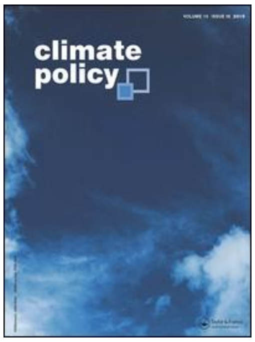

Climate Policy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tcpo20

Deepening transparency about value-laden assumptions in energy and environmental modelling: improving best practices for both modellers and non-modellers

John Bistline , Mark Budolfson & Blake Francis

To cite this article: John Bistline , Mark Budolfson & Blake Francis (2020): Deepening transparency about value-laden assumptions in energy and environmental modelling: improving best practices for both modellers and non-modellers, Climate Policy, DOI: 10.1080/14693062.2020.1781048

To link to this article: https://doi.org/10.1080/14693062.2020.1781048

Published online: 26 Jun 2020.

CLIMATE POLICY
https://doi.org/10.1080/14693062.2020.1781048

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Check for updates

Deepening transparency about value-laden assumptions in energy and environmental modelling: improving best practices for both modellers and non-modellers

John Bistline [a], Mark Budolfson [b] and Blake Francis [c]

[a]Electric Power Research Institute, Palo Alto, CA, USA; [b]Environmental and Occupational Health, Population-Level Bioethics, and Philosophy, Rutgers University, New Brunswick, NJ, USA; [c]Department of Philosophy, University of Maryland, Baltimore County, Baltimore, MD, USA

ABSTRACT

Transparency and openness are broadly endorsed in energy and environmental modelling and analysis, but too little attention is given to the transparency of value-laden assumptions. Current practices for transparency focus on making model source code and data available, documenting key equations and parameter values, and ensuring replicability of results. We argue that, even when followed, these guidelines are insufficient for achieving deep transparency, in the sense that results often remain driven by implicit value-laden assumptions that are opaque to other modellers and researchers, and that may not be understood by wider audiences to be controversial. This paper identifies additional best practices for achieving transparency by highlighting issues where disagreement over value judgements will persist for the foreseeable future. Recommendations for deepening transparency are developed by learning from successes and ongoing challenges represented in three case studies. We provide recommendations to accelerate the adoption of additional best practices for deepening transparency of energy and environmental modelling in policy-relevant domains, increasing stakeholder participation with non-modellers, and encouraging interdisciplinary dialogue.

Key policy insights

- Achieving all of the goals associated with transparency requires more than current practices of providing open source data, code, and model documentation.
- Greater interdisciplinary dialogue could improve transparency beyond current practices, including in model development, application, and communications.
- Better practices for addressing contentious and value-laden assumptions include providing accessible documentation for non-specialists, increasing policymaker participation to ensure that model outputs can inform questions, and performing sensitivity analyses that cover a range of reasonable views about value-laden assumptions.
- Energy and environmental modellers should account for audience-specific considerations to promote transparency, especially accounting for needs of non-modellers such as policymakers.

ARTICLE HISTORY
Received 28 January 2020
Accepted 4 June 2020

KEYWORDS
Transparency; modelling; economic analysis; uncertainty; values; energy and environmental policy

1. Introduction

Important societal issues increasingly require interdisciplinary collaborations and are being informed by quantitative modelling. For instance, with growing pressure to mitigate climate change while achieving the broader UN Sustainable Development Goals (SDGs), many climate and energy modelling groups are making their data and equations public (Hurrell et al., 2013; Pfenninger et al., 2018). While this added transparency can improve the scientific process as well as public trust in science, this paper argues that deepening

CONTACT John Bistline jbistline@epri.com Electric Power Research Institute, 3420 Hillview Avenue, Palo Alto, CA 94304, USA

transparency – beyond making data, equations, and model documentation merely openly available – is needed to address value-laden assumptions that may not be apparent to non-modellers. It presents three case studies to both illustrate best practices and identify opportunities for confronting value-laden assumptions in models. For our purposes, value-laden modelling assumptions involve taking concrete answers to contentious ethical questions for granted. Ethical questions could include whether and how to discount future goods, how to make decisions under uncertainty, and how to account for inequality, as discussed in Section 2.

Transparency is widely endorsed as a goal in energy and environmental modelling (Pfenninger et al., 2018; Schneider, 1997), as well as in interdisciplinary research across the natural and social sciences. Current practices for transparency focus on making model source code open access, documenting key equations and parameter values, providing data, and ensuring replicability of results (DeCarolis et al., 2012, 2017). Several recent papers recommend how quantitative researchers in energy, sustainability, and other scientific fields can make data, equations, and major assumptions transparent to other researchers (e.g. Bistline & Merrick, 2020; Craig et al., 2002; DeCarolis et al., 2012, 2017; Ha-Duong, 2001; Ince et al., 2012) as well as to policymakers and the public (e.g. Pfenninger et al., 2018).[1] Open-source software and crowdsourcing are more familiar within ecology (Benz et al., 2001; Grimm et al., 2006) and genetics, but are becoming increasingly common in climate science (Hurrell et al., 2013), energy (DeCarolis et al., 2012), and social sciences (Christensen & Miguel, 2018; Gertler et al., 2018). While several studies stress that these recommendations are but a first step towards improving transparency, the energy and environmental modelling literature tends to focus on transparency for other modellers (DeCarolis et al., 2012; Ince et al., 2012; Peng et al., 2014). Additionally, the U.S. Environmental Protection Agency's proposed rule 'Strengthening Transparency in Regulatory Science' aims to 'ensure that the data underlying [scientific studies] are publicly available in a manner sufficient for independent validation' (U.S. EPA, 2018). However, the proposed rule focuses on the availability of data and source code, which may be insufficient for making value judgments explicit.[2]

Indeed, several key contributions to best practices in energy modelling rarely, if ever, mention value-laden judgments in recommending improved transparency in energy and economic modelling (DeCarolis et al., 2012, 2017; Pfenninger et al., 2018; Strachan et al., 2016). Although these first steps toward improving transparency can be informative for modellers and help build social confidence, we argue that they are insufficient for achieving what might be called 'deep transparency', which requires making structural assumptions explicit, creating opportunities for interdisciplinary engagement, and explicitly communicating value-laden assumptions to stakeholders. Because value-laden assumptions are not merely hidden assumptions that can be made explicit by providing model source code and documentation, as DeCarolis et al. (2012, p. 1847) suggest, we argue for deepening transparency. We use 'deepening transparency' to signify improvements along multiple dimensions that are relevant not only to modellers but to broader stakeholders who use model results to make policy decisions, run businesses, allocate resources, or communicate risk.

Consumers of model outputs are increasingly audiences such as policymakers, industry, researchers in other disciplines, and foundations who may not necessarily have the background or resources to explore source code or run models (Rose & Scott, 2018). Yet, these audiences have a stake in understanding the structural and value-laden assumptions that inform model results. This limitation underscores the importance of deepening transparency strategies such as increasing stakeholder participation (Reed, 2008) and comprehensive sensitivity analysis with large ensembles of scenarios across a range of uncertain parameters (Morgan & Henrion, 1990). Recommendations to deepen transparency are important not just because they provide insights about 'black box' models (and how model outputs change as assumptions change). Deepening transparency also reminds audiences that the model building and analysis process is a series of choices that embody values, malleable parameters, and uncertain system dynamics.

This paper proceeds as follows. Section 2 describes case studies for deepening transparency about value-laden assumptions related to discounting, uncertainty, and distributional impacts. Section 3 builds on insights from the case studies to provide recommendations to accelerate transparency best practices. Section 4 concludes by discussing policy implications.

2. Case studies of deepening transparency for value-laden assumptions

The literatures on energy and environmental policy analysis provide many examples of opaque modelling assumptions with consequential impacts. We draw on work on transparency in energy and environmental modelling that predates the now influential open-source movement (Funtowicz & Ravetz, 1993; Oreskes et al., 1994; Proctor & Proctor, 1991; Schneider, 1997). Recommendations for deepening transparency are developed by identifying successes and opportunities for improvement in contentious areas of debate where opacity remains even when guidance for code availability and documentation are followed. This section examines three case studies that provide illuminating examples of how opaque and contentious assumptions can be key drivers of outputs but may be unknown to many stakeholders: discounting, uncertainty, and socioeconomic distribution of impacts.

These case studies highlight how best practices for analysts could differ from strategies for other stakeholders, which motivates recommendations in Section 3. They also illustrate opportunities for progress and engagement between disciplines. Each is in a different state in its development of modelling standards, providing useful points for comparison and contrast. The first case – discounting and intergenerational equity – is a longstanding debate where multidisciplinary interchanges and practices, like sensitivity analysis, have emerged that are worthy of emulation. The second example – representing uncertainty – has been recognized as influential but the modelling community still lacks robust guidelines to support analysis, despite a growing literature offering practical guidance (e.g. Yue et al., 2018). The final example – the socioeconomic distribution of impacts of climate damages – is an emerging issue where awareness is limited and opportunities exist to accelerate dialogue.

A common trait of the three cases is that disagreements about values and appropriate model representations will persist for the foreseeable future. The anticipated lack of convergence is a feature of these debates and makes the goals of transparency to general audiences (not only modellers) and availability of sensitivity analysis imperative. An iterative approach to model development and application acknowledges that insights are provisional, learning can provide openings for model refinement, and difficult-to-quantify dimensions of decision problems should not be ignored. A second shared feature is that basic transparency (exemplified by open access code and equation documentation) would fall short of the ultimate goal of deepening transparency, including identification of important assumptions that drive recommendations, accessibility to interested decisionmakers and stakeholder participation. In this way, these issues provide examples of how progress has been made, and what is needed for deepening transparency beyond open-source guidelines.

2.1. Discounting: weighing costs and benefits over time and across generations

Evaluating policies with long-term effects requires comparing the immediate benefits and costs (relative to an appropriate baseline) with those in the future. For benefit–cost analysis, these temporal tradeoffs are formalized through the discount rate, which allows payoffs or impacts occurring at different times to be compared. Debates over discounting in public policy have been contentious, especially in environmental economics where issues like climate change have long time horizons that may span centuries (Goulder & Stavins, 2002; Nordhaus, 2007; Stern & Taylor, 2007). A key question is whether the discount rate should be based on the rate of return on private investments or derived from ethical principles about the relative weights of well-being at different times (Arrow et al., 1996; Kelleher, 2017; Mintz-Woo, 2018). In energy models, discount rates often include both ethical and investment-related discount rates or aggregate many effects like time preference, opportunity cost of capital, risk, and/or financing, and differ by decisionmaker (Bistline et al., 2019). Over several decades, debates about discounting have drawn contenders across academic disciplines from philosophy, to the natural sciences, to economics (e.g. Arrow et al., 2014; Broome, 1994; Lind et al., 1982; Portney & Weyant, 1999). Given the large extent to which discounting parameters and selection principles affect outcomes and policy prescriptions, transparency could not be more important.

The discounting debate offers lessons about the benefits of interdisciplinary exchange and transparency in modelling decisions. This debate is a success story, but not because it yielded consensus about context-dependent discounting practices. Rather, it has been successful because discounting has been extensively investigated for many decades across disciplines. Progress has been made in clarifying implicit assumptions, conducting sensitivity analyses across a range of reasonable alternatives, and communicating implications of alternatives to a range of stakeholders. These practices have the potential to increase the political legitimacy of decisions made on the basis of these models (Fleurbaey et al., 2018; Wilkerson et al., 2013).

Highlighting lessons learned in the discounting debate can guide the move toward transparency and interdisciplinarity for other issues (Klein, 1990). For instance, consider the workshop held by the U.S. Environmental Protection Agency, Department of Energy, Resources for the Future, Stanford University, and the Electric Power Research Institute which led to the 1999 book edited by Portney and Weyant entitled '*Discounting and Intergenerational Equity*' (Portney & Weyant, 1999). The workshop brought together leading experts on discounting from different disciplines, and the book also offers one form of lasting documentation defining the contours of the debate. Exchanges in this context have helped to uncover embedded assumptions that, while valid in other settings, may be problematic for long-term policy issues with the complexity of climate change (e.g. neglecting overlapping agents, problems of time inconsistency, limitations of a unitary decisionmaker with a well-defined objective function).

A related example is the 2017 U.S. National Academies of Sciences report on models that estimate the social cost of carbon (SC-CO_2), such as those used by the U.S. Interagency Working Group (IWG) (NAS, 2017). Researchers provided commentary on the viability of the discounting assumptions used in U.S. government SC-CO_2 estimates, which is a metric used in regulatory impact analysis to account for net damages from climate change associated with an incremental change in emissions. The discussion allowed for shortcomings in the IWG's discounting assumptions to be quickly diagnosed and provided guidance on the selection of parameters, such as the pure rate of time preference and utility curvature (i.e. the relationship between additional wealth and additional wellbeing), treating uncertainty in producing and applying ranges of SC-CO_2 estimates, and combining SC-CO_2 values from different integrated assessment models (IAMs) with other cost and benefit estimates. The report also underscored the value in making models available and modular to encourage the 'transparent articulation of the inputs, outputs, uncertainties, and linkages' (NAS, 2017, p. 9).

Despite progress in improving the transparency of discounting assumptions, many opportunities for progress still exist. For instance, a recent Intergovernmental Panel on Climate Change (IPCC) special report featured IAM scenarios to evaluate mitigation pathways compatible with limiting warming to 1.5 °C above pre-industrial levels (IPCC, 2018). Although a publicly available scenario database was established for the special report, the models did not conduct sensitivities to test the robustness of results to alternate assumptions about discounting, which could alter the time pathway of mitigation options.[3]

Overall, discounting represents a success story of acting on salient ethical and behavioural issues in energy and climate analysis. Disagreement in the research community encouraged standards for interdisciplinary engagement, documenting assumptions, providing guidance on interpreting analysis, and testing the robustness of results to a range of inputs. Importantly, to deepen transparency, it was not sufficient to merely satisfy open-source guidelines – instead, much additional work involving interdisciplinary collaboration was needed. Improving collaboration between disciplines by organizing interdisciplinary conferences and projects is reproducible in other cases involving value-laden assumptions.

2.2. Uncertainty: assessing the impact of the unknown

Like discounting, uncertainty is a central facet in many energy and environmental economics problems. Unlike discounting, however, the treatment of uncertainty may not be as simple as varying a parameter value but may require structural changes to models and potentially a different conceptual paradigm, requiring analysts to quantify values that are not traditionally in their purview like probabilities associated with outcomes (Bistline, 2014; Lempert et al., 2003; Pindyck, 2007). Uncertainty analysis illustrates why transparency is important not only for the assumptions made about particular parameters but also for the presuppositions about structural features of models, including omitted dynamics.

The characterization and analysis of uncertainty have received increased attention in the climate change economics community in recent years. Although insuring against low-probability, high-impact risks has been a longstanding argument for emissions mitigation (Manne & Richels, 1993), more recent work (e.g. Weitzman,

2009) considers how uncertainty could impact the degree and type of precautionary action warranted to avoid climate change impacts. Notwithstanding the importance of risk in climate policy and in long-lived investments in the energy system, quantification and implementation of uncertainty in the modelling community has lagged behind its acknowledged importance (Bistline, 2015a). Integrated assessment and energy models are largely deterministic and focus on mean-valued outcomes, rarely incorporating uncertainty in the analysis to determine hedging strategies. Although the assumption of risk neutrality is recognized to be problematic in regulatory contexts, guidance for benefit–cost analyses through the U.S. Office of Management and Budget directs agencies to ignore these considerations (Kaufman, 2014).

Debates about work like Weitzman's have led to reappraisals of the treatment of uncertainty in energy and environmental models and, more generally, of how to conceptualize uncertainty and risk and their relation to public policy (Bistline, 2015b; Hausman et al., 2016; Stirling, 2010; Tol, 2003: Walker et al., 2003; Weitzman, 2009).[4] Trutnevyte et al. (2016) advocate for a scenario technique to capture multi-dimensional uncertainty, which involves qualitative studies of model users to test scenerio effectiveness. Although these dialogues have thus far been between economists, other domains of expertise can inform these deliberations, given how decision-making under uncertainty involves important economic, ethical, and technical dimensions. For instance, formulating, developing, and applying a stochastic IAM to inform climate policy might involve economists (to frame the decision problem), operations researchers (to overcome the computationally burdensome nature of such complex problems), social scientists (to elicit probabilities and understand possible heuristics and biases), philosophers (to inform the ethical and decision theoretic elements of the model and its use), and climate scientists (to improve the representation of the climate system).

Although the importance of uncertainty is being increasingly recognized, the research community has not settled on best practices, including:

- Selecting frameworks for uncertainty analyses like one-way sensitivity, scenario analysis, uncertainty propagation, sequential decision-making under uncertainty (Kann & Weyant, 2000) and methods to test the sensitivity of results to alternate frameworks before creating resource-intensive models (Bistline, 2015a)
- Selecting plausible scenarios and uncertainties, since a range of biases and heuristics may lead to a narrow list of potentially relevant factors (Trutnevyte et al., 2016)
- Assigning probabilities to outcomes, which may depend on a specific decision problem but can be distorted by biases (Kahneman et al., 1982)
- Communicating assumptions and results to stakeholders (Kahneman & Tversky, 2000)

2.3. Socioeconomic distribution of impacts: incorporating and evaluating effects across socioeconomic boundaries

As discussed above, discounting provides a familiar example of the sensitivity of models' recommendations to the intergenerational distribution of costs and benefits, i.e. the distribution across time. Another distributional issue is accounting for intra-generational costs and benefits within time periods across different countries and across different income levels. Our focus in this discussion is on equity considerations between and within countries related to large-scale climate policy.[5] Distributional assumptions are embedded through a social welfare function, which assigns value to outcomes within the model – e.g. when an optimal policy is calculated, this is the objective function that is maximized. Within the social welfare function, the choice of its specific form and parameters manifest a choice of attitudes toward the importance of equity and justice within and across generations. For example, a social welfare function that focuses only on maximizing global gross domestic product would ignore the greater importance to poor and wealthy households of additional dollars of income, whereas alternative social welfare functions have the ability to explicitly represent the diminishing marginal utility of income to the poor vs. rich, and thereby manifest concern for distributional inequality. The latter welfare functions allow parameterization of a concern for equity – when such a parameter is included, it is often called an 'equity weight' (Anthoff & Tol, 2010).

Despite the significance of these assumptions, distributional issues, unlike uncertainty, have not been the focus of rigorous investigation and debate. Indeed, best practices have not yet emerged (as they have in the case of discounting). For example, estimates of the SC-CO_2 aggregate monetized climate impacts over populations in different jurisdictions and incomes. When IAMs use different assumptions about the concern for and liability toward citizens of other countries, marginal damage estimates from CO_2emissions vary significantly and hence offer different recommendations for climate policy (Anthoff & Tol, 2010; Budolfson et al., 2017; Dennig et al., 2015). Although two of three models used to estimate the SC-CO_2 by the U.S. government have the structural ability to incorporate equity weighting between nations (Rose et al., 2014), equity is not given important weight in any official government modelling reports, and equity receives little attention in the SC-CO_2 estimation literature (NAS, 2017).

Further, all of the leading SC-CO_2 models lack the structure to perform equity weighting *within* nations, because they lack the representation of inequalities in costs and benefits for a nation's own citizens. As a result, all of these models *implicitly* assume that mitigation costs and future climate damages are distributed proportional to income within nations (Budolfson et al., 2017; Dennig et al., 2015). Identifying this implicit assumption and quantifying its effects are important for deepening transparency.

3. Recommendations for deepening transparency

Building on these examples, we offer recommendations to inform and improve analysis by creating shared standards to encourage transparency for both modellers and non-modellers. The goal is to create opportunities to accelerate the adoption of best practices to achieve results described in the discounting section sooner for a range of audiences, especially in relation to contentious and value-laden assumptions. These proposed standards are modular and complementary to facilitate adoption, and we propose reforms that encourage transparency, but balance researchers' time and effort with the value to the research community. Given the decision relevance of energy and environmental modelling, each recommendation would be enhanced through stakeholder participation (Reed, 2008). Strachan et al. (2016) provide suggestions on tailoring model-based analysis to policymakers and creating iterative engagement structures between modellers, stakeholders, and other experts. While Strachan et al. (2016) focus on improving the energy-modelling policy interface through transparency, our recommendations call for a deeper level of transparency that draws out value-laden assumptions through interdisciplinary engagement as well as sensitivity analysis. Table 1 summarizes these recommendations, which are detailed further below.

3.1. Recommendation 1: Additional sensitivity analysis and model diagnostics (rows 1–3 in Table 1)

Consumers of models, such as policymakers, industry, foundations, and other stakeholders, tend to look to published research for guidance on policy issues. Although these are often knowledgeable and motivated audiences, they are unable to devote considerable time and resources to learning and applying new models. In light of this omission, modellers should routinely run additional sensitivities to test robustness of results to parameters and structural assumptions, especially for value-laden or contentious model assumptions. Models should be run with all possibilities for uncertain or contentious assumptions, which has become a standard practice for issues like discounting (Section 2.1) but not for ones like equity weights (Section 2.3). Such model runs would allow consumers of model results to understand how alternate assumptions can impact recommendations (Morgan, 2015; Morgan & Henrion, 1990; Pannell, 1997). These stress tests can increase confidence in results and allow audiences with different values and beliefs to assess the reasonableness of model behaviour under different domains. They also steer model improvements toward areas that have the greatest impact on results.

DeCarolis et al. (2017) highlight how uncertainty analysis is a best practice in energy system optimization modelling given uncertainties about future technologies and policies, but decisions about incorporating uncertainty also have ethical dimensions, as discussed in Section 2.2. Encouraging a range of sensitivity analyses also promotes long-run norms of flexibility in model development, and the ability to conduct many sensitivities becomes a central design element, as the social cost of carbon example in Section 2 illustrates. As others have pointed out (e.g. Strachan et al., 2016), models tend toward increasing complexity as features are added but rarely removed. However, the expectation that reviewers for peer-reviewed articles and policymakers

Table 1. Examples of best practices and levels of engagement to encourage transparency. Column categories of 'Shallower' through 'Deeper' are ordered by increasing depth of transparency.

	Shallower	Deep	Deeper
1. Additional Sensitivity Analysis	Analysis performs sensitivities to the most important unknowns and acknowledges sensitivities that are not performed	Analysis provides a range of sensitivities to important parameters and structural assumptions, including potentially value-laden or contentious ones	Analysis provides a range of sensitivities; value-laden assumptions highlighted, defended, and discussed; other venues for broader engagement (e.g. interactive websites, clarifying policy implications)
2. Model Diagnostics	Limited model diagnostics performed to test model robustness but not made public	Limited model diagnostics performed and made public	Comprehensive diagnostics performed for all reasonable alternatives, publicly available report, and updated periodically
3. Uncertainty	Uncertainties clearly acknowledged but not formally characterized or analyzed	Uncertainty characterized (through statistical methods, elicited expert judgments, or other approaches) but not formally incorporated	Uncertainty characterization, analysis, and communication
4. Methods Transparency	White paper	Peer-reviewed	Peer-reviewed with periodic updates; other venues for broader engagement (e.g. crowdsourcing and collaborative websites) with value judgements highlighted and discussed
5. Data Transparency	Documentation cites sources for data and mentions whether they are available	Sources cited and data available by contacting authors	Sources cited and data posted to trusted repository with guide to non-modellers for use and possible limitations; quality assurance procedures
6. Code Transparency	Detailed documentation of model equations and implementation	Model code available on a trusted repository, enabling independent replication	Model code publicly available, extensively commented for other users, and documentation available for less experienced modellers with value judgments discussed; external verification
7. Interdisciplinary Engagement	Incorporate information from multiple disciplines	Multidisciplinary teams conduct analysis	Interdisciplinary research teams (including social scientists and ethicists) conduct analysis, elicit feedback, and communicate results to stakeholders; robust exchange about contested assumptions, with emphasis on value-laden assumptions (e.g. conferences, special journal issues)
8. Model Intercomparison Projects	Qualitative comparison of model structures and assumptions	Qualitative comparison coupled with a quantitative comparison of model outputs, including harmonized input assumptions (to isolate impacts of model structure)	Qualitative and quantitative comparisons with harmonized inputs; comprehensive assessment of differences in model inputs and outputs, including ethical implications; periodic reassessments with interdisciplinary teams; venues for broader engagement

will ask for many dozens of cross-sensitivities should encourage modelling teams to lower the computational cost of running their models by creating modular code where complexity can be tailored to the analysis (Cole et al., 2017).[6]

Sensitivity analysis, when paired with probability distributions that reflect expert judgment or empirical estimates, can provide guidance about the value of more detailed research (Bistline, 2015a). Given how modelling approaches that explicitly incorporate uncertainty into decision-making are resource intensive to create and computationally costly to run, it is important to use metrics like the value of the stochastic solution to test the sensitivity of results to alternate frameworks before creating more complex models (Bistline, 2015a), as described in Section 2.2. Modellers should ensure that a reasonably complete range of uncertainty is captured. For example, a one-way sensitivity, where a single parameter is varied at a time, is the de facto choice for many modellers, but a multi-dimensional sensitivity analysis is often required to diagnose interactions between parameters when varied simultaneously. Additional sensitivities also help to 'future proof' an analysis as conditions change and increase the relevance of results even when technological or policy-related surprises unfold. For instance, solar power costs have plummeted in recent years and have exceeded even optimistic forecasts (Creutzig et al., 2017). However, the scenarios in Working Group III of the IPCC's Fifth Assessment Report did not span ranges that included these trends (Kriegler et al., 2014). Large ensembles of scenarios can enable more robust decisions by offering a wide range of perspectives rather than an ex ante specification of a limited set of likely futures and can avoid conveying undue precision in the ability to understand structural and parametric uncertainties (Bistline & Young, 2020; Eshraghi et al., 2018). As discussed in Section 2.2, sensitivity analysis should be complemented by techniques for uncertainty analysis such as stochastic or robust optimization.

Ideally, source code should be made available with sufficient documentation to allow other researchers to run the code to explore the robustness of conclusions. When such arrangements are infeasible, modellers should provide a basic set of sensitivities themselves. Modellers should always be sensitive to potential blind spots in their analysis. Allowing other researchers to conduct verification exercises and stress test their work should be a welcomed method for efficient progress, replication, and self-correction. One way to encourage more robust uncertainty analysis is to create standard scenarios that modelling groups routinely run and update each year to track how model insights change with model development and as input assumptions change. Lessons learned are identified by comparing outputs across time and with actual market dynamics (NREL, 2018).

Sensitivity analysis can be helpful for communicating the role of value-based assumptions in models to non-experts. However, altering the discount rate, for example, can yield a broad range of modelling outcomes, seriously complicating the communication of model insights. Exposure to these complications may lead non-experts to question the utility of model analysis in the first place. In reply to this objection, we think that using sensitivity analysis can help to educate non-experts about the usefulness of models for helping to understand the implications of disagreements over value judgments on outcomes. As Strachan et al. (2016) point out, policymakers struggle to assess different insights from competing models as it is. Using sensitivity analysis to distinguish different value-based assumptions could actually help non-experts gain confidence in the utility of models for making decisions, especially in the context of contentious debates.

3.2. Recommendation 2: Better model documentation (beyond equations) to enable open communication across audiences and disciplines (rows 4–6 in Table 1)

One purpose of a model is to convey expert knowledge without the expert present. This objective cannot be advanced without proper documentation. Public engagement and deliberation require information, which is a particularly important component of the energy and environmental modelling process given strong public policy dimensions related to climate change, energy poverty, and economic development. As the first line of defense against sloppy analysis or misinterpretation of insights, modellers should adopt a leading role in communicating not only the assumptions of what is in a model but also what is left out, especially for issues whose importance is not yet sufficiently recognized. Deliberations about model design decisions are frequently omitted even in detailed documentation, and the existence and consequences of these decisions may be unknown to consumers and users of models. These omissions may have serious consequences when design decisions involve widely disagreed upon questions about ethics and value. Documentation allows audiences to understand modeller judgments and drivers of results to evaluate the reasonableness of assumptions and structures.

Journal policies can encourage transparency and rigorous documentation (Nosek et al., 2015). Top-tier journals in economics and political science now often require software code and data availability (Gertler et al., 2018), which is a critical step but does not go far enough to convey limitations to non-expert audiences and fails to highlight contentious value-laden assumptions. Thorough documentation is not necessarily a substitute for making model code available, which is needed to replicate numerical results. However, in light of competitiveness and intellectual property concerns, thorough documentation combined with sensitivity analysis can provide a basis for independent verification of research findings (i.e. replication of higher-level insights using the same assumptions but different modelling platforms). Furthermore, publishing analysis in peer-reviewed journals helps to strengthen the link between modelling choices and accountability. Replication using different platforms encourages feedback and sometimes collaboration between research groups. In addition, scholars from other disciplines can then more easily leverage existing models to explore further interdisciplinary research that addresses further questions with novel extensions of existing methods. For example, William Nordhaus's open-source posting of his models with extensive documentation, including of conceptually and philosophically contentious assumptions, has allowed other researchers to efficiently test the sensitivity of its policy recommendations to alternative assumptions (see for example Budolfson et al. (2019); Budolfson et al. (2017); and Scovronick et al. (2017)). Independent replications could be incentivized by creating special issues and dedicated journals that publish these studies (which have been emerging trends in recent years) and by having replications rewarded by funding agencies.[7]

Documentation of a model's structure and assumptions should be provided at different levels of detail to match the needs of different audiences. Current guidance focusing on model formulations may be useful for other modellers but not for other stakeholders. These different audiences require additional targeted documentation, as DeCarolis et al. (2012) and Pfenninger et al. (2017) point out. But the degree to which assumptions are discussed (and sensitivities tested) should be application-specific and should coincide to the degree to which these assumptions change decisions (Füssel, 2007). For instance, as discussed in Sections 2.2 and 2.3, omitted model considerations like risk and distributional impacts can be as important as features included in models, so documentation should apportion greater space to highlighting these issues even where they are not addressed directly. Documentation should also point out which structural and parametric uncertainties are due to unknowns about future values (e.g. solar PV capital costs), system dynamics and behavioural assumptions (e.g. consumer tradeoffs between vehicle purchase price and lifetime fuel/maintenance costs), and societal choices/preferences (e.g. policy design and implementation).

Another function of transparency is to provide recourse and accountability for the quality of model development decisions. Many energy and sustainability models cannot be validated in the same way that physical sciences models can (DeCanio, 2003). This means model quality is not determined by predication accuracy (i.e. discrepancies between select scenario outputs and observed states of the world). Rather, assessments of model value require technical domain expertise as well as transparency in model structure and assumptions. For instance, electric sector models determine generation shares of different technologies under assumptions about future markets, policies, and technologies. Although forecasts of parameters impact model outcomes, results are also driven by structural features of models like their spatial and temporal resolutions (Mai et al., 2018).

Documentation should not be cloaked in domain-specific jargon, as the over-reliance on technical language erects barriers to interdisciplinary collaboration. Howells et al. (2011) argue for documentation with a 'plain English description' to explain engineering and economic concepts of energy systems models to non-modelling audiences, but such recommendations are just as applicable for value-laden issues and ethically contentious assumptions. For instance, Negishi weights, when mentioned in model documentation at all, are often described in economic terms (e.g. equalizing the shadow price of capital) instead of in relation to their equity implications. A 'plain English description' of Negishi weights would include discussion that locates their ethical assumptions among alternative answers to contentious questions about equity, and highlights their implications when used within models, such as valuing lives lost in poorer nations less than lives lost in richer nations (Stanton, 2011).

A linked recommendation is not only to produce clear model documentation but also to communicate findings through several channels that are tailored to the needs of different audiences. This strategy can be as simple as creating an executive summary to accompany a lengthy report, a video abstract to accompany a journal article, or a summary for policymakers to supplement modelling efforts. Recent examples of such outreach are the targeted summaries of the IPCC Special Report on Global Warming of 1.5 °C, such as the summary for teachers (OCE, 2018). Researchers at University College London, in the UK, have developed animations to explain UK energy models to broad audiences (UCL Energy Institute, 2017).

3.3. Recommendation 3: Greater discussion across modelling groups and other disciplines, especially through model intercomparison projects (rows 7–8 in Table 1)

So far, our recommendations concentrate on how single modelling groups can deepen transparency. The third recommendation explicitly requires increasing communication across the modelling community and other disciplines. One encouraging option is the development of an entirely new discipline, 'Macro-Energy Systems' (Levi et al., 2019), which would pull into a single field researchers who currently work on energy systems in multiple disciplines and often publish in separate journals. The creation of Macro-Energy Systems as a discipline also has promise for improving discussion on ethical issues related to energy, as it presents an institutional platform to directly involve researchers in the humanities and social scientists in modelling projects.

Model intercomparison projects (MIPs) are coordinated multi-model exercises used to understand differences across data, methodologies, and outputs and have been used in a variety of fields like climate science and energy modelling (Weyant, 2017). MIPs can identify robust insights that hold across a range of models and analytical strategies and highlight areas of disagreement to guide future research (Parson, 1995). They provide a means for benchmarking and self-correction to adopt best practices and to follow the most productive research directions. These multi-model studies also establish a credible expectation among modellers that their research will be replicated across independent platforms and discourage decisionmakers and journalists from succumbing to 'single-study syndrome'. MIPs provide a transition from static, fragmented studies published by a single research group to open-ended collaborative commentaries that are dynamically updated, extended, and linked, incorporating perspectives from a range of relevant disciplines. A virtue of modelling is to provide a platform for communicating across disciplinary boundaries by transparently harmonizing and organizing consistent mathematical representations of disciplinary knowledge into a common accounting framework. Formal MIPs also increase incentives for collaboration with the potential to provide better alignment between individual and communal incentives, including increasing citations for model intercomparisons, replications, code, and null results (Nosek et al., 2015). In encouraging broad participation, multi-model studies benefit from having diverse models and perspectives represented, especially when groups are willing to spend time systematically investigating inputs and outputs to understand the degree to which differences are due to input assumptions or to model structure (Mai et al., 2018).

For energy and environmental models, as noted in Section 2.1 above, venues like Stanford University's Energy Modeling Forum have historically facilitated meetings and studies to focus on specific research topics (Huntington et al., 1982). Interdisciplinary collaborative commentaries encourage model transparency and can investigate the sensitivity of recommendations to a range of assumptions and structures while discussing issues behind the assumptions, which can lead to the identification and clarification of questions for future research (Edenhofer et al., 2006). The optimal size of these MIPs depends on the particular research question and available resources, as smaller comparisons facilitate deeper investigation of factors driving results (Cole et al., 2017) while larger studies provide a broader 'lay-of-the-land' (Bistline et al., 2018). Although larger studies are often open to any interested modelling groups, financial resources may constrain participation from a wider variety of institutions and fields, so to encourage broader perspectives, public funding should be earmarked for their participation in collaborative studies.

Researchers both in multi-model studies and single-model efforts can deepen transparency by finding creative approaches to engage stakeholders and increase accessibility. For instance, crowdsourcing and collaborative websites offer open-ended forums for resources, discussion, and documentation and allow stakeholders with diverse backgrounds to deliberate alongside modellers (Bazilian et al., 2012). These repositories also provide important training functions for analysts to be alert to ethically sensitive dimensions of models, and the collaborative norms fostered by crowdsourcing could have positive spillover effects for other researchers and the public. A second approach to enhance communications is to create interactive websites for the public to allow interested laypeople to see how results change with different assumptions (MIT, 2009; U.S. EIA, 2018). These websites allow audiences to ask better questions of experts and modellers, while testing the robustness of conclusions.

4. Conclusion and policy implications

Important societal issues are increasingly informed by quantitative modelling that requires interdisciplinary collaborations. For instance, in the post-Paris Agreement era, as pressure mounts to mitigate climate change while simultaneously achieving the broader SDGs, more and more climate and energy modelling groups are taking centre stage and making their data and equations public (Hurrell et al., 2013; Pfenninger et al., 2018). Although this transparency can improve the scientific process, policy decisions, and dissemination of insights, we argue that it is especially in the context of difficult policy decisions that the need for transparency extends beyond making data, equations, and model documentation merely openly available. Deeper transparency requires making structural assumptions explicit, creating opportunities for interdisciplinary engagement, and communicating value-laden assumptions to policymakers, foundations, and other stakeholders.

Ultimately, energy and environmental modellers must allocate their time and resources across a portfolio of activities, such as documenting models and making their code available, new model development (and enhancing model fidelity to conduct cutting-edge research), applying models in new settings, and communicating results. Some efforts are synergistic, while others create competing claims on scarce time, but all of these activities can foster positive norms within the modelling community in distinct ways. Individual modelling groups must find a balance that works in light of their unique strengths, audiences, preferences for how they spend their time, as well as public and private returns to additional research effort. There is unlikely to be a single correct portfolio of transparency practices for all modelling groups in all contexts. As we describe, transparency can be enhanced through model documentation, application, and communication. But given the range of stakeholders who do not have the time or expertise to delve into code, it is important to consider how model application and communication can make science more accessible to these audiences.

Deepening transparency about modelling assumptions – especially for potentially value-laden assumptions – is urgently needed in scientific contexts that use modelling to advise policy decisions, including energy and environmental models that inform decisions on climate change, as well as local pollution abatement, biodiversity conservation, and agriculture-related policymaking. Making analytical tools as transparent as possible can help to communicate insights to a broader range of stakeholders while appropriately conveying strengths, limitations, and assumptions about value to avoid misinterpreting results.

Notes

1. Recommendations include getting consent from intellectual property holders, choosing a platform, identifying publishable content, choosing a license, considering programming languages and frameworks, and building a community of users.
2. The replication crisis in disciplines like the social sciences have led to recommendations for better practices and changing research norms (Gertler et al., 2018), and although many suggestions are applicable for energy and environmental research, replicability is only a subset of transparency-related issues associated with prospective structural models that are the focus of this manuscript.
3. See Emmerling et al. (2019) for further discussion on the effect of the discount rate for setting climate change policy targets.
4. This discussion focuses on quantitative evaluations of uncertainty. However, we acknowledge much could be learned by improving collaborations between quantitative and qualitative research projects. Qualitative methods designed to engage broader stakeholder perspectives in the context of multiple uncertainities through, for example, NUSAP (Numeral Unit Spread Assessment Pedigree) workshops can also improve transparency and make value-laden assumptions more evident to in energy and environmental assessments (Van Der Sluijs et al., 2005; Petersen et al., 2011; Pye et al., 2018). However, we would argue that these projects fall short of deepening transparency by focusing more on stakeholder involvement than on collaboration between research groups and across disciplines.
5. For discussion of intra-country scale distributional analysis concerning the transition to low carbon energy, see Fell et al. (2019) and Oswald et al. (2020).
6. A compelling case could be made that there are too many energy models (Sarewitz, 2018) and that requiring open-source code will increase the cost of creating new models and therefore decrease the supply. However, while larger and well-resourced incumbents would likely be able to comply with these restrictions (Lange & Redlinger, 2019), the models that are more likely to be materially impacted would be smaller groups with more limited resources, and there may be strong equity arguments in favour of having these perspectives included in broader dialogues about energy and sustainability.
7. Governments may also have a role to play in improving model documentation requirements. For example, the UK has created a quality assurance manual that applies to government analysis (H.M. Treasury, 2015).

Acknowledgments

The authors wish to thank three anonymous reviewers and the editor for their helpful suggestions. The views expressed in this paper are those of the authors alone and do not necessarily state or reflect those of EPRI, Rutgers University, or University of Maryland, Baltimore County, and no official endorsement should be inferred.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

John Bistline http://orcid.org/0000-0003-4816-5739

References

Anthoff, D., & Tol, R. S. J. (2010). On international equity weights and national decision making on climate change. *Journal of Environmental Economics and Management*, *60*(1), 14–20. https://doi.org/10.1016/j.jeem.2010.04.002

Arrow, K. J., Cline, W. R., Goran-Maler, K., Munasinghe, M., Squitieri, R., & Stiglitz, J. E. (1996). Intertemporal equity, discounting, and economic efficiency. In J. P. Bruce, H. Lee, & E. F. Haites (Eds.), *Climate change 1995: Economic and social dimensions of climate change, contribution of Working Group III to the second assessment report of the intergovernmental panel on climate change*. Cambridge University Press.

Arrow, K. J., Cropper, M. L., Gollier, C., Groom, B., Heal, G. M., Newell, R. G., Nordhaus, W. D., Pindyck, R. S., Pizer, W. A., Portney, P. R., Sterner, T., Tol, R. S. J., & Weitzman, M. L. (2014). Should governments use a declining discount rate in project analysis? *Review of Environmental Economics and Policy*, *8*(2), 145–163. https://doi.org/10.1093/reep/reu008

Bazilian, M., Rice, A., Rotich, J., Howells, M., DeCarolis, J., Macmillan, S., Brooks, C., Bauer, F., & Liebreich, M. (2012). Open source software and crowdsourcing for energy analysis. *Energy Policy*, *49*, 149–153. https://doi.org/10.1016/j.enpol.2012.06.032

Benz, J., Hoch, R., & Legović, T. (2001). ECOBAS — Modelling and documentation. *Ecological Modeling*, *138*(1), 3–15. https://doi.org/10.1016/S0304-3800(00)00389-6

Bistline, J. E. (2014). Energy technology expert elicitations: An application to natural gas Turbine efficiencies. *Technological Forecasting and Social Change*, *86*, 177–187. https://doi.org/10.1016/j.techfore.2013.11.003

Bistline, J. E. (2015a). Electric sector capacity planning under uncertainty: Climate policy and natural gas in the US. *Energy Economics*, *51*, 236–251. https://doi.org/10.1016/j.eneco.2015.07.008

Bistline, J. E. (2015b). Fat-tailed uncertainty, learning, and climate policy. *Climate Change Economics*, *6*(2), 1550009. https://doi.org/10.1142/S2010007815500098

Bistline, J. E., Hodson, E., Rossmann, C. G., Creason, J., Murray, B., & Barron, A. R. (2018). Electric sector policy, technological change, and U.S. emissions reductions goals: Results from the EMF 32 model intercomparison project. *Energy Economics*, *73*, 307–325. https://doi.org/10.1016/j.eneco.2018.04.012

Bistline, J., James, R., & Sowder, A. (2019). Technology, policy, and market drivers of (and barriers to) advanced nuclear reactor deployment in the United States after 2030. *Nuclear Technology*, *205*(8), 1075–1094. https://doi.org/10.1080/00295450.2019.1574119

Bistline, J. E. T., & Merrick, J. H. (2020). Parameterizing open-source energy models: Statistical learning to estimate unknown power plant attributes. *Applied Energy*, *269*, 114941. https://doi.org/10.1016/j.apenergy.2020.114941

Bistline, J. E. T., & Young, D. T. (2020). Emissions impacts of future battery storage deployment on regional power systems. *Applied Energy*, *264*, 114678. https://doi.org/10.1016/j.apenergy.2020.114678

Broome, J. (1994). Discounting the future. *Philosophy & Public Affairs*, *23*(2), 128–156. https://doi.org/10.1111/j.1088-4963.1994.tb00008.x

Budolfson, M., Dennig, F., Fleurbaey, M., Scovronick, N., Siebert, A., Spears, D., & Wagner, F. (2019). Optimal climate policy and the future of world economic development. *The World Bank Economic Review*, *33*(1), 21–40. https://doi.org/10.1093/wber/lhx016

Budolfson, M., Dennig, F., Fleurbaey, M., Siebert, A., & Socolow, R. H. (2017). The comparative importance for optimal climate policy of discounting, inequalities, and catastrophes. *Climatic Change*, *145*(3-4), 481–494. https://doi.org/10.1007/s10584-017-2094-x

Christensen, G., & Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature*, *56*(3), 920–980. https://doi.org/10.1257/jel.20171350

Cole, W., Frew, B., Mai, T., Sun, Y., Bistline, J., Blanford, G., Young, D., Marcy, C., Namovicz, C., Edelman, R., Meroney, B., Sims, R., Stenhouse, J., & Donohoo-Vallett, P. (2017). *Variable renewable energy in long-term planning models: A multi-model perspective*. National Renewable Energy Laboratory.

Craig, P. P., Gadgil, A., & Koomey, J. G. (2002). What can history teach us? A retrospective examination of long-term energy forecasts for the United States. *Annual Review of Energy and the Environment*, *27*(1), 83–118. https://doi.org/10.1146/annurev.energy.27.122001.083425

Creutzig, F., Agoston, P., Goldschmidt, J. C., Luderer, G., Nemet, G., & Pietzcker, R. C. (2017). The underestimated potential of solar energy to mitigate climate change. *Nature Energy*, *2*(9), 17140. https://doi.org/10.1038/nenergy.2017.140

DeCanio, S. (2003). *Economic models of climate change: A critique*. Palgrave Macmillan.

DeCarolis, J. F., Daly, H., Dodds, P., Keppo, I., Li, F., McDowall, W., Pye, S., Strachan, N., Trutnevyte, E., Usher, W., Winning, M., Yeh, S., & Zeyringer, M. (2017). Formalizing best practice for energy system optimization modelling. *Applied Energy*, *194*, 184–198. https://doi.org/10.1016/j.apenergy.2017.03.001

DeCarolis, J. F., Hunter, K., & Sreepathi, S. (2012). The case for repeatable analysis with energy economy optimization models. *Energy Economics*, *34*(6), 1845–1853. https://doi.org/10.1016/j.eneco.2012.07.004

Dennig, F., Budolfson, M. B., Fleurbaey, M., Siebert, A., & Socolow, R. H. (2015). Inequality, climate impacts on the future poor, and carbon prices. *Proceedings of the National Academy of Sciences*, *112*(52), 15827–15832. https://doi.org/10.1073/pnas.1513967112

Edenhofer, O., Lessmann, K., Kemfert, C., Grubb, M., & Köhler, J. (2006). Induced technological change: Exploring its implications for the economics of atmospheric stabilization: Synthesis report from the innovation modeling comparison project. *The Energy Journal*, (Special Issue #1). https://doi.org/10.5547/ISSN0195-6574-EJ-VolSI2006-NoSI1-3

Emmerling, J., Drouet, L., Wijst, K.-I. v. d., Vuuren, D. v., Bosetti, V., & Tavoni, M. (2019). The role of the discount rate for emission pathways and negative emissions. *Environmental Research Letters*, *14*(10), 104008. https://doi.org/10.1088/1748-9326/ab3cc9

Eshraghi, H., de Queiroz, A. R., & DeCarolis, J. F. (2018). US energy-related greenhouse gas emissions in the absence of federal climate policy. *Environmental Science and Technology*, *52*(17), 9595–9604. https://doi.org/10.1021/acs.est.8b01586

Fell, M. J., Pye, S., & Hamilton, I. (2019). Capturing the distributional impacts of long-term low-carbon transitions. *Environmental Innovation and Societal Transitions*, *35*, 346–356. https://doi.org/10.1016/j.eist.2019.01.007

Fleurbaey, M., et al. (2018). International Panel on Social Progress. https://www.ipsp.org/wp-content/uploads/2018/10/IPSP-Executive-Summary.pdf

Funtowicz, S. O., & Ravetz, J. R. (1993). The emergence of post-normal science. In *Science, politics and morality* (pp. 85–123). Springer.

Füssel, H.-M. (2007). Methodological and empirical flaws in the design and application of simple climate-Economy models. *Climatic Change*, *81*(2), 161–185. https://doi.org/10.1007/s10584-006-9154-y

Gertler, P., Galiani, S., & Romero, M. (2018). How to make replication the norm. *Nature*, *554*(7693), 417–419. https://doi.org/10.1038/d41586-018-02108-9

Goulder, L. H., & Stavins, R. N. (2002). Discounting: An eye on the future. *Nature*, *419*(6908), 673–674. https://doi.org/10.1038/419673a

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., … DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1-2), 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Ha-Duong, M. (2001). Transparency and control in engineering integrated assessment models. *Integrated Assessment*, *2*(4), 209–218. https://doi.org/10.1023/A:1013384932008

Hausman, D., McPherson, M., & Satz, D. (2016). *Economic analysis, moral philosophy, and public policy* (3rd ed.). Cambridge University Press.

H.M. Treasury. (2015). *The Aqua Book: Guidance on producing quality analysis for government*. HM Government. https://www.gov.uk/government/publications/the-aqua-book-guidance-on-producing-quality-analysis-for-government

Howells, M., Rogner, H., Strachan, N., Heaps, C., Huntington, H., Kypreos, S., Hughes, A., Silveira, S., DeCarolis, J., Bazillian, M., & Roehrl, A. (2011). OSeMOSYS: The open source energy modeling system: An introduction to its ethos, structure, and development. *Energy Policy*, *39*(10), 5850–5870. https://doi.org/10.1016/j.enpol.2011.06.033

Huntington, H. G., Weyant, J. P., & Sweeney, J. L. (1982). Modeling for insights, not numbers: The experiences of the energy modeling forum. *Omega*, *10*(5), 449–462. https://doi.org/10.1016/0305-0483(82)90002-0

Hurrell, J. W., Holland, M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., … Marshall, S. (2013). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *84*(9), 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1

Ince, D. C., Hatton, L., & Graham-Cumming, J. (2012). The case for open computer programs. *Nature*, *482*(7386), 485–488. https://doi.org/10.1038/nature10836

Intergovernmental Panel on Climate Change. (2018). *Special report on global warming of 1.5 °C*. https://www.ipcc.ch/sr15/

Kahneman, D., Slovic, S. P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Kahneman, D., & Tversky, A. (2000). *Choices, values, and frames*. Cambridge University Press.

Kann, A., & Weyant, J. P. (2000). Approaches for performing uncertainty analysis in large-scale energy/economic policy models. *Environmental Modeling and Assessment*, *5*(1), 29–46. https://doi.org/10.1023/A:1019041023520

Kaufman, N. (2014). Why is risk aversion unaccounted for in environmental policy evaluations? *Climatic Change*, *125*(2), 127–135. https://doi.org/10.1007/s10584-014-1146-8

Kelleher, J. P. (2017). Descriptive versus prescriptive discounting in climate change policy analysis. *Georgetown Journal of Law & Public Policy*, *15*, 957–977.

Klein, J. T. (1990). *Interdisciplinarity: History, theory, and practice*. Wayne State University Press.

Kriegler, E., Weyant, J. P., Blanford, G. J., Krey, V., Clarke, L., Edmonds, J., Fawcett, A., Luderer, G., Riahi, K., Richels, R., Rose, S. K., Tavoni, M., & van Vuuren, D. P. (2014). The role of technology for achieving climate policy objectives: Overview of the EMF 27 study on global technology and climate policy strategies. *Climatic Change*, *123*(3-4), 353–367. https://doi.org/10.1007/s10584-013-0953-7

Lange, I., & Redlinger, M. (2019). Effects of stricter environmental regulations on resource development. *Journal of Environmental Economics and Management*, forthcoming.

Lempert, R. J., Popper, S. W., & Bankes, S. C. (2003). *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis*. RAND Corporation.

Levi, P. J., Kurland, S. D., Carbajales-Dale, M., Weyant, J. P., Brandt, A. R., & Benson, S. M. (2019). Macro-energy systems: Toward a new discipline. *Joule*, *3*(10), 2282–2286. https://doi.org/10.1016/j.joule.2019.07.017

Lind, R. C., Arrow, K. J., Corey, G. R., Dasgupta, P., Sen, A. K., Stauffer, T., Stiglitz, J. E., & Stockfisch, J. A. (1982). *Discounting for time and risk in energy policy*. Resources for the Future.

Mai, T., Bistline, J., Sun, Y., Cole, W., Marcy, C., Namovicz, C., & Young, D. (2018). The role of input assumptions and model structures in projections of variable renewable energy: A multi-model perspective of the U.S. electricity system. *Energy Economics*, *76*, 313–324. https://doi.org/10.1016/j.eneco.2018.10.019

Manne, A. S., & Richels, R. G. (1993). *Buying greenhouse insurance: The economic costs of carbon dioxide emission limits*. MIT Press.

Mintz-Woo, K. (2018). *Moral uncertainty over policy evaluation*. PhD dissertation. University of Graz.

MIT. (2009). Greenhouse gamble. https://globalchange.mit.edu/research/research-tools/risk-analysis/greenhouse-gamble

Morgan, M. G. (2015). Our knowledge of the world is often not simple: Policymakers should not duck that fact, but should deal with it. *Risk Analysis*, *35*(1), 19–20. https://doi.org/10.1111/risa.12306

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press.

NAS. (2017). *Valuing climate damages: Updating estimation of the social cost of carbon dioxide*. The National Academies Press.

Nordhaus, W. (2007). ECONOMICS: Critical assumptions in the Stern review on climate change. *Science*, *317*(5835), 201–202. https://doi.org/10.1126/science.1137316

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

NREL. (2018). *2018 standard scenarios report: A U.S. electricity sector outlook*. National Renewable Energy Laboratory.

Office for Climate Education. (2018). *IPCC special report "global warming of 1.5 °C" Summary for teachers*. https://www.ipcc.ch/site/assets/uploads/sites/2/2018/12/ST1.5_OCE_LR.pdf

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, *263*(5147), 641–646. https://doi.org/10.1126/science.263.5147.641

Oswald, Y., Owen, A., & Steinberger, J. K. (2020). Large inequality in international and intranational energy footprints between income groups and across consumption categories. *Nature Energy*, *5*(3), 231–239. https://doi.org/10.1038/s41560-020-0579-8

Pannell, D. J. (1997). Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. *Agricultural Economics*, *16*(2), 139–152. https://doi.org/10.1111/j.1574-0862.1997.tb00449.x

Parson, E. A. (1995). Integrated assessment and environmental policy making: In pursuit of usefulness. *Energy Policy*, *23*(4/5), 463–475. https://doi.org/10.1016/0301-4215(95)90170-C

Peng, R. D., Stodden, V., & Leisch, F. (Eds.). (2014). *Implementing reproducible research*. CRC Press.

Petersen, A. C., Cath, A., Hage, M., Kunseler, E., & van der Sluijs, J. P. (2011). Post-normal science in practice at the Netherlands Environmental Assessment Agency. *Science, Technology, & Human Values*, *36*(3), 362–388. https://doi.org/10.1177/0162243910385797

Pfenninger, S., DeCarolis, J., Hirth, L., Quoilin, S., & Staffell, I. (2017). The importance of open data and software: Is energy research lagging behind? *Energy Policy*, *101*, 211–215. https://doi.org/10.1016/j.enpol.2016.11.046

Pfenninger, S., Hirth, L., Schlecht, I., Schmid, E., Wiese, F., Brown, T., Davis, C., Gidden, M., Heinrichs, H., Heuberger, C., Hilpert, S., Krien, U., Matke, C., Nebel, A., Morrison, R., Müller, B., Pleßmann, G., Reeg, M., Richstein, J. C., … Wingenbach, C. (2018). Opening the black box of energy modelling: Strategies and lessons learned. *Energy Strategy Reviews*, *19*, 63–71. https://doi.org/10.1016/j.esr.2017.12.002

Pindyck, R. S. (2007). Uncertainty in environmental economics. *Review of Environmental Economics and Policy*, *1*(1), 45–65. https://doi.org/10.1093/reep/rem002

Portney, P. R., & Weyant, J. (1999). *Discounting and intergenerational equity*. Resources for the Future.

Proctor, R. N., & Proctor, R. W. (1991). *Value-free science: Purity and power in modern knowledge*. Harvard University Press.

Pye, S., Li, F. G. N., Petersen, A., Broad, O., McDowall, W., Price, J., & Usher, W. (2018). Assessing qualitative and quantitative dimensions of uncertainty in energy modelling for policy support in the United Kingdom. *Energy Research & Social Science*, *46*, 332–344. https://doi.org/10.1016/j.erss.2018.07.028

Reed, M. S. (2008). Stakeholder participation for environmental management: A literature review. *Biological Conservation*, *141*(10), 2417–2431. https://doi.org/10.1016/j.biocon.2008.07.014

Rose, S., Diaz, D., Blanford, G., Bistline, J., de la Chesnaye, F., & Wilson, T. (2014). *Understanding the social cost of carbon: A technical assessment*. (EPRI Report #3002004657). EPRI.

Rose, S., & Scott, M. M. (2018). *Grounding decisions: A scientific foundation for companies considering global climate scenarios and greenhouse gas goals*. (EPRI Report #3002014510). EPRI.

Sarewitz, D. (2018). *Of cold mice and isotopes: Or should we do less science*, Bonn Summer School in Higher Education Research and Science Studies. https://sfis.asu.edu/sites/default/files/should_we_do_less_science-revised_distrib.pdf

Schneider, S. H. (1997). Integrated assessment modeling of global climate change: Transparent rational tool for policy making or opaque screen hiding value-laden assumptions? *Environmental Modeling and Assessment*, *2*(4), 229–249. https://doi.org/10.1023/A:1019090117643

Scovronick, N., Budolfson, M. B., Dennig, F., Fleurbaey, M., Siebert, A., Socolow, R. H., Spears, D., & Wagner, F. (2017). Impact of population growth and population ethics on climate change mitigation policy. *Proceedings of the National Academy of Sciences*, *114*(46), 12338–12343. https://doi.org/10.1073/pnas.1618308114

Stanton, E. A. (2011). Negishi welfare weights in integrated assessment models: The mathematics of global inequality. *Climatic Change*, *107*(3–4), 417–432. https://doi.org/10.1007/s10584-010-9967-6

Stern, N., & Taylor, C. (2007). ECONOMICS: Climate change: Risk, ethics, and the Stern review. *Science*, *317*(5835), 203–204. https://doi.org/10.1126/science.1142920

Stirling, A. (2010). Keep it complex. *Nature*, *468*(7327), 1029. https://doi.org/10.1038/4681029a

Strachan, N., Fais, B., & Daly, H. (2016). Reinventing the energy modelling-policy interface. *Nature Energy*, *1*(3), 16012. https://doi.org/10.1038/nenergy.2016.12

Tol, R. S. J. (2003). Is the uncertainty about climate change too large for expected cost-benefit analysis? *Climatic Change*, *56*(3), 265–289. https://doi.org/10.1023/A:1021753906949

Trutnevyte, E., Guivarch, C., Lempert, R., & Strachan, N. (2016). Reinvigorating the scenario technique to expand uncertainty consideration. *Climatic Change*, *135*(3-4), 373–379. https://doi.org/10.1007/s10584-015-1585-x

U.S. Energy Information Administration. (2018). Annual energy outlook 2018: Interactive table viewer. https://www.eia.gov/outlooks/aeo/data/browser/

U.S. Environmental Protection Agency. (2018). Strengthening transparency in regulatory science. EPA–HQ–OA–2018–0259, Federal Register Vol. 83 No. 83. https://www.federalregister.gov/documents/2020/03/18/2020-05012/strengthening-transparency-in-regulatory-science

UCL Energy Institute. (2017). UCL-Energy creates animation explaining model of UK energy system. https://www.ucl.ac.uk/bartlett/energy/news/2017/jan/ucl-energy-creates-animation-explaining-model-uk-energy-system

Van Der Sluijs, J. P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005). Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: The NUSAP system. *Risk Analysis*, *25*(2), 481–492. https://doi.org/10.1111/j.1539-6924.2005.00604.x

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, *4*(1), 5–17. https://doi.org/10.1076/iaij.4.1.5.16466

Weitzman, M. L. (2009). On Modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics*, *91*(1), 1–19. https://doi.org/10.1162/rest.91.1.1

Weyant, J. (2017). Some contributions of integrated assessment models of global climate change. *Review of Environmental Economics and Policy*, *11*(1), 115–137. https://doi.org/10.1093/reep/rew018

Wilkerson, J. T., Cullenward, D., Davidian, D., & Weyant, J. P. (2013). End use technology choice in the National Energy Modeling System (NEMS): An analysis of the residential and commercial building sectors. *Energy Economics*, *40*, 773–784. https://doi.org/10.1016/j.eneco.2013.09.023

Yue, X., Pye, S., DeCarolis, J., Li, F. G. N., Rogan, F., & Gallachóir, B. Ó. (2018). A review of approaches to uncertainty assessment in energy system optimization models. *Energy Strategy Reviews*, *21*, 204–217. https://doi.org/10.1016/j.esr.2018.06.003