Estimating the likelihood of GHG concentration scenarios from probabilistic IAM simulations

David Huard[1], Jeremy Fyke[2], Iñigo Capellán-Pérez[3], H. Damon Matthews[4], and Antti-Ilari Partanen[5]

[1]Ouranos, Montréal, QC, Canada
[2]Canadian Centre for Climate Services, ECCC, Canada
[3]Research Group on Energy, Economy and System Dynamics, Escuela de Ingenierías Industriales, University of Valladolid, Valladolid, Spain
[4]Concordia University, Montréal, Qc, Canada
[5]Climate System Research, Finnish Meteorological Institute, Helsinki, Finland

Key Points:

- Probabilistic IAM simulations of future emissions are used to evaluate the relative probability of carbon dioxide concentration scenarios.
- Coordinated IAM and climate modeling experiments should be designed to enable probabilistic climate change risk assessments.

Corresponding author: David Huard, `huard.david@ouranos.ca`

Abstract

Climate change adaptation under resource constraints and future climate uncertainties would benefit from fully probabilistic climate risks assessments. Conducting such risk analyses requires assigning probabilities to the future greenhouse gases (GHG) and land-use scenarios used by global climate models. This paper proposes an approach to estimate the relative likelihood of carbon dioxide (CO_2) concentration scenarios used in key climate change modeling experiments. The approach relies on the comparison of CO_2 emissions from probabilistic simulations of Integrated Assessment Models (IAM) with compatible CO_2 emissions diagnosed by global climate models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5) and 6 (CMIP6). The approach is demonstrated with five emission simulations from four IAMs, leading to independent estimates of the relative likelihood of CMIP5 Representation Concentration Pathways and CMIP6' Shared Socioeconomic Pathways (SSP) up to 2100. Results suggest that SSP5-8.5 is an unlikely scenario for the second half of the century, but there is no clear consensus on the most likely scenario. Scenario likelihood is affected by a number of potential errors, including sampling errors, differences in emission sources simulated by the IAMs, and the lack of a common experimental framework for IAM simulations. These errors, along with the small IAM ensemble size, limit the applicability of the results. The delivery of fully probabilistic climate risk assessments would benefit from a coordinated probabilistic IAM experiment jointly designed with a coordinated climate modeling experiment where Earth System Model are driven by representative emission pathways.

Plain Language Summary

Climate model simulations are being increasingly used to understand future trends in the severity and frequency of impactful climate hazards and associated physical and socioeconomic risks. To run climate model simulations as part of large scale, coordinated climate projection exercises, climate models are provided with scenarios of greenhouse gases and land-use change over coming decades. However, the scenarios most widely used today for adaptation planning have no assigned probabilities; rather, they are explicitly intended to span a range of arbitrary climate futures. This reduces the applicability of resulting hazards projections for cost/benefit analysis of adaptation investments. To support risk-based adaptation decision making, this study combines five sets of probabilistic carbon dioxide emissions simulated by four Integrated Assessment Models (IAM) with

CMIP5 and CMIP6 climate model ensemble results to estimate the probability of future GHG concentration and associated climate scenarios. The results are IAM-dependent, though the majority of individual IAM-based analyses suggest that the high-emissions scenario SSP5-8.5 becomes unlikely as we reach the second half of the century. Based on lessons learned in this exercise, we propose that new sets of IAM and climate model experiments be appropriately designed at their initial stages to better support probabilistic climate change risk assessments.

1 Introduction

Future climate change impacts are captured in multi-model climate experiments designed and coordinated through the Coupled Model Intercomparison Project (CMIP). These climate modeling experiments explore, among other topics, the climate consequences of rising greenhouse gases (GHG) concentrations in the atmosphere (Taylor et al., 2012). Many decision-makers are now using these climate projections to assess hazards and make consequential planning and investment decisions. In many instances however, risk assessments are carried out without the benefit of a probabilistic framework quantifying the leading sources of uncertainties affecting projections: climate and carbon cycle sensitivity, natural variability, and future GHG emission and concentration scenarios.

While the climate community has been diligent in assessing and quantifying climate modelling uncertainties and natural variability (Lehner et al., 2020), and integrated assessment studies routinely evaluate socio-economic uncertainties (Pastor et al., 2020; Capellán-Pérez, 2016), there is very little guidance available regarding the relative probabilities of GHG scenarios underpinning the climate change simulations typically used to assess the impacts of climate change. In CMIP3, these transient climate change experiments were driven by a family of GHG scenarios called SRES. Despite significant differences across GHG SRES scenarios, the climate community has, by and large, avoided commenting on their respective likelihood; a common stance has been to consider all emission scenarios as "equally valid with no assigned probabilities of occurrence" (Nakicenovic et al., 2000). Murphy et al. (2009) explains that SRES scenarios have, by design, no assigned probability.

This reluctance to assign probabilities to GHG scenarios has carried into the following generations of GHG scenarios. In CMIP5, GHG concentration scenarios are defined by Representative Concentration Pathways (RCPs), and "no likelihood or preference is attached to any of the individual scenarios in the set" (van Vuuren et al., 2011). The same non-commitment holds for CMIP6' Shared Socioeconomic Pathways (SSPs) (Riahi et al., 2017).

From the point of view of users of climate projections, this lack of guidance on the probability of emission scenarios is however a serious impediment to judgment formation, risk analysis and ultimately, effective decision-making (Schneider, 2001; King et al., 2015; Hieronymus, 2020). According to Morgan and Keith (2008): "If judgments about likelihood are not supplied with the scenarios, they will be assumed by the users either explicitly or implicitly. The convention of not communicating information about the relative likelihood of scenarios therefore muddies communication between analysts and users." A concrete example of this are stakeholders' frequent requests for climate impacts based on "business-as-usual" scenarios. IAM modelers may have no preference for one scenario over the other, but most people will naturally assume that the continuation of historical trends is more likely than a change in the world's socioeconomic dynamics. This insistence on scenario-agnostism leaves decision-makers, with no special expertise in GHG scenarios, effectively responsible for assigning implicit or explicit likelihoods to future scenarios in order to craft high-cost, high-consequences adaptation plans (Ho et al., 2019).

To be fair, the probability of GHG emission scenarios is not a question climate modelers are well qualified to answer. The evolution of anthropogenic GHG emissions is influenced by policy, demography, economy, geopolitics and technology, topics well outside the expertise of climate science. The description of these factors and their interactions are captured by another class of model called Integrated Assessment Models (IAMs) (Sokolov et al., 2005; Moss et al., 2010; Agrawala et al., 2011; Koomey et al., 2019)). The IAM community generates hundreds of different scenarios, predicated on assumptions regarding future climate policies, technological advances, demography and energy markets. It is from these IAM simulations that the climate science community has drawn the GHG and land-use scenarios underlying RCPs and SSPs. The selection of model inputs was not meant to capture the most plausible scenarios, but rather to generate a sample of *representative* pathways exploring the "full range of emission scenarios available in the current scientific literature, with and without climate policy" (van Vuuren et al., 2011).

The IAMs used in CMIP experiments describe different *storylines*, but there are other avenues to assess emission scenario uncertainties. Indeed, van Vuuren et al. (2008) distinguish between storyline-based alernative scenarios and fully probabilistic scenarios. Storylines embody fundamentally different, yet internally consistent, representations of the future that can be represented by an IAM. Fully probabilistic scenarios are created by assigning probability distributions to key IAM input parameters, and sampling from those distributions to create a set of probabilistic emission pathways. The "conditional probability approach" combines both storylines and probabilistic scenarios, arguing that it is easier to define probability distributions for IAM parameters in the context of a particular storyline.

Uncertainty analysis has been identified as one of the current key weaknesses of IAMs (Pastor et al., 2020; Rogelj et al., 2017). For example, one contentious topic relates to the parameterization of climate damages in cost-benefit IAMs. The family of cost-benefit IAMs traditionally relies on median damages, overlooking the low and high tails of the distribution for the climate sensitivity. A lower or higher climate sensitivity implies smaller or larger climate hazards, and costs, for the same CO_2 concentration. If the climate sensitivity distribution has "fat tails", using the median estimate could bias cost assessments (Ackerman et al., 2010; Keen, 2020; Stern, 2013; Weitzman, 2012). A related issue is the possibility of tipping points in the climate system and their impact on damage functions (Lontzek et al., 2015; Cai et al., 2016).ă

The need for quantitative probabilistic assessments of uncertainties was expressed back in 2000 in a guidance document to Intergovernmental Panel on Climate Change (IPCC) authors by Moss and Schneider (2000): "We believe it is more rational for scientists debating the specifics of a topic in which they are acknowledged experts to provide their best estimates of probability distributions and possible outliers based on their assessment of the literature than to have users less expert in such topics make their own determinations." This comment was followed by the expectation that Bayesian approaches would be most appropriate to describe inherently subjective degrees of belief in our assessment of the state of science.

This view on the need for a Bayesian interpretation of uncertainties is often cited in later papers looking into probabilistic emission scenarios. For example, M. D. Webster et al. (2002), M. Webster et al. (2003), M. Webster et al. (2008) and M. Webster et al. (2012) sampled an *a priori* parameter distribution of the Emissions Predictions and Policy Analysis (EPPA) model to generate probabilistic GHG emission trajectories. These emissions were then fed into the MIT Integrated Global System Model (IGSM) to compute the posterior distribution for resulting temperature changes.

Schneider and Mastrandrea (2005), Sokolov et al. (2009) and Repetto and Easton (2015) similarly assigned probability distribution to parameters of the DICE model to assess the probability of *dangerous* anthropogenic interference with the climate system and assess policy options. The authors stated: We do not recommend that our quantitative results be taken literally, but we suggest that our probabilistic framework and methods be taken seriously: they produce relative trends and general conclusions that better represent a risk-management approach than estimates made without probabilistic representation of outcomes."

Ward et al. (2012) define a simplistic supply-side model of fossil fuel production to generate resource-constrained CO_2 emissions. Fossil fuel production is broken down at the national level by fuel types. The model makes the hypothesis that known reserves of conventional and unconventional fuel can and will be extracted. Production growth rate is considered an uncertain parameter and sampled from a distribution, with an initial growth rate of 10% for all fuels not currenly in production. Model assumptions are deliberately biased toward high production growth, with the intent of outlining the upper structural limit to CO_2 emissions.

Gillingham et al. (2018) ran multiple IAMs to assess the relative contribution of model structure and parametric uncertainty to future temperature, CO_2 concentration and economic output. Model parameters for population, productivity and climate sensitivity are sampled from *a priori* probability distribution drawn from the literature. This multi-model approach allowed the authors to assess leading sources of uncertainty and provide probability distributions for output variables.

This paper builds on these ideas and leverages outputs from published probabilistic emission simulations to estimate the conditional probability for the CO_2 *concentration* scenarios within RCPs and SSPs. The paper targets the most popular CMIP transient climate change experiments, in which scenarios impose time-varying GHG concentrations to global climate models (GCM). Note that newer generation of Earth System Models (ESM) can simulate carbon cycle processes, allowing them to be driven directly with GHG *emissions*; such emission-driven experiments would lend themselves to a much simpler probabilistic analysis than what is described in this paper.

2 Data and Method

Contrary to a common misconception, coordinated climate change experiments typically used in climate impacts studies are not driven by GHG *emission* scenarios, but use *prescribed* GHG *concentration* scenarios. This is done in part to give climate models that cannot simulate the carbon cycle an opportunity to participate. For recent CMIP iterations, emission scenarios drawn from select IAM simulations are converted into concentrations using a reduced complexity model called MAGICC (MAGICC6 for CMIP5 (Meinshausen, Raper, & Wigley, 2011), and MAGICC7 for CMIP6 (Meinshausen et al., 2019)). These GHG concentrations are part of the boundary conditions prescribed to climate models, along with land-use scenarios, aerosol concentrations, etc. Note that there are other CMIP experiments where ESMs are driven by emission scenarios, but they usually count fewer participating models and are currently rarely used in impact assessments. This will likely change as more models include carbon cycle processes allowing them to be driven directly by emissions. Figure 1 illustrates how CMIP5 RCP experiments are tied to IAM emission scenarios.

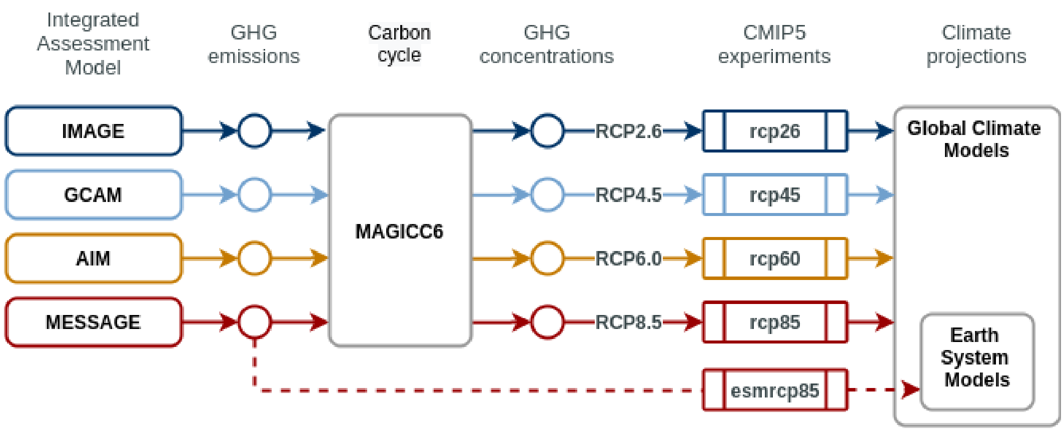

Figure 1. In CMIP5 RCP experiments, global climate models are prescribed greenhouse gas concentrations estimated by MAGICC6 from emission scenarios simulated by four different IAMs: IMAGE, GCAM, AIM and MESSAGE. In contrast, in the `esmrcp85` experiment, ESMs are prescribed GHG emissions directly and use their own carbon cycle processes to compute concentrations. Although details differ, the experimental setup for CMIP6 is conceptually similar.

An RCP or SSP scenario is a short-hand to describe an elaborate experimental design, of which the GHG concentration trajectories is just one component (Eyring et al., 2015). Climate modeling teams set up their model following this experimental design, run one or many simulations (realizations), and then archive model outputs according to precise data and metadata specifications meant to facilitate model intercomparisons. Model outputs include hundreds of different variables, from which many climate hazards can be derived: heatwaves, sea level rise, annual maximum precipitation, droughts, etc. To quantify uncertainties, climate impact studies typically include results from multiple realizations from multiple models driven by multiple GHG concentration scenarios. With these results in hand, a legitimate question by decision-makers could be for example: "considering known uncertainties, what is the probability of precipitation exceeding a given threshold over the period 2030–2050?"

2.1 Probabilistic Framework to Assess Climate Hazards

Let's denote a climate hazard as H. Risk analysts and decision makers are interested in $\mathrm{P}(H(t))$, the probability of occurrence of hazard H at some time t in the future. To lighten the presentation, time dependence t is implicitly assumed in the next equations. Note also that we're using the term *probability* to denote a subjective degree of belief in an hypothesis, not a frequency of occurrence.

What CMIP experiments can provide is a probability *conditional* on the experimental design or scenario. If we let S stand for a CMIP scenario, including GHG and aerosols concentrations, initial conditions, etc., then climate model simulation ensembles can be used to compute $\mathrm{P}(H \mid S)$, the probability of a climate hazard H conditional to the scenario S.

Although in principle CMIP offers multiple scenarios to draw from, the climate impact community has mostly relied on concentration-driven scenario experiments. For CMIP5, these include `rcp26`, `rcp45`, `rcp60`, `rcp85`, which we'll denote as $\mathbf{S}_{\mathrm{CMIP5}}$, the set of concentration-driven CMIP5 RCP experiments. For CMIP6, there are nine such ScenarioMIP experiments, but participating models are minimally expected to contribute simulations to Tier 1 experiments: $\mathbf{S}_{\mathrm{CMIP6}} = \{\mathtt{ssp126}, \mathtt{ssp245}, \mathtt{ssp370}, \mathtt{ssp585}\}$ (O'Neill et al., 2016). If we assume that the future climate will be captured by those sets of scenarios, then the hazard probability can be estimated by a weighted average of conditional hazard probabilities:

$$\mathrm{P}(H) \approx \sum_{S \in \mathbf{S}} \mathrm{P}(H \mid S)\mathrm{P}(S), \tag{1}$$

where the weights $\mathrm{P}(S)$ are the Bayesian *prior* for each scenario.

The first term on the right hand side of Eq. (1) is what climate impact studies routinely compute. Its computation can be as simple as a fit of a normal distribution to the hazards simulated by a multi-model ensemble, or can include considerations regarding model performance or model independence (Knutti et al., 2017). Our focus in this paper is with the second term, the probability of a given RCP or SPP scenario, or how likely are the future conditions described in each scenario S. A full answer to this question would require an evaluation of the joint probability of all scenario components: concentration trajectory for each individual GHG, land-use changes, aerosols, solar forcing, etc. To reduce the scope of the problem, we evaluate the scenario likelihood based only on the global CO_2 trajectory and ignore the influence of other scenario components.

This paper makes the claim that the scenario probability can be estimated conditionally to a set of probabilistic IAM emissions: $\mathrm{P}(S) \equiv \mathrm{P}(S \mid \mathbf{E}_{\mathrm{IAM}})$. Using Bayes' theorem, and assuming RCPs and SSPs can be represented solely by their global CO_2 concentrations C_S, we get

$$\begin{aligned}
\mathrm{P}(S \mid \mathbf{E}_{\mathrm{IAM}}) &= \frac{\mathrm{P}(\mathbf{E}_{\mathrm{IAM}} \mid S)\mathrm{P}(S)}{\mathrm{P}(\mathbf{E}_{\mathrm{IAM}})} \\
&\propto \mathrm{P}(\mathbf{E}_{\mathrm{IAM}} \mid C_S)\mathrm{P}(S). \qquad (2)
\end{aligned}$$

The second term on the right-hand side of Equation (2) is the prior for the scenario. It can be set subjectively by the risk analyst, or calculated based on other evidence. The next section discusses how to compute the first term, the likelihood of probabilistic emissions given the scenario CO_2 concentration.

2.2 Likelihood of Emission Trajectories

In Eq. (2), the likelihood term $\mathrm{P}(\mathbf{E}_{\mathrm{IAM}} \mid C_S)$ stands for the probability of CO_2 emission trajectories given a known CO_2 concentration scenario. Comparing concentrations and emissions implies a mechanism to convert one into the other. RCP and SSP scenarios use the MAGICC model to convert emissions into concentrations, but doing so here would dismiss carbon cycle uncertainties that are already partially neglected by concentration-driven experiments (van Vuuren et al., 2011).

The approach chosen here is to convert concentrations into *compatible emissions*. Compatible emissions are the anthropogenic fossil fuel emissions that balance carbon fluxes between the air, land and ocean when ESM are prescribed a given CO_2 pathway. Indeed, climate models simulate CO_2 exchanges between the atmosphere, land and ocean. If c_a, c_o and c_l stand respectively for the carbon stored in the atmosphere, the ocean and the land surface, then carbon emissions from fossil fuels e modify the total carbon budget:

$$e = \frac{\partial c_a}{\partial t} + \frac{\partial c_o}{\partial t} + \frac{\partial c_l}{\partial t}. \tag{3}$$

The first term on the right is the time derivative of the CO_2 concentration pathway used in the experimental design, and the other terms on the right are the carbon fluxes from the atmosphere into the ocean and land. These fluxes are simulated and archived by climate models, allowing us to compute e, the *compatible emissions* that would have led to the concentrations imposed by the experimental design. Differences in the carbon cycle representation of ESMs such as vegetation dynamics, fire-carbon interactions or ocean biochemistry, appear as inter-model spread in compatible emissions. Compatible emissions are computed for CMIP5 in Jones et al. (2013), and for CMIP6 in Liddicoat et al. (2021). In the latter, mean fluxes during the preindustrial period are removed from fluxes over the historical and future scenarios to correct for the fact that some ESMs have not reached equilibrium.

Carbon dioxide emissions compatible with prescribed concentrations are estimated for the CMIP5 and CMIP6 Tier 1 experiments following the methodology from Liddicoat et al. (2021). All model simulations for which both `fgco2`, the gas exchange carbon flux into the ocean, and `nbp`, the carbon flux from the atmosphere into the land, were available on the Earth System Grid Federation and free from defects are used (the number of simulations available per model are listed for CMIP5 and CMIP6 in Tables A1 and A2 respectively). Global fluxes are computed by multiplying ocean and land fluxes by the respective fractional ocean area (`sftof`) and land area (`sftlf`) as well as the respective grid cell area (`areacello`, `areacella`), and summing over the entire globe. Annual fluxes are entered into Equation (3) along with numerically differentiated RCP (Meinshausen, Smith, et al., 2011) and SSP concentrations (Riahi et al., 2017; Gidden et al., 2019), yielding annual compatible CO_2 emission time series shown in Figure 2. In cases where preindustrial simulations (`piControl`) are available, the mean compatible emissions over the last 30 years of the preindustrial period (see Tables B1 and B2) are subtracted from historical and future emissions.

For each RCP and SSP experiment, we thus have an ensemble of emissions that are compatible with the CO_2 concentration scenario. If we assume that the probability density of those compatible emissions can be described by a normal distribution, taking the mean $\mu_S(t)$ and variance $\sigma_S^2(t)$ of diagnosed emissions for each scenario S lets us define a time dependent functional form for the IAM emission likelihood:

$$\mathrm{P}(\mathbf{E}_{\mathrm{IAM}} \mid C_S) \equiv \mathcal{N}(\mathbf{E}_{\mathrm{IAM}} \mid \mu_S, \sigma_S). \quad (4)$$

To account for the varying number of realizations per model, ensemble statistics are first evaluated across realizations for each model, then across models within each experiment. More explicitly, if $e_{S,m,r}$ stands for compatible emission from experiment S, model m and realization r, and if $E_d[X]$ and $V_d[X]$ stand for the average and variance over dimension d respectively, then

$$\begin{aligned} \mu_S &= E_m[E_r[e_{S,m,r}]], \\ \sigma_S &= \sqrt{V_m[E_r[e_{S,m,r}]] + E_m[V_r[e_{S,m,r}]]}. \end{aligned}$$

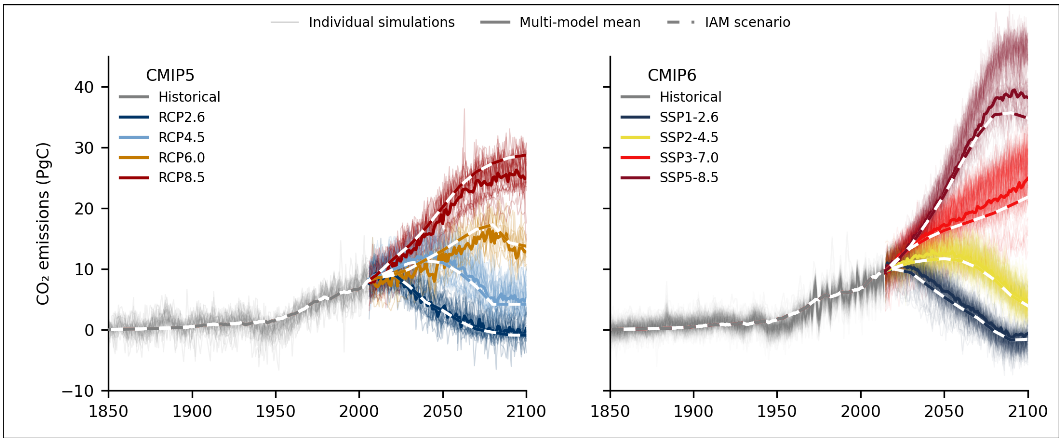

Figure 2. Compatible fossil CO_2 emissions diagnosed from CMIP5 (left) and CMIP6 (right) simulations from historical (gray) and future (color) scenarios. Thin lines denote individual model simulations, thick solid lines the multi-model mean for each experiment, and thick dashed lines the IAM scenario fossil fuel emissions. Note that the mean is not always centered relative to individual simulations, because the mean is first calculated over ensemble members, then over individual models. This is especially visible in the right panel for SSP5-8.5, where a cluster of 50 CanESM5 simulations reaches much higher compatible emission values than other ESMs.

Table 1 shows the total number of simulations and models available to compute those statistics for CMIP5 and CMIP6. Note that for RCP6.0, only seven different models were available.

Table 1. Number of simulations and models with at least one simulation available for each CMIP5 and CMIP6 experiment. For a break-down per model, see Tables A1 and A2.

Model name	Historical	RCP2.6	RCP4.5	RCP6.0	RCP8.5
Simulations	35	23	27	10	27
Models	13	11	13	7	13

Model name	Historical	SSP1-2.6	SSP2-4.5	SSP3-7.0	SSP5-8.5
Simulations	233	121	155	131	112
Models	22	17	16	18	16

2.3 Probabilistic IAM Emission Simulations

Probabilistic emission simulations are taken from five different papers, two of those using the same IAM. These papers were selected opportunistically based on two main criteria: ensembles of probabilistic CO_2 emission time series up to 2100 were available publicly or from the authors, and the simulations did not explicitly constrain emissions to meet policy ambitions. Note that these IAMs were not necessarily intended to yield *predictive* emissions, hence results derived from those simulations should be not be interpreted too literally. They are used here mainly to illustrate the potential of probabilistic IAM simulations to inform climate risks. To lighten the text, each paper is identified by an abbreviation.

Fyke and Matthews (2015) [FM15] have developed a reduced-form numerical carbon emission model based on differential equations describing the exchange of carbon between geological and exogenous (atmosphere, ocean and biosphere) reservoirs. The fluxes between those reservoirs depend on extraction and consumption rates, which in turn depend on availability and prices for fossil fuel and its alternatives. The model counts 28 parameters, 17 of which have significant uncertainty. For each of these uncertain parameter, a probability distribution describing its uncertainty was defined based on published estimates and expert judgment. The parameters were sampled (n=100,000) from their prior distributions using a latin hypercube sampler, and time series (2012–2100) of carbon emissions generated from the model equations (Figure 3). The parameters whose uncertainty had the greater impact on emissions are the minimum future non-fossil energy cost, the maximum size of potential fossil energy resources and the maximum potential carbon pricing.

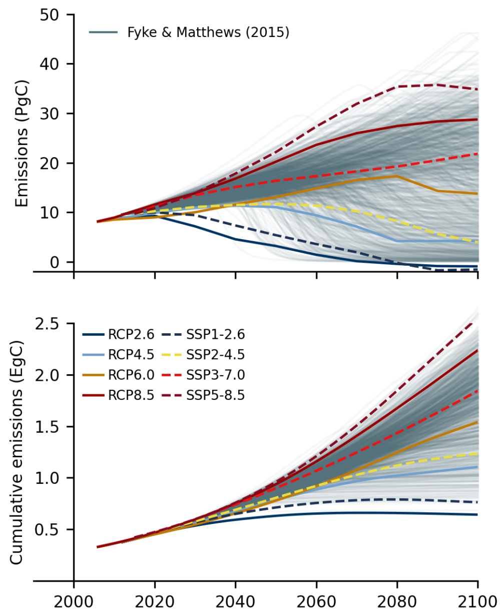

Figure 3. Stochastic emissions and cumulative emissions from Fyke and Matthews (2015) (gray), overlaid with IAM emissions from CMIP5 RCPs (solid lines) and CMIP6 SSPs (dashed lines).

Capellán-Pérez et al. (2016) [CP16] leverage the GCAM IAM, combined with a probabilistic assessment of recoverable energy resources, supply-cost curves and climate sensitivity, to analyze the relative importance of these factors in the temperature response at the end of the century. The study focuses on energy availability considerations, using the "remaining ultimately recoverable resources" (RURR) approach to estimate non-renewable energy sources. It uses a Monte Carlo approach, where uncertain parameters are sampled (n=1,000) from their respective *prior* distribution, and then fed into GCAM-MAGICC to obtain CO_2 emissions (Figure 4), total radiative forcing and the global temperature response over the period 2005–2100. GCAM-MAGICC is run in baseline mode, meaning that no climate policy are imposed. The results reveal that the coal RURR uncertainty is the determinant factor among fossil fuel resources considered. It also shows that not accounting for the values in the upper ranges of fossil fuel availability can lead to an underestimate of total radiative forcing by the end of the century.

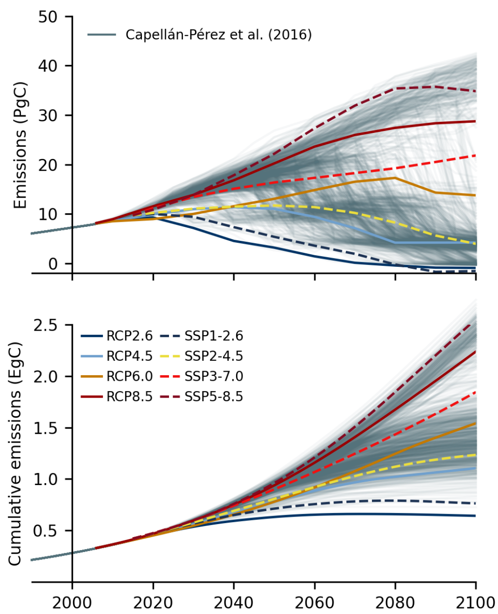

Figure 4. Stochastic emissions and cumulative emissions from Capellán-Pérez et al. (2016) (gray), overlaid with IAM emissions from CMIP5 RCPs (solid lines) and CMIP6 SSPs (dashed lines). Note that the abrupt jumps in the emission time series are likely due the lack of geological constraints to fossil fuel extraction rates. Non-renewables availability is modeled using supply-cost curves, and when a resource is depleted, its extraction drops to zero the following year.

Raftery et al. (2017) [R17] propose a model based on the country-level Kaya identity, where the future carbon emissions of a country are given by the product of population, gross domestic product (GDP) per capita and carbon intensity (carbon emitted by GDP). Probabilistic population projections up to 2100 are taken from United Nations (2015), reflecting data up to 2015. A joint Bayesian hierarchical model for the GDP and carbon intensity is calibrated on data from 1960–2010. The model assumes an evolving world technology frontier, to which countries' GDP converge at country-specific rates, and that all countries have reached a carbon intensity peak and are now on a declining trend. Model parameters are sampled by Monte Carlo (n=100,000) and the distribution of emissions (Figure 5) analyzed to assess the relative importance of population (2%), GDP (50%) and carbon intensity (48%). The model does not include explicit future climate policies, but the fact that model parameters are calibrated on historical data ensures that the influence of past and current policies are included in future projections.

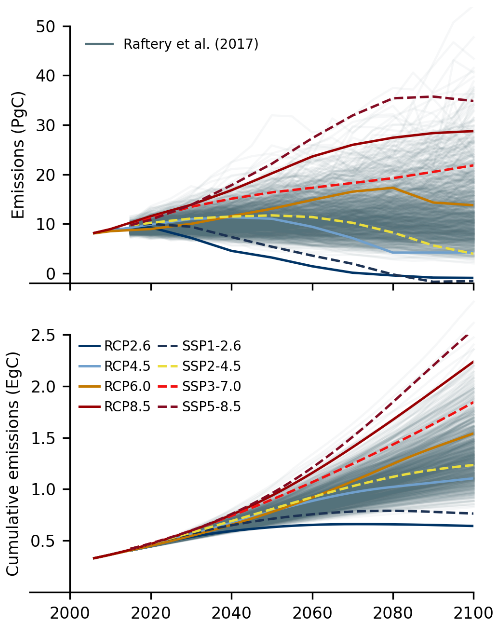

Figure 5. Stochastic emissions and cumulative emissions from Raftery et al. (2017) (gray), overlaid with IAM emissions from CMIP5 RCPs (solid lines) and CMIP6 SSPs (dashed lines).

Liu and Raftery (2021) [LR21] uses the same model as Raftery et al. (2017), but with five additional years of population (United Nations, 2019), economic and emission data. Slower growth in emissions between 2010–2015, compared to 1960–2010 period, explains the lower global annual emissions (Figure 6).

Capellán-Pérez et al. (2020) [CP20] describe the MEDEAS-W IAM developed in the course of the homonymous EU project (Modeling Energy System Development under Environmental and Socioeconomic constraints). The model counts nine modules: economy, energy demand, -availability, -infrastructures and -return on energy invested, minerals, land-use, water, climate/emissions, and social and environmental impact indicators. The model is meant to inform decision-making regarding the transition to sustainable energy systems, and grants considerable attention to biophysical constraints to growth

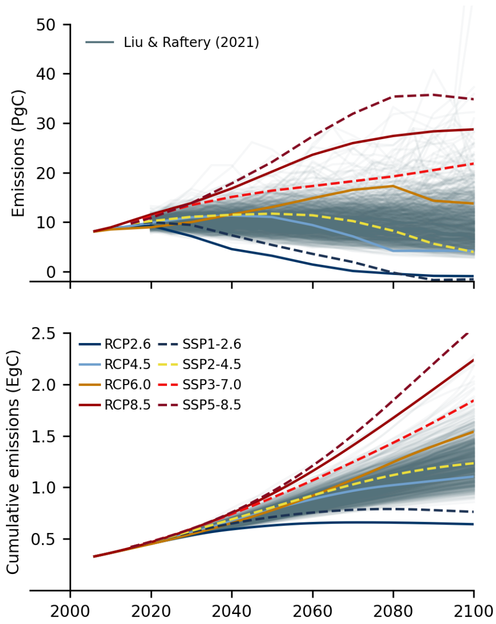

Figure 6. Stochastic emissions and cumulative emissions from Liu and Raftery (2021) (gray), overlaid with IAM emissions from CMIP5 RCPs (solid lines) and CMIP6 SSPs (dashed lines).

and energy availability, mineral and energy investments for energy shifts, sectoral economic structure, and climate change damages. The model accounts for the characteristics of 25 energy sources and technologies, including reserves, extraction rate, intermittency of some renewable energy sources and requirements for storage and overcapacity. The model assumes technological improvements at economic sectoral and technologies level, but also constraints to the potential for renewables due to for example lower quality siting and thermodynamical limits to generation efficiency. A non-linear climate damage function affects production and growth rate, consistent with the interpretation of "dangerous climate change" beyond CO_2 concentration thresholds. Monte-Carlo simulations (n=1,000) are run to perform an uncertainty analysis for 72 inputs. CO_2 emissions (1995–2100) simulated for the reference (Ref) scenario (a conditional probability approach based on the business-as-usual storyline) are shown in Figure 7.

Comparing and contrasting the above studies highlights several themes. Firstly, it is notable that each of the studies make contrasted model choices. For example, R17 and LR21 construct a data-driven statistical model framework that applies the simple Kaya identity at multiple points of time to develop emission time series. In contrast, FM15, CP16 and CP20 use more sophisticated time stepping numerical models with inherent

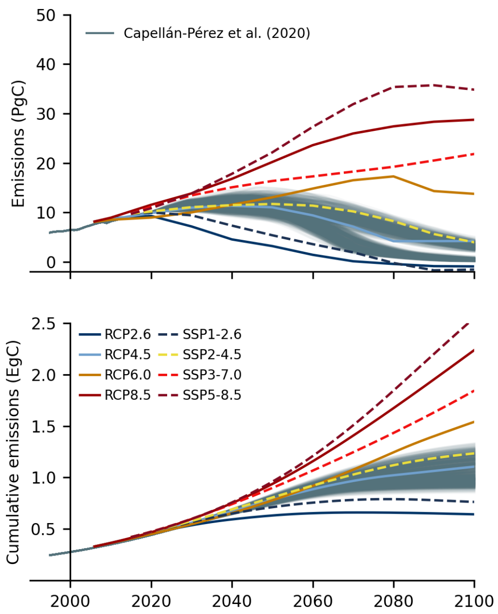

Figure 7. Stochastic emissions and cumulative emissions from Capellán-Pérez et al. (2020) (gray), overlaid with IAM emissions from CMIP5 RCPs (solid lines) and CMIP6 SSPs (dashed lines).

system dynamics, albeit with very different numerical methods: FM15 treat the global fossil fuel resource volume as an endogenous model component, whereas CP16 treat the same as an exogenous input variable and focus numerical efforts on estimating prices consistent with cleared markets, a concept which is entirely absent in the FM15 approach. In CP20, the total amount of reserves is fixed, but the extraction rate is subject to physical constraints (maximum extraction curves).

Secondly, the studies described here apply notably different methods to sample exogenous inputs to the respective model frameworks. R17 arguably dedicate the most effort towards this aspect, in applying a hierarchical (multi-level) Bayesian model framework approach for estimating modeled input parameters. Indeed, the term "model" in their study primarily applies to the statistical models which derive the inputs for the Kaya Identity relationship, rather than the functional equation itself. In contrast, FM15 use relatively simpler sampling of a larger number of uncertain, scalar, model parameters based on a normal distribution-weighted latin hypercube sampling approach that does not discriminate potential interdependencies between parameters. Providing further contrast, CP16 sample input parameters from empirical distribution functions for key input parameters, with distribution functions for important input parameters developed from available published literature values. In CP20, 72 uncertain parameters are sampled from uniform distributions, with ranges that go from ś20% around the reference scenario for most parameters, to ś50% or even [-50%, +100%] for the most uncertain parameters. As with model design diversity, there is significant long-term potential for convergent evolution towards optimal input parameter sampling practices, but also short-term challenges in comparing inter-study results because of differences in both sampled parameters, and statistical sampling methods.

A critical difference across studies is the definition of CO_2 emissions and the processes that contribute to them. FM15 simulate emissions from fossil fuel combustion only, while R17 and LR21 also include cement production. CP16 emissions account for industrial processes, fossil fuel combustion and land-use change, and in CP20 losses from energy transformation and distribution are included as well.

Despite model methodological differences, there are also several fundamental similarities between the studies. Per selection criteria, all papers adopted methodologies for model design and simulation production that avoided any predefinition of final results to which the model simulations are forced to meet (so-called "perfect foresight" or "policy optimization" modeling). This contrasts fundamentally with the RCP and SSP scenario families which are constrained by design to match radiative forcing levels in 2100. It also contrasts with scenarios exploring pathways consistent with avoidance of exceedance of particular temperature thresholds (*e.g.* 2řC above preindustrial temperature). Avoiding the "perfect foresight" approach is a necessary precondition for any fully-scoped probabilistic assessment of future emissions, because it allows probabilistic assessments to develop in a free-running manner without a priori constraints on final emission levels, radiative forcing anomalies, or temperature targets.

A second similarity between studies is their focus on evaluating the likelihood of exceeding global temperature thresholds, and its sensitivity to various factors. FM15 describe the long-term likelihood of exceeding various global temperature change levels. Beyond the climate response to cumulative emissions, these likelihoods are most sensitive to minimum non-fossil energy prices, maximum fossil energy resources and maximum carbon price. R17 focus on the likelihood of exceeding the 2řC threshold by 2100, with results that are largely influenced by the gross domestic product per capita and carbon intensity, and much less by population growth uncertainty. LR21 conduct a country-level analysis, estimating the likelihood that countries will meet their nationally determined contributions, and keep warming below 2řC. CP16 describe the likelihood of surpassing RCP emissions levels by 2100 and crossing the 2řC warming threshold in 2100. A sensitivity analysis covering resource related uncertainties indicate the dominant source of uncertainty are the remaining ultimately recoverable coal resources. CP20 is primarily devoted to the description of the MEDEAS modeling framework, and uses its Monte Carlo simulations to assess the robustness of the results, rather than estimate the probability of outcomes.

2.4 Annual vs Cumulative Emissions

Key climate change features, such as global air temperature change, surface ocean temperature rise and sea level rise are nearly linearly related to cumulative emissions (Matthews et al., 2009; Williams et al., 2012). Given our objective is to assess the probability of such climate impacts, it makes sense then to compute Eq. (4) on cumulative emissions rather than annual emissions. This however complicates the analysis, because results now hinge on the year from which we start cumulating emissions.

Historical CMIP5 and CMIP6 experiments start in 1850, while future scenarios start in 2006 for CMIP5 and 2015 for CMIP6. On the other hand, the five probabilistic IAM emissions simulations presented above start in 1990, 1995, 2010, 2012 and 2015. To align all results to a common starting point, all emissions are accumulated from 1750 onward, using observations taken from the Global Carbon Budget project (Friedlingstein et al., 2020) to fill gaps. For example, probabilistic cumulative emissions from FM15, starting in 2012, are incremented by the 2011 observed cumulative emission. Diagnosed compatible cumulative emissions for `historical` simulations are similarly incremented by observations from the year prior to the start of the experiment (1849 for CMIP6, and 1861 for CMIP5 to account for missing data in some simulations). Future simulations are matched with historical simulations from the same model, and whenever possible, the same realization.

2.5 Data Access and Analysis Software

CMIP5 and CMIP6 CO_2 land and ocean fluxes, grid cell areas and land-sea fractions were downloaded from ESGF using Synda (Nasser et al., 2020). Probabilistic emission simulations were obtained from authors (FM15, CP16, CP20), or reproduced by running publicly available code (R17, LR21, see Appendix B). Computations were carried out in the Python programming language using xarray (Hoyer & Hamman, 2017), pandas (McKinney, 2010), NumPy (Harris et al., 2020) and SciPy (Virtanen et al., 2020), and graphics created with Matplotlib (Hunter, 2007). Analysis-ready data and code to reproduce results from this paper are available at the Federated Research Data Repository (`https://doi.org/10.20383/102.0549`).

3 Results

The following sections discuss different types of emissions, and to avoid confusion, we use the following terminology. *Scenario emissions* refer to CO_2 emissions pathways defined in RCP and SSP scenarios. *Compatible emissions* are inferred from carbon fluxes simulated by CMIP models using Eq. (3). Finally, *probabilistic emissions* denote ensembles of CO_2 emission trajectories simulated by the probabilistic IAMs described in section 2.3.

The mean and standard deviation of compatible cumulative emissions for RCPs and SSPs are shown in Figure 8. For CMIP5 experiments RCP6.0 and 8.5, compatible cumulative emissions are considerably smaller than scenario emissions. In other words, less carbon emissions are needed in CMIP models than in MAGICC6 to reach the same CO_2 concentration. This could suggest that positive carbon feedbacks might be more powerful in GCMs than in MAGICC6, or that MAGICC6 has more effective ocean or land carbon sinks. See Jones et al. (2013) and Friedlingstein et al. (2014) for further discussions. The inverse is true for CMIP6, where the ensemble mean of compatible emissions are systematically larger than SSP scenario emissions. Note that although RCP8.5 and SSP5-8.5 have approximately the same radiative forcing in 2100, their CO_2 concentrations are different, and the differences in emissions are not unexpected. These differences have important repercussions for the interpretation of scenario probability.

The mean and standard deviations shown in Figure 8 parameterize the normal distribution of Eq. (4), which acts as the likelihood term in Eq. (2). To clarify, Eq. (4) es-

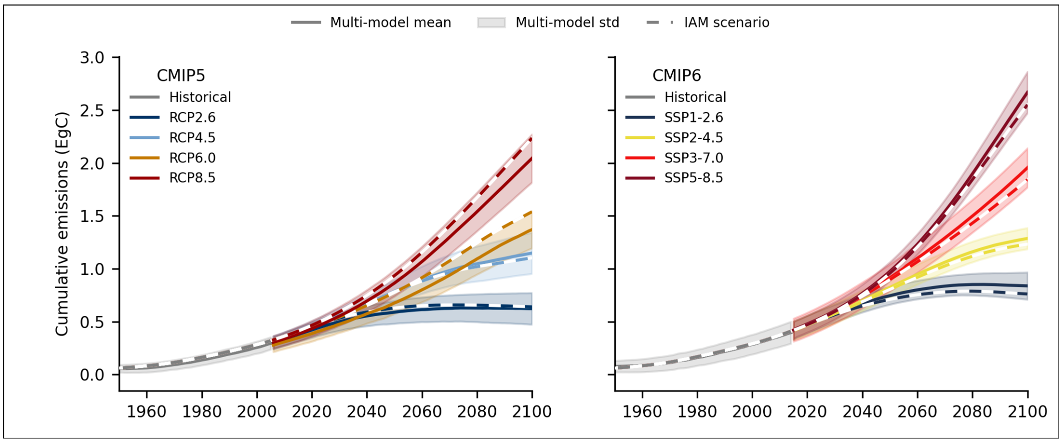

Figure 8. Mean and standard deviation of cumulative compatible emissions diagnosed from carbon fluxes in CMIP5 RCP (left) and CMIP6 SSP (right) Tier 1 experiments.

sentially evaluates the overlap between the distribution of compatible cumulative emissions for each scenario, and the distribution of probabilistic IAM cumulative emissions. This is illustrated in Figure 9 for one ensemble of probabilistic emission and one year.

The resulting relative likelihoods computed for CMIP5 and CMIP6 GHG scenarios and the five probabilistic emission simulations are presented in Figure 10. A *posterior* distribution can be obtained by multiplying these likelihoods with a *prior* for each scenario (Eq. 2). The likelihood time series are included in the supplementary material to enable assessments with different subjective *priors*. One notable feature is that high-end emission scenario RCP8.5 remains reasonably likely in all five IAMs until around 2060, despite the fact that it lies in the upper tail of probabilistic emissions from R17, LR21 and CP20. This is due to Eq. (4) evaluating scenario likelihood against compatible emissions, which for RCP8.5 are smaller than scenario emissions (see Fig. 2). Another feature worth highlighting is the low likelihood of RCP2.6 in all IAMs. Interestingly, for R17, LR21 and CP20, SSP1-2.6 is more likely than SSP3-7.0 by the end of the century.

As mentioned earlier, results from this probability assessment should not be interpreted too literally. For one, probabilistic CO_2 emissions are not directly comparable among the different IAMs. The CO_2 emission processes simulated by each probabilistic IAMs are different, some including cement production and industrial uses, while others do not.

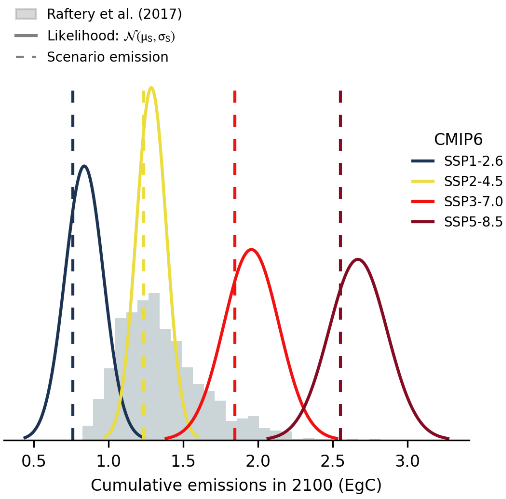

Figure 9. Illustrative histogram of cumulative emissions from Raftery et al. (2017) compared to the likelihood term (full line) for the four different SSP scenarios in 2100. SSP scenario emissions (dashed line) are shown for reference.

Also, emissions from CP16 and CP20 account for land-use changes, while compatible emissions do not. Those differences directly affect the estimated likelihoods.

Secondly, each IAM deals with policies very differently. Capellán-Pérez et al. (2020) include numerous policies regarding low carbon technologies, energy efficiency, recycling, transportation and afforestation. Fyke and Matthews (2015) includes policy-related parameters such as a maximum carbon price, a carbon tax or non-fossil energy unit cost. In contrast, Capellán-Pérez et al. (2016), Raftery et al. (2017) and Liu and Raftery (2021) include no explicit parameterization for climate policies.

Finally, the relatively small number of IAMs and the fact that multi-model ensembles are not homogeneous across experiments artificially distorts the likelihood estimation. For instance, in CMIP5 we would expect each RCP's likelihood to start at 25%, because in 2006 the four RCPs have exactly the same CO_2 concentration. With identical CO_2 concentration, the diagnosed emissions and their statistics should be very similar, except for small variations due to natural variability. Small ensemble sizes and differences in model make-up can shift the mean and variance of compatible emissions compared to the other RCPs, artificially perturbing the likelihood. This example gives a sense

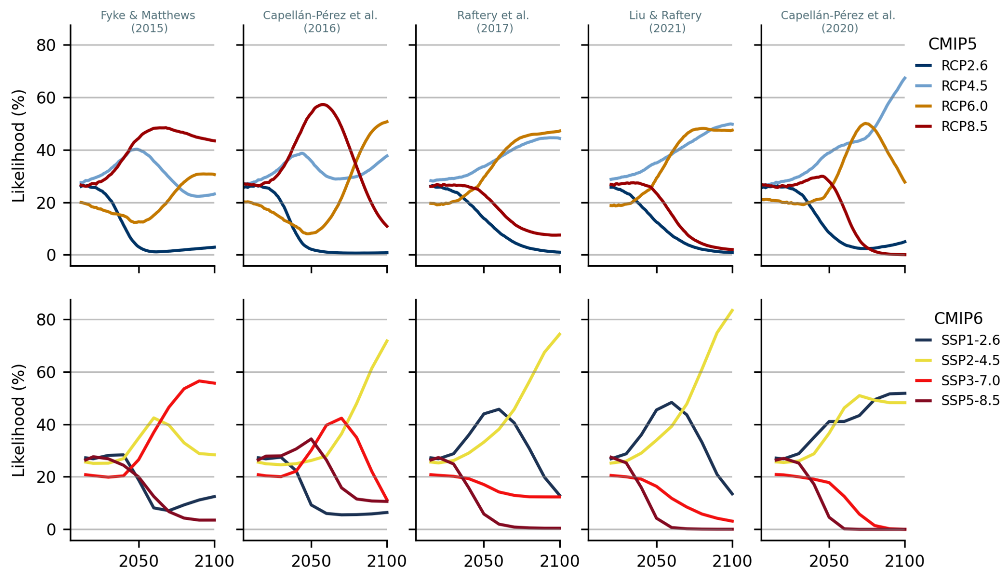

Figure 10. Bayesian likelihood for CMIP5 RCPs (top) and CMIP6 SSPs (bottom) CO_2 concentration pathways conditional on the probabilistic IAM CO_2 cumulative emission simulations ensembles from FM15, CP16, R17, LR21 and CP20. In theory, all scenarios should start with a likelihood close to 25% because the CO_2 concentration are identical at the start of each experiment. Discrepancies seen in CMIP5 RCP6.0 and SSP3-70 are due to differences in the CMIP ensemble make-up.

of the magnitude of the sampling error's influence on the results, and cast doubts on the applicability of these likelihoods for actual decision-making.

4 Discussion

The motivation for this paper is grounded in the day-to-day experience of climate service providers. Climate services strive to translate the best available climate science into actionable information. Because key climate projection experiments are run with GHG concentration scenarios, derivative products such as climate impact assessments or risk analyses are also conditional on GHG concentrations. In applications such as engineering or flood zone mapping, where a single value reflecting a level of risk is required, the lack of guidance on scenario probability acts as a barrier to climate adaptation.

One argument against assigning probabilities to emission scenarios is that it would be affected by *reflexive uncertainty* originating from human feedback to new information (Dessai & Hulme, 2004; van Vuuren et al., 2008). That is, the act of assigning probabilities to scenarios would change the probability of these scenarios, making future climate change "unquantifiable". Although this argument might hold in the abstract, the idea that an academic paper would have a significant influence on global carbon emissions can nowadays only be met with irony.

The reluctance of the climate community to assign probabilities to future GHG emission scenarios has led to the study of alternative decision-making approaches, often referred to as "Decision Making under Deep Uncertainty" (Stanton & Roelich, 2021). One suggestion stemming from these efforts is to switch the focus of discussions from *agreement-on-assumptions*, e.g. climate modeling assumptions, to *agreement-on-decisions* in order to find solutions that perform well under a wide array of future conditions and minimize regret (Kalra et al., 2014). Although valuable and illuminating, it is not clear how these concepts apply to decisions bound by strict regulatory frameworks, such as engineering or flood safety. Asking practitioners to overhaul laws, professional norms and regulations to account for climate change' deep uncertainties is sure to delay adaptation actions.

In the absence of a fully probabilistic decision-making framework, real-world adaptation decisions are made by non-experts, based on *ad hoc* selection of climate scenarios based on data availability, the precautionary principle, personal opinions or hearsay. Even among experts, debates around the relative likelihood of climate change scenarios often struggle with the emission *vs* concentration aspects of climate change experiments (Hausfather & Peters, 2020a; Schwalm et al., 2020a; Hausfather & Peters, 2020b; Schwalm et al., 2020b). The argument that "RCP8.5 is unlikely because it requires an implausible increase in coal use" may be true for the scenario's emissions (Ritchie & Dowlatabadi, 2017), but on its own doesn't imply that impacts derived from the concentration-driven `rcp85` CMIP experiment are also unlikely. As long as climate hazards are determined by concentration-driven modeling experiments, arguments referring to emissions' likelihood will have to account for uncertainties in carbon cycle simulations and carbon feedbacks.

Ideally, climate change experiments would be carried out by Earth System Models (ESM) driven by an ensemble of representative *emission* pathways, instead of representative *concentration* pathways. Carbon cycle feedbacks would be free to fully play their role, instead of being curtailed by prescribed CO_2 concentrations. The diagnosis of compatible emissions would become unnecessary, because ESM emission pathways could be directly compared with IAM simulated emissions.

Similarly, with only five probabilistic emission ensembles to draw from (two sharing the same model structure), it is clear that inter-model spread is not an accurate proxy for prediction uncertainties. Ideally, a coordinated multi-IAM experiment would be carried out, where dozens of independent modeling teams would contribute *predictive* simulations covering the same period, starting from the same initial conditions, and archiving their outputs using standardized variable definitions. IAM simulation outputs should map to ESM driving variables to facilitate the assignment of probabilities to representative *emission* pathways.

Together, these ESM and IAM coordinated experiments would provide researchers with the basic materials to conduct probabilistic climate change impact assessments, and answer a long-standing request from the climate service community and its stakeholders.

5 Conclusion

This paper adopts the argument that "it is very unhelpful to presume that all futures are equally likely" (Mckibbin et al., 2004), and suggests an approach to estimate GHG scenario probability using probabilistic emissions simulated by IAMs. Because climate models are driven by *concentration* pathways, CO_2 emissions compatible with those concentrations are estimated from CMIP5 and CMIP6 simulated carbon fluxes. For each RCP and SSP, these compatible emissions are compared with the probabilistic CO_2 emissions from five IAMs to estimate scenario's relative likelihood.

Although IAMs vary considerably in their structure and assumptions, the likelihoods obtained share similar traits. All rank RCP2.6 as the least likely until 2075. Although RCP8.5 depicts extremely high emissions, it remains relatively likely up until 2060. This is due to the fact that compatible emissions for high-end scenarios are considerably lower than their corresponding scenario emissions. For three IAMs out of five, SSP1-2.6 ends up more likely than SSP3-7.0 and SSP5-8.5 by the end of the century, with SSP5-8.5' likelihood dropping to zero in three IAMs.

The approach presented suffers from a number of caveats that cast doubts on the reliability of results. The objective is not necessarily for these likelihoods to be used in practice, but rather to illustrate the potential of probabilistic IAMs to inform scenario probability. Hopefully new climate and IAM experiments can be designed that will better address the need for fully probabilistic climate change risk assessments.

Appendix A CMIP Simulations Availability

Search requests on ESGF for CMIP5 and CMIP6 simulations were last updated in August 2021. A simulation is considered available if both variables `fgco2` and `nbp` are present for the historical period and at least one future experiment. In CMIP5, models MRI-ESM1 and CMCC-CESM were not considered due to the presence of abnormalities in the data. Also, INMCM4 was kept out of the analysis because it does not represent land-use changes.

Table A1. Number of simulations storing land and ocean carbon fluxes for each CMIP5 transient scenario experiment.

Model name	Historical	RCP2.6	RCP4.5	RCP6.0	RCP8.5
CanESM2	5	5	5	0	5
GFDL-ESM2G	1	1	1	1	1
GFDL-ESM2M	1	1	1	1	1
HadGEM2-CC	3	0	1	0	3
HadGEM2-ES	4	4	4	4	4
IPSL-CM5A-LR	6	4	4	1	4
IPSL-CM5A-MR	3	1	1	0	1
IPSL-CM5B-LR	1	0	1	0	1
MIROC-ESM	3	1	1	1	1
MIROC-ESM-CHEM	1	1	1	1	1
MPI-ESM-LR	3	3	3	0	3
MPI-ESM-MR	3	1	3	0	1
NorESM1-ME	1	1	1	1	1

Appendix B Preindustrial mean emissions

Even in the absence of anthropogenic carbon emissions, some models exhibit non-zero carbon fluxes. This may be due to models not having reached equilibrium, or in other cases to how they account for river outgassing. To try to avoid attributing these fluxes to fossil fuel emissions, the mean preindustrial compatible emissions are subtracted from compatible emission time series of the historical and future periods. Tables B1 and B2 present the mean compatible emissions computed over the last 30 to 50 years of the `piControl` simulation.

Table A2. Number of simulations storing land and ocean carbon fluxes for each CMIP6 Tier 1 ScenarioMIP experiment.

Model name	Historical	SSP1-2.6	SSP2-4.5	SSP3-7.0	SSP5-8.5
ACCESS-ESM1-5	29	10	12	10	10
CESM2	11	3	3	3	3
CESM2-FV2	3	0	0	0	0
CESM2-WACCM	3	1	5	3	5
CESM2-WACCM-FV2	3	0	0	0	0
CMCC-ESM2	1	1	1	1	1
CNRM-ESM2-1	9	5	10	5	5
CanESM5	65	50	50	50	50
CanESM5-CanOE	3	3	3	3	3
EC-Earth3-CC	1	0	1	0	0
GFDL-ESM4	1	1	0	1	0
INM-CM4-8	1	1	1	1	1
INM-CM5-0	3	1	1	5	1
IPSL-CM5A2-INCA	1	1	0	1	0
IPSL-CM6A-LR	32	6	11	11	6
MIROC-ES2L	31	10	30	10	10
MPI-ESM-1-2-HAM	3	0	0	3	0
MPI-ESM1-2-LR	9	10	10	10	10
MRI-ESM2-0	1	0	0	0	1
NorESM2-LM	3	1	3	3	1
NorESM2-MM	3	1	1	1	1
UKESM1-0-LL	17	16	13	10	4

Table B1. Mean preindustrial compatible emissions from the CMIP5 `piControl` experiment.

Model name	Member	Emissions (PgC)
CanESM2	r1i1p1	-0.07
GFDL-ESM2G	r1i1p1	0.22
HadGEM2-CC	r1i1p1	0.34
HadGEM2-ES	r1i1p1	-0.29
IPSL-CM5A-MR	r1i1p1	0.00
IPSL-CM5B-LR	r1i1p1	-0.10
MIROC-ESM	r1i1p1	0.09
MPI-ESM-LR	r1i1p1	-0.06
NorESM1-ME	r1i1p1	0.17

Table B2. Mean preindustrial compatible emissions from the CMIP6 `piControl` experiment.

Model name	Member	Emissions (PgC)
ACCESS-ESM1-5	r1i1p1f1	-0.21
CESM2	r1i1p1f1	-0.06
CESM2-FV2	r1i1p1f1	-0.07
CESM2-WACCM	r1i1p1f1	-0.09
CNRM-ESM2-1	r1i1p1f2	-0.71
CanESM5	r1i1p1f1	-0.09
CanESM5-CanOE	r1i1p2f1	-0.07
INM-CM4-8	r1i1p1f1	1.12
INM-CM5-0	r1i1p1f1	1.15
IPSL-CM6A-LR	r1i1p1f1	-0.01
MIROC-ES2L	r1i1p1f2	0.11
MPI-ESM1-2-LR	r1i1p1f1	-0.03
MPI-ESM1-2-LR	r2i1p1f1	0.15
MRI-ESM2-0	r1i2p1f1	0.41
UKESM1-0-LL	r1i1p1f2	-0.09

Open Research

Compatible emissions, probabilistic emissions, scenario emissions and observations used in this paper are available at the Federated Research Data Repository at `https://doi.org/10.20383/102.0549`, along with code to compute the likelihood, create graphics and tables.

CMIP data was downloaded from the Earth System Grid Federation using Synda.

SSP scenario emissions shown in figures 2 to 9 are based on data from the SSP database hosted by the IIASA Energy Program at https://tntcat.iiasa.ac.at/SspDb

Time series of CO_2 concentrations for RCP and SSP scenarios were obtained from International Institute for Applied Systems Analysis (IIASA) RCP and SSP databases.

Observed CO_2 time series from the Global Carbon Budget where obtained from the Integrated Carbon Observing System (`https://www.icos-cp.eu/`).

Probabilistic emissions from FM15 are available at `https://github.com/JeremyFyke/CEPM/blob/results/results/carbon_emissions.h5`

Code to generate probabilistic emissions from R17 can be found at `https://github.com/PPgp/CO2projections`.

Code to generate probabilistic emissions from LR21 can be found at `https://github.com/PPgp/BayesianClimateProjections`.

Acknowledgments

We acknowledge the World Climate Research Programmes Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (whose models are listed in Table A1 and A2 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energys Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

D.H. work is funded in part by Transport Canada. Many thanks go to Ramón de Elía for the impetus to write this paper, Diane Chaumont for the opportunity to finish it, and Martin Leduc and Patrick Grenier for their suggestions.

The authors also thank David Álvarez Antelo for re-running the n=1000 Monte Carlo simulations for CP20 and extract CO_2 samples, Gaëlle Rigoudy and Roland Séférian for help regarding CNRM-ESM2.1 fluxes, Michio Kawamiya for help regarding MIROC-ESM, and Spencer Liddicoat for clarifications regarding compatible flux computations.

References

Ackerman, F., Stanton, E. A., & Bueno, R. (2010, jun). Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE. *Ecological Economics*, *69*(8), 1657–1665. Retrieved from `https://linkinghub.elsevier.com/retrieve/pii/S0921800910001096` doi: 10.1016/j.ecolecon.2010.03.013

Agrawala, S., Bosello, F., Carraro, C., De Bruin, K., De Cian, E., Dellink, R., & Lanzi, E. (2011, aug). Plan or React? Analysis of Adaptation Costs and Benefits Using Integrated Assessment Models. *Climate Change Economics*, *02*(03), 175–208. Retrieved from `http://www.worldscientific.com/doi/abs/10.1142/S2010007811000267` doi: 10.1142/S2010007811000267

Cai, Y., Lenton, T. M., & Lontzek, T. S. (2016, may). Risk of multiple interacting tipping points should encourage rapid CO2 emission reduction. *Nature Climate Change*, *6*(5), 520–525. Retrieved from `http://www.nature.com/articles/nclimate2964` doi: 10.1038/nclimate2964

Capellán-Pérez, I. (2016). *Development and application of environmental integrated assessment modelling towards sustainability* (Doctoral dissertation, University of the Basque Country). Retrieved from `https://addi.ehu.es/handle/10810/19438`

Capellán-Pérez, I., Arto, I., Polanco-Martínez, J. M., González-Eguino, M., & Neumann, M. B. (2016). Likelihood of climate change pathways under uncertainty on fossil fuel resource availability. *Energy and Environmental Science*, *9*(8), 2482–2496. doi: 10.1039/c6ee01008c

Capellán-Pérez, I., De Blas, I., Nieto, J., De Castro, C., Miguel, L. J., Carpintero, Ó., … Álvarez-Antelo, D. (2020). MEDEAS: A new modeling framework integrating global biophysical and socioeconomic constraints. *Energy and Environmental Science*, *13*(3), 986–1017. doi: 10.1039/c9ee02627d

Dessai, S., & Hulme, M. (2004). Does climate adaptation policy need probabilities? *Climate Policy*, *4*(2), 107–128. doi: 10.1080/14693062.2004.9685515

Eyring, V., Bony, S., Meehl, G. A., Senior, C., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2015). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation. *Geoscientific Model Development Discussions*, *8*(12), 10539–10583. Retrieved from http://www.geosci-model-dev-discuss.net/8/10539/2015/ doi: 10.5194/gmdd-8-10539-2015

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *27*(2), 511–526. doi: 10.1175/JCLI-D-12-00579.1

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., ... Zaehle, S. (2020, dec). Global Carbon Budget 2020. *Earth System Science Data*, *12*(4), 3269–3340. Retrieved from https://essd.copernicus.org/articles/12/3269/2020/ doi: 10.5194/essd-12-3269-2020

Fyke, J., & Matthews, H. D. (2015). A probabilistic analysis of cumulative carbon emissions and long-term planetary warming. *Environmental Research Letters*, *10*(11), 115007. doi: 10.1088/1748-9326/10/11/115007

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., ... Takahashi, K. (2019). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: A dataset of harmonized emissions trajectories through the end of the century. *Geoscientific Model Development*, *12*(4), 1443–1475. doi: 10.5194/gmd-12-1443-2019

Gillingham, K., Nordhaus, W., Anthoff, D., Blanford, G., Bosetti, V., Christensen, P., ... Reilly, J. (2018). Modeling uncertainty in integrated assessment of climate change: A multimodel comparison. *Journal of the Association of Environmental and Resource Economists*, *5*(4), 791–826. doi: 10.1086/698910

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... Oliphant, T. E. (2020, sep). Array programming with NumPy. *Nature*, *585*(7825), 357–362. Retrieved from https://www.nature.com/articles/s41586-020-2649-2 doi: 10.1038/s41586-020-2649-2

Hausfather, Z., & Peters, G. P. (2020a, jan). Emissions the business as usual' story is misleading. *Nature*, *577*(7792), 618–620. Retrieved from http://www.nature.com/articles/d41586-020-00177-3 doi:

10.1038/d41586-020-00177-3

Hausfather, Z., & Peters, G. P. (2020b). RCP8.5 is a problematic scenario for near-term emissions. *Proceedings of the National Academy of Sciences*, *117*(45), 202017124. doi: 10.1073/pnas.2017124117

Hieronymus, M. (2020, nov). Nonlinear Interactions and Some Other Aspects of Probabilistic Sea Level Projections. *Water*, *12*(11), 3212. Retrieved from `https://www.mdpi.com/2073-4441/12/11/3212` doi: 10.3390/w12113212

Ho, E., Budescu, D. V., Bosetti, V., van Vuuren, D. P., & Keller, K. (2019). Not all carbon dioxide emission scenarios are equally likely: a subjective expert assessment. *Climatic Change*, *155*(4), 545–561. doi: 10.1007/s10584-019-02500-y

Hoyer, S., & Hamman, J. J. (2017, apr). xarray: N-D labeled Arrays and Datasets in Python. *Journal of Open Research Software*, *5*. Retrieved from `http://openresearchsoftware.metajnl.com/articles/10.5334/jors.148/` doi: 10.5334/jors.148

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. Retrieved from `http://ieeexplore.ieee.org/document/4160265/` doi: 10.1109/MCSE.2007.55

Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., ... Tjiputra, J. (2013). Twenty-first-century compatible CO2 emissions and airborne fraction simulated by CMIP5 earth system models under four Representative Concentration Pathways. *Journal of Climate*, *26*(13), 4398–4413. doi: 10.1175/JCLI-D-12-00554.1

Kalra, N., Hallegate, S., Lempert, R., Brown, C., Fozzard, A., Gill, S., & Shah, A. (2014). *Agreeing on Robust Decisions New Processes for Decision Making Under Deep Uncertainty* (No. June). Retrieved from `http://hdl.handle.net/10986/18772`

Keen, S. (2020, sep). The appallingly bad neoclassical economics of climate change. *Globalizations*, 1–29. Retrieved from `https://www.tandfonline.com/doi/full/10.1080/14747731.2020.1807856` doi: 10.1080/14747731.2020.1807856

King, D., Schrag, D., Dadi, Z., Ye, Q., & Ghosh, A. (2015). *Climate change A risk assessment* (Tech. Rep.). doi: 10.1111/j.1740-9713.2010.00403.x

Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., & Eyring,

V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters*. Retrieved from `http://doi.wiley.com/10.1002/2016GL072012` doi: 10.1002/2016GL072012

Koomey, J., Schmidt, Z., Hummel, H., & Weyant, J. (2019). Inside the Black Box: Understanding key drivers of global emission scenarios. *Environmental Modelling and Software*, *111*(January 2018), 268–281. Retrieved from `https://doi.org/10.1016/j.envsoft.2018.08.019` doi: 10.1016/j.envsoft.2018.08.019

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E., Brunner, L., … Hawkins, E. (2020). Partitioning climate projection uncertainty with multiple Large Ensembles and CMIP5/6. *Earth System Dynamics Discussions*, 1–28. doi: 10.5194/esd-2019-93

Liddicoat, S. K., Wiltshire, A. J., Jones, C. D., Arora, V. K., Brovkin, V., Cadule, P., … Ziehn, T. (2021). Compatible fossil fuel CO2 emissions in the CMIP6 earth system models' historical and shared socioeconomic pathway experiments of the twenty-first century. *Journal of Climate*, *34*(8), 2853–2875. doi: 10.1175/JCLI-D-19-0991.1

Liu, P. R., & Raftery, A. E. (2021). Country-based rate of emissions reductions should increase by 80% beyond nationally determined contributions to meet the 2 řC target. *Communications Earth & Environment*, *2*(1), 1–10. Retrieved from `http://dx.doi.org/10.1038/s43247-021-00097-8` doi: 10.1038/s43247-021-00097-8

Lontzek, T. S., Cai, Y., Judd, K. L., & Lenton, T. M. (2015, may). Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy. *Nature Climate Change*, *5*(5), 441–444. Retrieved from `http://www.nature.com/articles/nclimate2570` doi: 10.1038/nclimate2570

Matthews, H. D., Gillett, N. P., Stott, P. A., & Zickfeld, K. (2009, jun). The proportionality of global warming to cumulative carbon emissions. *Nature*, *459*(7248), 829–832. Retrieved from `http://www.nature.com/doifinder/10.1038/nature08047` doi: 10.1038/nature08047

Mckibbin, W. J., Pearce, D., & Stegman, A. (2004). Long Run Projections for Climate Change Scenarios. (160).

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In (pp. 56–61). Retrieved from https://conference.scipy.org/proceedings/scipy2010/mckinney.html doi: 10.25080/Majora-92bf1922-00a

Meinshausen, M., Nicholls, Z., Lewis, J., Gidden, M., Vogel, E., Freund, M., ... Wang, H. J. (2019). The SSP greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development Discussions*, 1–77. doi: 10.5194/gmd-2019-222

Meinshausen, M., Raper, S. C. B., & Wigley, T. M. L. (2011, feb). Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 Part 1: Model description and calibration. *Atmospheric Chemistry and Physics*, *11*(4), 1417–1456. Retrieved from https://acp.copernicus.org/articles/11/1417/2011/ doi: 10.5194/acp-11-1417-2011

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., ... van Vuuren, D. P. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, *109*(1-2), 213–241. Retrieved from http://link.springer.com/10.1007/s10584-011-0156-z doi: 10.1007/s10584-011-0156-z

Morgan, M. G., & Keith, D. W. (2008). Improving the way we think about projecting future energy use and emissions of carbon dioxide. *Climatic Change*, *90*(3), 189–215. doi: 10.1007/s10584-008-9458-1

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., ... Wilbanks, T. J. (2010, feb). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20148028http://www.nature.com/doifinder/10.1038/nature08823 doi: 10.1038/nature08823

Moss, R. H., & Schneider, S. H. (2000). Uncertainties in the IPCC TAR: Recommendations to Lead Authors for More Consistent Assessment and Reporting. In R. Pachauri, T. Taniguchi, & K. Tanaka (Eds.), *Guidance papers on the cross cutting issues of the third assessment report of the ipcc.* WMO.

Murphy, J., Sexton, D., Jenkins, G., Boorman, P., Booth, B., Brown, C., ... Wood, R. A. (2009). *UK Climate Projections Science Report: Climate change projections.* Exeter: Met Office Hadley Center.

Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., ...

Others (2000). *Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change* (Tech. Rep.).

Nasser, B., Levavasseur, G., & Denvil, S. (2020). *Synda.* Retrieved from `https://github.com/Prodiguer/synda{\%}22`

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. doi: 10.5194/gmd-9-3461-2016

Pastor, A., Vieira, D., Soudijn, F., & Edelenbosch, O. (2020, mar). How uncertainties are tackled in multi-disciplinary science? A review of integrated assessments under global change. *CATENA*, *186*, 104305. Retrieved from `https://linkinghub.elsevier.com/retrieve/pii/S0341816219304473` doi: 10.1016/j.catena.2019.104305

Raftery, A. E., Zimmer, A., Frierson, D. M., Startz, R., & Liu, P. (2017). Less than 2 řc warming by 2100 unlikely. *Nature Climate Change*, *7*(9), 637–641. doi: 10.1038/nclimate3352

Repetto, R., & Easton, R. (2015). Analyzing Climate Uncertainty and Risk with an Integrated Assessment Model. *SAIS Review of International Affairs*, *35*(1), 47–59. doi: 10.1353/sais.2015.0005

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., ... Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. doi: 10.1016/j.gloenvcha.2016.05.009

Ritchie, J., & Dowlatabadi, H. (2017, dec). Why do climate change scenarios return to coal? *Energy*, *140*, 1276–1291. Retrieved from `https://linkinghub.elsevier.com/retrieve/pii/S0360544217314597` doi: 10.1016/j.energy.2017.08.083

Rogelj, J., Fricko, O., Meinshausen, M., Krey, V., Zilliacus, J. J., & Riahi, K. (2017). Understanding the origin of Paris Agreement emission uncertainties. *Nature Communications*, *8*(June), 1–12. Retrieved from `http://dx.doi.org/10.1038/ncomms15748` doi: 10.1038/ncomms15748

Schneider, S. H. (2001, may). What is 'dangerous' climate change? *Nature*, *411*(6833), 17–19. Retrieved from http://www.nature.com/doifinder/10.1038/35075167 doi: 10.1038/35075167

Schneider, S. H., & Mastrandrea, M. D. (2005). Probabilistic assessment of "dangerous" climate change and emissions pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(44), 15728–15735. doi: 10.1073/pnas.0506356102

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020a). RCP8.5 tracks cumulative CO2 emissions. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(33), 19656–19657. doi: 10.1073/PNAS.2007117117

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020b). Reply to Hausfather and Peters: RCP8.5 is neither problematic nor misleading. *Proceedings of the National Academy of Sciences*, *117*(45), 202018008. doi: 10.1073/pnas.2018008117

Sokolov, A. P., Schlosser, C. A., Dutkiewicz, S., Paltsev, S., Kicklighter, D. W., Jacoby, H. D., ... Cohen, J. (2005). *The MIT Integrated Global System Model (IGSM) Version 2 : Model Description and Baseline Evaluation* (Tech. Rep. No. 124). MIT Joint Program on the Science and Policy of Global Change. doi: 10.1002/jame.20044

Sokolov, A. P., Stone, P. H., Forest, C. E., Prinn, R., Sarofim, M. C., Webster, M., ... Jacoby, H. D. (2009). Probabilistic forecast for twenty-first-century climate based on uncertainties in emissions (without policy) and climate parameters. *Journal of Climate*, *22*(19), 5175–5204. doi: 10.1175/2009JCLI2863.1

Stanton, M. C. B., & Roelich, K. (2021, oct). Decision making under deep uncertainties: A review of the applicability of methods in practice. *Technological Forecasting and Social Change*, *171*, 120939. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0040162521003711 doi: 10.1016/j.techfore.2021.120939

Stern, N. (2013, sep). The Structure of Economic Modeling of the Potential Impacts of Climate Change: Grafting Gross Underestimation of Risk onto Already Narrow Science Models. *Journal of Economic Literature*, *51*(3), 838–859. Retrieved from https://pubs.aeaweb.org/doi/10.1257/jel.51.3.838 doi: 10.1257/jel.51.3.838

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485.

United Nations. (2015). *World Population Prospects: The 2015 Revision* (Tech. Rep.). New York, New York, USA: Department of Economic and Social Affairs, Population Division.

United Nations. (2019). *World Population Prospects: The 2019 Revision* (Tech. Rep.). New York, New York, USA.

van Vuuren, D. P., de Vries, B., Beusen, A., & Heuberger, P. S. (2008). Conditional probabilistic estimates of 21st century greenhouse gas emissions based on the storylines of the IPCC-SRES scenarios. *Global Environmental Change*, *18*(4), 635–654. doi: 10.1016/j.gloenvcha.2008.06.001

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., ... Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change*, *109*(1), 5–31. doi: 10.1007/s10584-011-0148-z

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... Vázquez-Baeza, Y. (2020, mar). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. Retrieved from `http://www.nature.com/articles/s41592-019-0686-2` doi: 10.1038/s41592-019-0686-2

Ward, J. D., Mohr, S. H., Myers, B. R., & Nel, W. P. (2012). High estimates of supply constrained emissions scenarios for long-term climate risk assessment. *Energy Policy*, *51*, 598–604. Retrieved from `http://dx.doi.org/10.1016/j.enpol.2012.09.003` doi: 10.1016/j.enpol.2012.09.003

Webster, M., Forest, C., Reilly, J., Babiker, M., Kicklighter, D., Mayer, M., ... Wang, C. (2003, dec). Uncertainty Analysis of Climate Change and Policy Response. *Climatic Change*, *61*(3), 295–320. Retrieved from `http://www.springerlink.com/openurl.asp?id=doi:10.1023/B:CLIM.0000004564.09961.9fhttp://link.springer.com/10.1023/B:CLIM.0000004564.09961.9f` doi: 10.1023/B:CLIM.0000004564.09961.9f

Webster, M., Paltsev, S., Parsons, J., Reilly, J., & Jacoby, H. (2008). Uncertainty in Greenhouse Gas Emissions and Costs of Atmospheric Stabilization. *MIT Joint Program on the Science and Policy of Global Change*(165), 1–81. Retrieved from `http://globalchange.mit.edu/pubs/abstract.php?publication{\`

_}id=974

Webster, M., Sokolov, A. P., Reilly, J. M., Forest, C. E., Paltsev, S., Schlosser, A., … Jacoby, H. D. (2012, jun). Analysis of climate policy targets under uncertainty. *Climatic Change*, *112*(3-4), 569–583. Retrieved from http://link.springer.com/10.1007/s10584-011-0260-0 doi: 10.1007/s10584-011-0260-0

Webster, M. D., Babiker, M., Mayer, M., Reilly, J. M., Harnisch, J., Hyman, R., … Wang, C. (2002). Uncertainty in emissions projections for climate models. *Atmospheric Environment*, *36*(22), 3659–3670. doi: 10.1016/S1352-2310(02)00245-5

Weitzman, M. L. (2012, mar). GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal of Public Economic Theory*, *14*(2), 221–244. Retrieved from https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9779.2011.01539.x doi: 10.1111/j.1467-9779.2011.01539.x

Williams, R. G., Goodwin, P., Ridgwell, A., & Woodworth, P. L. (2012, oct). How warming and steric sea level rise relate to cumulative carbon emissions. *Geophysical Research Letters*, *39*(19), 2012GL052771. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2012GL052771 doi: 10.1029/2012GL052771