, policy makers: and the Press.

Nicholls, N. (1999). Cognitive illusions, heuristics, and climate prediction. *Bulletin of the American Meteorological Society, 80*(7), 1385-1397.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making, 25*(4), 390-401.

Pappenberger, F., Stephens, E., Thielen, J., Salamon, P., Demeritt, D., Andel, S. J., . . . Alfieri, L. (2013). Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. *Hydrological Processes, 27*(1), 132-146.

Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety, 54*(2), 95-111.

Patt, A., & Dessai, S. (2005). Communicating uncertainty: lessons learned and suggestions for climate change assessment. *Comptes Rendus Geoscience, 337*(4), 425-441.

Patt, A. G., & Schrag, D. P. (2003). Using specific language to describe risk and probability. *Climatic change, 61*(1-2), 17-30.

Peters, E. (2008). Numeracy and the perception and communication of risk. *Annals of the New York Academy of Sciences, 1128*(1), 1-7.

Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: the impact of evaluative categories on decisions. *Journal of experimental psychology: applied, 15*(3), 213.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological science, 17*(5), 407-413.
Pidgeon, N., & Fischhoff, B. (2011). The role of social and decision sciences in communicating uncertain climate risks. *Nature Climate Change, 1*(1), 35-41.
Politi, M. C., Han, P. K., & Col, N. F. (2007). Communicating the uncertainty of harms and benefits of medical interventions. *Medical Decision Making, 27*(5), 681-695.
Ramos, M. H., Mathevet, T., Thielen, J., & Pappenberger, F. (2010). Communicating uncertainty in hydro-meteorological forecasts: mission impossible? *Meteorological Applications, 17*(2), 223-235.
Reyna, V. F. (2008). A theory of medical decision making and health: fuzzy trace theory. *Medical Decision Making, 28*(6), 850-865.
Risbey, J. S., & Kandlikar, M. (2007). Expressions of likelihood and confidence in the IPCC uncertainty assessment process. *Climatic change, 85*(1-2), 19-31.
Roulston, M. S., Bolton, G. E., Kleit, A. N., & Sears-Collins, A. L. (2006). A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather & Forecasting, 21*(1).
Roulston, M. S., & Kaplan, T. R. (2009). A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. *Meteorological Applications, 16*(2), 237-244.
Simon, H. A. (1957). Models of man; social and rational.
Slingsby, A., Lowe, R., Dykes, J., Stephenson, D., Wood, J., & Jupp, T. (2009). *A pilot study for the collaborative development of new ways of visualising seasonal climate forecasts.* Paper presented at the Proc. 17th Annu. Conf. of GIS Research UK, Durham, UK, 1â€“3 April 2009.
Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 544.
Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and human behavior, 24*(3), 271.
Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science, 333*(6048), 1393-1400.
Spiegelhalter, D. J., & Riesch, H. (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369*(1956), 4730-4750.
Stephens, E. M., Edwards, T. L., & Demeritt, D. (2012). Communicating probabilistic information from climate model ensembles—lessons from numerical weather prediction. *Wiley Interdisciplinary Reviews: Climate Change, 3*(5), 409-426.
Stirling, A. (2007). Risk, precaution and science: towards a more constructive policy debate. *EMBO reports, 8*(4), 309-315.
Stoverinck, F. (2011). Communication of climat through the decisio.
Stoverinck, F., Dubois, G., & Amelung, B. Communiquer le changement climatique: un voyage dans les stratégies et plans d’adaptation Communicating climate information:

travelling through the decision-making process.
Todini, E., Alberoni, P., Butts, M., Collier, C., Khatibi, R., Samuels, P., & Weerts, A. (2005). ACTIF best practice–understanding and reducing uncertainty in flood forecasting.
Taylor, A. and Dessai, S. (2014). Report on survey of end-user needs for improved uncertainty and confidence level information. http://euporias.eu/sites/default/files/deliverables/D33.1_Final.pdf
Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology, 5*(2), 207-232.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.
Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological science, 3*(5), 305-309.
U.S. Preventive Services Task Force. (2008). U.S. Preventive Services Task Force Grade definitions. Retrieved from http://www.uspreventiveservicestaskforce.org/uspstf/grades.htm
Vaessen, A. (2003). 1O Onzekerheid in de tekst. *Niet bang voor onzekerheid, 1*, 119.
Van Dijk, E., & Zeelenberg, M. (2003). The discounting of ambiguous information in economic decision making. *Journal of Behavioral Decision Making, 16*(5), 341-352.
Vercelli, A. (1999). The recent advances in decision theory under uncertainty *Uncertain Decisions* (pp. 237-260): Springer.
Wertheimer, M. (1938). *Gestalt theory*: Hayes Barton Press.