On the consensus in climate policy scenarios

Mark Dekker (✉ mark.dekker@pbl.nl)

PBL Netherlands Environmental Assessment Agency

Andries Hof

Netherlands Environmental Assessment Agency   https://orcid.org/0000-0002-7568-5038

Maarten van den Berg

PBL Netherlands Environmental Assessment Agency

Vassilis Daioglou

PBL Netherlands Environmental Assessment Agency   https://orcid.org/0000-0002-6028-352X

Rik van Heerden

PBL Netherlands Environmental Assessment Agency

Kaj-Ivar van der Wijst

Utrecht University

Detlef van Vuuren

PBL Netherlands Environmental Assessment Agency   https://orcid.org/0000-0003-0398-2831

On the consensus in climate policy scenarios

Mark Dekker[a,b,*], Andries Hof[a,b], Maarten van den Berg[a], Vassilis Daioglou[a,b], Rik van Heerden[a], Kaj-Ivar van der Wijst[a,b], Detlef van Vuuren[a,b]

[a] PBL Netherlands Environmental Assessment Agency, Den Haag, Netherlands
[b] Copernicus Institute of Sustainable Development, Utrecht Universiteit, Utrecht, Netherlands
[*] Corresponding author: mark.dekker@pbl.nl

Analysis of large scenario databases has become a critical tool for identifying climate mitigation strategies, as shown in the latest IPCC report[1]. However, key elements of these strategies differ significantly among scenarios. Possible reasons include differences in climate target, models used, and assumptions on behavioral, technological and socio-economic developments[2-4]. For policymaking, it is important to know which of these factors are the main cause of the spread, but quantification of this is still missing. Here, we aim to identify consensus in climate policy scenarios by analyzing how much of the variance in scenario outcomes can be explained by these factors, using Sobol decomposition[5]. Some results, e.g. concerning future use of fossil and renewable resources, are mainly determined by climate target, while others, such as the composition of different renewables in the electricity mix and key outcomes of end-use sectors are more model dependent. Scenario aspects beyond model and climate outcome determine only a limited part of the variation, e.g., in nuclear power and hydrogen use, which suggests the need of more cross-model scenario variation. The outcomes put mitigation strategies in a new perspective by identifying which findings on the energy transition are robust and reveal key areas for future scenario development, model improvement and research.

In recent reports of the IPCC (Intergovernmental Panel on Climate Change), model projections play a key role and are embedded in large databases: the CMIP6 database[6] with climate physics projections and the AR6 database[7] with climate policy scenarios. These databases show a wide range of potential climate and energy futures. A core question is to which extent this spread is associated with climate policy choices and resulting global warming. In physical sciences, attribution of projections to different levels of global warming is a major topic[7] and while the topic is also important for mitigation literature[8] – for instance, to inform policymakers on how the electricity mix differs between a 1.5 °C world and a 3 °C world – a comprehensive overview of robustness in this field is still missing.

Three main drivers of the spread in climate policy scenarios can be distinguished: climate targets, model characteristics and scenario assumptions. *Climate targets (*or more precisely *climate outcomes)* are an obvious driver: an energy system that achieves specific climate goals (such as the Paris Agreement) differs significantly from a system that does not. IPCC's WGIII of the 6th assessment report (AR6)[1] indicates that many key energy variables correlate with climate outcome. The approximately 1200 scenarios in the AR6 database are labelled with categories ranging from C1 (below 1.5 °C temperature change in 2100 with limited or no temperature overshoot) to C8 (above 4 °C in 2100). *Model differences* also cause spread in scenario outcomes[9]. Not only do models suffer from parametric uncertainties in the estimations of processes such as technology learning rates[10], model differences are also caused by fundamental structural differences[11,12] associated with model type (e.g., general versus partial equilibrium) or the role of cost-optimality. Finally, beyond climate outcome and model characteristics, there are several *scenario assumptions* on technology, socio-economic development, societal preferences, and mitigation strategies

that also influence the spread of energy futures. Notable examples are the timing of mitigation, the use of carbon capture and storage (CCS) and lifestyle changes.

Even though there is growing attention to multi-model comparisons[13,14], previous work on consensus in mitigation strategies has been conducted mainly within closed diagnostic experiments with a confined selection of variables (e.g., emission pathways or specific technologies such as bioenergy[13,14]), models[2], scenarios[15,16] or regions – complicating robustness analysis and potentially yielding contradictions upon comparison. For example, emission projections in some ensembles are found to be most sensitive to model choice[17], while in other ensembles, having a different focus or model set, emission projection variations across climate outcomes are found to supersede such model differences[18]. In addition, while the assessment of statistical significance of an outcome is common practice in multi-model studies[16,19], the quantification of the relative impacts of different drivers in determining the observed spread is not – which yields a more detailed perspective on agreement and uncertainty. Hence, in current literature, a comprehensive and quantitative analysis of consensus in mitigation strategies is still missing despite the strong reliance on these scenarios in IPCC reports.

In this paper, we fill this gap by identifying the *consensus* across climate policy scenarios: aspects of the energy transition that are robust and less driven by model differences and scenario assumptions. We do so by quantifying the contribution of the three drivers of the observed spread in mitigation strategies for a wide range of variables. These factors are expressed in three indices reflecting the proportion of the spread explained by them, using Sobol's method of variance decomposition[5,20]. Because of our focus on achieving the Paris agreement goals, we analyze variables associated with the energy transition: greenhouse gas emissions, the total energy mix, the primary energy mix of electricity generation and the energy mix of end-use sectors (i.e., transport, industry, and the residential and commercial sector).

Electricity generation

In 2020, the power sector accounted for approximately 20% of worldwide final energy consumption and 40% of global CO_2 emissions[21]. A rapid shift in electricity generation is crucial to achieve the Paris climate goals: in most climate policy scenarios, electrification of end-use sectors is commonly regarded as a key strategy and the power sector reaches net-zero emissions before other sectors[22]. Different technologies aid in achieving net-zero emissions: fossil-fuel electricity generation plants can be equipped with CCS, replaced by intermittent renewable technologies like solar and wind, or replaced by nuclear, hydropower and biomass[23,24] . The use of biomass – when combined with CCS – can potentially lead to negative emissions via the permanent sequestration of biogenic carbon[25]. The optimal mix of technologies not only depends on assumptions regarding future costs and potential of these technologies, but also on concerns surrounding reliability, and assumed projections of energy demand and concomitant technology preferences.

Fig. 1a shows the degree to which the variance of key electricity generation variables can be explained by climate outcome, model, and other scenario assumptions, in each 5-year increment between 2030 and 2100. Data points in the upper corner of panel (a) denote variables of which spread is determined by climate outcomes, meaning that model differences have a weaker effect: this means that there is sufficient consensus on these variables to interpret them as robust ingredients of mitigation. While many of the technologies are mostly model-dependent (lower-left), there is a robust outcome that the use of coal and gas without CCS is robustly linked with climate target. Indeed, for climate outcomes of C3 and

lower (less than 2 °C temperature increase in 2100 with 67% probability and higher), coal use without CCS will be phased out in practically all scenarios.

The use of renewables is a robust part of mitigation strategies up to 2050. Later in the century, most scenarios have high shares of renewables (see Fig. S5 in SI B), but the exact value is more determined by model differences than the stringency of the climate target. Note that not only the absolute renewable deployment differs greatly among models, but also their fraction of the total does (see SI B.2). Interestingly, the rollout of individual renewable technologies varies mostly between models even before 2050. These model differences are associated with assumptions on technology costs, potentials, overall energy demand and the integration of renewables into energy systems. Only in the beginning of the projection periods (around 2030), values of solar and wind power are robust across models, while later, when the volume differences of these renewables expand significantly and competition intensifies, their projections become more model dependent.

Electricity from biomass (without CCS), nuclear and hydropower is already in 2030 mostly model dependent[26]. While in most models, nuclear power increases with climate ambition, large model differences remain in estimating costs and perspectives on nuclear risk factors[19,27]. Hydropower, being one of the cheaper renewable electricity technologies, is already used to (close to) its maximum potential in baseline scenarios with limited mitigation, making the spread mainly driven by uncertainties around potentials[26] (see panel (d)).

Compared to renewable technologies, the overall variance in CCS technologies from coal, gas and biomass (BECCS) is large and generally driven by model differences caused by competition between CCS technologies and highly different estimates of total CCS use in general. For BECCS, the spread is also driven for a significant part by climate outcome, reflecting some consistency among models on its importance in reaching climate goals. Its slowly decaying climate dependence later in the century reflects that the final BECCS use reaches maximum biomass potentials, yielding model-dependency in the level of these potentials and suppressing increase with increasing climate ambition (panel (c)).
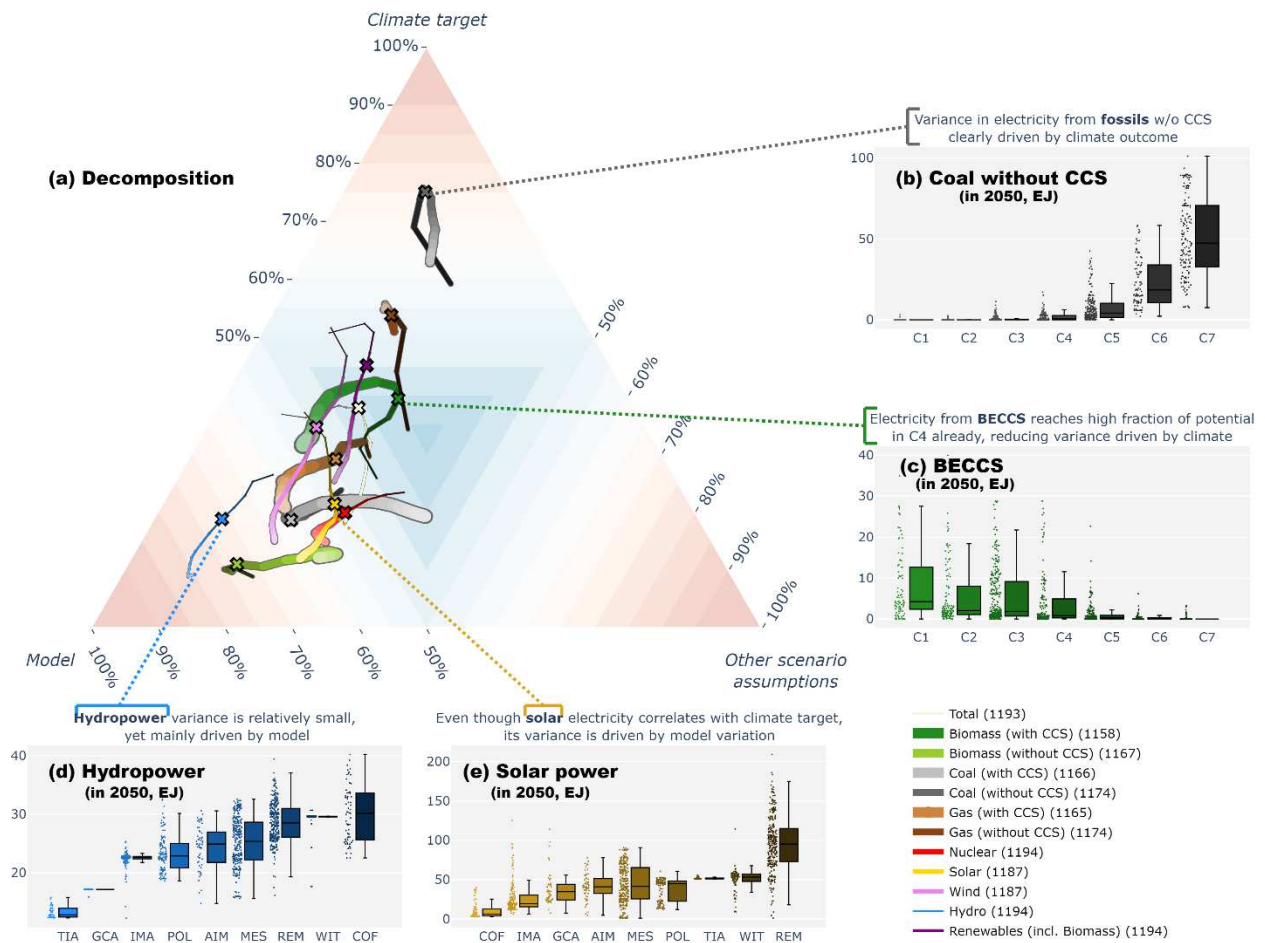
*Figure 1. **Panel (a)**: Variance decomposition of electricity generation variables (secondary energy): fraction of variance explained by differences in climate outcome (upper corner), models (lower-left corner) and other scenario assumptions (lower-right corner). The position of datapoints denotes the strength of each driver in the variable's spread: points in the upper corner of panel (a) indicate a robust link with climate outcome and that the variable's spread is less affected by model differences and other scenario assumptions. Lines indicate values between 2030 and 2100, with 5-year increments, where time is indicated in color shades (from dark to light shades). For each variable, the year 2050 is marked with a cross-sign. Line width reflect coefficient of variation (standard deviation divided by mean) of the increment with respect to any other increment in the figure. Number of scenario entries used for each variable is displayed in legend. **Panels (b-e)**: results split by model (abbreviated, see SI for reference) or climate category in 2050 in EJ for: (b) electricity from coal, (c) electricity from biomass with CCS, (d) hydropower and (e) solar power. Boxplots indicate quartiles, individual scenario entries are shown in dots. Interactive versions of all figures can be found as supplementary material to this paper. The same figure using fractions of the total electricity generation (instead of absolute values) in the variance decomposition is shown in Fig. S2 in SI B.*

Transport energy demand

The transport sector accounts for around 20% of global CO_2 emissions and 25% of total final energy consumption in 2020[21]. Like electricity, this sector also plays a key role in reaching net-zero emissions[28] and many routes exist to decarbonize the transport sector[29-31]. Fig. 2 shows the variance decomposition of transport energy carriers.

In 2020, the main energy carrier in transport was oil, accounting for over 90% of global final energy consumption[21]. Figs. 2a-b show that the future spread is dominated by climate outcomes. In other words, there is strong consensus that the oil will be replaced in scenarios with stringent climate targets. However,

there is less consensus on how it is replaced. There are three main substitutes for oil in transport: electricity, hydrogen[21,32] and bioenergy.

Although electricity in transport has a significant climate-dependency early in the century (2030-2035), model differences dominate. This is especially true later in the century. The key reason is that in several models electrification of transport also happens without stringent climate targets due its projected increasing competitiveness (see Fig. S6 in SI B), making model differences relatively larger.

A similar conclusion can be drawn for hydrogen. Estimates for hydrogen use in transport vary greatly, marked by the large marker size and visible in panel (e) across different models in 2050. Observing the wide and for some models dichotomous distributions of hydrogen use in transport, this can partially be explained by economics-of-scale for transport technologies[33]: either hydrogen rolls out significantly (at levels different across the models), or it does only very little and other technologies take over. These model differences are not superseded by differences across climate outcomes – in fact, even among the most ambitious scenarios (C1), multiple models project no hydrogen use at all in 2100 (see Fig. S7 in SI B) – yielding no overall robustness on hydrogen use in mitigation. In addition, hydrogen is sensitive to specific scenario assumptions, marked by the relatively right-ward position of the 2050-value. Possible explanations could be changes in the role of hydrogen in different versions of the same model and possible assumptions on the development of hydrogen and electricity infrastructure development.

Out of the three main substitutes for oil in transport, bioenergy is the most model-dependent in 2050 (panel (d)). Currently comprising only 3.5% of the transport energy use[21], it starts off in 2030 with relatively low variance, increasing over time and moving in Fig. 2a far into the lower-left corner. Its model-dependency is exemplified in the fact that only seven out of nine models report bioenergy in transport at all and that one particular model has a 4-40 times higher median bioenergy use in transport than other models in 2050, as observed in previous work[34] (see Fig. S8 in SI B). The high dependance of bioenergy deployment on model structure is in line with previous literature[13,14,34] and is also observed for other sectors (see below). Potential bioenergy supply and use is very heterogenous, with multiple possible feedstocks, conversion routes, and the wide range of possible end uses.
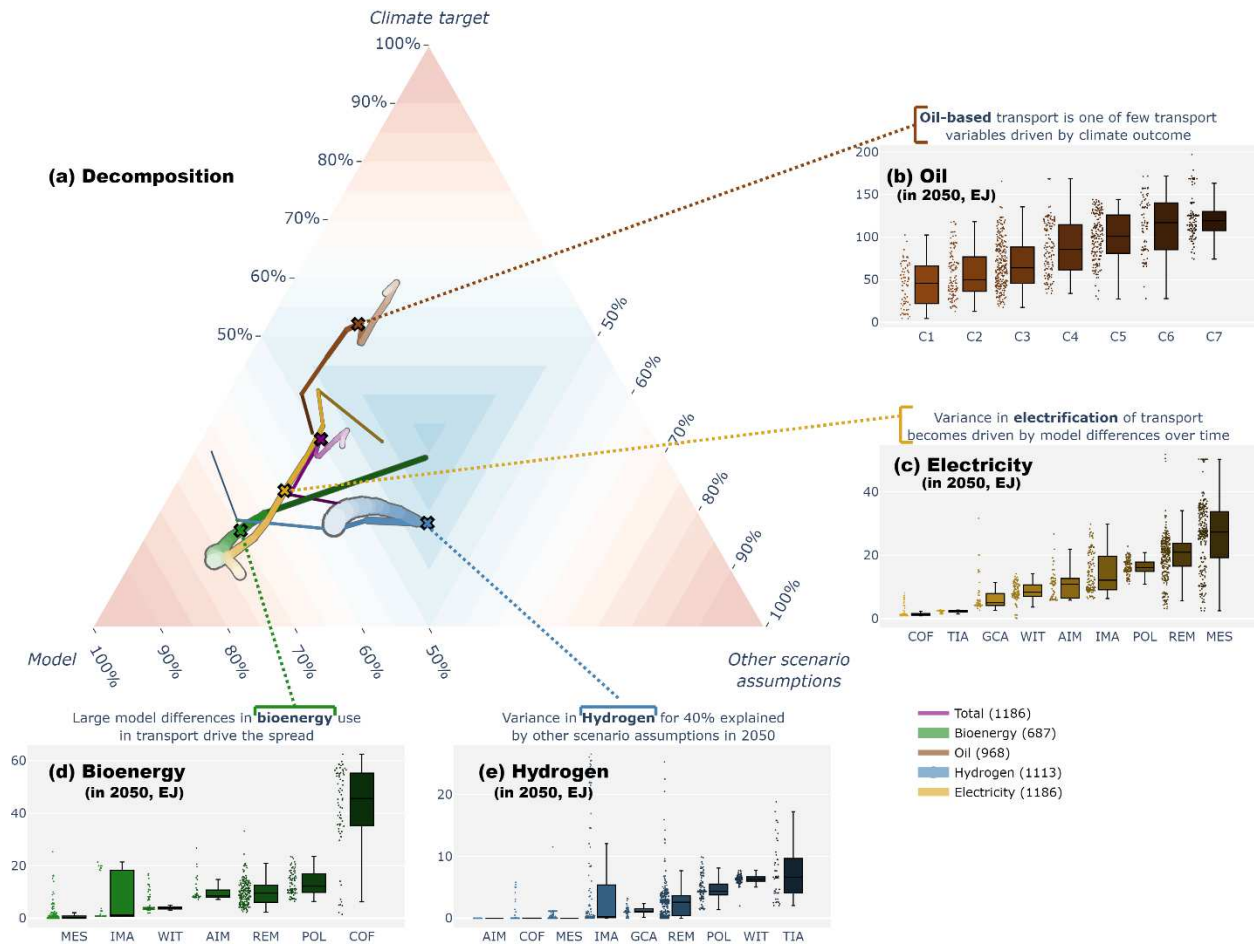
Figure 2. **Panel (a)**: *similar to Figure 1a, but for energy use in the transport sector. Line widths are on the same scale as line widths in Figure 1.* **Panels (b-e)**: *similar to Figure 1b-e, for (b) oil use in transport, (c) electricity use in transport, (d) bioenergy use in transport and (e) hydrogen use in transport.*

Overview: key variables in 2050

We extend the analysis to a larger group of key variables of the energy transition in Fig. 3, split into six categories: emissions (pink-cloud), primary energy (green-sun), the industry sector (brown-cog), the transport sector (teal-truck), the building sector (grey-house) and electricity generation (yellow-flash).

Clear differences can be recognized between the six groups of variables displayed in Fig. 3. In general, variance in emissions is dominated by climate outcomes, reflecting that emission reduction is a robust aspect of mitigation strategies. This can be expected for Kyoto gases and CO_2, but this is also found for methane (CH_4). In contrast, nitrous oxide (N_2O) is more model-dependent: while each individual model shows lower N_2O emissions in deep climate ambition scenarios, the disagreement on the absolute level supersedes that signal. Most sectoral emissions levels are also determined by climate outcomes, with the weakest signal in the buildings sector (48%). For primary energy and electricity generation, the use of most fossil fuels without CCS is also identified as robustly related to climate outcomes: in addition to overall use, also oil in transport, gas in buildings, and to a lesser extent, coal and gas in industry. The same is true for the total use of renewables, including biomass.

In contrast, spread in individual renewable technologies and CCS technologies are mostly driven by model differences, with biomass being neither clearly model nor climate dependent. However, when looking at *fractions* rather than absolute values of the primary energy sources (shown in Fig. S2 in SI B), biomass, wind and hydro energy turn out to be much more robustly quantified across climate outcomes. A key reason for the difference in robustness of results between the role of renewables in primary energy and the electricity mix, respectively, is that the former is reduced in more ambitious scenarios (as a result of increased efficiency). In contrast, total electricity demand increases because of electrification of end-use sectors. Although renewables use increases with climate ambition in both the primary energy and the electricity mix, this effect is stronger in a relative sense and more robustly linked with climate ambition in primary energy use than in electricity generation.

While end-use sector emissions and fossil use are climate-driven, the projection spread of non-fossil fuel compositions in these sectors are significantly affected by model differences. In particular, the spread in electricity use in all three sectors is for approximately 60% explained by model differences (see markings in Fig. 3), which is partly associated with differences in projecting the *total* energy consumption in these sectors[35]. However, the same analysis on electricity as a fraction of the total energy consumption per sector shows a much lower model dependency (down to approximately 40%, see Fig. S4 in SI B.2). Even more consensus is identified for the total electricity use as a fraction of the total final energy (not shown, peaking in 2050 with 53%), as well as total final energy itself (49% in 2050).

Figure 3 also reveals that the variance in most variables is dominated by differences in either climate outcome or model. Other scenario assumptions play a varying role but rarely dominate the variance. Example variables for which they do play a (still limited but) relatively large role are associated with specific on/off assumptions in models (e.g., hydrogen in transport and industry), lifestyle and policy-sensitive variables (e.g., total energy consumption in the residential and commercial sector or nuclear energy).
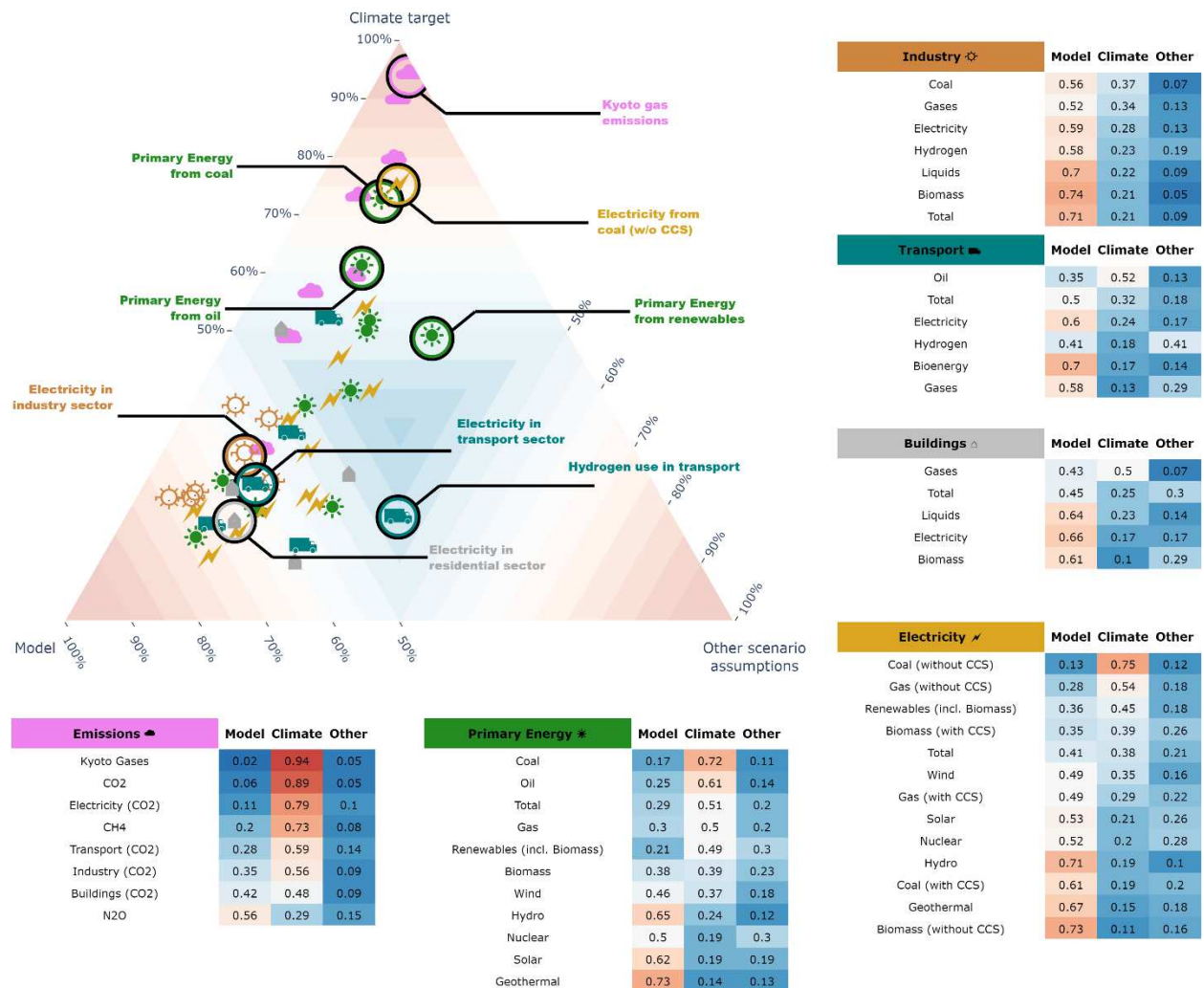
Industry ☼	Model	Climate	Other
Coal	0.56	0.37	0.07
Gases	0.52	0.34	0.13
Electricity	0.59	0.28	0.13
Hydrogen	0.58	0.23	0.19
Liquids	0.7	0.22	0.09
Biomass	0.74	0.21	0.05
Total	0.71	0.21	0.09

Transport 🚌	Model	Climate	Other
Oil	0.35	0.52	0.13
Total	0.5	0.32	0.18
Electricity	0.6	0.24	0.17
Hydrogen	0.41	0.18	0.41
Bioenergy	0.7	0.17	0.14
Gases	0.58	0.13	0.29

Buildings ⌂	Model	Climate	Other
Gases	0.43	0.5	0.07
Total	0.45	0.25	0.3
Liquids	0.64	0.23	0.14
Electricity	0.66	0.17	0.17
Biomass	0.61	0.1	0.29

Electricity ⚡	Model	Climate	Other
Coal (without CCS)	0.13	0.75	0.12
Gas (without CCS)	0.28	0.54	0.18
Renewables (incl. Biomass)	0.36	0.45	0.18
Biomass (with CCS)	0.35	0.39	0.26
Total	0.41	0.38	0.21
Wind	0.49	0.35	0.16
Gas (with CCS)	0.49	0.29	0.22
Solar	0.53	0.21	0.26
Nuclear	0.52	0.2	0.28
Hydro	0.71	0.19	0.1
Coal (with CCS)	0.61	0.19	0.2
Geothermal	0.67	0.15	0.18
Biomass (without CCS)	0.73	0.11	0.16

Emissions ☁	Model	Climate	Other
Kyoto Gases	0.02	0.94	0.05
CO_2	0.06	0.89	0.05
Electricity (CO_2)	0.11	0.79	0.1
CH_4	0.2	0.73	0.08
Transport (CO_2)	0.28	0.59	0.14
Industry (CO_2)	0.35	0.56	0.09
Buildings (CO_2)	0.42	0.48	0.09
N_2O	0.56	0.29	0.15

Primary Energy ☀	Model	Climate	Other
Coal	0.17	0.72	0.11
Oil	0.25	0.61	0.14
Total	0.29	0.51	0.2
Gas	0.3	0.5	0.2
Renewables (incl. Biomass)	0.21	0.49	0.3
Biomass	0.38	0.39	0.23
Wind	0.46	0.37	0.18
Hydro	0.65	0.24	0.12
Nuclear	0.5	0.19	0.3
Solar	0.62	0.19	0.19
Geothermal	0.73	0.14	0.13

Figure 3. **Upper-left**: overview of the variance decomposition of 2050 variables (for 2100 see Fig. S1 in SI B.1), for a selection of variables shown in the tables to the sides (identified by the symbols and color of the headings in the tables). Red-blue shading indicates the maximum magnitude across the presented indices: if a variable is in a more reddish area, its variance is dominated more strongly by a single factor. **Tables** indicate the indices as discussed in section 2 and are sorted by the index reflecting the variance explained by climate outcome differences. The values are highlighted by blue (low) to red (high) shades. The same figure using fractions of the total electricity generation (instead of absolute values) in the variance decomposition is shown in Fig. S3 and S4 in SI B.

Implications for policy and research

For both policymakers and researchers, identifying consensus on key variables in mitigation scenarios is vital to effective decision-making. Based on the largest available set of climate policy scenarios[7], we provide a comprehensive overview of robust ingredients in mitigation scenarios by quantifying the drivers of their spread in the IPCC AR6 database.

A number of aspects of climate policy scenarios that are found to be robust ingredients of mitigation (i.e., mainly driven by climate outcome) are a quantitative confirmation of what is already known[36], notably the decrease in emissions, with the interesting finding that the same holds for CO_2 emissions in individual end-use sectors and also for CH_4 emissions. While the energy mix of end-use sectors shows large

differences between models, also the decrease in fossil energy use (mostly in transport and buildings) can be regarded robust. More novel findings of robustness are found in early-century use of renewables – contrasting with high model dependency of individual renewable technologies – together with, to some extent, BECCS in the electricity mix, the fraction of wind, hydro and biomass in the primary energy mix, and electricity in end-use sectors (as fraction of the total energy use). In smaller-scale studies, some of these variables are found to be still highly model-dependent[2,37,38], which makes these findings shed new light on the robustness of their role in mitigation strategies.

Apart from the exceptions mentioned above, most individual technologies that could replace fossil fuels lack consensus, which is in line with previous work on multi-model robustness of energy technology projections[16,19]. Notable is the high model-dependence of solar energy (both primary as well as in the power mix), which is commonly argued to be of high importance in near-term mitigation[39,40]. CCS technologies in electricity generation are also highly model dependent, especially CCS in coal and gas plants. The high model differences in replacement in transport reflect that more research is needed into this direction: no robust cross-model mitigation alternative to fossil use in transport is found, apart from total electricity use and only as a fraction of the total final energy.

While models tend to agree that bioenergy use expands from current levels in mitigation scenarios, the exact supply and use of this resource differs a lot across models, particularly for the transport sector[29,41]. As highlighted in Daioglou et al. (2020)[34], technology characterization and coverage varies a lot across models, particularly concerning technology deployment constraints. Furthermore, it is stressed that bioenergy deployment in mitigation scenarios is largely driven by the energy system context, particularly the availability and costs of alternative mitigation options in and across end-uses, the availability and need of carbon dioxide removal, the speed with which large scale changes in the makeup of energy conversion facilities and their integration can take place, and the relative demand for different energy services.

Finally, little variance is caused by factors other than model or climate outcome, which does lead to a question regarding the representativeness of the AR6 database to cover a wide range of futures: in fact, the vast majority of scenarios in the database are based on SSP2 and cost-optimal assumptions. In other words, while theoretically one could think of various different future energy economies adhering to the same climate outcome, our results show that the cross-model scenario variation that addresses this wider solution space is limited. This finding highlights the need for a more systematic effort to better represent different scenario assumptions in studies: the consideration of a wider range of possible futures, but also a more consistent exploration of normative assumptions and storylines across models is needed to better reflect this kind of uncertainty in the database. Important elements to explore could include different economic growth rates (including post-growth scenarios[42]), levels of globalization and international co-operation, lifestyle and behavioral change and a distinction between "technology-focused" and "sufficiency" responses to climate and environmental challenges. A better representation of scenario variables could provide understanding of the driving forces determining energy futures, while model-based variability (lower-left corner) mostly reflects a lack of consensus.

This paper contains novel insights in the spread of key elements of climate policy scenarios and identifies consensus among them. This is of vital importance for policymaking and points out future research avenues. Moreover, our methodology of quantifying the drivers of the database spread is an innovation, which could be used for multi-model scenario studies and in future IPCC assessments.

References

1   IPCC. Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. (2022).

2   Kriegler, E. *et al.* Will economic growth and fossil fuel scarcity help or hinder climate stabilization? *Climatic Change* **136**, 7-22 (2016). https://doi.org:10.1007/s10584-016-1668-3

3   van Vuuren, D. P., de Vries, B., Beusen, A. & Heuberger, P. S. C. Conditional probabilistic estimates of 21st century greenhouse gas emissions based on the storylines of the IPCC-SRES scenarios. *Global Environmental Change* **18**, 635-654 (2008). https://doi.org:https://doi.org/10.1016/j.gloenvcha.2008.06.001

4   van Vuuren, D. P. *et al.* The costs of achieving climate targets and the sources of uncertainty. *Nature Climate Change* **10**, 329-334 (2020). https://doi.org:10.1038/s41558-020-0732-1

5   Sobol, I. M. Sensitivity Estimates for Nonlinear Mathematical Models. *Mathematical Modelling and Computational Experiments* **4**, 407-414 (1993).

6   Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937-1958 (2016). https://doi.org:10.5194/gmd-9-1937-2016

7   Byers, E. *et al.* AR6 Scenarios Database. (2022). https://doi.org:10.5281/ZENODO.5886912

8   van Beek, L., Hajer, M., Pelzer, P., van Vuuren, D. & Cassen, C. Anticipating futures through models: the rise of Integrated Assessment Modelling in the climate science-policy interface since 1970. *Global Environmental Change* **65**, 102191 (2020). https://doi.org:https://doi.org/10.1016/j.gloenvcha.2020.102191

9   van Vuuren, D. P. *et al.* How well do integrated assessment models simulate climate change? *Climatic Change* **104**, 255-285 (2011). https://doi.org:10.1007/s10584-009-9764-2

10   Rubin, E. S., Azevedo, I. M. L., Jaramillo, P. & Yeh, S. A review of learning rates for electricity supply technologies. *Energy Policy* **86**, 198-218 (2015). https://doi.org:https://doi.org/10.1016/j.enpol.2015.06.011

11   Duan, H., Zhang, G., Wang, S. & Fan, Y. Robust climate change research: a review on multi-model analysis. *Environmental Research Letters* **14**, 033001 (2019). https://doi.org:10.1088/1748-9326/aaf8f9

12   Harmsen, M. *et al.* Integrated assessment model diagnostics: key indicators and model evolution. *Environmental Research Letters* **16**, 054046 (2021). https://doi.org:10.1088/1748-9326/abf964

13   Rose, S. K. *et al.* An overview of the Energy Modeling Forum 33rd study: assessing large-scale global bioenergy deployment for managing climate change. *Climatic Change* **163**, 1539-1551 (2020). https://doi.org:10.1007/s10584-020-02945-6

14   Köberle, A. C. *et al.* Can global models provide insights into regional mitigation strategies? A diagnostic model comparison study of bioenergy in Brazil. *Climatic Change* **170**, 2 (2022). https://doi.org:10.1007/s10584-021-03236-4

15   Kriegler, E. *et al.* Making or breaking climate targets: The AMPERE study on staged accession scenarios for climate policy. *Technological Forecasting and Social Change* **90**, 24-44 (2015). https://doi.org:https://doi.org/10.1016/j.techfore.2013.09.021

16   Schaeffer, R. *et al.* Comparing transformation pathways across major economies. *Climatic Change* **162**, 1787-1803 (2020). https://doi.org:10.1007/s10584-020-02837-9

17   Sognnaes, I. *et al.* A multi-model analysis of long-term emissions and warming implications of current mitigation efforts. *Nature Climate Change* **11**, 1055-1062 (2021). https://doi.org:10.1038/s41558-021-01206-3

18   Vrontisi, Z. *et al.* Enhancing global climate policy ambition towards a 1.5 °C stabilization: a short-term multi-model assessment. *Environmental Research Letters* **13**, 044039 (2018). https://doi.org:10.1088/1748-9326/aab53e

19   Kriegler, E. *et al.* 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**, 353-367 (2014). https://doi.org:10.1007/s10584-013-0953-7

20   Saltelli, A. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications* **145**, 280-297 (2002). https://doi.org:https://doi.org/10.1016/S0010-4655(02)00280-1

21   IEA. World Energy Outlook. (2021).

22   Luderer, G. *et al.* Impact of declining renewable energy costs on electrification in low-emission scenarios. *Nature Energy* **7**, 32-42 (2022). https://doi.org:10.1038/s41560-021-00937-z

23   Fell, H., Gilbert, A., Jenkins, J. D. & Mildenberger, M. Nuclear power and renewable energy are both associated with national decarbonization. *Nature Energy* **7**, 25-29 (2022). https://doi.org:10.1038/s41560-021-00964-w

24   Sovacool, B. K., Schmid, P., Stirling, A., Walter, G. & MacKerron, G. Differences in carbon emissions reduction between countries pursuing renewable electricity versus nuclear power. *Nature Energy* **5**, 928-935 (2020). https://doi.org:10.1038/s41560-020-00696-3

25   Hanssen, S. V. *et al.* The climate change mitigation potential of bioenergy with carbon capture and storage. *Nature Climate Change* **10**, 1023-1029 (2020). https://doi.org:10.1038/s41558-020-0885-y

26   Gernaat, D. E. H. J., Bogaart, P. W., Vuuren, D. P. v., Biemans, H. & Niessink, R. High-resolution assessment of global technical and economic hydropower potential. *Nature Energy* **2**, 821-828 (2017). https://doi.org:10.1038/s41560-017-0006-y

27   Kim, S. H., Wada, K., Kurosawa, A. & Roberts, M. Nuclear energy response in the EMF27 study. *Climatic Change* **123**, 443-460 (2014). https://doi.org:10.1007/s10584-014-1098-z

28   van Soest, H. L., den Elzen, M. G. J. & van Vuuren, D. P. Net-zero emission targets for major emitting countries consistent with the Paris Agreement. *Nature Communications* **12**, 2140 (2021). https://doi.org:10.1038/s41467-021-22294-x

29   Leblanc, F. *et al.* The contribution of bioenergy to the decarbonization of transport: a multi-model assessment. *Climatic Change* **170**, 21 (2022). https://doi.org:10.1007/s10584-021-03245-3

30   Santarromana, R., Mendonça, J. & Dias, A. M. The effectiveness of decarbonizing the passenger transport sector through monetary incentives. *Transportation Research Part A: Policy and Practice* **138**, 442-462 (2020). https://doi.org:https://doi.org/10.1016/j.tra.2020.06.020

31   Zhang, R. & Hanaoka, T. Cross-cutting scenarios and strategies for designing decarbonization pathways in the transport sector toward carbon neutrality. *Nature Communications* **13**, 3629 (2022). https://doi.org:10.1038/s41467-022-31354-9

32   Oshiro, K. & Fujimori, S. Role of hydrogen-based energy carriers as an alternative option to reduce residual emissions associated with mid-century decarbonization goals. *Applied Energy* **313**, 118803 (2022). https://doi.org:https://doi.org/10.1016/j.apenergy.2022.118803

33   Odenweller, A., Ueckerdt, F., Nemet, G. F., Jensterle, M. & Luderer, G. Probabilistic feasibility space of scaling up green hydrogen supply. *Nature Energy* (2022). https://doi.org:10.1038/s41560-022-01097-4

34   Daioglou, V. *et al.* Bioenergy technologies in long-run climate change mitigation: results from the EMF-33 study. *Climatic Change* **163**, 1603-1620 (2020). https://doi.org:10.1007/s10584-020-02799-y

35   Riahi, K. *et al.* The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change* **42**, 153-168 (2017). https://doi.org:https://doi.org/10.1016/j.gloenvcha.2016.05.009

36   IPCC. Summary for Policymakers. In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. (2022).

37   Lu, Y., Cohen, F., Smith, S. M. & Pfeiffer, A. Plant conversions and abatement technologies cannot prevent stranding of power plant assets in 2 °C scenarios. *Nature Communications* **13**, 806 (2022). https://doi.org:10.1038/s41467-022-28458-7

38   Biilgen, S., Keles, S. & Kaygusuz, K. The Role of Biomass in Greenhouse Gas Mitigation. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* **29**, 1243-1252 (2007). https://doi.org:10.1080/00908310600623629

39   Cherp, A., Vinichenko, V., Tosun, J., Gordon, J. A. & Jewell, J. National growth dynamics of wind and solar power compared to the growth required for global climate targets. *Nature Energy* **6**, 742-754 (2021). https://doi.org:10.1038/s41560-021-00863-0

40   Creutzig, F. *et al.* The underestimated potential of solar energy to mitigate climate change. *Nature Energy* **2**, 17140 (2017). https://doi.org:10.1038/nenergy.2017.140

41   Bauer, N. *et al.* Global energy sector emission reductions and bioenergy use: overview of the bioenergy demand phase of the EMF-33 model comparison. *Climatic Change* **163**, 1553-1568 (2020). https://doi.org:10.1007/s10584-018-2226-y

42   Keyßer, L. T. & Lenzen, M. 1.5 °C degrowth scenarios suggest the need for new mitigation pathways. *Nature Communications* **12**, 2676 (2021). https://doi.org:10.1038/s41467-021-22884-9

Methods

General

The analysis is performed using Sobol's method of variance decomposition[5,20]. By sampling datasets where models and scenarios are uniformly represented, a measure of the intrinsic variance of each variable v is obtained and decomposed into first- and second-order variance contributions of the aforementioned factors. This yields three indices for the variance explained by the climate category (F_c), model (F_m) and other scenario assumptions (F_o) that add up to 1. Clearly, if $F_c \rightarrow 1$, there is statistical consensus or robustness about v being related to climate outcomes: e.g., even though there may be strong model differences in projecting v, the differences between climate outcomes supersede that signal (i.e., $F_c \gg F_m$). In this study, we treat these three indices as coordinates moving over time in a triangular variance decomposition "landscape" (Figs. 1-3), revealing how each variable's variance is determined, whether there is statistical consensus about its relation to climate outcomes and how this changes over time in our projections of the next century.

In this paper, we decompose variance of energy variables, which is intimately linked with the actual values of the variable: whether the variance of the *absolute* values of the variable is decomposed, or rather the variance of its *fraction* of the total energy consumption can therefore strongly affect the results, yielding a richer interpretation if both are analyzed. In the main text, we have chosen to show only variance decomposition results of each variable's *absolute* values, because not for all variables, their fractional counterpart is clearly defined or of interest (e.g., for emission variables). It is more intuitive and consistent to take a single approach for all variables. Nevertheless, both the absolute and fractional values of many energy variables can be policy relevant and insights from both are therefore utilized in the discussion. The results of the same analysis on the fractional values of several variables of which the results are shown in SI B.2.

Database

The AR6 database[7] is a product from the IPCC AR6 WGIII report on Mitigation of Climate Change[1]. In this analysis, we focus on scenarios that passed the historical vetting. This way we exclude scenarios with historical values that are highly different from observations and use only those that have a climate assessment of the resulting emissions. This yields a subset of 1202 scenarios, with 44 unique model versions, 13 unique model frameworks and 8 different climate categories. SI A discusses the models and climate categories in more detail. As explained below, we aggregate all model versions onto single 'model' labels, e.g., both IMAGE 3.0 and IMAGE 3.2 are identified as IMAGE.

Variance decomposition

For each variable v, a decomposition of the variance is performed at each time step t. In particular, we use Sobol's method, which is based on the reasoning that the value of a variable can be written as a function $f(x_1, x_2, \ldots, x_n)$ of a number of independent inputs x_i. In our case, we have two clearly identifiable inputs: the climate target x_c and the model x_m used to calculate the variable. Because that does not cover all variation we observe in the dataset and we do not have other information clearly distinguishing scenario entries apart from their model and climate target, we add a noise term ζ, leading to the following expression of this so-called Hoeffding-Sobol decomposition applied to our case:

$$v(t) = f(t, x_c, x_m) = f_0(t) + f_c(t, x_c) + f_m(t, x_m) + f_{cm}(t, x_c, x_m) + \zeta(t, x_c, x_m)$$

On the right-hand side, the first term is an overall average, f_i is a function of only factor i and f_{cm} is a function of both x_m and x_c. Note that the final (noise) term is also dependent on x_m and x_c: a-priorily, we do not know whether this noise is independent from climate target and/or model. Following the line of Sobol's theory, this noise term in fact contains both the first-order impact of other scenario assumptions on v, as well as potential second- and even third-order terms between these assumptions with climate target and models. While $v(t)$ is inherently a function of time, the variance decomposition is done for individual moments in time t_0. For clarity purposes, in the sequel, we therefore drop the t in the equations.

While f_0 is merely the overall average of v, the higher-order functions f_x are expressed as conditional expected values: e.g., $f_m = E(v|x_m) - f_0$, and $f_{mc} = E(v|x_m, x_c) - f_0 - f_c - f_m$. Taking the square integral of the above equation over $\vec{x}$ and dividing by the total variance of v ultimately yields:

$$1 = \frac{\text{var}(f_c(x_c))}{\text{var}(v)} + \frac{\text{var}(f_m(x_m))}{\text{var}(v)} + \frac{\text{var}(f_{cm}(x_c, x_m))}{\text{var}(v)} + \frac{\text{var}(\zeta(x_c, x_m))}{\text{var}(v)}$$

The last term can be interpreted as the total variance (including both first and higher order terms) explained by scenario assumptions *other* than climate target and model choice. Because it is not an actual first or second order term, we write it as S_o'. For the other terms, we use the definitions of Sobol indices $S_c = \text{var}(f_i)/\text{var}(v)$, yielding:

$$S_c(v) + S_m(v) + S_{cm}(v) + S_o'(v) = 1$$

The inputs x_c and x_m are, if taken from the dataset directly, not independent: some models have much more entries for C4 than for C2, for example, while other models have the opposite. This makes the Hoeffding-Sobol decomposition invalid: terms S_c and S_m would partly cover the same variance due to covarying labels. We solve this by not determining the variances directly from the database entries, but by creating sampled datasets such that all climate categories and models are uniformly represented, and in turn apply the variance decomposition on these sets. In practice, this works as follows. Per model-climate category pair (e.g., REMIND-C1, see table in SI A.3), we draw $p_{\text{sample size}}$ = 3000 scenarios. Because the combination REMIND-C1 does not have 3000 scenarios, we do this with replacing. Doing this for all climate-model pairs, we obtain a large dataset where x_c and x_m are perfectly orthogonal. From this set, we calculate the Sobol indices. Because this process is stochastic, we redo this process (p_{resample} =) 100 times and report the average. Indeed, this process involves two parameters $p_{\text{sample size}}$ and p_{resample}. If these parameters are taken too small, the results may be prone to stochasticity and the sample may lack sufficient uniformity, yielding errors in the indices. The values of 3000 and 100, respectively, are found to be high enough and approximating a deterministic result: when performing the same analysis using values of $p_{\text{sample size}} = 1000$ and $p_{\text{resample}} = 30$, the values of the indices F_c, F_m and F_o (defined below) changed on average with only 0.0007, 0.0008 and 0.0009, respectively, in 2050. In previous work[20], the calculation of the Sobol indices has been rewritten in matrix form, which is also what we use in favor of computational efficiency.

The sampling method removes bias stemming from differences in abundance of models. The method assumes that the scenario entries per climate category and model are representative for these labels, which arguably holds better for model-label combinations with many entries, than for those with just a few. Important to note is that TIAM-ECN does not have C1 and C2 entries, meaning that these entries are empty in the sampling and the sample is not perfect.

Our aim is to decompose the total variance into terms attributable to each individual input. However, in contrast to a similar approach in earlier work[2], the second-order Sobol term S_{cm} cannot be neglected, as for most variables v: $0.05 < S_{cm}(v, t = 2050) < 0.30$ (average 0.16). From an intuitive point-of-view, second-order variations also matter: when all models show variation among climate targets in their output, but in different magnitudes or at a different base level – so that it is less pronounced in the first-order term – this is of interest. Moreover, if the second-order term would be excluded, we cannot interpret indices as fractions of the total variance anymore, as they do no longer add up to 100%. For these reasons, we construct three indices based on the calculated first- and higher order Sobol terms, in which we add S_{cm} to S_c and S_m:

$$
\begin{cases}
F_c := S_c + \dfrac{1}{2} S_{cm} \\
F_m := S_m + \dfrac{1}{2} S_{cm} \\
F_o := S_o' = S_o + S_{om} + S_{oc} + S_{ocm}
\end{cases}
$$

which add up to 1. The indices S_o, S_{om}, S_{oc} and S_{omc} are mentioned here for interpretation purposes, but cannot explicitly calculated in the analysis: together, they form a 'rest' term and do not govern a clearly defined set of assumptions. We cannot calculate F_o directly because we lack labels and enough, well-distributed data to create a sample from which we can explicitly compute these individual terms. Therefore, we deduce F_o from F_c and F_m. Note that the expression for F_o also contains terms with 'model' and 'climate' subscripts. Even though it is not possible to compute them separately with available data anyway, we think adding them together is appropriate as long as the interpretation is clear: F_o does not only take into account model- and climate ambition-independent variations due to assumptions differences on for example GDP or technological advancement, but also takes into account how these assumption differences vary across models and climate ambition – varying GDP may have a different or more pronounced effect in model X than it has in model Y. However, for the identification of robust aspects, these model differences are less of interest, which legitimizes the choice of grouping them into F_o. Thus, for both practical reasons as well as interpretation reasons, we have chosen for these definitions of the indices.

Note that this means that the *total fraction of variance explained* by model differences may therefore be slightly higher than F_m (idem for F_c). This way, F_o acts as an upper bound on the relative effect of other scenario assumptions, both in lower and in higher order. Still, because the relative magnitude of F_o turned out to be so small, other scenario assumptions are not systematic in labels and quantification among model entries (idem climate outcomes) we expect that the *total* sensitivity towards model (climate) differences is already approximated by F_m (F_c) – even though the mathematical definition of F_m (F_c) only includes the first order and second-order term between model and climate.

Summarized, the interpretation of the resulting indices $F_i(v, t)$ is the percentage of the variance of variable v at year t that is explained due to differences in factor i (i.e., model, climate target or other scenario assumptions), where in the first two indices the effects of other scenario assumptions are aggregated. All higher-order terms are accounted for.

Pre-processing
A number of pre-processing steps are implemented prior to the variance decomposition itself. The first step concerns cleaning the data with a few significant outliers. In particular, two scenarios

(EN_NPi2020_800 and EN_NPi2020_900 by WITCH 5.0) are removed because they showed unrealistically high values for hydrogen use in transport (order 1e3 EJ while the total energy use in transport was a factor 10 lower and most other entries of hydrogen in transport are a factor 100 or even 1000 lower).

The next step is that model versions are aggregated into single models. For example, we do not distinguish IMAGE 3.0 from IMAGE 3.2. The reason is that if we would keep them separate, an unrealistic model-similarity would be the result: the "model" category would explain suddenly less variance because models seem to be more alike to each other (i.e., S_m drops significantly), while in reality this is not caused by unique models, but by unique model versions. Note that some models have over 10 versions of themselves reported as unique model versions. In this paper, we are interested in how different modelling perspectives or modelling groups project variables differently. We deem that this is best illustrated when distinguishing model frameworks from each other rather than mere model versions. In the SI, a table describes the exact translations between models and model versions. Of course, model versions may still differ, which are now not recognized as "model differences", but as differences among scenarios of the same model group, mostly contributing to F_o.

Scenarios with climate category C8 (exceed warming of 4°C by 2100 with over 50% probability) are removed, because only 4 out of 9 models (that in general have sufficient entries) report C8 scenarios, leaving a model bias in how C8 is reported. Note that these models cover a small fraction of the total set anyway (about 2% for most variables). After this pre-processing step, 1152 scenarios are left.

It is important to note that not all scenarios contain the same sets of variables. Especially some more detailed variables (e.g., hydrogen use as fuel for specifically passenger transport) are covered by only a few hundreds of scenarios. To keep the number of scenarios used in the analysis of each variable as high as possible, we create separate databases for each unique variable v, containing all scenarios (out of the aforementioned 1152) that contain v. In each such database, we remove all scenario entries of models that have less than 10 entries in total. In practice, the results are not sensitive to the value of this parameter for a broad range of its values because a clear separation of small-abundance models can already be recognized (see table later), the numbers being: 1, 7, 45, 47, 55, 65, 113, 114, 142, 266, 297. Clearly, the models having only 1 (MERGE-ETL) and 7 (EPPA) scenario entries are much lower than the rest and are therefore removed in this analysis – in particular because we also aim to distinguish climate categories within the model entries.

In the majority of the scenarios, data is provided in 5-year increments. For some scenarios, however, data may be missing or only 10-year increments are reported in the second half of the century. For this study, the temporal resolution must be fully equal among the scenarios, which is why we fill in these gaps using linear interpolation within the scenario entry such that all scenario entries have 5-year increments.

Limitations
A number of limitations to this methodology can be mentioned, some of which are already mentioned before. First, the number of entries with combined model and climate category labels vary greatly, as seen in the table in SI A.2. This was the reasoning behind applying the sampling method, but raises questions on the respective representativeness of each model-climate combination. For example, there is only one single COFFEE-C1 scenario: does this single scenario represent this combined label enough? Arguably, the 84 REMIND-C3 scenarios are a better representation of their combined label. Second, other scenario assumptions (beyond climate outcome) are not systematically varied along models and climate categories, making the sampling for this not well interpretable other than being a 'rest' term after climate

and model dependencies. In future research, similar analyses could be performed on databases where the SSPs are well represented among all model-climate combinations. Third, as mentioned, other scenario assumptions may have nonzero higher-order terms involving climate and model differences, as well. Fourth and final, the sampling method takes care of bias towards high-abundance models and climate outcomes, but is limited by the database itself: potential biases in the full scientific IAM community cannot be filtered, which are potentially not negligible.

A few final considerations about the term *consensus* should be noted. In this paper, we refer to 'consensus' about a variable's link to mitigation strategies when its variance is mainly or significantly driven by climate outcome and less so by other drivers. However, this does not necessarily mean that this variable has no uncertainty or spread anymore: it can very well be that the exact value of the respective variable still covers a broad range for a single climate target. Also, when a variable is significantly driven by other scenario assumptions (lower-right corner of the triangular panels in the figures), other forms of 'consensus' may be present. For example, when there is agreement ('consensus') on a *plural* set of energy futures, depending on a set of scenario assumptions, that are all possible under similar climate outcomes and similar models. This type of consensus can be an interesting future research avenue, also in light of the finding that only very little (relative) variance is captured by scenario assumptions.

Data and code availability

All output data of the analysis will be available on Zenodo upon publication. The AR6 scenario data is open source and can be found at data.ece.iiasa.ac.at/ar6. All code will be available on Github at https://github.com/MarkMDekker/variancedecomposition.

Author contributions

All authors conceived the study. MMD performed the analysis, generated the figures and wrote the first draft. MMD, KvdW and RvH contributed to the methods. AH, DvV, VD and MMD analyzed the output. All authors contributed to the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Supplementary Information

SI A. Additional methodological information

A.1 Model details

See table below for details on which models and model versions are used in the analysis. Models colored in orange are excluded from this analysis because of their small abundance in the dataset. For more information on each model, the reader is referred to the Integrated Assessment Modelling Community (IAMC) wiki: https://www.iamcdocumentation.eu/index.php/IAMC_wiki.

ABBR	MODEL	VERSIONS	INSTITUTION
IMA	IMAGE	IMAGE 3.0 IMAGE 3.0.1 IMAGE 3.0.2 IMAGE 3.2	Netherlands Environmental Assessment Agency (PBL)
TIA	TIAM-ECN	TIAM-ECN 1.1	Netherlands Organization for applied scientific research (TNO)
GCA	GCAM	GCAM 4.2 GCAM 5.2 GCAM 5.3 GCAM-PR 5.3	Joint Global Change Research Institute (JGCRI)
POL	POLES	POLES ADVANCE POLES CD-LINKS POLES EMF30 POLES EMF33 POLES ENGAGE POLES GECO2019	Joint Research Council (JRC)
AIM	AIM	AIM/CGE 2.0 AIM/CGE 2.1 AIM/CGE 2.2 AIM/Hub-Global 2.0	National Institute for Environmental Studies (NIES)
MES	MESSAGE(ix-GLOBIOM)	MESSAGE-GLOBIOM 1.0 MESSAGEix-GLOBIOM 1.0 MESSAGEix-GLOBIOM_1.1 MESSAGEix-GLOBIOM_1.2 MESSAGEix-GLOBIOM_GEI 1.0	International Institute for Applied Systems Analysis (IIASA)
REM	REMIND	REMIND 1.6 REMIND 1.7 REMIND 2.1 REMIND-Buildings 2.0 REMIND-MAgPIE 1.5 REMIND-MAgPIE 1.7-3.0 REMIND-MAgPIE 2.0-4.1 REMIND-MAgPIE 2.1-4.2 REMIND-MAgPIE 2.1-4.3 REMIND-Transport 2.1	Potsdam Institute for Climate Impact Research (PIK)
WIT	WITCH	WITCH 4.6 WITCH 5.0 WITCH-GLOBIOM 3.1 WITCH-GLOBIOM 4.2 WITCH-GLOBIOM 4.4	Euro-Mediterranean Center for Climate Change (CMCC)
COF	COFFEE	COFFEE 1.1	Alberto Luiz Coimbra Institute for Graduate Studies and Research in Engineering (COPPE/UFRJ)
CRO	C-ROADS	C-ROADS-5.005	Climate Interactive
EPP	EPPA	EPPA 6	Massachusetts Institute of Technology (MIT)
GEM	GEM-E3	GEM-E3_V2021	Institute of Communication and Computer Systems (ICCS)

A.2 Overview of climate categories

CLIMATE CATEGORY	MEANING
C1	Limit warming to 1.5 °C (>50%) with no or limited overshoot
C2	Return warming to 1.5 °C (>50%) after a high overshoot
C3	Limit warming to 2 °C (>67%)
C4	Limit warming to 2 °C (>50%)
C5	Limit warming to 2.5 °C (>50%)
C6	Limit warming to 3 °C (>50%)
C7	Limit warming to 4 °C (>50%)
C8	Exceed warming of 4 °C (>=50%)

A.3 Model and climate category prevalence in the database

As an example, the below table provides the number of historically-vetted scenarios per model and climate category for primary energy from coal. Orange cells have been removed from the dataset because of the low abundance of C8, EPPA and MERGE-ETL scenarios: including them would make well representative samples impossible. Also note the blue cells: these empty entries for TIAM-ECN, which *are* included, make the sampling method slightly imperfect. However, because it only concerns 2 out of 63 entries, the term is still small and the interpretation of the indices is approximately the same. The alternative would have been to fully drop the TIAM-ECN model as well, which in turn also decreases the representativeness of the sample. Hence, we have chosen to keep this model in and accept these empty entries.

	C1	C2	C3	C4	C5	C6	C7	C8	Total
AIM	4	3	17	8	13	4	6	0	55
COFFEE	1	4	14	15	21	9	1	0	65
EPPA	0	0	1	3	0	1	2	0	7
GCAM	6	5	13	9	6	1	6	1	47
IMAGE	7	9	34	18	22	16	34	2	142
MERGE-ETL	0	0	0	1	0	0	0	0	1
MESSAGE(ix-GLOBIOM)	20	43	59	39	57	20	28	0	266
POLES	4	10	26	24	20	11	19	0	114
REMIND(-MAgPIE)	41	44	84	16	34	19	48	11	297
TIAM-ECN	0	0	20	6	10	4	5	0	45
WITCH	9	2	29	14	24	9	12	14	113
Total	92	120	297	153	207	94	161	28	1152

SI B. Additional results

B.1 Results for 2100

While the main text focuses on results for 2050, which is of high relevance to policymakers in achieving the Paris climate goals, the same analysis can be done for the second half of the century. Being further away in the future, it is expected that model uncertainties expand, but still, it can provide insights in current robust aspects of energy systems in 2100. Fig. S1 shows an overview of the results for 2100, analogous to Fig. 3 in the main text.

Industry ☼	Model	Climate	Other
Coal	0.51	0.37	0.12
Gases	0.5	0.3	0.2
Liquids	0.56	0.3	0.14
Biomass	0.6	0.27	0.13
Electricity	0.58	0.2	0.22
Total	0.7	0.12	0.18
Hydrogen	0.74	0.05	0.21

Transport 🚍	Model	Climate	Other
Oil	0.28	0.57	0.15
Total	0.47	0.31	0.22
Hydrogen	0.55	0.18	0.27
Bioenergy	0.76	0.11	0.12
Electricity	0.74	0.09	0.17
Gases	0.77	0.08	0.15

Buildings △	Model	Climate	Other
Gases	0.45	0.43	0.13
Biomass	0.48	0.18	0.34
Liquids	0.73	0.16	0.11
Total	0.64	0.11	0.25
Electricity	0.73	0.1	0.17

Electricity ⚡	Model	Climate	Other
Coal (without CCS)	0.18	0.63	0.19
Gas (without CCS)	0.29	0.55	0.16
Biomass (with CCS)	0.53	0.32	0.16
Renewables (incl. Biomass)	0.51	0.25	0.24
Gas (with CCS)	0.61	0.21	0.18
Geothermal	0.71	0.19	0.1
Coal (with CCS)	0.41	0.19	0.4
Total	0.53	0.18	0.29
Wind	0.65	0.15	0.2
Nuclear	0.58	0.15	0.27
Solar	0.63	0.12	0.25
Biomass (without CCS)	0.6	0.12	0.28
Hydro	0.8	0.09	0.11

Emissions ☁	Model	Climate	Other
Kyoto Gases	0.02	0.9	0.07
CO2	0.05	0.87	0.08
Industry (CO2)	0.15	0.71	0.14
Electricity (CO2)	0.18	0.69	0.13
CH4	0.22	0.66	0.12
Transport (CO2)	0.26	0.63	0.12
Residential & Commercial (CO2)	0.4	0.48	0.13
N2O	0.48	0.38	0.15

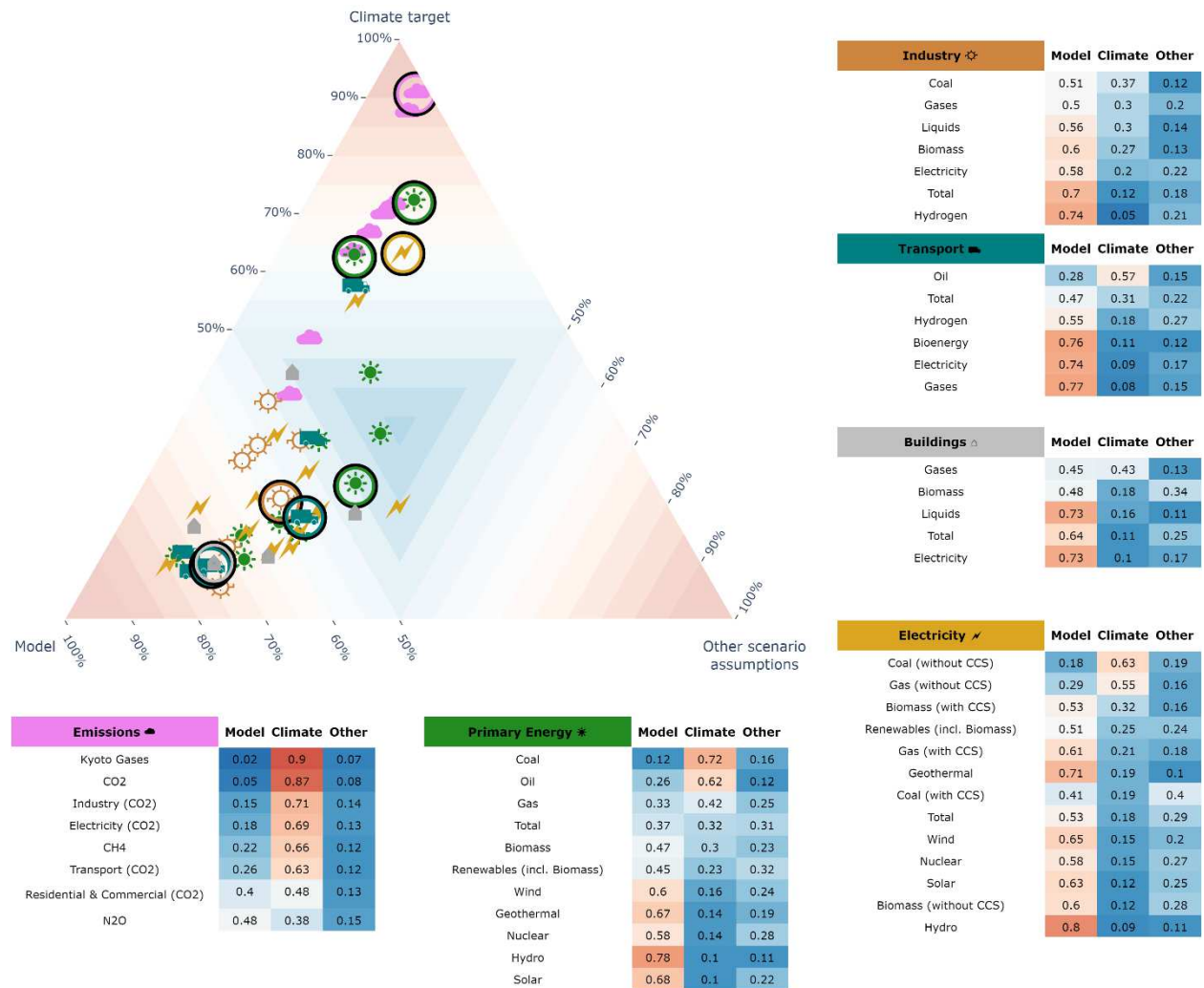| Primary Energy ✱	Model	Climate	Other
Coal	0.12	0.72	0.16
Oil	0.26	0.62	0.12
Gas	0.33	0.42	0.25
Total	0.37	0.32	0.31
Biomass	0.47	0.3	0.23
Renewables (incl. Biomass)	0.45	0.23	0.32
Wind	0.6	0.16	0.24
Geothermal	0.67	0.14	0.19
Nuclear	0.58	0.14	0.28
Hydro	0.78	0.1	0.11
Solar	0.68	0.1	0.22

Figure S1. As in Fig. 3 in the main text, but for the year 2100 rather than 2050. Highlighted variables in upper-left panel are those that are highlighted in Fig. 3 as well.

B.2 Absolute versus fractional variables

Figure S2 shows the fractions-based counterpart of Fig. 1 in the main text: variance decomposition results for energy carriers used for electricity generation. No major differences can be observed, making the conclusions we draw in the main text robust under this transformation. In particular, renewable technologies such as solar (yellow) and wind (violet) power, and even the aggregated renewable variable (purple) are all even less climate-dependent and more model dependent than in Fig. 1. This rules out the potential explanation of (e.g.) solar power being model dependent purely because of total absolute

energy consumption being model dependent. Interestingly, hydropower now seems rather static across the century with approximately 70% of its variance being determined by model differences, while its variance of absolute values grows from 60% to 80% model-determined in Fig. 1.
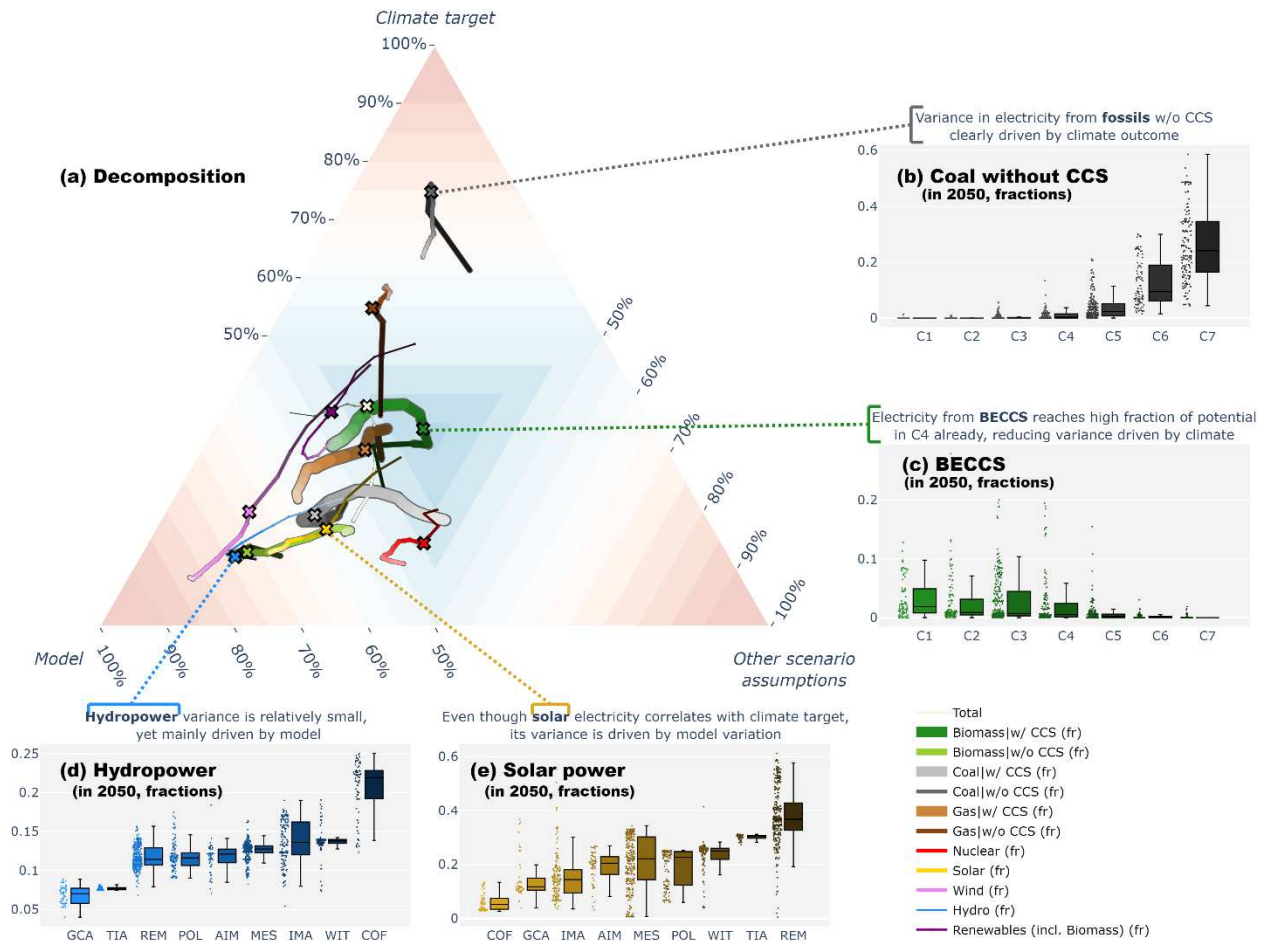

Figure S2. Variance decomposition results for energy mix used for electricity generation as in Fig. 1 of the main text, but not using the absolute value of these variables (in EJ), but their fraction of the total energy used. Barplots (b)-(e) are also changed accordingly.

Figure S3 shows the comparison between variance decomposition results based on the absolute and fractional values, respectively. For most variables, the approach does not change the results significantly, which again strengthens our conclusions in the main text. One exception is primary hydro use, being much less model-dependent in the fractional case (40%) in 2050 than it is in the absolute-value case (65%), which points to that hydropower is sensitive to differences in the total absolute electricity generation. Interestingly, the use of hydropower in electricity generation does not have this effect: even slightly the opposite (71% -> 74%). Nuclear primary energy use and nuclear power both have a decrease in significant model dependency when looking at fractions of the total rather than their absolute values.
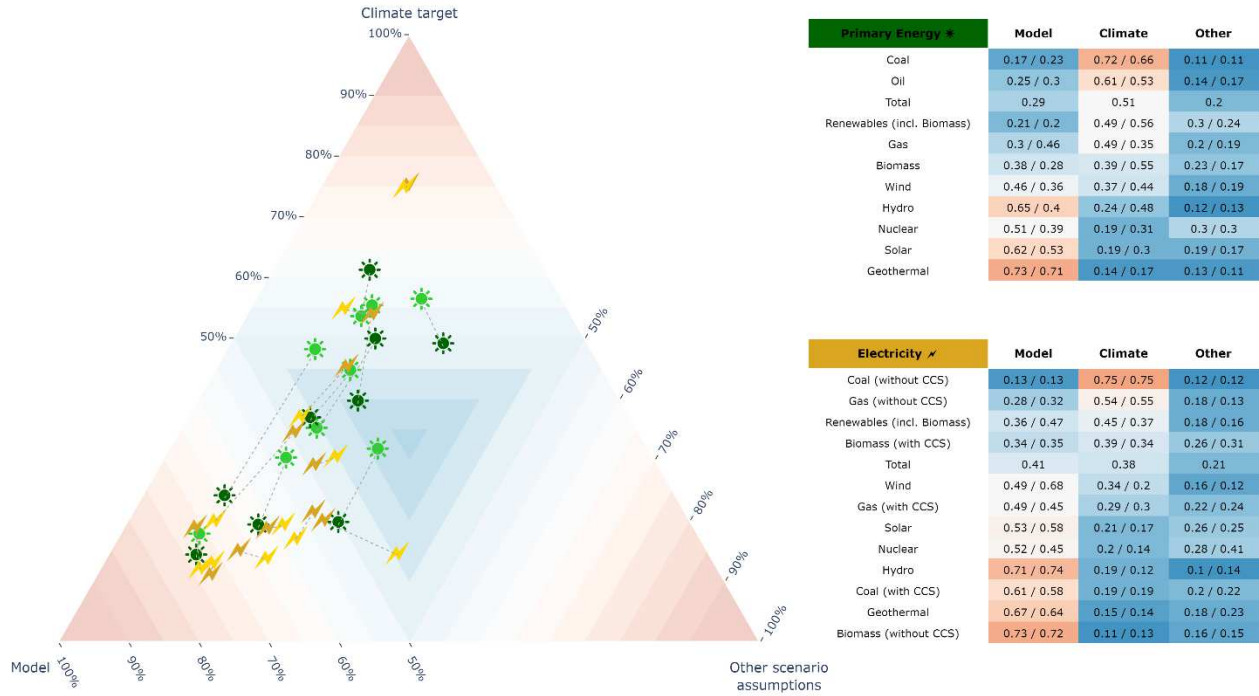
*Figure S3. Variance decomposition results in 2050 for the primary energy mix (green shades) and the energy mix used for electricity generation (yellow shades). **Left:** dark (green/yellow) shades indicate the decomposition based on absolute values, while light (green/yellow) shades indicate the decomposition based on fractions of the total. Dashed lines are drawn between absolute / fractional pairs that belong to the same variable. **Right:** tables showing the fractions of the variance explained by each factor (as in Fig. 3 of the main text). Tuples in tables indicate indices by performing the analysis on "absolute values / fractional values". Sorting and blue/red shading of the tables is (still) based on the absolute-value analysis.*

Similar to Fig. S3, we also compare the variance decomposition versions of variable magnitude versus variable fractions of the total for the energy carriers in the end-use sectors. This is shown in Fig. S4. A number of observations can be made. As mentioned in the main text, electricity use in these sectors become much less model dependent in the fractional case: apparently, the absolute value of electricity use is highly debated, but its relative role is more robust. In industry, hydrogen is 10%-point less model dependent, and in transport, a notable outlier is bioenergy, which moves from 70% of its variance determined by model differences, to only 52%. No notable changes in the "other scenario assumptions" dimension are visible.
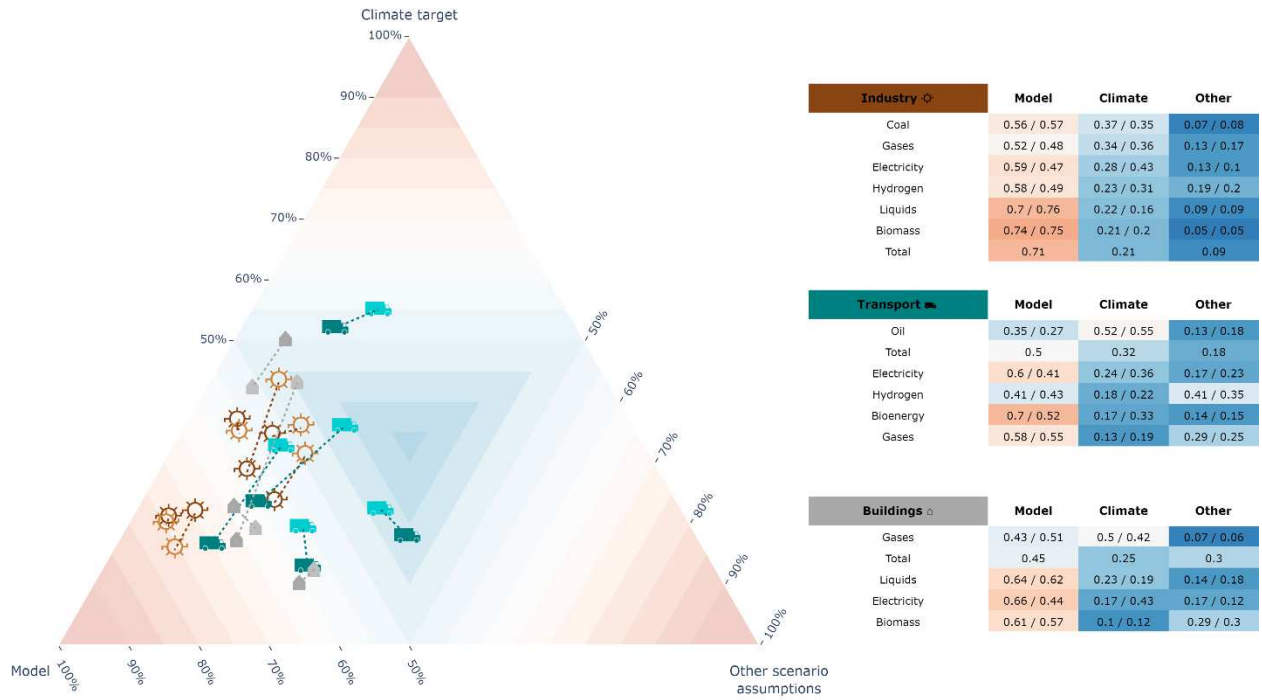
Industry ⚙	Model	Climate	Other
Coal	0.56 / 0.57	0.37 / 0.35	0.07 / 0.08
Gases	0.52 / 0.48	0.34 / 0.36	0.13 / 0.17
Electricity	0.59 / 0.47	0.28 / 0.43	0.13 / 0.1
Hydrogen	0.58 / 0.49	0.23 / 0.31	0.19 / 0.2
Liquids	0.7 / 0.76	0.22 / 0.16	0.09 / 0.09
Biomass	0.74 / 0.75	0.21 / 0.2	0.05 / 0.05
Total	0.71	0.21	0.09

Transport 🚚	Model	Climate	Other
Oil	0.35 / 0.27	0.52 / 0.55	0.13 / 0.18
Total	0.5	0.32	0.18
Electricity	0.6 / 0.41	0.24 / 0.36	0.17 / 0.23
Hydrogen	0.41 / 0.43	0.18 / 0.22	0.41 / 0.35
Bioenergy	0.7 / 0.52	0.17 / 0.33	0.14 / 0.15
Gases	0.58 / 0.55	0.13 / 0.19	0.29 / 0.25

Buildings ⌂	Model	Climate	Other
Gases	0.43 / 0.51	0.5 / 0.42	0.07 / 0.06
Total	0.45	0.25	0.3
Liquids	0.64 / 0.62	0.23 / 0.19	0.14 / 0.18
Electricity	0.66 / 0.44	0.17 / 0.43	0.17 / 0.12
Biomass	0.61 / 0.57	0.1 / 0.12	0.29 / 0.3

Figure S4. As in Fig. S3, but for energy consumption in end-use sectors.

B.3 Detailed data visualizations

In support of some the conclusions in the main text, here we show the projected values of a number of variables in more detail, sorted by model and climate category.
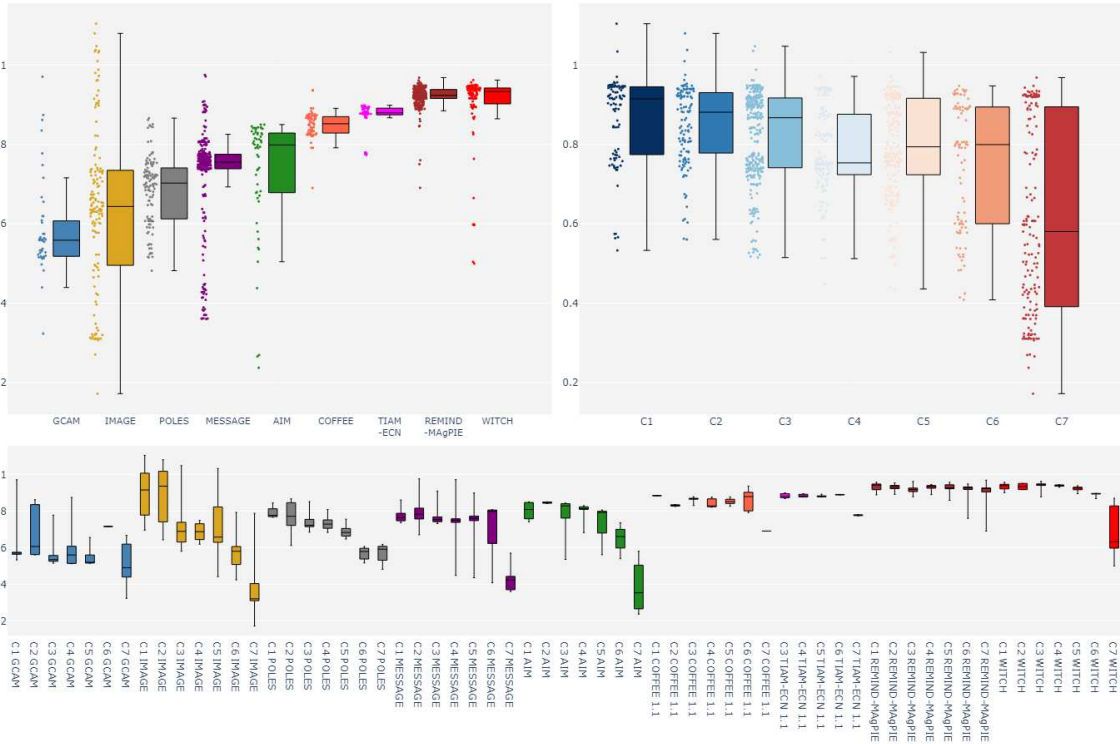

*Figure S5. Data of renewables (including biomass) used for electricity generation in 2100 in fractions of the total electricity generation, sorted by model (**upper-left**), climate category (**upper-right**) and both (**bottom**). Boxplots indicate quartiles (excluding outliers) and each dot reflects a single scenario projection.*
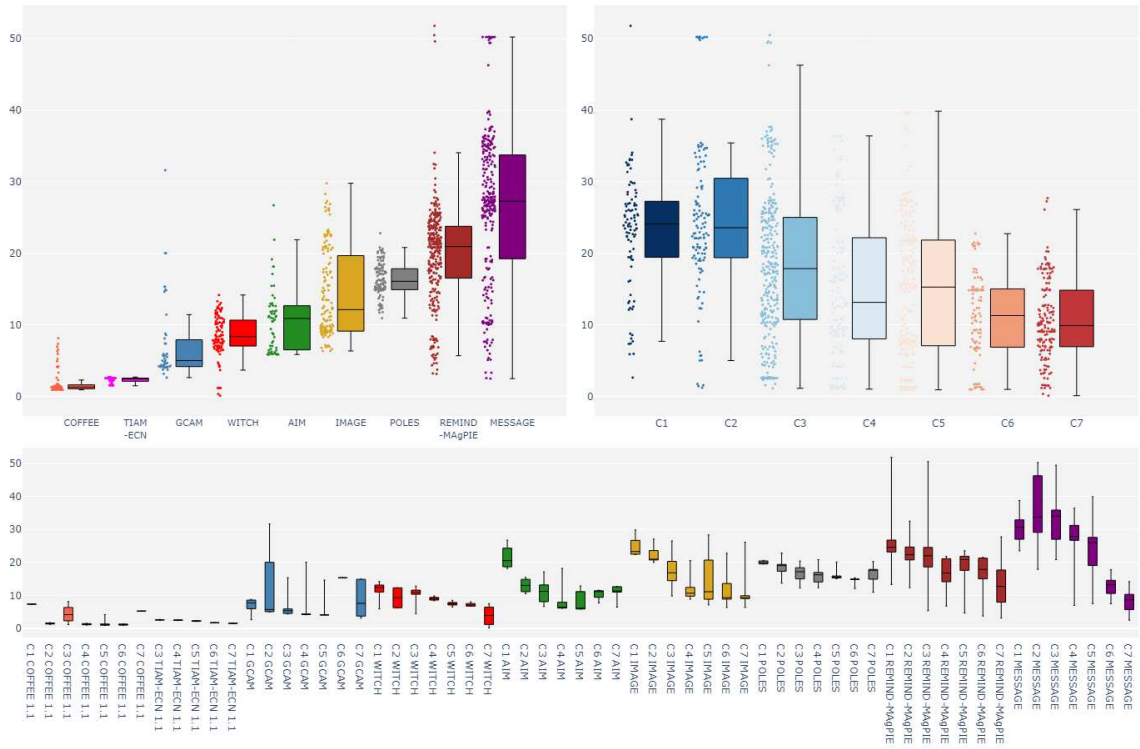
Final Energy|Transportation|Electricity in 2050

Figure S6. As in Fig. S5, but for electricity use in transport in 2050 (in absolute values).

Final Energy|Transportation|Hydrogen in 2100

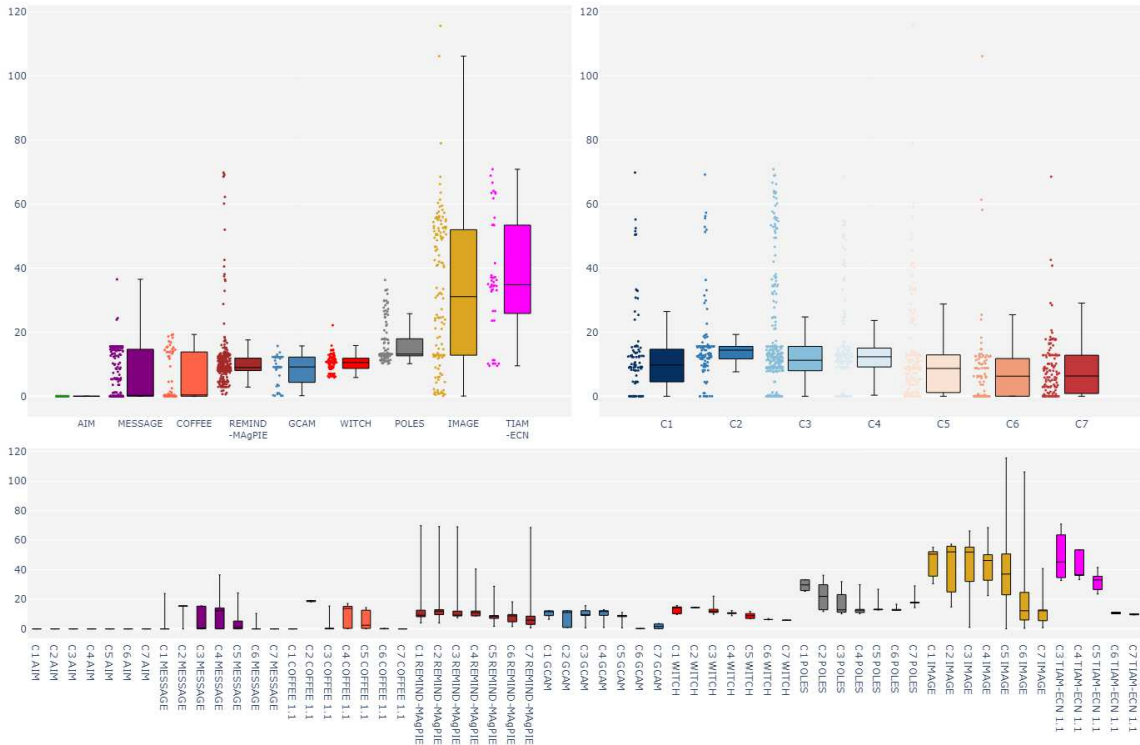*Figure S7. As in Fig. S5, but for hydrogen use in transport in 2100 (in absolute values).*

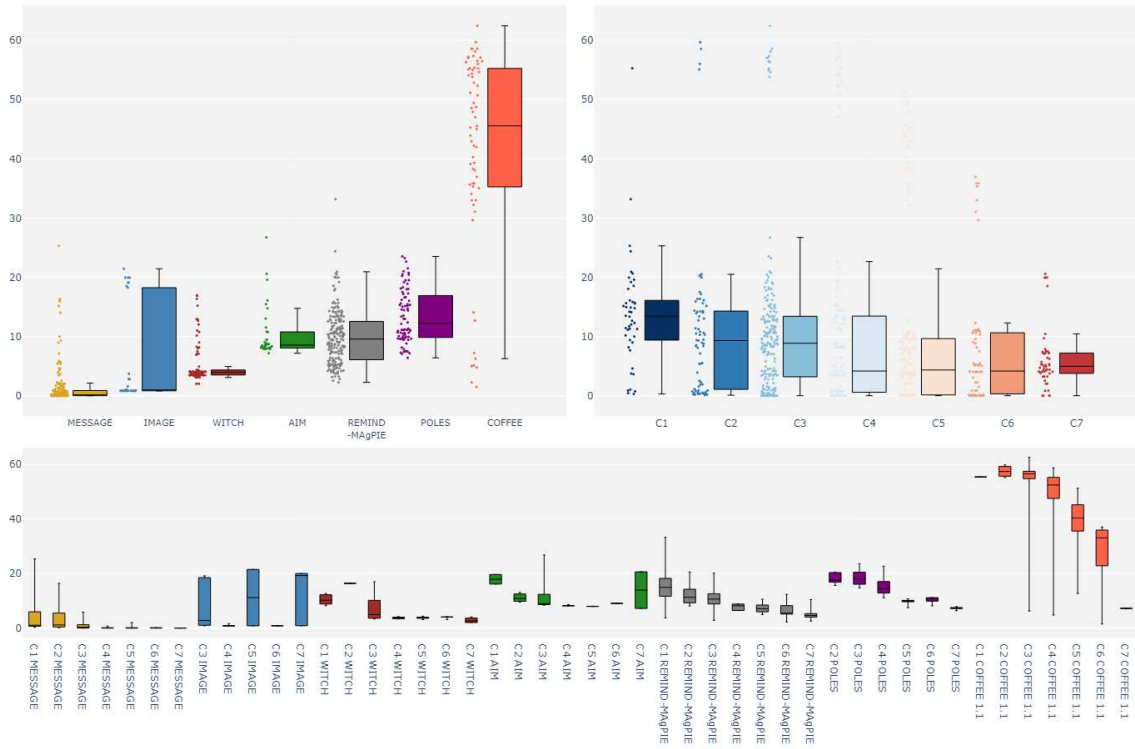Final Energy|Transportation|Liquids|Bioenergy in 2050

Figure S8. As in Fig. S5, but for bioenergy use in transport in 2050 (in absolute values).