The International Journal of Life Cycle Assessment (2024) 29:328–343
https://doi.org/10.1007/s11367-023-02251-0

CARBON FOOTPRINTING

Evaluating metrics for quantifying the climate-change effects of land-based carbon fluxes

Miguel Brandão[1] · Miko U. F. Kirschbaum[2] · Annette L. Cowie[3]

Received: 18 July 2022 / Accepted: 12 November 2023 / Published online: 5 December 2023

Abstract

Purpose Growing concern over climate change has increased interest in making use of the biosphere to reduce net greenhouse gas emissions by replacing fossil energy with bioenergy or increasing land-based carbon storage. An assessment of the effectiveness of these options requires detailed quantification of their climate-change mitigation potential, which must employ appropriate metrics to translate biophysical changes into climate-change impacts. However, the various currently available metrics use different proxy measures (e.g. radiative forcing, temperature changes, or others) as surrogates for climate-change impacts. Use of these different proxies can lead to contradictory conclusions on the most suitable policy options. We aim to provide criteria for the objective evaluation of metrics to build understanding of the significance of choice of metric and as a step towards building consensus on the most appropriate metric to use in different contexts.

Methods We compared fifteen available metrics that represent conceptual differences in the treatment of biospheric carbon fluxes and the proxies used to approximate climate-change impacts. We proposed a set of evaluation criteria related to the metrics' relevance, comprehensiveness, ease of application and acceptance by the research and policy community. We then compared the different metrics against these criteria.

Results and conclusions The different metrics obtained scores from 10 to 21 (out of 30). The Climate-Change Impact Potential scored highest against the criteria, largely because it relates climate-change impacts to three different aspects of temperature changes; thus, it most comprehensively covers the different aspects of climate-change impacts. Therefore, according to our evaluation criteria, it would be the most suitable metric for assessing the effect of different policy options on marginal climate-change impacts. We demonstrated that the proposed evaluation criteria successfully differentiated between the fifteen metrics and could be used as a basis for selecting the most appropriate metric for specific applications.

Keywords Bioenergy · Characterisation factor · Climate-change policy · Global change · Mitigation · Temperature

1 Introduction

In 2019, fossil fuels comprised 80.9% of the world total energy supply of 606 EJ year^{-1}. Nuclear and hydropower contributed 5.0% and 2.5%, respectively, but bioenergy was the most significant non-fossil energy source, supplying around 10% of global energy needs (IEA 2021). Policy initiatives in various countries are aiming to further bolster the contribution of bioenergy (EU 2009). Newer technologies, such as combined cycle gasification for combined heat and power generation, have made energy recovery from biomass more efficient and cost-effective (Breeze 2017). Efforts to reduce CO_2 emissions from the burning of fossil fuels have led to increased use of bioenergy in some countries that have available biomass resources and a commitment to reduce the environmental impact from fossil-fuel use, such as Sweden and the UK (Cross et al. 2021).

However, the mitigation benefits of bioenergy have been questioned. In particular, doubts have been raised about the benefits of bioenergy use in cases where there is a long-time lag between initial emissions from logging a mature forest and the eventual accrual of benefits from fossil-fuel

Communicated by Michael Z. Hauschild.

✉ Miguel Brandão
miguel.brandao@abe.kth.se

1 Sustainable Development, Environmental Science and Engineering (SEED), KTH–Royal Institute of Technology, Stockholm, Sweden

2 Manaaki Whenua-Landcare Research, Palmerston North, New Zealand

3 NSW Department of Primary Industries, University of New England, Armidale, NSW, Australia

substitution (Fargione et al. 2008; Cherubini et al. 2013). It has been suggested that systems with very long payback times may not contribute to climate-change mitigation within relevant time frames (Gibbs et al. 2008). If there is a period during which atmospheric CO_2 is increased, even if only temporarily, it will have a warming impact that needs to be included in any overall assessment. Payback time is not a concern in bioenergy systems that involve a sequestration phase prior to harvesting biomass, such as where energy crops are planted on abandoned agricultural land. In that instance, the sequestration phase should be included in quantifying the overall mitigation benefit of such options.

Many current climate-change policies and initiatives provide incentives to delay emissions or sequester C temporarily. These policies implicitly assume that there would be climate benefits from temporary C sequestration or emission reductions that occur sooner rather than later, as they will deliver immediate CO_2 reductions and, therefore, reduce warming in the short term (e.g. Matthews et al. 2023; Parisa et al. 2022; Galik et al. 2022). However, there is no consensus whether such temporary storage delivers climate-change mitigation benefits in the long term (e.g. Korhonen et al. 2002; Kirschbaum 2006; Galik et. al, 2022; Matthews et al. 2023; Groom and Venmans 2023). Here, we briefly outline the main benefits and adverse consequences of temporary carbon storage.

Possible benefits of temporary storage:

- Reduced warming for the period over which CO_2 is removed from the atmosphere
- Slowing the short-term rate of warming, thereby allowing time for society to decarbonise energy and industrial systems
- Allowing time for ecosystems and societies to adapt to climate change
- Delay in reaching tipping points (but see the further discussion on tipping points below)
- Better capacity to cope with climate-change impacts in the future, assuming more wealth in the future

Neutral outcomes of temporary storage:

- To achieve equilibrium temperature targets (e.g. 1.5 °C above pre-industrial levels), the exact timing of CO_2 emissions and removals is not overly important. Instead, to mitigate long-term climate change, it is most important to restrict total cumulative GHG emissions (IPCC, 2021).
- The potential role of 'climate tipping points' (e.g. Lenton et al. 2019; Armstrong McKay et al. 2022) is very uncertain. While tipping points are clear issues of concern, there is no consensus on their exact timing and irreversibility. Even if the timing and nature of tipping points could be established with certainty, it is unclear what would be gained by delaying the crossing of a threshold by a short period of time. The most useful policy aim would be to avoid the crossing of irreversible thresholds altogether rather than just delaying it.

Possible adverse impacts of temporary storage:

- There may be C-cycle feedback effects that counteract any short-term benefit (e.g. Korhonen et al. 2002; Gillett and Matthews 2010). The rate of ocean CO_2 uptake is correlated with the atmospheric CO_2 concentration so reduced short-term CO_2 concentrations will reduce ocean uptake. When temporarily stored C is then re-released in future, the atmospheric CO_2 concentration and resultant global temperatures will reach higher values than without temporary C sequestration (Korhonen et al. 2002; Kirschbaum 2003).
- The same level of temperature increase is likely to have greater impacts if it occurs at a time with higher background temperatures. An extra unit of warming may be inconsequential under current conditions but may have greater adverse impacts under higher background temperatures in the future (Kirschbaum 2014).
- Although temporary C storage might mitigate short-term climatic problems, it forces the problem of high future GHG concentrations onto future generations.

Time-dependent assessments can be particularly important for bioenergy systems, where there may be an asynchrony between emissions and sequestration (Helin et al. 2013). Kirschbaum (2017), for example, analysed a complex bioenergy scenario with consecutive losses and gains of carbon. These changes interacted with changing background CO_2 and temperature to lead to ultimate climate-change impacts that could not have been predicted without analysing the cascade of processes in the climate system that stretched from initial perturbations of the carbon cycle to damages to human and natural systems.

In national inventory reporting under the UNFCCC, CO_2 emissions from bioenergy are captured in the land use, land-use change and forestry sector through changes in carbon stock that are reported on an annual basis. They are excluded from the energy sector emissions to avoid double counting (IPCC 2019; Volume 1, Chapter 1).

Direct land-use change (LUC) can be an important factor in the net GHG balance of bioenergy, leading to emissions during establishment of a bioenergy system, or, conversely, net C uptake if a bioenergy plantation is established on a site with low prior C stocks. The ISO standard 14067 (ISO 2018) requires the inclusion of direct LUC emissions in the C footprint of products and provides guidance on assigning LUC emissions and removals to products (see Annex E.2: Assigning biogenic GHG emissions and removals from land-use change and land use to products). The ISO standard suggests as possible options the averaging of these emissions over the forest or crop rotation period, or over the anticipated lifetime of the processing plant or applicable policy programme (ISO 2018). Indirect land-use change, although recognised as potentially significant, is excluded from ISO 14067 due to lack of an agreed method for its quantification.

The climate-change effects of bioenergy systems are commonly quantified through life cycle assessments (LCA; Helin et al. 2013). LCA methodologies follow the principles, framework, requirements and guidelines in the ISO 14040 and 14044 standards (ISO 2006a, b) and are built into voluntary standards and regulatory frameworks for assessing the sustainability of bioenergy use (e.g. EU 2009; ISO 2015). Key indicators are the GHG savings from avoided fossil-fuel emissions and sequestered C per unit of energy product (e.g. $kgCO_2eq\ MJ^{-1}$).

Ultimately, the benefit of any activity aimed at climate-change mitigation is related to the extent to which the activity can reduce marginal climate-change impacts. The challenge is to quantify this reduction in climate-change impacts by specific actions. Climate-change metrics quantify the effect of emitting a unit of a given greenhouse gas (GHG) on a key proxy measure related to climate change (IPCC, 2021). Different metrics use various proxies to approximate those ultimate impacts, and a key task is to understand the limitations of the use of those proxies and determine the valid applications and interpretation of their use.

Most LCA studies have adopted the Global Warming Potential (GWP) with a characterisation time horizon of 100 years (GWP_{100}) as the preferred metric to convert the emissions of different GHGs into a common unit, kilogram CO_2 equivalents (IPCC 1990). These studies generally do not acknowledge the timing of emissions and removals (Helin et al. 2013; Brandão et al. 2013), so do not capture climate impacts resulting from any asynchrony between emissions and sequestration.

However, some studies use different metrics, and the application of different climate-change metrics can significantly affect the results (see Ahlgren et al. 2015; Brandão et al. 2019; Garcia et al. 2020; Matustík et al. 2022). The Life Cycle Initiative recommended the use of both the GWP_{100} and Global Temperature change Potential (GTP_{100}), introduced by Shine et al. (2005), as the most appropriate metrics to reflect short- and long-term climate-change impacts, respectively (see Frischknecht et al. 2016; Levasseur et al. 2016; Cherubini et al. 2016; Jolliet et al. 2018).

As most climate-change metrics are insensitive to the timing of emissions and removals, LCA studies typically sum net emissions over the entire life cycle. Metrics (typically GWP_{100}) are then applied subsequently to the summed emissions based on the implicit assumption that the timing of emissions and removals does not affect climate-change impacts. However, the importance of timing has increasingly been recognised (e.g. Ericsson et al. 2012; Zetterberg and Chen 2014). The ISO standard for the C footprinting of a product (ISO 14067; ISO 2018), therefore, not only specified that the C footprint should be calculated as the sum of GHG emissions and removals over the life cycle of a product, with no modification related to timing, but the standard also required that the life cycle inventory includes the time profile of emissions and removals to enable supplementary calculations that include timing. While the ISO standard provided no metrics for such supplementary calculations, various metrics have been proposed for application in LCAs that would allow the impacts of time to be included in calculations (Brandão et al. 2013). Some metrics exclude emissions that occur after more than 100 years, or quantify such long-term emissions in a separate category (Helin et al. 2013), such as the PAS2050 specification for the assessment of the life cycle greenhouse gas emissions of goods and services (BSI 2011).

In previous work, Brandão et al. (2019) reviewed 15 metrics that have been applied to assess the climate-change impacts of bioenergy systems and illustrated the sensitivity of results to the choice of metrics. In that work, all metrics were applied to the same carbon-stock changes, and the metrics used different proxies or calculation routines to quantify changes in climate-change impacts. These different approaches resulted in significantly different estimates of net climate-change effects, with differences being particularly stark for a scenario that involved an initial loss of forest carbon stocks combined with an eventual compensation by large fossil-fuel substitution benefits. Using the calculated carbon-stock changes from that scenario with different assessment metrics yielded estimated climate-change mitigation effects that ranged widely from very beneficial to highly detrimental (Brandão et al. 2019). Similar findings have been reported in other work (e.g. Garcia et al. 2020; Matustík et al. 2022). The divergence of assessment can be particularly stark when different greenhouse gases are compared, especially gases with short and long atmospheric persistence. For CH_4 emissions, for example, GWP_{100} is 27 whereas GTP_{100} is just 4.7 (Forster et al. 2021).

These discrepancies are deeply concerning. If decision-makers use different metrics in their assessment of the climate-change mitigation potential of different policy options, their decisions will be strongly influenced by the arbitrary choice of the metrics they employ to make their assessments. This is likely to result in suboptimal decisions about the effectiveness of different options of mitigative action (Tanaka et al. 2021; Edwards and Trancik 2022).

Thus, the current application of different metrics leads to ambiguous results that cause confusion for practitioners, decision-makers and the public and is likely to result in adoption of suboptimal policies for climate-change mitigation. At the same time, there is an urgency to take robust steps towards real climate-change mitigation. There is, therefore, a clear need to evaluate the existing metrics and recommend a way forward to enable clear and unambiguous assessment of the mitigation potential of different mitigation options.

The work presented here aims to contribute to the deliberations of the global research community on refining the development and application of different metrics with the ultimate aim of supporting the assessment and implementation of more effective climate policy. We specifically aim to generate and demonstrate criteria to support the selection of appropriate metrics to enable objective assessment of the contributions of different strategies for climate-change mitigation.

We build on the earlier work of Brandão et al. (2019) by providing a set of criteria to evaluate the performance of fifteen climate-change metrics with respect to scientific and application aspects. We first provide an overview of climate-change metrics (Sect. 2), then summarise the specific metrics reviewed (Sect. 3.1), before presenting a set of evaluation criteria (Sect. 3.2). Section 4 presents the results of our evaluation of the metrics according to those criteria, and Sect. 5 discusses the evaluation results and implications. We highlight the particular significance of the timing of emissions and sequestration for the assessment of the climate-change mitigation potential of bioenergy systems.

2 Metrics for characterising GHG flows

Climate-change metrics express relationships between GHG emissions with respect to their effect on key proxy measures related to climate impacts. These metrics thereby express relationships between physical perturbations, such as net GHG fluxes, and suitable proxies that can be used as surrogates for ensuing climate-change impacts (e.g. IPCC 2013; Brandao et al. 2013; Levasseur et al. 2016). Proxy measures can comprise midpoint indicators (e.g. cumulative radiative forcing) or endpoint indicators (e.g. human health, sea-level rise) along the cause-effect chain. These proxies are then used to estimate the climate-change impact related to unit emissions of GHGs of interest to aid consumers or policy makers to choose between different options. It is essential that metrics are underlain by a transparent and meaningful theory that clearly relates the metric to the climate-change impacts that ultimately matter for humans and natural systems (e.g. EC, 2010).

Given the different atmospheric residence times and radiative efficiencies of different GHGs (Table 1; Fig. 1), the evaluation must adopt a specified characterisation time horizon so that GHGs with different atmospheric lifetimes or mitigation strategies, such as bioenergy systems, with different timing of emissions and removals can be compared (Myhre et al. 2011). This is particularly important for the comparison between emissions of short-lived GHGs, such as CH_4, and those of long-lived GHGs such as CO_2 and N_2O. CO_2 and N_2O exert ongoing radiative forcing even centuries after their emission. CH_4, on the other hand, exerts high initial radiative forcing, but has a short atmospheric lifetime of around 12 years before it is oxidised (Forster et al. 2021). The early radiative forcing of CH_4 is, however, retained by any metric based on cumulative radiative forcing, such as GWP. The metric GWP* has been developed to better reflect the warming effect of short-lived climate forcers such as methane, by determining CO_2 equivalence from changes in their emission rate (Lynch et al. 2020).

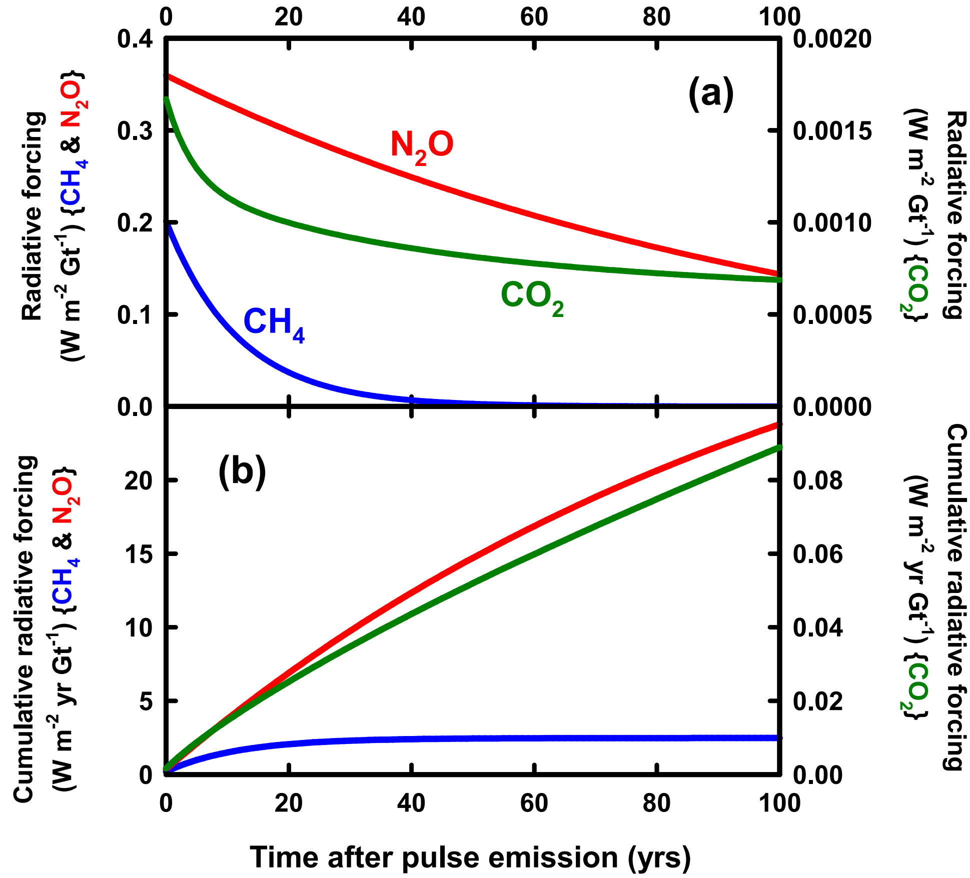
Different metrics use different proxy measures along the cause-effect chain (Fig. 2; Levasseur et al. 2016). Any emissions or removals of GHGs to or from the atmosphere change their atmospheric concentrations. Changing GHG concentrations then cause radiative forcing which will affect the climate, initially by changing surface temperatures, with subsequent effects on precipitation patterns, sea-level rise and extreme weather events, such as cyclones and heatwaves (Fig. 2). Ultimately, climate change impacts all species—human and non-human—through direct health effects and indirectly through changes in food production and ecosystem functioning. All metrics evaluated here fall somewhere between assessing the GHG fluxes (step 1) and ultimate impacts on human and natural systems (step 5).

Our work focuses on the 15 climate-change metrics that were previously studied by Brandão et al. (2019). The key features of these 15 metrics are described in Supporting Information (Appendix A) and summarised in Sect. 3 and Table 2. The most widely used metric for estimating climate-change impacts is the GWP (IPCC 1990; Fuglestvedt et al.

Table 1 Properties of the GHGs relevant to bioenergy systems (Forster et al. 2021)

Common name	Chemical formula	Atmospheric Lifetime (years)	Atmospheric concentration (ppmv) increase		Radiative efficiency (W m^{-2} $ppmv^{-1}$)	Radiative forcing in 2019 (W m^{-2})
			1750	2019		
Carbon Dioxide	CO2	Multiple	278	410 (47%)	0.013 ± 0.002	2.16 ± 0.26
Methane	CH4	11.8 ± 1.8	0.729	1.866 (156%)	0.57 ± 0.14	0.54 ± 0.11
Nitrous Oxide	N2O	109 ± 10	0.270	0.332 (23%)	2.8 ± 1.1	0.21 ± 0.03

Fig. 1 Radiative forcing after emissions of the three main GHGs, showing (**a**) instantaneous radiative forcing and (**b**) cumulative radiative forcing over 100 years. The scale for radiative forcing of CH_4 and N_2O is shown on the left axis and for CO_2 on the right axis. Figure computed from radiative efficiencies and decay functions in Table 1 (Forster et al. 2021)



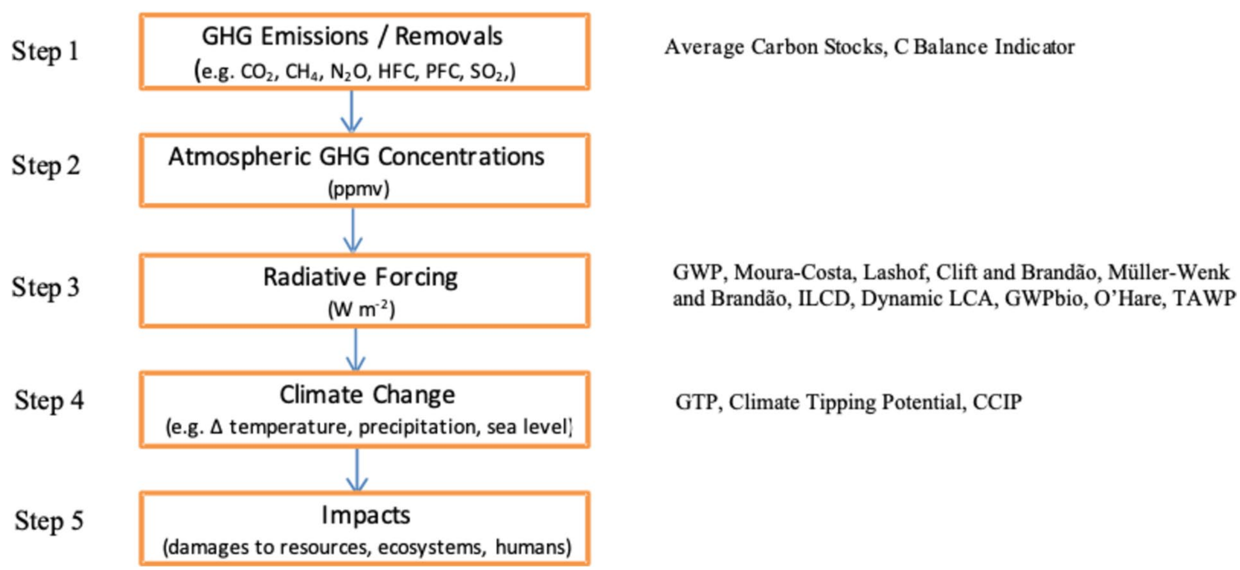


Fig. 2 The cause-effect chain illustrating the effects of GHG flows on climate change and associated impacts and ultimate damages (adapted from IPCC 2013). Different metrics (right hand side) use different proxies along the cause-effect chain. Policy relevance increases further down the chain, but so does uncertainty. The figure shows climate-change impacts resulting from effects of GHG emissions/removals acting via radiative forcing, but impacts can also result from changing the earth's energy balance in other ways (e.g. albedo), and conversely, some GHG can also have important direct effects, especially CO_2 and CH_4 (e.g. UNEP and CCAP 2021)

Table 2 Underlying features of the metrics

Metric	Reference	Effects measured	Impacts	Time sensitivity[a]	Sliding (S) / Truncated (T) Window[b]	Fossil / biospheric distinction[c]
GWP	IPCC (1990)	Radiative forcing	Cumulative	No	S	No
GTP	Shine et al. (2005)	Temperature	Point in time	Yes	T	No
Moura-Costa 1	Moura Costa and Wilson (2000)	Radiative forcing	Cumulative	No	S	Yes
Moura-Costa 2	Brandão et al. (2019)	Radiative forcing	Cumulative	Yes	T	No
Lashof	Fearnside et al. (2000)	Radiative forcing	Cumulative	Yes	T	Yes
Average Carbon Stocks	Kirschbaum et al. (2001)	C stocks	Cumulative	No	N/A	Yes
Müller-Wenk & Brandão	Müller-Wenk and Brandão (2010)	Radiative forcing	Cumulative	No	S	Yes
C Balance Indicator	Pingoud et al. (2016)	C flows	Cumulative	No	N/A	No
Clift & Brandão	Clift and Brandão (2008)	Radiative forcing	Cumulative	Yes	T	No
ILCD	EC (2010)	Radiative forcing	Cumulative	No	S	No
TAWP	Kendall et al. (2009)	Radiative forcing	Cumulative	Yes	T	No
Dynamic LCA	Levasseur et al. (2010)	Radiative forcing	Cumulative	Yes	T	No
GWPbio	Cherubini et al. (2011)	Radiative forcing	Cumulative	No	S	Yes
O'Hare	O'Hare et al. (2009)	Radiative forcing	Cumulative	Yes	T and discounted	No
CCIP	Kirschbaum (2014)	3 measures of temperature	Cumulative	Yes	T	No
Climate Tipping Potential	Jørgensen et al. (2014)	Cumulative C emissions	Point in time	Yes	T	No

[a]Time sensitivity describes whether the same activity occurring at different times within the assessment horizon leads to different calculated impacts

[b]Time window denoted as 'truncated' means that a fixed assessment period is used, whereas 'sliding' means that the assessment period expands with the timing of any future emissions

[c]The distinction between fossil and biospheric refers to the treatment of C stocks and whether use or saving of fossil fuels are treated differently from biospheric C-stock changes

2003; Shine et al. 2005). It quantifies the cumulative radiative forcing that results from a pulse emission of a unit of a greenhouse gas relative to that of CO_2 (step 3 in the cause-effect chain, Fig. 2). The Global Temperature change Potential (GTP) has also gained prominence over recent years (e.g. Stohl et al. 2015; Levasseur et al. 2016; Jolliet et al. 2018). It quantifies the temperature change at a specified future time.

Metrics, such as GWP and GTP, enable emissions of different GHGs to be weighted relative to a reference GHG emission. For example, GWP_{100} is used for national reporting under the UNFCCC to quantify the warming contributions from all country-wide net emissions of different GHGs from all sources and sinks. Metrics enable the net emissions of different gases to be expressed in the common units of CO_2 equivalents (CO_2eq).

Metrics require different assumptions and modelling choices that introduce some subjectivity, such as the choice of proxy measures to use, or the characterisation time horizon, the period over which to integrate radiative forcing or quantify other proxy measures. The sixth IPCC Assessment Report presented values of GWPs and GTPs for characterisation time horizons of 20, 50, 100 and 500 years (Forster et al. 2021) with very different numeric values. For example, the relative value for methane, expressed in CO_2 equivalents, varied between 4.7 ± 2.9 (for GTP_{100} with climate-change feedbacks) and 82.5 ± 25.8 (for GWP_{20}). This shows that the use of different metrics and characterisation time horizons can result in different C footprint assessments of identical bioenergy systems, and Cherubini et al. (2013) demonstrated that the use of GTP or GWP, and different characterisation time horizons, resulted in widely contrasting apparent benefits of the assessed bioenergy systems.

A key difference between metrics is whether they use cumulative or instantaneous proxies (Levasseur et al. 2016). GWP quantifies cumulative radiative forcing over the specified characterisation time horizon. In contrast, the GTP calculates the effect of a pulse emission of a GHG on the temperature change at a specific point in time and expresses this with respect to a pulse emission of CO_2. This fundamental difference between metrics leads to profound differences in the assessed importance of different GHGs.

Both cumulative and instantaneous measures reflect important biogeochemical aspects underlying different climate-change impacts (Fuglestvedt et al. 2003; Tanaka et al. 2010; Levasseur et al. 2016). Metrics that focus on only one of those impact categories, based on either cumulative or instantaneous measures, therefore, inevitably omit some aspects of actual impacts. That has been recognised in the development of the Climate-Change Impact Potential (CCIP) that explicitly calculates instantaneous temperature changes and rates of warming and cumulative warming and estimates overall climate-change impacts from the average of the three contributing components. It thus provides an indicator that reflects a broader range of ultimate climate-change impacts than metrics based on only one or the other impact category (Kirschbaum 2014).

With respect to the timing of biospheric CO_2 fluxes, it has been argued that any delayed emissions should be discounted as their impacts happen in the future (e.g. O'Hare et al. 2009; Groom and Venmans 2023). Analogous to the economic discounting of flows of income and capital, biophysical accounting could see emissions in the future assigned less weight than near-term emissions. If impacts are discounted at an annual rate of 5%, for example, it would reduce their value in year 100 to less than 0.8% of their value without discounting. The rationale for this discounting is partly based on a time preference of people. It is common to give greater importance to factors that will affect us in the near future than those that affect us later or only affect future generations. However, giving preference to our current generation over that of our children and grandchildren is difficult to justify ethically. It is also inconsistent with sustainable development, defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (Brundtland 1987).

Application of discount rates could also be justified through an expectation that future societies will be better able to deal with climate change through adaptation and the development of new technologies for mitigation and the production of less-GHG-intensive goods (e.g. energy and food). However, this subjective expectation is open to criticism, and some writers have suggested that the planet's future environmental buffering capacity may well be reduced compared to the present so that negative discount rates may actually be more appropriate than positive rates (e.g. Sterner and Persson 2008). Only the O'Hare metric includes an explicit discount rate.

3 Criteria for evaluating climate-change metrics

3.1 Brief summary of the considered metrics

Table 2 summarises the main features of the 15 metrics we analysed. A fuller description of each metric is given in the Supporting Information (Appendix A).

All metrics other than the Climate-Change Impact Potential calculate single proxy variables as an implicit measure of climate-change impacts. Two metrics (average carbon stocks and the carbon balance indicators) restrict their assessment to an evaluation of C stock changes, which is equated with an emission or removal of CO_2 (step 1 in Fig. 2). All other metrics go further and, at least, assess effects on radiative forcing. The GTP goes as far as calculating temperature changes at a future date, and the CCIP goes yet another step further by calculating three differently quantified temperature perturbations of the climate system and by multiplying each change by a severity term to get a step closer to impacts.

All metrics other than the GTP use cumulative measures as their proxies for climate-change impacts. They differ in their considered time windows, with either a fixed or sliding window for assessing impacts. A fixed window excludes any radiative forcing or other impact proxy beyond the end of that characterisation time horizon, whereas a sliding window expands the characterisation time horizon by the timing of emissions/removals. The GWP, for example, employs a sliding window, whereby emissions occurring in year 1 or year 50 are integrated over the same subsequent characterisation time horizon of typically 100 years.

Some metrics treat biospheric and fossil CO_2 net emissions differently (e.g. Helin et al. 2013) based on an interpretation of C-cycle feedbacks quantified through the Bern model (Joos et al. 2013) even though it is unclear whether such a distinction between fossil and biospheric CO_2 is useful and warranted (Brandão et al. 2013).

Other than the CCIP, none of the metrics consider the links between temperature perturbations and resultant impacts that are generally considered to be non-linear. Larger temperature increases are usually considered to lead to disproportionately larger impacts (Weitzman 2012, 2013; Howard and Sterner 2017). Any temperature change through mitigative action would then have much greater climate-change impacts if the temperature change occurs at a time with higher background temperatures.

3.2 Detailed criteria for evaluation of metrics

We developed a set of criteria for evaluating the climate-change metrics, adapted from the International Reference Life Cycle Data System (ILCD) handbook (EC 2010), and specifically tailored them to the needs of an evaluation of metrics for assessing climate-change impacts. In total, the assessment allowed for a total of 30 points, based on four broad categories, with one category further subdivided into three subcategories:

(A) Extent to which the metric reflects the ultimate impacts on humans and the environment (see Fig. 2): 5 points
(B) Comprehensiveness with which the proxies used by the metric cover the biogeochemical processes in the cause-effect chain that lead to climate-change impacts:
(C) B_1: instantaneous effects of increased temperature (5 points)
(D) B_2: effects of cumulative warming (5 points)
(E) B_3: inclusion of additional features, such as dependence on future base conditions or time discounting (5 points)
(F) Ease of application (5 points)
(G) Extent to which the metric has been peer reviewed (5 points)

These four criteria provide a detailed, transparent and reproducible assessment of the different metrics and can be abbreviated as (A) relevance, (B) comprehensiveness, (C) ease of application and (D) acceptance by the research and policy community. This procedure is similar to the approaches used in similar previous exercises. The ILCD Handbook used five scientific criteria for the evaluation of metrics (EC 2010): (i) completeness of scope; (ii) environmental relevance; (iii) scientific robustness and certainty; (iv) documentation, transparency and reproducibility; and (v) applicability. Additionally, it used one stakeholder criterion: degree of stakeholder acceptance and suitability for communication in a business and policy context.

For our assessment, we summed scores across the four criteria, with 5 points assigned each to criteria A, C and D and 15 points to criterion B, which implicitly made criteria A, C and D equally important, but assessed criterion B to be as important as all other criteria put together.

Criterion A represents the depth within the cause-effect chain from net GHG emissions to ultimate impacts on human well-being and the natural environment (Fig. 2). The further down the cause-effect chain a proxy measure is targeted, the higher the assigned score. Criterion B_1 relates to quantification of impacts as an instantaneous consequence of future temperature changes. Criterion B_2 complements Criterion B_1 by representing cumulative warming in the metric. Criteria A, B_1 and B_2 are considered to be the most important criteria to describe the representation of the key processes linked to ultimate climate-change impacts, and, thus, each carried a maximum of five points.

Criterion B_3 relates to the coverage of five other factors, but because each is considered to be less important than those under the first three criteria, only one point was assigned for each. Criterion B_3 (1) relates to the inclusion of direct effects of changing CO_2 (including its effect on radiative efficiency); Criterion B_3 (2) relates to inclusion of the effect of changing background temperatures on the calculation of marginal impacts; Criterion B_3 (3) relates to the inclusion of timing effects, such as discounting or critical thresholds; Criterion B_3 (4) relates to inclusion of non-climate impacts of GHG emissions such as ocean acidification; and Criterion B_3 (5) relates to the inclusion of the rate of warming as a type of climatic impact.

Criterion C relates to the ease of applying each metric. Some metrics present assessments through a simple table of numeric values (such as the tabulation of GWPs of different gases available from Forster et al. 2021), while other metrics require calculations to be carried out by the user, for specified conditions. Finally, Criterion D assesses whether critical parts of the model have been peer reviewed, and whether model documentation has been published and is readily accessible. Appendix B provides further details on these criteria and the reasons for the specific scores given to each metric.

4 Metric evaluation

Table 3 shows the scores awarded to each of the 15 metrics. Under criterion (A), all metrics address at least GHG emissions and removals (step 1 in Fig. 2), but the average carbon stocks and the carbon balance indicators are better thought of as inventory indicators and thus score poorly under this criterion. The GTP uses a more advanced proxy measure on the cause-effect chain, but only the CCIP uses proxies that come close to targeting endpoint impacts. The proxies used for climate-change impacts by other metrics are based on calculations at an early stage in the cause-effect chain (see the Supporting Information for further elaboration).

Criterion B considers the comprehensiveness with which impact pathways (Fuglestvedt et al. 2003; Kirschbaum 2014) are addressed, with Criterion B_1 focusing on the representation of impacts linked to the instantaneous effect of elevated temperatures. This criterion is given a weight of 5 points because comprehensiveness in the coverage of the environmental mechanism (i.e. cause-effect chain or impact pathway) is indeed the most important criterion for a metric intended to represent the impacts of climate change. Metrics that use atmospheric CO_2 as their proxies were assigned low scores as these metrics disregard the various steps between CO_2 concentration and eventual temperature-related impacts. Other metrics, including the GWP, use cumulative radiative forcing as

Table 3 Evaluation of the different metrics

	A	B1	B2	B3	C	D	Total
Global Warming Potential (GWP)	2	2	3	0	5	5	17
Global Temperature change Potential (GTP)	3	3	1	0	5	5	17
Moura-Costa 1	1	2	3	0	3	5	14
Moura-Costa 2	1	2	3	0	3	2	11
Lashof	1	2	3	0	3	3	12
Average C Stocks	0	1	1	0	3	5	10
Müller-Wenk & Brandão	1	2	3	0	3	5	14
C Balance Indicator	0	1	1	0	4	5	11
Clift & Brandão	1	2	3	0	3	3	12
International Reference Life Cycle Data System (ILCD)	1	2	3	0	3	3	12
Time-Adjusted Warming Potential (TAWP)	2	2	3	0	3	5	15
Dynamic LCA	1	2	3	0	3	5	14
Global Warming Potential bio (GWPbio)	1	2	3	0	2	5	13
O'Hare	1	1	3	1	3	5	14
Climate Change Impact Potential (CCIP)	4	4	4	3	1	5	21
Climate Tipping Potential	2	3	1	1	1	5	13

Each criterion is assessed on a range from '0' (not included at all) to '5' (achieved to its fullest). The metrics were given the following indicative scores: 5: Full compliance; 4: Compliance in all essential aspects; 3: Compliance in some aspects; 2: Little compliance; 1: No compliance

proxies (with various adjustments), which are not tightly correlated with maximum temperature increases. Short-lived GHGs, like CH_4, for example, exert radiative forcing for a short time after emissions, which is retained in measures of cumulative radiative forcing (see Fig. 1), but they will contribute little to temperatures experienced 100 years later (Fuglestvedt et al. 2003).

A score of ‘3’ was assigned to the GTP as it explicitly uses future temperatures as its impact proxy. A higher score of ‘4’ was assigned to the CCIP as it goes another step further by calculating an impact term as an explicit function of temperature changes in conjunction with varying background temperatures. None of the metrics was assigned a value of ‘5’ as none of them goes as far as explicitly assessing the impacts that arise from the different climate perturbations.

Criterion B_2 is similar to criterion B_1 but focuses on cumulative warming instead of the instantaneous effect of warming. The various metrics that use cumulative warming or cumulative radiative forcing as proxies for impacts are all assigned higher scores, especially the GWP. In contrast, the GTP is assigned a low score as it is solely based on an instantaneous rather than cumulative measure. The CCIP is assigned a ‘4’ as it explicitly calculates cumulative warming and thereby takes the calculations further than the GWP and other metrics that use cumulative radiative forcing as a proxy.

Criterion B_3 aims to capture any additional aspects related to climate-change impacts that are not covered under the first two categories. We considered five possible factors, with each one assigned only a single point as they are not considered to be as important as the primary factors considered under categories B_1 and B_2. They are described in more detail in the Supporting Information. The first of these is the inclusion of changing background CO_2 concentrations. Increasing background CO_2 will reduce the radiative efficiency of any additional CO_2 added to the atmosphere (Reisinger et al. 2011). This factor is only considered in the CCIP calculations, assigning it one point.

The second factor is the dependence of calculated impacts on background perturbations, such as temperature. The resultant impact from any temperature perturbation is greater if background conditions are already elevated. This factor, too, is only included in the CCIP calculations, assigning it a further point.

The third factor is the timing of impacts. Apart from the O’Hare metric and the Climate Tipping Potential, all metrics consider impacts at all times to be equally important so that an impact occurring 1 year after a pulse emission is assigned the same importance as an impact occurring 100 years after an emission. Deciding on the appropriateness of time discounting is an ethical rather than scientific issue, with no clear right and wrong approaches, but only the O’Hare metric explicitly includes time preference through a selected discount rate (O’Hare et al. 2009). The Climate Tipping Potential considers time by assessing the impact of an emission based on its proximity to a critical climate threshold. The O’Hare metric and the Climate Tipping Potential were each assigned a point here.

The fourth factor covers inclusion of (positive or negative) non-climate-change impacts of GHG emissions, such as CO_2 fertilisation, ocean acidification or the effects of CH_4 on tropospheric ozone (UNEP and CCAC 2021). None of the metrics analysed here includes non-climate-change impacts, so no points were assigned to any metric.

The fifth and final factor covers the inclusion of the rate of warming as a measure of climate-change impacts. It is a measure of the capacity for adaptation by both natural and socio-economic systems (Peck and Teisberg 1994; Jump and Peñuelas 2005; Millington et al. 2019) but generally received less attention than instantaneous warming or cumulative warming. Of the assessed metrics, only the CCIP includes calculation of the rate of warming, thus receiving one additional point here.

Most metrics did not obtain high scores against criterion B. Scores were mostly low because metrics either use very simple sets of calculations based on proxies calculated early in the cause-effect chain, or because they focus on only one impact category (instantaneous or cumulative) but ignore the other. Of the assessed metrics, the CCIP is the only one to explicitly include both cumulative and instantaneous measures, takes calculations further down the cause-effect chain than other metrics and includes other important features, such as a dependence on changing background CO_2 and temperature.

For criterion C, some metrics, like GWP and GTP, have look-up tables so that users can simply read required values from those tables. These metrics were assigned a score of 5. Other metrics require values to be calculated by users under specified conditions. Those calculations can be quite simple while other impact proxies require more complex calculations. The simplest metric is the carbon balance indicator, which is consequently assigned 4 points. Most other metrics, like those based on calculated radiative forcing, are only slightly more complex with few extra assumptions and calculations and are assigned 3 points. The GWP_{bio} requires more complex calculations, so that it was assigned only 2 points. The most complex calculations are required for the CCIP as it requires calculations of several steps along the cause-effect chain. Scores under criterion C tend to be inversely related to scores under criterion B, as more comprehensive representation of the details of climate-change impacts and an ability to consider the timing of emissions require greater calculation complexity and consequently reduced transparency.

For Criterion D, most metrics scored 5 points as the metrics have been described in peer-reviewed literature or other accessible publications. The ILCD, Lashof and Clift and Brandão metrics score lower as they have not been described in peer-reviewed literature, and details of the calculations of Moura-Costa 2 have been described formally only in the Appendix of Brandão et al. (2019).

Overall, considering Criteria A–D, the CCIP scored highest with an overall score of 21, mostly on the strength of its explicit inclusion of all three impact types. GWP and GTP obtained the next highest scores of 17. The Time-Adjusted Warming Potential obtained the next highest scores of 15, further followed by a range of other metrics with scores of 14 (Table 3).

5 Discussion

For comparing the role of different gases or policy options, most studies currently use GWP_{100}. Over recent years, use of GTP has also gained greater prominence as an alternative metric, often suggested as a complementary metric, to be presented along with GWP_{100} (e.g. Stohl et al. 2015; Levasseur et al. 2016; Jolliet et al. 2018). However, these two options provide strongly contrasting assessments for the same GHG fluxes, especially in the comparison of short-lived and long-lived GHGs (Forster et al. 2021). Other available metrics provide an even wider range of assessments of the same policy scenario that can range from positive to negative outcomes, even when only CO_2 is considered (Brandão et al. 2019; Garcia et al. 2020; Matustík et al. 2022). There can be a valuable role of different metrics for addressing specific aspects of the climate-change problem and highlighting important aspects or interactions. In a research context, or to inform policy development, there is, therefore, a justifiable role for different and complementary approaches that utilise a variety of available proxy measures. For these applications, multiple metrics could be used in a complementary manner, each representing and focusing on different aspects of climate-change impacts.

However, for devising optimal approaches to climate-change mitigation, the multitude of possible results from the wide range of available metrics leads to an untenable situation. It prevents a clear and unambiguous understanding of the effectiveness of different policy options for climate-change mitigation. In attempts to rectify that problem, various initiatives have sought to improve available GHG accounting metrics (e.g. Plattner et al. 2009; Jolliet et al. 2018). However, the understanding of metrics is still incomplete, their usage is inconsistent, possibly because guidance is contradictory, and consensus on desirable steps forward remains elusive. This situation has provided the rationale for the work reported here, where we aim to provide a perspective on the set of available choices that is complementary to those reported by Plattner et al. (2009) and Jolliet et al. (2018) and to contribute to the discussion of available metrics and their respective strengths and weaknesses.

Choosing the most appropriate metric for assessing climate-change impacts is challenging as all metrics have their merits and capture important aspects of the atmospheric and climatic changes that lead to ultimate impacts. In terms of relevance, the metrics that use proxies closer to impact endpoints would be preferable. However, complexity and uncertainty tend to increase with proximity to the endpoint as many ecosystem and human-system variables are impacted, and practitioners have to make an increasing number of value choices. Conversely, if transparency and ease of application are assigned higher priority in the selection criteria, then a midpoint indicator may be more suitable (Levasseur et al. 2016).

A key difference between metrics is their use of either instantaneous or cumulative measures as proxies of impacts, which each quantify different important components of climate-change impacts (e.g. Fuglestvedt et al. 2003; Tanaka et al. 2010; Levasseur et al. 2016). An instantaneous measure could be the temperature in future years, which would be an appropriate measure to quantify impacts such as damage from heat waves. A cumulative measure could be cumulative radiative forcing or cumulative warming, which would be an appropriate measure for impacts such as sea-level rise, in particular (e.g. Rahmstorf et al. 2012).

The large difference between metrics based on instantaneous and cumulative measures can be clearly seen in the comparison between warming potentials of the short-lived GHG CH_4, calculated with GWP_{100} as 27 and with GTP_{100} as 4.7, respectively (Forster et al. 2021). The reason for that difference is readily identified: GWP is based on a cumulative measure, while the GTP is based on an instantaneous measure. These two measures lead to vastly different assessments of ultimate impacts, even though both reflect climate-change impacts that are relevant to decision-making. It is clear that this dilemma can only be overcome if both instantaneous and cumulative measures are explicitly included in any assessment. That explicit inclusion of both instantaneous and cumulative factors has so far only been attempted in the CCIP.

A vexed issue relates to the question of whether expected future background conditions of GHG concentration and expected warming should be considered by the metric. Previous work has shown that such background conditions can be important and affect calculated impacts of marginal emissions. For example, the impact potentials of biogenic CH_4 calculated with the CCIP from 2010 to 2110 varied from 29 calculated under constant (2010) background conditions to 14 under RCP 8.5 (Kirschbaum 2014). Would it then be more appropriate to include the most relevant expected future background conditions in metrics calculations or would it be better to use constant conditions in the interest of transparency and reproducibility? There is no easy answer to that question as it is important to both derive accurate values and also be clear and unambiguous in a policy context, and future climatic background conditions are clearly uncertain.

Climate policy generally aims to mitigate climate-change impacts, and choosing a metric that comes closest to that ultimate policy goal is generally more important than simple operational considerations. For simple emissions choices, the selection of a metric is of limited importance since all steps along the emission-mpact chain (Fig. 2) are correlated. So, avoiding a unit of CO_2 emissions would have useful mitigation benefits regardless of the metric through which it is assessed. However, scenarios become more difficult to assess when they include both removals and emissions of GHGs at different times. For assessing the impact of such scenarios, it is important to use a metric that provides answers that correspond to ultimate climate-change impacts. Brandão et al. (2019), Garcia et al. (2020) and Matustík et al. (2022) showed that different metrics can assess the same scenario very differently, which makes the choice of metric critically important.

In practice, the GWP has been the most widely applied metric and has been mandated in national inventory reporting under the UNFCCC, but it is not without its limitations. For example, as a measure of cumulative radiative forcing, GWP is not directly related to a temperature limit (e.g. the 1.5 °C Paris target) nor does it provide a direct assessment of damages caused by future extreme temperatures (Shine et al. 2005). The Life Cycle Initiative (e.g. Jolliet et al. 2018), therefore, recommended the use of the GTP to complement the use of the GWP so that both metrics used in combination could reflect different kinds of impacts. However, it is questionable whether an analysis that produces two strongly contrasting numeric assessments of climate-change impacts is actually helpful. Instead, it is likely to create confusion, particularly for routine applications such as assessing compliance with prescribed emission limits.

The evaluation presented here has identified the Climate-Change Impact Potential as the most complete metric for providing a broad assessment of climate-change impacts. It focuses on the important endpoint concerns and includes assessments of three aspects of temperature perturbation reflecting three kinds of climate-change impacts: impacts related directly to future temperatures (e.g. heat-wave impacts), impacts related to the rate of warming (e.g. capacity for ecological or societal adaptation) and impacts related to cumulative warming (e.g. sea-level rise).

It is generally acknowledged that all three kinds are important (e.g. Fuglestvedt et al. 2003; Tanaka et al. 2010; Levasseur et al. 2016), yet all metrics apart from the CCIP focus on only one or other of these kinds of impacts. The CCIP explicitly includes all three kinds and assigns equal weighting to all three (Kirschbaum 2014). To our knowledge, there has been no comprehensive study that compared the relative importance of these three kinds of impact to provide an objective basis for weighting each component. While the weighting of the three kinds of impacts has not yet been established, it is clear that all three kinds are important. Metrics that do not try to include all three kinds are, therefore, necessarily incomplete in their assessments and cannot adequately describe the breadth of climate-change impacts.

Previous studies have compared alternative metrics (e.g. Peters et al. 2011; Azar and Johansson 2012; Tanaka et al. 2013, 2021; Edwards and Trancik 2022), and frameworks have been established to recommend use of specific metrics (e.g. Plattner et al. 2009; European Commission, 2010). These frameworks aimed to harmonise practice and ensure consistency and reproducibility. However, global consensus on metrics for use in policy development and implementation, and business decision-making, has not, yet, been reached.

Evaluating metrics against a common set of criteria is not free from value choices. Some choices are necessarily arbitrary, such as the characterisation time horizons and time discount rates, all of which affect the weight given to different GHGs relative to that of CO_2. The standard characterisation time horizon used in most studies and applied in most policies and voluntary schemes is 100 years. This characterisation time horizon is long enough to capture relevant climate-change impacts and is still within a long-term planning horizon.

As climate-change impacts are becoming apparent, and a need for urgent action is widely recognised, there have been calls to adopt shorter assessment horizons, such as 20 years, that would greatly shift the focus to short-lived GHGs such as CH_4. However, if that were to lead to preferential emission reductions of short-lived GHGs while ongoing emission of long-lived GHGs such as CO_2, it could be detrimental to achieving eventual climate stabilisation in the longer term (Balcombe et al. 2018). In terms of total climate-change impacts over 100 years, such a change in emphasis toward increased focus on short-lived GHGs could be counter-productive for reducing overall climate-change impacts.

Discount rates are another vexed issue. Most metrics use no discount rates, thus implicitly assigning the same importance to impacts occurring in 99 years, when they may affect the lives of our grandchildren, as to impacts that might occur next year. However, it could equally be argued that future generations might be better (e.g. Nordhaus 1997; Caney 2016) or less well (Sterner and Persson 2008) able to cope with adverse conditions in the future. Of the metrics assessed here, only the O'Hare metric (O'Hare et al. 2009) includes an explicit mechanism to adjust future impacts through a system of time preference. The choice of time preference cannot be resolved scientifically but requires resolution based on ethical arguments.

Currently, the formulation of evidence-based climate-change policy is hindered by the application of alternative metrics that lead to conflicting results about the effect of possible policy options. The usefulness of any proposed mitigation option cannot be determined unambiguously as the assessment will differ depending on the chosen assessment metric. Therefore, we urge the climate-science, climate policy and LCA communities to work towards reaching a consensus on the metrics used for assessment of climate-change impacts.

To assist in that process, we have proposed and demonstrated a set of criteria for objective evaluation of different climate-change metrics. These criteria could inform future evaluations towards tangible and quantifiable assessment of the strengths and weaknesses of climate-change metrics.

Our assessment was based on different criteria over four broad areas: (A) relevance, (B) comprehensiveness, (C) ease of application and (D) acceptance. All four areas are important, but relevance and comprehensiveness are clearly the most important. If a metric relies on proxies that do not comprehensively cover the relevant factors related to climate-change impacts, that shortcoming could not be compensated by ease of use or wide acceptance. We, therefore, awarded 20 out of a possible 30 selection points to relevance and comprehensiveness, but it could be argued that even that numeric condition might not have been strong enough and that only metrics should be considered that reach minimum thresholds for relevance and comprehensiveness.

Within the criterion of comprehensiveness, we considered inclusion of instantaneous and cumulative measures as being the most important for covering the range of different impacts that require different quantifications. There is general acceptance that the different kinds of impacts require different quantifications (e.g. Fuglestvedt et al. 2003; Tanaka et al. 2010; Levasseur et al. 2016). For example, sea-level rise is related to cumulative warming (e.g. Rahmstorf et al. 2012). It, therefore, cannot be quantified adequately by a metric based on future temperature increases. The aim of the Paris Accord to restrict future temperature increases to 2 °C is, therefore, unlikely to be sufficient to prevent ongoing sea-level rise for centuries. Conversely, dangerous impacts related to future heat waves would not be adequately quantified by cumulative measures, such as cumulative radiative forcing as used in the GWP. Application of the GWP, therefore, would be a poor quantification for impacts related instantaneously to temperatures in 100 years.

The distinction between criteria B_1 (focused on instantaneous impact measures) and B_2 (focused on cumulative impact measures) is, therefore, a necessary and critically important distinction and inclusion. In criterion B_3(5), we focus also on the rate of warming related to additional kinds of impacts (e.g. Peck and Teisberg 1994; Jump and Peñuelas 2005; Millington et al. 2019). In the development of CCIPs, Kirschbaum (2014) had assigned it the same importance as the other two kinds of impact categories (instantaneous and cumulative), but it generally receives less attention in impacts work (e.g. Levasseur et al. 2016), and we, therefore, allowed only one point for its inclusion.

Under criterion B_3, we also considered inclusion of background CO_2 concentrations, $B_3(1)$; background temperatures, $B_3(2)$; the timing of impacts, $B_3(3)$; and non-climate-change impacts, $B_3(4)$, awarding 1 point, each, to metrics that include these factors, but the validity of their inclusion can be questioned. It has been shown that changing background CO_2 concentrations and temperatures can affect calculated impacts (e.g. Reisinger et al. 2011; Kirschbaum 2014, 2017), which would suggest that these factors should be included in any metric calculation, but it makes calculated impacts dependent on the assumed path of these background conditions. On balance, we have considered that inclusion of these background conditions is warranted in order to get the truest possible reflection of the impact of any current day emissions, but lack of transparency and certainty are necessary costs of that inclusion.

Inclusion of a mechanism for accounting for the timing of impacts is even more controversial (O'Hare et al. 2009). Inclusion of a discount rate leads to a devaluation of impacts occurring in a more distant future. Would it, indeed, be appropriate to include a weighting for impacts quantified through any relevant proxy (Sterner and Persson 2008; Caney 2016)? Here, we awarded 1 point for the only available metric that included a mechanism for time discounting, but it could also be argued that such weighting is inappropriate and that points should even be deducted, instead.

All metrics considered here deal only with climate-change impacts, but the environment is also affected by the emission of GHGs in other ways, including air pollution with direct human-health effects, especially caused by methane (Sarofim et al. 2017; Shindell et al. 2017; UNEP and CCAC 2021). Increasing CO_2 concentration also has various non-climatic effects that may be positive, such as increasing plant production (e.g. Norby et al. 2005; Long et al. 2006) or negative, such as causing ocean acidification (e.g. Doney et al. 2009; Fox-Kemper et al. 2021). We potentially awarded 1 extra point for metrics that included non-climatic effects, but none of our considered metrics qualified for that.

The set of evaluation criteria presented here breaks the comparison between metrics into more explicitly identifiable components. This makes it easier to follow and question the reasoning for the ranking of metrics. One could take the analysis further and apply a mathematical framework for the assessment, such as a multi-criteria decision analysis (MCDA) (e.g. Wang et al. 2009; Kumar et al. 2017). Applying such a formal analysis based on input of selection criteria and their relative weighting could be a desirable refinement of the work initiated here.

If the relevant global communities can agree on such a unified approach towards evaluating, recommending or further developing appropriate assessment metrics, it will enable decision-makers to reach more robust decisions on climate-change mitigation strategies and support policy implementation. Agreed metrics can then be applied to provide consistent assessment of alternative bioenergy systems or assess other climate-change mitigation options that vary with respect to timing of emissions and removals or that involve trade-offs between different greenhouse gases with different atmospheric lifetimes.

6 Conclusions

Previous work has shown that the climate-change effects of a defined land-based activity could be assessed very differently through application of different metrics. Different metrics also result in very different assessed importance of different greenhouse gases, with stark differences between gases with short or long atmospheric residence times. Such divergence through the choice of metrics is deeply concerning as it leads to ambiguous results about the usefulness of different mitigation options and prevents rational decision-making about the most appropriate and cost-effective mitigative action.

This creates an urgency to assess available greenhouse gas metrics and adopt or develop a metric that most appropriately reflects the effect of different emission-related activities on ultimate climate-change impacts. To that end, we developed a set of criteria to assess fifteen available metrics based on their relevance, comprehensiveness, ease of application and acceptance by the research and policy community. We assigned scores under these headings to each of the metrics for a potential maximum score of 30 points.

According to our set of criteria, the Climate-Change Impact Potential (CCIP) scored highest with 21 points, followed by the Global Warming Potential (GWP) and Global Temperature change Potential (GTP) each with 17 points. The high points scored by the CCIP were mainly due to its more complete coverage of three different aspects of temperature changes, namely instantaneous warming, cumulative warming and rate of warming, that together define the range of relevant climate-change impacts. We demonstrated that the proposed evaluation criteria successfully differentiated between the fifteen metrics and could be used as a basis for selecting the most appropriate metric for specific applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11367-023-02251-0.

Acknowledgements We would like to acknowledge support from IEA Bioenergy.

Funding Open access funding provided by Royal Institute of Technology. The work of MUFK was supported by the Strategic Science Investment Fund (SSIF) of New Zealand's Ministry of Business, Innovation and Employment.

Data availability All relevant data have been provided in Appendix B.

Declarations

Conflict of interest The authors declare no competing interests.

References

Ahlgren S, Björklund A, Ekman A, Karlsson H, Berlin J, Börjesson P, Ekvall T, Finnveden G, Janssen M, Strid I (2015) Review of methodological choices in LCA of biorefinery systems-key issues and recommendations. Biofuels Bioprod Biorefin 9:606–619

Armstrong McKay DIA, Staal A, Abrams JF, Winkelmann R, akschewski B, Loriani S, Fetzer I, Cornell SE, Rockström J, Lenton TM (2022) Exceeding 1.5° C global warming could trigger multiple climate tipping points. Science 377: 1171-+

Azar C, Johansson DJA (2012) On the relationship between metrics to compare greenhouse gases – the case of IGTP, GWP and SGTP. Earth Syst Dynam 3:139–147

Balcombe P, Speirs JF, Brandon NP, Hawkes AD (2018) Methane emissions: choosing the right climate metric and time horizon. Environ Sci Processes Impacts 20:1323–1339

Brandão M, Levasseur A, Kirschbaum MUF, Weidema BP, Cowie AL, Jorgensen SV, Hauschild MZ, Pennington DW, Chomkhamsri K (2013) Key issues and options in accounting for carbon sequestration and temporary storage in life cycle assessment and carbon footprinting. Int J Life Cycle Assess 18:230–240

Brandão M, Kirschbaum MUF, Cowie AL, Hjuler SV (2019) Quantifying the climate change effects of bioenergy systems: comparison of 15 impact assessment methods. GCB Bioenergy 11:727–743

Breeze P (2017) Combined heat and power. Academic Press

Brundtland GH (1987) Our common future—call for action. Environ Conserv 14:291–294

BSI (British Standards Institute) (2011) Specification for the assessment of the life cycle greenhouse gas emissions of goods and services (PAS 2050). British Standards Institution, London

Caney S (2016) Climate change, intergenerational equity, and the social discount rate. In The Ethical Underpinnings of Climate Economics (pp. 53–76). Walsh A, Hormio S, Purves D (eds). Routledge. London

Cherubini F et al (2011) CO2 emissions from biomass combustion for bioenergy: atmospheric decay and contribution to global warming. GCB Bioenergy 3(5):413–426

Cherubini F, Bright RM, Strømman AH (2013) Global climate impacts of forest bioenergy: what, when and how to measure?. Environ Res Lett 8: Article #014049

Cherubini F, Fuglestvedt J, Gasser T, Reisinger A, Cavalett O, Huijbregts MA, Johansson DJ, Jørgensen SV, Raugei M, Schivley G, Strømman AH (2016) Bridging the gap between impact assessment methods and climate science. Environ Sci Policy 64:129–140

Clift R, Brandão M (2008) Carbon storage and timing of emissions – a note by Roland Clift and Miguel Brandao. Centre for Environmental Strategy (Working Paper 2(08))

Cross S, Welfle AJ, Thornley P, Syri S, Mikaelsson M (2021) Bioenergy development in the UK & Nordic countries: a comparison of effectiveness of support policies for sustainable development of the bioenergy sector. Biomass Bioenergy 144:Article#105887

Doney SC, Fabry VJ, Feely RA, Kleypas JA (2009) Ocean acidification: the other CO_2 problem. Ann Rev Mar Sci 1:169–192

Edwards MR, Trancik JE (2022) Consequences of equivalency metric design for energy transitions and climate change. Clim Change 175:Article #4

Ericsson N, Porsö C, Ahlgren S, Nordberg Å, Sundberg C, Hansson PA (2012) Time-dependent climate impact of a bioenergy system – methodology development and application to Swedish conditions. GCB Bioenergy 5:580–590

EU, (2009) Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC. Off J Eur Union 5:2009

European Commission (2010) International Reference Life Cycle Data System (ILCD) handbook—general guide for life cycle assessment—detailed guidance. Joint Research Centre—Institute for Environment and Sustainability. Luxembourg: Publications Office of the European Union

Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P (2008) Land clearing and the biofuel carbon debt. Science 319:1235–1238

Fearnside PM, Lashof DA, Moura-Costa P (2000) Accounting for time in mitigating global warming through land-use change and forestry. Mitig Adapt Strateg Glob Chang 5:239–270

Forster P, Storelvmo T, Armour K, Collins W, Dufresne JL, Frame D, Lunt DJ, Mauritsen T, Palmer MD, Watanabe M, Wild M, Zhang H (2021) The earth's energy budget, climate feedbacks, and climate sensitivity. In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou B (eds) Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, Cambridge, United Kingdom and New York, NY, USA, pp 923–1054

Fox-Kemper B, Hewitt HT, Xiao C, Aðalgeirsdóttir G, Drijfhout SS, Edwards TL, Golledge NR, Hemer M, Kopp RE, Krinner G, Mix A, Notz D, Nowicki S, Nurhati IS, Ruiz L, Sallée J-B, Slangen ABA, Yu Y (2021) Ocean, cryosphere and sea level change. In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou B (eds) Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp 1211–1362

Frischknecht R, Fantke P, Tschümperlin L, Niero M, Antón A, Bare J, Boulay AM, Cherubini F, Hauschild MZ, Henderson A, Levasseur A (2016) Global guidance on environmental life cycle impact assessment indicators: progress and case study. Int J Life Cycle Assess 21:429–442

Fuglestvedt JS, Berntsen TK, Godal O, Sausen R, Shine KP, Skodvin T (2003) Metrics of climate change: assessing radiative forcing and emission indices. Clim Change 58:267–331

Galik CS, Baker JS, Daigneault A, Latta G (2022) Crediting temporary forest carbon: retrospective and empirical perspectives on accounting options. Front For Glob 5:Article #933020

Garcia R, Alvarenga RAF, Huysveld S, Dewulf J, Allacker K (2020) Accounting for biogenic carbon and end-of-life allocation in life cycle assessment of multi-output wood cascade systems. J Clean Prod 275:Article #122795

Gibbs HK, Johnston M, Foley JA, Holloway T, Monfreda C, Ramankutty N, Zaks D (2008) Carbon payback times for crop-based biofuel expansion in the tropics: the effects of changing yield and technology. Environ Res Lett 3:Article #034001

Gillett NP, Matthews DH (2010) Accounting for carbon cycle feedbacks in a comparison of the global warming effects of greenhouse gases. Environ Res Lett 5:Article #034011

Groom B, Venmans F (2023) The social value of offsets. Nature 619:768–773

Helin T, Sokka L, Soimakallio S, Pingoud K, Pajula T (2013) Approaches for inclusion of forest carbon cycle in life cycle assessment - a review. GCB Bioenergy 5:475–486

Howard PH, Sterner T (2017) Few and not so far between: a meta-analysis of climate damage estimates. Environ Resour Econ 68:197–225

IEA (2021) Key world energy statistics 2021. International Energy Agency, Paris

IPCC (1990) Climate change: the Intergovernmental Panel on Climate Change scientific assessment. Cambridge University Press, Cambridge, UK, The Intergovernmental Panel on Climate Change (IPCC)

IPCC 2019, 2019 Refinement to the 2006 IPCC guidelines for national greenhouse gas inventories. Calvo Buendia E., Tanabe K., Kranjc A., Baasansuren J., Fukuda M., Ngarize S., Osako A., Pyrozhenko Y., Shermanau P., Federici, S. (eds). Published: IPCC, Switzerland.

IPCC, 2021: Summary for policymakers. In: Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte V., Zhai P., Pirani A., Connors S.L., Péan C., Berger S., Caud N., Chen Y., Goldfarb L., Gomis M.I., Huang M., Leitzell K., Lonnoy E., Matthews J.B.R., Maycock T.K., Waterfield T., Yelekçi O., Yu R., Zhou B. (eds.)]. Cambridge University Press. In Press.

IPCC, 2013: Annex III: glossary [Planton S. (ed.)]. In: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker T.F., Qin D. Plattner G.-K. Tignor M. Allen S.K. Boschung J. Nauels A. Xia Y., Bex V., Midgley P.M. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

ISO (2006a) 14040:2006. Environmental management -- life cycle assessment -- principles and framework. International Organization for Standardization (ISO). 20 p

ISO (2006b) 14044:2006. Environmental management -- life cycle assessment -- requirements and guidelines. International Organization for Standardization (ISO). 46 p

ISO (2015) 13065:2015 - Sustainability criteria for bioenergy. International Organization for Standardization (ISO). 57 p

ISO 14067 (2018) Greenhouse gases – Carbon footprint of products – requirements and guidelines for quantification and communication. International Organization for Standardization, Geneva

Jolliet O, Antón A, Boulay AM, Cherubini F, Fantke P, Levasseur A, McKone TE, Michelsen O, Canals LMI, Motoshita M, Pfister S (2018) Global guidance on environmental life cycle impact assessment indicators: impacts of climate change, fine particulate matter formation, water consumption and land use. Int J Life Cycle Assess 23:2189–2207

Joos F, Roth R, Fuglestvedt JS, Peters GP, Enting IG, von Bloh W, Brovkin V, Burke EJ, Eby M, Edwards NR, Friedrich T, Frölicher TL, Halloran PR, Holden PB, Jones C, Kleinen T, Mackenzie F, Matsumoto K, Meinshausen M, Plattner G-K, Reisinger A, Segschneider J, Schaffer G, Steinacher M, Strassmann K, Tanaka K, Zimmermann A, Weaver AJ (2013) Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. Atmos Chem Phys 13:2793–2825

Jørgensen SV, Hauschild MZ, Nielsen PH (2014) Assessment of urgent impacts of greenhouse gas emissions—the climate tipping potential (CTP). Int J Life Cycle Assess 19:919–930. https://doi.org/10.1007/s11367-013-0693-y

Jump AS, Peñuelas J (2005) Running to stand still: adaptation and the response of plants to rapid climate change. Ecol Lett 8:010–1020

Kendall A, Chang B, Sharpe B (2009) Accounting for time-dependent effects in biofuel life cycle greenhouse gas emissions calculations. Environ Sci Technol 43:7142–7147

Kirschbaum MUF (2003) Can trees buy time? An assessment of the role of vegetation sinks as part of the global carbon cycle. Clim Change 58:47–71

Kirschbaum MUF (2006) Temporary carbon sequestration cannot prevent climate change. Mitig Adapt Strateg Glob Chang 11:1151–1164

Kirschbaum MUF (2017) Assessing the merits of bioenergy by estimating marginal climate-change impacts. Int J Life Cycle Assess 22:841–852

Kirschbaum MUF (2014) Climate change impact potentials as an alternative to global warming potentials. Environ Res Lett 9:Article#034014

Kirschbaum MUF, Schlamadinger B, Cannell MGR, Hamburg SP, Karjalainen T, Kurz WA, Prisley S, Schulze E-D, Singh TP (2001) A generalised approach of accounting for biospheric carbon stock changes under the Kyoto protocol. Environ Sci Policy 4:73–85

Korhonen R, Pingoud K, Savolainen I, Matthews R (2002) The role of carbon sequestration and the tonne-year approach in fulfilling the objective of climate convention. Environ Sci Policy 5:429–441

Kumar A, Sah B, Singh AR, Deng Y, He XN, Kumar P, Bansal RC (2017) A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. Renew Sust Energ Rev 69:596–609

Lenton TM, Rockström J, Gaffney O, Rahmstorf S, Richardson K, Steffen W, Schellnhuber HJ (2019) Climate tipping points—too risky to bet against. Nature 575:592–595

Levasseur A, Lesage P, Margni M, Deschênes L, Samson R (2010) Considering time in LCA: dynamic LCA and its application to global warming impact assessments. Environ Sci Technol 44(8):3169–3174

Levasseur A, Cavalett O, Fuglestvedt JS, Gasser T, Johansson DJ, Jørgensen SV, Raugei M, Reisinger A, Schivley G, Strømman A, Tanaka K (2016) Enhancing life cycle impact assessment from climate science: review of recent findings and recommendations for application to LCA. Ecol Indic 7:163–174

Long SP, Ainsworth EA, Leakey ADB, Nösberger J, Ort DR (2006) Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. Science 312:1918–1921

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(4):044023

Matthews HD, Zickfeld K, Koch A, Luers A (2023) Accounting for the climate benefit of temporary carbon storage in nature. Nature Comm 14:Article#5485

Matustík J, Pohorely M, Koci V (2022) Is application of biochar to soil really carbon negative? The effect of methodological decisions in Life Cycle Assessment. Sci Total Environ 807:Article#151058

Millington R, Cox PM, Moore JR, Yvon-Durocher G (2019) Modelling ecosystem adaptation and dangerous rates of global warming. Emerging Top Life Sci 3:221–231

Moura Costa P, Wilson C (2000) An equivalence factor between CO2 avoidedemissions and sequestration–description andapplications in forestry. Mitig Adapt Strateg Glob Change 5(1):51–60
Müller-Wenk R, Brandão M (2010) Climatic impact of land use in LCA—carbon transfers between vegetation/soil and air. Int J Life Cycle Assess 15:172–182
Myhre G, Fuglestvedt JS, Berntsen TK, Lund MT (2011) Mitigation of short-lived heating components may lead to unwanted long-term consequences. Atmos Environ 45:6103–6106
Norby RJ, DeLucia EH, Gielen B, Calfapietra C, Giardina CP, King JS, Ledford J, McCarthy HR, Moore DJP, Ceulemans R, De Angelis P, Finzi AC, Karnosky DF, Kubiske ME, Lukac M, Pregitzer KS, Scarascia-Mugnozza GE, Schlesinger WH, Oren R (2005) Forest response to elevated CO_2 is conserved across a broad range of productivity. PNAS 102:18052–18056
Nordhaus WD (1997) Discounting in economics and climate change: an editorial comment. Clim Change 37:315–328
O'Hare M, Plevin RJ, Martin JI, Jones AD, Kendall A, Hopson E (2009) Proper accounting for time increases crop-based biofuels' greenhouse gas deficit versus petroleum. Environ Res Lett 4:Article#024001
Parisa Z, Marland E, Sohngen B, Marland G, Jenkins J (2022) The time value of carbon storage. For Policy Econ 144:Article#102840
Peck SC, Teisberg TJ (1994) Optimal carbon emissions trajectories when damages depend on the rate or level of global warming. Clim Change 28:289–314
Peters GP, Aamaas B, Lund MT, Solli C, Fuglestvedt JS (2011) Alternative "global warming" metrics in life cycle assessment: a case study with existing transportation data. Environ Sci Technol 45:8633–8641
Pingoud K, Ekholm T, Soimakallio S, Helin T (2016) Carbon balance indicator for forest bioenergy scenarios. GCB Bioenergy 8:171–182
Plattner GK, Stocker P, Midgley P, Tignor M (2009) IPCC expert meeting on the science of alternative metrics
Rahmstorf S, Perrette M, Vermeer M (2012) Testing the robustness of semi-empirical sea level projections. Clim Dyn 39:861–875
Reisinger A, Meinshausen M, Manning M (2011) Future changes in global warming potentials under representative concentration pathways. Environ Res Lett 6:Article#024020
Sarofim MC, Waldhoff ST, Anenberg SC (2017) Valuing the ozone-related health benefits of methane emission controls. Environ Resour Econ 66:45–63
Shindell DT, Fuglestvedt JS, Collins WJ (2017) The social cost of methane: theory and applications. Faraday Discuss 200:429–451
Shine KP, Fuglestvedt 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
Sterner T, Persson UM (2008) An even sterner review: introducing relative prices into the discounting debate. Rev Environ Econ Pol 2:61–76
Stohl A, Aamaas B, Amann M, Baker LH, Bellouin N, Berntsen TK, Boucher O, Cherian R, Collins W, Daskalakis N, Dusinska M, Eckhardt S, Fuglestvedt JS, Harju M, Heyes C, Hodnebrog O, Hao J, Im U, Kanakidou M, Klimont Z, Kupiainen K, Law KS, Lund MT, Maas R, MacIntosh CR, Myhre G, Myriokefalitakis S, Olivie D, Quaas J, Quennehen B, Raut JC, Rumbold ST, Samset BH, Schul M, Seland O, Shin KP, Skeie RB, Wang S, Yttri KE, Zhu T (2015) Evaluating the climate and air quality impacts of short-lived pollutants. Atmos Chem Phys 15:10529–10566
Tanaka K, Peters GP, Fuglestvedt JS (2010) Policy update: multicomponent climate policy: why do emission metrics matter? Carbon Manag 1:191–197
Tanaka K, Johansson DJA, O'Neill BC, Fuglestvedt JS (2013) Emission metrics under the 2 °C climate stabilization target. Clim Change 117:933–941
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:Article#eabf9020
UNEP and CCAC (2021) Global methane assessment: benefits and costs of mitigating methane emissions. Nairobi: United Nations Environment Programme
Wang JJ, Jing YY, Zhang CF, Zhao JH (2009) Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renew Sustain Energy Rev 13:2263–2278
Weitzman ML (2012) GHG targets as insurance against catastrophic climate damages. J Public Econ Theory 14:221–244
Weitzman ML (2013) A precautionary tale of uncertain tail fattening. Environ Resour Econ 55:159–173
Zetterberg L, Chen D (2014) The time aspect of bioenergy - climate impacts of solid biofuels due to biogenic carbon dynamics. GCB Bioenergy 7:785–796