

PERSPECTIVE

Unpacking the modeling process for energy policy making

Samuele Lo Piano¹  | Máté János Lőrincz¹ | Arnald Puy² | Steve Pye³ |
Andrea Saltelli^{4,5} | Stefán Thor Smith¹ | Jeroen van der Sluijs⁵

¹School of the Built Environment, University of Reading, Reading, Berkshire, UK

²School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

³Bartlett School Env, Energy & Resources, University College London, London, UK

⁴UPF-Barcelona School of Management, Barcelona, Catalonia, Spain

⁵Centre for the Study of the Sciences and the Humanities, University of Bergen, Bergen, Norway

Correspondence

Samuele Lo Piano, School of the Built Environment, TSBE Centre, JJ Thomson Building, University of Reading, Reading RG6 6ED, Berkshire, UK.
Email: s.lopiano@reading.ac.uk

Funding information

EU Horizon Europe, Grant/Award Number: 101076165; UK Research and Innovation, Grant/Award Numbers: EP/R035288/1, EP/Y02463X/1

Abstract

This article explores how the modeling of energy systems may lead to an undue closure of alternatives by generating an excess of certainty around some of the possible policy options. We retrospectively exemplify the problem with the case of the International Institute for Applied Systems Analysis (IIASA) global modeling in the 1980s. We discuss different methodologies for quality assessment that may help mitigate this issue, which include Numeral Unit Spread Assessment Pedigree (NUSAP), diagnostic diagrams, and sensitivity auditing (SAUD). We illustrate the potential of these reflexive modeling practices in energy policy-making with three additional cases: (i) the case of the energy system modeling environment (ESME) for the creation of UK energy policy; (ii) the negative emission technologies (NETs) uptake in integrated assessment models (IAMS); and (iii) the ecological footprint indicator. We encourage modelers to adopt these approaches to achieve more robust, defensible, and inclusive modeling activities in the field of energy research.

KEYWORDS

energy systems modeling, non-neutrality of methods, NUSAP, postnormal science, sensitivity auditing, uncertainty

1 | INTRODUCTION

Models are representations of a system. In this very feature resides their power: encoding the relations of the system that is represented and decoding the relations to understand the properties of the system from the model (Rosen, 1991).

Energy system modeling has become a mainstay of many policy processes. For instance, it has been extensively used in the United Kingdom since 2003 to inform the different energy and climate strategies, as well as legislative target-setting (Williams et al., 2018). However, energy policy is informed and often led by small sets of models and system representations. This is the case, for instance, in the energy policy making of the European Union, which has been extensively based only on the use of the findings of the price-induced market equilibrium system model (Clark, 2011; European

Commission, 2021). The methodological approach, or technical stance, adopted in a modeling activity is not neutral; however, it actually conditions the narrative produced within an analysis and consequently the decision-making it is meant to inform (Di Fiore & Saltelli, 2023; Saltelli, Benini et al., 2020). This aspect can lead to important controversies in a decision-making setting, especially when not transparently disclosed.

On top of this, the use of quantification has significantly increased over the last decades with the inflation of metrics, indicators, and scores to rank and benchmark options (Muller, 2018). The case of energy policy making in the European Union is again an effective example. The European Union's recent energy strategy has been underpinned by the *Clean Energy for all Europeans* packages, which are in turn supported by a number of individual directives, each

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Risk Analysis* published by Wiley Periodicals LLC on behalf of Society for Risk Analysis.

one characterized by a series of quantitative goals (European Commission, 2023). The quantification of the impact (impact assessment) is customarily required to successfully promote new political measures (European Commission, 2015a) and is in turn based on quantification, often from mathematical models (Saltelli et al., 2023). The emphasis on producing exact figures to assess the contribution of a new technology, political or economic measure has put many models and their users into contexts of decision-making that at times extends beyond their original intent (Saltelli, Bammer et al., 2020). At the same time, the efforts to retrospectively assess the performance of energy models have been extremely limited, one example being the *Energy Modeling Forum* in the United States (Huntington et al., 1982). In spite of this, retrospective assessments can be very helpful in understanding the sources of mismatch between a forecast and the actual figures reported a posteriori (Kooimey et al., 2003). For example, long-range forecast models are typically based on the assumption of gradual structural changes, which are at stake with the disruptive events and discontinuities occurring in the real world (Craig et al., 2002). This dimension is especially important in terms of the nature and pace of technology change (Bistline et al., 2023; Weyant & Olavson, 1999). A further critical element in this approach is the cognitive bias in scenario analysis that naturally leads to overconfidence in the option being explored and results in an underestimate of the ranges of possible outcomes (Morgan & Keith, 2008).

Additionally, in their quest for capturing the features of the energy systems represented, models have increased their complicatedness and/or complexity. In this context, the need to appraise model uncertainty has become of paramount importance, especially considering the uncertainty due to propagation errors caused by model complexification (Puy et al., 2022). In ecology, this is known as the O'Neil conjecture, which posits a principle of decreasing returns for model complexity when uncertainties come to dominate the output (O'Neill, 1989; Turner & Gardner, 2015). Capturing and apportioning uncertainty is crucial for a healthy interaction at the science–policy interface, including energy policy making, because it promotes better informed decision-making. Yet Yue et al. (2018) found that only about 5% of the studies covering energy system optimization models have included some form of assessment of stochastic uncertainty, which is the part of uncertainty that can be fully quantified (Walker et al., 2003). When it comes to adequately apportioning this uncertainty onto the input parameters and hypotheses through sensitivity analysis, the situation is even more critical: Only very few papers in the energy field have made the use of state-of-the-art approaches (Lo Piano & Benini, 2022; Saltelli et al., 2019). Further to that, the epistemic part of uncertainty, the one that arises due to imperfect knowledge and problem framing, has been largely ignored in the energy modeling literature (Pye et al., 2018). For instance, important sources of uncertainties associated with regulatory lag and public acceptance have typically been overlooked.¹

In this contribution, we discuss three approaches to deal with the challenges of non-neutrality and uncertainty in models: The numerical unit spread assessment pedigree (NUSAP) method, diagnostic diagrams, and sensitivity auditing (SAUD). These challenges are especially critical when only one (set of) model(s) has been selected to contribute to decision-making. One practical case is used to showcase in retrospective the relevance of the issue and the associated problems: the International Institute for Applied Systems Analysis (IIASA) global modeling in the 1980s.

We illustrate how the approaches we propose can assist in dealing with the challenges of decision-making in modeling by systematically examining their use in three case studies: the UK energy modeling for policy-making support, the uptake of negative emissions technologies (NETs) in integrated assessment models (IAMs), and the ecological footprint (EF). Finally, we draw conclusions from the lessons learned and their implications for modeling activities and their use at the policy-making interface.

2 | METHODS

The approaches we discuss have been inspired by the epistemological reflexivity of postnormal science (PNS). PNS becomes relevant in the presence of disputed values, high stakes, urgent decisions, and uncertain facts, all aspects that characterize energy policy making (Funtowicz & Ravetz, 1993). This characterization of energy policy is justified by factors such as trade-offs of socio-environmental aspects, possible important environmental impacts, and unknown relations and causalities in the socio-technical and environmental domain.

We selected NUSAP, diagnostic diagrams, and SAUD for this contribution due to their capability to address crucial aspects of the assumptions and modeling relations underpinning quantification. The appraisal of the assumptions is of paramount importance at the science–policy interface, a fortiori in a controversial policy domain such as on energy matters.

The NUSAP system for communication and management of uncertainty assesses the broader dimensions of uncertainty in quantitative analysis (Funtowicz & Ravetz, 1990; van der Sluijs, 2017; van der Sluijs, Craye et al., 2005). This approach retains the strengths of quantitative uncertainty assessment but adds a focus on the assessment of the quality or “pedigree” of the underlying model assumptions. It broadens the critical appraisal of knowledge to include several dimensions and locations of uncertainty in the modeling approach, including model structure (relationships embedded in equations), and model inputs (data, system boundaries, and problem frames) (Petersen et al., 2013; van der Sluijs & Petersen, 2008; Walker et al., 2003).

Pedigree is judged against multiple criteria based on a structured scoring system (van der Sluijs, Risbey et al., 2005). Possible choices of these criteria include the following:

¹ We thank one anonymous reviewer for pointing us to this issue.

TABLE 1 Numeral unit spread assessment pedigree (NUSAP) score against criteria (adapted from van der Sluijs, Craye et al. (2005)).

Criterion	Score				
	4	3	2	1	0
Proxy (representation)	Exact	Good	Moderate or acceptable	Weak	Poor
Empirical basis	Observation	Mix of observations and model-based estimates	Model estimates only	Educated guess	Crude speculation
Rigor	Best available practice	Reliable method; very few concerns	Acceptable method but questions on reliability	Preliminary or experimental method with no clear view of reliability	No discernible rigor in the method, grave concerns
Validation	Huge database of reliable sources	Compared with numerous reliable sources	Limited validation with only a few reliable sources	Weak validation, questions on reliability of sources	No validation
Theoretical understanding	Universally agreed theory	Accepted theory	Accepted theory but lack of consensus	Preliminary theory	Crude speculation
Choice space	No alternatives	Only a few plausible/acceptable alternatives	Small or limited range of alternatives	Moderate range of alternatives	Extremely wide range of alternatives
Justification	Fully justified	Strong justification	Acceptable justification	Weak justification	Completely speculative
Agreement amongst peers	Complete or near-complete agreement	High degree of agreement, with some variation	Some disagreement possible, there are a few competing schools of thought	Low degree of agreement, contentious subject	No agreement or almost no agreement, extremely controversial

- The *proxy* representation of the real-world system (how good or close a measure of the modeled quantity is to the actual quantity represented).
- The *empirical basis* of the numbers used (the degree to which direct observations, measurements, and statistics are used to estimate the parameter).
- The *rigor* of the methods to derive numbers (the norms for methodological rigor in this process applied by peers in the relevant disciplines).
- *Validation* (the degree to which one has been able to cross-check the data and assumptions used to produce the value of the parameter against independent sources).
- Level of *theoretical understanding* of the systems being modeled (the extent and partiality of the theoretical understanding that was used to generate the value of that parameter).
- The *choice space* (the degree to which alternative choices of plausible/acceptable assumptions could be made).
- The *justification* of the approximations made in the model (their reasonability, plausibility, or acceptability given one's understanding of reality).
- The level of *agreement amongst peers* (the coincidence between the model assumptions and those of other experts in the field).

The scoring system as per the criteria detailed is found in Table 1.

The pedigree scores can be combined with the quantitative indices obtained from a sensitivity analysis on how the output uncertainty is affected by a given input/assumption

(Saltelli et al., 2008). These two dimensions of uncertainty—stochastic (sensitivity indices) and epistemic (pedigree score)—can be visualized in relation to each other using a *diagnostic diagram* (see Section 3.2.2).

However, capturing the quantitative part of uncertainty in a technical sensitivity analysis does not suffice to address the epistemic background and scope of quantification, a crucial aspect in policy-relevant modeling studies. To this end, Saltelli and Funtowicz (2014) sought to enhance sensitivity analysis with SAUD, which has also been recently taken as one of the core ingredients of a manifesto for responsible quantification (Lo Piano et al., 2022, 2023; Saltelli, Bammer et al., 2020). SAUD is based on a checklist that allows for a systematic check of the scope and background of quantification, including modeling activities and indicators, in a policy-making context. The checklist covers the following points:

- **Rhetorical use:** Check against a rhetorical use of mathematics—Are large models being used where simpler ones would suffice?
- **Assumption hunting:** What assumptions were made? Were these explicit or implicit?
- **Detect GIGO:** Detect garbage in, garbage out (GIGO)—Was the uncertainty in the input artificially constrained to boost the model's certainty? Or, conversely, was it bloated to, for example, prevent regulators from making decisions in a case of harmful products?
- **Anticipate criticism:** Find sensitive assumptions before they find you—It is better to anticipate criticism by under-

TABLE 2 Case studies discussed in this article.

Case study	Energy policy issue	Method
Limitations in energy-related modeling		
International Institute of Applied Analysis (IIASA) global modeling	Political stance of the modeling activity implemented, uncertainty appraisal	Historic reference case
Improved approaches to modeling energy-related matters		
Energy modeling support for UK policy	Stochastic and epistemic uncertainty	NUSAP and diagnostic diagram
Negative emission technologies uptake in IAMs	Stochastic and epistemic uncertainty	NUSAP and (adapted) diagnostic diagram
Ecological footprint accounting	Uncertainty, scope of the model for decision-making	Sensitivity auditing

Abbreviations: numerical unit spread assessment pedigree, NUSAP; integrated assessment models, IAMs.

taking robust uncertainty and sensitivity analyses (UA and SA) before publishing one's results.

- **Aim for transparency:** Black box models do not play well in a public debate.
- **Do the right sum:** Do the right sums, not just the sums right—Is the issue properly identified or does the model address the “wrong” problem? Or is the model addressing a closed definition of what the problem might be instead of including multiple perspectives/stakeholders?
- **Perform UA and SA:** Perform thorough and state-of-the-art UA and SA.

The scrutiny and the analytical angle of a modeling activity within a policy study will depend on the engaged stakeholders and on the dialog developed with the modelers. Ideally, the SAUD checklist is to be applied recursively (Saltelli & Fun-towicz, 2014), over multiple rounds of iteration between the modelers and the engaged stakeholders, progressively adjusting the modeling exercise until it is considered suitable to contribute to the decision-making process. It can nonetheless also be used in an adversarial context with other stakeholders not directly involved in the deliberation process that would be interested in a deep scrutiny of the quantification/model adopted to inform the decision-making process (Fjelland, 2016; Saltelli et al., 2013).

Globally, several bodies working and the science–policy interface promoted the use of NUSAP and SAUD. For instance, SAUD has been recommended in guidelines for impact assessment, including those of the European Commission (2015b). The Science Advice for Policy by European Academies (2021) also recommended the use of both NUSAP and SAUD.

3 | CASE STUDIES

In this section, we discuss relevant issues in energy policy making along with the approaches proposed to tackle them in practical case studies. The full list of cases is shown in Table 2.

3.1 | Limitations in energy-related modeling: the cases of IIASA

3.1.1 | International Institute for Applied Systems Analysis (IIASA) global modeling

We selected the IIASA global modeling in the 1980s as historic reference case because it constitutes a major crisis in energy modeling for policy, and it triggered the development of quality auditing methods such as NUSAP and SAUD. The goal of the IIASA's Energy Project in the 1980s was to “understand the factual basis of the energy problem, i.e., to identify the facts and conditions for any energy policy” (Häfele et al., 1981). Table 3 summarizes the project's breadth and the magnitude of the IIASA energy modeling. The European Commission used the IIASA scenarios to develop a joint energy policy, which also played an important role in national decision-making committees (Häfele et al., 1981).

The IIASA energy model was built from a collection of interconnected submodels:

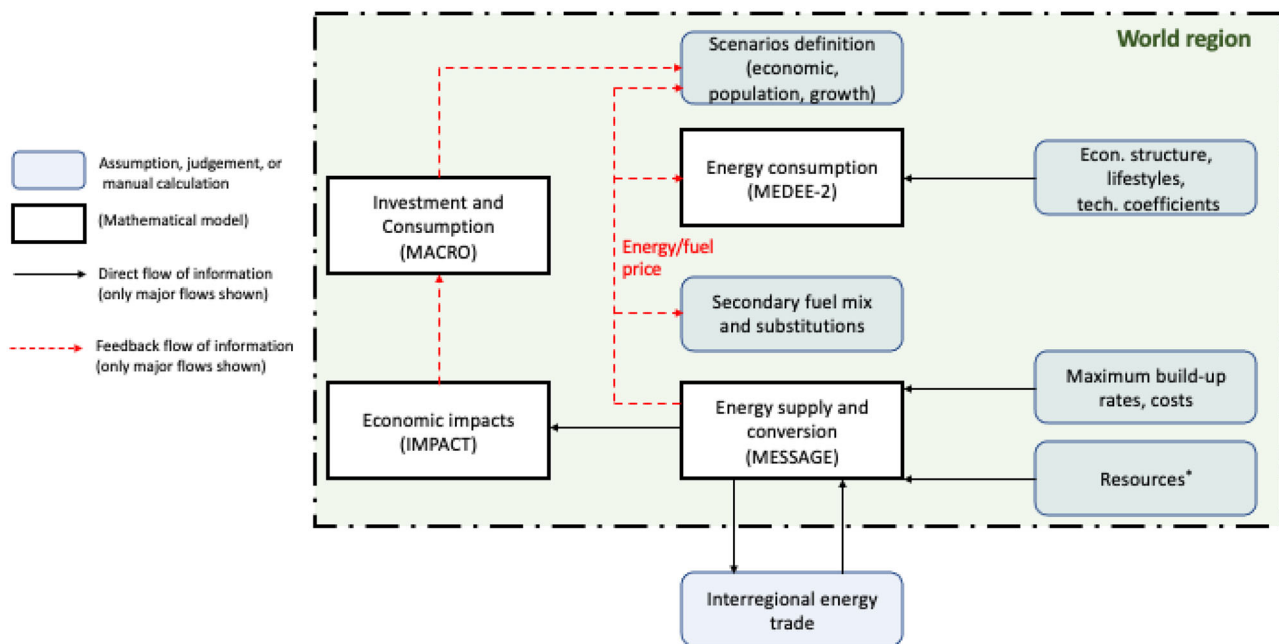
- MEDEE-2, an accounting framework-based energy demand model developed by the University of Grenoble.
- MESSAGE, a dynamic linear programming–based energy supply and conversion system model developed at IIASA.
- IMPACT, an input–output model for calculating the impacts of alternative energy scenario origins from the Siberian Power Institute.
- MACRO, a macroeconomic model developed in Canada and the USA.
- An oil trade gaming model produced by the Siberian Power Institute (Häfele et al., 1981; Keepin, 1984).

The IIASA's set of energy models is shown in Figure 1, along with the most relevant linkages among the many parts composing the system. Most of the feedback is provided by manual calculations: Changes to one set of inputs do not automatically propagate across models (Basile, 1980).

TABLE 3 Overview of International Institute for Applied Systems Analysis (IIASA) Energy System Program.

Duration	7 Years
Number of researchers involved	More than 140
Background of the research team group members	Economics, physics, engineering, geology, mathematics, psychology, psychiatry, and ethnology
Number of countries involved	19 (from developed and developing countries)
Number of partners	35
Institution involved	United Nations Environment Programme, International Atomic Energy Agency, National Center for Atmospheric Research, Electric Power Research Institute, Stanford Research Institute, Nuclear Research Center Karlsruhe, Institut Economique et Juridique de l'Énergie, Volkswagen Foundation, Federal Ministry of Research and Technology, Meteorological Office, National Coal Board, Austrian National Bank, and Siberian Power Institute
Number of research reports and conference proceedings	60

Source: Häfele et al. (1981).

**FIGURE 1** The International Institute for Applied Systems Analysis (IIASA)'s set of energy models. Source: Adapted from Basile (1980, p. 6).

We critically discuss the case study of the IIASA modeling activity by reflecting retrospectively under the guidance of relevant points of the SAUD checklist.

Among the significant problems encountered in the IIASA modeling activity was the failure of the feedback link between IMPACT and MEDEE-2. This meant that economic variables such as costs and resource extraction limits were not incorporated into the iteration process (Wynne, 1984). IMPACT was involved in only a few of the major "iterations" and was directed by a range of informal judgments (e.g., about future capital/output ratios).

This indicates that the models did not account for capital, land, labor, technology, water, and material investment. The basic assumptions of the IIASA analysis were that

the global population and average per capita income would have doubled by 2030, which may be doubtful as per point 2 of the SAUD checklist (assumption hunting). A practical implication of this scenario is that a sustainable future would only be possible through the expansion of adequate energy resources. Thereby, Häfele (1976) primarily focused on large-scale nuclear technology in developing future energy strategies as per the following criteria:

- CO₂ emissions reduction on a global scale by gradually substituting nuclear energy for fossil fuels.
- Maximum utilization of finite nuclear fuels via closed-loop recirculation and breeder technology, resulting in a virtually infinite source of energy.

- Diversification of resources with all energy sources within their optimal operating ranges.
- Secure energy supply with the energy mix base, reserve, and buffer capacity.
- The use and integration of “small-scale technologies” into the power grid with an ever-present nuclear power.
- Total CO₂ avoidance and substitution of fossil fuels utilizing coal gasification, methanation, and hydrogen technology.

In the end, the main conclusion of the IIASA's Energy Project was that the transition to fast-breeder reactors and large-scale solar and “coal synfuels” had to be made and could be achieved by 2030 if these power plants were largely deployed (Keepin, 1984). In the IIASA projections, the share of nuclear power was estimated to rise to 77% in 2030. In retrospect, we can observe that this forecast was considerably off the mark, but why was it so? By acknowledging possible sources of uncertainty, Keepin (1984) showed in a “sensitivity test” that minor changes to assumptions produced high effects on the results, which contradicted the “robust conclusions” of this modeling activity. Keepin developed an alternative scenario showing that an increase of nuclear costs by 16%, and of the coal extraction limit by 7%, would have resulted in phasing out the nuclear path. The model developers could have tested the ranges of these assumptions themselves, thus anticipating this possible criticism as per the fourth point of the SAUD checklist.

On an epistemic level, the IIASA modelers adopted from relational ecology the concept of resilience (Holling, 1973), where it was used to demonstrate that complex natural systems develop distinct strategies to respond to perturbations such as environmental change. To compute the shifting boundaries of points of equilibrium in the IIASA modeling activity, a reliable method for measuring potential fields and basins of attraction was required. Häfele (1976) saw resilience as a way to boost the credibility of his scenarios. He attempted to replicate Holling's resilience by dividing it into smaller “resilience basins.” However, Holling opposed the use of the resilience concept in nuclear energy due to the high risk and uncertainty of the technology entailed. As a result, Holling dismissed the Häfele team's proposals for massive nuclear parks, energy islands, and hydrogen pipelines as “bad science fiction,” that is, a forcing of one's desired option into another's concept, without recognizing the loss of information inherent in the process. Hulme (2011) defined this as “epistemological slippage.” This aspect resonates with the potential criticism about whether the model was “doing the right sum” as per the sixth point of the SAUD checklist. Holling also noted that one aspect of resilience is the avoidance of lock-ins and path-dependencies. On this score, resilience does not appear to favor nuclear energy.

The modeling exercise performed by IIASA is an illustration of how modeling can be used to attempt to influence the trajectories of technology (Keepin, 1984).

This IIASA modeling activity needs to be understood in the context of the 1970s energy crisis. Energy policy required accessible science, which IIASA set out to deliver. However, this case study shows that building a convincing modeling activity may require broadening the perspective from relying solely on one class of scientific experts and explore more possible options (Morgan & Keith, 2008); this would have opened and exposed the social and institutional assumptions embedded in the modeling activity as per the sixth point of the SAUD checklist.

3.2 | Improved approaches to modeling energy-related matters

3.2.1 | UK energy modeling as a support to policy-making

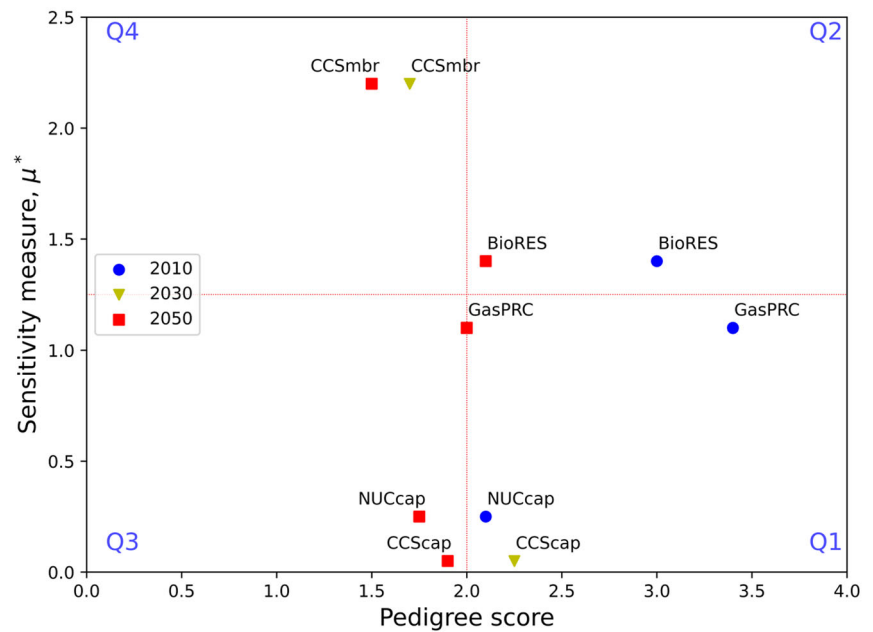
In the United Kingdom, scenario analysis using energy models has often suffered from deterministic thinking. A view of uncertainty has indeed been adopted, however, based on a storyline-and-simulation approach. Over the years, practitioners have increasingly realized that this is insufficient when dealing with complex and uncertain transitions that develop over relatively short timescales (Usher & Strachan, 2012). Based on ex post analysis of modeled energy futures, this approach has shown to be limited, with real-world developments completely outside of the anticipated range (Craig et al., 2002; Smil, 2008; Trutnevyte et al., 2016).

As a result, modeling practice in the United Kingdom has (to some extent) evolved and shifted toward a range of quantitative approaches to dealing with uncertainty, from probabilistic analysis (Pye et al., 2015) all the way to stochastic programming (Usher & Strachan, 2012) and modeling-to-generate-alternatives (Li & Trutnevyte, 2017; Trutnevyte, 2016). There is also recognition that the government requires more information on uncertainty, as outlined in the UK's Aqua Book on analytical quality assurance (HM Treasury, 2015).

Although these approaches push in the right direction, they are likely to overlook uncertainties that are not easily quantifiable (van der Sluijs, Risbey et al., 2005). These include the strength of the underlying knowledge base underpinning the modeling, or the degree to which the many assumptions made by modelers are value laden.

To broaden the assessment of uncertainties in energy modeling, the NUSAP approach was applied to a UK-based modeling exercise. This concerned ESME, a key model used for research informing UK government on energy issues (McGlade et al., 2018; Pye et al., 2015). The approach to the exercise was based on the following steps: (i) identify assumptions that affect the model results through global sensitivity analysis (GSA) and expert elicitation; (ii) determine criteria against which to assess pedigree; (iii) run the stakeholder workshop to generate the scores; and (iv) compare pedigree results to quantitative model results using a diagnostic diagram (Pye et al., 2018).

FIGURE 2 Diagnostic diagram to compare qualitative (aggregated pedigree scores, horizontal axis) against quantitative (sensitivity measure, vertical axis) uncertainties (yellow triangles: score/measure for the assumptions related to the model base year 2010; blue circles: score/measure for the assumptions related to the model base year for CSS, 2030; red square: score/measure for the assumptions related to the long-term future, 2050). The sensitivity measure (based on the elementary effects method [Saltelli et al., 2008] or Morris (1991) method) highlights the influence of the modeled uncertainty on the variance across the model objective function, which is the total discounted system costs.



In the ESME exercise, GSA was used to determine to which input parameters the model solution was most sensitive. In other words, if a policy-maker is looking for a cost-effective strategy, the sensitivity analysis seeks to identify the input assumptions that had the greatest influence on the costs of that strategy. The influence of different factors (representing quantitative uncertainty) was then plotted against the pedigree scores of those same assumptions using a diagnostic diagram (Funtowicz & Ravetz, 1990; van der Sluijs, Risbey et al., 2005). The pedigree scores were garnered in the stakeholder workshop by making use of the scoring criteria detailed in Table 1. Figure 2 shows the aggregate pedigree across the different categories.

Figure 2 indicates that some of the technology assumptions that are important for UK energy and climate policy have a weak aggregated pedigree. These land in Q4, a quadrant termed the “danger zone,” where assumptions have high sensitivity scores but weak pedigree. To this category belong bioenergy resource assumptions, which is crucial for biofuels in sectors such as international aviation and bioenergy with carbon capture and storage (BECCS), and carbon capture and storage (CCS) deployment (CCSmbr), which is again important for BECCS but also for hard-to-mitigate sectors such as iron and steel and cement. The value of such information for decision-makers is that they should proceed with caution when drawing policy conclusions from model solutions that rely heavily on bioenergy and CCS.

In addition to pedigree scoring, stakeholders were asked to score assumptions in terms of the extent to which an assumption was *justifiable and defensible*, and whether a specific assumption would likely find *agreement amongst peers*. The results highlighted that there are often many reasonable and possible choices for different assumptions. This emphasizes the need for transparency around modeling choices and a debate on, and scrutiny of, assumptions with broad stakehold-

ers' input. Given that previous modeling has been viewed as a black box, the process itself can enable critical scrutiny and contribute to the process of making energy systems analysis more transparent for decision-makers (Cao et al., 2016; Pfenninger, 2017).

3.2.2 | Negative emission technologies

Routes to meeting the targets of the 2015 Paris Agreement imply a commitment to reduce anthropogenic greenhouse gases emissions. To achieve this goal, two broad energy-technology approaches are considered: (i) real reduction of emissions from renewable energy technology (primary focus); and (ii) NET as an abatement of continued emissions. The use of NETs falls under a broader category of geoengineering, which are debated due to the potential unforeseen consequences on the environment, as well as reducing the collective commitment of society to environmental sustainability.

NETs can make the requirements regarding emission cuts less stringent by enhancing the planetary CO₂ sink capacity (van Vuuren et al., 2017). For instance, reductions of anthropogenic CO₂ emissions have been estimated in IAM activities at 60%–85% or 70%–95% for 2050, relative to 2010 figures, dependent on whether BECCS (a type of NET) is being deployed or not, respectively (van Vuuren et al., 2017). BECCS plays a prominent role in the NET literature (Anderson & Peters, 2016), to the point that large uptakes of BECCS have been posited in IPCC IAMs scenarios. These have been questioned in the literature on the basis of the following: first, there being too few existing plants (Babin et al., 2021); second, the capability of delivering negative emissions over the time span of the cultivations (Hanssen et al., 2020); third, the effectiveness of the coupled afforestation/reforestation strate-

gies (Krause et al., 2018; Turner et al., 2018); fourth, the important amount of land required (Field & Mach, 2017); and finally, the pace in increased land cultivation required to contribute to meeting the climate goals set for 2100 (Krause et al., 2018; Turner et al., 2018). Such a massive implementation could conflict with other fundamental sustainability goals such as food security and biodiversity conservation (Dooley et al., 2018).

The general picture that emerges from this criticism is that BECCS deployment would seemingly entail substantial stakes in return for very uncertain benefits. Presenting the outcome of these models with crisp figures results in leaving out other potential options. If uncertainties were acknowledged, other options may become comparable and worth investigating.

As regards the case of BECCS uptake in IAMs, Workman et al. (2020) identified the following points:

- The certainty on key assumptions (such as feasibility, cost, and deployment rates) over several decades was overestimated.
- Values beyond monetary proxies were excluded.
- Representatives of a single community define goals for climate policy rather than having these resulting from a dialog among multiple stakeholders.

Butnar et al. (2020) added criticisms on the scarce transparency of IAMs models relative to modeling assumptions, as well as on the treatment of the sociocultural and institutional dimension.

Public acceptance could be facilitated by opening the IAMs modeling activity to multi-criteria assessments, which are capable of including values beyond the monetary proxies used in cost optimizations (Stephens et al., 2021; Workman et al., 2020). However, this would require significant effort to update these modeling activities, their scope, and their theoretical background (Braunreiter et al., 2021).

Tavoni et al. (2017) investigated the potentially problematic nature of the underpinning assumptions of NETs uptake in IAMs. Based on a consultation with experts from the field, these authors identified as particularly difficult to modeling dimensions of governance, public acceptance, external costs, and impacts. By contrast, modeling was more propitious to appreciate the importance of operational costs and effectiveness.

In this highly debated context, Vaughan and Gough (2016) resorted to the NUSAP method by engaging 18 experts in a workshop to scrutinize several key assumptions fed into IAM models. The engaged experts identified nine key assumptions related to bioenergy (available land area, future yield, and proportion of energy), storage capacity, technology uptake, and capture rate (CCS), and cutting across several aspects (policy framework, social acceptability, and net negative emissions). The authors also made use of a diagnostic diagram whereby the pedigree score was assessed against a qualitatively estimated influence on results

through a dedicated pedigree score (in lieu of other sensitivity metrics). As in the case study on ESME, several of the discussed assumptions ended up in the danger zone of high influence on the result coupled to a weak pedigree score.

This case study illustrates how an issue in modeling activity identified in the literature can be brought to the fore and negotiated with relevant peers. The process enables mutual learning, while placing under the spotlight potential criticalities in the modeling activity.

3.2.3 | Ecological footprint

The EF is a successful sustainability indicator proposed by the Global Footprint Network. Diverse sources have advocated for its use as an indicator to lead energy policy making (Abbas et al., 2021; Metcalf, 2003). Energy consumption accounts for the most important part of the EF measure (Giampietro & Saltelli, 2014a). This part is bound to increase in the future because of BECCS deployment and further land allocation for energy uses. EF measures human demand on natural capital, which is understood as the quantity of natural land (expressed in global hectares equivalents) required to support an individual or economic activity. The “Earth overshoot day” is the date by which humanity will have used all available natural resources from the Earth’s yearly natural budget. The systematic anticipation of the Earth overshoot day over the years is widely recognized as a sign of humanity’s unsustainable pattern of economic development (Giampietro & Saltelli, 2014b). We make use here of the seven-point SAUD checklist, presented in the methods section, to evaluate whether EF is an adequate indicator to capture this aspect.

- Rhetorical use: According to Giampietro and Saltelli (2014a), the EF has been systematically overinterpreted in terms of representing the planet’s biocapacity. What is presented in the EF as a measure of what can be produced within the planet’s ecological limits is merely an accounting of agricultural productivity. Several other dimensions are excluded from EF accounting as per other points examined below.
- Assumption hunting: A potentially misleading feature in EF accounting concerns its bioenergy dimension. For instance, the question of how the CO₂ absorbing capacity decreases with the aging of forests is neglected. The same caveat applies to the paradox that replacing natural ecosystems with more productive human-made vegetation would lead to an improvement of the planet’s biocapacity rather than to an impoverishment due to a loss of biodiversity and natural habitats (Giampietro & Saltelli, 2014a).
- Detect GIGO: Several potential sources of uncertainty remain unaddressed in EF accounting. No error in terms of biocapacity is considered, nor is the variable of accuracy discussed at the local, national, and global levels. A data quality score is the only proxy included at the country

level. This leads to an issue in terms of how the information is processed and aggregated across scales, as examined in more detail below.

- Anticipate criticism: The rounding of values and the cascading of uncertainty across scales is one of the factors contributing to the fragility of EF accounting. To the best of our knowledge, this uncertainty has not been accounted for, let alone apportioned through sensitivity analysis, in the modeling adopted in EF accounting.
- Aim for transparency: The documentation on EF accounting is available, but some technical coefficients are not openly traceable. This is the case for the equivalence factors, which reflect the relative productivity of world average hectares of different types of land use. How these quantities were arrived can only be retrieved from a satellite workbook just available upon request.
- Do the right sum: The EF accounting does not help in defining whether types of land allocation actually contribute to sustainability (Galli et al., 2016). One example is that of a landfill, which is crucial on the waste sink side, but whose importance is entirely missed in the EF biocapacity accounting (Galli et al., 2016).
- Perform UA and SA: As previously discussed, uncertainty in the accounting is largely overlooked, except for a data quality score as proxy at the country level. Hence, the space of assumptions has not been explored, which leaves unaddressed the responsiveness of the EF indicator to its uncertainty sources.

The usefulness of SAUD stems from its capacity to highlight the limitations of quantification, in this case the EF, and therefore its suitability to practical policy-making problems and how the quantification could be improved to this end.

4 | CONCLUSIONS

A quantification, mathematical model, or any indicator can be thought of as a cathedral, an ancient *fabrica*, which is never finished; new bits are added or modified over time, bugs are solved, and new questions are posed. In this construction, choices are made all the time; these choices may concern the use of a physical or heuristic law, the value of variables, or which algorithm to choose among the many available to tackle a problem. We use the word “choice,” or “assumption,” as more than one possible item or value could be selected, but eventually just one enters the model construction. It is only normal that, with time, not even the model’s developers will be in condition to remember all that was chosen. This sedimentation of modeling assumptions, which enables the model to answer the questions asked from it, also constitutes an obstacle to its transparency. The only way out of this predicament is for modelers to make many choices, many assumptions, and propagate them through the model. Instead of a prediction, of a single point in the multidimensional space of model outputs, we now have a cloud.

This process of retracing one’s steps to rediscover the forgotten choices and assumptions, and to perform the analysis just described—which would appear technically as uncertainty quantification—is facilitated if one incorporates in the analysis the philosophies of NUSAP and SAUD that we illustrated in this contribution. Additionally, when engaging stakeholders, interfacing the two approaches in a diagnostic diagram can offer a thorough view of the uncertainty at play, as shown in this contribution’s case study of UK energy modeling for support to policy-making. It is also noteworthy that this process of going back in order to go forward plays an important role in system thinking (Koestler, 1989). This process, *the modeling of the modeling process*, is also valuable in unearthing path dependences and lock-ins, to identify stages where a given issue became frozen in a dead-alley configuration of conflict among stakeholders (e.g., large-scale development of NETs). Going forward, in turn, may take the form of a broadening of the spectrum of the policy options, as well as acknowledging the blind spots and limitations of a quantification.

In the present work, we have seen how several issues may impact quantification in energy systems in producing scenarios, indicators, and modeling activities. Untested assumptions and implicit political stances may result in assessments that are not “politically robust” (i.e., cannot be shared by all stakeholders), as shown in the case of the IIASA modeling activity with the controversy around proposing to largely resort to nuclear power as a means to achieve a resilient energy system.

Given the importance of uncertainty assessment in modeling for public policy, the degree to which model assumptions are value-laden and have “pedigree” needs to be considered by the community of modelers. Engaging stakeholders in workshops offers the chance to explore the logic, perspective, and framework that lies at the base of critical model assumptions, to demonstrate how to better generate trust in the analytical process, and to broaden expert input into the exercise, as in the UK energy policy making case studies. The renegotiation of the assumptions that inspired a move toward quantification can take place through direct interaction between the involved stakeholders, relevant experts, and policy-makers, ideally in a setting that allows experimentation with the socio-institutional roles “normally” entrusted to restricted communities of experts/regulators at the science–policy–society interface. All these aspects emerged from the ESME modeling activity, in which the group of experts/stakeholders/policy makers engaged in the workshop put under scrutiny the key modeling assumption. However, this approach could also be extended to SAUD with the joint deconstruction of a form of quantification. We have seen how the application of SAUD to the EF indicator highlighted its limits as a useful metric for sustainability. These limits may lead to potentially ill-conceived policy goals. The objection that these approaches are costly or encourage paralysis by analysis is unfounded for the case of energy policy, given the important stakes and long-time horizons covered by the policy. We invite the reader to imagine, for instance,

how the planning and construction of the Nord Stream 2 pipeline directly connecting Germany and Russia could have been developed differently in case more uncertainty and options were explored. Germany recently suspended the certification of the project upon the recent diplomatic tensions between Russia and Ukraine that resulted in the present war (at the time of writing). As an attempt to foster domestic energy security, was it impossible to explore the potential embitterment of international relations between the European Union and Russia in the modeling activities underpinning the project?

The applications of these methods to currently active topics such as the modeling of NETs in IAMs appear also promising. Having lawmakers demand the use of the analytical lenses suggested in the present work would help to strengthen the quantifications that underpin energy policy making, leading to tangible benefits for the overall policy-making process.

Finally, one should not forget that modeling is only “one of the voices around the table” when it comes to making policies. Local and experiential knowledge, historical insights and social values are other dimensions of prominent importance to take into account that modeling to inform policy should go hand-in-hand (Science Advice for Policy by European Academies, 2019; Saltelli et al., 2023).

ACKNOWLEDGMENTS

This research received funding from UK Research and Innovation grant agreement EP/R035288/1 as part of the Centre for Research into Energy Demand Solutions (CREDS). Andrea Saltelli acknowledges the funding of i4Driving, an EU Horizon Europe R&I project (Grant Agreement ID 101076165). Arnald Puy acknowledges funding by the UK Research and Innovation (UKRI) under the UK government’s Horizon Europe funding guarantee [grant number EP/Y02463X/1, project DAWN].

CONFLICT OF INTEREST STATEMENT

We declare that we have no known conflicts of interest or personal relationships that could have appeared to influence the work reported in this article.

DATA AVAILABILITY STATEMENT

No new data were created in the study.

ORCID

Samuele Lo Piano  <https://orcid.org/0000-0002-2625-483X>

REFERENCES

Abbas, S., Kousar, S., & Pervaiz, A. (2021). Effects of energy consumption and ecological footprint on CO₂ emissions: An empirical evidence from Pakistan. *Environment, Development and Sustainability*, 23(9), 13364–13381. <https://doi.org/10.1007/s10668-020-01216-9>

Anderson, K., & Peters, G. (2016). The trouble with negative emissions. *Science*, 354(6309), 182–183. <https://doi.org/10.1126/science.aah4567>

Babin, A., Vaneckhaute, C., & Iliuta, M. C. (2021). Potential and challenges of bioenergy with carbon capture and storage as a carbon-negative energy

source: A review. *Biomass and Bioenergy*, 146, 105968. <https://doi.org/10.1016/j.biombioe.2021.105968>

Basile, P. S. (1980). *The IIASA set of energy models: Its design and application*. IIASA.

Bistline, J., Blanford, G., Brown, M., Burtraw, D., Domeshek, M., Farbes, J., Fawcett, A., Hamilton, A., Jenkins, J., Jones, R., King, B., Kolus, H., Larsen, J., Levin, A., Mahajan, M., Marcy, C., Mayfield, E., McFarland, J., McJeon, H., ... Zhao, A. (2023). Emissions and energy impacts of the Inflation Reduction Act. *Science*, 380(6652), 1324–1327. <https://doi.org/10.1126/science.adg3781>

Braunreiter, L., van Beek, L., Hajer, M., & van Vuuren, D. (2021). Transformative pathways—Using integrated assessment models more effectively to open up plausible and desirable low-carbon futures. *Energy Research & Social Science*, 80, 102220. <https://doi.org/10.1016/j.erss.2021.102220>

Butnar, I., Li, P.-H., Strachan, N., Pereira, J. P., Gambhir, A., & Smith, P. (2020). A deep dive into the modelling assumptions for biomass with carbon capture and storage (BECCS): A transparency exercise. *Environmental Research Letters*, 15(8), 084008. <https://doi.org/10.1088/1748-9326/ab5c3e>

Cao, K.-K., Cebulla, F., Gómez Vilchez, J. J., Mousavi, B., & Prehofer, S. (2016). Raising awareness in model-based energy scenario studies—A transparency checklist. *Energy, Sustainability and Society*, 6(1), 28. <https://doi.org/10.1186/s13705-016-0090-z>

Clark, P. (November 7, 2011). Credibility of EU energy review questioned. Financial Times. Retrieved from <https://www.ft.com/content/9cf8f93e-0865-11e1-bc4d-00144feabdc0>

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.27.1.22001.083425>

Dooley, K., Christoff, P., & Nicholas, K. A. (2018). Co-producing climate policy and negative emissions: Trade-offs for sustainable land-use. *Global Sustainability*, 1, e3. <https://doi.org/10.1017/sus.2018.6>

European Commission. (2023). *Clean energy for all Europeans package*. European Commission. https://energy.ec.europa.eu/topics/energy-strategy/clean-energy-all-europeans-package_en

European Commission. (2015a). *Guidelines on impact assessment*. European Commission.

European Commission. (2015b). *Better regulation toolbox*. European Commission.

European Commission, Directorate-General for Climate Action, Directorate-General for Energy, Directorate-General for Mobility and Transport, De Vita, A., Capros, P., Paroussos, L. et al., EU reference scenario 2020 – Energy, transport and GHG emissions – Trends to 2050, Publications Office, 2021, <https://data.europa.eu/doi/10.2833/35750>

Field, C. B., & Mach, K. J. (2017). Rightsizing carbon dioxide removal. *Science*, 356(6339), 706–707. <https://doi.org/10.1126/science.aam9726>

Fjelland, R. (2016). When laypeople are right and experts are wrong: Lessons from love canal. *International Journal for Philosophy of Chemistry*, 22(1), 105–125.

Funtowicz, S., & Ravetz, J. R. (1990). *Uncertainty and quality in science for policy*. Kluwer.

Funtowicz, S., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, 25(7), 739–755. [https://doi.org/10.1016/0016-3287\(93\)90022-L](https://doi.org/10.1016/0016-3287(93)90022-L)

Galli, A., Giampietro, M., Goldfinger, S., Lazarus, E., Lin, D., Saltelli, A., Wackernagel, M., & Müller, F. (2016). Questioning the ecological footprint. *Ecological Indicators*, 69, 224–232. <https://doi.org/10.1016/j.ecolind.2016.04.014>

Giampietro, M., & Saltelli, A. (2014a). Footprints to nowhere. *Ecological Indicators*, 46, 610–621. <https://doi.org/10.1016/j.ecolind.2014.01.030>

Giampietro, M., & Saltelli, A. (2014b). Footworking in circles. *Ecological Indicators*, 46, 260–263. <https://doi.org/10.1016/j.ecolind.2014.06.019>

Häfele, W. (1976). *Second status report of the IIASA project on energy systems* [RR-76-001]. IIASA.

Häfele, W., Anderer, J., McDonald, A., & Nakicenovic, N. (1981). *Energy in a finite world: Paths to a sustainable future* (Vol. 1). Ballinger.

Hanssen, S. V., Daioglou, V., Steinmann, Z. J. N., Doelman, J. C., Van Vuuren, D. P., & Huijbregts, M. A. J. (2020). The climate change mit-

- igation potential of bioenergy with carbon capture and storage. *Nature Climate Change*, 10(11), 1023–1029. <https://doi.org/10.1038/s41558-020-0885-y>
- HM Treasury. (2015). *The aqua book: Guidance on producing quality analysis for government*. HM Treasury.
- Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4(1), 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
- Hulme, M. (2011). Reducing the future to climate: A story of climate determinism and reductionism. *Osiris*, 26(1), 245–266. <https://doi.org/10.1086/661274>
- 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](https://doi.org/10.1016/0305-0483(82)90002-0)
- Keepin, B. (1984). A technical appraisal of the IASA energy scenarios. *Policy Sciences*, 17(3), 199–276. <https://doi.org/10.1007/BF00138708>
- Koestler, A. (1989). *The ghost in the machine*. Arkana.
- Koomey, J., Craig, P., Gadgil, A., & Lorenzetti, D. (2003). Improving long-range energy modeling: A plea for historical retrospectives. *The Energy Journal*, 24(4), 75–92. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol24-No4-4>
- Krause, A., Pugh, T. A. M., Bayer, A. D., Li, W., Leung, F., Bondeau, A., Doelman, J. C., Humpenöder, F., Anthoni, P., Bodirsky, B. L., Ciais, P., Müller, C., Murray-Tortarolo, G., Olin, S., Popp, A., Sitch, S., Stehfest, E., & Arneth, A. (2018). Large uncertainty in carbon uptake potential of land-based climate-change mitigation efforts. *Global Change Biology*, 24(7), 3025–3038. <https://doi.org/10.1111/gcb.14144>
- Li, F. G. N., & Trutnevyte, E. (2017). Investment appraisal of cost-optimal and near-optimal pathways for the UK electricity sector transition to 2050. *Applied Energy*, 189, 89–109.
- Lo Piano, S., & Benini, L. (2022). A critical perspective on uncertainty appraisal and sensitivity analysis in life cycle assessment. *Journal of Industrial Ecology*, 26(3), 763–781. <https://doi.org/10.1111/jiec.13237>
- Lo Piano, S., Sheikholeslami, R., Puy, A., & Saltelli, A. (2022). Unpacking the modelling process via sensitivity auditing. *Futures*, 144, 103041. <https://doi.org/10.1016/j.futures.2022.103041>
- Lo Piano, S., Sheikholeslami, R., Puy, A., & Saltelli, A. (2023). Sensitivity auditing: A practical checklist for auditing decision-relevant models. In M. Di Fiore & A. Saltelli (Eds.), *The politics of modelling: Numbers between science and policy*. Oxford University Press.
- Di Fiore, M., & Saltelli, A. (Eds.). (2023). *The politics of modelling*. Oxford University Press.
- McGlade, C., Pye, S., Ekins, P., Bradshaw, M., & Watson, J. (2018). The future role of natural gas in the UK: A bridge to nowhere? *Energy Policy*, 113, 454–465. <https://doi.org/10.1016/j.enpol.2017.11.022>
- Metcalfe, G. (2003). *The ecological footprint of energy*. SPUR.
- Morgan, M. G., & Keith, D. W. (2008). Improving the way we think about projecting future energy use and emissions of carbon dioxide. *Climatic Change*, 90(3), 189–215. <https://doi.org/10.1007/s10584-008-9458-1>
- Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2), 161–174. <https://doi.org/10.1080/00401706.1991.10484804>
- Muller, J. Z. (2018). *The tyranny of metrics*. Princeton University Press.
- O'Neill, R. V. (1989). Perspectives in hierarchy and scale. In J., Roughgarden, R. M., May, & S. A., Levin (Eds.), *Perspectives in ecological theory*. Princeton University Press.
- Petersen, A., Janssen, P., van der Sluijs, J. P., Risbey, J. S., Ravetz, J., Wardekker, A., & Hughes, H. (2013). *Guidance for uncertainty assessment and communication*. PBL Netherlands Environmental Assessment Agency.
- Pfenninger, S. (2017). Energy scientists must show their workings. *Nature*, 542(7642), 393. <https://doi.org/10.1038/542393a>
- Puy, A., Beneventano, P., Levin, S. A., Lo Piano, S., Portaluri, T., & Saltelli, A. (2022). Models with higher effective dimensions tend to produce more uncertain estimates. *Science Advances*, 8(42), 1–10. <https://doi.org/10.1126/sciadv.abn9450>
- 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>
- Pye, S., Sabio, N., & Strachan, N. (2015). An integrated systematic analysis of uncertainties in UK energy transition pathways. *Energy Policy*, 87, 673–684. <https://doi.org/10.1016/j.enpol.2014.12.031>
- Rosen, R. (1991). *Life itself: A comprehensive inquiry into the nature, origin, and fabrication of life*. Columbia University Press.
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., & Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, 114, 29–39. <https://doi.org/10.1016/j.envsoft.2019.01.012>
- Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D., Pielke, R. Jr., Portaluri, T., Porter, T. M., Puy, A., Rafols, I., Ravetz, J. R., Reinert, E. S., Sarewitz, D., Stark, P. B., ... Paolo, J. P. V. (2020). Five ways to ensure that models serve society: A manifesto. *Nature*, 582, 482–484. <https://doi.org/10.1038/d41586-020-01812-9>
- Saltelli, A., Benini, L., Funtowicz, S., Giampietro, M., Kaiser, M., Reinert, E. S., & van der Sluijs, J. P. (2020). The technique is never neutral. How methodological choices condition the generation of narratives for sustainability. *Environmental Science and Policy*, 106, 87–98.
- Saltelli, A., & Funtowicz, S. (2014). When all models are wrong. *Issues in Science and Technology*, 30(2), 79–85.
- Saltelli, A., Guimaraes Pereira, Â., van der Sluijs, J. P., & Funtowicz, S. (2013). What do I make of your latinorum? Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy*, 9(2/3/4), 213–234. <https://doi.org/10.1504/IJFIP.2013.058610>
- Saltelli, A., Kuc-Czarnecka, M., Piano, S. L., Lőrincz, M. J., Olczyk, M., Puy, A., Reinert, E., Smith, S. T., & van der Sluijs, J. P. (2023). Impact assessment culture in the European Union. Time for something new? *Environmental Science & Policy*, 142, 99–111. <https://doi.org/10.1016/j.envsci.2023.02.005>
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis: The primer*. Wiley.
- Science Advice for Policy by European Academies. (2019). *Making sense of science for policy under conditions of complexity and uncertainty*. Science Advice for Policy by European Academies.
- Smil, V. (2008). Long-range energy forecasts are no more than fairy tales. *Nature*, 453(7192), 154–154. <https://doi.org/10.1038/453154a>
- Stephens, J. C., Kashwan, P., McLaren, D., & Surprise, K. (2021). The risks of solar geoengineering research. *Science*, 372(6547), 1161–1161. <https://doi.org/10.1126/science.abj3679>
- Tavoni, M., Bosetti, V., Shayegh, S., Drouet, L., Emmerling, J., Fuss, S., Goeschl, T., Guivarch, C., Lontzek, T. S., Manoussi, V., Moreno-Cruz, J., Muri, H., Quaas, M., & Rickels, W. (2017). *Challenges and opportunities for integrated modeling of climate engineering*. Fondazione Eni Enrico Mattei (FEEM).
- Trutnevyte, E. (2016). Does cost optimization approximate the real-world energy transition? *Energy*, 106, 182–193. <https://doi.org/10.1016/j.energy.2016.03.038>
- Trutnevyte, E., McDowall, W., Tomei, J., & Keppo, I. (2016). Energy scenario choices: Insights from a retrospective review of UK energy futures. *Renewable and Sustainable Energy Reviews*, 55, 326–337. <https://doi.org/10.1016/j.rser.2015.10.067>
- Turner, M. G., & Gardner, R. H. (2015). Introduction to Models. In *Landscape ecology in theory and practice* (pp. 63–95). Springer New York.
- Turner, P. A., Field, C. B., Lobell, D. B., Sanchez, D. L., & Mach, K. J. (2018). Unprecedented rates of land-use transformation in modelled climate change mitigation pathways. *Nature Sustainability*, 1(5), 240–245. <https://doi.org/10.1038/s41893-018-0063-7>
- Usher, W., & Strachan, N. (2012). Critical mid-term uncertainties in long-term decarbonisation pathways. *Energy Policy*, 41, 433–444. <https://doi.org/10.1016/j.enpol.2011.11.004>
- van der Sluijs, J. P. (2017). The NUSAP approach to uncertainty appraisal and communication. In *Routledge handbook of ecological economics*, 301–310. Routledge.

- van der Sluijs, J. P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J. R., & 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>
- van der Sluijs, J. P., & Petersen, A. (2008). Exploring the quality of evidence for complex and contested policy decisions. *Environmental Research Letters*, 3(024008), 1–9.
- van der Sluijs, J. P., Risbey, J. S., & Ravetz, J. R. (2005). Uncertainty assessment of Voc emissions from Paint in the Netherlands using the NUSAP system. *Environmental Monitoring and Assessment*, 105(1–3), 229–259. <https://doi.org/10.1007/s10661-005-3697-7>
- van Vuuren, D. P., Hof, A. F., van Sluisveld, M. A. E., & Riahi, K. (2017). Open discussion of negative emissions is urgently needed. *Nature Energy*, 2(12), 902–904. <https://doi.org/10.1038/s41560-017-0055-2>
- Vaughan, N. E., & Gough, C. (2016). Expert assessment concludes negative emissions scenarios may not deliver. *Environmental Research Letters*, 11(9), 095003. <https://doi.org/10.1088/1748-9326/11/9/095003>
- Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Kraayer 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>
- Weyant, J. P., & Olavson, T. (1999). Issues in modeling induced technological change in energy, environmental, and climate policy. *Environmental Modeling & Assessment*, 4(2), 67–85. <https://doi.org/10.1023/A:1019012317160>
- Williams, M. L., Beevers, S., Kitwiroon, N., Dajnak, D., Walton, H., Lott, M. C., Pye, S., Fecht, D., Toledano, M. B., & Holland, M. (2018). *The energy systems model and future scenarios*. NIHR Journals Library.
- Workman, M., Dooley, K., Lomax, G., Maltby, J., & Darch, G. (2020). Decision making in contexts of deep uncertainty—An alternative approach for long-term climate policy. *Environmental Science & Policy*, 103, 77–84. <https://doi.org/10.1016/j.envsci.2019.10.002>
- Wynne, B. (1984). The institutional context of science, models, and policy: The IIASA energy study. *Policy Sciences*, 17(3), 277–320. <https://doi.org/10.1007/BF00138709>
- 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>

How to cite this article: Lo Piano, S., Lőrincz, M. J., Puy, A., Pye, S., Saltelli, A., Smith, S. T., & van der Sluijs, J. (2023). Unpacking the modeling process for energy policy making. *Risk Analysis*, 1–12. <https://doi.org/10.1111/risa.14248>