NAVIGATE

Next generation of advanced integrated assessment modelling to support climate policy making

Report on the workshop "Robustness and legitimacy of models for climate policy assessment" and further plans for the NAVIGATE stakeholder process

NAVIGATE Deliverable 1.4, 24 August 2020

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 821124.

Prepared by: Evelina Trutnevyte (University of Geneva)

With inputs from: Nico Bauer (Potsdam Institute for Climate Impact Research), Johannes Emmerling (European Institute on Economics and the Environment), Celine Guivarch (CIRED), Marc Jaxa-Rozen (University of Geneva), Volker Krey (International Institute for Applied Systems Analysis), Elmar Kriegler (Potsdam Institute for Climate Impact Research), Jean-Francois Mercure (University of Exeter), Bas van Ruijven (International Institute for Applied Systems Analysis), Michiel Schaeffer (Climate Analytics), Jessica Strefler (Potsdam Institute for Climate Impact Research), Detlef van Vuuren (PBL Netherlands Environmental Assessment Agency), Christopher Weber (WWF Global Science), Charlie Wilson (University of East Anglia), Sonia Yeh (Chalmers University of Technology)

To be reviewed by: Ramona Lorenz (Potsdam Institute for Climate Impact Research), Aleksander Śniegocki (WiseEuropa)

Status: final, 24 August 2020

Table of content

1. Background

Rapid decarbonization of societies around the world is required to meet the Paris goals to hold global warming well below 2°C and to pursue efforts to limit warming to 1.5°C. This rapid decarbonization needs to be embedded in a broader agenda of sustainable development as defined by the 17 Sustainable Development Goals (SDGs) of the UN 2030 agenda.

Integrated Assessment Models (IAMs) of climate change are popular tools that quantify climate change mitigation pathways on how industrialized, emerging and least developed countries can work towards the Paris climate goals. IAMs account for the evolution of coupled systems of economy, energy, land, water, and climate in the short term (2020-2030), mid term (2030-2050), and long term (2050-2100). Due to their integrative nature, IAMs play an important role in the assessments of the Intergovernmental Panel on Climate Change (IPCC) as well as regional and national decarbonization strategies, e.g. the EU, China, and Brazil.

European Union's Horizon 2020 project **NAVIGATE** aims to develop the next generation of IAMs. The project targets major advancements in several areas:

- Improving the representation of transformative change in interlinked social, technological and economic systems and in consumer goods and services;
- Developing new capabilities to capture spatial and social heterogeneity for assessing distributional implications of climate change impacts and climate policy, and the interactions with other SDGs;
- Improving robustness, legitimacy, and usability of IAM results so that policy makers, business, civil society organizations, as well as other climate policy scholars are sufficiently informed and empowered to use IAM results.

The NAVIGATE project includes a series of activities to create and sustain a continuous **stakeholder exchange** between NAVIGATE and the key groups of experts and stakeholders. The aim is to gather feedback on the design of the project's research activities, the choice of the stakeholder-relevant IAM outputs and policy-relevant scenarios, and on the documentation needs of the stakeholders. The exchange is primarily enabled by means of three workshops and, if needed, online surveys that bring together three groups of participants: (i) stakeholders from national and international climate policy domains, business, and civil society organizations, (ii) scholars from various disciplines that are relevant to climate policy, and (iii) NAVIGATE consortium members.

This report presents the outcomes of the first NAVIGATE stakeholder and expert workshop "Robustness and legitimacy of models for climate policy assessment" (Section 2) and delineates the plans for future stakeholder process in the project (Section 3). Presentation slides of the workshop are available on the NAVIGATE website: https://navigate-h2020.eu/first_stakeholder_workshop/

2. First stakeholder and expert workshop "Robustness and legitimacy of models for climate policy assessment"

2.1. Aims of the workshop

The first NAVIGATE stakeholder and expert workshop on **"Robustness and legitimacy of models for climate policy assessment"** took place on 26-27 May 2020. The workshop was originally prepared to be hosted by the University of Geneva and take place in Chamonix, but in April 2020 it was redesigned

as an online event due to the travel restrictions related to the COVID-19 outbreak.

The topic of the workshop was decided during the first NAVIGATE Consortium meeting on 11-13 September 2019 in Potsdam. The first NAVIGATE workshop focused on the critical issues of robustness and legitimacy of models for climate policy assessment. **Robustness** refers to the analytical and technical adequacy of the models as well as the validity of modeling results given deep uncertainties and limits to state-of-the-art knowledge. **Legitimacy** encompasses transparency, traceability and accessibility of the modeling itself as well as the design of participatory processes at the modeling-policy interface. Although often discussed in the last decade, robustness and legitimacy of models face new challenges today because, as in the case of NAVIGATE, the models need to account for increasingly richer structural, regional and distributional information as well as for transformative change in social, technological and economic systems.

The **aim of the workshop** was therefore to enable the participants to share their experiences and to reflect on the good-practice examples, expectations, and remaining challenges for robustness and legitimacy of models for climate policy assessment. Specifically, these questions were addressed:

- What defines robust models, modeling results, and climate policy recommendations?
- By what means can this robustness be assessed, ensured, documented, and communicated?
- What do models need in order to be legitimate tools to inform climate policy?

The workshop combined big-picture discussions on the topics of robustness and legitimacy with group work on two case studies:

- **emissions gap** (Figure 1) between the current pathways of Nationally Determined Contributions until 2030 and the 2°C or 1.5°C targets without or with limited overshoot;
- informing **climate policy after the pandemic** that brought rapid change in trade, structure of the economy, employment, transport, lifestyles, and inequality.

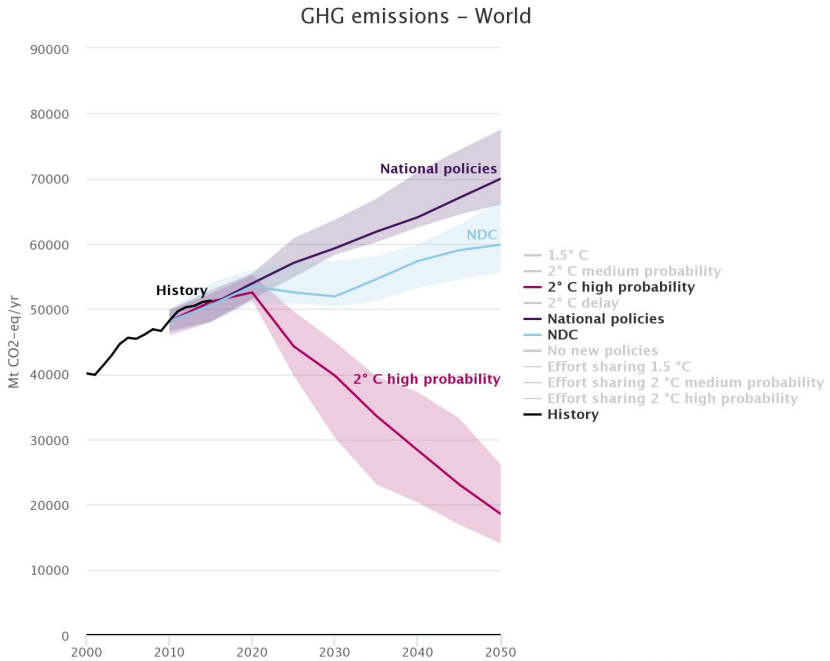

Figure 1. Case study for the workshop on emissions gap. The figure was adapted from https://themasites.pbl.nl/global-stocktake-indicators/take-indicators/

2.2. Program of the workshop

Day #1, Tuesday, 26 May 2020
Central European Summer Time

14:00 – 14:10 **Welcome and introduction**
Elmar Kriegler (Potsdam Institute for Climate Impact Research)
Evelina Trutnevyte (University of Geneva)
Philippe Tulkens (EC Directorate-General for Research and Innovation)

14:10 – 15:40 **Plenary session "Robustness and legitimacy: setting the scene**," moderated by *Massimo Tavoni (European Institute on Economics and the Environment)*

10 min for presentations and 5 min for questions:

- NAVIGATE project and climate policy after the pandemic: *Elmar Kriegler (Potsdam Institute for Climate Impact Research)*
- Robustness and legitimacy: *Evelina Trutnevyte (University of Geneva)*
- Insights from risk analysis: *Roger Cooke (Resources for the Future)*
- Robust decision making under uncertainty: *Julie Rozenberg (World Bank)*
- Evaluating IAMs: *Charlie Wilson (University of East Anglia)*

15:40 – 16:05 **Plenary discussion with all speakers**, moderated by *Massimo Tavoni (European Institute on Economics and the Environment)*

16:05 – 16:10 **Introduction to group work** by *Evelina Trutnevyte (University of Geneva)*

16:10 – 16:30 **Break**

16:30 – 17:45 **Group work "Robustness and legitimacy of evidence on emissions gap"**
Three break-out groups with 5-10 min input presentations:

1. Robustness of existing evidence and areas for future work, moderated by *Detlef van Vuuren (PBL Netherlands Environmental Assessment Agency)* and *Volker Krey (International Institute for Applied Systems Analysis)*

2. Modeling-policy interface for legitimacy, moderated by *Michiel Schaeffer (Climate Analytics)* and *Jessica Strefler (Potsdam Institute for Climate Impact Research)*

3. Gathering new robust evidence on carbon neutrality, moderated by *Elmar Kriegler (Potsdam Institute for Climate Impact Research)* and *Christopher Weber (WWF Global Science)*

Day #2, Wednesday, 27 May 2020
Central European Summer Time

14:00 – 15:15 **Group work "Informing climate policy after the pandemic"**
Three break-out groups with 5-10 min input presentations:

1. Recovery packages and structural change of the economy, moderated by *Nico Bauer (Potsdam Institute for Climate Impact Research)* and *Jean-Francois Mercure (University of Exeter)*

2. Distributional implications of climate policy after the pandemic, moderated by *Johannes Emmerling (European Institute on Economics and the Environment)* and *Celine Guivarch (CIRED)*

3. Role of lifestyles and behavior change, moderated by *Sonia Yeh (Chalmers University of Technology), Bas van Ruijven (International Institute for Applied Systems Analysis)* and *Charlie Wilson (University of East Anglia)*

15:15 – 15:30 **Break**

15:30 – 16:30 **Plenary session "Robustness and legitimacy: outlook**,**"** moderated by *Detlef van Vuuren (PBL Netherlands Environmental Assessment Agency)*

10 min for presentations and 5 min for questions:
- Experience in the finance sector: *Ryan Barrett (Bank of England)*
- Experience at an NGO: *Patrick Hofstetter (WWF Switzerland)*
- Experience at a foundation: *Seth Monteith (ClimateWorks Foundation)*
- Insights from climate science: *Sonia I. Seneviratne (ETH Zurich)*

16:30 – 17:00 **Plenary discussion with all speakers**, moderated by *Detlef van Vuuren (PBL Netherlands Environmental Assessment Agency)*

17:00 – 17:15 **Wrap up and closing**
Miles Perry (EC Directorate-General for Climate Action)
Philippe Tulkens (EC Directorate-General for Research and Innovation)
Evelina Trutnevyte (University of Geneva)
Elmar Kriegler (Potsdam Institute for Climate Impact Research)

2.3. Workshop's participants

Over the two days, the workshop attracted **99 registered participants (by invitation only)** from 16 countries and 49 stakeholder and expert organizations. Table 1 provides the list of the organizations present in the workshop.

2.4. Messages from the opening session

During the short welcome session, **Elmar Kriegler (Potsdam Institute for Climate Impact Research)** welcomed the participants and described the goals and research plans of the NAVIGATE project (Chapter 1) as well as its stakeholder dialogue activities. **Evelina Trutnevyte (University of Geneva)** then introduced the University of Geneva as the host institution of the workshop and gave an overview of the workshop's objectives, program, and the participants (Chapter 2.1-2.3).

Table 1. The list of organizations whose representatives were registered for the workshop

Organizations in climate policy, business and civil society	NAVIGATE consortium	Research institutions outside NAVIGATE
• Agora Energiewende • Bank of England • ClimateWorks Foundation • European Commission, Executive Agency for Small and Medium-sized Enterprises (EC EASME) • European Commission, Directorate-General for Climate Action (EC DG Clima) • European Commission, Directorate-General for Research and Innovation (EC DG RTD) • Intergovernmental Panel on Climate Change (IPCC) • International Energy Agency (IEA) • International Labour Organization (ILO) • International Renewable Energy Agency (IRENA) • SHURA Energy Transition Center • United National Economic Commission for Europe • United Nations Framework Convention on Climate Change • World Economic Forum • World Meteorological Organization • World Bank • WWF Global • WWF Switzerland	• CIRED and CNRS, France • Chalmers University of Technology, Sweden • Climate Analytics, Germany • COPPE, Federal University of Rio de Janeiro, Brazil • E3Modeling, Greece • International Institute for Applied Systems Analysis, Austria • National Center for Climate Change Strategy and International Cooperation, China • PBL Netherlands Environmental Assessment Agency, Netherlands • Potsdam Institute for Climate Impact Research, Germany • RFF-CMCC European Institute on Economics and the Environment, Italy • University of East Anglia, United Kingdom • University of Exeter, United Kingdom • University of Geneva, Switzerland • WiseEuropa, Poland	• CICERO Center for International Climate Research, Norway • Delft University of Technology, Netherlands • Electric Power Research Institute, United States • European Commission, Joint Research Centre • ETH Zurich, Switzerland • Imperial College London, United Kingdom • Lawrence Livermore National Laboratory, United States • National Renewable Energy Laboratory, United States • National Technical University of Athens, Greece • North Carolina State University, United States • Pacific Northwest National Laboratory, United States • RAND Corporation, United States • Resources for the Future, United States • SINTEF, Norway • Stanford University, United States • University of Miami, United States • Vietnam Initiative for Energy Transition, Vietnam

In his welcome speech, **Philippe Tulkens (EC Directorate-General for Research and Innovation)** suggested that the models have never been so popular because the COVID-19 pandemic improved the general awareness about their usefulness for policy making and increased public trust in the work of scientists. This was likely to have a positive spillover on climate modeling in terms of understanding and appreciating its value and strengthening the links to evidence-based policy making. In his view, the workshop's focus on robustness, legitimacy, and transparency of models for climate policy assessment was extremely timely and he thanked the consortium for looking into this matter that EC Directorate-General for Research and Innovation strongly supported. He then spoke about how the pandemic had a profound impact on the climate modeling community by altering the key baseline assumptions on gross domestic product, employment, structural change

of the economy, capital markets and key financial variables, distributional effects, and deeper changes across the society like lifestyles. He argued that this poses new questions on how to incorporate these impacts in modeling, especially when dealing with incomplete data, incorrect assumptions, and huge uncertainty about the path and duration of the pandemic with mostly "known unknowns" and a large amount of "unknown unknows". Finally, he suggested that the modeling work and its robustness will be critical for informing the recovery phase and therefore it is crucial for the modeling community to step up communication and outreach efforts towards governments, the public, and industry in order to ensure that the climate agenda is not forgotten in this critical moment. This could be done by providing timely and evidence-based arguments in support of green recovery packages and being more vocal about the risks of climate change.

During the first plenary session "Robustness and legitimacy: setting the scene," **Elmar Kriegler (Potsdam Institute for Climate Impact Research)** also argued that the approach of integrated assessment modeling is more relevant than ever today because it enables a quantitative exploration of response options to public policy challenges. He drew parallels between modeling for climate change and for the pandemic, where both types of models tackle policy challenges, but on very different time scales and with very different response options. Both types of modeling involve active participation from sciences and from policy, where scientific advice is perceived consequential, but in parts rejected as elitist and technocratic. Policies for climate change involve gradual, long-term economic and regulatory response, whereas the pandemic is characterized by mandatory and disruptive short-term behavioral policies. As the impacts of the pandemic can be expected to have long-term effects on the society, economy, policy, and hence climate change, he argued that climate policy and sustainable development considerations need to play central role for recovery policies and stimulus packages after the pandemic and for addressing the rise in inequalities due to both causes. The IAMs hence need good representations of structural change of the economy, shifts in consumption, transient processes after a shock, and inequality implications. Science-policy and science-society communication activities building on transparency and validity need to be strengthened too. He closed his presentation with a discussion about the various lines of connection between IAMs and social science research, for example, on adding the institutional dimension or policy sequencing into models.

Evelina Trutnevyte (University of Geneva) spoke about the relevance and key challenges for robustness and legitimacy of models in climate policy assessment. She started with a reflection on three decades of climate mitigation modeling, where the increasing number of modeling teams, more complex and higher resolution models, wider range of acknowledged uncertainties, and increasing computational power led to thousands of mitigation scenarios available for research and policy making. She demonstrated the challenge with preliminary results from an analysis of 1'550 systematically collected scenarios of global PV growth from two IPCC databases and 190 other scenario publications. Enormous uncertainty in PV growth by 2050 was discovered and a large portion of it could be linked to the general meta indicators of the scenarios (e.g. date, location and type of organization), assumptions on climate and technology policy, or modeling approach (e.g. type, boundaries, and constraints of the model), rather than to more specific modeling assumptions. She emphasized that finding ways to assess and document the robustness of models and scenarios and to ensure their legitimacy would help distill more robust findings given the enormous uncertainty and proliferation of scenarios. Then, she gave an overview of the survey results on robustness and legitimacy (Section 2.7).

Roger Cooke (Resources for the Future) introduced the risk perspective by asking whether you would put your children on an airplane in which the authorities had medium confidence that the chance of a crash was less than 10%. First of all, he contrasted three types of views to risk and

uncertainty: the so-called Pollyanna's view that is consistently overly positive, the Chicken Little's view that is overly pessimistic, and the BOGSAT approach ("Bunch of Guys/Gals Sitting Around a Table"). In his opinion, the approach of the IPCC 5th Assessment Report with its statements that communicate qualitative probabilities (e.g. extremely unlikely, very unlikely) and confidence statements (e.g. high confidence) on the basis of qualitative synthesis of the authoring team's judgment is the closest to the BOGSAT approach. He then gave an example of the case of future sea level rise, where a structured expert elicitation in 2018 revealed broad uncertainties by the year 2300 and these uncertainties were larger than the full model spread of the IPCC 5th Assessment Report. The elicitation of structured expert judgements proved valuable to cover the various types of views by Pollyanna, Chicken Little, and BOGSAT. He concluded with the need in climate policy assessment and modeling to look at the tails of distributions, to quantify the uncertainty, and then to communicate it.

Julie Rozenberg (World Bank) spoke about how models can be used to draw robust policy insights, especially in developing countries. She first argued that climate policy in the developing countries needs to be assessed together with the wider development objectives, and this assessment is especially difficult given rapid changes, competing objectives, and uncertain futures in these countries. She then contrasted two approaches: (i) "predict then act" where models are used to estimate future trends in order to define the best near-term decision and to evaluate how sensitive this decision is to the prediction, and (ii) the approach of decision-making under deep uncertainty that instead looks at the available strategies and decisions first, identifies the vulnerabilities of these strategies, and then aims to adapt the strategies to reduce the vulnerabilities. She then presented two studies where this approach of decision making under deep uncertainty was applied. The first study modeled the impacts of climate change on households and poverty and explored the uncertainty pertaining to the future demographic and social economic changes. The study revealed that most of the uncertainty was in the baseline, which is difficult to communicate. The analysis of the drivers of poverty reduction raised new interesting policy questions, e.g. if the drivers of poverty reduction interact with climate change impacts and policy. The second study focused on assessing how much countries need to spend in infrastructure by 2030, where they found that sometimes model uncertainties are bigger than the difference between the climate objectives. She then concluded that there is a need to explore more and to communicate better the interactions between climate objectives and the baseline, as well as to focus on the absolute outcomes rather than incremental changes compared to the baseline.

Charlie Wilson (University of East Anglia) argued that IAMs are increasingly influential, but their fitness for purpose is not visibly evaluated in an open-ended process of testing and improving the structural validity (if a model is an accurate representation of the system response being modeled) and behavioral validity (if the model outputs are consistent with observational data). He then presented an overview of various evaluation methods of IAMs since 1970s, including historical simulations, near-term observations, generalizable historical patterns, hierarchy of models, model inter-comparisons and diagnostics, sensitivity analyses, and improvements in model documentation and review. All these methods have their strengths and limitations. For example, some methods, like generalizable historical patterns or sensitivity analyses, can test overall model performance, but other methods, like historical simulations, only test a subset of causal mechanisms or spatial scales. Other methods, like hierarchy of models or model documentation and review, are open to the evaluation by the third parties, whereas model inter-comparisons and diagnostics involve learning within the modeling teams. After discussing the strengths and limitations of various methods, he proposed to move from single- to multiple-method concerted evaluation, and from ad hoc evaluation to synthesis, protocols, and visible reporting.

2.5. Output from the group work on emissions gap

2.5.1. Robustness of existing evidence and areas for future work

This group work session, moderated by **Volker Krey (International Institute for Applied Systems Analysis)** and **Detlef van Vuuren (PBL Netherlands Environmental Assessment Agency)**, looked into the robustness and legitimacy of the so-called emissions gap analysis, i.e. the difference between the emissions under current policies (*where are we going?*) and the emission trajectories consistent with the well-below 2°C or 1.5°C goals (*where do we need to be?*). The moderators first presented the key methods used and also some strong assumptions in the current analysis, related to the pathways linked to temperature goals (cost-optimal pathways, limited consideration of social-inertia, dependency on the assessment of negative emissions and discount rates) and related to the current policy scenarios (different methods focusing either on international consistency or national detail).

In terms of the robustness and legitimacy, multiple aspects were discussed. The initial discussion focused on the **use of national assumptions on national economic projections**. Nationally Determined Contributions (NDCs) and pledges are often based on specific country-level socio-economic developments, which are not taken into account in the long-term temperature pathways. National assessments tend to compare scenarios based on national economic projections with those based on Shared Socio-economic Pathways (SSPs), despite possible large differences in underlying assumptions. In the discussion, this was indeed seen as an issue, although projects like COMMIT and ENGAGE are encouraging cooperation between global and national teams. This leads to better understanding of the differences, and possibly also adjusting the national and maybe global projections. It would be good if global projections could be updated more often.

In terms of **usefulness of cost-optimal scenarios as a reference for the emissions gap analysis**, the opinion of the participants was divided with arguments on both sides. The proponents of cost-optimal scenarios talked about transparency of method and simplicity of scenario design as well as the normative nature of long-term pathways, which is consistent with the method. The ideal pathway, even if it cannot be reached, remains a good reference. The cost-optimal scenarios were judged suitable at the global scale, but clearly more problematic at the regional and sectoral scale due to equity issues. The opponents argued that cost-optimal scenarios present a false reference because what is in the policies is not necessarily cost-optimal, and cost-optimality can have many meanings that are not based on a single metric of costs and can include, for instance, biodiversity, climate, food, or air pollution. There is also evidence that historical trends were not cost-optimal. The issue from financial perspective is that different modeling methods lead to very different macro-economic implications, such as the scale of investments, increasing the debt or waiting for technology.

In terms of **legitimacy of current policy and NDC estimates**, a diverse range of suggestions were made, including improved equity considerations in modeling for effort sharing, ex-post evaluation exercises that are helpful for policy makers to assess past success, transparency of assumptions and methods, as well as the approval of the outcomes of the analysis. For instance, the IPCC reports are accepted by parties and therefore not disputed in the negotiations, in contrast to other

analyses, including the UNEP Gap report. At the same time, there is a need for independent assessments, as it is hard to discuss country level results in the more official arenas. The IGST and projects like COMMIT are important, but one might need to think of ways to get these into the official negotiations. The open access to models was discussed here too, including the considerations if the funders would need to change requirements of grant agreements, or if effective open access requires good documentation. It was acknowledged that these are significant multi-year investments that are needed to make complex IAMs open access and there is limited funding for maintenance of energy systems models and IAMs (with some recent improvements).

In terms of **uncertainties and re-evaluation after the pandemic**, the key uncertainties to be taken into account when assessing emissions outcomes of NDCs are on socio-economic development, international cooperation, technology development, lifestyle and behavioral change, and changes in policy and governance. The group work participants unanimously voted that re-evaluation of current policy and NDC assessments after the COVID-19 crisis is required. It was then discussed how models could be made more robust towards developments of this kind in the future, including economic development and recovery, change in globalization in response to pandemic, post-pandemic measures and stimulus packages, lifestyle and behavioral change (e.g. remote working), rapid disruptive changes, and the use of large scenario ensembles as a method.

2.5.2. Modeling-policy interface for legitimacy

This group work session, moderated by **Michiel Schaeffer (Climate Analytics)** and **Jessica Strefler (Potsdam Institute for Climate Impact Research)**, looked into processes for ensuring modeling-policy interface for legitimacy. The moderators first presented a brief introduction, highlighting an uncomplete set of five dimensions of legitimacy of models and their results in a policy making context: realism, relevance, fairness, transparency, and robustness. The discussion then focused on questions: (i) what additional dimensions or aspects are essential, and (ii) which, among those that were presented and discussed, are the highest priority and could potentially even disqualify a model or a result if missing. Most aspects brought forward by the participants were related to realism (representing state and trends of countries and sectors), relevance (timely and topical), and robustness with uncertainty. There are no general reasons to disqualify models, but there may be specific reasons for specific persons, and this could be further looked into using the results presented earlier in the opening session by Evelina Trutnevyte on solar PV growth scenarios (Section 2.4). Instead of disqualifying models, a practice of distinguishing different classes and levels of detail for different questions may be helpful.

In terms of **realism**, it was acknowledged that the models need to produce results that are recognizable, in particular in the national context. On the one hand, the models should reproduce the current state of and trends in energy system and land use for the stakeholders to recognize the data and to understand why the data is different from national inventories and other national and international sources. Models have to have the latest data and clearly show fundamental characteristics of the (national) economy, for instance, bioethanol in Brazil. Calibration to historic facts is usually useful, but can also be problematic: some trends change over time, for instance, the assumption that no technology can grow faster than 10% per year was invalidated by observed trends in renewable technologies. There are also differences to be understood in emissions inventories. Nonetheless, there is a window of opportunity in the background of the COVID-19 pandemic, where there seems to be raised confidence in models in general, from which the modeling groups could benefit. But the studies should be very carefully framed at a national level to “prove” suitability for a specific country.

In terms of **relevance**, the suggestions included the need for mitigation goals to be clear and timely, assessment of controversial and desirable mitigation options on the supply side and demand side, the explicitness about costs, (co-)benefits, investment needs, and returns on investment, as well as the assessment of impacts, damages, adaptation, together with the synergies and trade-offs. It was acknowledged that the policy makers and other stakeholders, including the financial sector, need timely results, which could be inconsistent with the desire for scientific robustness. However, such results may also be dismissed by stakeholders and countries that do not have full confidence in the institutions that produce them. Relevance is enhanced by speaking to the various sectors adequately. Indicators used may be robust (e.g. the effects of gross domestic product on mitigation actions), but not granular enough to be of use. This lack of granularity may also obscure the meaningful results and hide uncertainties apparent in more granular data that the stakeholders should be aware of. The modeling groups need to anticipate policy questions and prepare models and methodologies accordingly. The creation or loss of jobs due to policies is important information for policy makers, as are the distributional impacts. Results must relate to broader policy concerns and include co-benefits and synergies across policy domains.

In terms of **robustness and uncertainty**, a distinction was made among scientific basis of models (e.g. do models include the necessary technologies, policies, or granularity), understanding of models and the results (e.g. understanding the differences between models and robust patterns across models), and robustness of the results (e.g. how robust are the patterns, such as that the availability of nuclear power has little influence on mitigation costs). Many participants brought up a desired heterogeneity of models, and a need for honest assessment of how heterogenous is the selection of models and approaches in a specific study. If model structures are very similar, then the models agree and robustness in terms of models arriving at the same conclusions is very thin. Diversity does not only cover the selection of models applied in the study, but also the scenario space. Shared Socio-economic Pathways (SSPs) represent rather a sensitivity analysis than systematic exploration of scenario space, for example, when models do not cover different economic views, such as de-growth or at least no-growth.

In terms of **fairness,** three dimensions of interest were distinguished by the moderators: modeling of (i) the distribution of costs and benefits between countries and income groups, (ii) the fairness of mitigation action, including the distribution of economic costs, investments, other costs (e.g. jobs), co-benefits, and negative side-effects, and (iii) fairness of impacts and adaptation needs in terms of the distribution of climate risks, damages, adaptive capacity.

In terms of **transparency**, the moderators acknowledged the importance of a clear communication of assumptions and application of the sensitivity analysis, for instance, on which technologies are available and at which costs, potentials, and side effects. The key uncertainties for communication and scenario analysis were suggested to be socio-economic trends and SSPs, future costs and potentials, as well as the remaining carbon budgets. The open source and open access dimensions for model documentations were also mentioned.

2.5.3. Gathering new robust evidence on carbon neutrality

This group work session, moderated by **Elmar Kriegler (Potsdam Institute for Climate Impact Research)** and **Christopher Weber (WWF Global Science)**, focused on carbon neutrality, that is, the goal of net zero CO_2 emissions of a given entity (e.g. the World, country, region, company). This

goal can be based on direct emissions (production-based emissions; Scope 1) or based on direct and indirect emissions (consumption-based emissions; Scope 1-3). The first approach is more common. The virtue of net zero CO_2 targets is that they can be applied across scales, not only globally. The timing of the net zero is fairly robust against near-term policy assumptions, if temporary overshoot is allowed, but it can be earlier in delay scenarios, if peak warming limit is imposed. It is more robust than carbon budget estimates because +/- 200 $GtCO_2$ translate to +/- 10 years. National pathways to net zero CO_2 are explored in national energy models and IAMs, especially since the IPCC Special Report on 1.5^oC warming, and it was a key piece in the Conference of Parties (COP 26). This target has long been common for non-state actors, for instance, Science Based Targets Initiative for companies or Net Zero Asset Owner Alliance for investors. The issues with net zero targets for non-state actors and similarly for countries revolve around the types of gases covered (accounting for non-CO_2 gases or not; the use of global warming potential or dynamic accounting), timing (long-term, interim, and pathway as a target), offsetting (if allowable and to what extent), equity (who must move and at what rate), and technology dependency (e.g. carbon dioxide removal). It was also discussed how long the net zero CO_2 is to be maintained as it could also be seen as equilibrium or end state of society, possibly aiming for slightly below net zero CO_2 to compensate the remaining emissions of other long-lived greenhouse gasses. The role of carbon dioxide removal and how it relates to net zero CO_2 was also discussed.

In terms of **legitimacy of IAMs for exploring carbon-neutral futures** and pathways, the key modeling elements were perceived to be granularity (if necessary regions, sectors and technology details are covered to capture carbon-neutral systems at the scale of interest), systems and policy dynamics (if there is a sufficiently accurate description of the interaction between systems and policy pathways towards carbon neutrality), transparency and validity (if there is a publicly available and expert-reviewed model documentation at a high level and in detail, and what is the track record of applications), and uncertainty quantification (if there is an ability to explore parameter and scenario sensitivities of pathways towards carbon neutrality). The key variations in socio-economic, policy and technology assumptions that should be explored in scenarios for robust assessment of carbon neutrality were suggested to be as follows: socio-economic pathways and inputs (including transformative ones); policy coverage, effectiveness, timing and heterogeneity; technology parameters and limits; alternative management practices; energy system (including energy efficiency and conservation, deep electrification, and sector coupling, biofuels), industry and buildings (including carbon capture and utilization, carbon storage in materials and buildings), and land (plantation versus sustainable forestry, agricultural practices, and soil organic carbon enhancement).

In terms of the **technology and sector perspectives**, it was suggested that carbon storage in materials is important to include. It has been missing so far and thus the full flexibility and bottlenecks in achieving net zero CO_2 are not visible. For example, refineries are needed to produce petrochemicals, but those can be replaced by biochemicals when introducing new biomass conversion routes, lowering utilization factors of petroleum refineries, and making them uneconomic. This indirect effect of introducing biofuel is as important as the direct effect. Same holds for green H_2 routes, for instance, when H_2 from solar PV drives electrolysis, or CO_2 from carbon capture and storage or direct air carbon capture and storage, outputting olefins.

In terms of the **regional dimension**, countries are asked alongside their NDCs to provide long-term development strategy for low greenhouse gas emissions and every country then decide how it will achieve its target as there are many different pathways. The parties need to be specific with their interpretation of net (zero) emissions in declaring their ambition as otherwise it can have various meanings. The important topics for countries in the context of pathways towards net zero

emissions are cost reductions of renewable technologies and the role of biomass. Removals have not been discussed as much, but this is expected to change. In terms of how granular the regional dimension should be, the country level was perceived to be the most relevant for climate policy makers, but it is less clear for non-state actors. Other relevant dimensions were on how leapfrogging could be assessed for less developed countries on combining development pathways with net zero targets, and how much residual emissions are wanted to balance regionally and sectorally and who would be balancing in terms of distributional and equity considerations.

It was discussed that there is a need for multi-model sensitivity analysis to explore these questions and inform policy making and for regional and sectoral target setting, as well as for both global and national modeling teams working in concert to link global target with national climate strategies (via global stocktake to temperature targets). National-level models, even relatively simple ones, are very important to support countries in the design of their NDCs and decarbonization strategies. Global models have showed that aiming for carbon neutrality sometime around 2050 is a robust target, and countries are now looking for tools that help them understand what carbon neutrality means for them, and design their own strategy. Global models are still very useful to consolidate and monitor progress, but in the short run, the focus should be on having tools to explore efforts needed at a national level. A wish was expressed that these tools and models could be simple enough to be used for robust decision making, that is, stress-testing the impact of different policies and public investment strategies to a wide range of uncertainties. On the other hand, they should be detailed enough to test other policies than just carbon prices. An open question remained if these models should be sectoral models or if they can be integrated at national level to strike the right balance between robustness and accuracy.

In terms of **transparency and validity**, the discussion, first of all, focused on the extent to which the models are fit for purpose. The International Energy Agency's model was discussed as an example: it has good granularity, systems and policy dynamics, but focuses on energy. Hence, it is of more limited value for the analysis of net zero CO_2, which should include consideration of CO_2 storage in materials and terrestrial CO_2 storage. Further issues were raised on the transparency and uncertainty analysis of the model. Second, a trade-off between comprehensibility (and possibly robustness) as well as complexity was acknowledged as models cannot solve all key questions. There is a need for a combination and hierarchy of national and global models as well as a combination of "heavy" models representing necessary technological, sectoral, and policy processes in detail and "light" models for exploring key uncertainties. Analogy to the use of Earth system models and reduced climate models by climate modeling community was drawn here. Overall, it was seen important to invest in capacity building in countries and regions that are less advanced in research and modeling of transformation pathways to net zero CO_2.

2.6. Output from the group work on informing climate policy after the pandemic

2.6.1. Recovery packages and structural change of the economy

On the second afternoon of the workshop, the group work sessions focused on informing climate policy after the pandemic and the role of models. The first session on recovery packages and structural change, moderated by **Nico Bauer (Potsdam Institute for Climate Impact Research)** and

Jean-Francois Mercure (University of Exeter), started with the joint presentation by the two moderators on the key dimensions that are relevant for modeling. The first dimension was about the type of recession that could be expected after the pandemic, including the effects of the pandemic as well as of the lock down, supply and demand shocks, and disruptions in the supply chains and trade, and in the labor market and investments. The challenges before the pandemic were also acknowledged, such as partial recovery after 2008/2009 economic crisis, low labor productivity, youth unemployment and public debt in southern Europe, and interest rates that are lower than the growth of gross domestic product. The second dimension was on the shape of the macroeconomic recovery, where various possibilities are still open, such as V (temporary shock with fast recovery), W (long recovery with double dip recession), L (permanent reduction of economic activity), and J (recovery that is stronger than the shock). There are various drivers that feed into the shape of the recovery, such as monetary and fiscal policies, labor markets, and policy response, including Green New Deal. The largest structural changes could be expected in services, such as increasing productivity (e.g. home office, teleconferencing, e-payment), decreasing productivity (e.g. physical distancing in indoor spaces, mass gatherings), demand (e.g. tourism, transport, hospitality services), labor (e.g. health service jobs), as well as trends in manufacturing, such as shortening of the supply chains and unclear effects on globalization of consumer products. The shock, recovery and the structural change, can be implemented in models either by means of exogenous assumptions (e.g. anatomy of the shock, near-term trends and extensions, policy response) or as emerging recovery scenarios (e.g. dynamic response of the economy and structural change). Examples of previous and ongoing studies on modeling shocks and recovery were shown.

The discussion on the **shape of the recovery** revealed rather pessimistic expectations with most people expecting a long and bumpy road towards economic recovery (Figure 2). Some participants even expected a persistent reduction of economic growth and nobody expected a recovery that would lift economies above the previous growth potential. Most participants expected a reduction of service demand and a mixed picture on the productivity developments in the service sector. In the industry sector, most participants expected a shortening of international supply chains and trade. The Green New Deal was perceived as a credible political initiative that addresses the transition towards a sustainable economic system, but it was not seen as a mean to push the digitalization agenda.

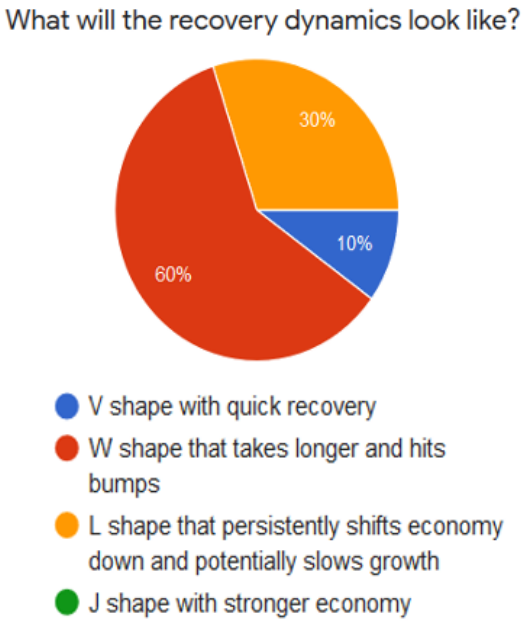

Figure 2. Poll with 20 participants of the group work on how the recovery dynamics would look like.

The participants brought up a number of issues that mostly addressed **near-term and fiscal issues**, such as labor developments and business default, the role of debt for economic growth, or differentiation of private and public sector investments that are not fully captured by models. It was highlighted that the sum of announced stimulus packages is twice as much as during the 2008/09 crisis and a comparison with IAM investment requirements for climate protection reveals a difference of a factor of 30. The questions were also posed whether there could be a green recovery without public spending programs (particularly in the land-use sector). Some participants saw the introduction of a carbon tax as a challenge in the current economic crisis phase.

Regarding **model legitimacy**, the question was brought up whether IAMs are appropriate tools given their high level of aggregation and long time steps. The models used particularly for the COVID-19 related issues in NAVIGATE are not typical IAMs with 5-year time steps, but are based on single-year time steps. For the NAVIGATE analysis, a strong recommendation was given that several different recovery scenarios should be developed, and the narratives of these recovery scenarios need to be well introduced because the outcome is still uncertain (Figure 2).

2.6.2. Distributional implications of climate policy after the pandemic

This group work session, moderated by **Johannes Emmerling (European Institute on Economics and the Environment)** and **Celine Guivarch (CIRED)**, started with a presentation on the key issues related to the coverage of inequality in climate policy assessment after the pandemic. Three dimensions of inequality were acknowledged: distribution (income, wealth, between and within countries, intergenerational issues), impacts and vulnerabilities, and policy incidences. Some channels of the pandemic's impact on inequalities can be expected through employment and wages, capital income and wealth distribution (financial market dynamics, continued stock participation, and the equity premium puzzle), capital-intensive carbon-neutral and digitalized consumption goods, education, and digital divide. There is emerging evidence that a significant and persistent long-term increase in inequality are anticipated. The poor households are hit the most, and unemployment spike affects the lower skills more severely. The policies of fiscal stimulus have strong distributive implications in terms of transfers, social benefits, assistance for small and medium enterprises, social insurance systems, support for companies and self-employed, and monetary policy impacts on interest rates and equity premium. In terms of inequality in modeling for climate policy assessment, the models could for example consider wage premia and capital income dynamics, progressive climate policy costs, alternative policy instruments with different incidences, and analyze households by deciles, as well as include distribution of climate impacts more broadly.

The discussion afterwards focused on two questions: how the pandemic has changed inequality issues in Europe, within and between countries, and how this is interacting with the distributional implications of climate change impacts and climate policies. In terms of **the impact of the pandemic on inequality issues**, the participants saw that many different dimensions of inequality are affected: income, labor, geographic inequality, and the environment. The dominating channel of impact was perceived to be employment, acknowledging the interplay between education, employment, and inequality. For instance, in terms of the future of education of middle- and lower-income households, ineffective online learning or homeschooling for some time could have trickle-down effect, although the longer-term impacts are yet to be understood. Another example was discussed for the case of Brazil in terms of the design of the recovery packages. In terms of amounts of financing, the policies seem to be strong enough, but their reach and efficiency are not so effective, given lots of overlaps between policy instruments and low access to financing. There is an ongoing research by the COVID-19 International Small Business Study that gives some perspective on these policies for Brazil as well as for other countries.

In terms of the **interactions between the pandemic and distributional implications of climate change impacts and climate policies**, there are inequalities at very different levels within and between countries to be considered: gender inequality, health inequality, impacts on trade and trade policies, localization of production and hence labor, geopolitical implications, asymmetric sectoral impacts (e.g. air travel, tourism), and increased vulnerability because of the lockdown.

There are lessons to be learnt from the pandemic for climate policies, for instance, that the policy responses can be much faster than previously thought or that air pollution and COVID-19 have brought attention to the pollution issue. It was perceived that now it could be the right moment for carbon tax because oil prices are low, but this would need to be done distributionally in a just and acceptable way. Policy packages for recovery that are being designed, often integrate elements of green recovery. Behavioral change during the pandemic and its policies has been unprecedentedly fast, throughout countries and populations, and shows how quick it can happen. There are still open questions how long the effects will last, and will the societies go back to normal. From a modeling perspective, IAMs are not able to model shocks, so other modeling methods should be considered, like agent-based models or network models. Overall, the evolution of the crisis is still largely unknown, and it is too early for precise or definitive conclusions. The least that can be said, is that the topic of inequality is taking even more importance than before. However, it remains unclear whether and how it affects the issue of distributive impacts of climate policies.

2.6.3. Role of lifestyles and behavior change

This group work session, moderated by **Sonia Yeh (Chalmers University of Technology)**, **Bas van Ruijven (International Institute for Applied Systems Analysis)**, and **Charlie Wilson (University of East Anglia)**, looked into the questions of what does climate policy need to know about lifestyles after the pandemic, what are the challenges for modeling lifestyles after the pandemic to inform climate policy, and what are the implications for the robustness, legitimacy, and usability of current modeling approaches. After short introductions by the moderators, the inputs of the participants were collected via three poll questions:

- What does climate policy need to know about lifestyles after the pandemic?
  1. The lasting effect of the pandemic on lifestyles and emissions [vote score = 6.6/13]
  2. The effect of the pandemic on public acceptability of climate policies which directly or indirectly affect lifestyles [vote score = 5.5/13]
  3. The effect of the pandemic on global convergence of lifestyles across countries [vote score = 4.5/13]

- What are the challenges for modeling lifestyles after the pandemic (to inform climate policy)?
  1. The same as it was before the pandemic: better endogenous representation of lifestyles [vote score = 7.4/12]
  2. Keeping track of empirical research about lasting effects of pandemic on lifestyles [vote score = 6.4/12]
  3. Capturing inter- and intra-regional heterogeneity in post-pandemic lifestyles [vote score = 5.7/12]

- How can IAM modeling of lifestyles improve robustness, legitimacy, and usability to be helpful for policymakers, business, and civil society?
  1. Need for a much wider exploration of the future possibility space in scenarios and modeling [vote score = 7.9/12]
  2. Ignore near-term reproducibility of IAM results and focus on capture long-term dynamics and path dependence [vote score = 5.1/12]
  3. Improve the transparency, traceability and accessibility of the modeling itself as well as the design of participatory processes [vote score = 8.5/12]

In terms of **improving the modeling of disruptions and extremes**, it was acknowledged that

scenarios and modeling have been caught unprepared in terms of pandemic impact on lifestyles (as have many other fields). IAMs were not designed to predict the future, but to explore the long-term implications. This has made it very clear that the IAM community needs to explore much more disruptive futures. The sharp shock nature of the pandemic requires modelers to think specifically about lasting impacts of short-term shocks. Modeling can and should pay more attention to futurists who explore future disruptions more systematically. However, there was no demand for lifestyle-shock scenarios prior to the pandemic. IAM community emphasizes consistency in scenarios based on observations in the past, but shocks reveal that our beliefs about consistency are often flawed or not very stable or robust.

In terms of **modeling lasting changes to lifestyles and related economic activity**, enduring impacts on lifestyles – particularly travel patterns – are relevant and should be modelable. There will be certainly enduring impacts of the pandemic which need accounting for. Modeling has looked at lifestyle changes (e.g. travel patterns), but not at short-term disruptions of this magnitude. Supermarket chains developing scenarios of consumer preferences and shopping behavior are not anticipating return to pre-lockdown and there is a clear anticipation of removing or reducing global supply chains. Economic slowdown and change in mobility practices (globalization) were happening over the past decade, with counteracting effects on energy, land, and emissions. The pandemic has intensified both trends with potential unintended impacts on emissions. The pandemic has also impacted behaviors which were common before as well as behaviors which are adaptive responses to the lockdown. It is unclear if the pandemic has changed the public appetite for more stringent climate policies, or for more regulatory top-down type policies.

In terms of **challenges for modeling lifestyles,** modelers treat demand often as a function of simple gross domestic product and price relationships. This is inadequate for capturing lifestyle change and disruption. Economic activity solely modeled by gross domestic product does not capture the real economy, particularly informal employment which is likely to rise significantly. Pandemic has sharpened inequalities of access to, for instance, digital infrastructure. Social heterogeneity is important for models to try and do a better job at capturing pandemic impacts. It is even more important for the most vulnerable to have a voice in the post-lockdown scenario and modeling efforts. There is a clear emphasis on risk, equity, impacts, and co-benefits. Social sciences are also playing an increasing role in UNFCCC discussions, and should do so in IAMs. Models optimize systems, whereas lifestyles are not amenable to optimization-based approaches. Any model solutions for a global optimum imply that regions deviating from this best trajectory are 'suboptimal' and this approach risks undermining the legitimacy of modeling, if regional lifestyles may not be optimal from a global emissions perspective.

In terms of **strengths and limitations of IAMs for modeling post-pandemic lifestyles**, the approach based on the hierarchy of models also recognizes that different models (used jointly) can be used for different purposes. IAMs cannot and should not do everything. Many different modeling tools are needed to explore response to the pandemic. Models which give rise to emergent behaviors are useful for defining scenario stories which IAMs can interpret and quantify. Pandemic has very clearly shown the interconnections between different activities, sectors, and countries. This highlights the importance of IAMs as system integration tools. Modelers often work in timesteps of 5-10 years, so they should distinguish whether these pandemic causes a short-term blip (which is barely noticeable and should not be modeled) or whether it causes structural change (which is definitely noticeable and should be modeled). Models have typically taken a longer-term interest in path dependence and inter-sectoral tradeoffs. But this is often mismatched with the needs of decision makers in particular sectors

(e.g. land-use planners) or climate policy makers. Regional IAMs are one useful tool for responding to the needs of policy makers. Therefore, the challenge is for modelers to move away from highly aggregated (in time and in space) tools to tools which are more closely aligned to the users' needs.

In sum, the group work concluded that enduring effects of the pandemic on lifestyles are likely, but not yet known. The effects of the pandemic on climate policy acceptability are possible, but not known either. The main research needs now are empirical (to track enduring effects) rather than modeling. There can be challenges for modeling if, and only if, the effects on lifestyles are long-term and/or if there is a need to better capture the underlying drivers, motivations and distributional impacts of lifestyle change. These modeling challenges simply add to a suite of existing challenges for better capturing lifestyles and other people-centered issues of equity, access, risk, co-benefits. But IAMs are not the only tools in the box, and should be used in conjunction with the other tools with relevant purposes and designs for lifestyle modeling. More generally, the IAM community should be more diligent and conscious in making models more transparent and communicate their limitations with respect to lifestyles modeling to make them useful for policy makers.

2.7. Messages from the closing session

The closing plenary session "Robustness and legitimacy: outlook" included presentations by the intended users of IAM outputs. **Ryan Barrett (Bank of England)** spoke about using the outputs of IAMs from the perspective of a central bank with the objective to assess impacts on the economy (e.g. inflation, growth, structural change), financial stability (e.g. whether there will be a financial crisis), and financial firms (e.g. how firms manage risks). He suggested that IAMs are suitable for macro-financial decision-making due to their energy and land use output granularity, standardization of climate outputs and modeling frameworks, and communication of the key scenario drivers and policy recommendations. However, in his view, the economic resolution and narrative are lacking in IAMs, including sectoral granularity (e.g. impact on gross value added, structural change), the range of outputs (e.g. impact on capital, labor, technology business investment, consumer spending), integration of transition and climate impacts, and effects of fiscal and monetary policy. IAMs also do not sufficiently cover regional granularity and disorderly pathways with policy fragmentation, delayed response, imperfect information, and frictions like unemployment, capital stranding, or inflation. In terms of legitimacy of IAMs, he argued that four elements are important: (i) addressing model uncertainty because a high agreement in modeling results could be due to similar structural assumptions in models; (ii) addressing parametric uncertainty, where carbon price and impact on gross domestic product could be also shown for different technology price assumptions and not only for the likelihood of reaching temperature targets, (iii) improving the understanding of outputs because outputs like carbon price are misunderstood and misapplied, (iv) improving transparency of model assumptions so that it is possible to understand what is driving the results. For the last point, he suggested that formats that are in between a high-level summary for policy makers and hundreds of pages of model equations would be useful. He finished his presentation by giving the example of the Network for Greening the Financial System that is working to produce new scenarios for macro financial analysis.

Patrick Hofstetter (WWF Switzerland) spoke from the view of climate advocacy, especially referring to the kinds of information that is necessary for national-level policy making. He gave examples of the complex knowledge gaps in the case of Switzerland, such as the impact of potentially abandoning herding cows in the Alps during the summer, avoiding reforestation on the

meadows below the tree line, and accounting for the effects of soil carbon, forest carbon, and change in albedo. He argued that there is a need to make decision makers more aware of the their direct responsibilities by relating marginal emissions to concrete outcomes: for instance, if a new legislation mitigates 10 million tons CO_2, how many premature deaths or how many climate refugees will this avoid. From his point of view, IAMs do not focus enough on the relevant policy level which is mostly national level. IAMs also do not allow policy makers to play with various applicable policy instruments to anticipate their potential impact. He argued that the existing models in Switzerland mostly focus on the economic and market-based instruments, but do not cover or poorly cover bans, standards, non-monetary incentives, or sufficiency. He also asked if barriers and loopholes as well as disruptions and shocks are considered in the models and whether they should be considered, especially if the models produce scenarios until 2060. He then argued that for policy makers distributional effects do not end at the level of household income and employment, and that there is a need to analyze the profits from an instrument and who carries the burden. He concluded his presentation suggesting that there is a need to be careful with the credibility of science, to avoid black boxes and be transparent about the models and parametric assumptions, to have decision makers in mind when selecting endpoints and scenario alternatives, to allow for creativity so that the sets of policy instruments and their features can be freely selected and modeled, and to include policy instruments that enable a fast transition.

Seth Monteith (ClimateWorks Foundation) spoke about the use of IAMs for amplifying climate philanthropy by means of education about the current situation and by means of discourse on where the funding should be allocated for mitigation. The level of funding in terms of spending per unit of abated tons of CO_2 varies between the continents as well as mitigation options (e.g. power, transport, forests, etc.) and it depends on the relationship with funding organizations, the capacity within the network of grantees, windows of opportunity for policy action, co-benefits for secondary goals, and financial or operational barriers. In terms of robustness and legitimacy of IAMs, he argued that expanding the use cases for these models comes with a risk and that there is a need to reflect whether IAMs are always the tool for new questions because these models need to remain true to their original use case and their development and maintenance are important for transparency. He spoke about the challenges in communicating results, especially when there is a risk of decision paralysis without guidance from those who understand the complexities. He also suggested that there is a need to update Shared Policy Assumptions because mixing and matching Representative Concentration Pathways, Shared Socio-economic Pathways, and Shared Policy Assumptions is not straightforward for generating a baseline that allows to assess policies. Further development of Shared Policy Assumptions and a continuous identification of policies to be included could result in more up-to-date baselines. He finished with suggestions for IAMs to improve sectoral representation for policy assessment, including industrial coverage and integration, aviation, shipping, and feedbacks between carbon dioxide removal technologies, land use, food system, and behavior.

Sonia Seneviratne (ETH Zurich) brought the perspective from climate science and in particular focused on including changes in extreme events in IAMs. She argued that the robustness and legitimacy of IAMs is strongly affected by the lack of inclusion of climate extremes because extremes behave differently from mean climate and the impact models may not cover them well. IAM land use scenarios affect climate extremes through biophysical feedbacks and not only through carbon cycle, but these processes are not integrated. Extremes affect ecosystems, for example, when a single fire or draught can destroy a large part of a forest and annihilate several years of CO_2 storage, or when extremes affect production of biofuels, food and realism of scenarios with Bio-Energy Carbon Capture and Storage. Extremes also affect people, for instance,

when a heat wave or a heavy rain associated with tropical cyclone can kill people and destroy likelihoods. Single extreme events can also contribute to instability, conflict and migration, such as the Syrian drought in 2006-2011. Extremes can also affect energy production, when there is a lack of cooling for nuclear reactors or there is an increase in energy demand for conditioning during heat waves. She then spoke about the relation between global temperature and stronger warming of extremes in land hot spots. She reflected whether the changes in the land use from current IAMs are realistic given changes in extremes and what is the role of biophysical effects like albedo and evapotranspiration. She argued that the main issue for integrating extremes, including impact and feedbacks, in IAMs is the present decoupling between IAMs and Earth System Models, and wondered whether the progress could be accelerated using an emulator for climate models, including variability and extremes. She presented the Modular Earth System Model Emulator with spatially Resolved output (MESMER) and suggested that follow-up version of this model could include extreme indices. She concluded that such an integration of extremes into IAMs would be a new frontier for research and collaboration among Working Groups 1, 2, and 3 of the IPCC.

For the final closing comments, **Miles Perry (EC Directorate-General for Climate Action)** reflected on the new European Green Deal and recovery after the pandemic, where he stated that the work on the climate transition is not slowing down. He argued that such a rapid drop in growth and emissions after the pandemic is not sustainable and spoke about the role of IAMs for demonstrating mitigation pathways that would be more economically and politically sustainable.

Philippe Tulkens (EC Directorate-General for Research and Innovation) reiterated that, while the workshop was planned before the pandemic, it was extremely timely due to its focus on robustness and legitimacy as well as the impact of the pandemic in IAMs. He argued that transparency is central and that there is a consensus in the community about its importance, so now concrete steps to implement it are needed. He perceived that the NAVIGATE consortium is well placed to make progress in this area, but should go beyond model documentation. The consortium could consider proposing concrete and workable business models, where their income is not related to the restrictions of the usage of their codes by others, and the EC can work with the consortium on this question. He asked if the NAVIGATE consortium would agree individually or collectively to be fully transparent to the EC at least because this would be a good sign, appreciated by policy makers. In terms of robustness, he spoke about the concerted multi-level evaluations of models, which are already funded, but need to become more systematic and have a broader scope. As one example, he suggested to broadly look at ex-post evaluation of European climate policies since their inception in 2000, where concerted multi-IAM evaluations could have their place. In terms of communication, he suggested that policy advice needs to be policy relevant and timely and that the modeling community in European Union-funded projects could be ready to provide the right input when it is needed. The pandemic gives an unprecedented opportunity for the community to provide input into recovery packages, making sure that they are green and sustainable, and supporting Nationally Determined Contribution and long-term strategies due under UNFCCC. This possibly requires going out of the comfort zone and involving in more active communication that is adapted to the audience, including society at large, and is not limited to academic-style papers. Finally, he spoke about the modeling work that is needed to make sure that models are fit for reflecting the post-pandemic reality, including improvements in the demand sector (behavior change, lifestyles), structural change, inequality, shocks and integration with health models, as well as more controversial issues related to the key economic assumptions, such as selective degrowth or stable-state economy moving away from full employment.

Evelina Trutnevyte (University of Geneva) then summarized her highlights of the two afternoons.

In terms of robustness and legitimacy, there was an overall consensus in the workshop that this is a timely topic. It has become even more timely since the pandemic due to the momentum and appreciation of modeling for policy making, but this also raises the requirements to ensure robustness and legitimacy of models. From the workshop, the topics for suggested future work to improve robustness and legitimacy could be grouped in three categories: (i) further development of models, (ii) new methods for uncertainty, robustness, and evaluation, and (iii) matters related to the process and broader participation. In terms of model development, there were multiple suggestions during the workshop on how IAMs could be improved for robustness, including better representation of the effects of the pandemic, equity and distributional considerations, sector granularity, extremes and climate feedbacks, as well as ensuring that IAMs produce policy-relevant outputs, especially at the national level. In terms of methods for robustness, the suggestions included the hierarchy of simple and more complex models, methods from decision making under deep uncertainty, expert elicitation, multi-method concerted evaluation exercises, and ex-post historic evaluations. In terms of the process and broader participation, the workshop again showed that there is an inevitable momentum towards open access. The robustness and legitimacy, on the one hand, are ensured through peer review, science committees, and the modeling community exchanges, but there is also a need for broader use of models for education and discourse to grow the community and to ensure there is a two-way exchange between policy, stakeholders, and the modelers. She finished by saying that this workshop was exactly the attempt to create such a two-way exchange and thanked everyone for their extensive participation.

Elmar Kriegler (Potsdam Institute for Climate Impact Research) concluded the workshop by saying that the workshop showed the richness of topics that can be discussed with IAMs and hence proved the integrating role of models. IAMs can quantify the key drivers and policy impacts of interest, and they can also help dive deeper into specific topics. In terms of legitimacy, he reiterated the importance of transparency, where models like GCAM, Message, or MAgPIE/Remind move in open-source direction in addition to high-level documentation. Open-source approach requires extensive time and specific funds though. For robustness, a flotilla of models is needed where IAMs are like large tankers and there are many other smaller boats that address specific questions. Large models with high complexity can be used for policy analysis, whereas simpler models can be used for uncertainty analysis or for education. He then argued that the modeling community needs to take up the question of the pandemic. Finally, he thanked the participants and organizers and closed the workshop.

2.8. Results of the online survey

In order to collect wider inputs of the workshop participants as well as external stakeholders and experts outside the workshop, an open-ended online survey was distributed in February-May 2020. On the one hand, the survey was sent to all registered participants and invitees of the workshop, the NAVIGATE advisory board, the NAVIGATE consortium members, other H2020 projects (e.g. Locomotion, Paris Reinforce), and through the internal lists of stakeholders of the NAVIGATE members. On the other hand, the survey was distributed via the established mailing lists of the communities of Integrated Assessment Modeling Consortium, OpenMod, Decision Making under Deep Uncertainty, Sustainability Transitions Research Network, Strommarkttreffen, Energy and Social Science Network, and World Economic Forum's Expert Network.

The survey received 50 responses from 16 countries in total, including 36 researchers and 6 policy, 4 business, and 4 other stakeholders. There were 24 responses from integrated assessment modelers and 14 responses from other modelers. Three quarters of the responses came outside the NAVIGATE consortium.

Tables 2-4 present the summary of the results from the survey on the three key themes of improving robustness and legitimacy: improving models and methodologies, data and reporting, and communication and process. There were multiple topics that repeatedly emerged in the survey responses: the models should be fit for purpose as one model cannot be robust for all questions; there is a needed for extended sensitivity and uncertainty analysis; the data should be of good quality, up to date, and open; the models should undergo ex-post validation; the models, methods, and assumptions should be well documented, open, and reproducible; the limitations should be documented and communicated; the conditionality of model outputs to the assumptions should be transparent; the models should be driven by user demands; and the models and publications should be peer-reviewed. Divergences were also observed. Some respondents argued in favor of simple models that are understood and run easily, whereas others suggested that the complexity of climate policy requires comprehensive and hence complex models. In terms of reporting the results, some called for simple and short messages on main drivers and their effects, whereas others argued that models should be used for insights and thus the results would need to be communicated together with more extensive information about the assumptions and limitations. While there was an overall agreement about acknowledging uncertainty, some respondents argued that uncertainty analysis should be the tool to expand thinking to avoid a false sense of certainty, whereas others expected that robust models should not be too sensitive to uncertain parametric and structural assumptions.

Table 2. Findings on improving models and methodologies for robustness and legitimacy (* marks the themes that repeatedly appeared in the survey results with high frequency)

Choice and setup of models	General features	Thematic features	Uncertainty and sensitivity
Purpose • ***Fit for purpose****; one model cannot be robust for all questions **Simple vs. complex** • Comprehensive models with greatest degree of complexity; serving to complicate rather than simplify ***VS.*** • Simple models, but not simplistic • Understood and run easily	**Decision options** • Wide variety of technological and behavioral options • Different types of climate policies **Transformation** • Transformative change and option for a sustainable economy • Non-linearities, tipping points, synergies, dynamic feedbacks, black swans **Granularity** • Sufficient spatial, temporal, and sectoral granularity	**Thematic improvements** • Structural changes in the human system • Demand side in much more detail • Distributional aspects • Political economy • Some aspects of energy, agriculture, and land systems • Biodiversity issues **Climate** • Permafrost thaw, nitrogen cycle and its impacts on vegetation, phosphorus cycle, and implications of carbon store • Effects of climate extremes • Afforestation and its effectiveness, e.g. droughts, fire • Ocean in the climatic component	**Uncertainty** • ***Extended sensitivity and uncertainty analysis**** • Should not eliminate the uncertainties in the future • Meaningful, diverse scenarios without anchoring • Trends, insights, and not numbers or predictions • Identify mitigation policies despite uncertainty **Other** • Stochastic representations, conditional projections • Stability under small fluctuations in model parameters • Not overly sensitive under different types of uncertainties

Table 3. Findings on improving data and reporting for robustness and legitimacy (* marks the themes that repeatedly appeared in the survey results with high frequency)

Data and validation	Openness and reproducibility	Reporting the modeling	Reporting the results
Data • ***Good quality data**** • ***Up-to-date data**** with the latest statistics and sectoral analyses • Good base-year calibration • Realistic assumptions **Validation** • ***Ex-post validation, matching the observed trends and impacts**** • Comparison with other models and studies • Multi-model comparisons; diagnostic work • Verified structural equations, plausible reactions to changes in input data	**Open** • ***Open data, code, documentation, model outputs, methodology**** • FAIR: findable, accessible, interoperable, and reusable • Documentation at two levels: expert user and high-level user **Reproducible** • ***Ability to reproduce**** the outcomes independently from the researcher	**Scope and limitations** • What the ***models can and cannot do**** and why • Communicate routine omissions **Methodology** • ***Transparent methodology and assumptions**** • Functioning of the model and why it gives the results it does • Normative assumptions and inherent judgments • Explicitly address controversies associated with prior assumptions • Model comparisons as tools to highlight the complementarities	**Short messages vs. insights** • Short messages on main drivers and their effects • Simple and yet capable to show the complexity behind the scenes **VS.** • Insights on causal relations, trade-offs, opportunities without adding own value judgement on a certain solution • Communicate ***limitations**** and sensitivities • Detailed ***matching of results with assumptions**** **Uncertainty vs. not too much of it** • Report ***uncertainty**** not to give a false sense of certainty • Uncertainty is not fragility; results are not predictions • Serve to expand rather than narrow the range of policies and approaches discussed **VS.** • ***Not too sensitive to uncertain**** parametric and structural assumptions and initial conditions

Table 4. Findings on improving communication and process for robustness and legitimacy (* marks the themes that repeatedly appeared in the survey results with high frequency)

Communication	Legitimacy	Process
Focus on the users • ***Driven by user demands**** in order to be relevant, e.g. effects on business, other sectors of the economy, and sectoral policies • Adapt communication strategy to the audience; empirical testing • Not only OECD and not only experts • Help interpretation; interactive platforms	**Academic rigour** • ***Peer reviewed models, data, and scenarios**** • Tested by a broader community • Developed and run according to the standards agreed by the community • Assessed by independent scientific committee • Verified by academic scholars and not by non-experts (including users) • Track record of successful applications to climate policy analysis	**One-way communication** • Direct and objective messages • Science first, independent of politics • Issued by impartial institution and supported by an authority of actors **Two-way engagement** • Driven by user demands; connected to policy debates • Ask stakeholders for the most relevant topics and discuss model results, e.g. in workshops

Table continues on the next page

Communication style • Do not overstate the scope and certainty • The concept of uncertainty should be better transmitted • Meaningful reasoning, not alarmistic • Results need to be contextualized to avoid misinterpretation **General** • More outreach and communication	• Open model and a broad and active community joining to send corrections and updates **Some caution** • Over-legitimacy: decision makers may refer to these models acritically • Closed community • Various models as only one source of evidence • Policy recommendations should be made by stakeholders who will balance their objectives with model results	• Broad participation, including governments, industries, academics, and the public • Include experts from all parts of the World • Policy makers having access to the modeling teams to interrogate them **Types of stakeholder inputs** • Agree on assumptions rather than results • Co-design qualitative narratives • User-relevant model outputs

3. Further plans for the stakeholder process

Besides being fully documented and publicly available, the outcomes of this workshop have been passed to the specific NAVIGATE tasks for reflection on how to integrate them in the project's activities. The relevant tasks are: Task 1.1 on critical reflection on IAM scenarios, Task 1.2 on robust insights from modeling, Task 1.4 on exchange and capacity building, Task 6.3 on model documentation and transparency, and Task 6.4 on communication of IAM outputs and insights (including IAM NAVIGATOR). The insights from the group work on climate policy after the pandemic were also passed on to the working groups on the pandemic in Work Packages 2 (structural change), 3 (behavior and lifestyles), and 4 (inequalities). The NAVIGATE Coordination Board will further discuss how the workshop's outcomes could further shape NAVIGATE research.

The main NAVIGATE tools for continuous stakeholder and expert engagement are the three workshops and this report presents the outcomes of the first workshop. The initial list of topics for the workshops was collected during the NAVIGATE kick-off meeting in September 2019 in Potsdam, Germany. From this list, the topic of the second workshop was then discussed in the consortium meeting in May 2020 and announced in June 2020. The **next NAVIGATE workshop** will focus on climate impacts, damages, adaptation, and adaptive capacity and their modeling, in line with the NAVIGATE Work Package 5. Due to this thematic focus, the workshop will primary target experts rather than stakeholders. The second workshop will take place in spring 2021 in Milan, Italy. If the pandemic permits, the workshop will be held on site, as originally planned, and would involve a smaller number of 7-9 experts who would travel to Milan.

The topic of the **third NAVIGATE workshop** to be held in Paris, France, in spring 2022 will be decided during the next consortium meeting in spring 2021. The remaining topics are: inequality and metrics beyond gross domestic product (Work Package 4); climate policies to be assessed and climate policy entry points (Work Package 6); multi-sector scenarios and structural change (Work Package 2); and socio-technical energy transitions in modeling and lumpy vs. granular investments (Work Package 2). If there will be a need for further stakeholder engagement, a possibility to organize an additional smaller workshop online is at the moment kept open, even if this workshop has not been foreseen in the initial NAVIGATE work plan.

In order to coordinate the stakeholder activities in NAVIGATE as well as in other related

consortiums, joint discussions and planning was initiated together with the Horizon 2020 projects ENGAGE and COMMIT, as well as Joint Programming Initiative project CHIPS. As one concrete outcome, NAVIGATE consortium will contribute to co-organizing three additional workshops on climate policy after the pandemic with the focus on Europe, Asia, and international dimension in autumn 2020, together with ENGAGE and COMMIT. Other ways for NAVIGATE members to participate in the workshops of the other projects, and vice versa, have been foreseen. If the topic of the third (main) NAVIGATE workshop in Paris would be ultimately decided to focus on inequality, the workshop could be potentially co-organized with the CHIPS project that has an explicit focus on inequality.

Initially, the NAVIGATE project foresaw to create a stakeholder and expert database for the purpose of ensuring that the consortium members can reach to a wider set of stakeholders on demand via an online survey or bilateral discussions. The proposed plan was to merge existing stakeholder databases from the Horizon 2020 project DEEDS and Joint Programming Initiative ERA4CS project SENSES, but the changes in the European Union's General Data Protection Regulation have proved the transfer of the databases between projects challenging. After the discussion in the NAVIGATE kick-off meeting in September 2019 in Potsdam, it was jointly decided not to pursue this database. As demonstrated in Section 2.8, other successful ways were found to reach a wider range of stakeholders and experts in an online survey.