IIASA

Learning from global emissions scenarios

Brian C. O'Neill and Nebojsa Nakicenovic

IIASA REPRINT

Learning from global emissions scenarios

Brian C. O'Neill

International Institute for Applied Systems Analysis, Laxenburg, Austria, and Institute for Study of Society and Environment, National Center for Atmospheric Research, Boulder, CO, USA

Nebojsa Nakicenovic

International Institute for Applied Systems Analysis, Laxenburg, Austria, and Vienna University of Technology, Vienna, Austria

Learning from global emissions scenarios

Brian C O'Neill[1,2] **and Nebojsa Nakicenovic**[2,3]

[1] Institute for the Study of Society and Environment, National Center for Atmospheric
Research, Boulder, CO, USA
[2] International Institute for Applied Systems Analysis, Laxenburg, Austria
[3] Vienna University of Technology, Vienna, Austria

E-mail: boneill@ucar.edu

Abstract
Scenarios of global greenhouse gas emissions have played a key role in climate change analysis
for over twenty years. Currently, several research communities are organizing to undertake a
new round of scenario development in the lead-up to the Fifth Assessment Report of the
Intergovernmental Panel on Climate Change (IPCC). To help inform this process, we assess a
number of past efforts to develop and learn from sets of global greenhouse gas emissions
scenarios. We conclude that while emissions scenario exercises have likely had substantial
benefits for participating modeling teams and produced insights from individual models,
learning from the exercises taken as a whole has been more limited. Model comparison
exercises have typically focused on the production of large numbers of scenarios while
investing little in assessing the results or the production process, perhaps on the assumption that
later assessment efforts could play this role. However, much of this assessment potential
remains untapped. Efforts such as scenario-related chapters of IPCC reports have been most
informative when they have gone to extra lengths to carry out more specific comparison
exercises, but in general these assessments do not have the remit or resources to carry out the
kind of detailed analysis of scenario results necessary for drawing the most useful conclusions.
We recommend that scenario comparison exercises build-in time and resources for assessing
scenario results in more detail at the time when they are produced, that these exercises focus on
more specific questions to improve the prospects for learning, and that additional scenario
assessments are carried out separately from production exercises. We also discuss the obstacles
to better assessment that might exist, and how they might be overcome. Finally, we recommend
that future work include much greater emphasis on understanding how scenarios are actually
used, as a guide to improving scenario production.

Keywords: scenarios, climate change, integrated assessment, learning

1. Introduction

Scenario development has a long history of use both in and
outside of the climate change field. Scenarios can serve a
variety of purposes. One key dimension of this variety is
the extent to which scenario exercises serve either process-
or product-oriented purposes (Wilkinson and Eidinow 2008,
Hulme and Dessai 2008b, O'Neill *et al* 2008). For some
exercises, the process of scenario development is itself the
primary goal, for example as a means to find commonalities
across different perspectives, achieve consensus on goals, or

come to a shared understanding of challenges. Scenarios
developed in the business community to promote strategic
thinking have frequently been process-oriented. For others,
the product—the content of the scenarios developed—is the
main goal. Once produced, these products then have a life
of their own that is generally divorced from the process
that generated them, serving many further purposes. Global
emissions scenarios, and in particular those that are the focus of
this paper, are regarded as mainly product-oriented exercises.
What is of most interest are the emissions pathways that are
produced, how they relate to the various factors driving them,

and what the results tell us about the prospects for future climate change, impacts, and mitigation.

Over the past decade there have been methodological developments in scenario production that have begun to merge approaches from the two different traditions. The integrated scenario communities have begun to combine the primarily interpretative and narrative-based scenario analyses undertaken by Royal Dutch/Shell (Wack 1985a, 1985b, Schwartz 1991) and other groups, which were mainly process-oriented, with primarily product-oriented analytical and quantitative integrated modeling work. The result has been new sets of scenarios that, while still clearly falling in the product-oriented category, draw on approaches from both camps by combining the development of detailed narrative storylines with their 'quantification' in various integrated models (Raskin *et al* 1998, Nakicenovic *et al* 2000). For example, the Special Report on Emissions Scenarios (SRES) (Nakicenovic *et al* 2000) of the Intergovernmental Panel on Climate Change (IPCC) represented one such early effort to include a process for scenario development that started with storylines that were translated into four sets of quantitative scenarios. The process was highly iterative in its attempt to combine interpretive storylines with descriptive modeling, differing sharply from previous IPCC scenario development exercises (Houghton *et al* 1990, Leggett *et al* 1992).

A number of subsequent scenario development efforts have expanded on and improved this approach. The Millennium Ecosystem Assessment scenarios (Carpenter *et al* 2005) also cut across the narrative and quantitative divide in their four scenario sets. The US Climate Change Science Program (CCSP) scenarios (Clarke *et al* 2007, Parson *et al* 2007) included both components as well, although not in an integrated form: one volume contained a detailed assessment of scenario development processes, including the role of interpretative dimensions, while the other described an exercise to develop more traditional quantitative scenarios by three integrated assessment modeling groups. The integration of process and content still remains as one of the major challenges to be overcome in the further development of emissions scenarios.

Little attention has been devoted to understanding the use, value, or success of scenarios. Hulme and Dessai (2008a, 2008b) propose that success might be judged in terms of the predictive value, contribution to decision-making, or facilitation of learning resulting from scenarios, and illustrate this possibility using as an example the scenarios developed as part of the UK Climate Impacts Program. Here we address the last of their three criteria for judging success—the facilitation of learning—by evaluating how much we have learned from a number of prominent global emissions scenario exercises. In contrast to Hulme and Dessai (2008a), our main focus is not the learning among participants in the scenario development processes that may have taken place (although we agree that this is one useful outcome). Rather, we focus on learning from evaluating the products: the scenarios produced. We measure this learning by identifying and evaluating the lessons that have been reported in published assessments of scenarios. We include assessments of the broad scenario literature such

as those found in IPCC reports, as well as assessments of scenarios that were produced as a set, often as part of model comparison exercises. Thus our goal is not to offer a new assessment of the primary scenario literature, drawing our own conclusions about what we can learn from scenarios so far developed. Rather, we assess existing assessments, in order to identify what learning can be documented to have taken place already from evaluating sets of scenarios as a group.

Better understanding what we have learned from scenarios, and what factors may have helped or hindered such learning, can be useful in maximizing learning from future exercises. This is particularly important given that the scenario communities—including climate and earth system modeling, integrated assessment, and impacts, adaptation and vulnerability communities—are about to embark on a new round of scenario development in advance of a Fifth Assessment Report by IPCC, expected to be published in 2013–2014. This new scenario development has been catalyzed by a series of IPCC expert meetings, and a basic framework outlining the process has been established (Moss *et al* 2008). However many details remain to be worked out by the research communities themselves. The remainder of the paper evaluates a number of past scenario development and assessment exercises, identifying conclusions drawn and strengths or weaknesses of the process from the point of view of learning. We then summarize findings and use them to make several recommendations for how future scenario processes might be designed to maximize the potential for learning, including a focus on more specific questions and greater investment in assessing results.

2. Emissions scenarios: assessing the assessments

We briefly examine six assessments or scenario exercises that included some element of assessment of its own results: the SRES, the post-SRES mitigation scenarios developed for the IPCC Third Assessment Report (Morita and Robinson 2001), two model comparisons from the Energy Modeling Forum (Weyant 2004a, de la Chesnaye and Weyant 2006), the chapter on mitigation scenarios included in the IPCC Fourth Assessment Report (Fisher *et al* 2007), and the scenarios developed for the US Climate Change Science Program (CCSP; Clarke *et al* 2007). This survey is not comprehensive. In particular, it does not focus on a few recent comparison exercises that would likely offer additional insights into learning from assessments, including the Low Carbon Society scenarios (Skea and Nishioka 2008, Strachan *et al* 2008) and the Innovation Modeling Comparison Project (Grubb *et al* 2006). Nonetheless, we believe our empirical basis is wide enough to draw an initial set of conclusions.

2.1. SRES (2000)

The SRES scenarios were produced in order to provide the climate change research communities with a common basis for climate change projections, mitigation analyses, and impact assessments. Given that goal, and the nature of the IPCC as an assessment body rather than a generator of new research, the

remit to the writing and modeling teams was to produce a set of scenarios that spanned the range found in the literature, for driving forces as well as emissions. The scenarios were also limited to describing possible futures in which no additional climate change policies were assumed to be adopted. The approach combined four broad narrative storylines with six different models, leading to a total of 40 quantitative scenarios.

Most of the SRES report is devoted to review of the scenario literature and an extensive description of the process of development of the SRES scenarios, and of the scenario outcomes themselves. Deriving lessons from the scenarios was a task largely left to intended users. Nonetheless, some initial descriptive conclusions were drawn in the report itself. Most of these describe individual aspects of driving force and emissions results, but there are three that derive from an assessment of the set of scenarios taken as a whole (see e.g. the SPM and Technical Summary, Box TS-4):

- there is no single central 'best guess' scenario either in the existing literature or in the SRES scenarios;
- technology is at least as important a driving force of GHG emissions as population and economic development across the set of 40 SRES scenarios; and
- scenarios with different driving forces can lead to similar cumulative emissions and those with similar driving forces can branch out into different categories of cumulative emissions.

The first conclusion strengthened a message from the previous set of IPCC scenarios, the IS92 scenarios. Although the IS92 scenarios were presented as a set of six alternative pathways with no single most likely outcome, most users adopted the central IS92a scenario as the 'business as usual' scenario and used it alone. The SRES report went to substantial further lengths to avoid this outcome, associating alternative storylines with each scenario family and providing no emissions path which appeared to be central. The judgment of the SRES authors was that, while some users might believe that one scenario or family was more likely than another, another user might have a differing but valid opinion. The authors therefore described the scenarios as equally sound but offered no judgment of their own as to relative likelihood.

The second conclusion was derived from a kind of model comparison exercise carried out within the overall SRES framework, addressing the question of how sensitive future emissions might be to alternative assumptions about technological development. A set of variants of one scenario family (the A1) was created that reflected different assumptions about whether the direction of technological change in the energy sector was toward or away from fossil fuel-intensive technologies. It showed that this assumption, holding other drivers such as population and economic growth the same, can lead to a range of emissions that is as large as the range across the full set of SRES outcomes. The third conclusion generalizes this result, and also shows the converse: that similar emissions can result from different combinations of driving forces. Among other things, this conclusion had important implications for impact assessments by pointing to

the necessity of considering a wide range of possible socio-economic conditions that could be consistent with a given climate change outcome.

The SRES report therefore did include useful general conclusions drawn from analysis of the scenario set as a whole. However the scope for drawing such conclusions was limited by the fact that the main task of the process was assessment of the scenarios in the literature and production of a set of scenarios that reflected that literature. This restricted scope limited potential contributions in two important areas. First, there was scant opportunity for assessment of the SRES scenarios themselves. Instead, this task was left to future research, which would presumably draw on the SRES scenarios, and to future assessment activities. Second, the scenarios did not aim to address specific research questions. For example, the climate change implications of the scenarios were not addressed, nor were questions about the highest or lowest plausible emissions scenarios in the absence of climate policy. Thus, even though users often assume that the SRES range can be taken to represent the full range of uncertainty in plausible emissions futures (and, by extension, in the climate change outcomes based on SRES that were later produced), that assumption is true only if one assumes that the underlying literature that SRES reflects already fully captured that range, an assumption that has not been examined. Similarly, the sensitivity analysis to technology assumptions was a useful exercise directed at a more specific question, but was not replicated for other driving forces. For example, no analysis of sensitivity to a full uncertainty range in population or economic growth assumptions was carried out. This was likely beyond the time and resources available to the SRES team, but nonetheless represents a lost opportunity for a fuller analysis.

2.2. IPCC Third Assessment Report—the post-SRES scenarios (2001)

The most substantial effort to draw conclusions from the body of emissions scenario literature occurs in the periodic IPCC assessments. However, given the fact that the IPCC cannot do new analysis, possibilities are limited. IPCC chapters assessing emissions scenarios describe the general features of scenarios as a group, and identify broad trends. It is rarely possible to evaluate individual scenarios; at best, a few outliers may be mentioned. A detailed analysis of both the process and content of different scenarios in the literature is not possible both because of the size of the task and the space limitation for any particular topic inherent in a comprehensive assessment like the IPCC reports.

Going beyond broad conclusions regarding groups of scenarios has required special efforts to undertake additional comparison exercises, reminiscent of the technology sensitivity analysis in SRES. For example, the chapter on mitigation scenarios in the Working Group 3 volume of the IPCC Third Assessment Report assessed several hundred scenarios in the literature. Conclusions regarding this large set were limited to descriptions of the range of scenario types, models used, and the gases and mitigation strategies included. To identify more specific lessons, assessment focused on a subset of 31

scenarios that all achieved stabilization at 550 ppm CO_2. One important result was the range of costs of stabilization across scenarios, which was found to be 0%–3.5% of GDP (including a few scenarios with negative costs, or benefits associated with stabilization). While indicative of the range of uncertainty in costs, this result is limited to describing the range of best estimates across models, but leaves out parameter and other uncertainties within any given model, which could substantially widen uncertainty ranges. In addition, it was observed that costs and mitigation strategies were strongly related to the assumed reference scenario, which varied widely across studies, but specific conclusions about the relationship could not be drawn because no systematic studies of this question had been carried out. Such studies can be complicated by the fact that it is not always possible to determine which baseline has been used to develop stabilization and mitigation scenarios.

Partly in order to address this gap in the literature, a process was organized by SRES and TAR authors to generate a set of 'post-SRES mitigation scenarios' that used the SRES scenarios as baselines. It involved nine modeling teams (including the six that participated in SRES) assessing stabilization at (generally) two or more stabilization levels between 450 and 750 ppm CO_2, using two or more SRES baselines. A total of 76 stabilization scenarios were produced. This exercise produced some useful results (Morita *et al* 2000), but again had to limit itself to broad conclusions. For example, it noted that a key manner in which the baseline scenario influences mitigation strategies was through the scale of reductions required. Concentration stabilization requires much larger reductions of CO_2 emissions under development paths with high emissions (such as the A1FI and A2 scenarios) than under development paths resulting in lower emissions (such as B1 and B2). These differences in reduction requirements result not only in different costs of stabilization, but also in selection of different technology and/or policy measures.

In addition, the comparison began to quantify the range in other aspects of mitigation scenarios such as the timing of reductions. It showed, for example, that scenarios stabilizing at 450 ppm CO_2 reached 20% reductions in CO_2 emissions within the next few decades, depending on the model and the baseline scenario, as compared to reaching the same reduction level between 2030 and 2090 for stabilization at 650 ppm or above. It also concluded that taken together, results suggested that stabilization at 450 ppm would require Annex I emissions reductions that go beyond Kyoto commitments by 2020, but that stabilization at 550 ppm would not. Furthermore, early in the century emissions reductions tended to come predominantly from reductions in energy intensity, while later in the century they came predominantly from decarbonizing energy supply.

These conclusions were important not because they were qualitatively new, but because they quantified the general features of a large set of results from a wide range of models and scenarios applied to particular cases (the stabilization levels). For example, the general tendency for reductions to occur earlier for lower stabilization levels was well known, but the assessment identified a specific timeframe. Similarly, the

comparison of reductions achieved in these scenarios to those called for by the Kyoto Protocol was suggestive of the degree of additional action that might be necessary.

These positive contributions notwithstanding, the force of the conclusions was limited by the fact that the comparison exercise was not directed specifically at these questions. For example, while the range of times at which 20% reductions in emissions were achieved could be identified for each stabilization level, one could not conclude that this timing was actually necessary, since alternatives in which timing was varied were not tested. Similarly, the level of Annex I emissions reductions made in the post-Kyoto period could be reported, but one could not conclude that these were necessary, or even optimal, since they were affected by differing assumptions across models about how much emissions would be reduced in non-Annex I countries. Thus, the post-SRES exercise did contribute to learning from emissions scenarios, but it fell short of what could be learned from a more focused activity with greater financial resources and more time. Such an activity would also offer an opportunity to develop a process for assessing a range of different scenario approaches and outcomes.

2.3. EMF-19 and -21 (2004, 2006)

The Energy Modeling Forum (EMF) has carried out semi-regular model comparison exercises for more than 25 years. Each study typically consists of a set of scenarios that participating modeling teams agree to produce (although not all teams necessarily produce all scenarios). Results are typically described in a collection consisting of individual papers reporting results from each modeling team along with a brief overview paper that examines results as a whole. This overview is not designed to be an in-depth examination of the collective results, but rather to highlight a few issues and key results to give the reader a feel for the study as a whole.

For example, the EMF-21 study of multi-gas mitigation involved 19 models running a number of scenarios aimed at the broad issue of the role of non-CO_2 greenhouse gases and sinks in climate policy (de la Chesnaye and Weyant 2006). The summary paper (Weyant *et al* 2006) points out main features of the results, including the fact that all models show that including non-CO_2 gases in a stabilization scenario reduces the need for true CO_2 reductions and reduces the total cost of mitigation. Although qualitatively this result was already known, the model comparison gave a quantitative sense of how much costs were reduced—15–70% reduction in carbon price in 2025, and 0–56% in 2100, for stabilizing radiative forcing at 4.5 W m^{-2} relative to a wide range of 'modeler's choice' baseline scenarios. The overview also points out that models using global warming potentials (GWPs), a physically-based index used to equate the emissions of different gases, produce more methane reductions early in the century compared to models that avoid the use of GWPs by calculating radiative forcing directly and producing a least-cost mix of emissions reductions across gases.

The EMF-19 study included 14 models that focused on 'Technology and Global Climate Change Policy' (Weyant

2004a). Its main goal was to better understand 'how models being used for global climate change policy analyses represent current and potential future energy technologies, and technological change.' Models ran reference scenarios (including a standardized SRES B1 scenario), a 550 CO_2 stabilization scenario, and scenarios in which carbon taxes reach 100 dollars per ton of carbon ($/tC). Although the standardized B2 scenario provided a means of controlling for the effects of the baseline scenario, the overview paper (Weyant 2004b) focused on the non-standardized reference scenarios (modeler's choice) and the 550 stabilization scenarios, and on results at the global level. It emphasized the importance of the reference scenario assumptions in determining the nature of the stabilization scenario results, driving for example large differences in the required carbon tax to reduce emissions, which in turn influence technological choices, a finding which echoed the TAR conclusions. It drew several other broad conclusions as well: that stabilizing CO_2 concentrations will require major transformations of energy technologies, that this transformation will take many decades, and that its costs, while large, can be lessened by starting the process sooner and pursuing many options in parallel.

Here again, results and broad conclusions from both EMF exercises are certainly useful, and strengthened by the fact that they are supported not by one model but by many. Yet the potential for learning from such exercises clearly exceeds these conclusions. Results could be used to identify important dependencies, for example regarding the costs and time required to develop and implement new technologies. What drives variation in results across models? To what extent are these variations driven by modeling issues (e.g., technologies included in some models but not others) versus reflections of real world uncertainties (such as evolution of costs over time)?

2.4. IPCC Fourth Assessment Report (2007)

As in the TAR, Working Group 3 of the Fourth Assessment Report included a chapter on mitigation scenarios (Fisher *et al* 2007). The content of this chapter might be broadly categorized as an assessment of ranges of inputs and outputs in the scenario literature, quantification of some key results such as costs, and identification of key sensitivities. For example, this chapter concluded that the range of emissions and their drivers, with a few exceptions, has not changed much since SRES was published in 2000. It also pointed out trends in the literature—toward more multi-gas scenarios, a small but growing number that include global land use, and more that incorporate endogenous technological change. As was the case in the EMF exercises, qualitative conclusions on scenario implications were not surprising—lower stabilization targets imply earlier mitigation and higher costs, the importance of technological development for costs, and the importance of multi-gas mitigation for reducing costs and providing more degrees of freedom in when and where to reduce GHG emissions. The value added was the quantification of the range of results available in the literature. In particular, the assessment provided ranges of emissions trajectories for

different stabilization targets. For example, the chapter concluded that the lowest stabilization scenarios (below 490 ppm CO_2eq) have global emissions peaks before 2015, while for the highest levels this turning point occurs around 2040. Regarding costs, it concluded for example that stabilizing at 3.5–4 W m^{-2} will have costs ranging from a benefit of 1% of GDP to a loss of 2% in 2050. The chapter also identified what costs are sensitive to—the baseline scenario, stabilization target, technologies assumed to be available, rate of technological change, inclusion of multi-gas or land use reduction options.

As was the case in the TAR, these descriptive findings were useful in characterizing existing scenarios, but cannot provide the same information as model comparison exercises aimed at answering specific questions. Global emissions tend to peak around 2015 to reach the lowest stabilization scenarios, but must they do so? If they peak in 2025, is stabilization no longer feasible? Or does this rather imply substantially 'negative' emissions in the second half of the century? Regarding costs, is the range of estimates in the literature a full accounting for uncertainty, or a range of best estimates that excludes the tails of the distribution, thus underestimating uncertainty?

2.5. US CCSP scenarios (2007)

The US Climate Change Science Program is in the midst of producing 21 'Synthesis and Assessment Products' (SAPs) aimed at evaluating climate change science to help inform policy and prioritize research. Product 2.1, released in 2007, focuses on scenarios and is split into two parts: a first report (2.1a; Clarke *et al* 2007) that develops and assesses a small set of emissions and concentration scenarios, and a second (2.1b; Parson *et al* 2007) that assesses the development and use of global change scenarios more broadly. We focus on SAP 2.1a, since it is a coordinated scenario development exercise and assessment in the style of previous EMF and IPCC exercises (see Parson 2008, for a discussion that draws on 2.1b).

The CCSP scenarios were developed by three modeling groups, each of which produced a reference scenario (with no climate policy) of its choice, and in addition four policy scenarios in which concentrations were stabilized at a range of levels in the long term. The stated aims of the exercise were to address broad questions similar to those in previous EMF exercises and IPCC assessments: what emissions paths are consistent with various stabilization goals, and what might be the costs and energy system characteristics required to achieve them? There is a particular emphasis on implications for the US, although global results are provided as well.

As in previous exercises, there is a detailed description of the outcomes of the various scenarios across the three models, which are interesting in and of themselves. For example, outcomes regarding the relative growth of Annex I and non-Annex I emissions under various conditions are certainly germane to current policy debates, and descriptions of energy system characteristics such as the percentage of energy from zero-carbon energy sources over time provide important insights into the possible scale of changes needed to achieve

various long-term goals. Results also showed that emissions reductions tend to occur first in the electric power system, and later in the transport, industry, and buildings sectors, reflecting the assumed relative costs of reductions across these sectors. In addition, it was found that carbon prices varied substantially across different models, an outcome that was attributed to alternative assumptions about baseline scenarios and about future technology availability and costs.

However, the scope for drawing conclusions based on the CCSP scenario results as a set was relatively limited, as in previous comparisons. The general conclusions reached echoed those identified previously: that energy use and emissions grow in the absence of policy; that in order to stabilize concentrations, emissions must peak and then decline (with timings similar to those found in the literature); that stabilization will eventually require a transformation of the energy system; and that the nature of the assumed baseline scenario strongly affects outcomes. Stronger or more specific conclusions were difficult to draw. The descriptive findings are indicative of possible outcomes, but say little about the robustness of the conclusions. If a given scenario indicates a certain percentage of zero-carbon energy by 2050, for example, it is not clear how to interpret the result. How much different could that percentage be and still achieve the same concentration outcomes, and at what cost?

As in earlier efforts, the intention seems to have been to develop scenarios whose real value would be to serve as a basis for later applications or as input to later assessments. To this end the report prominently lists further analysis that would be useful, including simulation of climate change consequences and impacts following from these emissions scenarios, more detailed analysis of the technological and economic implications of mitigation, or assessment in the context of other scenarios in the literature. It also indicates that the key value added of these scenarios is the use of updated economic and technological data, improved models, and a multi-gas approach, particularly as compared to models used in the SRES and TAR assessments.

3. Conclusions and recommendations

Based on the assessment of previous model comparison exercises and scenario assessments, we find that coordinated scenario development exercises often were not designed with the explicit purpose in mind of learning from evaluation of the outcomes as a set and have typically left most of the assessment of their results to future efforts. However, those future efforts, including most prominently the IPCC assessment reports, have had insufficient capacity to carry out detailed and thorough analyses, leaving much of the potential for learning from these scenarios untapped.

One might be tempted to conclude that, given the limited insight that has been derived from evaluating sets of scenarios as a group, resources would be better invested elsewhere and it would be advisable to reduce the number of such exercises that are carried out. We do not believe this is the case. There are many other benefits that flow from such activities, and the lessons learned from evaluating results as a set are just one of

them. Thus the total return on investment is high. However it could be higher, and a particularly valuable additional benefit obtained, if assessment of results were improved.

In principle, assessments could be advanced either by improving the assessment activities that are built-in to scenario development processes, or by improving assessments that take place after the fact. Here we make two specific recommendations regarding the former, and one regarding the latter. In either case, successful improvements will depend on having a good understanding of the factors that have hindered learning from scenario exercises so far. We have already suggested that in many cases limited time and resources are a contributing factor, but we offer additional possibilities here as well.

(1) Future scenario development exercises should invest much more time and effort in assessment of the set of results as a whole, at the time they are produced

The SRES, IPCC assessment chapters, and EMF and CCSP exercises have made important contributions to our understanding of future emissions and mitigation options. In addition, they have served as an impetus for generating a large number of scenarios in the literature, stimulated model development, and led to improvements in methodology within the scenario community. Nonetheless, much greater gains could be had from such activities by investing more time and effort in the assessment of results across modeling groups at the time the scenarios are generated. Specific authors could be designated from the outset to carry out detailed comparisons of scenario results in a manner designed to draw tangible conclusions regarding the question being addressed.

There are likely many reasons such an approach has so far been the exception rather than the rule, all of which represent obstacles to improvement. Ultimately, understanding what these obstacles are and how one might overcome them is an important research question that could be investigated through a range of social scientific approaches (Garb *et al* 2008). We offer a few possibilities that might serve as starting points. Most broadly, in many cases the goals of scenario generation and comparison exercises are not well aligned with drawing specific conclusions regarding a particular policy question. Rather, the focus is often on improvements in technical aspects of modeling, data and assumptions, something that is exceedingly important but typically mostly of interest to participating modeling groups only. Thus, it may be that little effort is expended on thoroughly assessing the full set of model results in order to draw conclusions because that is simply not a goal of the exercise in the first place. If this is the case, then improvements would require realigning the goals of such exercises to more directly address specific questions (a recommendation we make below).

The mix of incentives and disincentives for participating groups may also serve as an obstacle to better assessment of results. Modeling groups typically participate in comparison exercises voluntarily, and participating involves a major investment of time and resources. Thus to be viable these exercises need to provide many carrots and few sticks. The biggest carrots are the chance to publish results in a high profile outlet, and the chance to stay at the forefront of research

and model development by interacting closely with others in the field. A thorough comparison and assessment of results in the broader scenario contexts that go beyond technical aspects of different models might be seen as a potential stick, 'punishing' a group whose model was not perceived to perform as well as another. This obstacle might be overcome by structuring the assessment component in such a way that it minimized the potential for particular groups to be painted in an unfavorable light. Or, the potential disincentive might be offset by increasing the positive incentives, for example by providing funding or an even higher profile for published results.

Finally, cultural differences across disciplines may also serve as a barrier to better assessment within scenario exercises. It may that those individuals who would be best equipped to carry out an assessment have different academic backgrounds, peer groups, and views on modeling than do the individuals participating in the production of the scenarios. These differences could generate a potential for conflict if, for example, the assessors are viewed as not being 'of' the same community as the modelers and may not have the best interests of participants in mind. Avoiding this type of conflict might require careful choice of assessment teams, perhaps making sure that they contained members of modeling teams as well as those with broader purviews. Clear terms of reference can also help avoid fears that the assessment would focus on studying modeling teams and their scenario development approaches rather than helping to improve the process of scenario development and assessment.

(2) Scenario exercises should include a focus on more specific questions and communities

In our review of assessments we found that most exercises have been rather loosely focused on broad questions regarding possible characteristics of reference and mitigation scenarios, and less often focused on specific research or policy questions, or on sensitivity of outcomes to specific assumptions, model differences, or differences in scenario development processes. Scenario exercises would lead to more informative conclusions if they aimed from the outset to address more specific questions. As we described above, previous exercises have generated more informative conclusions when they focused on more constrained questions. The sensitivity analysis to technological assumptions within SRES, and the post-SRES mitigation scenarios, both go some way toward this goal. However, many more specific questions would make good candidates for future coordinated, multi-model analyses.

- What are the implications of new information? Most scenario analyses have been done within the context of a single 'story' over time—SRES storylines, for example, or scenarios in which the long-term stabilization goal is known from the beginning. What are the implications of learning—whether about climate risks or mitigation technologies—that induces a sudden change in mitigation strategy or long-term climate goals? How important is flexibility in response strategies, and what are the limits of responding to new information?
- What are the risks and opportunities of specific energy resources such as biofuels, methane hydrates, or bioengineered biomass?

- What conditions create 'barely feasible' scenarios? E.g., what is the lowest emissions or concentration outcome that appears feasible, defined by pushing the envelope of current knowledge? How does this lowest, barely feasible outcome change if mitigation responses are delayed?
- What is the full range of uncertainty in emissions over the next few decades? In particular, how rapidly might emissions grow in the absence of policy? Most global emissions scenarios have focused on 50–100 year timescales, with uncertainty assessments primarily relevant to the longer term as well (e.g., Webster *et al* 2002, Gritsevskyi and Nakicenovic 2000, Nakicenovic *et al* 2000). Shorter-term scenarios typically have not included a careful assessment of uncertainty, despite the fact that trends can change substantially over this time period (Sheehan 2008) and can have important consequences, including the possibility of making some long-term climate change goals infeasible (O'Neill and Oppenheimer 2002).

This recommendation complements those made in the US CCSP scenario assessment report (Parson *et al* 2007; see also Parson 2008). There, it was recommended that, because the diversity of users and their needs is so large, many additional scenario-related activities are needed to supplement 'core' global change scenarios that provide broad descriptions of future emissions and climate change outcomes. Additional activities would include interpretations of results at the regional level, downscaling exercises, and development of additional consistent scenarios at the national or sub-national level. We agree with this recommendation. Our point here is slightly different: that the core scenarios themselves should be expanded in number to include more targeted, less all-purpose scenarios. In other words, in addition to developing the capacity to adapt existing types of core scenarios to specific needs, we should also be producing different types of global emissions and climate change scenarios that are better suited to addressing specific questions and particular subsets of needs.

There may well be a trade-off between the specificity of the question addressed, and the shelf-life of a scenario. Today's priority questions will change, particularly as the climate debate turns increasingly toward response options. To be most relevant, the scenario community will need to interact more closely with user communities and develop scenarios, and conclusions based on them, on a more flexible timescale. It might seem that such an approach would be at odds with the enterprise of constructing long-term global scenarios. But because scenarios may extend 100 years into the future does not mean that they have to be relevant for 100 years. Scenarios can usefully inform our decisions today even if they are overtaken by events tomorrow. In other words, we recommend that a process be developed whereby the art and craft of developing scenarios can evolve along with advances in our basic understanding of the underlying processes. In this way, scenarios would be continuously updated and from time to time fundamentally overhauled depending on improvements in our understanding, scenario development advances, and needs of the salient 'user' communities.

*(3) Undertake more assessments outside of scenario develop-
ment or comparison exercises*

The periodic IPCC assessments cannot be relied on to
provide all of the assessment of scenario results; the job is
simply too big. They are limited to already published peer-
reviewed studies and have restricted space for assessment. In
addition, such assessments often take too long so that the
emissions scenarios on which they are based are rendered
obsolete, or it is not feasible to truly integrate impacts,
vulnerabilities and adaption needs after the mitigation efforts
are accounted for in the emissions scenarios. Ideally, IPCC
assessments should be able to synthesize results not only from
individual modeling studies, but also from a greater number
of assessments of model comparison results than currently
exists. While in many cases it would be preferable to produce
these assessments in conjunction with the exercises producing
the scenarios themselves, in some cases this will not be
possible. The obstacles we discuss above (or others we have
not identified) may not be surmountable, time or resources
may prohibit such assessment, or assessment of the full set
of results may simply be a low priority. Thus it would
also be valuable to increase the number of assessments that
take place after the fact, separate from any given scenario
development exercise. For example, Parson *et al* (2007)
recommend that the US CCSP support an increase in the
capacity not only to develop but also to assess scenarios. In
general, establishing fora in which regular scenario comparison
(and possibly development) can be carried out would be
valuable and would improve understanding of which scenario
differences can be attributed to different assumptions, which
to different modeling approaches, and which to differences
in the underlying processes of scenario development and
formulation.

Such assessments after the fact are difficult to carry out.
It can be difficult or impossible to obtain the necessary model
input assumptions and outputs in comparable formats to carry
out detailed comparisons. It can also be difficult to interact
with members of the different modeling groups, who will have
moved on to other projects and priorities, in order to understand
model differences and interpret results of the comparisons.
At a minimum, to facilitate later assessment, better databases
of scenario results and comprehensive scenarios inventories
should be a high priority during scenario development and
comparison exercises. In addition, the makeup of author
teams carrying out such assessments should include sufficient
representation of modeling teams to facilitate the comparisons.

All three categories of recommendations we make here
are particularly relevant to a key aspect of the next phase of
emissions scenario development: analysis of 'Representative
Concentration Pathways,' or RCPs. The RCPs are four
pathways of atmospheric concentrations and emissions of
greenhouse gases, reactive gases, and aerosols that were
identified through an IPCC expert meeting in order to facilitate
a new parallel process of scenario development (Moss *et al*
2008). Climate modeling groups will use the RCPs as input
to generate new simulations of future climate changes, while
emissions scenario modelers are to develop a range of socio-
economic and technological scenarios that would be consistent

with the RCPs. Both sets of outcomes would then be used
in assessments of impacts, adaptation, and vulnerability. The
RCPs clearly have a useful role to play, but past experience
suggests three important aspects of the scenario development
process that would lead to more useful outcomes. First, any
coordinated exercises organized as part of this process should
plan from the beginning to devote substantial time and effort
to the analysis of scenario results as a set, and not assume that
future IPCC (or other) assessments will have sufficient capacity
to do the job on its own.

Second, the development of emissions scenarios based
on RCPs should be done in a way that allows for more
useful comparisons to be made across models or methods
than has been done in the past. A large number of widely
differing models, scenarios, methods, and assumptions will
only lead to broad conclusions unlikely to be substantially
more informative than those drawn from past sets of multi-gas
scenarios. And third, emissions scenario development should
not focus overly much on RCPs at the expense of addressing
more specific questions. There are a wide range of important
questions that can be addressed through scenarios that are not
well served by casting them within the RCP framework (or,
in fact, that cannot be addressed within this framework at all).
If most of the emissions scenario community spends most of
its time on RCPs, an important opportunity for generating new
insights will be lost.

Finally, we recommend that the research community
improve its understanding of how scenarios are used. Our
analysis and recommendations have been focused on the
production of scenarios, and how best to learn from them. How
global emissions scenarios are used within the scientific field
is reasonably well known, for example in providing a common
framework for coordinating studies across climate change,
mitigation, and impacts and adaptation research communities.
However, we know woefully little about how they are used
outside of it. Improved awareness of scenario use will be an
essential component of choosing the most relevant questions
to address, and designing the best process to learn from the
scenarios that are produced.

References

Carpenter S R, Pingali P L, Bennett E M and Zurek M B (ed) 2005
    Ecosystems and human well-being: scenarios *Millennium
    Ecosystem Assessment* vol 2 (Washington, DC: Island Press)
Clarke L, Edmonds J, Jacoby H, Pitcher H, Reilly J and
    Richels R 2007 Scenarios of greenhouse gas emissions and
    atmospheric concentrations *Sub-report 2.1A of Synthesis and
    Assessment Product 2.1 by the U.S. Climate Change Science
    Program and the Subcommittee on Global Change Research*
    (Washington, DC: Department of Energy, Office of Biological
    and Environmental Research) p 154
de la Chesnaye F and Weyant J (ed) 2006 Multi-greenhouse gas
    mitigation and climate policy *Energy J.* **27** (Special issue)
Fisher B S *et al* 2007 Issues related to mitigation in the long term
    context *Climate Change 2007: Mitigation. Contribution of
    Working Group III to the Fourth Assessment Report of the
    Inter-governmental Panel on Climate Change* ed B Metz,
    O R Davidson, P R Bosch, R Dave and L A Meyer (Cambridge:
    Cambridge University Press)

Garb Y, Pulver S and VanDeveer S 2008 Scenarios in society, society in scenarios: toward a social scientific analysis of storyline-driven environmental modelling *Environ. Res. Lett.* **3** 045015

Gritsevskyi A and Nakicenovic N 2000 Modeling uncertainty of induced technological change *Energy Policy* **28** 907–21

Grubb M, Carraro C and Schellnhuber J 2006 Technological change for atmospheric stabilization: introductory overview to the innovation modeling comparison project *Energy J.* **27** 1–16 Special edition: endogenous technological change and the economics of atmospheric stabilisation

Houghton J T, Jenkins G J and Ephraums J J (ed) 1990 *Climate Change: The IPCC Scientific Assessment* (Cambridge: Cambridge University Press) p 365

Hulme M and Dessai S 2008a Negotiating future climates: a critical review of the development of climate scenarios for the UK *Environ. Sci. Policy* **11** 54–70

Hulme M and Dessai S 2008b Predicting, deciding, learning: can one evaluate the 'success' of national climate scenarios? *Environ. Res. Lett.* **3** 045013

Leggett J, Pepper W J and Swart R J 1992 Emissions scenarios for IPCC: an update *Climate Change 1992. The Supplementary Report to the IPCC Scientific Assessment* ed J T Houghton, B A Callander and S K Varney (Cambridge: Cambridge University Press)

Morita T and Robinson J 2001 Greenhouse gas emission mitigation scenarios and implications *Climate Change 2001: Mitigation* ed B Metz, O Davidson, R Swart and J Pan (Cambridge: Cambridge University Press)

Morita T, Nakicenovic N and Robinson J 2000 Overview of mitigation scenarios for global climate stabilization based on new IPCC emission scenarios (SRES) *Environ. Econ. Policy Stud.* **3** (2) 65–88

Moss R *et al* 2008 *Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies* (Geneva: Intergovernmental Panel on Climate Change) p 132

Nakicenovic N *et al* 2000 *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press)

O'Neill B C and Oppenheimer M 2002 Dangerous climate impacts and the Kyoto Protocol *Science* **296** 1971–2

O'Neill B C, Pulver S, VanDeveer S and Garb Y 2008 Where next with global environmental scenarios? An introduction and overview *Environ. Res. Lett.* **3** 045012

Parson E 2008 Useful global-change scenarios: current issues and challenges *Environ. Res. Lett.* **3** 045016

Parson E, Burkett V, Fisher-Vanden K, Keith D, Mearns L, Pitcher H, Rosenzweig C and Webster M 2007 Global change scenarios: their development and use *Sub-Report 2.1B of Synthesis and Assessment Product 2.1 by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research* (Washington, DC: Department of Energy, Office of Biological and Environmental Research) p 106

Raskin P, Gallopı'n G, Gutman P, Hammond A and Swart R 1998 *Bending the Curve: Toward Global Sustainability* (Stockholm: Stockholm Environment Institute) Available at http://www.tellus.org/seib/publications/bendingthecurve.pdf

Schwartz P 1991 *The Art of the Long View: Planning for the Future in An Uncertain World* (New York: Doubleday)

Sheehan P 2008 The new global growth path: implications for climate change analysis and policy *Clim. Change* **91** 211–31

Skea J and Nishioka S 2008 Policies and practices for a low-carbon society *Clim. Policy* **8** S5–16

Strachan N, Foxon T and Fujino J 2008 Policy implications from the low-carbon society (LCS) modelling project *Clim. Policy* **8** S17–29

Wack P 1985a Scenarios: uncharted waters ahead *Harv. Bus. Rev.* **63** 72–89

Wack P 1985b Scenarios: shooting the rapids *Harv. Bus. Rev.* **63** 139–50

Webster M D, Babiker M, Mayer M, Reilly J M, Harnisch J, Sarofim M C and Wang C 2002 Uncertainty in emissions projections for climate models *Atmos. Environ.* **36** 3659–70

Weyant J P (ed) 2004a EMF 19 alternative technology strategies for climate change policy *Energy Econ.* **26** 501–755

Weyant J P 2004b Introduction and overview *Energy Econ.* **26** 501–15

Weyant J P, de la Chesnaye F and Blanford G 2006 Overview of EMF–21: multigas mitigation and climate policy *Energy J.* **27** 1–32 (Multi-Greenhouse Gas Mitigation and Climate Policy) (Special issue No. 3)

Wilkinson A and Eidinow E 2008 Comparing scenarios for global environmental assessment with scenarios in organizational settings *Environ. Res. Lett.* **3** 045017