



U.S. DEPARTMENT OF  
**ENERGY**

PNNL-20788

Prepared for the U.S. Department of Energy  
under Contract DE-AC05-76RL01830

# Characterizing Uncertainty for Regional Climate Change Mitigation and Adaptation Decisions

S Unwin  
J Rice

R Moss  
M Scott

September 2011



**Pacific Northwest**  
NATIONAL LABORATORY

*Proudly Operated by **Battelle** Since 1965*

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY

*operated by*

BATTELLE

*for the*

UNITED STATES DEPARTMENT OF ENERGY

*under Contract DE-AC05-76RL01830*

**Printed in the United States of America**

**Available to DOE and DOE contractors from the  
Office of Scientific and Technical Information,  
P.O. Box 62, Oak Ridge, TN 37831-0062;  
ph: (865) 576-8401  
fax: (865) 576-5728  
email: [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)**

**Available to the public from the National Technical Information Service,  
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161  
ph: (800) 553-6847  
fax: (703) 605-6900  
email: [orders@ntis.fedworld.gov](mailto:orders@ntis.fedworld.gov)  
online ordering: <http://www.ntis.gov/ordering.htm>**



This document was printed on recycled paper.

(9/2003)

# **Characterizing Uncertainty for Regional Climate Change Mitigation and Adaptation Decisions**

S Unwin  
J Rice

R Moss  
M Scott

September 2011

Prepared for  
the U.S. Department of Energy  
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory  
Richland, Washington 99352

# Summary

This white paper describes the results of new research to develop an uncertainty characterization (UC) process to help address the challenges of regional climate change mitigation and adaptation decisions. This research is being carried out as part of the integrated Regional Earth System Model (iRESM) initiative, a new scientific framework developed at Pacific Northwest National Laboratory to evaluate the interactions between human and environmental systems and mitigation and adaptation decisions at regional scales. The framework integrates a regional climate model; a regional energy-economy model; and highly spatially-resolved models of crop productivity, building energy demands, electricity infrastructure operation and expansion, and water supply and management. The iRESM framework is intended to help regional stakeholders (scientists as well as decision makers) understand the consequences of climate change as well as the consequences of policies to mitigate or adapt to such change within regions.

The initiative has developed the following four science questions to guide its research:

- How do intrinsic regional characteristics shape, enhance, or constrain regional mitigation and adaptation opportunities?
- How do projected changes in mean climate versus climate extremes affect the development of adaptation and mitigation strategies?
- How might interactions between management decisions and natural processes contribute to rapid or nonlinear changes, and do they contribute to climate feedbacks?
- How will adaptation and mitigation strategies interact in the next few decades in terms of achieving their respective goals?

An important consideration for the iRESM initiative is that Earth system mechanisms and future changes are imperfectly understood and in some cases deeply uncertain—especially at the level of resolution required for regional analyses and decision making. The UC process developed for the initiative addresses uncertainty by first identifying through sensitivity analysis the key uncertainties in data inputs, individual model structures, and coupled models that are important for particular stakeholder questions and evaluation criteria. These key uncertainties are then characterized and propagated to determine the robustness of the framework’s results for the particular questions, thereby providing insights for researchers and decision makers alike. The process differs from many traditional applications of uncertainty quantification (UQ) because of its focus on stakeholder needs and its allowance for qualitative and semi-quantitative methods for describing uncertainty. The process not only permits the dimensionality of the UQ problem to be reduced, it also allows research efforts to be targeted at the uncertainties that really matter for the question in hand.

This decision-specific orientation for UC has multiple implications for the iRESM initiative that will continue to be explored, including: the importance of stakeholder interactions and the development of methods for communicating results; the development of a flexible model architecture that will facilitate the application of the iRESM model components relevant to particular applications; and the identification of approaches for reducing model run times to facilitate the UC process, including the development of surrogate models.

## Acronyms and Abbreviations

AOGCM	atmosphere-ocean general circulation model
BEND	Building ENergy Demand
DCLM	Distributed Community Land Model
DCLM-WM	Distributed Community Land Model-Water Management
D-S	Dempster-Shafer
EMF	Energy Modeling Forum
IPCC	Intergovernmental Panel on Climate Change
IRESM	integrated Regional Earth System Model
LHS	Latin Hypercube Sampling
MELD	electricity demand
MIT	Massachusetts Institute of Technology
OEM	electricity operations
RB	Robust Bayes
RCP	representative concentration pathways
REIF	Regional Energy Infrastructure Framework
RESM	Regional Earth System Model
R-GCAM	Regional-Global Change Assessment Model
RPS	renewable portfolio standards
UC	uncertainty characterization
UP	uncertainty propagation
UQ	uncertainty qualification

# Contents

Summary .....	ii
Acronyms and Abbreviations .....	iii
1.0 Introduction .....	1.1
2.0 UC: Taxonomies and Methods .....	2.1
2.1 Taxonomies of Uncertainty .....	2.1
2.2 Methods for UC.....	2.6
2.2.1 Non-Quantitative Methods for UC.....	2.7
2.2.2 UQ Methods .....	2.9
2.3 UP Methods.....	2.23
2.3.1 Quantitative UP.....	2.23
2.3.2 Qualitative Methods .....	2.26
3.0 Decision Making Under Uncertainty.....	3.1
3.1 Methods.....	3.1
3.1.1 Classical Decision Methods .....	3.1
3.1.2 Robust Strategies.....	3.2
3.1.3 Adaptive Strategies .....	3.2
3.1.4 Non-Algorithmic Decisions .....	3.3
3.1.5 Conclusions on Decision Methods .....	3.3
3.2 Interpretation of Results/Visualization.....	3.3
3.2.1 Sensitivity: Radar Graphs and Tornado Diagrams.....	3.3
3.2.2 Uncertainty Box and Whisker Plots, Scatter Plots .....	3.5
3.2.3 Spatial Variation.....	3.7
4.0 Recommended UC Process for iRESM.....	4.1
4.1 Identify Stakeholder Decision Support Needs .....	4.2
4.2 Establish Research Metrics/Decision Criteria.....	4.3
4.3 Identify Necessary Model Components of iRESM Framework.....	4.4
4.4 Identify Sources of Uncertainty in Models and Model Couplings.....	4.6
4.5 Select UC Methods.....	4.7
4.6 Perform Screening/Sensitivity Analysis.....	4.7
4.7 Characterize/Quantify Uncertainty in Sensitivity Drivers .....	4.8
4.8 Design UP Approach.....	4.8
4.9 Perform UP.....	4.9
4.10 Evaluate and Interpret Results.....	4.9
5.0 Conclusion.....	5.1
6.0 References .....	6.1

## Figures

2.1 The Relationships Between UC, UQ, and UP .....	2.1
2.2 Taxonomy for Uncertainty Source Identification in iRESM .....	2.6
2.3 Hypothetical D-S and Bayesian Representations of Uncertainty: 2095 CO <sub>2</sub> Emissions Level ...	2.13
2.4 Hypothetical Possibilistic, D-S, and Bayesian Representations of Uncertainty .....	2.15
2.5 Hypothetical p-Box Representations of Uncertainty: 2095 CO <sub>2</sub> Emissions Level .....	2.17
3.1 Examples of Radar Graphs that Depict Many Variables' Contribution to a Single Outcome and Alternative Strategies to Fulfilling Competing Objectives .....	3.4
3.2 Percentage of Variance in Carbon Prices in (a) 2020, (b) 2060 Explained by Uncertain Parameter Under a 550 ppm Stabilization Case.....	3.5
3.3 Comparison of Flat Carbon Tax, Hotelling Carbon Tax, and Contingent Carbon Tax in Controlling the CO <sub>2</sub> Concentration in the Atmosphere Relative to a Business-As-Usual Case...	3.6
3.4 Uncertainties in Model-Estimated CO <sub>2</sub> Concentrations for a Baseline and Four Policy Cases Computed by the MIT IGSM Model .....	3.6
3.5 Graphically Depicting Multiple Elements of Uncertainty .....	3.7
3.6 Variability in Geographic Distribution at the Year 2050 of Standard Deviations in Temperature Changes from AITMES Scenario Within the MAGICC/SCENGEN 5.3 Model ...	3.8
4.1 UC Process for iRESM .....	4.1
4.2 iRESM Framework .....	4.4
4.3 iRESM Component Models Determining Electricity Prices .....	4.5

## Tables

4.1 Intersection of Key Midwest Issues and iRESM Research Questions.....	4.3
4.2 Potential Decision Criteria for Regional RPS Decision.....	4.4
4.3 Summary of Uncertainty Source Identification for Electricity Prices .....	4.6

# 1.0 Introduction

All predictive science entails uncertainty. It is an awareness of this uncertainty that guides the prudent application of mathematical models in real-world decision-making environments. In decision making on climate change adaptation and mitigation at regional scales, uncertainties are profound and pervasive, and the stakes are high (*Nature* 2010). For example, the fates of ecosystems and species and human decisions such as those regarding food production, water resource management, and energy supply and demand are inextricably linked. Climate change impacts on natural and human systems and adaptation and mitigation actions will affect economic development and financial flows. Finally, the fact that mitigation and adaptation decisions will often require tradeoffs between and among human and ecological objectives completes the picture of the daunting challenges facing regional decision makers: risks, complexities, uncertainties, and tradeoffs.

This paper describes the results of new research to develop an uncertainty characterization (UC) process to help address the challenges of regional climate change mitigation and adaptation decisions. This research is being carried out as part of the integrated Regional Earth System Model (iRESM) initiative, a new scientific framework developed at Pacific Northwest National Laboratory to evaluate the interactions between human and environmental systems and mitigation and adaptation decisions at regional scales. The framework integrates a regional climate model; a regional energy-economy model; and highly spatially resolved models of crop productivity, building energy demands, electricity infrastructure operation and expansion, and water supply and management. The iRESM framework is intended to help regional stakeholders (scientists as well as decision makers) understand the consequences of climate change as well as the consequences of policies to mitigate or adapt to such change within regions. The initiative has developed the following four science questions to guide its research:

- How do intrinsic regional characteristics shape, enhance, or constrain regional mitigation and adaptation opportunities?
- How do projected changes in mean climate versus climate extremes affect the development of adaptation and mitigation strategies?
- How might interactions between management decisions and natural processes contribute to rapid or nonlinear changes, and do they contribute to climate feedbacks?
- How will adaptation and mitigation strategies interact in the next few decades in terms of achieving their respective goals?

As pointed out above, an important consideration for modeling any policy addressing either mitigation or adaptation is that the mechanisms linking such Earth system changes are imperfectly understood and in some cases deeply uncertain. Therefore, any model or set of models describing feedbacks and linkages must include methods for characterizing this uncertainty, quantifying it accurately to the extent possible, and depicting it in a manner understood by its intended audience. In a departure from the emphasis of much climate change science, the UC process under development does not focus not on improving regional predictions per se. Rather, it focuses on identifying and characterizing the key uncertainties in data inputs, individual model structures, and coupled models to determine the robustness of the framework's results, thereby providing insight for researchers and decision makers alike. In addition, to meet the twin requirements for scientific accuracy and policy relevance, the characterization of uncertainty must be accomplished in a computationally feasible, transparent, and defensible manner.

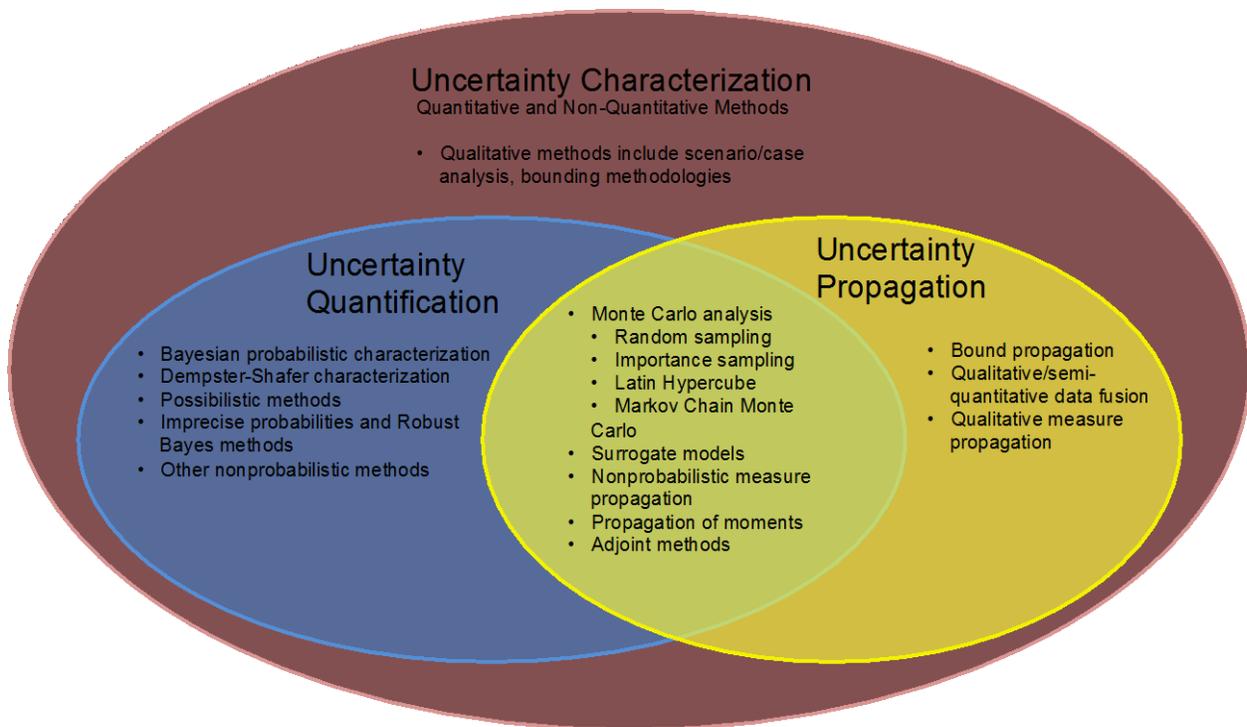
This white paper is intended as a resource to establish a common understanding of methodological options for UC among investigators from diverse scientific fields and to provide the foundation for a set of recommendations regarding an appropriate UC process for the initiative. The paper is structured as follows: Section 2 first describes a recommended taxonomy of uncertainty for the iRESM framework and then summarizes the wide range of methods available for UC, quantification, and propagation. Section 3 reviews alternate methods for utilizing UC results in decision making. Section 4 describes the recommended UC process for the iRESM initiative and the results to date from applying the process to the first iRESM pilot region.

## 2.0 UC: Taxonomies and Methods

### 2.1 Taxonomies of Uncertainty

Although all sources of uncertainty affect the stakeholder’s confidence in a prediction, some sources are more amenable than others to characterization—particularly quantitative characterization. Perhaps because of this, the phrase “uncertainty quantification” (UQ) is often used interchangeably with “uncertainty characterization,” (UC) when it is in fact a subset of the larger UC space (i.e., there are semi-quantitative and qualitative methods of UC). Uncertainty propagation (UP) methods are also a subset of the UC space, and some methods for UP are also considered UQ methods. The following definitions are provided to clarify the differences between UC, UQ, and UP. Figure 2.1 provides a Venn diagram to illustrate the organization within the UC space and lists some of the specific methods in each subset.

- UC is any proposition that measures, quantitatively or qualitatively, the degree of uncertainty associated with a parameter level, prediction, or other entity for which a true but unknown value or outcome exists.
- UQ is that subset of UC approaches in which quantitative measures are defined over the space of possibilities for uncertain parameters, predictions, etc. Probabilistic characterization is the most common means of UQ.
- UP is the mechanical means of making inferences about the UC of output parameters and predictions based on the UC established for input parameters and models.



**Figure 2.1.** The Relationships Between UC, UQ, and UP

In addition to clarifying the UC space as shown in Figure 2.1, distinguishing between uncertainty types is helpful for defining and communicating the UC process for the iRESM initiative. Foremost, a distinction made in numerous domains (Paté-Cornell 1996; Kaplan and Garrick 1981) is that of *aleatory* uncertainty versus *epistemic* uncertainty. Aleatory uncertainty, sometimes referred to as stochastic or random uncertainty, is that which is (as a practical matter) inherent in the system under study. This uncertainty is considered to be an attribute of the system itself and cannot be narrowed through increased knowledge on the part of the analyst. Aleatory *variability* is perhaps a more suitable phrase since this form of uncertainty reflects the random variability in the attributes of the system, such as the behavior of the members of a population, random variations between engineered/physical systems, or variations in meteorological conditions.

In contrast, epistemic uncertainty reflects a state of limited knowledge about the system on the part of the analyst. Epistemic uncertainty can, in principle, be narrowed or eliminated through acquisition by the analyst of additional information. For example, the value of a physical parameter (that has a precise, objective, but unknown value) entering a model can be subject to epistemic uncertainty.

As in most taxonomies, the line between these two varieties of uncertainty is not a bright one. As a practical matter, the uncertainty that is ultimately experienced by the analyst or decision maker has the same practical implications regardless of whether it is irreducible and inherent in the system or stems from limited knowledge. The treatment of these two types of uncertainty, in a mathematical sense, is often similar also. Probabilities, interpreted in a classical frequentist sense, provide a natural framework in which to accommodate aleatory uncertainty, and while non-probabilistic approaches to characterizing epistemic uncertainty have been proposed and sometimes adopted (see Section 2.2), probability theory also remains the most widely applied framework for epistemic uncertainty.

However, notwithstanding the dull line between these types of uncertainty, there do exist some forms of analysis that incorporate both types and in which the distinction needs to be carefully preserved. For example, climate modelers are aware that short-term forecasts are sensitive to initial conditions and inherent climate system variability, which are essentially aleatory, but that longer-term climate forecasts are dominated (for given climate forcing assumptions) by uncertainties in model structure and parameterizations, which are essentially epistemic. To account for epistemic uncertainty, perturbed physics ensembles are run of individual models with sampled parameter values, and multi-model ensembles are used to account for differences in model structure (Tebaldi and Knutti 2007). For an example from human systems modeling, the aleatory variability in the behavior of a population may need to be specified as an element of an economic model, yet there may exist epistemic uncertainty about the mean and shape of that aleatory distribution. In this case, uncertainty in the parameters of the aleatory distribution (mean and variance, say) may each need to be characterized by epistemic probability distributions.

Going beyond these classical concepts, other taxonomies have sought to address the uncertainties associated with use of physical, engineering, and social models to predict the behavior of complex, inhomogeneous, self-interacting systems (e.g., Paté-Cornell 1996; Budnitz et al. 1997). In some technical disciplines, uncertainty types and the means by which they are characterized have become so well established that uncertainty analysis techniques have become standardized and sometimes even proceduralized (ANS/IEEE/NRC 1983; NRC 1990; Budnitz et al. 1997). Where they exist, standardized methods have tended to associate UC methodologies with uncertainty types. Yet, notwithstanding the maturity and routine application of various methodologies in certain domains, there remain outstanding

conceptual questions that are frequently debated within technical communities about the most appropriate, defensible, and comprehensible means of UC (Shafer 1976; Zadeh 1981; Unwin 1988; Helton et al. 2008).

Climate science is among those domains in which growing attention is being focused on the conceptual and practical issues of uncertainty. Many of these issues mirror discourse in other science, engineering, and policy domains, while some have particular significance in the arena of climate mitigation and adaptation. There are several proposed taxonomies of uncertainty in the climate-related decision domain (Moss and Schneider 2000; Rotmans and Van Asselt 2001; Katz 2002; Peterson 2006; Wilby and Dessai 2010). For example, Rotmans and Van Asselt (2001) identify the following categories in the context of integrated assessment modeling:

- *Social uncertainties*: Those associated with describing social behavior of people in terms of demographics, consumption, migration, and urbanization.
- *Economic uncertainties*: Those associated with the production and consumption of resources, capital, and labor.
- *Environmental uncertainties*: Those associated with physical, biological, and chemical transformation of substances and their penetration of the natural environment.
- *Institutional uncertainties*: Those associated with the gamut of policy options and measures in terms of financial measures, legislative measures, education, and research and development programs.

Moss and Schneider (2000) suggest an uncertainty taxonomy that reflects the accumulation of uncertainties in the analytical chain associated with estimation of climate impacts. The uncertainty contributors that “cascade” to produce the overall uncertainty are those associated with: emission scenarios, carbon cycle response, global climate sensitivity, regional climate change scenarios, and impacts. Wilby and Dessai (2010) suggest a different cascading taxonomy of uncertainties comprising the following elements: knowledge of future societies, greenhouse gas emissions, climate modeling, defining regional scenarios, assessing impacts, assessing local impacts, and adaptation responses. Katz (2002) refers to a taxonomy that is more reflective of the fundamental uncertainty types discussed below for the iRESM framework: measurement error (classical issues of random and systemic bias in parameter estimation), variability (*aleatory* in the vernacular introduced above), model structure (validity and completeness), and scaling aggregation (modeling uncertainty associated specifically with application of a model at various spatial and temporal scales).

It is interesting that in the context of the graphical display of uncertainty, Moss and Schneider (2000) also suggest a de facto taxonomy of uncertainty that, while not explicitly tied to climate, does not derive from precedent in other domains. Rather than the typical reductionist/analytic approach of delineating sources of uncertainty associated with the elements of an analysis, these authors characterize the net uncertainty associated with a study “finding” (a discrete outcome) in terms of simple 10-point scales for each of four dimensions (paraphrased): (1) the extent and strength of theory supporting the finding; (2) the availability of supporting model results; (3) the extent of supporting real-world observations; and (4) the degree of consensus within the knowledgeable technical community on the underlying theories, models, and observations. Both of Moss and Schneider’s approaches were developed for application in expert judgment-based assessments—and while such dimensions would unlikely be distinguished explicitly in the analytic propagation of uncertainty, their second taxonomy may provide a conceptual framework for the *a posteriori* attribution of uncertainty factors relevant to the analytical predictions.

Finally, Moss and Schneider (2000) also suggest examining sources of uncertainty in continuous quantities, grouped according to their origins in data, models, and other miscellaneous sources such as definitional imprecision or inappropriate spatial/temporal units.

As for the examples above, the taxonomy of uncertainty recommended for the iRESM initiative needs to address the challenges of its specific domain. The iRESM framework will couple a regional Earth system model of atmosphere, land, and ocean processes (RESM) with a regional integrated assessment model of socioeconomics, the energy economy, water supply and demand, and land use/land cover change. Depending on the research question, these models will interact with one or more spatially explicit sectoral models. These sectoral models include climate-sensitive building energy demand, energy infrastructure expansion and operation, crop productivity, and water supply/management. Due to this architecture, the UC process must address the sources of uncertainty in the individual models as well as in the coupled models.

In exercising a model or set of coupled models—as in the iRESM framework—to make predictions about real-world events, there are several fundamental reasons for questioning the accuracy of those predictions for uncertainty. These reasons are associated with the following questions:

1. Are the model input parameters accurately quantified?

We can represent a model as a function or mapping  $f$  that associates the values of some vector of input parameter values  $\underline{X}$ , to some vector of output parameter values (the predictions)  $\underline{Y}$ :

$$\underline{Y} = f(\underline{X}). \quad (1)$$

To exercise this mapping, the input parameters ( $\underline{X}$ ) need to be quantified, perhaps empirically or based on upstream models. Uncertainty as to whether these parameters have been accurately estimated is the first general source of uncertainty about the accuracy of the model outputs ( $\underline{Y}$ ). Limited confidence in the quantification of input parameters may be rooted in the fact that the supporting data are sparse, noisy, ambiguous, or possibly processed through data analysis models of uncertain validity.

The aggregate uncertainty associated with quantification of the input parameters constitutes this first class of uncertainty. This, of all forms of uncertainty, is generally the most accessible to quantification and analysis (to be discussed below).

2. Is the model skillful?

The question underlying this contribution to uncertainty is: Given the scope of the analysis and the phenomena and factors it is intended to account for, is the model an accurate one? In terms of Equation 1, this can be reiterated as the question of whether  $f$  is an accurate mapping between  $\underline{X}$  and  $\underline{Y}$ , representing valid science or statistical relationships. In general, a model would be considered valid if the mapping accurately mirrors the relationship between  $\underline{X}$  and  $\underline{Y}$  that would, in principle, be produced experimentally or through real-world observation (although this is not often a practically implementable criterion for assessing validity). For instance, the validity of a model built on data generated under conditions that differ from those relevant to the predictive problem (e.g., historic conditions used to assess future scenarios) is uncertain. Another consideration is that, depending on the nature of the model, the effect of aleatory variability in the real-world system needs to be discounted before meaningful comparison can be made with model predictions. That is, where a model does not explicitly address aleatory variability

(through, say, the construction of stochastic probability distributions), then the uncertainties implicit in the predictions of the model must at least exceed the irreducible aleatory variability of the subject system, regardless of the model's skills. Insights into the skillfulness and even the completeness of models (the "known unknowns" part—see item 3, below), as well as the scope of aleatory variability versus epistemic uncertainty, can sometimes be obtained through intermodel comparisons that run "competitor" models of processes or phenomena with the closely matched assumptions. (See Section 2.2.2.6.)

The notion of prediction accuracy cannot always be completely disentangled from the vagueness inherent in the model outputs. That is, the precision with which a model output parameter can be identified with a real-world measurable influences the degree of uncertainty associated with the model's validity. So, for instance, a model that predicts, or is interpreted as predicting, "high market damage" would be considered to have uncertainty characteristics that differ from one intended to attach greater quantitative precision to the predicted degree of market damage.

### 3. Is the model complete?

This third form of uncertainty is associated with the question of whether the model (or model suite) has captured all the phenomena and factors necessary to produce accurate predictions, given the intended scope of application. The completeness issue has two components, which were memorably articulated in another context by a U.S. Secretary of Defense as the *Known Unknowns* and the *Unknown Unknowns*. The former category would capture phenomena and factors that are known to be absent from the model but that are perhaps not sufficiently understood to allow their incorporation into the analysis. The latter category, in effect, simply contains the acknowledgement that the modelers may not have thought of everything that could influence outcomes. This type of uncertainty is the most problematic from a UC perspective in that it is, in principle, beyond analytic resolution.

While these categories are generally accepted across technical communities as the principal contributors to uncertainty in system modeling, there is a strong case for one additional category, particularly in the context of large, complex computer-implemented models:

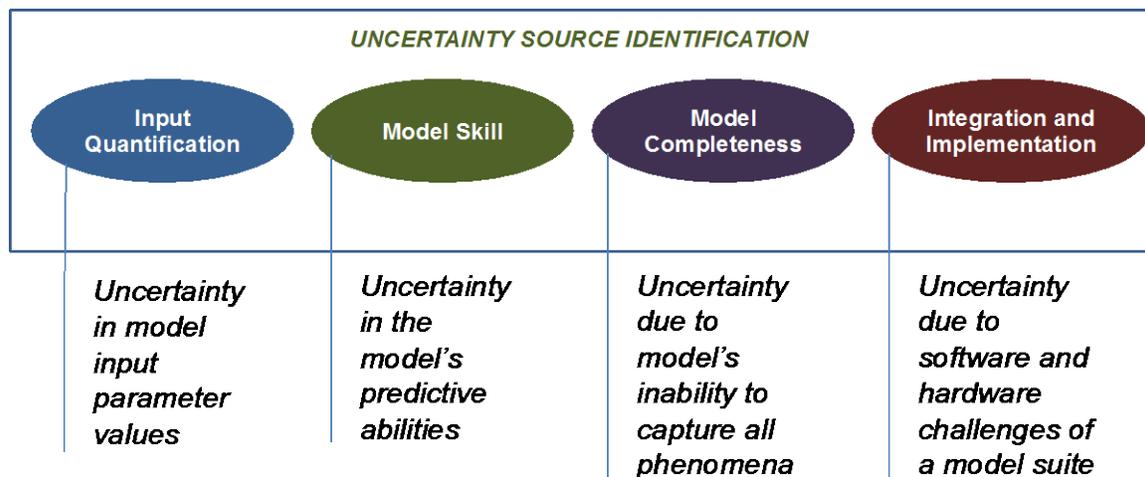
### 4. Is the model accurately implemented?

This form of uncertainty is rooted in the question of the veracity of the calculations underlying implementation of a model or model suite. That is, have the calculation elements implied by the model and by the bridging requirements between models been executed in accordance with the intent of the modelers and analysts? The more complex and extensive a set of models, the greater the likelihood that software errors, hardware faults, and analyst mistakes will occur. Inability to confidently verify all such aspects of a large, interacting set of calculations is the source of this fourth form of uncertainty.

These, then, are the four fundamental categories of uncertainty associated with system modeling. Note that they are germane to characterizing confidence in the values of the immediate modeling outputs and to derivative insights gained from comparison of model inputs and outputs. For instance, while we might expect insights such as "the value of  $X_4$  strongly influences the value of  $Y_6$ " or "the directional influence of  $X_2$  on  $Y_3$  is positive" to be more robust than predictions of output parameter levels, they are insights that can nevertheless be subject to the four sources of uncertainty. For this reason, even where the objective of a modeling effort is confined to gleaning insight concerning the relationships between elements of a system (such as complex human-environmental systems) rather than making predictions of system behavior, the robustness of those insights under uncertainty cannot generally be guaranteed.

Therefore, there are few circumstances under which analytic uncertainties need not be systematically characterized to gain valid and defensible insight, particularly for complex systems.

Having identified these four general uncertainty sources, it should be noted that, in a practical setting, there is not always a crisp boundary that separates them. For instance, uncertainty regarding the validities of competing sub-models may often be treated in an analysis by establishing a single parameter in the parent model, the value of which is varied to represent the competing sub-models. In this way, the modeling uncertainty is treated as a parametric uncertainty. Similarly, issues of model completeness are not always distinguishable from those of model validity. Despite these issues, these four sources of uncertainty represent the taxonomy recommended for the iRESM UC process as shown in Figure 2.2.



**Figure 2.2.** Taxonomy for Uncertainty Source Identification in iRESM

## 2.2 Methods for UC

Having identified a taxonomy of uncertainty for iRESM, the question arises of how to characterize these uncertainty types within the modeling framework and, ultimately, how to convey them to stakeholders. This section identifies prospective means of UC and discusses the candidate mathematical and structural schema in which uncertainty might be represented.

Again, there exist numerous reviews of prospective bases for UC (Apostolakis 1989; Unwin 1989; Dubois 2006; Helton et al. 2006), some specifically addressing the climate domain (Rotmans and Van Asselt 2001; Katz 2002; Webster et al. 2003; Morgan et al. 2009; Swart et al. 2009). Here, the intent is to provide a brief overview of UC methods that are potentially applicable to climate and integrated assessment modeling. Our ultimate objective is identification of the appropriate mix of UC methods for iRESM. The following is a list of questions to be answered during methods selection. Keeping these questions in mind, the description of each UC method to follow includes a brief summary of strengths, weaknesses, and other considerations germane to the iRESM environment.

- Which approaches to UC are more attuned to specific sources and types of uncertainty inherent in the illustrative policy decisions, and when is UQ a robust means of UC?

- What levels of arbitrariness (real or perceived), defensibility, and reproducibility are associated with a specific approach to UC? How transparent is a given approach to UC? How will it be comprehended and interpreted by a decision maker?
- What has been the experience with particular approaches to UC across domains for similar policy decisions, good or bad?
- What constraints does the need to integrate across multiple sources and characteristics of uncertainty place on the selection of UC methods?
- What are the computational implications of particular approaches to UC—is a given approach practical given the nature of the iRESM model suite? How can the computational and analytic requirements of UC help inform software engineering and model development choices?
- What methods are available for eliciting a reasonable range of values and probabilities for quantitative relationships in models, and on what bases should they be selected given the nature of iRESM?
- What decision strategies, frameworks, or methods complement each approach to UC for the illustrative policy decisions, and how might they address incomplete confidence on the part of a decision maker in a given approach to UC?

## 2.2.1 Non-Quantitative Methods for UC

### 2.2.1.1 Bounding Analysis

Bounding analysis is a simple, conservative, non-quantitative method for UC that requires an uncertainty range to be established for each input parameter to a model. These ranges are then propagated through the model to establish the corresponding ranges on the output parameters. No measures are attached to the ranges (such as probability densities), and propagation consists of establishing the full range of each output parameter that can be generated by varying the input parameters over their respective ranges (constrained by any dependences between inputs). Where the relationships between outputs and inputs are pairwise monotonic, establishment of the output ranges can be straightforward; otherwise, some sampling methodology is required. Cast in the notation of Equation 1, the upper and lower ends of the range of output parameter  $Y_i$ ,  $Y_{iu}$ , and  $Y_{il}$ , respectively, can be expressed as

$$Y_{iu} = \sup_{\underline{X}} f_i(\underline{X}) \quad (2a)$$

$$Y_{il} = \inf_{\underline{X}} f_i(\underline{X}) \quad (2b)$$

where the supremum (smallest upper bound) and infimum (greatest lower bound) operations are effected over the uncertainty ranges of the input vector  $\underline{X}$ . Where there is uncertainty regarding the validity of the model  $f(\underline{X})$  itself and competing models exist, then the evaluation of the output uncertainty ranges may incorporate switching between models also.

## Strengths

The limited amount of information needed to characterize uncertainty is the principal strength of this approach. It places no demand on an analyst to produce highly structured representations of uncertainty (such as a probability density), requiring only the establishment of uncertainty bounds on each parameter.

## Weaknesses

The weakness of the approach lies in the questionable interpretation of the uncertainty bounds. Such bounds on the input parameters are sometimes interpreted as the range that, with *reasonable* or high certainty, captures the true value of the parameter. However, since the propagation mechanism is a simple bounding process, then whatever interpretation is placed on the input parameter ranges is not preserved in the output parameter bounds. That is, if we hypothetically placed a probabilistic interpretation on *reasonable certainty* (0.95 probability, say, of capturing the true parameter value) then the nonprobabilistic nature of the UP precludes a similar interpretation of the output ranges. Indeed, this methodology tends to produce highly conservative (i.e., wide) output uncertainty ranges as a consequence of a process in which conservatisms are progressively compounded during the propagation process.

While bounding analysis can often form a basis for real-world, every-day decision making in noncomplex problems involving limited numbers of variables, this type of approach is seldom preferred in technical, analytical contexts, although there is some precedent for its use (see Lipinski et al. 1985). Nevertheless, there may be a rationale for blending aspects of simple bounding analysis into a broader framework of UC where, for example, there is limited basis for the assignment of probabilities to competing models.

### 2.2.1.2 Scenario Analysis

Scenario analysis is an approach to UC that is often used in the climate change domain. Promulgated by the Intergovernmental Panel on Climate Change (IPCC 2000), it involves the identification of a set of scenarios that, in combination, characterize a range of alternative futures. Each scenario is structured around a narrative, or *storyline*, that relates a coherent combination of future conditions. Elements of a storyline for the IPCC emission scenarios, for instance, include a projection of population size and demographics, the pace of economic growth, the rate of development of energy-efficient technologies, the distribution of energy resources, and the mix of land use. Each scenario is defined as a combination of quantitative point estimates for these factors and is intended to portray a consistent possible future. Subsequent analysis of climate impacts is then subject to the boundary conditions defined by each scenario.

A scenario analysis shares with bounding analysis the feature that there is no attempt to probabilistically weight the defining scenario conditions; however, it differs in important respects also. While a bounding analysis is intended to encapsulate the predictions associated with a large (in principle, infinite) set of combinations of model input parameters, the scenario approach is one in which specific combinations of projected conditions are selected *a priori* to, in some sense, span a range of possible futures. In fact, the probability associated with any one future is infinitesimally small since it reflects a combination of point values of key parameters. Nevertheless, the intent is that any conclusions drawn with regard to individual scenarios are interpolatable to address intermediate but unarticulated scenarios.

## Strengths

The strength of the scenario approach is, first, a computational one. Through the *a priori* identification of a manageable number of discrete scenarios, this approach avoids combinatorial problems associated with the range of possible futures, and thus the analytical burden in predicting future conditions is substantially reduced. A second advantage is that the storylines inherent in the scenario definitions are more easily communicated to stakeholders than future conditions that are the product of a sampling algorithm applied to the entire space of possibilities.

## Weaknesses

If viewed purely as a means of UC, as opposed to a narrative device for relating a range of possible futures, the scenario approach has significant flaws. Scenario definition is heuristic rather than systematic. Unless it is accompanied by a detailed review of the literature to identify ranges and distributions for key driving forces and outcomes, there is no basis for confidence that the resultant scenario set captures the range of possible climate impacts or uncertainty that would be necessary to support decision making. Moreover, in picking combinations of point estimates from a high-dimensional space of parameter possibilities, it is unlikely that the potential synergies between factors and the resultant range of predicted outcomes will be acceptably complete.

### 2.2.2 UQ Methods

There has been a uniform appreciation that the source and type of an uncertainty dictates the degree to which various means of quantitative characterization are useful and practical. For instance, uncertainty associated with the question of whether a model is a complete representation of potentially important phenomena is a form of uncertainty that is the most resistant to systematic characterization. With increasing mathematical sophistication, a variety of quantitative frameworks for representing uncertainty have been proposed and advocated over the past 40 years. Yet, while mathematical models such as the Dempster-Shafer (D-S) Theory of Evidence (Dempster 1968) and Zadeh's Possibility theory (Zadeh 1978) have been put forth to remedy the perceived shortcomings of probabilistic methods, it is probability theory that remains the prevalent and most familiar (in most technical communities) methodological basis for addressing uncertainties.

There are sound reasons for the retention of probabilistic methods (familiarity being among them); yet, there are aspects of uncertainty and its measurement that are not shared by the measurables typically encountered in other technical domains. While researchers seek to quantify uncertainty, it is a quantity that is not objectively measurable. Probability, as a characterization of uncertainty, is generally interpreted in the so-called Bayesian sense. That is, a probability ultimately measures an analyst's (or some aggregate of analysts') conception of uncertainty, interpreted as a *degree of belief* that some uncertain proposition is true, such as the proposition that the value of a physical parameter lies within some specified quantitative range. While there exist theoretical bases for the systematic incorporation of objective data into the formulation of probabilities, sparseness of available data or ambiguity in its interpretation generally renders probabilities highly subjective. It is certainly the case that the so-called *frequentist* interpretation of probability—in which probabilities are associated with the outcomes of repeated, random trials—finds little scope for application in most real-world decision environments, including the climate arena.

### 2.2.2.1 Probabilistic Analysis

Probabilistic approaches to uncertainty analysis are, by now, sufficiently prevalent and conventional that the term *uncertainty quantification* is often considered synonymous with probabilistic analysis. The history of probability theory as a basis for what we would now term *uncertainty analysis* spans four centuries; however, the advent of the so-called Bayesian interpretation of probability in the mid-20<sup>th</sup> century has resulted in a significant expansion of the use of probabilistic methods. Unlike classical statistical methods that revolve around the availability of stochastic sampling models to produce and propagate classical confidence intervals, the Bayesian interpretation (de Finetti 1972; Martz and Waller 1982) has permitted probabilities to be employed as the means of expressing informed opinions and uncertainty in problem domains for which classical statistical models are unavailable. Probabilities that reflect the beliefs of an informed individual, or set of individuals, and thus represent epistemic uncertainty, are the cornerstone of probabilistic uncertainty analysis. In brief (and with some caveats), a probability density  $P(x)$  defined over some parameter space reflects the relative weight of belief assigned by the expert(s) that the true value of the parameter  $X$  is about  $x$ . Similarly, probabilities can be assigned to competing models of a phenomenon (NRC 1990; Budnitz et al. 1997).

Expressed mathematically, the probability distribution reflecting uncertainty in a model output parameter  $Y_i$  is related to the joint distribution over model input parameters by (referring to Equation 1)

$$P(Y_i) = \int P(\underline{X}) \cdot \delta[Y_i - f_i(\underline{X})] \cdot d\underline{X} \quad (3)$$

where  $\delta$  is the Dirac delta function. However, this simple conceptual expression generally belies a resource-intensive, numerical process for its realization involving statistical sampling from the input joint probability distribution, propagation of inputs through a series of complex models, and reconstitution of the model's results to produce the output probability distributions. (Issues associated with the propagation of probabilities through complex models are addressed in Section 2.3.)

There also exist less-quantitative methodologies for UC that are, nevertheless, founded in probabilistic principles. Rather than seeking to establish full probability distributions or assign precise probability values to alternative hypotheses, these semi-quantitative approaches define coarse probability categories that may be defined on scales such as *Very High*, *High*, *Medium* ..., or as order-of-magnitude probabilities, such as *One in a Thousand Chance*. To allow such probabilistic characterizations to be propagated through models (generally simple models in these cases), combinatorial rules are established that reflect an underlying probabilistic logic. Such methods often find application in hazard studies of engineered systems and simplified risk assessments (CCPS 2008).

### Strengths

As noted, probabilistic uncertainty analysis has been a core methodology in numerous application domains, including climate and integrated assessment modeling (Reilly et al. 1987; Scott et al. 1999; Sokolov et al. 2009; Webster et al. 2009). A key strength of Bayesian probabilistic methodology—one on which its mathematical foundations rests—is the conceptual relationship between probability and the notion of event frequencies. Even though the Bayesian interpretation explicitly rejects the notion that event frequencies (pertaining to the outcomes of stochastic events, such as the toss of a coin) are the appropriate conceptual basis for interpreting probabilities, stochastic concepts nevertheless help standardize the measure of probability. So, for instance, a Bayesian probability of one-sixth can be still

be gauged by an appreciation of the chances of any one side of a die landing face up. So, while Bayesian probabilities reflect degree of belief but not the relative frequencies of the outcomes of random events, an association with frequentist notions helps standardize the meaning and appreciation of numerical probabilities. This ability to interpret probability is a critically important attribute both for the generation of probabilities by analysts and for their comprehension by stakeholders.

A further and pragmatic advantage of probabilistic methods lies in the wealth of methods and supporting tools for the elicitation of probabilities from subject matter experts (addressed in Section 2.2.2.5 ) and for the propagation of probabilities through complex models (Wojtkiewicz et al. 2001; Helton et al. 2006).

## **Weaknesses**

The principal weakness of Bayesian probabilistic methods resides in the common perception that there is a lack of defensibility and a degree of arbitrariness in the formulation of probabilities. This is an acute problem when the insights of an analysis are sensitive to the probability distributions selected.

There has always existed an uncomfortable tension between, on the one hand, the notions promulgated by Bayesian philosophers/theoreticians and, on the other, the practical needs of those who seek to incorporate probabilistic analysis into public policy and decision making. Viewed as a means for an individual to order and systematize his/her beliefs to support personal decision making, the defensibility and transparency of probability assignments is de-emphasized. Therefore, the prevalent belief among Bayesian philosophers—that probabilities cannot be *wrong* as they express personal weights of belief and that two individual confronted by the same evidence need not necessarily be expected to produce similar probabilities—are not serious indictments of the method. However, in the arena of public decision making, such views have serious implications for the acceptability and value of probabilistic methods.

The community of Bayesian analysts who put these methods to practical application in, for instance, risk analysis and related decision making (e.g., Kaplan 1990) tend to hold a view that contrasts with the community of more abstract thinkers. Particularly, the notion that a given body of objective evidence should lead all rational analysts to produce similar probabilistic characterizations of uncertainty is a necessary perspective if probabilistic methods are to have practical public value. Therefore, much of the research and methods development in uncertainty analysis revolves around the means of establishing defensible, transparent, and scrutable processes for formulating probability distributions and their dependences (correlations for the purposes of Monte Carlo sampling) based on the elicitation and processing of informed opinion (Unwin et al. 1989; Wheeler et al. 1989; Budnitz et al. 1997; see also Section 2.2.2.4.). Parallel tracks in research focus on the question of how the results of a probabilistic analysis may be most confidently applied such that the insights are robust to changes in the input probabilities (Regan et al. 2005; Lempert and Collins 2007). Also, much of the methodology development revolving around alternative, nonprobabilistic means of UC (to be addressed) has been motivated by these perceived weaknesses in Bayesian methods.

Finally, a practical weakness of probabilistic analysis lies in its computational demands. The numerical implementation of Equation 3 generally demands Monte Carlo techniques involving multiple implementations of underlying models (from hundreds to hundreds of thousands of model runs, depending on the number of variables and complexities of the models). When the models have

significant run times, random sampling of the inputs can be impractical—and while there exist more economic sampling methodologies (discussed in Section 2.3.1), there will likely remain computational challenges for iRESM given the nature of the model inventory.

### 2.2.2.2 D-S Theory of Evidence

The D-S theory of evidence, which emerged in the 1970s (Dempster 1968; Shafer 1990), is touted by many analysts as providing a more natural framework than the Bayesian one to represent informed opinion and epistemic uncertainty. To the mathematician, D-S is a generalization of probability theory, while to the practical analyst, it is a framework that has the benefit of being less demanding on the provider of informed opinion, not requiring probability distributions that even the provider may consider arbitrary. The mathematics of D-S theory will not be expounded here; rather, some of the salient features that may render it an attractive approach to UC are outlined below.

Consider some parameter that may take on one of several discrete values (based, perhaps, on the predictions of competing models) where there is uncertainty as to which is the *true* (or at least the best) value. Assume there is maximal ignorance; that is, no apparent reason to favor any one of the estimates. In probability theory, this would demand the assignment of equal probability to each estimate. However, this is an arbitrary characterization of uncertainty since the parameter space might easily have been discretized differently (e.g., grouping models of certain types together to produce averaged results over a given class of models). (Note that this arbitrariness would be manifested in continuous parameter spaces as the question of how to choose the space metric—a flat probability distribution differs from a log-flat distribution.)

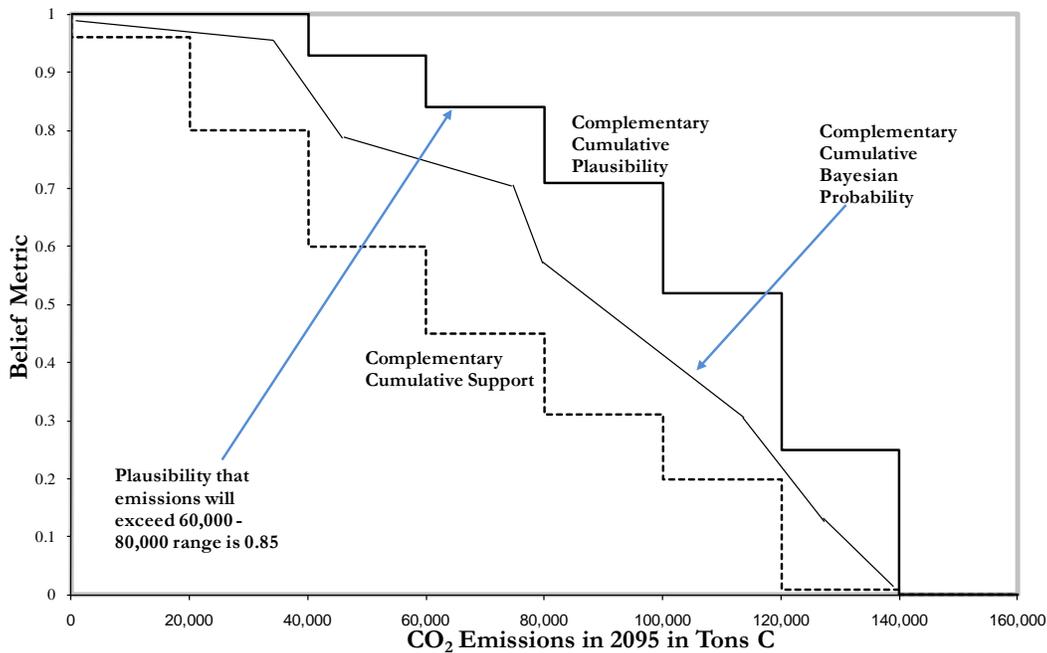
The fundamental quantity in D-S theory is the *probability mass*, in contrast to the more conventional probability of Bayesian analysis. Like a probability, a probability mass represents a weight of evidence. However, whereas conventional probability theory requires a normalized probability to be attached to every member of the set of possibilities (i.e., parameter values), D-S methods require only the attachment of normalized probability masses to selected subsets (i.e., combinations of possibilities) as dictated by the evidence. That is, there is no obligation to distribute the probability mass among the subset members. If there is complete ignorance as to which parameter value should be preferred, then all the probability mass can be applied to the set of values as a whole, without the need to distribute probability masses to the individual values.

Once probability masses have been assigned to selected subsets (these subsets are the so-called *focal elements* of the full set of possible values) via whatever inferences the experts make based on the evidence, then two new types of belief measures can be defined that apply to each individual value in the set of possibilities: *Support* and *Plausibility*. These two measures are calculated from the probability masses. Support of a given parameter value is interpreted as the degree to which the evidence points towards that specific value (or the model that produced it). Plausibility is the degree to which the evidence is *consistent* with that value; or, equivalently, the degree to which the evidence fails to point away from the value. This bifurcation of belief metrics is not present in Bayesian probability theory and is argued by proponents of D-S to more closely reflect the way we form uncertain beliefs.

Support and Plausibility are sometimes referred to as *lower and upper probability measures* (since, in theory, they bound the hypothetical probability of a parameter value that would have been generated from the same evidence), but this is inappropriate terminology in that they do not behave in accordance with

the mathematics of probability. Indeed, it is the so-called Dempster Combination Rule, applied to the underlying probability masses, that dictates the appropriate propagation mechanism in D-S theory.

Figure 2.3 shows a hypothetical D-S representation of uncertainty in CO<sub>2</sub> emission levels for the year 2095 in the form of complementary cumulative belief curves. For comparison, the figure also shows a hypothetical Bayesian representation of uncertainty. Note that the stepwise nature of the D-S representations reflects the (hypothetical) fact that probability masses have been attached to various input parameter ranges rather than to individual parameter levels within each range.



**Figure 2.3.** Hypothetical D-S and Bayesian Representations of Uncertainty: 2095 CO<sub>2</sub> Emissions Level

D-S has found application in numerous domains such as climate modeling (Luo and Caselton 1997; Sadiq et al. 2006), homeland security (Unwin and Fecht 2009), industrial reliability analysis (Parikh et al. 2001) and natural disaster risk assessments (Park 2010), to mention a few.

## Strengths

As already noted, the proponents of D-S theory cite as its principal strength the capability to more naturally (compared to Bayesian probability theory) accommodate the characterization of epistemic uncertainty (Shafer 1990; Dubois 2006). It does not demand the attachment of probabilities to every member of the set of possibilities or, equivalently, does not require the formulation of probability distributions over the parameter spaces. Thus, a principal concern about the probabilistic approach—that of the degree of structure required to characterize uncertainty and the arbitrariness associated with the resultant probability distributions—is obviated.

## Weaknesses

D-S is not a panacea. One critical drawback revolves around standardization of the D-S belief metrics. While the frequency analog of probability lends to the intuitive interpretation of a given probability level, this analogy does not exist for Support and Plausibility metrics (also measured on the interval 0 to 1). That is, an analysis that determines the Plausibility of a given outcome to be 0.75 is not easily interpretable, particularly for a stakeholder who was uninvolved in the analysis. This raises the more general concern of the degree to which a relatively arcane methodology will be understood and trusted by a decision maker.

There exist analytical challenges to D-S also. Since probability masses are defined on the set of subsets of the space of possibilities, and a set with  $n$  elements has  $2^n$  subsets, then the computational challenges of probabilistic analysis (noting that Monte Carlo methods are applicable to D-S also [Helton et al. 2006]) could potentially be exceeded in D-S approaches.

Finally, it should be noted that D-S theory is considerably less mature than Bayesian probability theory, and this is reflected not only in a general lack of familiarity among most technical communities with D-S methods, but also in aspects of the theory itself. So, while there is a well defined algebra associated with the mathematics of probability, there is more than one prospective mechanism for the combination and propagation of D-S measures. While the Dempster Combination Rule has traditionally been the basis for the propagation of probability masses, deficiencies in that rule have been identified (that is, nonintuitive properties) and a series of optional algorithms have been suggested (Senz and Ferson 2002; Smarandache 2004). Consequently, the basis for picking one rule over another would introduce a measure of arbitrariness to which probabilistic analysis is not subject.

### 2.2.2.3 Possibility Theory

Another framework for UC is *Possibility Theory* (Zadeh 1978; Dubois and Prade 1995). Unlike D-S theory, this method is not rooted in probability theory or its generalizations, but finds its basis in fuzzy set theory (Zadeh 1965). Like D-S, proponents of possibility theory argue that it can provide a more natural framework in which to accommodate epistemic uncertainty and informed opinion. Again, the argument in favor of possibility theory revolves around the notion that it places more realistic demands on the providers of informed opinion than the Bayesian approach. While, like probabilistic methods, it does require the construction of quantitative distributions over parameter spaces, a possibility measure—again defined on the interval of 0 to 1—does not seek to characterize the relative likelihood of a given parameter value being true but, rather, only reflects the consistency of that value with evidence. (Note that this is a similar notion to *plausibility* in D-S theory. Note also that there exists a fuzzy set-theoretic equivalent to the D-S *support* measure called *necessity*, which need not be discussed further here.) This metric is viewed as a more credible measure to elicit from experts than an opinion on the relative likelihood of a given parameter value. So, while there are some conceptual and theoretical connections between possibility and probability measures (Unwin 1986; Dubois 2006), the assignment of possibilities does not demand probabilistic thinking.

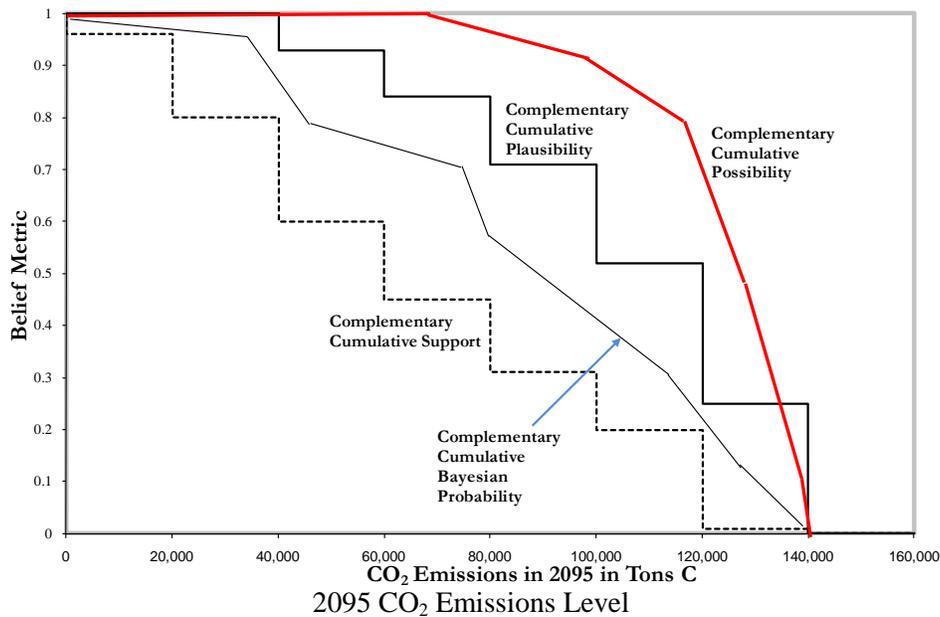
Reflective of the substantially different interpretations of possibility and probability measures, the mathematical means of propagating possibilities is very distinctive from that of probability (although Helton et al. [2008] have investigated the use of Monte Carlo methods in the context of possibility theory). Indeed, fuzzy set theory provides the underlying combinatorial logic for possibility theory. So,

for example, the possibility of the proposition “model A provides the best description of phenomenon 1 and Model B provides the best description of phenomenon 2” is equal to the lesser of: the possibility that “model A provides the best description of phenomenon 1” and, the possibility that “Model B provides the best description of phenomenon 2.” That is, reflective of its fuzzy set theoretic underpinnings, the propagation algebra of possibility measure is based on minmax logic. With reference to Equation 1, the overall propagation logic could be stated as:

$$Po(Y_i) = \max_c \min_{X_{ic}, Y_i=f(\underline{X}_c)} [Po(X_{1c}), \dots Po(X_{Nc})] \quad (4)$$

where  $Po$  is the possibility measure of its argument, the *min* operation is taken over the possibility measures associated with a combination of input parameter levels that produce the output value  $Y_i$ , and the *max* operation is taken over all such input combinations. This expression would be generalized to address uncertainty in the model,  $f$ , itself.

Figure 2.4 augments Figure 2.3 with a hypothetical possibilistic representation of uncertainty in year 2095 CO<sub>2</sub> emissions levels.



**Figure 2.4.** Hypothetical Possibilistic, D-S, and Bayesian Representations of Uncertainty

Possibility theory and variants based in fuzzy set theory have been applied in numerous domains to date (Dubois and Prade 2003; Darby 2009), including climate-related research (Duong 2003; Hall et al. 2007).

## Strengths

Like D-S theory, proponents of possibility theory cite as its principal strength the capability to more naturally (compared to Bayesian probability theory) accommodate the characterization of epistemic uncertainty. As noted earlier, possibility measures do not demand an assessment of relative likelihoods of events but, rather, require only that the consistency of a model or parameter value with evidence be assessed on a quantitative scale.

## Weaknesses

Possibility theory, like D-S theory, suffers from the absence of a sound basis for the standardization of measures. That is, how can uniform meaning be attached to some measure of possibility (what does a possibility of 0.65 mean?) and how can that meaning be conveyed to a stakeholder? Like D-S theory, possibility theory would be viewed as arcane in most technical communities, and certainly so in stakeholder communities.

While the touted strength of possibility theory lies in its distinctive character from Bayesian probability theory, this also presents disadvantages. In particular, there exists no mature, unique basis for updating possibility distributions in light of emerging statistical data. In contrast, Bayesian probability theory provides well established mechanisms for the fusion of informed opinion with statistical data.

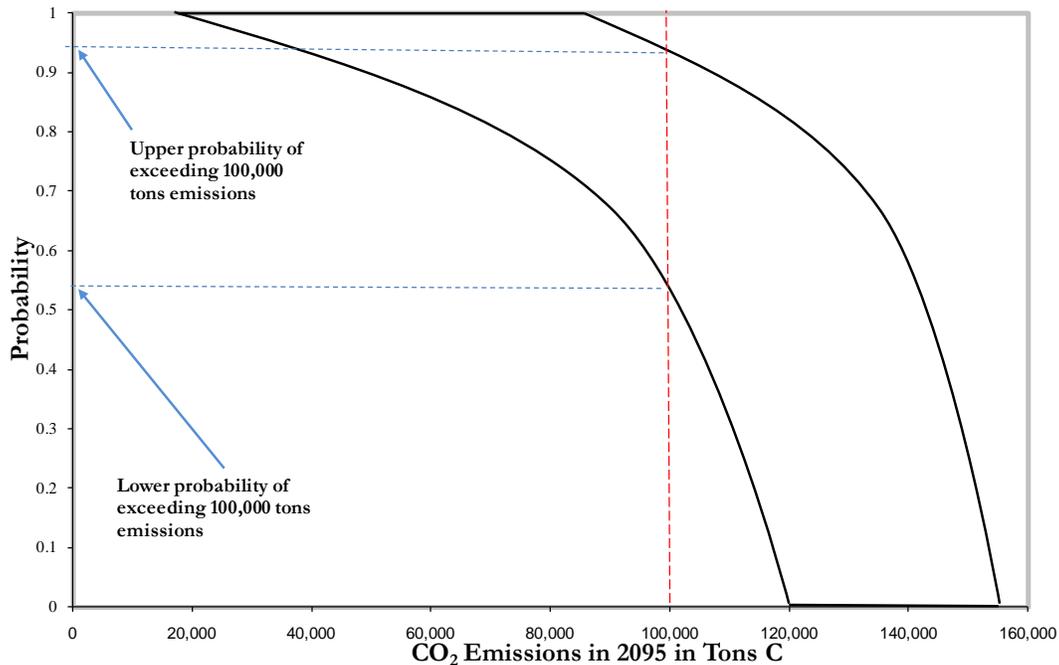
### 2.2.2.4 Imprecise Probabilities and Robust Bayes Methods

As noted under the *Weaknesses* of probabilistic methods, notwithstanding the Bayesian philosophical position that probabilities are themselves representations of uncertainty, concerns about the “uncertainty,” or perhaps more accurately the perceived arbitrariness, associated with the selection of probability distributions are often expressed in public decision-making environments. Against this background, the notion has been promulgated that where probabilities cannot be uniquely specified, then less-precise representations of uncertainty are preferable. One such representation revolves around the concept of *imprecise probability*. Here, rather specifying the probability (P) of an event as a single number, it is replaced by a doubleton that represents the bounds on that probability:  $[P_{\text{Lower}}, P_{\text{Upper}}]$ . This mathematical entity, representing the range of probabilities considered consistent with the underlying uncertainty, then possesses its own combinatorial algebra (Williamson and Downs 1990). (Note that other frameworks also adopt the term *imprecise probability*.)

There are several means of operationalizing such a UC framework (see, for example, Ferson et al. 2003). Among these is the Robust Bayes (RB) approach. The RB concept is that uncertainty in, for example, the magnitude of a physical parameter is represented not by a single probability distribution, but by a set of distributions. Members of this distribution set may be subject to one or more common constraints that reflect some state of knowledge about the uncertain parameter. For instance, all the distributions may be required to have the same mean if there is some *a priori* rationale for constraining the mean. The idea is that there is no basis by which to establish preferences among the members of this distribution set, and so the set itself represents the state of uncertain knowledge. Imagine now that there is a distribution set associated with each input parameter to a model and that each possible combination of input distributions is then propagated through the model. Since each single combination of input distributions produces a single probability distribution over a given output parameter, then executing all combinations of input distributions results in a set of probability distributions over that output parameter. This set then represents the uncertainty in that output.

The relationship between imprecise probabilities and the distribution sets generated by RB methods is as follows. Consider the distribution set associated with a single parameter (either input or output). If all the probability distributions in that set were plotted in, say, complementary cumulative form, then the bounding hull (that is, the envelope) of these distributions would represent the imprecise complementary cumulative probabilities associated with the parameter (see Figure 2.5). This figure is a so-called

probability box or *p-box* representation of uncertainty (the heavy lines representing the sides of the box). The RB approach can, in a sense, be viewed as a sensitivity analysis to alternative input distributions.



**Figure 2.5.** Hypothetical p-Box Representations of Uncertainty: 2095 CO<sub>2</sub> Emissions Level

Although beyond the scope of this review, it can be shown that the imprecise probability/p-box representation of uncertainty is closely related to the D-S representation (which might be intuited from a comparison of Figure 2.3 and Figure 2.5).

## Strengths

The principal strength of this approach is that it obviates concerns associated with the defensibility of selecting any one Bayesian probability distribution to represent an uncertainty. In this method, a set of distributions is selected, with each member of the set considered to be consistent with the underlying uncertainty. A second advantage associated with RB methods is that the core mechanism for propagating uncertainties is a familiar, probabilistic one to which conventional Monte Carlo methods can be applied.

## Weaknesses

While this method removes some of the arbitrariness associated with probabilistic UC, there nevertheless remains the residual issue of which probability distributions to include in the set. The more distributions selected, the greater the computational challenge associated with propagating uncertainty. That is, the familiar, practical impediments to Monte Carlo analysis associated with long model run times is exacerbated in this current approach by the requirement that multiple combinations of input probability distributions be propagated. It may be possible in some circumstances to alleviate such problems by the use of faster running surrogate models (such as response surfaces or reduced-order models) or the use of other computational approximations in UP.

### 2.2.2.5 Elicitation of Informed Opinion

Where historical data or statistical information is deemed incomplete or insufficient to represent uncertainty in a classical, statistical sense, the formal and structured elicitation of informed judgment has substantial precedent as a basis for augmenting sparse or ambiguous data. Where the means of UC is Bayesian, this can take the form of the direct elicitation of probabilities associated with, for example, the likelihoods of parameter values, events, scenario outcomes, or model skills. A subjective probability assessment translates an expert's understanding of uncertainty (i.e., his/her state of information) into a quantitative form—a probability distribution. Expert elicitation has been widely applied in engineering and environmental risk analyses, including climate change (see, e.g., Morgan and Henrion 1990; Morgan et al. 2001), and is recognized as a standard practice for quantifying uncertainty by the IPCC (IPCC 2001) although the IPCC has yet to actually apply the method in preparation of its reports. It should be noted, however, that expert elicitation is usually time consuming and costly (especially if multiple experts must be interviewed). In a standard Bayesian decision analysis process, expert elicitation is only undertaken for those parameters or events whose uncertainty has been identified as significant for the decision based on a deterministic sensitivity analysis of the underlying model or on preliminary assignments of tentative probabilities.

Expert elicitation methods were first developed in conjunction with the development of decision analysis techniques in the late 1960s and early 1970s (Spetzler and von Holstein 1975 is recognized as the seminal work). Since then, they have become standard—and in some domains, proceduralized—tools for UC (e.g., Wheeler et al. 1989; Budnitz et al. 1997). The methods have drawn extensively from research in cognitive psychology (Tversky and Kahneman 1974; Kahneman et al. 1982) showing that, when unaided, individuals tend to use various heuristics when making judgments about uncertainty, resulting in systematic biases in their assessment, such as overconfidence. If these biases are not identified and managed by the interviewer in the elicitation process, then the resulting probability assessment will suffer accordingly. Other challenges in expert elicitation arise from common mistakes made in causal reasoning and conditional probability, even by those trained in probability and statistics (e.g., the “Monty Hall Dilemma”; see Herbranson and Schroeder 2010). Awareness of these pitfalls has been shown to be insufficient for self-correction, although providing experts with some training in probability elicitation prior to a formal assessment has been shown to facilitate the elicitation process and increase expert confidence in the results (Keeney and von Winterfeldt 1991). The overall objective of elicitation is not necessarily to encode the beliefs of the specific expert participants but to use those participants as surrogates for the wider technical community with the goal of capturing the center and breadth of knowledgeable opinion.

While various practitioners have developed specific protocols for different research topics and numbers of experts (e.g., Morgan and Henrion 1990; Keeney and von Winterfeldt 1991; Hora 1996), the core of the individual expert interview has five steps designed to address systematic biases and to produce a distribution that the expert agrees properly represents his/her state of information:

1. motivating
2. structuring
3. preconditioning
4. encoding
5. verifying.

In the motivating step, the interviewer strives to uncover and then remove any motivational biases that may affect the expert's judgment. Motivational biases are typically caused by the fear of consequences, the promise of rewards, or the desire for objectivity. An expert might want to suppress any expression of uncertainty to appear knowledgeable, for example, or resist participating in the elicitation in favor of "objective models" or "hard data." The primary strategy in this step is to explain the importance of the elicitation (e.g., show the sensitivity analysis results) and stress that what is being asked for is the breadth of knowledge, not a prediction or a point estimate. Point estimates about the future are almost certain to be wrong, whereas a thoughtfully constructed probability distribution is very likely to contain the future outcome. In some cases, the interviewer may need to disqualify the expert if the motivational issues appear insurmountable.

In the structuring step, the variable to be assessed must be precisely defined and in terms the expert is comfortable with. For example, if the parameter is the capital cost of utility-scale wind farms in 2020, then the precise units must be clarified (e.g., \$/kW or \$/kW-yr), what the cost includes (e.g., equipment only, grid interconnection, etc.), whether the cost is in real or nominal dollars, and so on. It is also in this step that any key conditioning assumptions should be drawn out by the interviewer. For example, is the expert's judgment about future wind farm capital costs conditioned on an assumption that the U.S. Government will be subsidizing advanced turbine research over the next X years? Is it based on assumptions about the progress of wind turbine development in China and Europe? When such assumptions are identified, it may be necessary to elicit multiple distributions, conditional on the outcomes from these underlying uncertainties, unless the expert is also knowledgeable about these influential events. Influence diagrams are often used to assist in the decomposition of conditional probability.

The preconditioning step is the part of the process that addresses cognitive biases, that is, unintentional heuristics that can result in an assessment that does not reflect all of the expert's information. Four types are common: availability, adjustment and anchoring, representativeness, and implicit conditioning. Availability bias occurs when judgment is influenced by how "available" information is to recall. Information is more available when it is dramatic, recent, redundant, certified, imaginable, and so on. For example, a non-aided assessment of the annual likelihood of a Katrina-magnitude hurricane made during the month after the Katrina hurricane would likely have been far different than an assessment made the month before. The interviewer must ask the expert for his/her sources of information and engage in a discussion of complementary events if availability bias is suspected.

The adjustment and anchoring bias occurs when the expert anchors on an initial estimate and adjusts inadequately for uncertainty. This bias is almost always present to some extent, as most discussions of uncertainty tend to revolve around the "most likely," or "base case" value. The most effective interview technique to counter this bias is for the interviewer to avoid mentioning any possible parameter values to the expert and to ask the expert to describe extreme outcomes. Essentially, this technique takes advantage of the availability bias to compensate for anchoring.

Representativeness bias results from the heuristic that says an event is likely to the extent that it is consistent with the evidence. It reflects a situation in which the expert is not considering the full range of causes for an event and judges its likelihood based on the narrow evidence at hand. This bias can be fairly complex to resolve, as it may require the expert to consider the implications of Bayes' theorem in his/her probability assessment.

The final bias addressed in the preconditioning step is implicit conditioning. This occurs when the expert assigns a high probability to an event but is unconsciously conditioning the event on a set of unstated assumptions that, in reality, would have a low likelihood of simultaneous occurrence and should therefore reduce the likelihood of the event being assessed. As with the other types of biases, the interviewer must carefully question the expert and attempt to draw out such unstated assumptions.

Once the preconditioning step is complete, the interviewer proceeds to the actual encoding. Typically, the process is to assess points on the cumulative probability distribution. In keeping with the desire to combat biases, the interviewer begins the encoding by asking the expert to consider extreme values for the parameter in question and to describe the scenarios that could lead to these extreme events. This dialogue allows the interviewer to define the low and high extremes of the distribution. The interviewer then proceeds to assess individual points on the distribution, keeping the evolving distribution out of sight of the expert to guard against efforts to ensure consistency.

Depending on the expert's comfort level with probability distributions, several techniques and tools have been developed to facilitate this process. For experts accustomed to working with probability distributions, a direct assessment of the distribution is usually preferred and can represent a significant time savings (see Morgan et al. 2001). For others, the fixed value method and the fixed probability method are the most common assessment approaches. In the fixed value method, the interviewer assesses a cumulative probability for a specific (fixed) value. The most commonly used tool for this approach is the probability wheel. Its purpose is to provide a visual representation of the relative likelihood of an event. The probability wheel is like a roulette wheel but with two complementary colors (usually blue and orange) instead of a circle of numbers. Complementary colors are used so that neither color is visually dominant. The relative sizes of the two colors are adjustable. The back of the wheel indicates the numerical probability corresponding to the relative size of the adjustable color. Questions are posed to the expert about the cumulative probability of a specific point on the cumulative distribution in the following form: 'What do you think is more likely: that the value of the parameter is less than or equal to X or that a random spinner would land in the blue section of this wheel?' By adjusting the size of the blue section appropriately and asking the question repeatedly, the interviewer finds the point at which the expert is indifferent between the two events. The relative size of the blue section (numerically interpreted on the back of the wheel) therefore corresponds to the cumulative probability of  $x \leq X$ . This process is repeated for different parameter values, with the interviewer being careful to vary the section of the distribution that is being assessed to prevent the expert from trying to control the shape of the distribution. Typically, 10 or so points are assessed in this manner. In practice, some experts have found the wheel unnecessary after a few points and preferred to provide probabilities directly (Morgan and Henrion 1990).

The fixed probability method involves assessing values corresponding to specific fractiles on the distribution. The interval technique is the most common form of the fixed probability method. Like the probability wheel, this technique also uses a spatial representation of uncertainty, but it is specifically targeted at producing the median and the quartiles of the distribution and is a helpful check on the result produced with the probability wheel. It involves drawing a horizontal line and bisecting it with a tick mark and asking the expert for what parameter value is it equally likely that the future value could be less than or greater than that value. This value is then noted as the 50th percentile value on the cumulative distribution and written below the tick mark for the expert to see. The 25th percentile is then assessed by asking the expert to consider solely the area to the left of the median tick mark and to provide the value in this range for which the future value is equally likely to be above or below it. A parallel approach is taken for the 75th percentile.

The final step in the elicitation is verification. This all-important step involves asking the expert to verify that he/she agrees with the implications of the distribution. The distribution is either sketched out by hand, or one of various software tools can be used to create a continuous distribution based on the elicitation results. Verification questions are then created based on the resulting distribution. For example, “your distribution says you believe that the likelihood that this parameter is less than or equal to X is 80%. Do you agree?” and so on. Another important verification addresses the implications of the extremes of the distribution: “Do you believe that the probability this parameter (exceeds Y) (is less than or equal to Z) is zero?” Sometimes the verification process will reveal inconsistencies (e.g., the assessed points cannot be resolved into a continuous curve) and/or unstated assumptions (e.g., a bimodal distribution) and requires a new assessment.

When multiple experts are involved, subjective probability assessment provides a means to measure the importance of differences in expert opinion. Experts may disagree on point estimates but be shown to be largely in agreement when their probability distributions are compared. Sharing the underlying information and assumptions revealed through the elicitation process may result in a consensus, but if disagreements remain, their importance can be tested through probabilistic sensitivity analysis. The Delphi method is a process to manage communication among multiple experts. It involves circulating conflicting opinions anonymously and then providing the experts the opportunity to revise their views and comment on those of others. Successive iterations of this technique have been shown to produce consensus (Linstone and Turoff 1975). Variations of the Delphi method have been used to resolve expert inconsistencies resulting from the subjective probability assessment process (Morgan and Henrion 1990; Keeney and von Winterfeldt 1991; Hora 1996).

As an elicitation method set is developed for the iRESM, a principal consideration will be the trade-off between the level of elicitation formality, the resources required to implement the process (time, cost, and number of participants), and the benefits in terms of information quality.

#### **2.2.2.6 Model Inter-Comparison Methods**

Model inter-comparison projects can be a valuable tool for assessing the level of agreement or disagreement among similar models. These projects typically run a group of models under the same or similar assumptions to produce a common set of comparable outputs. The comparison may be done either with or without Monte Carlo methods. Monte Carlo assessments rarely have been performed with large climate models (coupled atmosphere-ocean general circulation models or AOGCMs) because of their high dimensionality and long run times. Careful assessments of the results often show how different specific empirical formulation of equations or slightly different choices of phenomena or variables among the suite of models influence results and point out key sensitivities.

Climate model inter-comparisons such as the World Climate Research Program’s Coupled Model Intercomparison Project provide a standard experimental protocol for studying the output of AOGCMs and a systematic way of testing and evaluating models. The Coupled Model Intercomparison Project provides climate community-based infrastructure in support of climate model diagnosis, validation, inter-comparison, documentation and data access (WCRP 2011). Of particular relevance to the iRESM initiative are inter-comparison projects focused on dynamically downscaled or regional climate models such as the North American Regional Climate Change Assessment Program (see <http://www.narccap.ucar.edu/>) and the COordinated Regional climate Downscaling EXperiment (see <http://www.meteo.unican.es/en/projects/CORDEX>), “a [World Climate Research Program]-sponsored

program to organize an international coordinated framework to produce an improved generation of regional climate change projections world-wide for input into impact and adaptation studies” (CORDEX).

There is widespread recognition that different models emphasize different parts of the climate system, that all models are somewhat sensitive to their initial boundary conditions, that some processes are well understood but must be approximated because of scaling and computational issues, and that for other phenomena, lack of understanding is the constraint. Knutti (2008) describes the existing set of AOGCMs as a family of coexisting models that sample to some extent the uncertainty in describing the system. As Knutti et al. (2010) observe,

the “multimodel” approach provides a sensitivity test to the models’ structural choices. Additionally, an implicit assumption exists that multiple models provide additional and more reliable information than a single model, and higher confidence is placed on results that are common to an ensemble [of runs of different models], although in principle all models could suffer from similar deficiencies.

Ensembles are used at least partly because the mean of results from an ensemble of models better matches present-day observed values than those of individual models in the ensemble. There has been a substantial effort to understand both the match of individual model and the mean of the ensemble of forecasts (weighted or unweighted) to observed data. Weighting the results of individual models using Bayesian model averaging has been done for probabilistic weather forecasting models (Raferty et al. 2003) and for AOGCMs (Tebaldi et al. 2005). Less formally, AOGCM results have also been combined using Reliability Ensemble weighted averaging, with weights based on bias in historical forecasts (Giorgi and Mearns 2002).

To some extent, the spread of forecasts from the individual models and a standardized group of scenarios has been used as a proxy for overall uncertainty in climate variables for the future. However, some individual modeling groups have shown that the spread of forecasts using a set of standard scenarios in a Monte Carlo experiment with a single model may have a different mean and a wider variance than the ensemble of non-Monte Carlo runs from multiple models using the same standard set of scenarios, thus calling both the mean and variance of these ensembles into question (Webster et al. 2009).

Including others, Knutti et al. (2010) caution that matching model results to present-day climate is only weakly connected to the ability to forecast well. Delsole and Shukla (2010) define “fidelity” as a measure of how well the climatology of forecasts replicate the observed climatology (a measure of the difference between two distributions over space—spatially averaged relative entropy) and “skill” as a test of the match to a forecast (a spatially averaged mutual information between the forecast and corresponding verification—e.g., from a short out-of-sample forecast). Because discrepancies in climatological means are typically much greater, models that more closely replicate the observed climatological mean tend to have better skill in forecasting.

Model inter-comparability studies of integrated assessment models have been done regularly by the Integrated Assessment Modeling Consortium and Energy Modeling Forum (EMF). The Integrated Assessment Modeling Consortium is focused on developing a new set of standard assumptions and scenarios to guide model analyses for both integrated assessment and climate modeling communities (representative concentration pathways or RCPs; see Moss et al. 2010). The consortium is also concerned with socioeconomic scenarios that are compatible with the RCPs and usable by the integrated modeling

and vulnerability, impact, and adaptation communities (IAMC 2011). EMF has conducted a series of studies, each with an ad hoc working group of 50-100 members, organized to examine a single topic to which many existing models can be applied. Topics besides integrated assessments related to climate change are studied; however, four of the five most recent studies (EMF 21, 22, 24, and 25) all addressed climate-related topics (EMF 2010). These studies include:

- EMF-21, Multigas Mitigation and Climate Change
- EMF-22, Climate Change Control Scenarios
- EMF-24, Technology Strategies for Achieving Climate Policy Objectives
- EMF-25, Energy Efficiency and Climate Change Mitigation.

Most commonly, these modeling exercises do not require Monte Carlo projections. The results, however, provide insights into the importance of certain model completeness, sensitivity, and structural differences, as well as likely difficulties in following certain policies or emission paths that they model. EMF 14 did examine approaches for performing uncertainty analysis in large scale policy/economic models and conducted a model comparison with three uncertain parameters (climate sensitivity, warm climate damages, and discount rate), but this appears to have been done using a few model runs rather than a Monte Carlo analysis (Manne 1996).

Another example of intermodel comparison with a sectoral and impact emphasis is the Agriculture Model Intercomparison and Improvement Project, “a distributed climate-scenario simulation exercise for historical model inter-comparison and future climate change conditions with participation of multiple crop and world agricultural trade modeling groups around the world” (Rosenzweig 2010). Both historical testing against observations and assessment of future mitigation and adaptation strategies will be tested.

## **2.3 UP Methods**

Means of UP can be either quantitative or qualitative in nature. In quantitative analysis, the general means of propagation is dictated by the algebra of the measures (such as probabilities) used to characterize uncertainty (already discussed in Section 2.2.2). It is the nature and complexity of the underlying models, however, that dictate the practical and specific means of quantitative UP. For instance, analytic propagation of uncertainty is generally impractical and Monte Carlo methods must be used.

### **2.3.1 Quantitative UP**

Quantitative methods have conventionally included Bayesian characterization of the uncertainties of inputs and key parameter values in the linked models used in the analysis. Uncertainty is propagated by Monte Carlo methods involving multiple implementations of the models using input values sampled from subjective probability distributions. These distributions are derived from scientific literature and directly elicited from domain experts. These distributions may be sampled independently, but it is generally considered better practice to account for known dependencies among input variables through sampling correlations.

Generally, a Monte Carlo approach that involves random sampling of input distributions can be problematic when the run time of the underlying model is substantial. In these cases, multiple implementations of the model may be impractical, particularly if the sample size required to achieve unbiased output distribution estimates is in the thousands, or tens of thousands. Therefore, much attention has been focused across numerous domains on the development of statistical methods that reduce the required sample size for a given model. Latin Hypercube Sampling (LHS; Helton and Davis 2003) is one of the more established methods for improved Monte Carlo analysis. It is a so-called *stratified* sampling approach that ensures the full breadths of the input distributions are sampled and that the output distributions are sufficiently unbiased for a sample size much smaller than would be required for the same output properties using conventional random sampling. LHS-based methodologies also exist to allow required correlations to be imposed on the input marginal distributions (Iman and Davenport 1982).

Importance sampling (Swiler and West 2010) is another alternative to random sampling in which the more *influential* portions of the input parameter space can be preferentially sampled to produce a more accurate UC for the output ranges of interest. For instance, the analyst's interest may be focused on extreme output temperature predictions, in which case importance sampling provides a means of focusing on the regions of the input space that most influence these extreme predictions. Again, this approach allows an economy of sample size, in this case by focusing on the parameter space region of most interest.

Another class of sampling methodologies is Markov Chain Monte Carlo (Gelfand and Smith 1990). Unlike the methods previously identified, a given sample member in Markov Chain Monte Carlo is randomly selected based on the value of the previous sample member. That is, the entire input sample is *not drawn a priori*, but is generated sequentially throughout the analysis process with the objective of reaching an equilibrium distribution that closely estimates the actual output distribution. The sampling process is continued until the stable output distribution is produced. Tao et al. (2009) have applied this sampling methodology in the context of climate and crop productivity.

An alternative to sampling economy as a means of addressing long model run times is to create surrogate models that emulate the behaviors of the detailed models while having substantially shorter run times. There are numerous approaches available for the development of surrogate models, the most conventional being the development of so-called *response surfaces*, which are generally constructed from statistical regression fits between model inputs and outputs based on a limited number of model runs (Iman and Conover 1980). These response surfaces, which have minimal run times, are then used as the basis for conducting the full uncertainty analysis in lieu of the original model. Note, however, that, depending on the goodness of fit of the surrogate, uncertainty can be introduced around the accuracy with which the source model is being emulated.

In recent years, improved bases for creating model surrogates for the purposes of Monte Carlo sampling have been developed. These include so-called *stochastic expansion methods*, such as Polynomial Chaos Expansion and Stochastic Collocation (Eldred et al. 2008). In these approaches, the surrogate model is expressed as an expansion in orthogonal polynomials of which the arguments are random variables. The coefficients in the expansion can be determined by limited sampling of the underlying science model. The advantage of these stochastic expansion methods is that they demand relatively small sample sizes as the basis for UP. As these new methodologies are explored, some disadvantages are also being addressing associated, for example, with the scalability of the surrogate models.

Among analysis packages that allow implementation of a broad range of sampling and surrogate model methods is the Design Analysis Kit for Optimization and Terascale Applications tool developed by Sandia National Laboratories (Adams et al. 2010).

The most commonly used sampling methodology in the climate community has been LHS, which might typically achieve about the same accuracy as random sampling for an LHS sample size of about 10% or less of the required random sample size. Advantages of LHS include its broad validation through extensive use, conceptual simplicity; the availability of associated global sensitivity analysis procedures (Iman and Conover 1980); and effectiveness as a model verification tool, especially as compared with decomposition of variance and other techniques (Helton and Davis 2003).

It is worth mentioning another class of methodologies that is oriented largely to model sensitivity analysis, but potentially usable in the context of UP. These involve so-called *intrusive* methods that demand modifications to the underlying science models (as opposed to Monte Carlo methods that leave the science models intact but require multiple implementations). One of the more established of such approaches is the Adjoint method (Cacuci and Schlesinger 1994). This involves the incorporation of *adjoint* equations into the underlying model that reflect relationships between the derivatives of the basic model parameters. The solutions to the adjoint equations then measure the model input/output sensitivities. These sensitivity measures can be used to form input/output Taylor expansions that are the basis for propagating distribution moments through the model. While such methods are well established for linear models, the basis for their application to nonlinear systems is more formative.

There is a growing body of literature on multi-model propagation of uncertainty in integrated assessment models that attempts to estimate the impacts of future climate, energy technology, and policy on certain human and natural systems. Recent efforts include Tomassini et al. (2010) for global climate, Groves et al. (2008) for water resources, Tao et al. (2009) for crop productivity, Heinrich et al. (2007) for electric system expansion, and Webster et al. (2008, 2009) and Sokolov et al. (2009) for impacts and costs of climate policy. Efforts to date have tended to be international or national in scope and have paid relatively little attention to regional impacts and regional energy, environmental, and climate policy.

For example, Tomassini et al. (2010) investigate the sensitivity of the probability of exceeding 2°C by the year 2100. They do this with IMAGE model version 2.3, an integrated assessment model, using five climate stabilization scenarios that span a variety of climate forcings (without probabilities attached). IMAGE output was used to feed the Bern2.5D climate model of “intermediate complexity” to perform probabilistic climate projections. The analysis considered only the uncertainty in climate model parameters but not uncertainty in the economic development (for a fixed scenario) or short-term climate uncertainty in emissions. Massachusetts Institute of Technology’s (MIT’s) Integrated Global System Model, which couples an emissions model with a climate model of intermediate complexity, has been used in a similar fashion with Monte Carlo methods. It has been used to investigate the uncertainty in global emissions and climate change (e.g., air temperature and sea level rise) in the absence of climate policy (Sokolov et al. 2009) and with the emissions part of the model turned off, the consequences for climate of specific standardized mitigation scenarios (Webster et al. 2009).

Among the specific sector impact analyses, Groves et al. (2008) presented a method for developing large ensembles of local daily weather that reflect a wide range of plausible future climate change scenarios while preserving many statistical properties of local historical weather patterns. The technique was applied to water utility service area in southern California, and water resources consequences of

climate change were evaluated. In the Washington State Climatic Change Impacts Assessment (CIG 2009; for a summary, see Miles et al. 2010), scenarios with statistically downscaled climate change from multiple AOGCMs (temperature and precipitation) were developed at the 1/8° level and run through a hydrology model and reservoir management model to determine water availability for irrigation. Water availability, precipitation and temperature forecasts were then used to generate yields from irrigated crops and these results are used in turn used to calculate farm profitability. All of the steps in the analysis used distributions of outputs from the upstream steps. Tao et al. (2009) ran a process-based general crop model with 10 climate scenarios (combinations of five AOGCMs and two emissions scenarios) and 60 sets of crop model parameters to obtain probabilistic projections of maize (corn) productivity on a grid of a large growing area in China. The effort was described as a super-ensemble-based probabilistic projection Bayesian probability inversion and a Markov Chain Monte Carlo technique was applied to the crop model. For evaluation, 55,000 sets of parameter values were compared with flowering dates, maturity dates, and yields; for forecasts, 18,000 sets of simulation results were derived for the 2020s, 2050s, and 2080s. Statistics of outputs and cumulative density functions were derived.

Heinrich et al. (2007) make use of Multi-Attribute Value Theory (MAVT) (a type of multi-criteria decision analysis) and Monte Carlo techniques to derive preferred expansion paths for electric energy generation in South Africa, giving explicit consideration to global impacts such as climate change and regional impacts such as local air quality (due to South Africa's high coal plant density region [Mpumalanga]) and water consumption (due to national water shortages). The analysis emphasized "robust" solutions (i.e., a solution will perform well under numerous future values of parameters) and "flexible" solutions (i.e., a solution can be adapted at a future point in time). An uncertainty distribution was defined for each technology, for each of the uncertainty parameters (investment cost, operations and maintenance cost, fuel cost, and emission factors). Performance was specified against alternative uncertain scenarios using cross-criterion indifference judgments to estimate partial value functions. This analysis only traded off cost against CO<sub>2</sub> emissions; however, it is possible to further rate the technologies in terms of water use or efficiency for different climate outcomes and develop an integrated solution. In either case, propagation of uncertainty from climate policy and climate constraints is straightforward.

### 2.3.2 Qualitative Methods

Qualitative means of UP also exist. In Section 2.2.1.1 the approach of simple bounding analysis was outlined, involving the intuitive covariation of model input parameters to assess the range of output values. The inherent weaknesses and problems of interpretation associated with this approach were also identified. Methods sometimes described as semi-quantitative or semi-qualitative also exist. In such approaches, uncertainty scales such as a 1 to 10 ranking, or *high*, *medium*, *low* are established and attached to elemental uncertainties of the analysis, and those uncertainties are then propagated or combined by means of a pre-established heuristic. For instance, an output parameter that results from the input of two input parameters, each assigned *medium* uncertainty, might be assigned a *high* uncertainty. The underlying models to which such uncertainty methods are applied are generally sufficiently simple that a UP heuristic can be intuited (see, for example, CCPS 2008).

## 3.0 Decision Making Under Uncertainty

Decision science is a mature and multi-faceted discipline that might even be argued to have been an antecedent of the mathematics of uncertainty. It has evolved as a richly structured set of methodologies that broach a wide variety of decision problems. The stable of methods address, to pick a few examples, the basis for meeting pre-established acceptance criteria (such as risk goals), the ranking of decision options by risk or uncertainty criteria, cost-benefit analysis, and comparison of decision options across disparate, multi-attribute criteria. Leveraging the accelerating capability in computational and communications technologies, decision science has more recently subsumed the challenge of creating information-rich environments that support shared and communal decision making, relying on advancements in data analytics, serious gaming, and data visualization (Sanfilippo et al. 2009). However, our focus here is not on the specific methods and technologies of decision science but rather on consideration of broad decision strategies—particularly those currently being considered as means of addressing uncertainties in climate policy—to make sure that the models, analytical capabilities, and UC methods developed by the iRESM team will meet the needs of policy-makers.

### 3.1 Methods

The issue of the subjectivity and defensibility of quantitative uncertainty measures—a question that extends beyond probabilistic approaches—has led the community of risk and decision analysts to consider decision methods that enjoy some measure of robustness to the specifics of UQ. That is, rather than focus exclusively on the means of enhancing the defensibility of a particular UQ, attention is shifted to devising decision strategies that are less sensitive to the choice in UQ. Thus, notions of *robust* and *adaptive* decision-making have been expounded, particularly in the climate arena, to manage risk while avoiding the placement of unjustified confidence in any one set of quantitative uncertainty characteristics.

Against this backdrop of evolving UQ methods and decision frameworks, the diverse set of policy questions, science uncertainties, and modeling issues associated with climate research provide a challenging test environment for methodological advances.

There have been several reviews of prospective decision strategies in the arena of climate mitigation and adaptation (Toth et al. 2001; Bell et al. 2003; Morgan et al. 2009; Means et al. 2010). We will not reproduce the insights and conclusions of those reviews here but, rather, will identify some of the key decision strategies that have been identified and then briefly discuss what constraints support of those strategies might place on the UC methods developed for iRESM.

#### 3.1.1 Classical Decision Methods

In this family of methods, the expected utilities (probabilistically weighted figures of merit) associated with the outcomes of each decision option are compared, and the option with the greatest expected utility is selected (see, for example, Aleskerov et al. 2007). Since the identities of the preferred decisions are sensitive to the input probability distributions (which can be contentious—see Section 2.2), there is often some reluctance in real-world settings to adopt such methods wholesale. Against that background, alternative strategies have been proposed (see Section 3.1.2 through Section 3.1.5).

Since these classical decision methods are founded on probabilistic principles, the prospect of using iRESM results to support such decision strategies favors the use of probabilistic UC methods—at least to the extent demanded by specific decision problems. However, the generalization of utility optimization models to non-probabilistic UC environments is an area that continues to receive attention (Dubois and Prade 1990; Caselton and Luo 1992), and anticipation of some level of adoption of classical decision concepts would not likely constrain the iRESM UC methodology to a purely probabilistic one.

### **3.1.2 Robust Strategies**

The intent underlying use of this family of decision methods is to produce preferred decisions that are less sensitive to the selected input probability distributions than in the case of classical decision methodology (Lempert et al. 1996; Regan et al. 2005; Lempert and Collins 2007). Robust strategies capture a range of methods. In the conservative extreme, the decision strategy can be completely decoupled from the choice of input probabilities, in which the preferred decision option is the one that results in the least severe worst-case outcome (this being the worst case given the range of outcome uncertainty associated with the option).

Intermediate approaches that draw a balance between this extreme and classical decision methods include one in which the classically optimal decision first is identified based on best-estimate probabilities. Next, it is determined how the expected outcomes of that decision would change if the input probabilities were adjusted in an adverse way. Alternative decision options are then identified that reduce the expected consequences associated with the more adverse set of input probabilities. In effect, this approach sacrifices maximized outcome utilities (with respect to best estimate probabilities) to hedge against the adverse consequences associated with more pessimistic input probabilities. Other robust methodologies, with the common feature of reduced sensitivity to the UC, have also been proposed.

While robust decision strategies have largely been defined in terms of probabilistic UC methodologies (reflecting the prevalence of probabilistic UC), the intent to support robust decision methods does not confine iRESM to any one means of UC. Indeed, the motivation for nonprobabilistic uncertainty methods has been the enhanced robustness of the UC (see Section 2.2), and thus of the decisions to which it is applied.

### **3.1.3 Adaptive Strategies**

Adaptive approaches to decision making can be viewed as a subset of the robust methodologies in which the basis for hedging against potentially contentious input probabilities is to distribute decisions along the timeline; that is, to base current decisions on the assumption that future corrective actions can be taken as conditions unfold and uncertainties narrow. Such approaches balance the risk of irrevocably adverse impacts associated with a delay in critical decisions against the advantage of postponing some decisions until the uncertainties have narrowed. Lempert et al. (1996) demonstrate that for a specific climate policy problem set, the performance of adaptive strategies is superior to that of a classical non-adaptive approach. Practically, we would expect any real-world decision policy, or at least decision practice, to be based partially on adaptive strategy. While the performance of adaptive decision making has been demonstrated in a probabilistic setting (Lempert et al. 1996), no particular UC methodology is inherent in such strategies, and thus the prospect of supporting adaptive decision making does not significantly constrain the iRESM choice in UC methods.

### **3.1.4 Non-Algorithmic Decisions**

Notwithstanding the abundance of literature and the sound theoretical bases for formal decision methodologies, the reality is that public and private decisions—by national policymakers, regulators, corporate officers, military commanders, or other decision makers—are seldom, if ever, based exclusively on decision algorithms. At best, the products of formal decision methods constitute one set of data points among those considered in the deliberations of a decision maker. Some of the reasons for reluctance to incorporate decision-theoretic methods wholesale into decision making—such as limited confidence in the completeness of the UC or perceived arbitrariness in the assignment of probabilities to future conditions—have already been discussed in Section 2.1 and Section 2.2. Another reason is that the decision makers are often ill at ease with the complexity of decision models and of the science models that support them, and this impedes their confident use in decision making. While such impediments can to some extent be overcome through effective stakeholder communication, it is nevertheless unlikely that decision algorithms will ever provide a comprehensive basis for decision making.

What is important, therefore, is that the results and insights of a scientific analysis, independently of the decision methods that might be proposed, be coherent, transparent (to the degree practical), and conveyable to the decision makers. That is, effective decision support by iRESM requires the use of a quantitative framework in which predictions, consequences, and associated uncertainties can be systemically and consistently cast in a comprehensible, defensible, and unambiguous way. These are among the criteria against which alternative UC methods were assessed in Section 2.2.

### **3.1.5 Conclusions on Decision Methods**

Section 3.1 briefly outlines the major categories of decision strategies that have been proposed and discussed in the climate mitigation and adaptation literature. While the bases for the preference of one class of decision model over another have been expounded by the proponents, formal decision models are seldom, if ever, the sole or principal basis for public decision making. Regardless of decision methodology, it is crucial that the approach to UC adopted by the iRESM team creates a decision environment that is information-rich, comprehensible, defensible, and, to the extent practical, complete. The considerations in Sections 2.2 and 2.3 can help make sure the iRESM approach to UC meets these criteria.

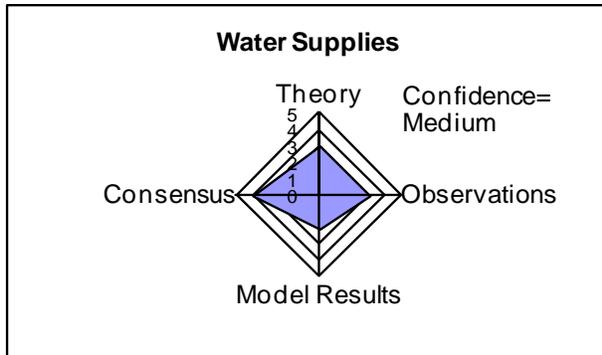
## **3.2 Interpretation of Results/Visualization**

If uncertainties in model results are used to inform policy, it is important that users and decision makers be provided a useful view of the uncertainties in model inputs and outputs and a sense of how decisions are likely to be affected by these uncertainties. This section provides a brief discussion of a few visual tools that have proven useful in depicting uncertainties in model inputs and outputs.

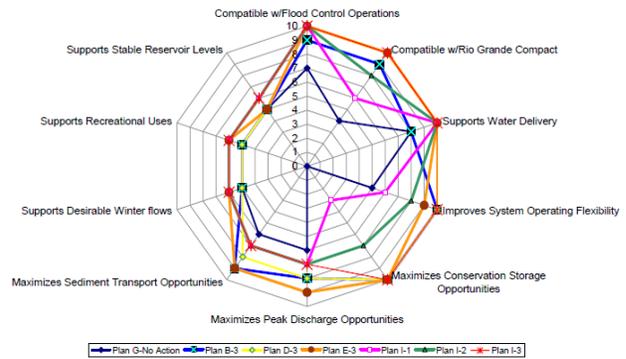
### **3.2.1 Sensitivity: Radar Graphs and Tornado Diagrams**

Radar graphs can be used to show how changes in a wide variety of variables or conditions contribute to a single outcome. Figure 3.1a is from the human settlements chapter of the IPCC third assessment report (McCarthy et al. 2001). It uses a subjective 5-point scale to estimate and combine four dimensions of expert confidence concerning the projected level of climate change impact on future water supplies.

The (nominally independent) rating dimensions were: 1) the degree of consensus among experts, 2) the strength of the underlying theory, 3) quality of model results predicting water shortages, and 4) the consistency of predicted impacts with observations. The overall impact confidence level was based on the area of the shaded polygon, computed from the individual confidence levels. Figure 3.1b is a different type of radar graph showing the impacts of a number of policy approaches on several competing policy objectives. This is a “likeliest impact” chart and could be repeated for “high” (0.95 probability) and “low” (0.05 probability) impacts on any of the dimensions.



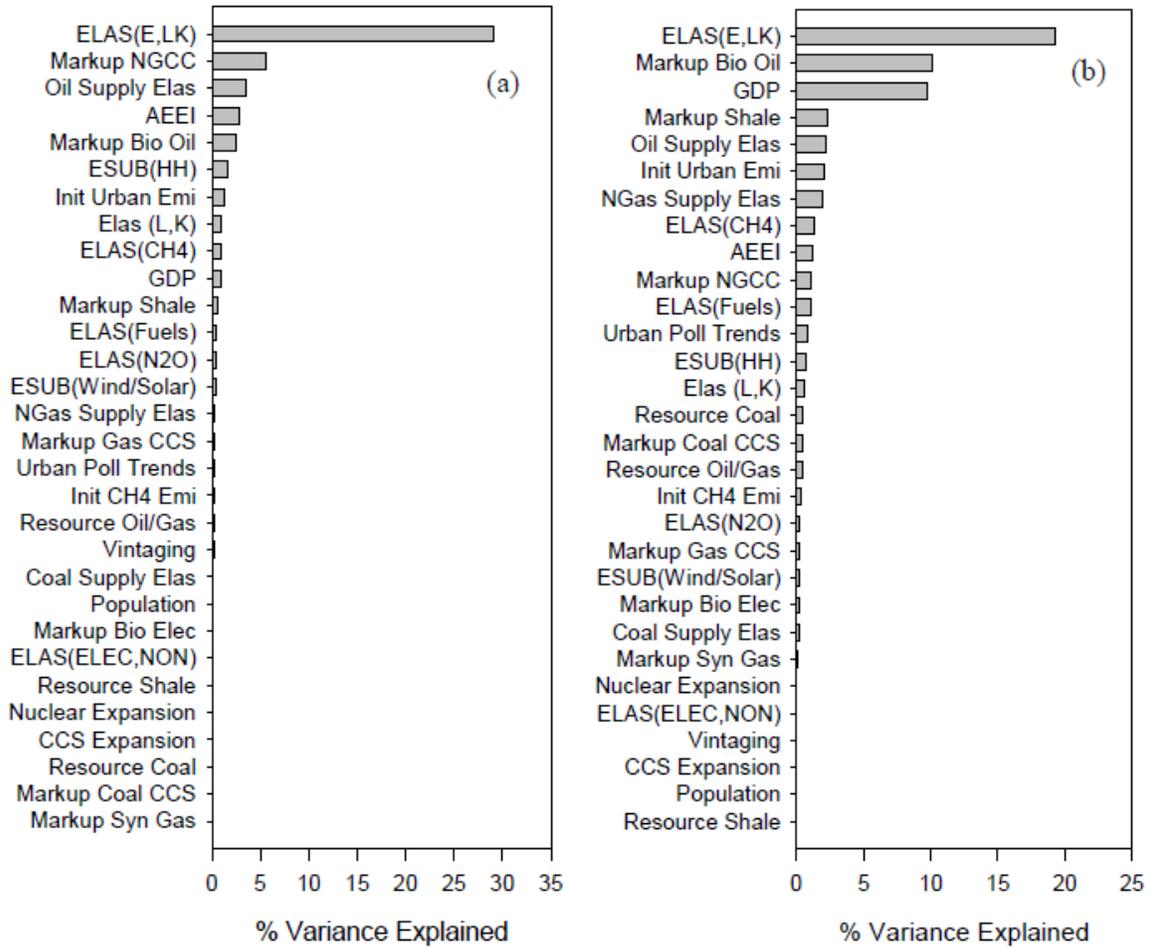
a. Confidence Ratings for Impacts of Climate Change on Human Settlements: Water Supplies. Source: McCarthy et al. 2001.



b. Multi-dimensional rating of alternative approaches to managing a riverine habitat. Source: USCOE/USBR/NMISC 2007.

**Figure 3.1.** Examples of Radar Graphs that Depict Many Variables’ Contribution to a Single Outcome and Alternative Strategies to Fulfilling Competing Objectives

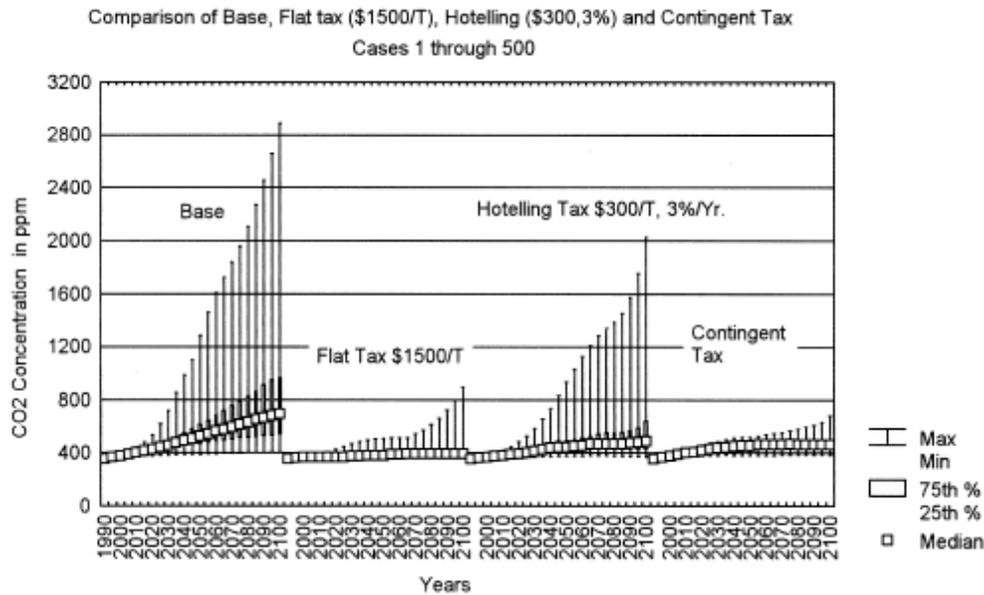
Tornado graphs are another means to show the relative contribution of uncertainty to overall uncertainty of a model outcome. Figure 3.2 (from Webster et al. 2008) shows the decomposition of variance in the MIT Emissions Prediction and Policy Analysis model and shows how variation in several groups of uncertain parameters explain the variance in the cost-controlling CO<sub>2</sub> emissions (carbon price) for two different forecast periods. Note that a handful of variables (<5) explain the overwhelming percentage of overall variance, but the ranking of the variables changes over time. Analyses of this type depict sensitivity of the model to various underlying assumptions and identify “what matters” for policy analysis. Sensitivity analyses can be done as in Figure 3.2 using relative variance or by using a common fixed percentage change across variables to isolate the marginal effect of changes in single variables as was done by Scott et al. (1999).



**Figure 3.2.** Percentage of Variance in Carbon Prices in (a) 2020, (b) 2060 Explained by Uncertain Parameter Under a 550 ppm Stabilization Case. The percentage variance is calculated as the ratio of the partial sum of squared errors to the total sum of squared errors from the results of an analysis of variance. Source: Webster et al. (2008).

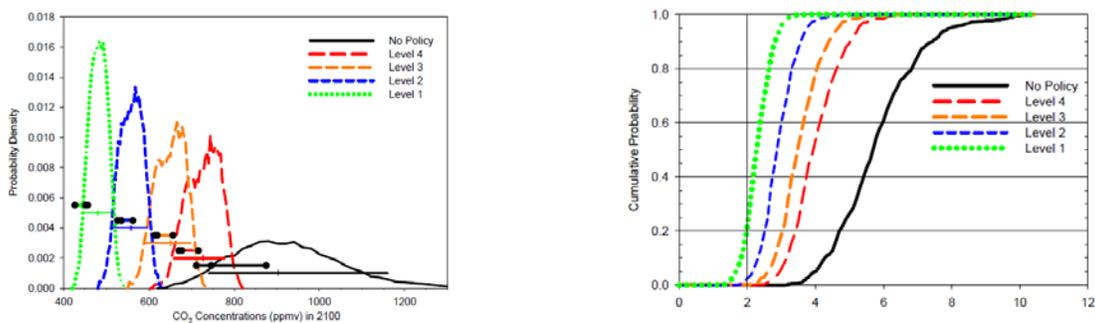
### 3.2.2 Uncertainty Box and Whisker Plots, Scatter Plots

Box and whisker plots are economical ways of summarizing and depicting results of uncertainty analyses. As illustrated in Figure 3.3, it is possible to efficiently show several elements of uncertainty at once. From an uncertainty analysis using the MiniCAM model (Scott et al. 1999), this figure simultaneously compares the impact of model uncertainty over time on atmospheric CO<sub>2</sub> for the IS92 base case and three policy scenarios. The figure demonstrates that simultaneous very low-probability “tail values” of certain input variables can result in very high CO<sub>2</sub> concentrations and, that while all three of the proposed control policies show similar median concentration pathways, two out of the three do a much better job of preventing worst-case scenarios.



**Figure 3.3.** Comparison of Flat Carbon Tax, Hotelling Carbon Tax, and Contingent Carbon Tax in Controlling the CO<sub>2</sub> Concentration in the Atmosphere Relative to a Business-As-Usual Case. Source: Scott et al. (1999); copyright permission pending.

Where there is a lot of information available concerning the specific shape of distributions of values, it can be useful to show the density function directly or the cumulative density function derived from it. The example from Figure 3.4 is a set of density distributions of CO<sub>2</sub> concentrations and the corresponding cumulative probability density function for temperature under a base case and four carbon policy scenarios from the MIT Integrated Global Systems Model (IGSM) (Webster et al. 2009). Moreover, where these outcomes vary by location within the iRESM region, these figures can be shown as a family of figures (e.g., 14 of them for individual states in the iRESM region).



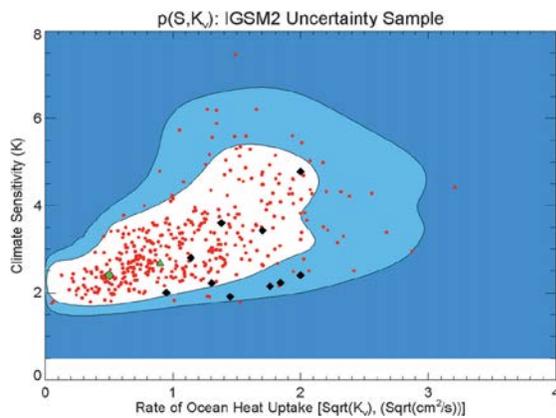
a. Frequency distributions, and medians and 95% bounds, of CO<sub>2</sub> concentrations averaged for the decade 2091-2100. Horizontal lines with three circles indicate range of reported results from Clarke et al. (2007), and circles indicate the point estimates from the three models.

b. Cumulative probability distributions of global mean temperature change from decadal average for 1861-1870 (preindustrial) to the decadal average for 2091-2100. Source: Webster et al. (2009)

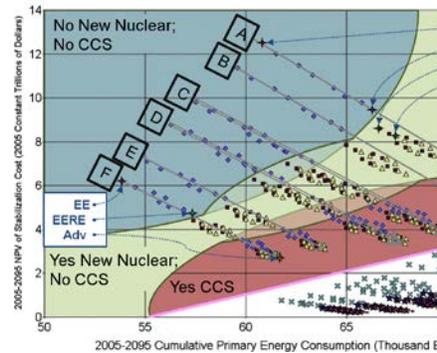
Source: Webster et al. (2009).

**Figure 3.4.** Uncertainties in Model-Estimated CO<sub>2</sub> Concentrations for a Baseline and Four Policy Cases Computed by the MIT IGSM Model

Scatter plots also are sometimes valuable in depicting the distribution of outputs or inputs and can be structured to simultaneously depict several elements or dimensions of information concerning output. For example, Figure 3.5a shows how inputs for a sample of 400 realizations of climate sensitivity and ocean heat uptake are associated with each other in the IGSM integrated assessment model; it also shows their range and the implicit values required to match outputs of global general circulation models. Figure 3.5b shows how control costs and energy consumption vary across portfolios of carbon control technologies and control levels in Pacific Northwest National Laboratory's Global Change Assessment Model. The methods illustrated in Figure 3.5a could be applied to show, for example, the impacts of different growth rates in regional population or trade coefficients for the iRESM region for total emissions of CO<sub>2</sub>, while the method in Figure 3.5b could be used to evaluate the impacts of regional technology combinations on CO<sub>2</sub> emissions under various policy assumptions.



a. The marginal posterior probability density function for climate sensitivity (S) and ocean heat uptake (Ky). The shading and thick contours denote rejection regions for significance levels of 10% and 1%, respectively. Green circles and triangles indicate mode and a median on the distribution, respectively. Black diamonds indicate values of the parameters of the MIT climate model needed to represent behavior of different AOGCMs in the simulations with 1% yr<sup>-1</sup> increases in the CO<sub>2</sub> concentration. Red dots show values for Ky and S from 400 samples. Source: Sokolov et al. 2009.



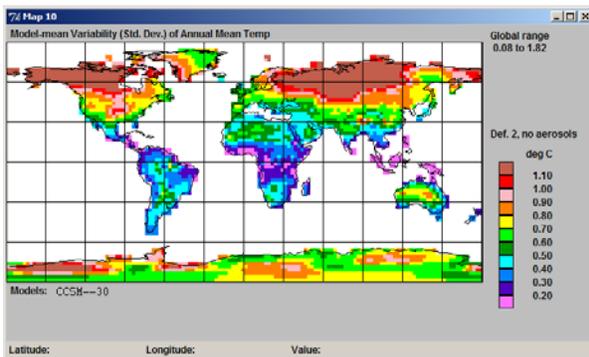
b. Cost and energy consumption results for 768 combinations of control technologies. Diagonal clusters A-F are formed by common end-use technology level combinations, combinations are noted with and without nuclear and with and without carbon capture and storage for control goals of 450 and 550ppm. Source: McJeon et al. 2010.

**Figure 3.5.** Graphically Depicting Multiple Elements of Uncertainty

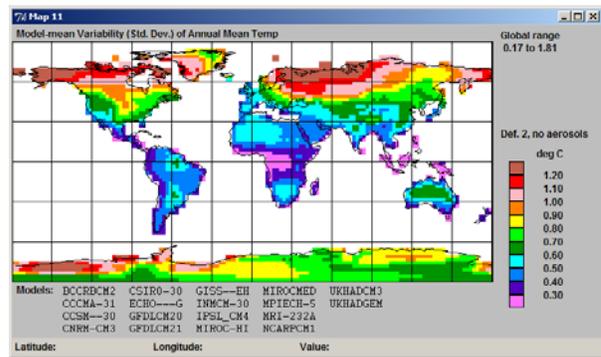
### 3.2.3 Spatial Variation

Spatial variation in output from forecasts adds an additional element of uncertainty. To pick one well known example, climate modelers know that different plausible models of atmospheric processes and/or geographic downscaling schemes can produce different changes in mean temperature and other climate variables at locations of interest on the ground. However, uncertainty within and between models concerning location of climate impacts could influence regional decision making. Figure 3.6 shows how uncertainty of regional allocation of climate impacts can be portrayed visually on a gridded map landscape. (In iRESM, the grid would be much finer [10-25 km] but over a much more limited geographic area.) The techniques involved can be applied in principle to other sub-regional geographic

entities, such as counties, states, or utility service nodes. Figure 3.6 is a figure generated by MAGICC/SCENGEN 5.3 (Wigley 2008), which allocates policy-driven changes in forecasted global temperature, precipitation, and other variables to a global  $2.5^\circ \times 2.5^\circ$  latitude-longitude grid according to the relative allocations implicit in 24 different coupled AOGCMs. The variation in model outcomes can be depicted by use of a color palette as simply the standard deviation in forecasted values for a large number of realizations for a single AOGCM (Figure 3.6a) or as the standard deviation of realizations for an average of allocation schemes among several models (Figure 3.6b). While MAGICC/SCENGEN itself depends on model-based geographic downscaling schemes (AOGCMs in particular), virtually any other set of spatial comparisons for any kind of geographically allocated variable (for example, the difference in standard deviations of total water consumption at each county) can be depicted in this manner.



a. Example of percent inter-annual standard deviations in geographic temperature change within a specific allocation scheme at the year 2050 (CCSM30 Model,  $1.64^\circ\text{C}$  global impact)  
Source: Wigley (2008)



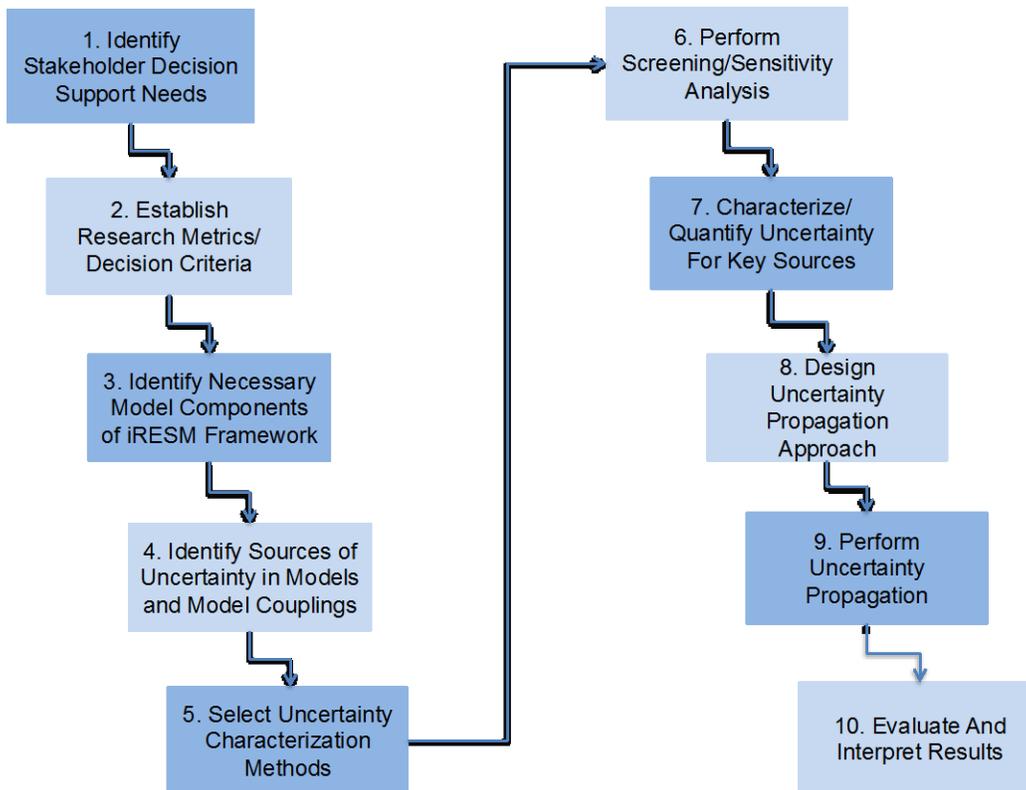
b. Example of inter-annual standard deviations in geographic temperature impact within an average of 18 specific climate allocation schemes.  
Source: Wigley (2008)

**Figure 3.6.** Variability in Geographic Distribution at the Year 2050 of Standard Deviations in Temperature Changes from AITMES Scenario Within the MAGICC/SCENGEN 5.3 Model

## 4.0 Recommended UC Process for iRESM

In the previous sections, technical approaches to UC have been outlined along with strengths and shortcomings considered relevant to iRESM. While the methodologies identified are not exhaustive, and there are variations of approach within methods sets, we have captured the principal methodological groups that have found application in practical problem domains. Ultimately, however, the merit of an approach to UC must be measured by the technical validity and level of stakeholder comfort associated with incorporation of that UC into a decision process.

The UC process developed for iRESM is predicated on the importance of helping stakeholders (i.e., researchers and decision makers) understand the robustness of results provided by iRESM, that is, which uncertainties have the biggest influence on key decision criteria or other research metrics. This means that the stakeholder context necessarily provides the organizing principles for UC. The UC process has the following ten steps represented in Figure 4.1 below.



**Figure 4.1.** UC Process for iRESM

The implications of this process represent a profound change from the typical UQ approach that focuses on characterizing all the uncertainty in all parameters of a single model. Not only is such an approach very likely to be computationally infeasible for an integrated model suite with significant run times, but the level of effort required to produce defensible UCs on every parameter is also very likely to be beyond any reasonable project scope. The remainder of this section describes each of these steps in detail. The first four steps of the process were exercised in FY 2011 in a limited application to develop the process, learn how to interact with policy stakeholders and modeling teams, and determine how to

approach the full design for UC going forward. A significant *inter alia* result of this activity is that the UC process with some slight modifications and additions has strongly influenced the overall design for the entire iRESM project, thereby demonstrating that the needs of the process for adequately characterizing uncertainty for robust decision making are very nearly identical to the needs for good model development and good analysis.

## 4.1 Identify Stakeholder Decision Support Needs

The first step in this process is to identify one or more decisions or research questions to be addressed with the framework. The research into stakeholder decision support needs in the 14-state Midwest pilot region has been initiated and indicates the following decisions and issues are of first importance (Rice et al. 2010):

- Mitigation-related:
  - The impacts of renewable portfolio standards (RPS) on regional electricity costs, land use, water use, and transmission needs
  - The consistency of regional RPS policy requirements across the region
  - The sustainability, cost, and performance of renewables and sequestration opportunities
- Adaptation-related:
  - Climate change impacts on water availability, temperature, and precipitation (changes in means and extremes)
  - Ecosystem and human health impacts due to these changes
- General:
  - Land and water use conflicts arising from mitigation and adaptation decisions
  - Energy-water nexus (e.g., the impact of increased renewables market penetration on water conflicts, and the interactions between water use and energy use).

This list will be updated as additional input is received from the continuing stakeholder interaction process that is part of the iRESM study, but it provides a useful starting point to begin characterizing relevant decision contexts in the pilot region.

The next step in the process is to consider this region-specific list in light of the broader research questions defined for the iRESM framework (see Section 1). The research team produced the synthesis shown in Table 4.1 indicating key intersections between Midwest stakeholder decision support needs and iRESM's research questions.

**Table 4.1.** Intersection of Key Midwest Issues and iRESM Research Questions

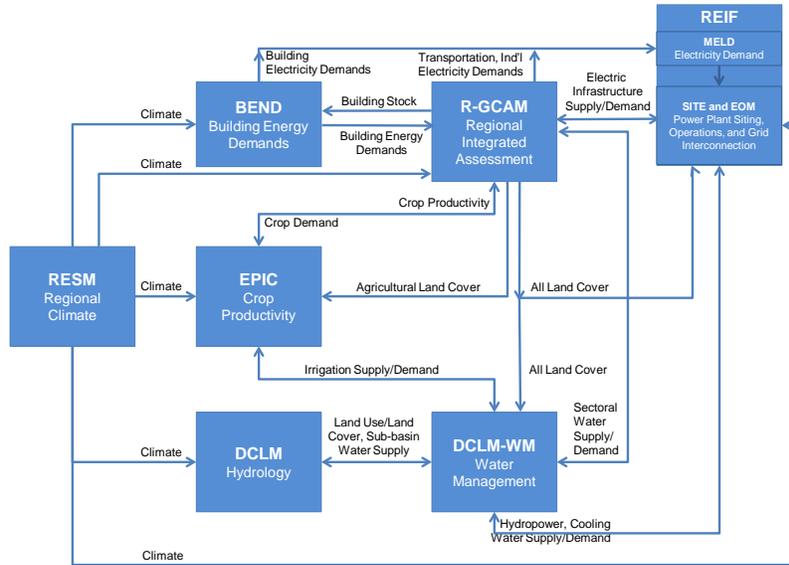
Stakeholder Issues	iRESM Science Questions			
	Regional Opportunities for Constraints on Mitigation and Adaptation	Impact of Uncertainty in Mean vs. Extremes of Climate Change on Mitigation and Adaptation	Mitigation and Adaptation Causing Non-Linear Changes and Climate Feedbacks	Intersections Between Mitigation and Adaptation Affecting Their Outcomes
Regional RPS Policy Requirements	X	X		
Climate Change Impacts on Water, Temperature and Precipitation		X		X
Sustainability, Cost of Renewables, Sequestration	X	X	X	X
Land and Water Use Conflicts	X	X	X	X
Energy-Water Nexus	X	X		X
Ecosystem and Human Health Impacts		X		

Based on an examination of this table, the iRESM team decided that the initial Midwest region pilot study should explore the sustainability, cost, and performance of renewables and sequestration alternatives necessary to meet RPS requirements and to focus on potential land and water use conflicts. These crosscutting issues address the breadth of the iRESM research questions. More specifically, the team decided to propose the following mitigation question for the pilot region’s first UC analysis:

*What are the impacts of different levels of renewable portfolio standards on key outcomes such as regional electricity prices, emissions, land use conflicts, water use conflicts, agricultural land use decisions, and electric infrastructure needs?*

## 4.2 Establish Research Metrics/Decision Criteria

This step involves establishing the research metrics or decision criteria of concern to stakeholders. These metrics/criteria should be quantitatively defined so that they can be mapped to output variables from the iRESM framework component models, shown in **Figure 4.2**. Table 4.2 lists a set of potential criteria consistent with the RPS decision described above and indicates the iRESM component models that produce outputs relevant to these criteria. To make this decision context more tractable for the initial demonstration of the UC process, the scope of outcomes was narrowed to focus exclusively on UC with respect to regional electricity prices.



**Figure 4.2.** iRESM Framework

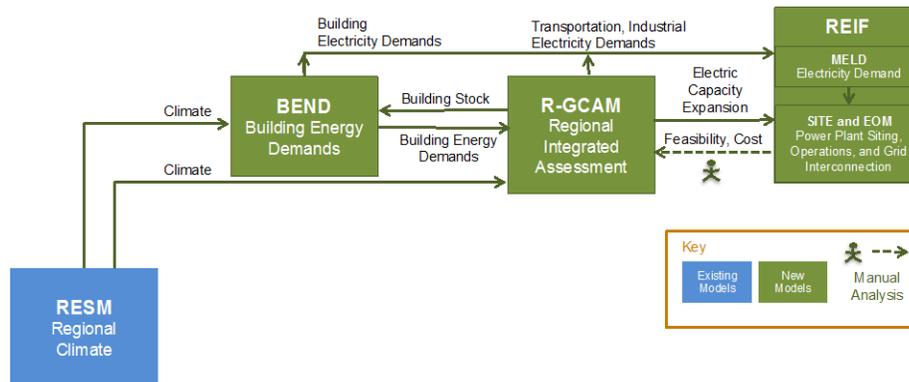
**Table 4.2.** Potential Decision Criteria for Regional RPS Decision

Potential Decision Criteria	Relevant iRESM Model(s)	Relevant Model Output
Regional GHG emissions (by species)	R-GCAM, REIF	R-GCAM: CO <sub>2</sub> and other long-lived GHG broken out by energy, land use. REIF: Electric sector CO <sub>2</sub> emissions
Fossil fuel consumption (oil, gas, coal)	R-GCAM	Consumption of coal and refined fuels by type
Electricity prices	R-GCAM, REIF	R-GCAM: Average marginal price of new generation, overall average price. REIF: Overall average price, locational marginal prices
Grid reliability	REIF	Unserved energy, locational marginal price
Cost of new transmission infrastructure requirements	REIF	Total miles and cost of new grid interconnections by voltage level; where major system upgrades are needed
Transportation fuel prices	R-GCAM	Fuel price index, TBD breakout for biofuels
Food prices	R-GCAM	Prices for major crops; aggregate prices for others
Land use/land cover change	R-GCAM	% change by category
Water availability	DCLM, DCLM-WM	Sub-basin water supply and allocation by use type

### 4.3 Identify Necessary Model Components of iRESM Framework

Figure 4.3 identifies the component models of the iRESM framework that would be involved in the determination of regional electricity prices. RESM provides the daily climatology utilized by the downstream models to drive energy demand and natural resource availability for energy supply needs.

BEND, the Building ENergy Demand model, calculates climate-dependent demands for heating, ventilation, and cooling services in the residential and commercial building sectors. BEND develops hourly energy demands based on detailed building energy use simulations for a wide range of different building types. These results are aggregated to the sub-regional, or utility zonal, spatial resolution so that they can be calibrated to actual utility loads and be used in the electricity operations simulation.



**Figure 4.3.** iRESM Component Models Determining Electricity Prices

BEND is integrated with R-GCAM, the Regional-Global Change Assessment Model, which simulates the impact of socioeconomic change, climate, climate policy, technological change, and resource availability on energy and agricultural technology selection, land use change, and the prices of primary and secondary energy sources (e.g., oil and electricity) and agricultural products through a market-based, economic equilibrium approach. R-GCAM provides BEND with the change in building stock characteristics over time along with energy prices, and BEND determines building energy demands in response. R-GCAM’s simulation of electricity markets is performed at the state level using annual load duration curves and annualized assumptions for electricity generation cost and performance. R-GCAM outputs annual average electricity prices by state.

The final model in the framework involved in the calculation of electricity prices is REIF, the Regional Energy Infrastructure Framework. REIF provides a more highly temporally and spatially resolved picture of the electricity system than is possible with R-GCAM. REIF is a check on the feasibility of the R-GCAM results as well as a more detailed examination of the impact of climate and climate policy on the electric grid. REIF has three components: the Model of Electricity Demand (MELD), a regional siting model for electric generation and transmission infrastructure (SITE), and the Electricity Operations Model (EOM). MELD receives the climate-dependent hourly building electricity demands from BEND and electricity demand from other sectors from R-GCAM, and it calculates total hourly electricity demand by utility zone. The GIS-based SITE model determines if and where the R-GCAM capacity expansion plan can be sited across the utility zones within the region. SITE addresses land suitability, water resource availability, grid interconnection costs, and the marginal value of new generation in each utility zone (provided by EOM) to simulate siting decisions. EOM optimizes power plant operations given the hourly supply, demand, and transmission constraints within the region. Among several other results, REIF outputs the hourly marginal price of electricity in each utility zone.

## 4.4 Identify Sources of Uncertainty in Models and Model Couplings

In this step, the UC team met with the RESM, BEND, R-GCAM, and REIF model development teams to identify the sources of uncertainty in their models with respect to the determination of electricity prices. The taxonomy of uncertainty presented earlier (see Figure 2.2) describing the four main sources of uncertainty (input quantification, model skill, model completeness, and integration and implementation) served as the basis for the discussions. Table 4.3 summarizes the results of these meetings and includes each team’s *a priori* judgment about which uncertainty sources are likely to be sensitive. Note that, due to the frequent overlap between model skill and model completeness issues, the results for these categories have been combined in the table. Also, the discussions focused on the individual models and did not address the integration aspects. Finally, the REIF discussion focused exclusively on the sources of uncertainty in the SITE model.

**Table 4.3.** Summary of Uncertainty Source Identification for Electricity Prices

Mode	Input Quantification	Model Skill/Model Completeness	<i>a priori</i> Judgment of Sensitive Uncertainty Sources
RESM	<ul style="list-style-type: none"> <li>• Different physics parameterizations</li> <li>• Boundary conditions (Community Earth System Model ensemble members)</li> <li>• RESM initial conditions</li> </ul>	<ul style="list-style-type: none"> <li>• Different physics representations</li> <li>• Bias correction may not be appropriate for future climate</li> <li>• Possible issues with interactions between the physics and the scale of the model</li> <li>• Unknown unknowns</li> </ul>	<ul style="list-style-type: none"> <li>• Alternate physics more important than aleatory uncertainty over long run (the opposite in the short run)</li> <li>• Land surface processes</li> <li>• Cloud processes</li> </ul>
BEND	<ul style="list-style-type: none"> <li>• Count of building types in a region</li> <li>• Building stock characteristics</li> <li>• Necessary spatial resolution</li> <li>• Inputs from RESM and R-GCAM</li> </ul>	<ul style="list-style-type: none"> <li>• Representativeness of locations chosen to represent the region</li> <li>• Count of buildings by building type is calibrated to reproduce base year total annual energy demand and hourly electricity profiles</li> </ul>	<ul style="list-style-type: none"> <li>• Uncertainty in climate likely to dominate uncertainties in base year building inventory</li> <li>• Demand response to price changes (interactions with R-GCAM)</li> <li>• Technological change (from R-GCAM)</li> </ul>
R-GCAM	<ul style="list-style-type: none"> <li>• Socioeconomics</li> <li>• Policy</li> <li>• Resource base</li> <li>• Technologies</li> <li>• Climate (from RESM)</li> </ul>	<ul style="list-style-type: none"> <li>• Market clearing approach not consistent with electric system regulatory framework or water rights framework</li> <li>• Model assumes decisions are made rationally</li> <li>• Electricity system not modeled explicitly</li> <li>• Industrial sector not modeled at same detail as energy, transportation, agriculture</li> </ul>	<ul style="list-style-type: none"> <li>• Policy</li> <li>• Resource base</li> <li>• Technology characteristics</li> <li>• Climate</li> </ul>

SITE	<ul style="list-style-type: none"> <li>• Criteria used to mask unsuitable land including policy and social issues</li> <li>• Parameters LMP-based siting and grid interconnection cost algorithms</li> <li>• GIS layers are from different dates and resolutions</li> <li>• Technology siting requirements</li> <li>• Inputs from upstream models</li> </ul>	<ul style="list-style-type: none"> <li>• Technology siting order, mix affects results</li> <li>• Choice of objectives for siting algorithms to simulate regional planning</li> <li>• May overstate or understate restrictions on suitable land</li> <li>• Choice of spatial resolution—finer resolution better represents land-based constraints</li> <li>• Details of local policies not captured</li> <li>• Future policies, technology requirements are unknown unknowns</li> </ul>	<ul style="list-style-type: none"> <li>• Suitability criteria</li> <li>• Parameters and structure of siting algorithm</li> <li>• Future land and technology policy</li> <li>• Water availability</li> </ul>
------	--	--	---

---

## 4.5 Select UC Methods

As discussed in Section 2, the appropriate UC methods depend upon the nature of the uncertainties identified and the nature of the decisions to be addressed by the iRESM models. The uncertainties have not yet been fully identified and are expected to emerge as experience is gained with running the individual models. However, examples can be provided concerning how the process of selection will work in the case of the models shown in Table 4.3. Research in FY 2011 shows that uncertainty in physics representations within climate models (RESM) and uncertainty in public policy toward carbon emissions (R-GCAM) will be major drivers of uncertainty in those models. The impact of physics uncertainties has been addressed by the climate community by running ensembles of multiple climate models and then developing distributions of output from the model runs (either unweighted or weighted by model performance in backcasting actual historical climate) to quantify and describe the uncertainties involved. The iRESM team is actively investigating how best to characterize uncertainties from this source. Multiple boundary conditions may be available to run RESM, and the results of those runs can be used to characterize the impacts of boundary conditions. To characterize the impacts of public policy toward carbon emissions, the iRESM team is discussing the development of a number of numerical experiments. These experiments will demonstrate the impacts on electricity demand of emissions scenarios that are consistent with IPCC RCPs. Uncertainties in this case may be partially quantified using descriptions of ranges or overlapping distributions of scatterplots of scenario outcomes as described in Section 3.2.2.

## 4.6 Perform Screening/Sensitivity Analysis

In FY12, a particular coupling of iRESM component models and one or more decision criteria will be selected for a demonstration of the UC process. After performing the uncertainty source identification and UC methods selection steps as described above in Section 4.1 through Section 4.5, the team will perform a structured sensitivity analysis to determine the subset of the uncertainty sources that are the biggest drivers of the uncertainty of each criterion (e.g., electricity prices). This is a deterministic process that requires the team to develop bounds for the sources of uncertainty amenable to quantification (e.g., high and low values for the U.S. population growth rate during 2050-2100 consistent with IPCC scenario RCP4.5 [Moss et al. 2010]). Ideally, these bounds would be estimated to represent the 5th and 95th percentiles on a cumulative probability distribution, but other criteria may be used. Once the bounds have

been developed, a separate simulation of the coupled models needs to be performed for each uncertainty bound.

As an example of a screening analysis, one energy-economy experiment will isolate the impact of climate change on R-GCAM. Two R-GCAM simulations for 2005-2100 will be run and compared: both will incorporate socioeconomic and policy assumptions consistent with RCP4.5, but one will be driven by the RCP4.5 climate and the other will have no climate drivers at all (essentially assuming a static climate over the simulation horizon). The goal of this deterministic experiment, however, is to understand the impact of including climate drivers (or not) on R-GCAM's simulation of the evolution of the regional energy-economy. A second experiment will couple BEND and R-GCAM to explore the results of their interactions under RCP4.5. This coupling will cause BEND's calculations to respond to technological change and population growth as well as to energy price changes (demand response). The BEND outputs will replace R-GCAM's calculation of the pilot region's BEND (based on heating and cooling degree-days).

## **4.7 Characterize/Quantify Uncertainty in Sensitivity Drivers**

After the simulations for the sensitivity analysis are complete, the results will be analyzed to rank the uncertainty sources in terms of their relative impact on the variation in the evaluation metrics, similar to the method shown in Figure 3.2. The uncertainties with the largest impact on the metric (i.e., the “sensitive variables”) are candidates for quantification with probability distributions. This process will draw as appropriate on the techniques and methods described in Section 2, quantifying uncertainties where possible.

## **4.8 Design UP Approach**

Methods for propagation are described in Section 2.3, with different methods being appropriate for different types of uncertainty. Uncertainty that can be quantified is typically propagated by Monte Carlo methods through multiple implementations of the coupled models. However, as described in Section 2.3, semi-quantitative and qualitative methods are also available and will be used where appropriate. Probability distributions may be sampled independently, but it is generally considered better practice to account for known dependencies among input variables through sampling correlations, and this has been done in uncertainty evaluation of predecessors to R-GCAM, for example (Scott et al. 1999). Latin Hypercube-based methodologies exist to allow required correlations to be imposed on the input marginal distributions (Iman and Davenport 1982; Swiler and West 2010). Propagation will require software to support and facilitate UP, as discussed in Section 2.3.

Due to the large number of simulations likely to be required for a sufficient sample size, it is possible that run time issues may arise in propagating uncertainties across the coupled models of the framework. Initial work done during FY 2011 suggests that run times may be extremely long (running to months on a parallel processing platform with ~2000 cores available) for fully integrated runs of the complete iRESM suite. The team will investigate flexible architectural approaches to shortening run times by running component codes on a coarser grid and/or using larger time steps. In addition, the team will investigate, as necessary, the use of fast-running surrogate models based on reduced-form processes or statistically derived response surfaces to emulate the behaviors of the detailed component models while having substantially shorter run times. In addition, the new computing resources available through a recent

institutional computing investment (14,000+ cores available) may resolve some or all run time issues for UP.

## **4.9 Perform UP**

This step in the process is the execution of the UP strategy designed in the previous step. In general, it is expected that the number of simulations and run times needed for UP will be highly variable and dependent on the relevant model couplings and the number of sensitive variables. The UC experiment defined in FY12 will move into the UP phase in FY13. The lessons learned from this experiment regarding the selection of UP engine, surrogate models, and so on, will be incorporated into the evolving iRESM platform architecture.

## **4.10 Evaluate and Interpret Results**

The results of the UP will be analyzed and reviewed to determine the most effective ways in which to depict and communicate them. There are a variety of known psychological barriers and communication challenges surrounding uncertainty and scenarios in general. The preliminary work done in FY 2011 on stakeholder needs (Rice et al. 2010) by the iRESM Focus Area 3 team is being expanded into a broader and deeper stakeholder engagement research project. Research is being done in the iRESM initiative to facilitate application of regional modeling frameworks to adaptation and mitigation decisions and to add to the range of tools available for facilitating communication of uncertainty to stakeholders through participatory scenario simulation exercises. Uncertainty regarding alternate/preferred means of communicating uncertainty analysis results will be incorporated into the analysis and visualization of the results.

## 5.0 Conclusion

This paper has described a new research approach to the process of UC to help address the challenges of regional climate change mitigation and adaptation decisions. The iRESM initiative is developing a prototype suite of models to facilitate integrated assessments related to sustainable economic and energy development in the context of climate change, specifically addressing potential conflicts between this development and various uses of land and water resources. The iRESM project has adopted as a fundamental principle that any model or set of models addressing either climate mitigation or adaptation must include as a fundamental element of its design methods for characterizing this uncertainty, quantifying it to the extent possible, and depicting it in a manner understood by its intended audience. The UC process under development focuses on identifying and characterizing the key uncertainties in data inputs, individual model structures, and coupled models to determine the robustness of the framework's results. A key aspect of the process is to focus both modeling and UC on specific information needs of stakeholders (for example, the decision criteria that will be considered by decision makers in evaluating different options). Without such focus, UC can devolve into an unbounded and unhelpful process that, in practice, is difficult if not impossible to complete. This decision-specific orientation for UC has multiple implications for the iRESM initiative that will continue to be explored in the work of iRESM Focus Area 3, including: the importance of stakeholder interactions and the development of methods for communicating results; the development of a flexible model architecture that will facilitate application of subsets of the iRESM model components relevant to particular applications; and the identification of approaches for reducing model run times to facilitate UQ, including development of surrogate models.

## 6.0 References

- Adams BM, KR Dalbey, MS Eldred, DM Gay, LP Swiler, WJ Bonhoff, JP Eddy, K Haskell, and PD Hough. 2010. *DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.0 User's Manual*. SAND2010-2183, Sandia National Laboratories, Albuquerque, New Mexico.
- Aleskerov F, D Bouyssou, and B Monjardet. 2007. *Utility Maximization, Choice and Preference (Studies in Economic Theory)*. 2nd edition, Springer, Berlin, Germany.
- ANS, IEEE, and NRC (American Nuclear Society, Institute of Electrical and Electronic Engineers, and U.S. Nuclear Regulatory Commission). 1983. *A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants, Final Report*. NUREG/CR-2300, U.S. Nuclear Regulatory Commission, Office of Regulatory Research, Washington, D.C.
- Apostolakis GE. 1989. "Uncertainty in Probabilistic Safety Assessment." *Nuclear Engineering and Design* 115:173-179.
- Bell ML, BF Hobbs, and H Ellis. 2003. "The use of multi-criteria decision-making methods in the integrated assessment of climate change: implications for IA practitioners." *Socio-Economic Planning Sciences* 37:289-316.
- Budnitz RJ, G Apostolakis, DM Boore, LS Cluff, KJ Coppersmith, CA Cornell, and PA Morris. 1997. *Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts*. NUREG/CR-6372, U.S. Nuclear Regulatory Committee, Washington, D.C.
- Cacuci DG and ME Schlesinger. 1994. "On the application of the adjoint method of sensitivity analysis to problems in the atmospheric sciences." *Atmosfera* 7:47-59.
- Caselton WF and W Luo. 1992. "Decision making with imprecise probabilities: Dempster-Shafer Theory and application." *Water Resources Research* 28(12): 3071-3083.
- CCPS (Center for Chemical Process Safety). 2008. *Guidelines for Hazard Evaluation Procedures*. 3rd ed. Wiley-AIChE, Hoboken, New Jersey.
- CIG (Climate Impacts Group). 2009. *The Washington Climate Change Impacts Assessment: Evaluating Washington's Future in a Changing Climate*. Eds. M McGuire Elsner, J Littell, and L Whitely Binder. Center for Science in the Earth System, Joint Institute for the Study of the Atmosphere and Oceans, University of Washington, Seattle, Washington. Accessed on September 23, 2011 at <http://cses.washington.edu/db/pdf/wacciareport681.pdf>.
- Clarke L, J Edmonds, H Jacoby, H Pitcher, J Reilly, and R Richels. 2007. *Scenarios of Greenhouse Gas Emissions and Atmospheric Concentrations*. Department of Energy, Office of Biological and Environmental Research, Washington, D.C.
- CORDEX (Coordinated Regional climate Downscaling Experiment) webpage. Accessed on September 23, 2011 at <http://www.meteo.unican.es/en/projects/CORDEX>.

- Darby JL. 2009. "Tools for Evaluating Risk of Terrorist Acts Using Fuzzy Sets and Belief/Plausibility." In 28th proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society, Cincinnati, Ohio.
- De Finetti B. 1972. *Probability, Induction and Statistics*. John Wiley & Sons, Ltd., London, United Kingdom.
- Delsole T and J Shukla. 2010. "Model Fidelity versus Skill in seasonal forecasting." *Journal of Climate* 23:4794-4806.
- Dempster AP. 1968. "A generalization of Bayesian inference." *Journal of the Royal Statistical Society B*(30):205-247.
- Dubois D and H Prade. 1990. "Aggregation of possibility measures." Chapter in *Multiperson Decision Making Using Fuzzy Sets and Possibility Theory*. Eds. J Kacprzyk and M Fedrizzi. Kluwer Academic Publishers, Dordrecht, the Netherlands.
- Dubois D and H Prade. 1995. "Possibility theory as a basis for qualitative decision theory." In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Quebec.
- Dubois D and H Prade. 2003. "Possibility Theory and its Applications: A Retrospective and Prospective View." In proceedings of *The 12th IEEE International Conference on Fuzzy Systems IEEE International Conference on Fuzzy Systems*, Vol. 2. May 25-28, 2003, St. Louis, Missouri.
- Dubois D. 2006. "Possibility Theory and Statistical Reasoning." *Computational Statistics and Data Analysis* 51(1) 47-69.
- Duong MH. 2003. "Imprecise Probability Bridges Scenario-Forecast Gap." In proceedings of the 1st International Workshop on Integrated Climate Models: An interdisciplinary assessment of climate impacts and policies. September 30-October 3, 2003. International Center for Theoretical Physics, Trieste, Italy.
- Eldred MS, CG Webster, and PG Constantine. 2008. "Design Under Uncertainty Employing Stochastic Expansion Methods." In proceedings of the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference. September 10-12, 2008, Victoria, British Columbia, Canada.
- EMF (Energy Modeling Forum). 2010. *About the Energy Modeling Forum at Stanford University*. Accessed on September 23, 2010 at [http://emf.stanford.edu/docs/about\\_emf/#leading](http://emf.stanford.edu/docs/about_emf/#leading).
- Ferson S, V Kreinovich, L Ginzburg, DS Myers, and K Sentz. 2003. *Constructing Probability Boxes and Dempster-Shafer Structures*. SAND2002-4015, Sandia National Laboratories, Albuquerque, New Mexico.
- Gelfand AE and AFM Smith. 1990. "Sampling-based approaches to calculating marginal densities." *Journal of the American Statistical Association* 85:398-409.
- Georgi F and LO Mearns. 2002. "Calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the 'reliability ensemble averaging' (REA) method." *Journal of Climate* 15(10): 1141-115.

- Groves DG, D Yates, C Tebaldi. 2008. "Developing and applying uncertain global climate change projections for regional water management planning." *Water Resources Research* 44(12):W12413.
- Hall J, G Fu, and J Lawry. 2007. "Imprecise Probabilities of Climate Change: Aggregation of Fuzzy Scenarios and Model Uncertainties." *Climate Change* 81:265-281.
- Heinrich G, L Basson, B Cohen, M Howells, and J Petrie. 2007. "Ranking and selection of power expansion alternatives for multiple objectives under uncertainty." *Energy* 32(12):2350-2369.
- Helton JC and FJ Davis. 2003. "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems." *Reliability Engineering and System Safety* 81(1):23-69.
- Helton JC, JD Johnson, CJ Sallaberry, and CB Storlie. 2006. *Survey of Sampling-Based methods for Uncertainty and Sensitivity Analysis*. SAND2006-2901, Sandia National Laboratories, Albuquerque, New Mexico.
- Helton, JC, JD Johnson, WL Oberkampf, and CJ Sallaberry. 2008. *Presentation of Analysis Results Involving Aleatory and Epistemic Uncertainty*. SAND2008-4379, Sandia National Laboratories, Albuquerque, New Mexico.
- Herbranson WT and J Schroeder. 2010. "Are Birds Smarter Than Mathematicians? Pigeons (*Columba livia*) Perform Optimally on a Version of the Monty Hall Dilemma." *Journal of Comparative Psychology* 124:1-13.
- Hora SC. 1996. "Aleatory and Epistemic Uncertainty in Probability Elicitation with an Example from Hazardous Waste Management." *Reliability Engineering and System Safety* 54:217-223.
- IAMC (Integrated Assessment Modeling Consortium). 2011. *Working Group on Representative Concentration Pathways (RCP)*. Accessed on May 26, 2011 at <http://iamconsortium.org/>.
- Iman RL and JM Davenport. 1982. "Rank Correlation Plots for Use with Correlated Input Variables." *Communications in Statistics: Simulation and Computation* 11(3):335-360.
- Iman RL and WJ Conover. 1980. "Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment." *Communications in Statistics: Theory and Methods* 1980 A9(17):1749-1842.
- IPCC (Intergovernmental Panel on Climate Change). 2000. *Special Report on Emission Scenarios*. Cambridge University Press, Cambridge, United Kingdom.
- IPCC (Intergovernmental Panel on Climate Change). 2001. "Quantifying Uncertainties in Practice." Chapter 6 in *IPCC Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. Intergovernmental Panel on Climate Change, Geneva, Switzerland. Accessed on September 27, 2011 at <http://www.grida.no/climate/gpg/english.htm>.
- Kahneman D, P Slovic, and A Tversky (eds.). 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge, United Kingdom.
- Kaplan S and BJ Garrick. 1981. "On The Quantitative Definition of Risk." *Risk Analysis* 1(1):11-27.

- Kaplan S. 1990. "Bayes is for Eagles." *IEEE Transactions on Reliability* 39(2): 130-131.
- Katz RW. 2002. "Techniques for Estimating Uncertainty in Climate Change scenarios and Impact Studies." *Climate Research* 20:167-185.
- Keeney RL and D von Winterfeldt. 1991. "Eliciting probabilities from experts in complex technical problems." *IEEE Transactions on Engineering Management* 38:191-201.
- Knutti R, R Furrer, C Tebaldi, J Cermak, and GA Meehl. 2010. "Challenges in combining projections from multiple climate models." *Journal of Climate* 23:2739-2758.
- Knutti R. 2008. "Should we believe model predictions of future climate change?" *Philosophical Transactions of the Royal Society A* 2008(366):4647-4664.
- Lempert RJ and M Collins. 2007. "Managing the risk of uncertain threshold responses: Comparison of robust, optimum, and precautionary approaches." *Risk Analysis* 27(4):1009-1026.
- Lempert RJ, ME Shlesinger, and SC Bankes. 1996. "When we don't know the costs or the benefits: Adaptive strategies for abating climate change." *Climate Change* 33(2):235-274.
- Linstone H and M Turoff. 1975. *The Delphi Method: Techniques and Applications*. Addison-Wesley, Reading, Massachusetts.
- Lipinski RJ, AK Mast, DA Powers, and JV Walker. 1985. *Uncertainty in Radiological Release Under Specific LWR Accident Conditions*. SAND84-0410, Sandia National Laboratories, Albuquerque, New Mexico.
- Luo WB and B Caselton. 1997. "Using Dempster-Shafer Theory to represent climate change uncertainties." *Journal of Environmental Management* 49:73-93.
- Manne, A. 1996. *Hedging Strategies for Global Dioxide Abatement: A Summary of Poll Results—EMF Subgroup Analysis for Decisions Under Uncertainty*. Working paper 14.2 Energy Modeling Forum. Accessed on September 26, 2011 at <http://emf.stanford.edu/files/pubs/22418/WP1402.pdf>.
- Martz HF and RA Waller. 1982. *Bayesian Reliability Analysis*. John Wiley & Sons, New York, New York.
- McCarthy JJ, OF Canziani, NA Leary, DJ Dokken, and KS White (eds.). 2001. *Climate Change 2001: Impacts, Adaptation, and Vulnerability*. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Figure 7-2. Cambridge University Press, New York, New York.
- McJeon H, L Clarke, P Kyle, M Wise, A Hackbarth, B Bryant, and RJ Lempert. 2010. *Technology Interactions among Low Carbon Energy Technologies: What Can We Learn from a Large Number of Scenarios?* Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, Maryland.
- Means E III, M Laugier, J Daw, L Kaatz, and M Waage. 2010. *Decision Support Planning Methods: Incorporating Climate Change Uncertainties into Water Planning*. Water Utility Climate Alliance, San Francisco, California.

- Miles EL, MM Elsner, JS Littell, LW Binder, and DP Lettenmaier. 2010. "Assessing regional impacts and adaptation strategies for climate change: the Washington Climate Change Impacts Assessment." *Climatic Change* 102(1-2):9-17.
- Morgan MG and M Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, Cambridge, United Kingdom.
- Morgan MG, H Dowlatabadi, M Henrion, D Keith, R Lempert, S McBride, M Small, and T Wilbanks. 2009. *Best Practice Approaches for Characterizing, Communicating and Incorporating Scientific Uncertainty in Climate Decision Making*. U.S. Climate Change Science Program, Washington, D.C.
- Morgan MG, LF Pitelka, and E Shevliakova. 2001. "Elicitation of Expert Judgments of Climate Change Impacts on Forest Ecosystems." *Climatic Change* 49:279-307.
- Moss R and SH Schneider. 2000. "Uncertainties in the IPCC TAR: Recommendations to lead authors for more consistent assessment and reporting." In *Guidance Papers on the Cross Cutting Issues of the Third Assessment Report of the IPCC*. R Pachauri, T Taniguchi, K Tanaka (eds.). World Meteorological Organisation, Geneva, Switzerland.
- Moss R H, JA Edmonds, , KA Hibbard, MR Manning, SK Rose, DP van Vuuren, TR Carter, S Emori, M Kainuma, T Kram, GA Meehl, JFB Mitchell, N Nakicenovic, K Riahi, S J Smith, R J Stouffer, AM Thomson, JP Weyant, and TJ Wilbanks. 2010. "The next generation of scenarios for climate change research and assessment." *Nature* 463(7282):747-756.
- Nature*. 2010. "Validation Required." 463(7283):849.
- NRC (U.S. Nuclear Regulatory Commission). 1990. *Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants*. NUREG-1150, Nuclear Regulatory Commission, Washington, D.C.
- Parikh CR, MJ Pont, and NB Jones. 2001. "Application of Dempster-Shafer theory in condition monitoring systems: A case study." *Pattern Recognition Letters* 22(6-7):777-785.
- Park N. 2010. "Application of Dempster-Shafer theory of evidence to GIS-based landslide susceptibility analysis." *Environmental Earth Sciences* 62(2):367-376.
- Paté-Cornell ME. 1996. "Uncertainties in risk analysis: Six levels of treatment." *Reliability Engineering and System Safety* 54:95-111.
- Peterson S. 2006. "Uncertainty and economic analysis of climate change: A Survey of Approaches and Findings." *Environmental Modeling and Assessment* 11:1-17.
- Raferty AE, F Balabdaoui, T Gneiting, and M Polakowski. 2003. *Using Bayesian Model Averaging to Calibrate Forecast Ensembles*. Technical Report No. 440, Department of Statistics, University of Washington, Seattle Washington.
- Regan HM, Y Ben-Haim, B Langford, WG Wilson, P Lunderberg, SJ Andelman, and MA Burgman. 2005. "Robust decision-making under severe uncertainty for conservation management." *Ecological Applications* 15(4):1471-1477.

Reilly JM, JA Edmonds, RH Garner, and AL Brenkert. 1987. "Uncertainty analysis of the IEA/ORU CO<sub>2</sub> emissions model." *The Energy Journal* 8 (3):1-29.

Rice J, R Moss, P Runci, K Anderson. 2010. *Understanding Decision Support Needs For Climate Change Mitigation And Adaptation-Midwest Region Stakeholders*. PNNL-20104, Pacific Northwest National Laboratory, Richland, Washington.

Rosenzweig C. 2010. *The Agriculture Model Intercomparison and Improvement Project (AgMIP) (Invited)*. Abstract #GC34A-01 from the American Geophysical Union, Fall Meeting. Accessed on May 27, 2011 at <http://adsabs.harvard.edu/abs/2010AGUFMGC34A..01R>.

Rotmans J and MBA Van Asselt. 2001. "Uncertainty in integrated assessment modeling: A labyrinthic path." *Integrated Assessment* 2:43-55.

Sadiq R, Y Kleiner, and BA Rajani. 2006. "Estimating risk of contaminant intrusion in water distribution networks using Dempster-Shafer theory of evidence." *Civil Engineering and Environmental Systems* 23(3):129-141.

Sanfilippo AP, AJ Cowell, EL Malone, RM Riensche, JJ Thomas, SD Unwin, PD Whitney, and PC Wong. 2009. "Technosocial Predictive Analytics in Support of Naturalistic Decision Making." In proceedings of *The 9th Bi-annual International Conference on Naturalistic Decision Making*. Ed. BLW Wong. June 23-26, 2009, London, United Kingdom. British Computer Society, Swindon, United Kingdom.

Scott MJ, RD Sands, J Edmonds, AM Liebetrau, and DW Engel. 1999. "Uncertainty in Integrated Assessment Models: Modeling with Minicam 1.0." *Energy Policy* 27:855-879.

Sentz K and S Ferson. 2002. *Combination of Evidence in Dempster-Shafer Theory*. SAND2002-0835, Sandia National Laboratories, Albuquerque, New Mexico.

Shafer G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey.

Shafer G. 1990. "Perspectives on the theory and practice of belief functions." *International Journal of Approximate Reasoning* 4(5-6):323-362.

Smarandache F. 2004. *An In-Depth Look at Information Fusion Rules and the Unification of Fusion Theories*. Accessed on September 27, 2011 at <http://arxiv.org/pdf/cs/0410033v2>.

Sokolov AB, PH Stone, CE Forest, R Prinn, MC Sarofim, M Webster, S Paltsev, CA Schlosser, D Kicklighter, S Dutkiewicz, J Reilly, C Wang, B Felzer, JM Melillo, and HD Jacoby. 2009. "Probabilistic forecast for Twenty-First-Century climate based on uncertainties in emissions (without policy) and climate parameters." *Journal of Climate* 22:5175-5204.

Spetzler CS and CA Stael von Holstein. 1975. "Probability encoding in decision analysis." *Management Science* 22:340-358.

Swart R, L Bernstein, M Ha-Duong, and A Peterson. 2009. "Agreeing to disagree: uncertainty management in assessing climate change, impacts and responses by the IPCC." *Climatic Change* 92(1):1-29.

Swiler LP and NJ West. 2010. "Importance Sampling: Promises and Limitations." In proceedings of the 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. April 12-15, 2010, Orlando, Florida.

Tao F, Z Zhang, J Liu, M Yokozawa. 2009. "Modeling the impacts of weather and climate variability on crop productivity over a large area: A new super-ensemble-based probabilistic projection." *Agricultural and Forest Meteorology* 149:1266-1278.

Tebaldi C and R Knutti. 2007. "The use of the multi-model ensemble in probabilistic climate projections." *Philosophical Transactions of the Royal Society A* 365:2053-2075.

Tebaldi C, RL Smith, D Nychka, and LO Mearns. 2005. "Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multi-model ensembles." *Journal of Climate* 18:1524-1540.

Tomassini L, R Knuth, GK Plattner, DP van Vuuren, TF Stocker, RB Howarth, and ME Borsuk. 2010. "Uncertainty and risk in climate projections for the 21<sup>st</sup> century: comparing mitigation to non-intervention scenarios." *Climatic Change* 103(4):399-422.

Toth FL, M Mwandosya, C Carraro, J Christensen, J Edmonds, B Flannery, C Gay-Garcia, H Lee, KM Meyer-Abich, E Nikitina, A Rahman, R Richels, Y Ruqiu, A Villavicencio, Y Wake, and J Weyant. 2001. "Decision-making Frameworks." In *Climate Change 2001: Mitigation*. Eds. B Metz, O Davidson, R Swart, and J Phan. Cambridge University Press, New York, New York.

Tversky A and D Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases." *Science* 185:1124-1131.

Unwin SD and BA Fecht. 2009. "Consumer choice and Dempster-Shafer models of threat prioritization for emerging dual-use technologies: Application to synthetic biology." *Defense and Security Analysis* 25(1):37-52.

Unwin SD, EG Cazzoli, RE Davis, M Khatib-Rahbar, M Lee, H Nourbakhsh, CK Park, and E Schmidt. 1989. "An information-theoretic basis for uncertainty analysis: application to the QUASAR Severe Accident Study." *Reliability Engineering and System Safety* 26(2):143-162.

Unwin SD. 1986. "A fuzzy set theoretic foundation for vagueness in uncertainty analysis." *Risk Analysis* 6:27.

Unwin SD. 1988. "In defense of the probabilistic decision." *Reliability Engineering and System Safety* 23(4):315.

Unwin SD. 1989. "Expert opinion and information theory." *Reliability Engineering and System Safety* 26:186.

USCOE/USBR/NMISC (U.S. Army Corps of Engineers, Albuquerque District, U.S. Department of Interior, Bureau of Reclamation, and New Mexico Interstate Stream Commission). 2007. *Upper Rio Grande Water Operations Review Final Environmental Impact Statement*. FES-07-05, Bureau of Reclamation, Albuquerque Area Office, Albuquerque, New Mexico.

Webster M, AP Sokolov, JM Reilly, CE Forest, S Paltsev, A Schlosser, C Wang, D Kicklighter, M Sarofim, J Melillo, RG Prinn, and HD Jacoby. 2009. *Analysis of Climate Policy Targets Under Uncertainty*. Report No. 180. Massachusetts Institute of Technology Joint Program on the Science and Policy of Global Change, Cambridge, Massachusetts.

Webster M, C Forest, J Reilly, M Babiker, D Kicklighter, M Mayer, R Prinn, M Sarofim, A Sokolov, P Stone, and C Wang. 2003. "Uncertainty analysis of climate change and policy response." *Climatic Change* 61:295-320.

Webster M, S Paltsev, J Parsons, J Reilly, and H Jacoby. 2008. *Uncertainty in Greenhouse Gas Emissions and Costs of Atmospheric Stabilization*. Report No. 165. Massachusetts Institute of Technology Joint Program on the Science and Policy of Global Change, Cambridge, Massachusetts.

Wheeler TA, SC Hora, WR Cramond, and SD Unwin. 1989. *Analysis of Core Damage Frequency from Internal Events: Expert Judgment Elicitation*. NUREG/CR-4550, Vol. 2, U.S. Nuclear Regulatory Commission, Washington, D.C.

Wigley TML. 2008. *MAGICC/SCENGEN 5.3: User Manual (Version 2)*. University Corporation for Atmospheric Research, Boulder, Colorado. Accessed on September 27, 2011 at <http://www.cgd.ucar.edu/cas/wigley/magicc/UserMan5.3.v2.pdf>.

Wilby RL and S Dessai. 2010. "Robust adaptation to climate change." *Weather* 65:180-185.

Williamson RC, and T Downs. 1990. "Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds." *International Journal of Approximate Reasoning* 4(2):89-158.

Wojtkiewicz SF, MS Eldred, RV Field, Jr., A Urbina, and JR Red-Horse. 2001. "Uncertainty quantification in large computational engineering models." In *Proceedings of the 42nd AIAA/ASME/ASCE/AHSASC Structures, Structural Dynamics, and Materials Conference*. April 16-19, 2001, Seattle, Washington.

WRCP (World Climate Research Program). 2011. *CMIP: Coupled Model Intercomparison Project - Overview*. Accessed on September 21, 2011 at <http://cmip-pcmdi.llnl.gov/>.

Zadeh LA. 1965. "Fuzzy sets." *Information and Control* 8:338-353.

Zadeh LA. 1978. "Fuzzy sets as a basis for a theory of possibility." *Fuzzy Sets and Systems* 1:3-28.

Zadeh LA. 1981. "Possibility theory and soft data analysis." In *Mathematical Frontiers of the Social and Policy Sciences*. Eds. L Cobb and RM Thrall. Westview Press, Boulder, Colorado.



**Pacific Northwest**  
NATIONAL LABORATORY

*Proudly Operated by **Battelle** Since 1965*

902 Battelle Boulevard  
P.O. Box 999  
Richland, WA 99352  
1-888-375-PNNL (7665)

[www.pnl.gov](http://www.pnl.gov)



U.S. DEPARTMENT OF  
**ENERGY**