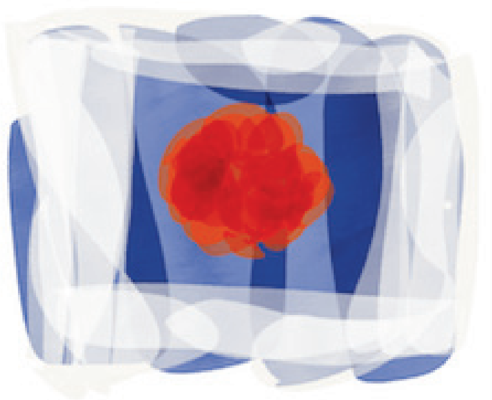
Deliverable D5.1

Initial Guidance Framework for Knowledge Quality Assessment in CoCliServ

Author(s) and affiliation(s)	Date	Version
Scott Bremer, UiB SVT Jeroen P. van der Sluijs, UiB SVT	15 Mar 2019	1.1



The CoCliServ project benefits from funding obtained through the ERA4CS Joint Call on Researching and Advancing Climate Services Development.

CoCliServ is funded by the following national funding agencies: Agence Nationale de la Recherche (**ANR**), France; Service public fédéral de programmation politique scientifique (**BELSPO**), Belgium; Deutsches Zentrum für Luft- und Raumfahrt EV (**DLR**), Germany; Nederlandse organisatie voor wetenschappelijk onderzoek (**NWO**), the Netherlands; Norges forskningsrad (**RCN**), Norway.

Table of contents

Executive summary/summary

In the practice of climate change adaptation, typically facts are uncertain, values in dispute, stakes high, and decisions urgent. This requires a so-called *post-normal science* approach. In a post-normal approach, the normal science task of fact-finding is still regarded as necessary, but no longer as fully feasible nor as sufficient to interface science and policy. It needs to be complemented with a task of exploring the relevance of deep uncertainty and ignorance that limit our ability to establish objective, reliable, and valid facts. To perform this task, Knowledge Quality Assessment (KQA) tools are central in post-normal science.

In exploring new modes of co-constructing climate services, CoCliServ explores and implements a novel tool for assessing knowledge quality, using a checklist-based approach. These checklists should be used to structure discussion and reflection on quality in a transdisciplinary co-construction collective, where each member of the community has a responsibility to contribute their own knowledge and to appraise the quality of the knowledge provided by others.

To this end, this deliverable presents two complimentary checklists for knowledge quality assessment of climate services. The first checklist focuses on an 'external' assessment in terms of a collaborative / joint assessment of climate services and knowledge by an actor group. The second checklist can serve as a self-reflexive and self-appraising 'internal' assessment. The checklists are for assessing knowledge quality relative to particular climate service projects, or instances when climate knowledge is used for responding to a specific problem or question or task. They assist in evaluating the relevance or fitness for purpose relative to that specific problem, question or task. The checklists are deliberation support tools. They are designed to support reflection and discussion about knowledge quality among a group of actors with an interest in a climate service project. The checklists are meant to trigger and structure a critical dialogue on knowledge quality of climate services within a co-construction collective.

Goal/Purpose of the document

This document aims to provide a guidance framework for dealing with uncertainty and quality in the production and use of climate knowledge for informing local climate adaptation. The goals of this guidance framework are to:

- Enable and promote conscious, context-aware, explicit, argued and well-documented options regarding the treatment of uncertainties and assumptions in the production and use of a shared knowledge base for place-based climate services;
- Provide guidance to climate scientists and local stakeholders and decision-makers on how to cope with inherently uncertain scientific information, to enable sensible responses to uncertain future scenarios and challenges in ways that enable and empower local communities to achieve their aspirations.

Relationship to the Description of Work (DOW)

This deliverable is part of work package 5 "Knowledge Quality Assessment" and results from Task 5.1: Operationalise the Guidance approach to knowledge quality assessment for the CoCliServ project.

1. Introduction to the guidance framework

Climate change adaptation under uncertainty exhibit a number of characteristics that make scientific advice too complex to tackle with normal scientific procedures. Problems with such characteristics require new ways of interfacing science and policy. Funtowicz & Ravetz (1993) have called this class of problems post-normal. The characteristics of post-normal problems are:

- Decisions will need to be made before conclusive scientific evidence is available;
- Potential impacts of 'wrong' decisions can be huge;
- Values are in dispute;

- The knowledge base is characterized by large (partly irreducible, largely unquantifiable) uncertainties, multi-causality, knowledge gaps, and imperfect understanding of the complex systems involved;
- While more research advances our knowledge, it does not always lead to less uncertainty because it tends to reveal unforeseen complexities;
- Scientific assessments are dominated by computer simulation models, scenarios, assumptions, extrapolations, most of which cannot be validated;
- Many (hidden) value loadings reside in problem frames, indicators chosen, and assumptions made.

Appendix A1.1 provides more background on the concept of post-normal science.

In a post-normal approach, the normal science task of fact-finding is still regarded as necessary, but no longer as fully feasible nor as sufficient to interface science and policy. It needs to be complemented with a task of exploring the relevance of deep uncertainty and ignorance that limit our ability to establish objective, reliable, and valid facts. To perform this task, Knowledge Quality Assessment (KQA) tools are central in post-normal science (Clark & Majone 1985; Funtowicz & Ravetz 1990; Walker et al 2003; Van der Sluijs et al 2005; Refsgaard et al 2006, 2007, Saltelli et al 2008; Van der Sluijs et al 2008, Kloprogge et al 2011; Maxim & Van der Sluijs 2014). KQA seeks to systematically reflect on the limits of knowledge in relation to its fitness for function. It comprises systematic analysis of, and critical reflection on uncertainty, assumptions and dissent in scientific assessments in its societal and institutional contexts.

Two particular strategies to deal with uncertainty dominate current practice of scientific fact-finding on complex risks: uncertainties are either downplayed to promote radical risk mitigation policies (enforced consensus/overselling certainty) or they are overemphasised to prevent government intervention in the economy. Both promote policy strategies that can result in extreme error-costs for society. A more sophisticated strategy to deal with uncertainty is urgently needed. Within CoCliServ such an approach is being pioneered.

In exploring new modes of co-constructing climate services, CoCliServ explores and implements criteria for assessing knowledge quality, using a checklist-based guidance approach inspired on the Guidance for Uncertainty Assessment and Communication that the work-package leader developed earlier for the Netherlands

Environmental Assessment Agency (Van der Sluijs et al, 2003, 2008; Petersen et al, 2011, 2013). These checklists should be used to structure discussion and reflection on quality in a transdisciplinary co-construction collective, where each member of the community has a responsibility to contribute their own knowledge and to appraise the quality of the knowledge provided by others.

Such processes of extended peer review (Funtowicz and Ravetz 1993) evolve through dialogue, with quality appraisal structured by formal criteria or tools agreed on within the community. These criteria of quality extend beyond narrow criteria of single scientific disciplines to encompass broader notions of what constitutes quality knowledge, including its fitness for *purpose*, the *people* producing it, the *process* used to produce it, and the final knowledge *product*. The KQA approach will also be used for critical self-reflection by the CoCliServ consortium about the activities in the case studies and work packages. This permits CoCliServ concepts and tools to be evaluated so that a statement can be made about their value in the case studies and their transferability to other settings. The guidance framework gives particular attention to uncertainties, their treatment and their shared understanding. The framework presented here enables and promotes systematic critical reflection on knowledge quality as a central activity in interfacing climate science and local governance through the co-creation of place-based climate services. The scientific background of the guidance framework is documented in the appendix 1.

1.1 A checklist approach for knowledge quality assessment of climate services

This deliverable puts forward two complimentary checklists for knowledge quality assessment of climate services. Three things are important to make clear up front. First, the checklists are for assessing knowledge *quality relative to particular climate service projects*, or instances when climate knowledge is used for responding to a discrete problem or question or task. They assist in evaluating the relevance or fitness for purpose relative to that specific problem, question or task. They are not suited to a general assessment of climate knowledge, at a national scale for instance, because knowledge quality here takes as its reference point the particular and contingent purpose or function for which climate knowledge is mobilised.

Second, these checklists are *deliberation support tools*. They are designed to support reflection and discussion about knowledge quality among a group of actors with an interest in a climate service project. In CoCliServ they need to be developed and applied collaboratively between the work package partners and case study stakeholders. They help ask, what constitutes high quality knowledge for this purpose, and for we who have an interest in it? And how do we, or could we better, fulfil these quality aspirations? They are NOT objective evaluative tools. Third, they can be used both *ex ante*, to guide climate service work from the beginning, or *ex post*, to look back over climate services produced and reflect on their strong and weak points.

In the following, we introduce the two checklists. The first checklist focuses on an 'external' assessment in terms of a collaborative / joint assessment of climate services and knowledge by an actor group. The second checklist can serve as a self-reflexive and self-appraising 'internal' assessment.

We recommend to spend some time in the beginning to clearly articulating the various *functions* that climate services play, seen from the various perspectives. Quality is in the first place fitness for function, so a shared understanding of the functions relative to which the fitness is assessed is a crucial step. Questions to ask the participants in a dialogue on knowledge quality include: What is the benefit for you using climate services? What function does it fulfil for your work? To assist in this step we drafted a list of candidate functions of climate services that might be of relevance (Table 1).

Table 1 Possible functions of knowledge in decision support

Enlightenment function

explore systems dynamics to gain understanding of complex interactions driven by scientific curiosity and desire to understand the mechanism that produce observed complex patterns of system behaviour

Integration function

to integrate knowledge from multiple disciplines

Pragmatic function

instrumental knowledge; dealing with if-then questions; indicating what the factual consequences are of a question;

Interpretive function

no longer driven by scientific curiosity but by the need to solve practical problems, to advise about how to manage the complex system at hand;

Forecasting function

to anticipate possible future system behaviour, stressors, and impacts

Legitimative function

to legitimate existing or new policies or to support a particular position in a policy debate

Depolitisation function

to use scientific knowledge or models as a device to create unchallengeable authority to settle societal conflict

The first checklist is a relatively blank framework, which is collaboratively filled out by actors interested in a climate service project. It is 'external' to all actors, because it can only be completed in cooperation with others, as a way of bridging knowledge quality expectations. It is inspired by work of Clark and Majone (1985), which is detailed further in the Appendix. In their seminal paper 'Critical Appraisal of Scientific Inquiries with Policy Implications', they noticed that what counts as high quality differs depending on one's perspective, depending on what they call their *critical role*. For instance, what constitutes a high quality climate service seen from the perspective of a scientist is different from a high quality climate service seen from the perspective of a policy maker. Therefore a first step in filling out this framework is to note the relevant *critical roles*; writing down the side those actors with an interest in a climate service project – each to a row. Ideally, all (categories of) actors will be present for filling out this framework, and can identify and represent themselves in completing the framework. For some categories of actors (e.g. animals, habitats, future generations), it may be necessary that some present actor represent them.

A second insight from Clark and Majone is that quality criteria differ according to the object of critical appraisal. They call this *critical modes*, which head up the columns, and are divided into: input (data; methods, people, competence, (im)matureness of field); process (good scientific practice, procedures for review, documenting etc.) and output (problem solved? hypothesis tested?). For the case of CoCliServ we add a forth critical mode/column: use, because our framework should not only address the step of the co-creation of climate services but should also include quality appraisal of their use in local climate change adaption. This creates a two by two matrix with critical roles heading the rows, and critical modes the columns (table 2).

In filling out the framework, actors discuss and register in the matrix cells their perspective on important quality criteria at each critical mode, or phase, of producing and using a climate service. That could be knowledge that attracts political support as an input, a process with regular face-to-face meetings, or an output that is web-based, for instance. In this way, quality criteria are contingent to the different actors' concrete expectations at each stage of the process. All are visible alongside each other in the matrix. Of course, as this is a deliberation support tool, it is not simply a task of collecting expectations in different cells. Actors present must justify quality criteria before they are recorded, and challenge others on their criteria. The completed matrix is a product of negotiation, not a collage.

Table 2: The deliberative knowledge quality assessment tool

Critical mode *Critical role*	input (and context)	process	output	use
Actor 1				
Actor 2				
Actor 3				
Actor 4				
....				

In an informal test of this framework during CoCliServ 2nd annual retreat in Dordrecht, 10-12 October 2018, we identified for instance as actors for the Dordrecht case: (climate) scientists, local government, water boards, Netherland's Delta program, Vogelnest ("Bird's nest", a social meeting place in the neighbourhood), province and the residents. Figure 1 shows a flipchart of that test-run.

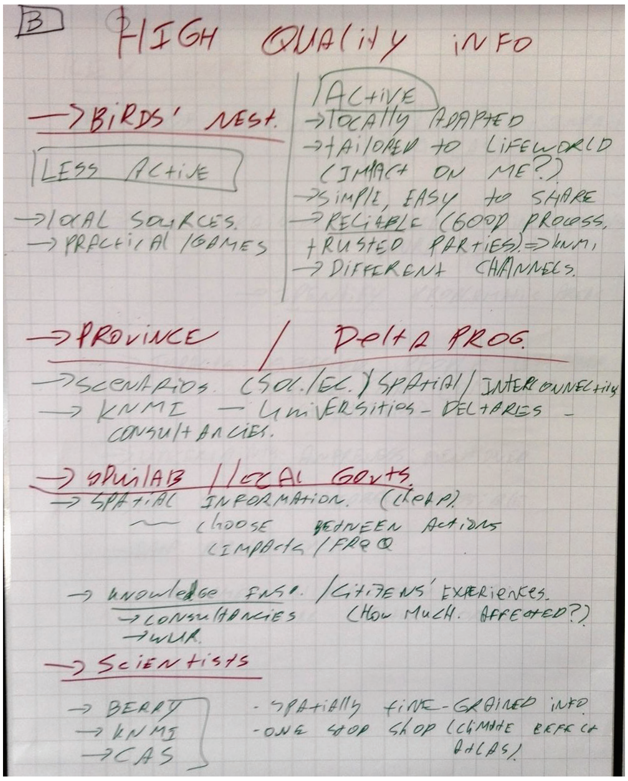


Figure 1: Informal brainstorm on what constitutes high quality information for the Dordrecht case study

The second checklist is complementary to the ‘external’ checklist. It is also reflexive and dialogic, but it can be used ‘internally’ by an actor group in self-appraising the quality of the knowledge they produce (even to the extent that an individual can have an internal dialogue). It substitutes the different actors in the matrix with a list of broad and relatively universal knowledge quality criteria assembled from a long and rich literature on the subject (see the Appendix). Table 3 lists these 10 principles, which each occupy a row on the matrix. An actor can then assess in what concrete way each of these principles can take form at each critical mode. For example, transparency might concretely mean having open meetings, which are filmed and put on a website. Actors can also select the criteria are most important from them or rank the list of criteria, or add criteria that are missing from their perspective (see also Meinke, 2017).

Of course, as it is, the generic principles render this matrix less contingent to each specific climate service project. But there are two ways to address this. First, the generic principles could be replaced with contingent criteria if this table for ‘internal’ reflection followed an exercise with the ‘external’ checklist. In this way, actors’ own concrete criteria, could be registered down the side of the matrix in place of the generic principles. Second, the internal and external checklists open up for an interesting comparison, of generic ideas of quality in the scholarship, compared to the ideas of quality in practice. So these two checklists have potential for being used separately, in series, or through comparison.

Table 3. Sort list of candidate quality criteria for climate services:

To what degree are the climate services:

1. Salient and fit for function
2. Uncertainty aware
3. Based on credible data
4. Inclusive and interactive
5. Legitimate and deliberative
6. Transparent and responsible
7. Intelligible and usable
8. Flexible and adaptable
9. Iterative and accounting for progressing insights
10. Encouraging of learning

Once the matrix is populated with criteria deemed relevant for the case at hand, it can serve as a checklist to trigger and structure a critical dialogue on knowledge quality of climate services within a co-construction collective.

Appendix 1: Scientific background of the guidance framework for knowledge quality assessment.

This appendix briefly presents the scientific background of the knowledge quality checklists presented in this deliverable. Section A.1.1 introduces the concepts of post-normal science, knowledge quality assessment and uncertainty. Section A.1.2. discusses the roles of uncertainty in the various phases of the knowledge generation cycle. Section A.1.3 reviews the various approaches to co-production of climate knowledge. Section A.1.4 provides an overview of state-of-the art approaches and guidelines to knowledge quality assessment and uncertainty appraisal.

A1.1 Post-normal science and Knowledge Quality Assessment

Climate change adaptation under uncertainty exhibit a number of characteristics that make scientific advice too complex to tackle with normal scientific procedures. Problems with such characteristics require new ways of interfacing science and policy (Funtowicz & Ravez 1990). Funtowicz & Ravetz (1993) have called this class of problems post-normal, where 'normal' refers to Kuhn's (1962) concept of normal science. Kuhn describes normal science both as "a strenuous and devoted attempt to force nature into the conceptual boxes supplied by professional education" (Kuhn 1962: 5) and as the practice of uncritical puzzle-solving within an unquestioned framework or 'paradigm'.

Funtowicz & Ravetz (1993) signalised that such a normal science approach runs into serious limitations when addressing societal issues (in that time, nuclear reactor safety) where scientific evidence is contested and plagued by uncertainties while decision stakes are high and values are in dispute. The available knowledge bases are typically characterised by imperfect understanding of the complex systems involved. Models, scenarios, and assumptions dominate assessment of these problems and many (hidden) value loadings reside in problem frames, indicators chosen, and assumptions made. Scientific assessments of complex risks are thus unavoidably based on a mixture of knowledge, assumptions, models, scenarios, extrapolations and known and unknown unknowns. Consequently, scientific assessments will unavoidably use expert judgements. It comprises bits and pieces of knowledge that differ in status, covering the entire spectrum from well-established knowledge to judgments, educated guesses, tentative assumptions and even crude speculations (Van der Sluijs et al 2005; 2008). Knowledge utilisation for risk governance requires a full and public awareness of the various sorts of uncertainty and underlying assumptions.

PNS provides a set of practical insights meant to assist scientists and the recipients of their research to work together fruitfully in situations defined by the so-called PNS mantra, where “the facts are uncertain, the values in dispute, the stakes high and the decisions urgent”. The theory of PNS is illustrated by its well-known ‘quadrant-rainbow’ with three areas (see Figure A1.1 below).

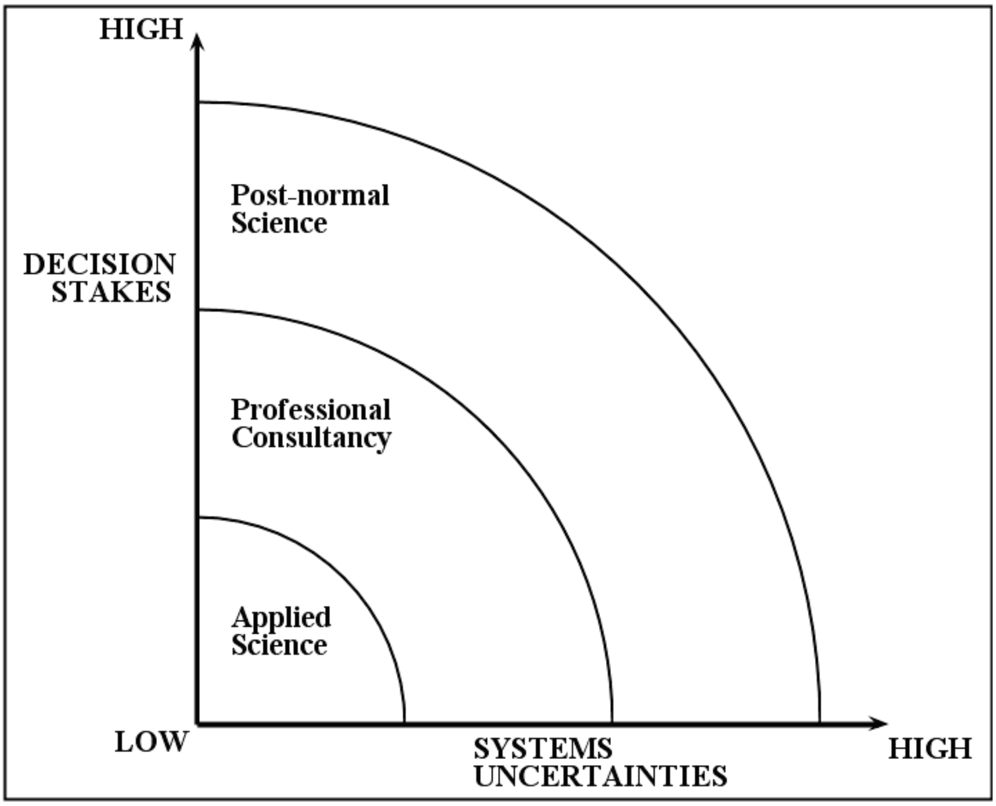


Figure A1.1 Post normal science diagram

The horizontal axis represents ‘Systems Uncertainties’ and the vertical one ‘Decision Stakes’. The three quadrants identify Applied Science, Professional Consultancy, and Post-Normal Science. Different standards of quality and styles of analysis are appropriate to different regions in the diagram, i.e. Post-normal science does not claim relevance and cogency on all of science's application but only on those defined by the PNS's mantra just described: ‘facts uncertain, values in dispute, stakes high and decisions urgent’. For applied research, science’s own peer quality control system will suffice (or so was assumed at the moment PNS was formulated in the early nineties), while professional consultancy was considered appropriate for these settings which cannot be ‘peer-reviewed’, and where the skills and the tacit knowledge of a practitioner are needed at the forefront, e.g. in a surgery room, or in a house on fire. Here a surgeon or a fireman takes a difficult technical decision based on her or his training and appreciation of the situation (the Greek concept of ‘metis’).

More often than not, the post-normal nature of a problem has implications for what ought not to be done, what pitfalls should be avoided, and what should make us suspicious in appraising scientific evidence. Thus, PNS embraces complexity and warns against the dangers of reductionism - the idea that every practical issue problem can be decomposed into a sum of simpler technical problems, or against arbitrary separations between facts and values, especially at the science-policy interface, or against science as a truth-machine.

PNS is foremost concerned about the quality of the scientific process, and addresses this process, seen as recursive (e.g., participatory and iterative) and reflexive (the analyst strives to see herself as part of the analysis). In the present situation of the ‘crisis of science’ (problems of reproducibility, retraction, fraud, and so on) and of expertise (questions about the legitimacy of experts’ inputs), and of run-away innovation (e.g., gene-drives, algorithms, and other cases), PNS can provide useful diagnoses and suggestions for action, if not always therapies.

Situating its interest at the science-policy interface, PNS focuses on the centrality of quality, avoiding the trap of useless controversies about truth, and stressing recognition of the plurality of publics that typify wicked problems – i.e. those problems where stakeholders disagree even on the definition of what the problem is. PNS encourages - against artificial separations – the integration of facts and values, and is thus in tension with visions of science as neutral, such as e.g. upheld in scientism.

PNS is both a critical concept and an inspiration for a new style of research practice. The dichotomous nature of PNS can be described as both descriptive (describing urgent decision problems – post-normal issues – characterized by incomplete, uncertain or contested knowledge and high decision stakes and how these characteristics change the relationship between science and governance) and normative (proposing a style of scientific inquiry and practice that is reflexive, inclusive and transparent in regards to scientific uncertainty and moving into a direction of democratisation of expertise) (Strand 2017). It is based on three defining features (Funtowicz & Ravetz 1993; Petersen et al 2011):

- The management of uncertainty. Post-normal science acknowledges that uncertainty is more than a number-range. Ambiguous knowledge assumptions and ignorance give rise to deep uncertainties;
- The acknowledgement of a plurality of legitimate perspectives - both cognitive and social. Complex problem solving requires interdisciplinary teamwork including expertise from outside science (NGOs, stakeholders, citizens). Scientists from different backgrounds often have irreconcilable and conflicting, yet tenable and legitimate scientific interpretations of the same body of evidence;
- The management of quality. An extended peer community includes representatives from social, political and economic domains who openly discuss on various dimensions of uncertainties, strengths, weaknesses and ambiguities in the available body of scientific evidence and its implications for all stakeholders with respect to the issue at hand.

In a post-normal approach, the normal science task of “getting the facts right” is still regarded as necessary, but no longer as fully feasible nor as sufficient to interface science and policy. It needs to be complemented with a task of exploring the relevance of deep uncertainty and ignorance that limit our ability to establish objective, reliable, and valid facts. To perform this task, **Knowledge Quality Assessment** (KQA) tools are central in post-normal science (Clark & Majone 1985; Funtowicz & Ravetz 1990; Walker et al 2003; Van der Sluijs et al 2005; Refsgaard et al 2006, 2007, Saltelli et al 2008; Van der Sluijs et al 2008, Kloprogge et al 2011; Maxim & Van der Sluijs 2014). KQA seeks to systematically reflect on the limits of knowledge in relation to its fitness for function. It comprises systematic analysis of, and critical reflection on uncertainty, assumptions and dissent in scientific assessments in its societal and institutional contexts.

Two particular strategies to deal with uncertainty dominate current practice of scientific fact-finding on complex risks: uncertainties are either downplayed to promote radical risk mitigation policies (enforced consensus/overselling certainty) or they are overemphasised to prevent government intervention in the economy. Both promote policy strategies that can result in extreme error-costs for society. A more sophisticated strategy to deal with uncertainty is urgently needed. Within CoCliServ such an approach is being pioneered.

A1.1.1 Conceptualisation of uncertainty

Van der Sluijs (2012) distinguishes three understandings of scientific uncertainty that different players in a scientific discourse over evidence may have: deficit view, evidence evaluation view, and complex systems view. Each way of seeing the phenomenon of scientific uncertainty leads to a different approach to uncertainties and each has its own drawbacks (Box A1.1).

Box A1.1 **Three understandings of uncertainty** (Van der Sluijs 2012)

Deficit view: Uncertainty as imperfections in the knowledge - 'truth with error bars'

The deficit view sees uncertainty as a temporary shortcoming in knowledge. It assumes that science will ultimately provide certainty. The approach is to reduce uncertainty, among other things, by creating increasingly complex models. As long as this is unsuccessful, the uncertainty is expressed numerically, for example, an error bar around a best guess.

The related science–policy interface model assumes a role of 'speaking truth to power' and remaining imperfections captured in a quantified error bar. This approach runs into the limitation that by far not all uncertainties can be expressed quantitatively in a reliable way. What's more, in practice, uncertainties do not become reduced with more research (e.g. Trenberth 2010): the problem appears to become ever more complex. The drawback of this approach is that there is a semblance of certainty because the magic numbers coming from the increasingly complex models suggest that there is more knowledge than is actually the case.

Evidence evaluation view: Uncertainty as lack of unequivocalness: 'consensus as proxy for truth'

The second view sees uncertainty as a problematic lack of unequivocalness. One scientist says this, the other says that. It is unclear who is right. It requests comparative evaluation of research results, focused on building scientific consensus in multi disciplinary expert panels (e.g. the IPCC). This approach is geared towards generating robust findings representing 'the best of our knowledge'. The corresponding science–policy interface model acknowledges that available knowledge is inconclusive and that the truth cannot yet be established - and it solves this limitation by assuming a role of 'speaking consensus to power', where consensus is a proxy for truth and is established through a negotiated (amongst a broad group of peers) widely shared interpretation of the yet inconclusive body of scientific evidence.

The drawback of this approach is that issues over which there is no consensus remain underexposed. It promotes anchoring towards previously established consensus positions, hiding diversity of perspectives and leads to under-appreciation of dissent (e.g., Sarewitz 2011). Often, it is precisely this dissent which tends to be extremely relevant to policymaking, for instance, dissent on how close we are to possible tipping points in the climate system that may produce catastrophic outcomes.

<u>*Complex systems view:*</u> *Uncertainty and dissent as facts of life: 'joint exploration of uncertainties and ignorance'*

The third view sees uncertainty as a mere fact of life, something which unavoidably plays a role in complex and politically sensitive topics. We accept the fact that uncertainty and dissent are not temporary but permanent, and recognise that not all uncertainties can be quantified. For example, models, scenarios, and extrapolations, all critically depend on the validity of the assumptions that unavoidably need to be made (Oreskes et al 1994; Pilkey & Pilkey 2007; NRC 2007). This post-normal view demands a culture that openly addresses uncertainty and that recognises that there are many things that science cannot yet provide conclusive answers for. It acknowledges plurality of scientific and societal perspectives (Burgess et al 2007; Stirling 2007). Ignorance and the influence of values are focused on here. It takes a reflective approach to uncertainty, using KQA techniques (Clark & Majone 1995; Van der Sluijs et al 2005; Van der Sluijs et al 2008; Kloprogge et al 2011) and deliberative risk governance (Pellizzoni 2001, Pereira & Funtowicz 2009). Knowledge production and use are seen as deliberative or participative social processes. Robustness is sought here primarily in policy strategy and not in the knowledge base: which policy is useful regardless of which of the diverging scientific interpretations of the knowledge is correct. The assumed science–policy interface model is one of ‘working deliberatively within imperfections’ where scientists, policymakers and other societal actors jointly explore the relevance of ignorance and uncertainties.

The drawback of this approach is that uncertainty and minority interpretations are so much in the spotlight that we forget how much we do know about these risks and which items actually enjoy broad consensus, which highlights the importance of developing and promoting in parallel more sophisticated non-paralyzing strategies to take uncertainty on board in decision making. Another drawback is that you need a lot of discussion between the parties involved, and each one has to showcase his thinking, which requires "negotiation in good faith" (Ravetz 2006).

The quantification (deficit view) and consensus (evidence evaluation) approaches are well developed and widely used. CoCliServ pioneers the third view as an alternative approach that is better fit to meet the challenges of developing high quality climate services to support local communities to meet local climate adaptation challenges. This view sees uncertainty and ignorance as intrinsic to complex systems and partly irreducible and seeks to transform the epistemic conceptualisation of uncertainty by integrating insights from amongst others STS and historic epistemology into a framework for knowledge quality assessment of climate services.

A1.2 Uncertainty and the knowledge generation cycle

According to Maxim and van der Sluijs (2011), uncertainty sources affecting knowledge production processes can be classified according to location, type as well as position within the knowledge generation cycle.

A1.2.1 Location of uncertainty

According to Maxim and van der Sluijs (2011) the existing ‘positivistic’ uncertainty analysis models such as those used in climate change modelling run the risk of lacking social relevance when they fail to adequately inform negotiations between stakeholders. From the perspective of the science-policy interface, uncertainty must also include quality criteria which are relevant to the political, social, and economic contexts. Regarding the location of uncertainty in a given body of knowledge the typology distinguishes three main categories:

1. Uncertainty related to the content of knowledge;
2. Uncertainty related to the process of knowledge production; and
3. Uncertainty related to the context of knowledge production.

'Content uncertainty' is related to data selection and curation, models' construction and quality assurance, and statistical procedures. It also includes conceptual uncertainty, understood as ignorance about qualitative relationships between phenomena.

'Process uncertainty' relates to the procedural quality of the process of knowledge construction. Under this domain falls considerations of completeness, credibility, transparency, saliency, credibility, legitimacy, and fairness.

'Context uncertainty' relates to the socio-economic and political factors influencing the knowledge production process. Context means identifying the boundaries of the real world to be modelled at the moment that the problem is framed.

Process and Context uncertainties have similarities with the categories of Assessment and Pedigree in NUSAP (Funtowicz and Ravetz, 1990).

A1.2.2 Key uncertainty types

While lack of knowledge and natural variability are the more widely treated types of uncertainty (often under the heading of epistemic and stochastic uncertainties) other uncertainty types need to be considered in science for policy. These include

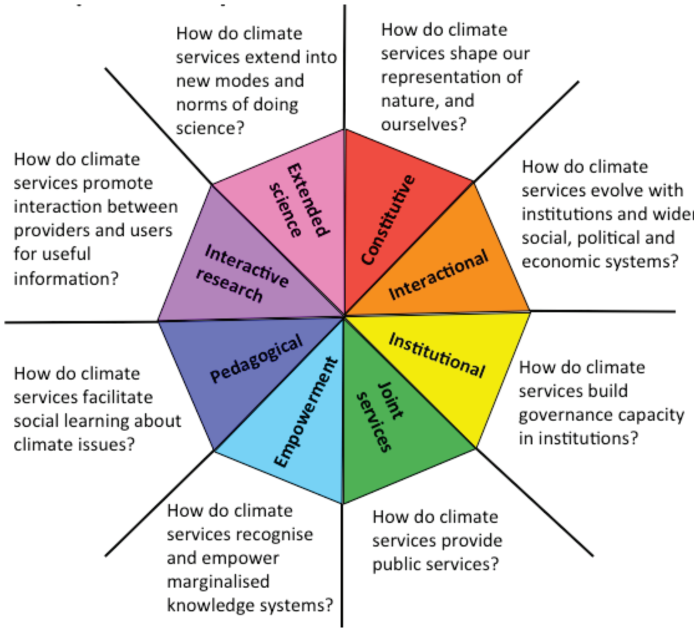
- Expert subjectivity, which may be due to philosophical, or professional orientation, and conflict of interest (or even fraud).
- Communication uncertainty, which may be associated with ambiguity (lack of clarity about the intended meaning of a word), context dependence (failure to specify the context), underspecificity (overly general statements), and vagueness.

A1.2.3 Position within the knowledge cycle

The relative impact of substantive, contextual and procedural uncertainty is highly dependent of where one is located in the context of the knowledge cycle: ‘problem framing’, ‘knowledge production’, and ‘knowledge communication and use’ all have their distinctive characteristics. The concept of quality as applied to scientific evidence is not the same as quality in the deployment of the evidence for policy – e.g. an excellent scientific quality may simple escape, or be ignored, or miscommunicated accidentally or instrumentally, with a resulting poor policy decision, while – at the opposite extreme – a modest scientific quality may end up being sufficient for the purpose of reaching a desirable policy compromise. The interplay between substantive, procedural, and contextual uncertainties may produce unforeseen effects. For example, in a regulatory context, regulators may imposed the use of risk assessment methods that are inadequate for the nature of the risk, or the experts involved may lack the relevant competence, or enough time to critically review the knowledge. In post normal science quality is often defined in relation to fitness for purpose. What should be avoided is the positivistic – reductionist approach of taking uncertainty as a mere attribute of the evidential basis. ‘Lack of knowledge’ is only a part of ‘lack of knowledge quality’, and not necessarily the most important part.

A1.3. Unpacking different approaches to climate knowledge co-production: Bremer & Meisch (2017)

Another key facet of knowledge quality relates to how it is co-created or ‘co-produced’ by different actors. As seen in many of the generic knowledge quality principles in Table 2, there is a long scholarship arguing that quality knowledge is rarely produced by one actor in isolation. Knowledge production is, intended or not, a social process, and there are normative reasons why we should deliberatively attend to the social facet of knowledge production. This introduces the work on co-production. But here again, there are numerous ways, based on numerous traditions and justifications, by which we can set out to produce knowledge with others. Bremer and Meisch (2017) distilled at least eight different traditions of climate knowledge co-production, which they arranged as a ‘prism’ of different lenses on co-production.



By employing a richer understanding of climate service co-production, there are also opportunities for building more comprehensive evaluation frameworks. Most current frameworks are structured around interactive research (Lemos & Morehouse 2005), which has usability as the central criterion of quality. This is a good starting point, but with a regard for the other perspectives on co-production comes a greater appreciation for the other disparate criteria that underpin the quality and value of climate services; where quality is socially construed and value socially constructed. Usefulness alone is arguably insufficient. This literature suggests that at the very least, issues like scientific credibility, legitimacy of co-production processes and ensuring sustained commitment of actors should be critically scrutinised (Armitage et al. 2011; Cash et al. 2003; Dessai & van der Sluijs 2011; Hegger et al. 2012). Conceptually, Bremer and Meisch (2017) also engage with how each different lens evaluates ‘good’ co-production, highlighting the need for more multi-faceted evaluation frameworks of climate services too (Table A.1).

Table A.1: Criteria for evaluating good or successful co-production of climate services (from Bremer & Meisch, 2017)

Co-production lens	*Evaluation criteria*
Constitutive lens	The diagnosis of climate services role in rebuilding representations of climate, and the social orders for living with this climate
Interactional lens	The exposure and critical challenge to dominant social forces steering climate services
Iterative interaction lens	The usability of climate information products in a decision-making context
Extended science lens	The social robustness, accountability, and legitimacy of climate information in the face of uncertainty
Public services lens	The efficient and effective provision of public services
Institutional lens	The building of adaptive capacity in institutions
Social learning lens	The creation of a setting for learning to learn
Empowerment lens	The empowerment of marginalised knowledge systems for governance

Methodologically, assembling multi-faceted evaluation frameworks is itself an exercise in co-production. We need to co-produce frameworks for evaluating co-production if we are going to be true and consistent with the concept. This implies going back to climate service communities and exploring with them how they appraise quality and value. For instance, if we opt for post-normal science, this means interrogating all stakeholders about the principles, processes, people, purposes and pedigrees that together determine quality and value in a context (Funtowicz & Ravetz 1993; Kloprogge & Van der Sluijs 2006). Related to quality, co-production concepts can render insights into how value is layered on climate services, and how we can evaluate value according to different criteria. Understanding direct and indirect co-production processes, and layers of value, is a critical step towards mainstreaming climate services and developing its markets.

A1.4. Overview of approaches and guidelines to knowledge quality assessment in the literature

Without pretending to be complete, we discuss four illustrative state-of-the-art approaches to knowledge quality assessment: the seminal 1985 quality appraisal framework by Clark and Majone; the post-normal science-inspired Guidance approach of the Netherlands Environmental Assessment Agency; the uncertainty guidance of the Intergovernmental Panel on Climate Change; and the Numeral Unit Spread Assessment Pedigree (NUSAP) approach to uncertainty assessment and communication.

A1.4.1 Clark and Majone's 1985 framework for Critical Appraisal of Scientific Inquiries with Policy Implications

Under the title "The Critical Appraisal of Scientific Inquiries with Policy Implications", Clark and Majone (1985) presented one of the first comprehensive frameworks for quality assessment in the science policy interface. The framework acknowledges that each actor that has a stake in quality control has a different role in the process of critical evaluation. For instance, scientists will emphasize other criteria in quality control than policy-makers. Further, Clark and Majone’s taxonomy distinguishes three general modes of critical appraisal: the input, the output and the process by which inquiry is conducted. Input refers to data; methods, people, competence, (im)matureness of field etc. Output relates to questions as whether the problem is solved and the hypothesis tested. Process concerns issues such as good scientific practice, procedures for review, documenting etc.

The resulting framework is displayed below.

Table 1. Critical criteria. **(Clark & Majone, 1985)**

Critical Role	Critical Mode Input	Output	Process
Scientist	Resource and time constraints; available theory; institutional support; assumptions; quality of available data; state of the art.	Validation; sensitivity analyses; technical sophistication; degree of acceptance of conclusions; impact on policy debate; imitation; professional recognition.	Choice of methodology (e.g., estimation procedures); communication; implementation; promotion; degree of formalization of analytic activities within the organization.
Peer Group	Quality of data; model and/or theory used; adequacy of tools; problem formulation. Input variables well chosen? Measure of success specified in advance?	Purpose of the study. Are conclusions supported by evidence? Does model offend common sense? Robustness of conclusions; adequate coverage of issues.	Standards of scientific and professional practice; documentation; review of validation techniques; style; interdisciplinarity.
Program Manager or Sponsor	Cost; institutional support within user organization; quality of analytic team; type of financing (e.g., grant vs. contract).	Rate of use; type of use (general education, program evaluation, decisionmaking, etc.); contribution to methodology and state of the art; prestige. Can results be generalized, applied elsewhere?	Dissemination; collaboration with users. Has study been reviewed?
Policymaker	Quality of analysts; cost of study; technical tools used (hardware and software). Does problem formulation make sense?	Is output familiar and intelligible? Did study generate new ideas? Are policy indications conclusive? Are they consonant with accepted ethical standards?	Ease of use; documentation. Are analysts helping with implementation? Did they interact with agency personnel? With interest groups?
Public Interest Groups	Competence and intellectual integrity of analysts. Are value systems compatible? Problem formulation acceptable? Normative implications of technical choices (e.g., choices of data).	Nature of conclusions; equity. Is analysis used as rationalization or to postpone decision? All viewpoints taken into consideration? Value issues.	Participation; communication of data and other information; adherence to strict rules of procedure.

In addition, Clark and Majone proposed four meta quality criteria: adequacy, value, effectiveness an legitimacy. Adequacy covers issues such as reliability, reproducibility, uncertainty analysis etc.

Value has three aspects: Internal: how well is the study carried out? and external: fitness for purpose or fitness for function. Personal value has to do with subjectivity, preferences, choices, assumptions and biases. Effectiveness is about the question whether it help to solve practical problems. Legitimacy has two aspects: numinous, which is about natural authority, independence, credibility and competence, and civil, which has to do with compliance with agreed procedures.

A1.4.2 The Guidance approach of the Netherlands Environmental Assessment Agency

The ‘Guidance for Uncertainty Assessment and Communication’ approach to knowledge quality assessment (the Guidance) (Janssen et al., 2005; Petersen et al.,

2013) is a comprehensive framework that covers both the substantial and the societal dimensions of quality. It is a proven tool that has previously been successfully applied in various contexts and several members of the CoCliServ team (Van der Sluijs, Wardekker, Bremer) have expertise and experience on its application and use. The Guidance was developed in 2002 through a partnership between the Netherlands Environmental Assessment Agency and Utrecht University, and has become widely used in that Agency. It has reportedly stimulated co-learning processes among scientific advisors and policy makers for a deeper understanding and awareness of uncertainty and its policy implications (van der Sluijs et al., 2008; Petersen et al, 2011).

The Guidance tool adopts a checklist approach, designed to transparently highlight and communicate uncertainties along a scientific assessment process as a way of structuring informed public and policy debate; be it for an Environmental Impact Assessment for a particular project, or a broader assessment of the body of knowledge used to inform a policy programme (see e.g. Janssen et al. (2005), van der Sluijs et al. (2008) and Petersen et al. (2013)). It does not limit its focus to formal, quantitative methods for sensitivity and uncertainty analysis, but extends its scope to the social context of knowledge production, including assumptions and value-loadings. In this way, it systematically guides scientists in an exploration of deeper uncertainties that reside, for instance, in problem framings, expert judgments, and assumed model structures. ‘‘It provides a heuristic that encourages self-evaluative systematic critical reflection in order to become aware of pitfalls in knowledge production and use. It also provides diagnostic help as to where uncertainty may occur and why.” (van der Sluijs et al., 2008). The Guidance focuses on six elements of knowledge production and use (Table A.2).

Table A.2: Criteria and key issues for knowledge quality in the Guidance (Petersen et al., 2013).

Phase in the assessment	**Key uncertainty & quality issues to critically reflect upon**
Problem framing	(i) Existing frames of the problem, other than that of end users; (ii) the interconnections with other problems; (iii) any other relevant aspects of the problem not addressed in the research questions; (iv) the role of the study in the policy process; and (v) the way in which the study connects to previous studies on the subject
Stakeholder involvement	(i) The relevant stakeholders; (ii) their views, roles, stakes and involvement with respect to the problem; and (iii) the aspects of the problem on which they disagree
Indicator/visualization selection	(i) Adequate backing for selection; (ii) alternative indicators; and (iii) support for selection in science, society, and politics
Appraisal of knowledge base	(i) The quality that is required; (ii) the current state of knowledge; and (iii) the gap between these two
Mapping and assessing relevant Uncertainties	(i) The relative importance of statistical uncertainty, scenario uncertainty and recognized ignorance with respect to the problem at hand; (ii) the uncertainty sources that are most relevant to the problem; and (iii) the consequences of these uncertainties for the conclusions of this study
Communication of uncertainty Information	(i) Context of reporting; (ii) robustness and clarity of main messages; (iii) policy implications of uncertainty; (iv) balanced and consistent representation in progressive disclosure of uncertainty information; and (v) traceability and adequate backing

For a recent application to knowledge for climate adaptation, see Haque et al, (2017) from which the above summary presentation of the guidance approach is taken.

A1.4.3 The IPCC guidance Note on Consistent Treatment of Uncertainties

The IPCC guidance note (IPCC, 2010) mainly gives advice on the use of standardized, calibrated language to express levels of precision and degrees of confidence. It is not clear in how expert judgements should be drafted, it leaves a lot of freedom to the author teams to take their own approach. As a result, the uncertainty terminology is standardized across the IPCC report but the underlying method that each author team has used to arrive at their judgements is not harmonized and not always transparent or reproducible.

It relies on two metrics for communicating the degree of certainty in key findings:

- Confidence in 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 the degree of agreement. Confidence is expressed qualitatively.
- Quantified measures of uncertainty in a finding expressed probabilistically (based on statistical analysis of observations or model results, or expert judgment).

For all key-findings reported, the validity of the finding should be expressed (summary terms: “limited,” “medium,” or “robust”), as well as the degree of agreement (summary terms: “low,” “medium,” or “high”). Further, a level of confidence needs to be expressed using five qualifiers: “very low,” “low,” “medium,” “high,” and “very high.” It synthesizes the author teams’ judgments about the validity of findings as determined through evaluation of evidence and agreement. Figure A.2 depicts summary statements for evidence and agreement and their relationship to confidence.

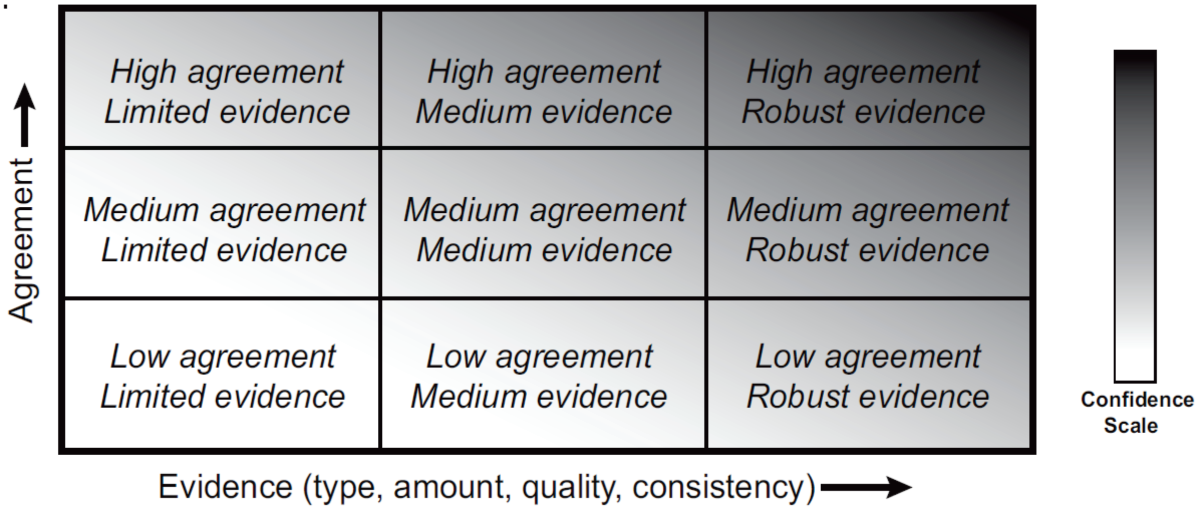


Figure A2: A depiction of evidence and agreement statements and their relationship to confidence. Confidence increases towards the top-right corner as suggested by the increasing strength of shading.

The IPCC guidance further requires authors to consistently use a likelihood terminology, as defined in Table A.3, which provides calibrated language for describing quantified uncertainty. It can be used to express a probabilistic estimate of the occurrence of a single event or of an outcome (e.g., a climate parameter, observed trend, or projected change lying in a given range).

Table 1. Likelihood Scale	
Term*	**Likelihood of the 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

Table A.3: IPCC likelihood scale

Finally, the IPCC guidance presents Risbey & Kandlikar (2007)'s ladder of quantification, but without any justification IPCC has left out of the ladder the most relevant category, namely effective ignorance. Risbey & Kandlikar (2007) addressed the epistemic question: what format is in accordance with the level of knowledge on the quantity? Their proposed ladder includes following categories:

- Full probability density function
  - Robust, well defended distribution
- Bounds
  - Well defended percentile bounds
- First order estimates
  - Order of magnitude assessment
- Expected sign or trend
  - Well defended trend expectation
- Ambiguous sign or trend
  - Equally plausible contrary trend expectations
- Effective ignorance [MISSING in the IPCC Guidance note!]
  - Lacking or weakly plausible expectations

From an epistemological point of view we strongly advise that everyone who uses the IPCC guidance notes and modifies it by re-including the category "effective ignorance" which should be added under their heading 11 before their category A in their guidance note and should read:

"Effective ignorance (Lacking or weakly plausible expectations): In most cases we know quite a bit about the outcome variable. Yet despite this, we may not know much about the factors that would govern a change in the variable of the type under consideration. As such, it may be difficult to outline plausible arguments for how the variable would respond. If the arguments used to support the change in the variable are so weak as to stretch plausibility, then this category is appropriate. Selecting this category does not mean that we know nothing about the variable. Rather, it means that our knowledge of the factors governing changes in the variable in the context of interest is so weak that we are effectively ignorant in this particular regard. If this category is selected, describe any expectations, such as they are, and note problems with them." (Risbey & Kandlikar 2007)

A1.4.4 The NUSAP approach: Numeral Unit Spread Assessment Pedigree

NUSAP is a notational system proposed by Funtowicz and Ravetz (1990), which aims to provide an analysis and diagnosis of uncertainty in science for policy (see Van der Sluijs 2017 for a recent overview). It captures both quantitative and qualitative dimensions of uncertainty and enables one to display these in a standardized and self-explanatory way. It promotes criticism by clients and users of all sorts, expert and lay and will thereby support extended peer review processes.

The basic idea is to qualify quantities using the five qualifiers of the NUSAP acronym: Numeral, Unit, Spread, Assessment, and Pedigree. By adding expert judgment of reliability (Assessment) and systematic multi-criteria evaluation of the production process of numbers (Pedigree), NUSAP has extended the statistical approach to uncertainty (inexactness) with the methodological (unreliability) and epistemological (ignorance) dimensions. By providing a separate qualification for each dimension of uncertainty, it enables flexibility in their expression. By means of NUSAP, nuances of meaning about quantities can be conveyed concisely and clearly, to a degree that is quite impossible with statistical methods only.

We will discuss the five qualifiers. The first is **Numeral**; this will usually be an ordinary number; but when appropriate it can be a more general quantity, such as the expression "a million" (which is not the same as the number lying between 999,999 and 1,000,001). Second comes Unit, which may be of the conventional sort, but which may also contain extra information, as the date at which the unit is evaluated (most commonly with money).

The middle category is **Spread**, which generalizes from the "random error" of experiments or the "variance" of statistics. Although Spread is usually conveyed by a number (either ±, % or "factor of") it is not an ordinary quantity, for its own inexactness is not of the same sort as that of measurements. Methods to address Spread can be statistical data analysis, sensitivity analysis or Monte Carlo analysis possibly in combination with expert elicitation.

The remaining two qualifiers constitute the more qualitative side of the NUSAP expression. **Assessment** expresses qualitative judgments about the information. In the case of statistical tests, this might be the significance level; in the case of numerical estimates for policy purposes, it might be the qualifier "optimistic" or "pessimistic". In some experimental fields, information is given with two ± terms, of which the first is the spread, or random error, and the second is the "systematic error" which must estimated on the basis of the history of the measurement, and which corresponds to our Assessment. It might be thought that the "systematic error" must always be less than the "experimental error", or else the stated "error bar" would be meaningless or misleading. But the "systematic error" can be well estimated only in retrospect, and then it can give surprises.

Finally there is P for **Pedigree**, which conveys an evaluative account of the production process of information, and indicates different aspects of the underpinning of the numbers and scientific status of the knowledge used. Pedigree is expressed by means of a set of pedigree criteria to assess these different aspects. Assessment of pedigree involves qualitative expert judgment. To minimize arbitrariness and subjectivity in measuring strength, a pedigree matrix is used to code qualitative expert judgments for each criterion into a discrete numeral scale from 0 (weak) to 4 (strong) with linguistic descriptions (modes) of each level on the scale. Each special sort of information has its own aspects that are key to its pedigree, so different pedigree matrices using different pedigree criteria can be used to qualify different sorts of information. Table A.4 gives an example of a pedigree matrix for emission monitoring data. An overview of pedigree matrices found in the literature is given in the pedigree matrices section of http://www.nusap.net. Examples of questionnaires used for eliciting pedigree scores can be found at http://www.nusap.net.

Table A.4 Pedigree matrix for emission monitoring data – Risbey et al, 2001.

Score	Proxy representation	Empirical basis	Methodological rigour	Validation
4	An exact measure of the desired quantity	Controlled experiments and large sample direct measurements	Best available practice in well established discipline	Compared with independent measurements of the same variable over long domain
3	Good fit or measure	Historical/field data uncontrolled experiments small sample direct measurements	Reliable method common within est. discipline Best available practice in immature discipline	Compared with independent measurements of closely related variable over shorter period
2	Well correlated but not measuring the same thing	Modelled/derived data Indirect measurements	Acceptable method but limited consensus on reliability	Measurements not independent proxy variable limited domain
1	Weak correlation but commonalities in measure	Educated guesses indirect approx. rule of thumb est.	Preliminary methods unknown reliability	Weak and very indirect validation
0	Not correlated and not clearly related	Crude speculation	No discernible rigour	No validation performed

We will briefly elaborate the four criteria in this example pedigree matrix.

Proxy representation. Sometimes it is not possible to measure directly the thing we are interested in or to represent it by a parameter, so some form of proxy measure is used. Proxy refers to how good or close a measure of the quantity that we measure or model is to the actual quantity we seek or represent. Think of first order approximations, over simplifications, idealizations, gaps in aggregation levels, differences in definitions, non-representativeness, and incompleteness issues.

Empirical basis. Empirical basis typically refers to the degree to which direct observations, measurements and statistics are used to estimate the parameter. Sometimes directly observed data are not available and the parameter or variable is estimated based on partial measurements or calculated from other quantities.

Parameters or variables determined by such indirect methods have a weaker empirical basis and will generally score lower than those based on direct observations.

Methodological rigour. Some method will be used to collect, check, and revise the data used for making parameter or variable estimates. Methodological quality refers to the norms for methodological rigour in this process applied by peers in the relevant disciplines. Well-established and respected methods for measuring and processing the data would score high on this metric, while untested or unreliable methods would tend to score lower.

Validation. This metric refers to the degree to which one has been able to crosscheck the data and assumptions used to produce the numeral of the parameter against independent sources. In many cases, independent data for the same parameter over the same time period are not available and other data sets must be used for validation. This may require a compromise in the length or overlap of the data sets, or may require use of a related, but different, proxy variable for indirect validation, or perhaps use of data that has been aggregated on different scales. The more indirect or incomplete the validation, the lower it will score on this metric.

Appendix 2: Glossary

Deep uncertainty: stems from the notion that uncertainty is more than a number-range. Ambiguous knowledge assumptions and ignorance give rise to deep uncertainties. Deep uncertainty refers to all uncertainty that cannot be quantified in a reliable way, so that its treatment in scientific assessments heavily relies on subjective expert judgements, which makes it prone to co-shaping. For instance, models, scenarios, and extrapolations used in risk assessment all critically depend on the validity of the assumptions that unavoidably need to be made, while most of them can intrinsically not be validated (Oreskes et al 1994; Pilkey & Pilkey 2007; NRC 2007). An essential element in understanding controversy is that uncertainty can also be artificially manufactured as a deceitful tactic through “merchants of doubt strategies” (Michaels 2005; Oreskes & Conway 2010).

Context validation: Context validity refers to the probability that an estimate has approximated the true but unknown range of (causally) relevant aspects and rival hypotheses present in a particular policy context. Context validation thus is minimizing the probability that one overlooks something of relevance. It can be performed by a participatory bottom-up process eliciting from stakeholders those aspects considered relevant as well as rival hypotheses on underlying causal relations, and rival problem definitions and problem framings. See Dunn, 1998, 2000.

Extended facts: Knowledge from other sources than science, including local knowledge, citizens’ surveys, anecdotal information, and the results of investigative journalism. Inclusion of extended facts in environmental assessment is one of the key principles of Post-Normal Science. (Funtowicz and Ravetz, 1993)

Extended peer community: a community involving all those with stakes or interest in an issue being debated/decided at the science policy interface. The community includes experts from various disciplines as well as lay experts.

GIGO: Literally, Garbage In, Garbage Out, refers to models where the uncertainties of the inputs must be suppressed lest the outputs become completely indeterminate (Futowicz and Ravetz, 1990). A variant formulation is ‘Garbage In, Gospel Out’ referring to a tendency to put faith in glossy outcomes of computer simulations, regardless of the limited quality of the inputs and model assumptions.

NUSAP: Acronym for Numeral Unit Spread Assessment Pedigree Notational system developed by Silvio Funtowicz and Jerry Ravetz to better manage and communicate uncertainty in science for policy. In NUSAP, the increasing severity of uncertainty is marked by the three categories of uncertainty, Spread for technical uncertainty (or error-bar), Assessment for methodological (or unreliability) and Pedigree for border with ignorance (or the essential limitations of a particular sort of scientific practice). (Funtowicz and Ravetz, 1990; Van der Sluijs, 2017)

Pitfall: A pitfall is a characteristic error that commonly occurs in assessing a problem. Such errors are typically associated with a lack of knowledge or experience, and thus may be reduced by experience, by consultation of others, or by following procedures designed to highlight and avoid pitfalls. In complex problems we sometimes say that pitfalls are 'dense', meaning that there is an unusual variety and number of pitfalls.

Post-normal science (PNS): is both a critical concept and an inspiration for a new style of research practice. On the one hand PNS refers to a class of problems − post-normal issues − characterized by: decisions urgent, knowledge incomplete, uncertain or contested, values in dispute and decision stakes high. PNS theory highlights how these characteristics change the relationship between science and governance. At the same time PNS inspires a new style of scientific inquiry and practice that is reflexive, inclusive and transparent in regards to scientific uncertainty and moves into a direction of democratisation of expertise (Strand, 2017).

Sensitivity analysis: The study of the relative importance of different input factors on the model output. A global sensitivity analysis: studies how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input.

Sensitivity auditing: An extension of sensitivity analysis where the model is no longer 'given' but is instead discussed in relation to the motivation of its developers, its framing questioned and possible interest and power relationships investigated. This extension becomes useful when a model is used for policy evaluation or deliberation. (Saltelli et al., 2013)

Uncertainty analysis: Focuses on quantifying the uncertainty in model output, often achieved via an uncertainty propagation analysis using Monte Carlo methods.

Value-ladenness: Value-ladenness refers to the notion that value orientations and biases of an analyst, an institute, a discipline or a culture can co-shape the way scientific questions are framed, data are selected, interpreted, and rejected, methodologies are devised, explanations are formulated and conclusions are formulated. Since theories are always underdetermined by observation, the analysts' biases will fill the epistemic gap which makes any assessment to a certain degree value-laden.

Acknowledgements

The authors thank Birgit Gerkensmeier and Florentin Breton for valuable comments form a WP3 perspective on the 15 June 2018 version of this deliverable. We tank all participants of CoCliServ's second annual retreat in Dordrecht 10-12 October 2018 for participating in an informal test of the KQA checklists presented in this deliverable.

References

Armitage D, Berkes F, Dale A, Kocho-Schellenberg E, Patton E. (2011) Co-management and the co-production of knowledge: learning to adapt in Canada's Arctic. *Glob Environ Change* 2011, 21:995–1004.

Bremer, S., & Meisch, S. (2017). Co-production in climate change research: reviewing different perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 8(6), e482.

Burgess J, Stirling A, Clark J, et al. (2007) Deliberative mapping: A novel analytic-deliberative methodology to support contested science-policy decisions, *Public Understanding of Science* 16(3): 299-322.

Cash, D. W., Clark, W. C., Alcock, F., Dickson, N. M., Eckley, N., Guston, D. H., ... & Mitchell, R. B. (2003). Knowledge systems for sustainable development. *Proceedings of the national academy of sciences*, 100(14), 8086-8091.

Clark WC and Majone G (1985) The Critical Appraisal of Scientific Inquiries with Policy Implications. *Science, Technology & Human Values* 10 (3):6-19.

Dessai S. and van der Sluijs J.P. (2011) Modelling climate change impacts for adaptation assessments, in Stephen Senn, Philip Dawid, Mike Christie and Kenneth Andrew Cliffen (2011). *Simplicity, complexity and modelling*, John Wiley & Sons, Chichester, 83-102. ISBN: 978-0-470-74002-6, Funtowicz SO and Ravetz JR (1990) Uncertainty and Quality in Science for Policy, Kluwer, Dordrecht.

Funtowicz SO and Ravetz JR (1993). Science for the post-normal age. *Futures* 25(7): 739-755.

Haque, M.M., Bremer, S., Aziz, S.B. and van der Sluijs J.P. (2017) A critical assessment of knowledge quality for climate adaptation in Sylhet Division, Bangladesh. *Climate Risk Management* 16:43-58.

Hegger D, Lamers M, Van Zeijl-Rozema A, Dieperink C. 2011. Conceptualising joint knowledge production in regional climate change adaptation projects: success conditions and levers for action. *Environmental Science & Policy* 18:52–65.

Janssen, P.H.M., Petersen, A.C., Van der Sluijs, J.P., Risbey, J., Ravetz, J.R. (2005), A guidance for assessing and communicating uncertainties. *Water science and technology*, 52 (6) 125–131

Kloprogge P. & Van der Sluijs J.P. (2006), The inclusion of stakeholder knowledge and perspectives in integrated assessment of climate change. *Climatic Change* 75 (3):359-389.

Kloprogge P., Van der Sluijs J.P. and Petersen A.C. (2011) A method for the analysis of assumptions in model-based environmental assessments, *Environmental Modelling & Software* 26 (3): 289–301.

Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago, 172 pp.

Lemos, M.C., Morehouse, B.J., 2005. The co-production of science and policy in integrated climate assessments. *Global Environmental Change* 15:57–68.

Maxim L. and van der Sluijs J.P. (2011) Quality in environmental science for policy: assessing uncertainty as a component of policy analysis. *Environmental Science & Policy*, 14, (4) 482-492.

Maxim L and Van der Sluijs JP (2014) Qualichem in vivo: A tool for assessing the quality of in vivo studies and its application for Bisphenol A, *PLOS One* 9 (1) e87738. doi:10.1371/journal.pone.0087738

Meinke, I. (2017) 'Stakeholder-based evaluation categories for regional climate services – a case study at the German Baltic Sea coast', *Adv. Sci. Res.*, 14, 279-291.

Michaels D. (2005) Doubt is their product. Industry groups are fighting government regulation by fomenting scientific uncertainty. *Scientific American* 292:96-101.

NRC (2007) *Models in Environmental Regulatory Decision Making*, National Research Council, Washington DC: National Academies Press.

Oreskes N., Shrader-Frechette K. and Belitz K. (1994) Verification, validation, and conformation of numerical models in the Earth sciences. *Science* 263:641–646.

Oreskes N., Conway E. (2010) *Merchants of Doubt, How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. New York: Bloomsbury Press.

Pellizzoni L. (2001) Democracy and the governance of uncertainty: The case of agricultural gene technologies, *Journal of Hazardous Materials* 86, 205–222

Pereira Â.G., Funtowicz S. (2009) *Science for Policy*. Oxford: Oxford UniversityPress.

Petersen A.C., A. Cath, M. Hage, E. Kunseler, and J.P. van der Sluijs (2011). Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency. *Science Technology & Human Values* 36 (3), 362-388.s.

Petersen A.C., Janssen P.H.M., van der Sluijs J.P., Risbey J.S., Ravetz J.R., Wardekker J.A., Martinson Hughes H. (2013) *Guidance for Uncertainty Assessment and Communication*, second edition. Netherlands Environmental Assessment Agency, Bilthoven.

Pilkey O.H. and Pilkey-Jarvis L. (2007) *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*, New York: Columbia University Press.

Ravetz, J. R. (2006). Post-normal science and the complexity of transitions towards sustainability. *Ecological complexity* 3(4), 275-284.

Refsgaard J-C, Van der Sluijs J.P., Brown J., and van der Keur P. (2006), A Framework For Dealing With Uncertainty Due To Model Structure Error. *Advances in Water Resources* 29 (11):1586–1597.

Refsgaard J-C., Van der Sluijs J.P., Højberg A.L., Vanrolleghem P.A. (2007), Uncertainty in the environmental modelling process: A framework and guidance, *Environmental Modelling & Software* 22 (11):1543-1556.

Risbey, J. S., & Kandlikar, M. (2007). Expressions of likelihood and confidence in the IPCC uncertainty assessment process. *Climatic change* 85(1-2), 19-31.

Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M. and Tarantola S. (2008) *Global Sensitivity Analysis: The Primer* (Chichester: Wiley)

Saltelli, A., Guimarães Pereira, Â., van der Sluijs, J.P. and Funtowicz S. (2013) What do I make of your Latinorum? Sensitivity auditing of mathematical modelling, *International Journal of Foresight and Innovation Policy* 9(2/3/4): 213-234.

Sarewitz D. (2011) The voice of science: let's agree to disagree. *Nature* 478, 7, doi:10.1038/478007a.

Stirling A. (2007) A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface* 4(15): 707-719.

Strand R. (2017) Post Normal Science. Chapter 28 in: Spash CL (Ed.) *Routledge Handbook of Ecological Economics*: *Nature and Society*. Routlegde: London. ISBN-13: 978-1138931510.

Trenberth K. (2010) More knowledge, less certainty. *Nature Reports Climate Change* 4:20-21.

Van der Sluijs J.P. (2012) Uncertainty and dissent in climate risk assessment, a post-normal perspective. *Nature and Culture* 7 (2) 174-195.

Van der Sluijs, J.P. (2017) The NUSAP approach to uncertainty appraisal and communication. Chapter 29 p.301-310 in C.L. Spash (Ed.) *Routledge Handbook of Ecological Economics: Nature and Society*. Routlegde: London. ISBN-13: 978-1138931510.

Van der Sluijs, J.P, Risbey, J.S., Kloprogge, P., Ravetz, J.R., Funtowicz, S.O., Corral Quintana, S., Guimarães Pereira, Â, De Marchi, B., Petersen, A.C., Janssen, P.H.M. Hoppe, R., and Huijs S.W.F. (2003). *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Detailed Guidance*. Utrecht University & RIVM, Bilthoven.

Van der Sluijs J.P., M. Craye, S. Funtowicz, P. Kloprogge, J. Ravetz, and J. Risbey (2005), Combining Quantitative and Qualitative Measures of Uncertainty in Model based Environmental Assessment: the NUSAP System. *Risk Analysis* 25 (2):481-492

Van der Sluijs J.P., A.C. Petersen, P.H.M. Janssen, J.S. Risbey and J.R. Ravetz (2008) Exploring the quality of evidence for complex and contested policy decisions, *Environmental Research Letters* 3:024008 (9pp)

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