



Consequences of equivalency metric design for energy transitions and climate change

Morgan R. Edwards^{1,2,3} · Jessika E. Trancik^{3,4}

Received: 1 October 2019 / Accepted: 25 September 2022 / Published online: 11 November 2022
© The Author(s) 2022

Abstract

Assessments of the climate impacts of energy technologies and other emissions sources can depend strongly on the equivalency metric used to compare short- and long-lived greenhouse gas emissions. However, the consequences of metric design choices are not fully understood, and in practice, a single metric, the global warming potential (GWP), is used almost universally. Many metrics have been proposed and evaluated in recent decades, but questions still remain about which ones perform better and why. Here, we develop new insights on how the design of equivalency metrics can impact the outcomes of climate policies. We distill the equivalency metric problem into a few key design choices that determine the metric values and shapes seen across a wide range of different proposed metrics. We examine outcomes under a hypothetical 1.5 or 2 °C policy target and discuss extensions to other policies. Across policy contexts, the choice of time parameters is particularly important. Metrics that emphasize the immediate impacts of short-lived gases such as methane can reduce rates of climate change but may require more rapid technology changes. Differences in outcomes across metrics are more pronounced when fossil fuels, with or without carbon capture and storage, play a larger role in energy transitions. By identifying a small set of consequential design decisions, these insights can help make metric choices and energy transitions more deliberate and effective at mitigating climate change.

Keywords Equivalency metrics · Energy systems · Climate change · Life cycle assessment

✉ Morgan R. Edwards
morgan.edwards@wisc.edu

✉ Jessika E. Trancik
trancik@mit.edu

¹ La Follette School of Public Affairs, University of Wisconsin—Madison, Madison, WI 53706, USA

² Nelson Institute Center for Sustainability and the Global Environment, University of Wisconsin—Madison, Madison, WI 53706, USA

³ Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁴ Santa Fe Institute, Santa Fe, NM 87501, USA

1 Introduction

Efforts to mitigate climate change, such as replacing coal with natural gas in the energy sector (Brandt et al. 2014; McJeon et al. 2014; Zhang et al. 2016; Alvarez et al. 2012; Alvarez et al. 2018; Tanaka et al. 2019; Mallapragada and Mignone 2017; Klemun and Trancik 2020), may involve a tradeoff between short-lived greenhouse gases (e.g., methane, CH₄) and long-lived carbon dioxide (CO₂). Equivalency metrics are frequently used to convert emissions of various gases into equal mass emissions of CO₂ (in CO₂-equivalent units). These metrics enable decision-makers to assess the climate impacts of technologies and policies on a single scale. For example, they are used to evaluate the climate impacts of different energy technologies and infrastructure investments, communicate and compare policy proposals to reduce emissions, and establish exchange rates between gases in emissions trading schemes.

One metric, the global warming potential (GWP) with a time horizon of 100 years (Lashof and Ahuja 1990; Rodhe 1990), is used almost universally for these purposes (Cherubini et al. 2016; Levasseur et al. 2016). It has also been adopted as the official metric for implementing the Paris Agreement (United Nations 2019), with plans to review and potentially revise this choice in the future. The GWP(100) was initially proposed as a placeholder over 30 years ago to illustrate the difficulties in comparing greenhouse gases to one another (Schmalensee 1993; O'Neill 2000; Shine 2009). Over the intervening decades, critics have questioned its choice of impact measure and arbitrary time horizon, which captures neither the near-term impacts of short-lived gases such as CH₄ nor the long-term impacts of CO₂ (Eckaus 1992; Ocko et al. 2017; Fesenfeld et al. 2018; Daniel et al. 2012). They argue the GWP(100) was not designed to inform major decisions, and that it is past time to reevaluate its use and consider alternatives (Shine 2009; Plattner et al. 2009).

A variety of metrics have been proposed to address these critiques and provide alternatives for operationalizing climate policies and evaluating the energy technologies currently attracting hundreds of billions of dollars in investments (International Energy Agency 2019). Some of these metrics compare emissions over different time horizons or based on different physical impacts (Shine et al. 2005; Gillett and Matthews 2010; Peters et al. 2011; Sterner et al. 2014; Shine et al. 2015; Mallapragada and Mignone 2020; Sarofim and Giordano 2018). Others are inspired by a climate policy goal (Edwards and Trancik 2014; Shine et al. 2007; Johansson 2012) or calculated using forecasts of the costs (and, occasionally, benefits) of mitigation using an integrated assessment model (Reilly and Richards 1993; Kandlikar 1996; Manne and Richels 2001). As with the GWP(100), these alternative metrics express non-CO₂ emissions in CO₂-equivalent units, allowing direct comparisons between technologies and policies with different CO₂ and non-CO₂ impacts.

In lieu of proposing drop-in replacements for the GWP(100), some have called for restructuring the ways in which emissions are exchanged in climate policy to account for the different lifetimes of greenhouse gases. Discussions in this area have focused on the relative benefits of single- versus multi-basket emissions policies, where multi-basket policies group emissions based on lifetimes and limit or prohibit trading between baskets (Godol and Fuglestedt 2002; Daniel et al. 2012). Multi-basket policies can help ensure emissions reductions across all climate forcing agents and reduce uncertainty in the climate outcomes of policies. However, prohibiting trade also reduces flexibility in the mitigation pathway, potentially increasing policy costs (Tanaka et al. 2010).

More recently, researchers have proposed a middle ground where a change in pulse emissions of long-lived greenhouse gases is equated to a permanent change in the

emissions rate of a short-lived greenhouse gas (Smith et al. 2012). A metric called the GWP* (and various approximations; Collins et al. 2020) has been proposed as an alternative to the traditional GWP for comparing these pulse and rate emissions (Allen et al. 2016, 2018; Cain et al. 2019). The GWP* can be directly used in assessments that focus on emissions pathways, for example in evaluating emissions reduction pledges to the Paris Agreement (Lynch et al. 2020). However, because it requires knowledge of the future stream of short-lived emissions, the policy contexts in which it is applicable are constrained to those where future emissions can be controlled. It is less applicable to assessing energy technology choices or evaluating and implementing less durable emissions reduction policies.

Despite these many proposals for new metrics, a comprehensive understanding of the consequences of metric design choices is missing. Many existing analyses focus on cataloging metric design choices and the intentions behind them (Schmalensee 1993; Tanaka et al. 2010; Manning and Reisinger 2011; Tol et al. 2012; Boucher 2012; Deuber et al. 2013; Mallapragada and Mignone 2017, 2020; Fuglestedt et al. 2003; O'Neill 2003; Plattner et al. 2009). For example, the existing literature has shown how different metric designs relate to one another and to the Global Cost Potential (GCP) and Global Damage Potential (GDP) (Tol et al. 2012; Deuber et al. 2013; Mallapragada and Mignone 2020). However, relating the design of metrics to one another does not necessarily provide insight on the consequences of applying those metrics. Such insights are important for allowing decision-makers to select an equivalency metric to match their objectives.

Some studies have begun to fill this gap in understanding by comparing different metrics in specific use cases. For example, researchers have studied how metrics can influence assessments of the relative attractiveness of technologies with different CO₂ and non-CO₂ emissions intensities for climate change mitigation. Life cycle assessments have highlighted the particularly critical role the choice of equivalency metric plays in evaluating the climate impacts of natural gas, biofuels, and other energy sources with high CH₄ emissions (Edwards and Trancik 2014; Alvarez et al. 2012; Frank et al. 2012; Edwards et al. 2017; Boucher and Reddy 2008; Fuglestedt et al. 2009; Cherubini et al. 2016; Levasseur et al. 2016).

Other studies have begun to explore the consequences of different metric choices by simulating the effects of applying them in various climate policies. This research has shown that “equivalent” emissions of CO₂ and CH₄ can result in very different climate impacts over time under a single-basket emissions policy, including overshoots of climate policy thresholds and different temperature outcomes and definitions of “net-zero” (Fuglestedt et al. 2000; Edwards and Trancik 2014; Edwards et al. 2016; Daniel et al. 2012; Smith et al. 2013; Fuglestedt et al. 2018; Tanaka et al. 2021; Wigley 2021). Other work has shown that the choice of metric can influence the costs of mitigation (Ekholm et al. 2013; Reisinger et al. 2013; Harmsen et al. 2016; van den Berg et al. 2015) and the distribution of mitigation burdens across regions and sectors of the economy (Brennan and Zaitchik 2013; Strefler et al. 2014).

A growing subset of this literature uses integrated assessment models to explore the costs of applying various metrics, typically with an exogenously applied radiative forcing or temperature limit. Because climate outcomes are constrained, these studies focus on the extent to which different metrics increase the costs of mitigation, compared to a case with optimal emissions reductions for individual gases. The estimated additional costs (relative to the optimal solution) of choosing GWP metrics with shorter time horizons (e.g., 20 years) are generally small (less than 5% of total mitigation costs, though this amount can still be large in absolute terms) (Godol and Fuglestedt 2002; Johansson et al. 2006; Strefler et al. 2014), whereas metrics with longer time horizons (and thus lower impact values assigned to CH₄) can lead to higher costs and major differences in mitigation choices (Brennan and Zaitchik 2013; Harmsen et al. 2016; Ekholm et al. 2013; Reisinger et al. 2013).

Here we aim to contribute new insight to this literature by systematically exploring the connection between metric design choices and policy consequences. We examine the consequences of metrics in terms of climate outcomes and energy transitions under an emissions policy designed to limit global temperature change. First, we ask which metric design choices are most influential in determining metric values. We identify a few key metric design levers and vary these to generate a large set of potential equivalency metrics. This set covers metrics that have been discussed previously in the literature and also reveals new designs. Second, we ask how the choice of metric affects the outcomes of climate policies, focusing on differences in temperature outcomes (without an exogenous constraint) and the underlying technology transitions.

Specifically, we simulate the energy technology choices and temperature changes that result when different metrics are applied in emissions policies designed to meet the Paris climate goals. We do not explicitly simulate the costs of different transitions because of the large uncertainties involved in predicting supply-side technology costs and demand-side changes. Rather, we design our simulation to span a wide range of different possible transition pathways and draw conclusions about metric consequences that apply across those different pathways. These generalizable insights can help inform metric selection based on societal goals, such as reducing climate risks, supporting consumption or economic activity, and incentivizing energy transitions.

2 Methods

2.1 Metric design

2.1.1 Overview of design choices

The standard GWP compares greenhouse gases based on the cumulative heat trapping (or radiative forcing) effects of pulse emissions,¹ from a time of emission $t=0$ over a fixed time horizon τ ,

$$\text{GWP}(\tau) = \frac{\int_0^\tau f_X(t')c_X(t')dt'}{\int_0^\tau f_K(t')c_K(t')dt'}, \quad (1)$$

where $c_i(t')$ represents the concentration of a gas i at time t' , $f_i(t')$ is the heat trapping ability per unit concentration (i.e., radiative efficiency), X represents a generic greenhouse gas, and K represents the reference gas (CO_2 , if emissions are expressed in CO_2 -equivalent units). Both c and f may also vary over time due to the saturation of carbon sinks with continued emissions and the nonlinear relationship between concentration and radiative forcing for major greenhouse gases.

A number of alternatives to the GWP have also been proposed (see Table 1 for an overview of metrics from the literature). These metrics use different physical impacts, time-frames, and other weighting factors to compare emissions and, unlike the GWP, may change with the time of emission t . Taken together, these metric designs $\mu(t)$ can be represented by the general equation

¹ Early discussions raised the question of whether the GWP should represent the integrated effect of unit mass (or pulse) emissions or the instantaneous effects of sustained (or step) emissions changes (O'Neill 2000). Others have suggested that metrics should compare emissions over the lifetimes of technologies or investment decisions (Harvey 1993; Alvarez et al. 2012). We focus our discussion on metrics for comparing pulse emissions.

Table 1 Summary of equivalency metric designs in the literature in terms of the choice of impact measure, time horizon and weight, and resulting metric shape (when calculating equivalency values for short-lived greenhouse gases such as CH₄). Time horizons may be either cumulative or instantaneous and either fixed (i.e., taking on the same value regardless of the time of emission) or variable (i.e., changing with the time of emission, for example in relation to a particular time of interest). Fixed time horizons result in static metrics while variable time horizons result in time-dependent metrics. Metrics that have the same formula are listed together under the most common name in the literature. For example, the formula for the stabilization warming potential (SWP) (Edwards et al. 2016) is the same as the GWP, but the time horizon is calculated based on an intended timeline for stabilizing radiative forcing under a climate policy

Metric	Impact measure	Time horizon and weight	Metric shape
Global warming potential (GWP) (Lashof and Ahuja 1990; Rodhe 1990; Edwards et al. 2016)	Radiative forcing	Cumulative, fixed	Static
Global temperature potential (GTP) (Shine et al. 2005)	Temperature	Instantaneous, fixed	Static
Integrated global temperature potential (IGTP) (Gillett and Matthews 2010; Peters et al. 2011)	Temperature	Cumulative, fixed	Static
Climate change impact potential (CCIP) (Kirschbaum 2014)	Temperature ^a	Cumulative, fixed	Static
Temperature proxy index (TEMP) (Tanaka et al. 2009)	Temperature	Weighted, ^b fixed	Static
Global sea level rise potential (GSP) (Stern et al. 2014)	Sea level rise	Instantaneous, fixed	Static
Integrated global sea level rise potential (IGSP) (Stern et al. 2014)	Sea level rise	Cumulative, fixed	Static
Global precipitation potential (GPP) (Shine et al. 2015)	Precipitation	Instantaneous, fixed	Static
Global damage potential (GDP) (Reilly and Richards 1993; Kandlikar 1996)	Economic damages	Cumulative, fixed	Static
Instantaneous climate impact (ICI) (Edwards and Trancik 2014)	Radiative forcing	Instantaneous, variable	Increasing
Cumulative climate impact (CCI) (Edwards and Trancik 2014)	Radiative forcing	Cumulative, variable	Increasing
Optimized climate impact (OCI) (Edwards et al. 2016)	Radiative forcing	Weighted, ^c variable	Increasing
Dynamic global temperature potential (DGTP) (Shine et al. 2007)	Temperature	Instantaneous, variable	Increasing
Cost-effective temperature potential (CETP) (Johansson 2012)	Temperature	Weighted, ^d variable	Increasing
Global cost potential (GCP) (Mame and Richels 2001)	Temperature	Weighted, ^c variable	Increasing
Forcing equivalency index (FEI) ^e (Wigley 1998; Manning and Reisinger 2011)	Radiative forcing	Weighted, ^b variable	Decreasing

^a Impact is defined as the average of instantaneous temperature, cumulative temperature, and the rate of temperature change

^b Metrics are defined to best match the climate impacts of a historical or future emissions pathway over a specified time period

^c Weights are the output of an optimization model with a constraint on climate impacts (e.g., radiative forcing or temperature)

^d Weights are equal to zero before the expected time of temperature stabilization and discounted after stabilization is reached

^e We note that the FEI is calculated to match historical or projected radiative forcing from different greenhouse gases and thus does is not directly represented in Eq. 2; however, Eq. 2 can be used to calculate metrics with similar shapes and values to the FEI

$$\mu(t) = \frac{\int_{t_1}^{t_2} w(t', t) I_X(t', t) dt'}{\int_{t_1}^{t_2} w(t', t) I_K(t', t) dt'}, \quad (2)$$

where $I(t', t)$ represents the impact at time t' of a gas emitted at time t and $w(t', t)$ is the weight assigned to these impacts at time t' (which may vary with t). The choice of time parameters t_1 and t_2 determines the time horizon over which impacts are evaluated and may also vary with t . The impact measure I is most commonly expressed in physical units (e.g., radiative forcing, temperature, sea level rise, or precipitation change) but may also be in economic units (e.g., climate damages). The weighting function w is usually equal to one but may vary for metrics that use discounting, model output, or other methods to determine weights (Manne and Richels 2001; Edwards et al. 2016; Reilly and Richards 1993; Kandlikar 1996; Johansson 2012; Wigley 1998; Manning and Reisinger 2011).² (For the GWP, I in Eq. 2 is a combination of f and c in Eq. 1 and $w = 1$.)

The parameters in Eq. 2 can be varied to create many metric designs, including those previously discussed in the literature as well as new ones. These metrics may be static or time-dependent. A *static metric* assigns the same value to emissions regardless of when they occur. Static metrics define equivalency in various ways. For example, the GWP compares emissions based on their integrated radiative forcing impacts over a fixed duration (typically 100 years) (Stocker et al. 2013). The most commonly discussed alternative, the global temperature potential (GTP), is defined based on the relative temperature impacts of emissions a fixed length of time after they occur (Shine et al. 2005). Other static metrics have also been designed on the basis of various physical impacts, including radiative forcing, temperature (Gillett and Matthews 2010; Peters et al. 2011; Kirschbaum 2014), sea level rise (Stern et al. 2014), and precipitation (Shine et al. 2015), or based on economic factors such as the forecasted costs and benefits of reducing emissions of different gases (Schmalensee 1993; Reilly and Richards 1993; Deuber et al. 2013; Tol et al. 2012). Since the impacts of short-lived greenhouse gases such as CH_4 decay more quickly (relative to CO_2), metrics that emphasize impacts close to the time of emission assign a higher equivalency value to these gases.³

Other metrics allow the equivalency value assigned to gases to depend on the time of emission, resulting in a *time-dependent metric*. With physical metrics, this is accomplished by defining the year or range of years when impacts are evaluated relative to a particular time of interest, such as when a climate threshold is expected to be reached. These metrics compare emissions based on their impacts, typically radiative forcing or temperature, at a particular point in time (Edwards and Trancik 2014; Shine et al. 2007; Tanaka et al. 2013; Abernethy and Jackson 2022) or summed up to a point in time (Edwards and Trancik 2014). Other impact measures and weighting schemes can also be used. For example, with time-dependent economic metrics, emissions are typically compared based on the relative cost-effectiveness (or, less commonly, costs and benefits) of mitigation, given a simulation with a specified policy objective and a set of mitigation cost assumptions (Manne and

² Alternatively, the integration bounds could be defined from 0 to infinity, and the weight could be used to determine the time horizon by setting $w = 1$ for impact years within the time horizon and $w = 0$ for other impact years.

³ The opposite is generally true when comparing the climate impacts of very long-lived gases to CO_2 ; metrics that emphasize immediate impacts assign a lower equivalency value to these gases.

Richels 2001). For many simulations, it has been shown that these metrics may be well approximated using physical parameters and a discount rate (Johansson 2012).

Time-dependent metrics can take on a variety of shapes. When formulated around a climate threshold, both physical and economic metrics increase the value assigned to short-lived gases as the threshold approaches. After this time is reached, a metric may continue to emphasize the immediate impacts of emissions (Edwards and Trancik 2014; Shine et al. 2007; Abernethy and Jackson 2022) (taking on a constant value). We refer to metrics with this design as *increasing metrics*. Among time-dependent metrics, increasing metrics have received the most attention in previous literature. Increasing metrics in Table 1 include physical metrics such as the ICI, CCI, and DGTP, and economic target-based metrics such as the CETP and GCP. In contrast, a metric that places a high initial value on short-lived gases and decreases over time can also be used. These *decreasing metrics* (and static approximations) have been proposed as a way to closely follow a particular climate pathway (Wigley 1998; Tanaka et al. 2009; Manning and Reisinger 2011; Wigley 2021), although previous discussion of these metrics is limited.

Alternatively, a metric may increase initially but then relax its emphasis on the immediate impacts of emissions and evaluate impacts over increasingly long time periods. This results in a metric that decreases the value assigned to short-lived gases after the time of interest. We refer to these metrics, which increase as a policy target is approached and decrease afterwards, as *hybrid metrics*, because they combine increasing and decreasing shapes. The hybrid metric is a new proposal made in this study and addresses some of the motivations for both increasing and decreasing metrics.⁴ We hypothesize that hybrid metrics can enable decision-makers to reach a target level of climate change and remain at this level after the target is reached, without permanently assigning high penalties to short-lived gases. This feature may be especially useful for sectors where completely eliminating CH₄ emissions is infeasible (e.g., agriculture and small residual uses of fossil fuels).

Discounting, if applied, would affect the weighting function w in Eq. 2. Discounting in metrics can be used with the goal of guiding emissions decisions toward a set of reductions that are forecasted to be optimal, based on a simulated mitigation pathway. Such metrics necessarily rely on a set of assumptions about future changes in technology costs, market responses, or the costs of damages. Depending on the integrated assessment model used for the simulation, these assumptions may be represented more explicitly (e.g., as technology cost changes with time and production) or less explicitly as an economy-wide marginal abatement cost curve or damage function. In this paper we focus on a wider set of metric values (rather than solving for an optimal metric), recognizing the uncertainty in these assumptions. We ask whether conclusions can be drawn that cover many possible metric values.⁵

Various sources of uncertainty can affect the parametric and structural features of different metric designs (i.e., the parameter values and functions employed in specific realizations of Eq. 2). These uncertainties may be physical, including uncertainty in

⁴ A recent paper has also demonstrated that cost-effective metrics can take on an increasing-and-decreasing shape for pathways that include substantial overshoots of temperature targets (Tanaka et al. 2021), but it takes a different approach methodologically than that employed here, solving for an optimal metric as the output of a simulation of the cost-effective mitigation pathway under a set of assumptions about the future costs of mitigation.

⁵ Numerically, the effect of changing the time horizon in physical metrics such as the GWP can achieve similar effects on metric values to applying different discount rates to physical impacts (Sarofim and Giordano 2018; Mallapragada and Mignone 2020), as seen in Eq. 2.

the removal rates of greenhouse gases or the temperature response to radiative forcing (Reisinger et al. 2010; Olivié and Peters 2013). They can also be definitional, such as ambiguity in what background state to use when modeling the climate impacts of pulse emissions (Reisinger et al. 2011) and which indirect effects and feedbacks to include (Shindell et al. 2009; Gasser et al. 2017; Sterner and Johansson 2017). Uncertainty in impacts tends to increase along the causal chain from emissions to concentrations, radiative forcing, temperature and other physical climate changes, and ultimately climate damages (Fuglestedt et al. 2003; Boucher et al. 2009). Some metrics use external information to select the time horizon, such as the time when climate impacts are expected to peak under a climate policy (Edwards et al. 2016; Abernethy and Jackson 2022), or to determine the marginal cost of abating greenhouse gas emissions, optimal discount rate (and discount function), or cost of climate damages (Boucher 2012). These inputs are also subject to uncertainty.

2.1.2 Metric examples

We vary the design choices reviewed in Section 2.1.1 to calculate a set of equivalency metrics for consideration in this study. These examples are chosen to span the wide range of shapes and values in the equivalency metric design space, to enable us to draw generalizable conclusions about metric design choices and their consequences that apply across a range of decarbonization pathways. These metric examples can all be calculated by varying the time parameters and impact measures in Eq. 2 and setting the weights w to one. We consider a total of sixteen metrics as well as the GWP(100). These metrics include four shapes (static, increasing, decreasing, and hybrid), two treatments of time (instantaneous and cumulative), and two indicators of impact (radiative forcing and temperature change). We discuss additional metric designs in Section D of the Supplementary Information. We focus on metrics for CH_4 due to its shorter lifetime (compared to CO_2) and large contribution to the climate impacts of energy systems and other sectors of the economy.

First, we consider a static instantaneous metric (μ_{SI}) and static cumulative metric (μ_{SC}),

$$\mu_{\text{SI}}(\tau) = \frac{I_X(\tau, t_0)}{I_K(\tau, t_0)}$$

$$\mu_{\text{SC}}(\tau) = \frac{\int_{t_0}^{\tau} I_X(t', t_0) dt'}{\int_{t_0}^{\tau} I_K(t', t_0) dt'},$$

where t_0 is the present day, τ is the time horizon or time of interest (expressed in years, e.g., 2050), $I(a, b)$ is the impact at time a of a gas emitted at time b , X is a generic greenhouse gas (e.g., CH_4), and K is the reference gas (e.g., CO_2). Second, we consider an increasing instantaneous metric (μ_{II}) and increasing cumulative metric (μ_{IC}),

$$\mu_{\text{II}}(\tau, t) = \begin{cases} I_X(\tau, t)/I_K(\tau, t) & t < \tau \\ I_X/I_K, & t \geq \tau \end{cases}$$

$$\mu_{\text{IC}}(\tau, t) = \begin{cases} \int_t^{\tau} I_X(t', t) dt' / \int_t^{\tau} I_K(t', t) dt' & t < \tau \\ I_X/I_K, & t \geq \tau \end{cases}$$

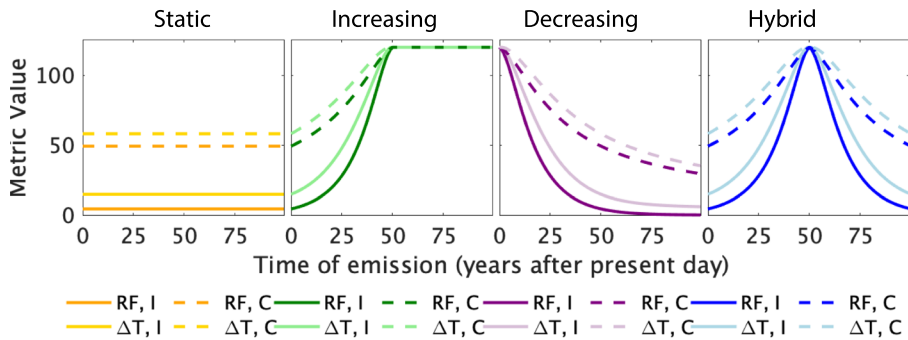


Fig. 1 Examples of metrics for comparing the climate impacts of CH₄ and CO₂ emissions, in units of grams CO₂-equivalent per gram CH₄. We present static, increasing, decreasing, and hybrid metric shapes that use cumulative (C, dashed lines) or instantaneous (I, solid lines) designs, calculated using radiative forcing (RF) or temperature change (ΔT) as the indicator of impact. These metrics cover the spectrum of weighting approaches in previously proposed designs (see Table 1) as well as new ones. Static, increasing, and hybrid metrics are calculated for a time of interest 50 years after the present day; decreasing metrics define the present day as the time of interest

where t is the time of emission, I_X/I_K is the ratio of the instantaneous impacts of gases X and K , and other variables are defined above. Third, we consider a decreasing instantaneous metric (μ_{DI}) and decreasing cumulative metric (μ_{DC}),

$$\mu_{DI}(t) = \begin{cases} I_X/I_K & t = t_0 \\ I_X(t, t_0)/I_K(t, t_0), & t > t_0 \end{cases}$$

$$\mu_{DC}(t) = \begin{cases} I_X/I_K & t = t_0 \\ \int_{t_0}^t I_X(t', t_0) dt' / \int_{t_0}^t I_K(t', t_0) dt'. & t > t_0 \end{cases}$$

Finally, we consider a hybrid instantaneous metric (μ_{HI}) and hybrid cumulative metric (μ_{HC}),

$$\mu_{HI}(\tau, t) = \begin{cases} I_X(\tau, t)/I_K(\tau, t) & t < \tau \\ I_X/I_K & t = \tau, \\ I_X(t, \tau)/I_K(t, \tau), & t > \tau \end{cases}$$

$$\mu_{HC}(\tau, t) = \begin{cases} \int_{\tau}^t I_X(t', t) dt' / \int_{\tau}^t I_K(t', t) dt' & t < \tau \\ I_X/I_K & t = \tau \\ \int_{\tau}^t I_X(t', \tau) dt' / \int_{\tau}^t I_K(t', \tau) dt'. & t > \tau \end{cases}$$

Figure 1 presents values for these eight metric design choices using both radiative forcing and temperature impacts and for an example time of interest τ that is 50 years after the present day. We describe the calculation of radiative forcing and temperature impacts I from emissions in Section 2.2 and Section A.4 of the Supplementary Information. The time of interest τ can also be selected based on the time when climate impacts are expected

to near a threshold as defined by a particular policy and set of mitigation scenarios. In Figs. 3 and 4, for example, τ is defined by the peak radiative forcing and temperature change expected under a 1.5 and 2 °C policy (see Section 2.2 and Section A.4 of the Supplementary Information for a discussion of these mitigation scenarios). (See also Edwards and Trancik (2014), Edwards et al. (2016), and Abernethy and Jackson (2022) for further discussion of threshold-based metrics.)

Temperature metrics take on higher values than radiative forcing metrics for a given time horizon (as seen in Fig. 1). However, radiative forcing also peaks earlier than temperature change for a given mitigation scenario. As a result, the time parameters for the radiative forcing metrics are shorter than those for the temperature metrics, when these parameters are calculated based on the timing of peak climate (radiative forcing or temperature) impact. Since these two effects partially cancel, the radiative forcing and temperature metrics are more similar in Figs. 3 and 4 (which use a time of interest of 2049 for radiative forcing and 2078 for temperature change under a 1.5 °C policy and 2059 and 2095, respectively, for 2 °C) than in Fig. 1 (which uses a single time of interest of 50 years after the present day).

Except for the role the choice of impact may play in selecting time parameters, variation in metric values due to the choice of impact measure is somewhat constrained (Peters et al. 2011; Johansson 2012; Shine et al. 2005). This is explained by the fact that, with physical metrics, the ratio of different greenhouse gas impact measures (e.g., temperature change and sea level rise) generally approaches the ratio of their radiative efficiencies as the time of evaluation t' approaches the time of emission t . Further from t , moving along the causal chain from radiative forcing to temperature change to sea level rise places increasingly more value on short-lived gases, because the latter processes have more inertia (Solomon et al. 2010; Sterner et al. 2014).⁶ Economic cost-effectiveness metrics typically approach a lower maximum value than purely physical metrics, but only because they include discounted impacts beyond the target year, a form of weighting (Tanaka et al. 2013). Metrics based on economic damages depend on the damage function, which is difficult to determine with confidence but is often estimated as a simple function of temperature change with economic discounting (Hammitt et al. 1996). Perhaps due to these large uncertainties, climate damage metrics are largely discussed from a theoretical perspective and are not typically proposed for real-world applications.

In contrast, the choice of time parameters t_1 and t_2 (and the weighting function w , which we do not vary in these example metrics) strongly influence metric values (Tanaka et al. 2010; Mallapragada and Mignone 2017; Peters et al. 2011). In fact, a large number of metric designs, including many metrics previously discussed in the literature as well as new ones, can be well represented by choosing a single impact measure and varying t_1 and t_2 . A metric that uses a constant time horizon will be static, whereas one that varies with the time of emission will be time-dependent (increasing, decreasing, or hybrid). For example, for increasing metrics, the time horizon shrinks as the time of emission approaches the time of interest (e.g., the climate impacts are expected to peak). Within each shape, instantaneous metrics evaluate impacts at a point in time, and cumulative metrics aggregate impacts up to that point in time. Instantaneous metrics generally assign lower impact values to short-lived gases such as CH_4 than cumulative metrics (except when impacts are evaluated at the time of interest, at which point they are equal). For time-dependent metrics, instantaneous

⁶ Precipitation changes are more complicated because they result from different counterbalancing impacts from radiative forcing and temperature (Shine et al. 2015).

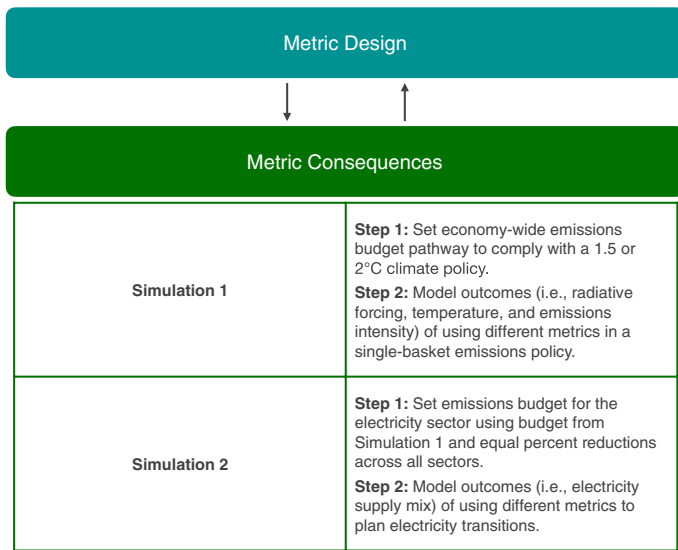


Fig. 2 Schematic overview of methods in this paper, including metric designs and two simulations for assessing metric consequences economy wide and in the electricity sector. Improved understanding of metric consequences can improve metric design (represented by the feedback arrow in the diagram)

metrics also have a larger rate of change than cumulative metrics, which take on higher values and change more slowly over time.

2.2 Simulations of metric consequences

We use two simulations to explore the consequences of equivalency metric design choices in terms of the potential effects of different metric formulations on energy system transitions and climate outcomes (see Fig. 2 for an overview). We focus on simulating the case of a single-basket emissions policy designed to limit global temperature change, in which a budget is set for all emissions (in CO₂-equivalent units) and a metric is used to establish exchange rates among different gases. This setup follows the approach introduced in the Montreal Protocol and later adopted in the Kyoto Protocol and also reflects the standard format for communicating Nationally Determined Contributions to the Paris Agreement (Daniel et al. 2012). We also arrive at several conclusions about the consequences of metric design choices that are more generally applicable to other policies, as we indicate in our discussion of the results and summarize in Box 1 in Section 4.

The first simulation (“Simulation 1,” Fig. 2) examines the effects of applying various metrics on the radiative forcing, temperature, and emissions intensity outcomes of a global, economy-wide policy constraining CO₂-equivalent emissions, while the second simulation (“Simulation 2,” Fig. 2) further resolves the underlying technology transitions involved in meeting an emissions constraint applied to the electricity sector. In the first simulation, we consider changes in emissions intensity from economy-wide consumption, while the second simulation focuses in greater detail on a particular sector

(electricity) in order to allow for an investigation of the potential range of underlying technology portfolios that could be used to reduce the emissions intensity of that sector (i.e., emissions per unit electricity generated). Both simulations are highly simplified and contain only as much detail as is needed to examine the potential consequences of different metric designs in terms of climate outcomes and technology transitions.

These two simulations extend a framework for testing equivalency metrics developed in Edwards et al. (2016) (which focuses on radiative forcing) to examine emissions policies designed to limit temperature change to 1.5 or 2 °C. The framework consists of two steps (see Fig. 2): (1) calculate a CO₂-equivalent emissions budget pathway that would meet the policy target if all emissions were CO₂ and (2) simulate technology or emissions choices to meet the emissions budget using a given metric. We perform these tests using various metrics (described in Section 2.1.2) to examine how the choice of metric influences the impacts of climate policies. Section A of the Supplementary Information discusses our methods for examining metric design consequences in further detail. We describe the process for calculating emissions budgets (Section A.1) and simulating the consequences of metric design globally (Section A.2) and in the electricity sector (Section A.3). We also discuss our simple climate model (Section A.4).

Step 1 selects a climate policy target and calculates a CO₂-equivalent emissions budget pathway to meet that target. We focus here on both 1.5 and 2 °C climate policies. To calculate annual emissions budgets for each of these policies, we use the GWP(100) to determine CO₂-equivalent emissions in the starting year and solve for the future reductions that, if all emissions were CO₂, would satisfy a 1.5 and 2 °C temperature constraint. Emissions budgets $e_B(t)$ evolve from initial levels e_0 according to a changing exponential growth rate $g(t')$ (see also Allen et al. (2009) and Section A.1 of the Supplementary Information),

$$e_B(t) = e_0 \exp \left[\int_{t_0}^t g(t') dt' \right]. \quad (3)$$

The result is a unique CO₂ emissions budget pathway that specifies annual allowed CO₂-equivalent emissions that, if they were entirely CO₂, would meet the temperature target (1.5 or 2 °C). Under our modeled scenarios, radiative forcing and temperature are expected to peak in 2049 and 2059, respectively, for the 1.5 °C policy and 2078 and 2095 for the 2 °C policy, and these peak years are used to define τ (see Section 2.1.2) for the metrics presented in Figs. 3 and 4.⁷

We use impulse response functions to estimate the impacts of greenhouse gas emissions on concentrations, radiative forcing, and temperature change. These functions take the general form

⁷ We note that many modeled scenarios for meeting a 1.5 °C end of century target involve a temporary overshoot of the temperature target (Tanaka and O'Neill 2018). Our emissions reduction scenarios are more ambitious than what would be needed in such a peak-and-decline scenario. We also do not simulate scenarios with net-negative greenhouse gas emissions. The approach we develop to define the budget pathway could be easily extended to cases where the temperature reaches 1.5 °C at the end of the century after a temporary overshoot or those with significant use of negative emissions technologies (e.g., by modifying Eq. 3 to allow for net negative, rather than net zero, emissions and updating Eq. 5 to include a term for CO₂ removals).

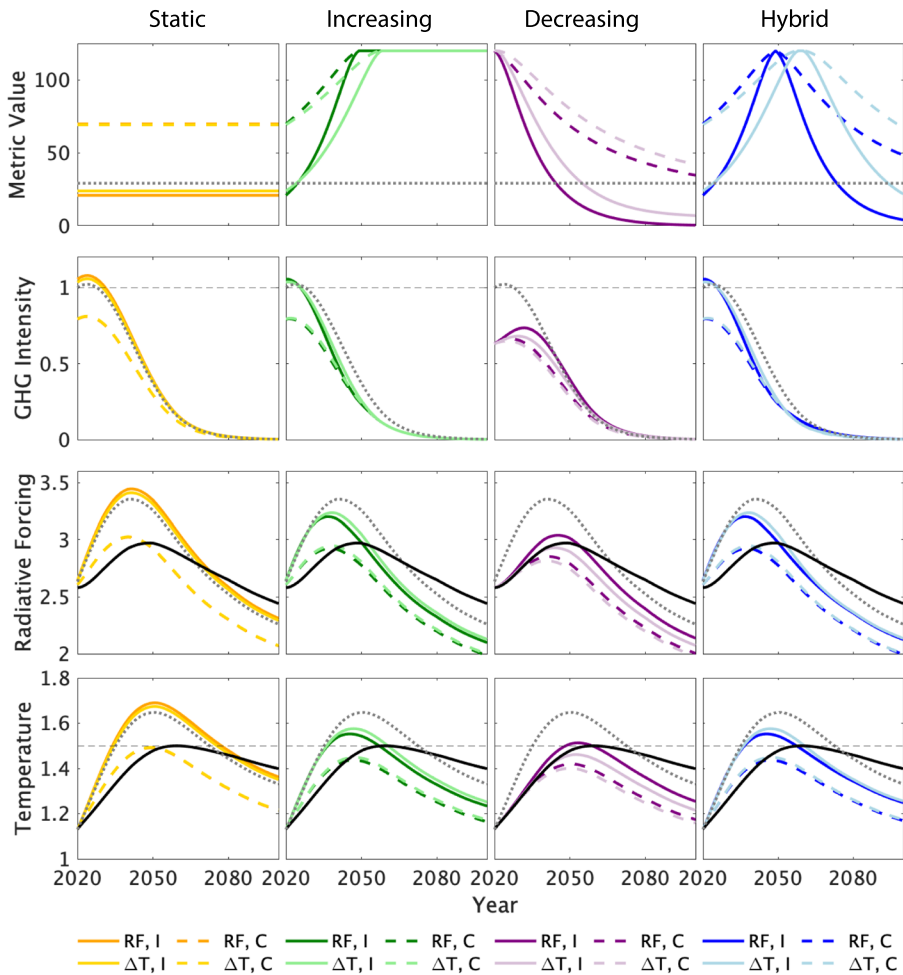


Fig. 3 Metric values (grams CO₂-equivalent per gram CH₄), emissions intensity relative to present day (unitless), radiative forcing (Wm⁻²), and temperature (°C) outcomes of alternative metrics designed around a 1.5 °C policy target. We present metrics with static, increasing, decreasing, and hybrid shapes; cumulative (C) and instantaneous (I) formulations; and radiative forcing (RF) and temperature (ΔT) impact measures. Dotted gray lines show results for the GWP(100), and solid black lines show the climate outcomes of the budget scenario if all emissions were CO₂. Dashed gray lines in the second row indicate present-day emissions intensity levels, and dashed gray lines in the bottom row show the 1.5 °C temperature threshold. CH₄ and CO₂ intensities (as well as emissions, since consumption is constant) follow the same trends as the greenhouse gas (GHG) emissions intensity since mass reduction percentages for both gases are equal

$$\text{IRF}_x(t) = a_{x,0} + \sum_{i=1}^n a_{x,i} \cdot \exp\left(\frac{-t}{\tau_{x,i}}\right), \quad (4)$$

where $\text{IRF}_x(t)$ is the impulse response function, $\tau_{x,i}$ are time scales, and $a_{x,i}$ are fractions whose sum is one. We use parameters from the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (Stocker et al. 2013) for concentration and radiative forcing and parameters from an ensemble model study for temperature change

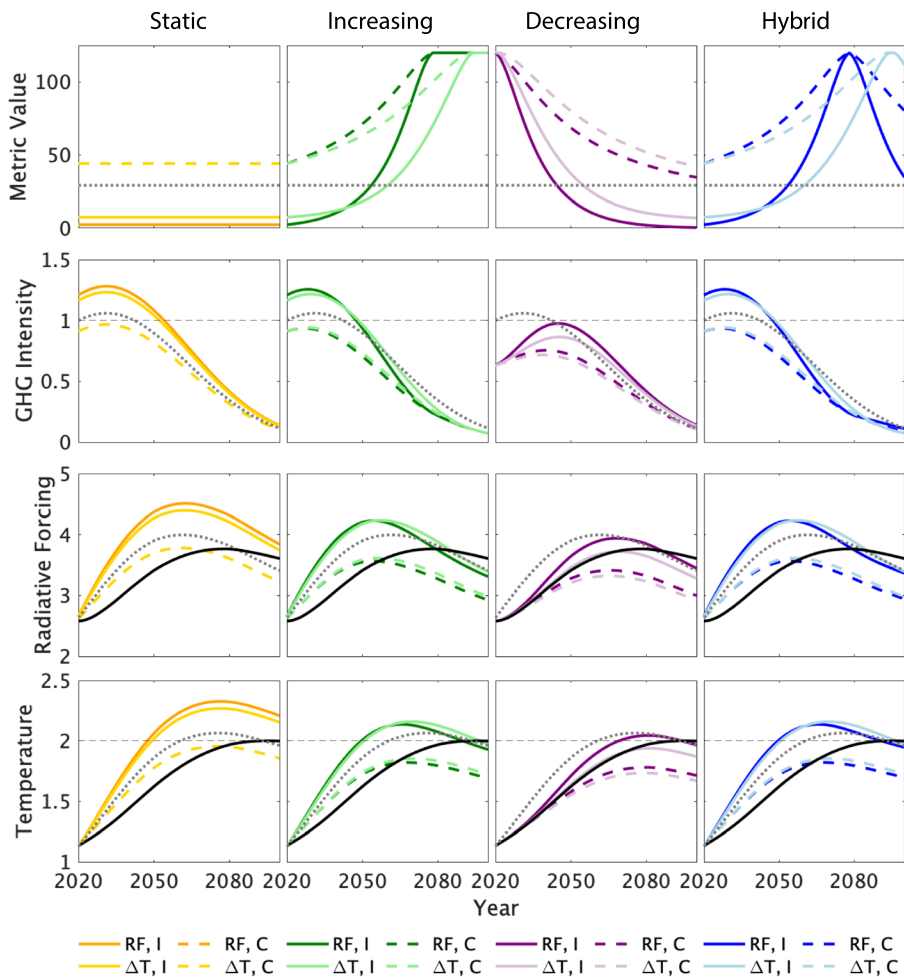


Fig. 4 Metric values (grams CO₂-equivalent per gram CH₄), emissions intensity relative to present day (unitless), radiative forcing (Wm⁻²), and temperature (°C) outcomes of alternative metrics designed around a 2 °C policy target. We present metrics with static, increasing, decreasing, and hybrid shapes; cumulative (C) and instantaneous (I) formulations; and radiative forcing (RF) and temperature (ΔT) impact measures. Dotted gray lines show results for the GWP(100), and solid black lines show the climate outcomes of the budget scenario if all emissions were CO₂. Dashed gray lines in the second row indicate present-day emissions intensity levels, and dashed gray lines in the bottom row show the 2 °C temperature threshold. CH₄ and CO₂ intensities (as well as emissions, since consumption is constant) follow the same trends as the greenhouse gas (GHG) emissions intensity since mass reduction percentages for both gases are equal

(Caldeira and Myhrvold 2013) (see Section A.2 of the Supplementary Information). We use these same functions in estimating both metric values and climate outcomes in our simulations, in order to examine the consequences of metric design choices without introducing variation due to differences in how the metrics and our model represent climate variables. The method we present here could be adapted using more detailed climate models or other representations of the climate response to emissions (Smith et al. 2018; Millar et al. 2017).

Step 2 simulates the use of different equivalency metrics to comply with the emissions budget designed above, both economy-wide and then in the electricity sector. For our economy-wide simulations, we focus on total global emissions. Our analysis focuses on CO₂ and CH₄ emissions, which together contribute over 90% of CO₂-equivalent emissions from well-mixed greenhouse gases using the GWP(100). We present a case in the main text where the emissions budget $e_B(t)$, defined by

$$e_B(t) = c(t) \cdot p(t) \cdot (e_{K0} + e_{M0} \cdot \mu(t)), \quad (5)$$

is met with equal percent mass reductions in CO₂ and CH₄, where e_{K0} and e_{M0} are the initial CO₂ and CH₄ emissions, $\mu(t)$ is the metric, $c(t)$ is consumption, and $p(t)$ is the emissions intensity of consumption (i.e., emissions per unit consumption). We study outcomes under different relative CO₂ and CH₄ reduction scenarios in Section C.2 of the Supplementary Information.

We solve for the value of $c(t) \cdot p(t)$ in each period to satisfy the emissions budget $e_B(t)$, given a chosen metric $\mu(t)$ and initial emissions e_{K0} and e_{M0} . Both $c(t)$ and $p(t)$ are normalized by the simulation starting year. The climate outcomes we examine are the resulting radiative forcing and temperature trends (see Step 1 and Section A.4 of the Supplementary Information for the method used to calculate climate impacts). We also examine the resulting emissions intensity pathway, which represents changes in the underlying supply-side and demand-side technology mix. Consumption is defined here to include the use of services that emit greenhouse gases across all sectors of the economy.⁸ We do not explicitly resolve the contributions to emissions reductions from changes in $c(t)$ and $p(t)$; a lower value for $c(t)$ shifts the feasible $p(t)$ value upwards and vice versa. For simplicity, we present results for changes in $p(t)$ for a case where $c(t)$ remains constant.

We also examine in greater detail how the choice of equivalency metric may affect technology transitions in the electricity sector. In this simulation, specific technologies and their emissions intensities are represented. We calculate an emissions budget for electricity globally by multiplying the economy-wide emissions budget by the fraction of current emissions that come from electricity generation. We then calculate changes in the electricity generation mix required to meet demand and to satisfy the emissions budget using one of three technology decision rules, selected to cover a wide range of possible decarbonization transitions:

1. *Natural Gas Bridge*, where the electricity mix transitions, in order of preference, from coal to natural gas and then to low-carbon energy sources (e.g., solar, wind, and nuclear) as needed to meet the emissions constraint,
2. *Low-Carbon Choices*, where it transitions directly from fossil fuel generation (first coal and then natural gas) to low-carbon energy sources, and
3. *Carbon Capture*, where it transitions from coal to natural gas and then to fossil fuels with carbon capture and storage (CCS), with a further transition to lower-emissions energy sources if required to avoid exceeding the emissions budget.

Electricity demand is assumed to grow at 1% per year (International Energy Agency 2019). We use present-day emissions intensity estimates for all technologies. Similar to the

⁸ This approach differs from the simulations presented in Edwards et al. (2016), which included two hypothetical “technologies” — one with higher CO₂ and lower CH₄ intensities and the other with higher CH₄ and lower CO₂ intensities — and solved for the technology mix that would maximize consumption given the emissions budget and chosen equivalency metric.

first simulation, changes to these assumptions would increase or decrease the rate of the transition but are not expected to change the qualitative trends reported. Natural gas, and to a lesser extent coal, results in significant life cycle CH_4 emissions (see Section A.3 of the Supplementary Information for emissions intensity values and Section C.3 for a sensitivity analysis for fossil fuel CH_4 intensities). CH_4 intensities may be reduced through policies targeting emissions throughout the supply chain, which would tend to reduce differences across metric outcomes. If the emissions budget cannot be met with given residual emissions from low-carbon technologies, negative emissions from bioenergy with carbon capture and storage (BECCS) are introduced in later years.

Our simulations do not involve solving for the cost-optimal decarbonization transition as many previous studies do when examining the economic costs of different metrics or identifying optimal metric values (e.g., Godol and Fuglestad 2002; Johansson et al. 2006; Streffler et al. 2014; Brennan and Zaitchik 2013; Harmsen et al. 2016; Ekholm et al. 2013; Reisinger et al. 2013). Such studies typically use integrated assessment models and solve for optimal supply-side changes such as the electricity mix and demand-side changes affecting consumption. These models necessarily involve forecasts of technology costs and the response of the economy to an emissions constraint or the effects of climate change. However, there is significant uncertainty, for example around the evolution of technology and other economic responses to climate policy. Given the large uncertainties involved, we focus instead on examining a range of possible decarbonization transitions and draw qualitative conclusions about metric consequences based on results from a range of transition scenarios.

The modeling approach presented here is designed to provide insight for a variety of policy contexts, including country- or sector-specific policies. Our results apply most directly to policies that limit long-term climate change by applying a CO_2 -equivalent emissions budget designed around a physical climate target (e.g., a temperature threshold). This budget can be designed to keep long-term climate changes below an upper bound that is defined by the case in which all emissions in the CO_2 -equivalent budget are CO_2 . If a policy is designed without reference to a physical climate target limiting long-term climate change (and thus CO_2 emissions), using equivalency metrics that assign a high impact value to CH_4 and other short-lived greenhouse gases can lead to substantially greater long-term temperature changes because they permit higher CO_2 emissions. We return to this important issue in the discussion section.

3 Metric design consequences

3.1 Emissions intensity and climate change

We examine outcomes from our economy-wide simulation (Simulation 1) for a set of sixteen equivalency metrics, representing a range of possible design choices (see Section 2), as well as the GWP(100). These metric designs include four shapes (static, increasing, decreasing, and hybrid), two formulations (instantaneous and cumulative), and two impact measures (radiative forcing and temperature change). The time parameters are based on the year when climate impacts reach their maximum value under the budget scenario (where all emissions are CO_2). Radiative forcing and temperature

peak in 2049 and 2059, respectively, for the 1.5 °C policy and 2078 and 2095 for the 2 °C policy. The resulting metric values are presented in Figs. 3 and 4.

Climate outcomes vary across the four metric shapes and between cumulative and instantaneous formulations (see Figs. 3 and 4, third and fourth rows). Radiative forcing and temperature change can temporarily exceed policy targets by a large margin for some metric designs. Static instantaneous metrics lead to the largest overshoots. The GWP(100) leads to large overshoots for the 1.5 °C policy but somewhat lower overshoots for the 2 °C policy, since the timeline for reaching this higher policy target is longer (and closer to 100 years). Decreasing metrics most closely follow the intended pathway of climate impacts in early years. Increasing and hybrid metrics lead to more rapid rates of climate change early on but limit peak climate impacts. Overshoots of intended maximum impact levels are higher for radiative forcing than for temperature change. Static instantaneous metrics exceed radiative forcing targets by approximately 16% for a 1.5 °C policy and 20% for a 2 °C policy, whereas they exceed temperature targets by 11% and 13%, respectively. Overshoots for increasing and hybrid instantaneous metrics are comparatively smaller, reaching 8% and 12% of radiative forcing for a 1.5 and 2 °C policy, respectively, and 5% and 8% of temperature change.

Metric design choices also carry consequences for the technology changes required to comply with an emissions policy. We estimate CO₂ and CH₄ emissions intensity over time (in units of mass per unit consumption), relative to emissions intensity in the first year of the simulation (see Figs. 3 and 4, second row). Requirements vary significantly across metric designs. Instantaneous increasing and static metrics all lead to values close to present levels early on, whereas decreasing metrics (and to a lesser extent all cumulative metrics) lead to significantly lower values. This suggests that higher metric values, while they constrain rates of climate change, can require rapid near-term technology changes when they are applied without a ramp-up period. This is especially true in the case of decreasing metrics, which can require an immediate step change decrease in emissions intensities. A large reduction in emissions intensity is required across all metrics in later years. However, metrics that place a lower value on CH₄ in later years, including decreasing and hybrid metrics, permit higher long-term emissions intensities. While these differences appear small compared to current levels, they can be significant as a fraction of future (lower) emissions intensities.

Since our baseline scenario models equal percent mass reductions in CO₂ and CH₄, the relative emissions intensity pathways are the same for both gases. Real-world deviations from this assumption will depend on the availability and attractiveness of climate change mitigating technologies, as well as the design of the climate policy. Some emissions reductions may be achieved with improvements in current technologies (for example, fixing leaks in natural gas systems (Edwards et al. 2021)), while others will be due to changes in the portfolio of supply-side technologies (for example, a transition from natural gas to low-carbon energy sources). Reductions may also be traded across different sectors of the economy with different baseline CO₂ and CH₄ emissions. If CO₂ reductions are prioritized over CH₄, the relative CH₄ intensity of the economy increases, and the differences in the outcomes across metrics are greater (see Section C.2 of the Supplementary Information). Or, if CH₄ reductions are sufficiently prioritized over CO₂, a policy risks being less effective at limiting long-term climate change (Edwards and Trancik 2014). Policies that stipulate CO₂-specific reduction targets or cumulative CO₂ budgets can mitigate these risks by ensuring that CO₂ reductions are equal to or greater than these limits.

Our analysis examines a situation where an emissions policy is first determined, and later one of a set of possible metrics is selected to establish equivalency between greenhouse gases. However, emissions policies may be set in other ways, with implications for metric consequences. If emissions budgets are designed to meet a climate policy threshold, or calculated using an exogenously-determined percent reduction target, changing the calculation of initial CO₂-equivalent emissions will also impact the trajectory of the emissions budget. The implications of using different metrics to calculate emissions budgets can be counterintuitive. Consider as an example a climate policy framed as a percentage reduction in CO₂-equivalent emissions relative to a base year. If the chosen metric is static, and if the emissions constraint is met by equal percent reductions in CO₂ and CH₄ emissions, the policy would lead to the same outcome regardless of the metric value. However, outcomes differ if the metric is time-dependent or if the choice of metric changes the relative mix of CO₂ and CH₄ reductions. A time-dependent metric with a low value in the base year and high values in later years will result in the lowest allowed emissions overall and the most stringent restrictions on CH₄ emissions. As a result, an instantaneous increasing metric leads to the smallest overall emissions under an exogenous percent reduction policy (see Section C.1 of the Supplementary Information for further discussion).

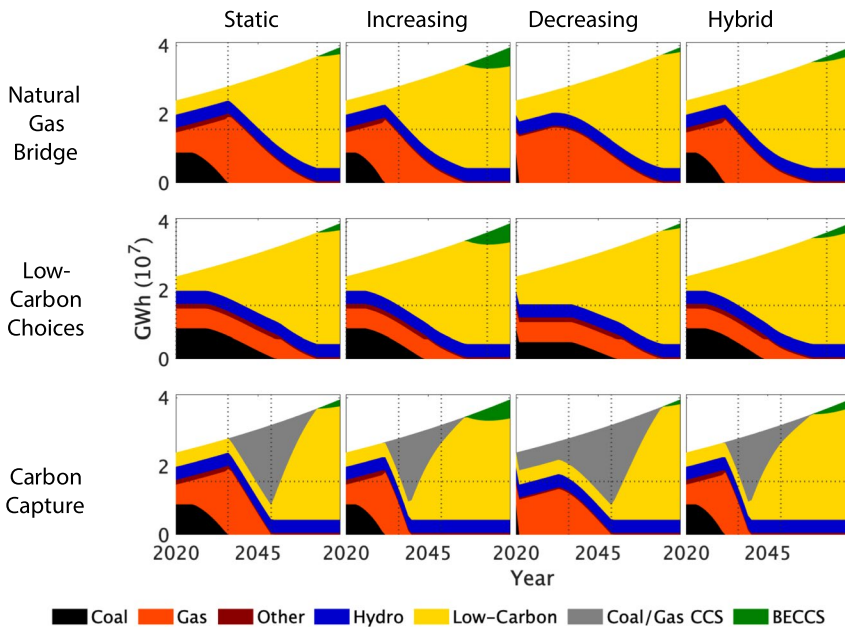


Fig. 5 Electricity generation mix for a 1.5 °C policy using static, increasing, decreasing, and hybrid instantaneous radiative forcing metrics, under three scenarios: *Natural Gas Bridge*, *Low-Carbon Choices*, and *Carbon Capture*. Differences across metrics are initially more pronounced for scenarios that emphasize high-CH₄-emitting fossil fuel generation (with or without CCS). Dashed horizontal lines indicate the initial level of fossil fuel generation, which immediately and significantly changes when applying decreasing metrics. Dashed vertical lines indicate the timing of, first, peak fossil fuel generation and, second, complete phaseout of unabated fossil fuel use using static metrics. This timing is similar (with the exception of the initial drop in fossil fuel generation) for decreasing metrics and earlier for increasing and hybrid metrics. Note that the axes in this figure are shorter (2020–2070) than in Figs. 3 and 4 (2020–2100)

3.2 Electricity system transitions

We further examine the consequences of different metrics on technology transitions by focusing in on the electricity sector (Simulation 2). The purpose of this simulation is to study the technological changes that may be required to support reductions in emissions intensity, such as those shown in Figs. 3 and 4 (labeled “GHG intensity”), under different metrics. In this simulation, we allocate a portion of the overall emissions budget pathway to the electricity sector under a 1.5 and 2 °C climate policy as described in Section 2.2. We examine a range of possible supply-side technology mixes to draw conclusions about metric differences that are robust to uncertainty about which technologies will ultimately be favored in a decarbonization transition. Specifically, we present electricity mixes over the next 50 years under three technology decision rules and for instantaneous static, increasing, decreasing, and hybrid radiative forcing metrics (see Section B of the Supplementary Information for cumulative metrics).

Significant changes to the electricity generation mix are required across all metric designs and technology scenarios, but certain key differences in metric consequences are apparent (see Figs. 5 and 6). Decreasing metrics, which place a large emphasis on

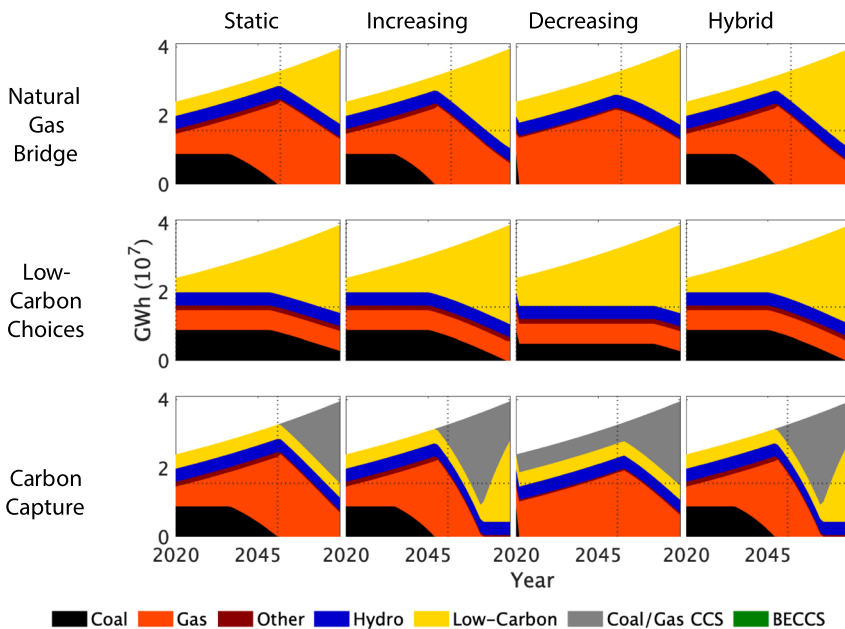


Fig. 6 Electricity generation mix for a 2 °C policy using static, increasing, decreasing, and hybrid instantaneous radiative forcing metrics, under three scenarios: *Natural Gas Bridge*, *Low-Carbon Choices*, and *Carbon Capture*. Differences across metrics are initially more pronounced for scenarios that emphasize high-CH₄-emitting fossil fuel generation (with or without CCS). Dashed horizontal lines indicate the initial level of fossil fuel generation, which immediately and significantly changes when applying decreasing metrics. Dashed vertical lines indicate the timing of peak fossil fuel generation using static metrics. This timing is similar (with the exception of the initial drop in fossil fuel generation) for decreasing metrics and earlier for increasing and hybrid metrics. (Complete phaseout of unabated fossil fuel use does not occur within the time frame shown in this figure.) Note that the axes in this figure are shorter (2020–2070) than in Figs. 3 and 4 (2020–2100)

CH₄ emissions today, require an immediate 45% reduction in coal generation, whereas transitions under static, increasing, and hybrid metrics are less abrupt initially. However, phaseout of unabated natural gas use occurs earlier for increasing and hybrid metrics than for static or decreasing metrics under a 1.5 °C policy (by 5–10 years). (Complete phaseout of unabated fossil fuel use does not occur for a 2 °C policy within the time frame considered.) Static and decreasing metrics also permit more prolonged use of fossil fuels with CCS. Increasing and hybrid metrics lead to similar generation mixes since their values do not begin to diverge until after the metric time horizon; however, they do lead to differences in required BECCS generation toward the end of the simulation period under a 1.5 °C policy. In general, metrics that place a high impact value on CH₄ in later years also result in larger deployment of BECCS.

The differences across metrics are either muted or amplified depending on the electricity supply mix. These differences are more significant for scenarios that emphasize energy sources with high life cycle CH₄ emissions (e.g., natural gas and fossil fuels with CCS). These include the *Natural Gas Bridge* and *Carbon Capture* scenarios. The *Low-Carbon Choices* scenario, which transitions directly from the current generation mix to energy sources that are both low-CO₂ and low-CH₄, depends less strongly on the choice of metric, since the CH₄ intensity of these technologies is lower. Differences across metrics also depend on the climate policy and are more significant for ambitious policies (e.g., the 1.5 °C policy considered here). The above observations are generally consistent with the trends in economy-wide emissions intensity in Figs. 3 and 4 but provide further insight into how those trends may be reflected in technology changes in a particular sector.

4 Conclusion

Over the past thirty plus years, a substantial literature has emerged on equivalency metric design, yet the potential consequences of choosing one metric over another in climate policy have not yet been comprehensively characterized. Many proposed metrics are inspired by reasonable objectives, yet their ability to achieve them has not been probed extensively. Perhaps as a result of the limited information available on the benefits or drawbacks of different metrics, alternative metrics have not seen significant uptake. However, more recently they have been discussed and applied in a small but growing set of engineering and policy analyses (e.g., Levasseur et al. 2016). The consequences, and even the motivation, behind using an alternative metric in these cases is not always stated. It may be a desire to use a different impact measure, for example by replacing radiative forcing with temperature change, or to emphasize a particular problem of interest, for example by using a shorter time horizon to highlight the near-term impacts of CH₄ emissions. By understanding how design choices change metric values, and the consequences of these values in use, decision-makers can make more informed metric choices that are linked to climate policy objectives.

In this paper we aim to contribute new understanding on the consequences of different metric designs. We examine in detail the impacts of metric design for a single-basket emissions policy, while also discussing the implications in other policy contexts. When used to evaluate short-lived greenhouse gases, we find that cumulative metrics lead to lower overshoots of climate policy thresholds, compared to instantaneous metrics with the same formulation and time horizon. Static, instantaneous metrics lead to

the largest overshoots. Increasing, decreasing, and hybrid instantaneous metrics reduce these overshoots. Decreasing metrics can theoretically reduce rates of climate change, relative to other shapes, by imposing immediate and dramatic emissions cuts in early years which may be difficult to implement in practice. These differences across metrics are magnified under more ambitious climate policies and if technologies with high CH₄ intensities (i.e., coal and natural gas with or without CCS) play a large role in decarbonization transitions.

Box 1. Consequences of metric design when used to assess short-lived gases (e.g., CH₄) under a climate policy.

1. Metrics that assign higher values to short-lived gases such as CH₄ can lead to larger reductions in CH₄ and CO₂ under a single-basket emissions policy. High metric values can be used to reduce rates of climate change and overshoots of climate thresholds but can require more rapid technology transitions.
2. The shapes of metrics (i.e., changes to metric values over time) also determine metric consequences:
 - Decreasing metrics can theoretically eliminate initial deviations from an intended climate pathway (i.e., the pathway used in designing the emissions policy) but place significantly greater immediate penalties on technologies with high CH₄ emissions that may not be possible to implement in practice.
 - Increasing metrics can lead to higher near-term impacts (compared to the climate pathway intended in designing the emissions policy) but can be designed to limit overshoots of policy targets (e.g., radiative forcing or temperature thresholds).
 - Hybrid metrics can balance reducing the risks of overshoots and providing time for technology transitions in the near term while avoiding over-penalizing CH₄ emissions in the long term.
 - Simpler, static metrics can still address some of the concerns about the GWP(100). For example, a GWP metric with a time horizon based on a climate policy threshold (i.e., a static cumulative radiative forcing metric) decreases overshoots of climate policy thresholds, compared to the GWP(100), but retains its simplicity.
3. The consequences of different metric designs can depend on the stringency of the climate policy and availability and attractiveness of different energy technologies. The choice of metric is particularly consequential under an ambitious climate policy and if technologies with high CH₄ intensities (i.e., coal and natural gas with or without CCS) play a large role in energy systems.

Metric choices also affect changes to the technology mix that are required by a mitigation scenario, as we show for the example of the electricity sector. Overall, metrics that place a high impact value on CH₄ early on require an accelerated phaseout of coal and

natural gas, whereas those with a high value later on place limits on sustained use of fossil fuels with CCS at current CH₄ intensities of fossil fuel infrastructure. For metrics that increase the impact value assigned to CH₄ over time, deploying new CH₄-intensive technologies (e.g., CCS) may not be economical, as these assets may be stranded ahead of their design lifetime to meet emissions targets (Edwards et al. 2022).

Information about metric consequences can guide metric design, by identifying formulations that are suitable for different contexts. Static integrated metrics are simple and can perform well when the time horizon is selected based on climate policy goals. They may be well-suited for cases where simplicity is favored over more precisely regulating impacts. Increasing and decreasing metrics each have different contexts where they may be appropriate. Increasing metrics allow more time for technology transitions because they result in less stringent emissions intensity targets today, while limiting overshoots of policy thresholds. Decreasing metrics are effective at limiting initial rates of climate change, and permit less aggressive long-term mitigation, but can require rapid near-term technology change (and thus might be applied in end-use sectors where immediate emissions cuts are technically feasible and economical). Hybrid metrics, which increase initially and later decrease, combine some of the attractive features of both increasing and decreasing metrics. The approach presented here for selecting metrics based on their consequences can also be applied to other metrics and evaluation criteria.

Our simulations focus on the impacts of metric design for a single-basket emissions policy, but we also arrive at more generalizable insights that can inform decision-making on the selection of metrics across a variety of policy contexts. A user of our results may observe, for instance, that a decreasing metric emphasizes the immediate impacts of emissions and thus leads to lower rates of temperature change, but that it requires significantly more rapid technology changes. With these types of generalizable insights, different user communities can better choose between decreasing metrics and other metric shapes in a variety of decision contexts, such as a sectoral or economy-wide emissions policy, a clean energy standard, or in technology evaluation and investment decisions. For example, a policymaker concerned about limiting rates of warming might prefer a metric that rapidly approaches a peak value that captures the immediate impacts of CH₄ emissions, whereas one concerned about managing retirement of existing fossil fuel fleets while meeting long-term temperature targets might prefer a more slowly increasing metric.

Our results also help demonstrate how the implications of the choice of metric can depend on the design of the emissions policy. We design our emissions budgets such that they would meet a given temperature threshold if all emissions were CO₂. This approach is conservative in the sense that it limits long-term temperature changes, which are driven by cumulative CO₂ emissions (Matthews et al. 2009). Other approaches to designing emissions budgets may lead to a potential for higher long-term temperature change. One way to address this challenge in emissions markets is via a “do no harm” rule that uses different metrics for offsetting CO₂ and CH₄ emissions (Allen et al. 2021). Interactions between the budget design and metric selection can also lead to initially counterintuitive results, especially if metrics are time-dependent. For example, a policy that defines emissions limits based on a percent reduction relative to a base year would lead to the greatest emissions reductions if there is a large difference between initial metric values (when the base year is defined) and later metric values (when emissions must be reduced) (Klemun and Trancik 2020).

The choice of metric is important across sectors and scales. Our results focus on global impacts because climate change depends on the total emissions resulting from many decentralized technology and policy decisions. However, metric values also impact the

distribution of mitigation requirements (and therefore costs) across countries and communities, by changing the relative emphasis placed on short- versus long-lived greenhouse gas reductions. Similarly, the choice of metric may influence the pace of mitigation across energy, industry, and other sectors in an economy-wide policy. For example, food systems (and especially livestock; Gerber et al. 2013) have high life cycle CH₄ emissions and comprise a significant fraction of total CO₂-equivalent emissions (Crippa et al. 2021). Metrics that place a high impact value on CH₄ both penalize actions that emit high levels of CH₄ today and provide greater rewards for efforts to mitigate these emissions in the future. How the choice of metric affects CH₄-intensive activities therefore depends on the level of CH₄ emissions as well as the ease of mitigating these emissions. Different actors may favor particular metrics based on their assessment of the expected mitigation costs they imply. Challenges in reducing CH₄ emissions also vary across sectors and actors, and prospects for adopting new metrics in these contexts may differ.

The choice of metric is important today, since many conventional technologies have high life cycle CH₄ intensities, and this choice may become increasingly important if the global economy transitions to technologies with higher life cycle CH₄ intensities and lower CO₂ intensities, or if policies focus heavily on CO₂ reductions (Klemun and Trancik 2020). Energy systems in particular represent over 70% of total emissions, and they pose a central challenge for climate change mitigation where many technologies (especially natural gas) have high CH₄ intensities that may be difficult to reduce (International Energy Agency 2019; Edwards et al. 2021). Accelerating climate impacts mean there may be less time for course correction (Masson-Delmotte et al 2018), making it even more urgent to revisit the choice of metric now. With an understanding of their consequences, decision-makers can choose metrics that are more aligned with societal goals when applied at a variety of scales, from technology- and sector-specific to economy-wide policies (Roy et al. 2015; Edwards et al. 2016; Edwards et al. 2017).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10584-022-03442-8>.

Acknowledgements The authors thank John Deutch for discussions on hybrid metrics and related topics.

Author contribution MRE and JET developed the concept and designed the methods for this study. MRE developed and performed the modeling. MRE and JET analyzed the results and wrote the paper.

Funding Open Access funding was provided by the MIT Libraries. This work received funding from the MIT Environmental Solutions Initiative and the National Science Foundation Graduate Research Fellowship Program (Award #1122374).

Data and code availability Data and code are available on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abernethy S, Jackson RB (2022) Global temperature goals should determine the time horizons for greenhouse gas emission metrics. *Environ Res Lett* 17:024019
- Allen MR et al (2009) Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature* 458:1163–1166
- Allen MR et al (2016) New use of global warming potentials to compare cumulative and short-lived climate pollutants. *Nature Clim Change* 6:773
- Allen MR et al (2018) A solution to the misrepresentations of CO₂-equivalent emissions of short-lived climate pollutants under ambitious mitigation. *NPJ Clim Atmos Sci* 1:1–8
- Allen MR et al (2021) Ensuring that offsets and other internationally transferred mitigation outcomes contribute effectively to limiting global warming. *Environ Res Lett* 16:074009
- Alvarez RA, Pacala SW, Winebrake JJ, Chameides WL, Hamburg SP (2012) Greater focus needed on methane leakage from natural gas infrastructure. *Proc Natl Acad Sci USA* 109:6435–6440
- Alvarez RA, et al. (2018) Assessment of methane emissions from the US oil and gas supply chain. *Science* 361:186–188
- Boucher O (2012) Comparison of physically- and economically-based CO₂-equivalencies for methane. *Earth Syst Dynam* 3:49–61
- Boucher O, Friedlingstein P, Collins B, Shine KP (2009) The indirect global warming potential and global temperature change potential due to methane oxidation. *Environ Res Lett* 4:044007
- Boucher O, Reddy MS (2008) Climate trade-off between black carbon and carbon dioxide emissions. *Energy Policy* 36:193–200
- Brandt AR, et al. (2014) Methane leaks from North American natural gas systems. *Science* 343:733–735
- Brennan ME, Zaitchik BF (2013) On the potential for alternative greenhouse gas equivalence metrics to influence sectoral mitigation patterns. *Environ Res Lett* 8:014033
- Cain M, et al. (2019) Improved calculation of warming-equivalent emissions for short-lived climate pollutants. *NPJ Clim Atmos Sci* 2:1–7
- Caldeira K, Myhrvold N (2013) Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration. *Environ Res Lett* 8:034039
- Cherubini F, et al. (2016) Bridging the gap between impact assessment methods and climate science. *Environ Sci Policy* 64:129–140
- Collins WJ, Frame DJ, Fuglestedt JS, Shine KP (2020) Stable climate metrics for emissions of short and long-lived species—combining steps and pulses. *Environ Res Lett* 15:024018
- Crippa M, et al. (2021) Food systems are responsible for a third of global anthropogenic GHG emissions. *Nature Food* 2:198–209
- Daniel JS, et al. (2012) Limitations of single-basket trading: Lessons from the Montreal Protocol for climate policy. *Clim Change* 111:241–248
- Deuber O, Luderer G, Edenhofer O (2013) Physico-economic evaluation of climate metrics: a conceptual framework. *Environ Sci Policy* 9:37–45
- Eckaus RS (1992) Comparing the effects of greenhouse gas emissions on global warming. *Energy J* 13:25–35
- Edwards MR, McNerney J, Trancik JE (2016) Testing emissions equivalency metrics against climate policy goals. *Environ Sci Policy* 66:191–198
- Edwards MR, Trancik JE (2014) Climate impacts of energy technologies depend on emissions timing. *Nat Clim Change* 4:347–352
- Edwards MR et al (2017) Vehicle emissions of short-lived and long-lived climate forcers: trends and trade-offs. *Faraday Discuss* 200:453–474
- Edwards MR et al (2021) Repair failures call for new policies to tackle leaky natural gas distribution systems. *Environ Sci Technol* 55:6561–6570
- Edwards MR et al (2022) Quantifying the regional stranded asset risks from new coal plants under 1.5°C. *Environ Res Lett* 17:024029
- Eklholm T, Lindroos TJ, Savolainen I (2013) Robustness of climate metrics under climate policy ambiguity. *Environ Sci Policy* 31:44–52
- Fesenfeld LP, Schmidt TS, Schrodle A (2018) Climate policy for short- and long-lived pollutants. *Nat Clim Change* 1:933–936
- Frank ED, Han J, Palou-Rivera I, Elgowainy A, Wang MQ (2012) Methane and nitrous oxide emissions effect the life-cycle analysis of algal biofuels. *Environ Res Lett* 7:014030
- Fuglestedt JS, Berntsen TK, Godal O, Skodvin T (2000) Climate implications of GWP-based reductions in greenhouse gas emissions. *Geophys Res Lett* 27:409–412

- Fuglestedt JS, et al. (2003) Metrics of climate change: Assessing radiative forcing and emission indices. *Clim Change* 58:267–331
- Fuglestedt JS, et al. (2009) Transport impacts on atmosphere and climate: Metrics. *Atmos Environ* 44:4648–4677
- Fuglestedt J, et al. (2018) Implications of possible interpretations of ‘greenhouse gas balance’ in the Paris Agreement. *Philos Trans Royal Soc A* 376:20160445
- Gasser T, et al. (2017) Accounting for the climate-carbon feedback in emission metrics. *Earth Syst Dynam* 8:235–253
- Gerber PJ, et al. (2013) Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities. Food and Agriculture Organization of the United Nations (FAO), Rome
- Gillett NP, Matthews HD (2010) Accounting for carbon cycle feedbacks in a comparison of the global warming effects of greenhouse gases. *Environ Res Lett* 5:034011
- Godol O, Fuglestedt J (2002) Testing 100-year global warming potentials: Impacts on compliance costs and abatement profile. *Clim Change* 52:93–127
- Hammit JK, Jain AK, Adams JL, Wuebbles DJ (1996) A welfare-based index for assessing environmental effects of greenhouse-gas emissions. *Nature* 381:301–303
- Harmen MJHM, et al. (2016) How climate metrics affect global mitigation strategies and costs: a multi-model study. *Clim Change* 136:203–216
- Harvey LDD (1993) A guide to global warming potentials (GWPs). *Energy Policy* 21:24–34
- International Energy Agency (2019) World Energy Investment 2019. International Energy Agency, Paris
- Johansson DJA (2012) Economics- and physical-based metrics for comparing greenhouse gases. *Clim Change* 110:123–141
- Johansson DJA, Persson UM, Azar C (2006) The cost of using global warming potentials: Analysing the trade off between CO₂, CH₄ and N₂O. *Clim Change* 77:291–309
- Kandlikar M (1996) Indices for comparing greenhouse gas emissions: Integrating science and economics. *Energy Econ* 18:265–281
- Kirschbaum MUF (2014) Climate-change impact potentials as an alternative to global warming potentials. *Environ Res Lett* 9:034014
- Klemun MM, Trancik JE (2020) Timelines for mitigating methane impacts of using natural gas for carbon dioxide abatement. *Environ Res Lett* 14:124069
- Lashof DA, Ahuja DR (1990) Relative contributions of greenhouse gas emissions to global warming. *Nature* 344:529–531
- Levasseur A, et al. (2016) Enhancing life cycle impact assessment from climate science: Review of recent findings and recommendations for application to LCA. *Ecol Indic* 71:163–174
- Lynch J, Cain M, Pierrehumbert R, Allen M (2020) Demonstrating GWP*: a means of reporting warming-equivalent emissions that captures the contrasting impacts of short-and long-lived climate pollutants. *Environ Res Lett* 15:044023
- Mallapragada D, Mignone BKA (2017) Consistent conceptual framework for applying climate metrics in technology life cycle assessment. *Environ Res Lett* 12:074022
- Mallapragada DS, Mignone BK (2020) A theoretical basis for the equivalence between physical and economic climate metrics and implications for the choice of global warming potential time horizon. *Clim Chang* 158:107–124
- Manne AS, Richels RG (2001) An alternative approach to establishing trade-offs among greenhouse gases. *Nature* 410:675–677
- Manning M, Reisinger A (2011) Broader perspectives for comparing different greenhouse gases. *Philos Trans R Soc Lond A* 369:1891–1905
- Masson-Delmotte V et al (eds) (2018) Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development and efforts to eradicate poverty. Intergovernmental Panel on Climate Change, Geneva
- Matthews HD, Gillett NP, Stott PA, Zickfeld K (2009) The proportionality of global warming to cumulative carbon emissions. *Nature* 459:829–832
- McJeon H et al (2014) Limited impact on decadal-scale climate change from increased use of natural gas. *Nature* 514:482–485
- Millar RJ, Nicholls ZR, Friedlingstein P, Allen MR (2017) A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions. *Atmos Chem Phys* 17:7213–7228
- O’Neill BC (2000) The jury is still out on global warming potentials. *Clim Change* 44:427–443
- O’Neill BC (2003) Economics, natural science, and the costs of global warming potentials. *Clim Change* 58:251–260
- Ocko IB, et al. (2017) Unmask temporal trade-offs in climate policy debates. *Science* 356:492–493

- Olivié DJL, Peters GP (2013) Variation in emissions metrics due to variation in CO₂ and temperature impulse response functions. *Earth Syst Dynam* 4:267–286
- Peters GP, Aamaas B, Berntsen T, Fuglestad JS (2011) The integrated global temperature change potential (iGTP) and relationships between emission metrics. *Environ Res Lett* 6:044021
- Plattner G-K, Stocker T, Midgley P, Tignor M (eds) (2009) IPCC Expert Meeting on the Science of Alternative Metrics. Intergovernmental Panel on Climate Change, Geneva
- Reilly JM, Richards KR (1993) Climate change damage and the trace gas index issue. *Environ Resour Econ* 3:41–61
- Reisinger A, Meinshausen M, Manning M (2011) Future changes in global warming potentials under representative concentration pathways. *Environ Res Lett* 6:024020
- Reisinger A, Meinshausen M, Manning M, Bodeker G (2010) Uncertainties of global warming metrics: CO₂ and CH₄. *Geophys Res Lett* 37:2–7
- Reisinger A, et al. (2013) Implications of alternative metrics for global mitigation costs and greenhouse gas emissions from agriculture. *Clim Change* 117:677–690
- Rodhe H (1990) A comparison of the contribution of various gases to the greenhouse effect. *Science* 248:1217–1219
- Roy M, Edwards MR, Trancik JE (2015) Methane mitigation timelines to inform energy technology evaluation. *Environ Res Lett* 11:04024:10
- Sarofim MC, Giordano MR (2018) A quantitative approach to evaluating the GWP timescale through implicit discount rates. *Earth Syst Dyn* 9:1013–1024
- Schmalensee R (1993) Comparing greenhouse gases for policy purposes. *Energy J* 14:245–255
- Shindell DT et al (2009) Improved attribution of climate forcing to emissions. *Science* 326:716–718
- Shine KP (2009) The global warming potential – the need for an interdisciplinary retrieval. *Clim Change* 96:467–472
- Shine KP, Allan RP, Collins WJ, Fuglestad JS (2015) Metrics for linking emissions of gases and aerosols to global precipitation changes. *Earth Syst Dynam* 6:525–540
- Shine KP, Berntsen TK, Fuglestad JS, Skeie RB, Stuber N (2007) Comparing the climate effect of emissions of short- and long-lived climate agents. *Philos Trans R Soc Lond A* 365:1903–1914
- Shine KP, Fuglestad JS, Hailemariam K, Stuber N (2005) Alternatives to the global warming potential for comparing climate impacts of emissions of greenhouse gases. *Clim Change* 68:281–302
- Smith SJ, Karas J, Edmonds J, Eom J, Mizrahi A (2013) Sensitivity of multi-gas climate policy to emission metrics. *Clim Change* 117:663–675
- Smith SM, et al. (2012) Equivalence of greenhouse-gas emissions for peak temperature limits. *Nat Clim Change* 2:8–11
- Smith CJ, et al. (2018) FAIR v1.3: a simple emissions-based impulse response and carbon cycle model. *Geosci Model Dev* 11:2273–2297
- Solomon S, et al. (2010) Persistence of climate changes due to a range of greenhouse gases. *Proc Natl Acad Sci USA* 107:18354–18359
- Sterner EO, Johansson DJA (2017) The effect of climate–carbon cycle feedbacks on emission metrics. *Environ Res Lett* 12:034019
- Sterner E, Johansson DJA, Azar C (2014) Emission metrics and sea level rise. *Clim Change* 127:335–351
- Stocker TF et al (eds) (2013) Climate Change 2013: The Physical Science Basis. Intergovernmental Panel on Climate Change, Geneva
- Streffer J, Luderer G, Aboumahboub T, Kriegler E (2014) Economic impacts of alternative greenhouse gas emission metrics: a model-based assessment. *Clim Change* 125:319–331
- Tanaka K, Boucher O, Ciais P, Johansson DJA, Morfeldt J (2021) Cost-effective implementation of the Paris agreement using flexible greenhouse gas metrics. *Sci Adv* 7:eabf9020
- Tanaka K, Cavaletto O, Collins WJ, Cherubini F (2019) Asserting the climate benefits of the coal-to-gas shift across temporal and spatial scales. *Nature Clim Change* 9:389–396
- Tanaka K, Johansson DJA, O'Neill BC, Fuglestad JS (2013) Emissions metrics under a 2°C stabilization target. *Clim Change* 117:933–941
- Tanaka K, O'Neill BC (2018) The Paris agreement zero-emissions goal is not always consistent with the 1.5°C and 2°C temperature targets. *Nature Clim Change* 8:319–324
- Tanaka K, O'Neill BC, Rokityanskiy D, Obersteiner M, Tol RSJ (2009) Evaluating global warming potentials with historical temperature. *Clim Change* 96:443–466
- Tanaka K, Peters GP, Fuglestad JS (2010) Multicomponent climate policy: Why do emission metrics matter? *Carbon Manag* 1:191–197
- Tol RSJ, Berntsen TK, O'Neill BC, Fuglestad JS, Shine KP (2012) A unifying framework for metrics for aggregating the climate effect of different emissions. *Environ Res Lett* 7:044006

- United Nations (2019) Report on the Conference of the Parties serving as the meeting of the Parties to the Paris Agreement on the third part of its first session, held in Katowice from 2 to 15 December 2018 (FCCC/PA/CMA/2018/3/Add.2)
- van den Berg M, Hof AF, van Vliet J, van Vuuren DP (2015) Impact of the choice of emission metric on greenhouse gas abatement and costs. *Environ Res Lett* 10:24001
- Wigley TML (1998) The Kyoto Protocol: CO₂, CH₄, and climate implications. *Geophys Res Lett* 25:2285–2288
- Wigley TML (2021) The relationship between net GHG emissions and radiative forcing with an application to Article 4.1 of the Paris Agreement. *Clim Change* 169:1–9
- Zhang X, Myhrvold NP, Hausfather Z, Caldeira K (2016) Climate benefits of natural gas as a bridge fuel and potential delay of near-zero energy systems. *Appl Energy* 167:317–322

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.