CO$_2$-equivalence metrics for surface albedo change based on the radiative forcing concept: A critical review

Ryan M. Bright$^1$ and Marianne T. Lund$^2$

$^1$Department of Forest and Climate, Norwegian Institute of Bioeconomy Research (NIBIO), PO Box 115, 1431-Ås, Norway

$^2$Centre for International Climate Research (CICERO), 0349 Oslo, Norway

Correspondence to: Ryan M. Bright (ryan.bright@nibio.no)

Abstract. Management of Earth’s surface albedo is increasingly viewed as an important climate change mitigation strategy both on (Seneviratne et al., 2018) and off (Field et al., 2018; Kravitz et al., 2018) the land. Assessing the impact of a surface albedo change involves employing a measure like radiative forcing (RF) which can be challenging to digest for decision-makers who deal in the currency of CO$_2$-equivalent emissions. As a result, many researchers express albedo change ($\Delta\alpha$) RFs in terms of their CO$_2$-equivalent effects, despite the lack of a standard method for doing so, such as there is for emissions of well-mixed greenhouse gases (WMGHGs; e.g., IPCC AR5, Myhre et al. (2013)). A major challenge for converting $\Delta\alpha$ RFs into their “CO$_2$-equivalent” effects in a manner consistent with current IPCC emission metric approaches stems from the lack of a universal time-dependency following the perturbation (perturbation “lifetime”). Here, we review existing methodologies based on the RF concept with the goal of highlighting the context(s) in which the resulting CO$_2$-equivalent metrics may or may not have merit. To our knowledge this is the first review dedicated entirely to the topic since the first CO$_2$-eq. metric for $\Delta\alpha$ surfaced 20 years ago. We find that, although there are some methods that sufficiently address the time-dependency issue, none address or sufficiently account for the spatial disparity between the climate response to CO$_2$ emissions and $\Delta\alpha$ – a major critique of $\Delta\alpha$ metrics based on the RF concept (Jones et al., 2013). We conclude that considerable research efforts are needed to build consensus surrounding the RF “efficacy” of various surface forcing types associated with $\Delta\alpha$ (e.g., crop change, forest harvest, etc.), and the degree to which these are sensitive to the spatial pattern, extent, and magnitude of the underlying surface forcings.

1 Introduction

The albedo at Earth’s surface helps to govern the amount of solar energy absorbed by the earth system and is thus a relevant physical property shaping weather and climate (Cess, 1978; Hansen et al., 1984; Pielke Sr. et al., 1998). On average, Earth reflects about 30% of the energy it receives from the sun, of which about 13% may be attributed to the surface albedo (Stephens et al., 2015; Donohoe and Battisti, 2011). In recent years it has become the subject of increasing research interest amongst the scientific community, as measures to increase Earth’s surface albedo are increasingly viewed as an integral component of climate change mitigation and adaptation, both on (Seneviratne et al., 2018) and off (Field et al., 2018; Kravitz et al., 2018) the
Surface albedo modifications associated with large-scale carbon dioxide removal (CDR) can detract from the effectiveness of such mitigation strategies (Boysen et al., 2016). Like emissions of GHGs and aerosols, perturbations to the planetary albedo via perturbations to the surface albedo represent true external forcings of the climate system and can be measured in terms of changes to Earth’s radiative balance – or radiative forcings (Houghton et al., 1995). The radiative forcing (RF) concept provides a first-order means to compare surface albedo changes (henceforth \( \Delta \alpha \)) to other perturbation types, thus enabling a more comprehensive evaluation of human activities altering Earth’s surface (Houghton et al., 1995; Pielke Sr. et al., 2002).

The RF concept, however, is a backward-looking measure, does not express the actual temperature response to the perturbation, and requires complex calculations or modeling tools. To enable aggregation of emissions of different gases to a common scale, the concept of CO\(_2\)-equivalent emissions is commonly used in assessments, decision making, and policy frameworks. While initially introduced to illustrate the difficulties related to comparing the climate impacts of different gases, the field of emission metrics -- i.e., the methods to convert non-CO\(_2\) radiative constituents into their CO\(_2\)-equivalent effects – has evolved and presently includes a suite of alternative formulations, including the Global Warming Potential (GWP) adopted by the UNFCCC (O’Neill, 2000; Fuglestvedt et al., 2003; Fuglestvedt et al., 2010). Today, CO\(_2\)-equivalent metrics form an integral part of UNFCC emission inventories and climate agreements (e.g. The Kyoto Protocol) – in addition to the fields of Life Cycle Assessment (Heijungs and Guineév, 2012) and Integrated Assessment Modeling (O’Neill et al., 2016) – despite much debate around GWP as the metric of choice. As such, many researchers seek to convert RF from \( \Delta \alpha \) into a CO\(_2\)-equivalent effect, which is particularly useful in land use forcing research when perturbations to terrestrial carbon cycling often accompany the \( \Delta \alpha \). Although seemingly straight-forward at the surface, the procedure is complicated by two key fundamental differences between \( \Delta \alpha \) and CO\(_2\): additional CO\(_2\) becomes well-mixed within the atmosphere upon emission, and, the resulting atmospheric perturbation persists over long time scales and cannot be fully reversed by human interventions over those time scales. In other words, CO\(_2\)’s RF is both temporally- and spatially-extensive with the ensuring climate response being independent of the location of emission, whereas the RF and ensuing climate response following \( \Delta \alpha \) is more localized and can be fully reversed on short time scales.

These challenges have led researchers to adopt a variety of diverging methods for converting albedo-change RFs (henceforth \( RF_{\Delta \alpha} \)) into CO\(_2\)-equivalence. Unlike for conventional GHGs, however, there has been little concerted effort by the climate metric science community to build consensus or formalize a standard methodology for \( RF_{\Delta \alpha} \) (as evidenced by IPCC 4AR and AR5). Here, we review existing CO\(_2\)-equivalent metrics for \( \Delta \alpha \) and their underlying methods based on the RF concept. To our knowledge this is the first review dedicated to the topic since the first \( \Delta \alpha \) metric surfaced 20 years ago. Herein, we compare and contrast existing metrics both quantitatively and qualitatively, with the main goal of providing added clarity surrounding the context in which the proposed metrics have (de)merits. We start in Section 2 by providing an overview of the methods conventionally applied in the climate metric context for estimating radiative forcings following CO\(_2\) emissions and surface
albedo change. We then present the reviewed $\Delta \alpha$ metrics in Section 3 and systematically evaluate them quantitatively in Section 4 and qualitatively in Section 5. In Section 6 we review the interpretation challenges of a CO$_2$-eq. measure for $\Delta \alpha$ based on the RF concept, and in Section 7 we conclude with a discussion about the limitations and uncertainties of the reviewed metrics while providing recommendations and guidance for future application.

2 Radiative forcings from CO$_2$ emissions and surface albedo change

IPCC emission metrics are based on the stratospherically-adjusted RF at the tropopause in which the stratosphere is allowed to relax to the thermal steady state (Myhre et al., 2013; IPCC, 2001). Estimates of the stratospheric RF for CO$_2$ (henceforth $RF_{CO_2}$) are derived from atmospheric concentration changes imposed in global radiative transfer models (Myhre et al., 1998; Etminan et al., 2016). For shortwave RFs there is no evidence to suggest that the stratospheric temperature adjusts to a surface albedo change (at least for land-use and land cover change (LULCC), Smith et al. (2020); Hansen et al., (2005); Huang et al. (2020)) and thus the instantaneous shortwave flux change at TOA is typically taken as $RF_{\Delta \alpha}$, consistent with Myhre et al. (2013).

One of the major critiques of the instantaneous or stratospherically-adjusted RF is that they may be inadequate as predictors of the climate response (i.e., changes to near surface air temperatures, precipitation, etc.). The climate may respond differently to different perturbation types despite similar RF magnitudes, as feedbacks are not independent of the perturbation type (Hansen et al., 1997; Joshi et al., 2003). Alternative RF definitions that include tropospheric adjustments (Shine et al., 2003) or even land surface temperature adjustments (Hansen et al., 2005) have been proposed with the argument that such adjustments are more indicative of the type and magnitude of feedbacks underlying the climate response (Sherwood et al., 2015; Myhre et al., 2013). Alternatively, climate “efficacies” can be applied to adjust instantaneous or stratospherically-adjusted RF – where efficacy is defined as the temperature response to some perturbation type relative to that of CO$_2$. The implications of applying efficacies for spatially heterogenous perturbations like $\Delta \alpha$ are discussed further in Section 6.

2.1 CO$_2$ radiative forcings

Simplified expressions for the global mean $RF_{CO_2}$ (in W m$^{-2}$) due to a perturbation to the atmospheric CO$_2$ concentration are based on regression fits of radiative transfer model outputs (Myhre et al., 1998; Myhre et al., 2013):

$$RF_{CO_2}(\Delta C) = 5.35 \ln \left( \frac{C_0 + \Delta C}{C_0} \right)$$

where $C_0$ is the initial concentration and $\Delta C$ is the concentration change. Because of the logarithmic relationship between RF and CO$_2$ concentration, CO$_2$’s radiative efficiency – or the radiative forcing per unit change in concentration over a given
background concentration – decreases with increasing background concentrations. When $\Delta C$ is 1 ppm and $C_0$ is the current concentration, we may then refer to the solution of Eq. (1) as CO$_2$’s current global mean radiative efficiency – or $\alpha_{CO_2}$ (in W m$^{-2}$ ppm$^{-1}$).

Updates to the $RF_{CO_2}$ function (Eq. 1) were given in Etminan et al. (2016) where the constant 5.35 (or $RF_{2xCO_2}/ln(2)$) was replaced by an explicit function of CO$_2$, CH$_4$, and N$_2$O concentrations. However, this update is only important for very large CO$_2$ perturbations and is unnecessary to consider for emission metrics that utilize radiative efficiencies for small perturbations around present-day concentrations (Etminan et al., 2016).

For emission metrics, it is more convenient to express CO$_2$’s radiative efficiency in terms a mass-based concentration increase:

$$k_{CO_2} = \frac{\alpha_{CO_2} \varepsilon_{air} 10^6}{\varepsilon_{CO_2} M_{atm}}$$  \hspace{1cm} (2)

where $\alpha_{CO_2}$ is the radiative efficiency per 1 ppm concentration increase, $\varepsilon_{CO_2}$ is the molecular weight of CO$_2$ (44.01 kg kmol$^{-1}$), $\varepsilon_{air}$ is the molecular weight of air (28.97 kg kmol$^{-1}$), and $M_{atm}$ is the mass of the atmosphere ($5.14 \times 10^{18}$ kg). The solution of Eq. (2) thus yields CO$_2$’s global mean radiative efficiency with units in W m$^{-2}$ kg$^{-1}$.

The global mean radiative forcing over time following a 1 kg pulse emission of CO$_2$ can be estimated with an impulse-response function describing atmospheric CO$_2$ removal in time by Earth’s ocean and terrestrial CO$_2$ sinks:

$$RF_{CO_2}(t) = k_{CO_2} \int_{t=0}^{t} y_{CO_2}(t) dt$$ \hspace{1cm} (3)

where $y_{CO_2}$ is the multi-model mean CO$_2$ impulse-response function described in (Joos et al., 2013; Myhre et al., 2013) for a CO$_2$ background concentration of 389 ppmv, $t$ is the time step, and $k_{CO_2}$ is the radiative efficiency per kg CO$_2$ emitted upon the same background concentration (i.e., $1.76 \times 10^{15}$ W m$^{-2}$ kg$^{-1}$) which is assumed constant and time-invariant for small perturbations and for the calculation of emission metrics (Joos et al., 2013; Myhre et al., 2013). The pulse-response function ($y_{CO_2}$) comprises four carbon pools representing the combined effect of several carbon-cycle mechanisms rather than directly corresponding to individual physical processes. Although considered ideal for metric calculations in IPCC AR5, state-dependent alternatives exist in which the carbon cycle response is affected by rising temperature or CO$_2$ accumulation in the atmosphere (Millar et al., 2017).
For an emission (or removal) scenario, $RF_{CO_2}(t)$ is estimated from changes to atmospheric CO$_2$ abundance computed as a convolution integral between emissions (or removals) and the CO$_2$ impulse-response function:

$$RF_{CO_2}(t) = k_{CO_2} \int_{t=0}^{t} e(t') \gamma_{CO_2}(t - t') dt'$$

(4)

where $t$ is the time dimension, $t'$ is the integration variable, $e(t')$ is the CO$_2$ emission (or removal) rate (in kg).

### 2.2 Shortwave radiative forcings from surface albedo change

The time step of Eq. (3) is typically one year, thus it is convenient to utilize an annually averaged $RF_{\Delta \alpha}$ when deriving a CO$_2$-equivalent metric. Given the asymmetry between solar irradiance and the seasonal cycle of surface albedo in many extratropical regions, a more precise estimate of the annual $RF_{\Delta \alpha}$ is one based on the monthly (or even daily) $\Delta \alpha$ (Bernier et al., 2011).

The local annual mean instantaneous $RF_{\Delta \alpha}$ (in W m$^{-2}$) following monthly surface albedo changes (unitless) can be estimated with radiative kernels derived from global climate models [e.g., (Pendergrass et al., 2018; Smith, 2018; Soden et al., 2008; Block and Mauritsen, 2015)], although it should be pointed out that kernels are model- and state-dependent. Bright & O’Halloran (2019) recently presented a simplified $RF_{\Delta \alpha}$ model allowing greater flexibility surrounding the prescribed atmospheric state, given as:

$$RF_{\Delta \alpha}(t) = \frac{1}{12} \sum_{m=1}^{12} - SW_{\Delta \alpha, m, t}^{sfc} / T_{m, t}^{\alpha} \Delta \alpha_{m, t}$$

(5)

where $\Delta \alpha_{m, t}$ is a surface albedo change in month $m$ and year $t$, $SW_{\Delta \alpha, m, t}^{sfc}$ is the incoming solar radiation flux incident at surface level in month $m$ and year $t$, and $T_{m, t}^{\alpha}$ is the all-sky monthly mean clearness index (or $SW_{\alpha, m, t}^{sfc} / SW_{\alpha, t}^{toa}$; unitless) in month $m$ and year $t$.

It is important to reiterate that the $RF_{\Delta \alpha}$ defined with either Eq. (5) or GCM-based kernels strictly represents the instantaneous shortwave flux change at TOA and is not directly comparable to other definitions of $RF$ based on net (downward) radiative flux changes at TOA following atmospheric adjustments. A perturbation to $\Delta \alpha$ will result in a modification to the turbulent heat fluxes leading to radiative adjustments in the troposphere (Laguë et al., 2019; Huang et al., 2020; Chen and Dirmeyer, 2020). However, in the context of emission metrics, both $RF_{\Delta \alpha}$ and $RF_{CO_2}$ have merit given that they do not require coupled climate model runs of several years to compute.
3 Overview of CO₂-equivalent metrics for $RF_{\alpha\alpha}$

Over the past 20 years, a variety of metrics and their permutations have been employed to express $RF_{\alpha\alpha}$ as “CO₂-equivalence”, as evidenced from the 27 studies included in this review (Table 1).

Table 1. Studies included in this review.

<table>
<thead>
<tr>
<th>Study</th>
<th>Metric</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betts (2000)</td>
<td>$EESF$</td>
<td>AF = 0.5</td>
</tr>
<tr>
<td>Akbari et al. (2009)</td>
<td>$EESF$</td>
<td>AF = 0.55</td>
</tr>
<tr>
<td>Montenegro et al. (2009)</td>
<td>$EESF$</td>
<td>AF = 0.5</td>
</tr>
<tr>
<td>Thompson et al. (2009a)</td>
<td>$EESF$</td>
<td>AF = 0.5</td>
</tr>
<tr>
<td>Thompson et al. (2009b)</td>
<td>$EESF$</td>
<td>AF = 0.5</td>
</tr>
<tr>
<td>Muñoz et al. (2010)</td>
<td>$EESF$</td>
<td>AF based on C-cycle model and $TH = 20, 100, and 500$ yrs.</td>
</tr>
<tr>
<td>Menon et al. (2010)</td>
<td>$EESF$</td>
<td>AF = 0.55</td>
</tr>
<tr>
<td>Georgscu et al. (2011)</td>
<td>$EESF$</td>
<td>AF = 0.50</td>
</tr>
</tbody>
</table>
| Cherubini et al. (2012) | $GWP$  | Based on effective $RF$ estimated with a climate efficacy of 1.94\(^b\)
| Bright et al. (2012) | $GWP$  | TH = 20; 100; 500 yrs.                                              |
| Susca, T. (2012b)    | $\sum TDEE^a$ |                                                                    |
| Susca, T. (2012a)    | $\sum TDEE^a$ |                                                                    |
| Guest et al. (2013)  | $GWP$  |                                                                     |
| Zhao & Jackson (2014)| $EESF$ | AF = 0.5; Based on effective $RF$ estimated with a climate efficacy of 0.52\(^c\)
| Caiazzo et al. (2014)| $EESF$ | AF based on C-cycle model and $TH = 100$ yrs.                       |
| Singh et al. (2014)  | $GWP$  | TH = 100 yrs.                                                        |
| Bright et al. (2016) | $TDEE$ |                                                                     |
|                     | $\sum TDEE$ |                                                                    |
| Mykleby et al. (2017)| $EESF$ | AF based on C-cycle model and $TH = 80$ yrs.                        |
| Fortier et al. (2017)| $EESF$ | AF based on C-cycle model and $TH = 100$ yrs.                       |
| Carrer et al. (2018) | $EESF/TH$ | AF based on C-cycle model and $TH = 100$ yrs.                      |
| Carrer et al. (2018) | $GWP/TH$ | TH = 100 yrs.                                                        |
| Favero et al. (2018) | $EESF$ | AF based on C-cycle model and $TH = 100$ yrs.                       |
| Sieber et al. (2019) | $GWP$  | TH = 100 yrs.                                                        |
| Sieber et al. (2020) | $GWP$  | TH = 100 yrs.                                                        |
| Genesio et al. (2020)| $EESF$ | AF = 0.47                                                            |
Chiefly differentiating the methods behind the metrics shown in Table 1 – described henceforth – is how time is represented with respect to both the $\Delta \alpha$ and the reference gas (i.e., CO2) perturbations. Among the most common approaches is to relate $RF_{\Delta \alpha}$ to the $RF$ following a CO2 emission imposed on some atmospheric CO2 concentration background, but with a fraction of the emission instantaneously removed by Earth’s ocean and terrestrial CO2 sinks by an amount defined by one minus the so-called “airborne fraction” (AF) – or the growth in atmospheric CO2 relative to anthropogenic CO2 emissions (Forster et al., 2007).

This method – or the “emissions equivalent of shortwave forcing (EESF)” – was first introduced by Betts (2000) and may be expressed (in kg CO2-eq. m$^{-2}$) as:

$$EESF = \frac{RF_{\Delta \alpha}}{k_{CO2} A_E AF}$$

(6)

where $RF_{\Delta \alpha}$ is the local annual mean instantaneous $RF$ from a prescribed monthly $\Delta \alpha$ scenario (in W m$^{-2}$), $k_{CO2}$ is the global mean radiative efficiency of CO2 (e.g., Eq. (2); in W m$^{-2}$ kg$^{-1}$), $A_E$ is Earth’s surface area ($5.1 \times 10^{14}$ m$^2$), and AF is the airborne fraction. Because AF appears in the denominator in Eq. (6), the CO2 equivalent estimate will be highly sensitive to the choice of AF. Figure 1 plots AF since 1959 which, as can be seen, can fluctuate considerably over short time periods, ranging from a high of 0.81 in 1987 and low of 0.20 in 1992.
Figure 1. 1959-2018 airborne fraction (AF), defined here as the growth in atmospheric CO$_2$ – or the atmospheric CO$_2$ remaining after removals by ocean and terrestrial sinks – relative to anthropogenic CO$_2$ emissions (fossil fuels and LULCC). “Uncertainty” is defined as $AF \pm |BI|/E$, where $E$ are total anthropogenic CO$_2$ emissions and $BI$ is the budget imbalance – or $E$ minus the sum of atmospheric CO$_2$ growth and CO$_2$ sinks. Underlying data are from the Global Carbon Project (Friedlingstein et al., 2019).

185 More importantly, use of AF in Eq. (6) means that time-dependent atmospheric CO$_2$ removal processes following emissions are not explicitly represented. However, using the AF may be justifiable in some contexts – such as when $\Delta \alpha$ has no time dependency (on inter-annual scales). For example, the pioneering study by Betts (2000) – to which almost all CO$_2$-eq. literature for $\Delta \alpha$ may be traced (Table 1) – made use of AF when estimating CO$_2$-equivalence of $RF_{\Delta \alpha}$ because the research objective was to compare an albedo contrast between a fully grown forest and a cropland (i.e., $\Delta \alpha$) to the stock of CO$_2$ in the forest – a stock that had been assumed to accumulate over 80 yrs. which is the approximate time frame over which Earth’s CO$_2$ sinks
function to remove atmospheric CO₂ to a level conveniently similar to the AF. Had a transient or interannual Δα scenario been modeled, however, applying the EESF method at each time step of the scenario would have severely overestimated CO₂-equivalent emissions.

For this reason, Bright et al. (2016) argued that for time-dependent Δα scenarios (i.e., when Δα evolves over interannual time scales), the time-dependency of CO₂ removal processes (atmospheric decay) following emissions should be taken explicitly into account when estimating the effect characterized in terms of CO₂-equivalent emissions (or removals), thus proposing an alternate metric termed “time-dependent emissions equivalence” -- or TDEE:

\[
TDEE = A_E^{-1}k_{CO}_2^{-1}Y_{CO}_2^{-1}RF_{\Delta\alpha}^* \tag{7}
\]

where \(TDEE\) is a column vector of CO₂-equivalent emission (or removal) pulses (i.e., one-offs) with length defined by the number of time steps (e.g., years) included in the Δα time series (in kg CO₂-eq. m² yr⁻¹), \(RF_{\Delta\alpha}^*\) is a column vector of the local annual mean instantaneous \(RF_{\Delta\alpha}\) (in W m⁻²) corresponding to the Δα time series (or \(RF_{\Delta\alpha}(t)\)), and \(Y_{CO}_2\) is a lower triangular matrix with column (row) elements as the atmospheric CO₂ fraction decreasing (increasing) with time (i.e., \(y_{CO}_2(t)\)). The elements in vector \(TDEE\) thus give the CO₂-equivalent series of emission (or removal) pulses in time yielding the instantaneous \(RF_{\Delta\alpha}\) time profile (\(RF_{\Delta\alpha}(t)\)) corresponding to the temporally-explicit Δα scenario (\(Δα(t)\)). Summing all elements in \(TDEE\) (i.e., \(\sum TDEE\)) gives a measure of the accumulated CO₂-eq. emissions (removals) over time.

Time-dependent metrics like the well-known Global Warming Potential (GWP) (Shine et al., 1990; Rogers and Stephens, 1988) have also been applied to characterize \(Δα(t)\), which accumulates \(RF_{\Delta\alpha}(t)\) over time (temporally-discretized) up to some policy or metric time horizon (TH) which is then normalized to the temporally-accumulated radiative forcing following a unit pulse CO₂ emission over the same TH:

\[
GWP_{\Delta\alpha}(TH) = \frac{\sum_0^{t=TH}RF_{\Delta\alpha}(t)}{A_Ek_{CO}_2Y_{CO}_2(TH)} \tag{8}
\]

where \(TH\) is the temporal accumulation or metric time horizon. Because it is a cumulative measure, studies making use of GWP often divide by the number of time steps (TH) to approximate an annual CO₂ flux (e.g., Carrer et al., (2018)). The result of Eq. (8) can be interpreted a single pulse of CO₂ (in kg CO₂-eq. m²) at \(t = 0\) which approximately gives the same time-integrated RF at TH as that from a time-dependent Δα scenario.
3.1 Metric permutations

Some studies have applied various permutations of the three metrics presented above. For instance, some have applied definitions of the airborne fraction (AF) based on CO$_2$’s pulse-response function (i.e., $y_{CO_2}(t)$) when estimating $EESF$ on the grounds that the analysis required a long and forward-looking time perspective (Caiazzo et al., 2014; Favero et al., 2018; Mykleby et al., 2017; Muñoz et al., 2010; Sciusco et al., 2020). A consequence is that the magnitude of the CO$_2$-eq. calculation is highly sensitive to the subjective choice of the $TH$ chosen as the basis for the AF (typically taken as the mean atmospheric fraction for the period up to $TH$ – or $TH^{-1} \int_{t=0}^{t=TH} y_{CO_2}(t) \, dt$). Other permutations include the normalization of $EESF$ or $GWP(TH)$ by $TH$ to arrive at a uniform time series of CO$_2$-eq. pulses (Carrer et al., 2018), or the summing of $TDEE$ up to $TH$ to obtain a CO$_2$-eq. stock perturbation measure (Bright et al., 2020; Bright et al., 2016).

3.2 Metric decision tree

Their relative merits and drawbacks (further discussed in Sections 4 & 5) notwithstanding, Figure 2 presents a decision tree for differentiating between the reviewed $\Delta \alpha$ metrics presented heretofore.

**Figure 2. Decision-tree for $\Delta \alpha$ metrics included in this review.**
A principle differentiator after the time-dependency distinction is whether CO₂-equivalence corresponds to a single emission (removal) pulse or a time series of multiple CO₂-equivalent emission (removal) pulses. For the time-dependent metrics (Fig. 2, right branch), further distinction can be made according to whether the CO₂-equivalent effect is an instantaneous effect (in the case of the time-series measures) and whether IPCC compatibility is desired by the practitioner (in the case of the single pulse measures). By “IPCC compatibility”, we mean that the metric computation and physical interpretation aligns with emission metrics presented in previous IPCC climate assessment reports and IPCC good practice guidelines for national emission inventory reporting. A second or alternate distinction can be made for the time-dependent and single pulse measures according to whether the CO₂-equivalent effect corresponds to the present (\( t = 0 \)) or the future (\( t = TH \)).

### 3.3 Δα vs. emission metrics

All metric application entails subjective user decisions, such as type of metric (i.e., instantaneous vs. accumulative; scalar vs. time-series) and time horizon for impact evaluation. CO₂-equ. metrics for \( \Delta \alpha \) require additional decisions by the practitioner affecting both their transparency and uncertainty, which are highlighted in Table 3.

<table>
<thead>
<tr>
<th>Radiative forcing agent</th>
<th>RF Metric</th>
<th>Initial Perturbation (emission or ( \Delta \alpha ))</th>
<th>Perturbation time-dependency</th>
<th>RF model</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHGs</td>
<td>GWP</td>
<td>Unit pulse</td>
<td>IPCC</td>
<td>IPCC</td>
</tr>
<tr>
<td>( \Delta \alpha ), time-dependent</td>
<td>( T\text{D}EE; \sum T\text{D}EE; GWP )</td>
<td>User defined</td>
<td>User-defined</td>
<td>User defined</td>
</tr>
<tr>
<td>( \Delta \alpha ), time-independent</td>
<td>( E\text{E}SF )</td>
<td>User defined</td>
<td>None</td>
<td>User defined</td>
</tr>
</tbody>
</table>

First among these is the need to quantify the initial physical perturbation (i.e., \( \Delta \alpha \)) which is irrelevant for IPCC emission metrics where the initial perturbation is a unit pulse emission. For \( \Delta \alpha \) metrics, uncertainty surrounding estimates of the initial (or reference) and perturbed albedo states is introduced. Second, for the time-dependent metrics (Table 3, second row) additional uncertainty is introduced by the metric practitioner when defining the time-dependency of the \( \Delta \alpha \) perturbation, which may be contrasted to IPCC emission metrics where the temporal evolution of the perturbation (i.e., atmospheric concentration change) is pre-defined (or rather, lifetimes and decay functions of the various forcing agents). Likewise, the RF models employed to give radiative efficiencies for various forcing agents are pre-defined by the IPCC -- models having origins linked to standardized experiments employing rigorously evaluated radiative transfer and/or climate models, which may be
contrasted to the models applied to estimate $RF_{\Delta \alpha}$ which can vary widely in their complexity and uncertainty (for a brief review of these, see Bright & O’Halloran (2019)).

4 Quantitative metric evaluation

The metrics presented in Section 3 are systematically compared quantitatively henceforth by deriving them for a set of common cases, starting first with the metrics applied to yield a time series of CO$_2$-eq. pulse emissions (or removals) in time. For all calculations, the assumed climate “efficacy” (Hansen et al., 2005) – or the global climate sensitivity of $RF_{\Delta \alpha}$ relative to $RF_{CO_2}$ – is 1.

4.1 CO$_2$-eq. pulse time series measures

Let us first consider a geoengineering case where a rooftop is painted white during the first year of a 100-year simulation which increases the annual mean surface albedo (Fig. 3 A) for the full simulation period resulting in a constant negative $RF_{\Delta \alpha}$ (Fig. 3 B). The objective is to estimate a series of CO$_2$-eq. fluxes associated with $RF_{\Delta \alpha}(t)$. 
Figure 3. Example application of metrics yielding a complete time series of CO$_2$-eq. pulse emissions or removals. A) Time-dependent $\Delta \alpha$ scenario ("$\Delta \alpha$" = $\alpha_{\text{new}}$ - $\alpha_{\text{old}}$); B) The corresponding global mean instantaneous shortwave radiative forcing over time ($RF_{\Delta \alpha}(t)$); C) The derived metrics TDEE, GWP(100)/100, and EESF/100 for a range of airborne fractions (AF), ; D) The reconstructed $RF_{\Delta \alpha}(t)$ based on the values shown in panel C) and Eq. (4). Note that the legend in panel D) also applies to panel C).

Figure 3 C presents the results after applying the relevant metrics to the common $RF_{\Delta \alpha}$ and time-dependent $\Delta \alpha$ scenario. To assess their fidelity or "accuracy", the resulting CO$_2$-eq. series of annual CO$_2$ pulses (in this case removals) are used with Eq. (4) to re-construct the $RF_{\Delta \alpha}$ time profile (Fig. 3 B). Unsurprisingly, annual CO$_2$-eq. removals estimated with the TDEE approach (Fig. 3 C) exactly reproduce $RF_{\Delta \alpha}$, and thus the two red curves shown in Figures 2B and D are identical (note the difference in scale). Figure 3 C illustrates the sensitivity of the EESF-based measure derived using an AF of 0.47 (mean of
the last seven years based on the most recent global carbon budget, e.g. Friedlingstein et al., (2019); Fig. 1) relative to a broad range of AF values (note that the result obtained using AF = 1 is referred to as the “time-independent emissions equivalent (TIEE)” presented in Bright et al. (2016)). Irrespective of the AF value that is chosen, when applied in a forward-looking analysis utilizing a time-dependent Δα scenario with a time horizon of 100-yrs., the EESF approach underestimates the magnitude of the annual CO₂-eq. pulse occurring in the short-term relative to TDEE (Fig. 3 C) and hence also RF_{Δα} in the short-term (Fig. 3 B & D). For higher AF values, annual CO₂-eq. removals estimated using the EESF-based approach will underestimate RF_{Δα} at each time step (Fig. 3 D), despite the higher magnitude CO₂-eq. estimate (relative to TDEE) seen in the longer-term (Fig. 3 C).

For TH = 100, the EESF-based estimate will always be lower in magnitude in the short-term and higher in magnitude in the longer-term relative to TDEE (Fig. 3 C). The same is also true for the annual GWP-based CO₂-eq. estimate, although at least the reconstructed RF_{Δα} value at t = TH will always be identical to the actual RF_{Δα} value at t = TH (Fig. 3 D). In general, EESF- and GWP-based estimates of annualized CO₂-eq. emissions (or removals) are sensitive to the chosen TH and will always exceed (in magnitude) estimates based on TDEE. This is demonstrated in Figure 4.

![Sensitivity of CO₂-eq. flux to TH](https://doi.org/10.5194/acp-2020-1109)

**Figure 4.** Magnitude of the annual CO₂-eq. emission (removal) pulse as a function of metric TH for the EESF and GWP measures relative to TDEE which is insensitive to TH.
The EESF-based estimate in this example is higher (in magnitude) than the GWP-based estimate because the assumed AF of 0.47 is lower than the mean atmospheric fraction following pulse emissions (i.e., $y_{CO_2}(t)$) over the range of time horizons shown (the mean atmospheric fraction at $TH = 100$ when applying Joos et al. (2013) function is 0.53). In contrast to the EESF- and GWP-based approaches, the magnitude of the annual CO$_2$-eq. removals estimated with TDEE is insensitive to the chosen $TH$.

4.2 Single CO$_2$-eq. pulse measures

Turning our attention now to measures yielding a single CO$_2$-eq. emission or removal pulse, let us now consider a forest management case where managers are considering harvesting a deciduous broadleaf forest to plant a more productive evergreen needleleaved tree species. It is known that when the evergreen needleleaf forest matures in 80-years its mean annual surface albedo will be about 2% lower than the deciduous broadleaved forest. The corresponding annual local $RF_{\Delta a}$ at year 80 is 1.8 W m$^{-2}$, and we wish to associate a CO$_2$-equivalence to this value in order to weigh it against an estimate of the total CO$_2$ stock difference between the two forests after 80 years (i.e., $TH = 80$). Assuming we have no information about how the albedo evolves a priori in the two forests before year 80, we have no choice but to apply the EESF measure.

Figure 5 presents the CO$_2$-eq. estimate based on EESF for an AF range of 0.1 – 1, shown together with an estimate in which the AF is obtained using the mean fraction of CO$_2$ remaining in the atmosphere at 80 years following an emission pulse, obtained from the latest IPCC impulse-response function ($y_{CO_2}(t)$), and with the highest and lowest airborne fractions of the last 7 years.

![Figure 5. Sensitivity of EESF to the airborne fraction (AF).](https://doi.org/10.5194/acp-2020-1109)
Figure 5 illustrates EESF’s sensitivity to the assumed AF. For instance, EESF with AF = 0.3 is double that estimated with AF = 0.6 – a normal AF range for the past 60 years (Fig. 1). EESF estimated using AF from 2015 (Fig. 5, green diamond) is 44% lower than EESF using AF from the previous year (Fig. 5, magenta diamond). If surface albedo is ever to be included in forestry decision-making – as some have proposed (Thompson et al., 2009a; Lutz and Howarth, 2014) – the subjective choice of the AF becomes problematic given this large sensitivity. For instance, if the decision-making basis in this example depends on the net of the CO$_2$-eq. of $\Delta \alpha$ and a difference in forest CO$_2$ stock of 4.5 kg CO$_2$ m$^{-2}$, adopting an AF of 0.5 might lead to a decision to plant the new tree species given that the stock difference would exceed the EESF estimate (i.e., CO$_2$ sinks dominate), whereas adopting an AF of 0.4 might lead to a decision to forego the planting given that the CO$_2$-eq. of $\Delta \alpha$ would exceed the stock difference (i.e., surface albedo dominates).

Now let’s assume the decision maker does have insight into how the surface albedos of both forest types will evolve over the full rotation period. In this new example, harvesting the deciduous broadleaf forest to plant an evergreen needleleaf species will first increase the surface albedo in the short-term, yet as the evergreen needleleaf forest grows and tree canopies begin to close and mask the surface, the albedo difference ($\Delta \alpha$) reverts to negative and stays negative for the remainder of the rotation. This results in an annual mean local $RF_{\Delta \alpha}(t)$ profile that is first negative and then positive, which is depicted in Figure 6 A (blue solid curve).
Figure 6. Example application of metrics yielding a single CO₂-eq. emission (or removal) pulse following a hypothetical forest tree species conversion. A) $RF_{Δα}(t)$ and corresponding $TDEE$ (left y-axis) and the time-integrated or cumulative $RF_{Δα}(t)$ and $RF_{CO₂}(t)$ following a 1 kg pulse emission (right y-axis); B) $EESF$ estimated for the $Δα$ (and $RF_{Δα}$) occurring at $TH = 80$ shown in relation to $GWP(TH)$ and $\sum TDEE$ estimated for all $THs$. 

Converting the $RF_{Δα}(t)$ time profile first to a time series of CO₂-eq. emission/removal pulses (i.e., $TDEE$, Fig. 6 A, dashed blue curve) then summing to year 80 gives a measure of the total quantity of CO₂-eq. emitted (or removed) at year 80 – or $\sum TDEE$ (Fig. 6 B, blue curve). $\sum TDEE$ thus “remembers” the negative $Δα$ in the early phases of the rotation period (short-term), leading to a lower CO₂-eq. estimate at year 80 relative to $EESF$ estimates computed with airborne fractions of 0.66 and lower. Similarly, the $GWP$-based estimate “remembers” the negative $Δα$ occurring in the short-term; however, $GWP$ is a normalized measure, meaning that the time-evolving radiative effects of $Δα$ and CO₂ are first computed independently from...
each other prior to the CO$_2$-equivalence calculation, whereas for TDEE (and hence $\sum$TDEE) CO2-equivalence depends directly on the time-evolving radiative effect of $\Delta \alpha$. Framed differently, $\sum$TDEE remembers prior CO$_2$-eq. fluxes yielding the radiatively equivalent effect of the time-dependent $\Delta \alpha$ scenario, whereas the “memories” of $RF_{\Delta \alpha}$ and $RF_{CO2}$ underlying the GWP-based CO$_2$-equivalent estimate are first considered in isolation (Fig. 6 A, red curves). Hence the GWP-based CO$_2$-eq. estimate in this example is much lower than the $\sum$TDEE-based estimate since the temporally-accumulated $RF_{CO2}$ following a unit pulse emission at $t = 0$ (or $CRF_{CO2}$, also known as $AGWP_{CO2}$) is significantly larger than the temporally-accumulated $RF_{\Delta \alpha}$ (or $CRF_{\Delta \alpha}$) when evaluated at $TH = 80$ yrs.

When scalar metrics are required, Figure 6 illustrates the large inherent risk of applying a static measure like EESF to characterize $\Delta \alpha$ in dynamic systems. Moreover, for dynamic systems in which $\Delta \alpha$’s time-dependency is defined a priori, Fig. 6 illustrates the importance of clearly defining the time horizon at which the physical effects of $\Delta \alpha$ and CO$_2$ are to be compared: GWP gives an effect measured in terms of a present-day CO$_2$ emission (or removal) pulse, while $\sum$TDEE gives an effect measured in terms of a future CO$_2$ emission (or removal). In other words, internal consistency between the ecological and metric time horizons is relaxed with GWP but preserved with $\sum$TDEE.

5 Qualitative metric evaluation

The reviewed metrics and underlying methods for converting shortwave radiative forcings from $\Delta \alpha$ (i.e., $RF_{\Delta \alpha}$) into their CO$_2$ equivalent effects – summarized in Table 4 – can primarily be differentiated by the physical interpretation of the derived measure and by whether or not a time-dependency (inter-annual) for $\Delta \alpha$ was defined a priori.

### Table 4. Overview of distinguishing attributes, methodological differences, drawbacks, and merits of the six metrics included in this review.

<table>
<thead>
<tr>
<th>$\Delta \alpha$ Metric</th>
<th>CO$_2$-equivalence Interpretation</th>
<th>Time-dependent $\Delta \alpha$ scenario</th>
<th>Drawbacks</th>
<th>Merits</th>
</tr>
</thead>
<tbody>
<tr>
<td>EESF</td>
<td>Single pulse</td>
<td>No</td>
<td>- Sensitive to choice of airborne fraction (AF) - Not forward-looking - No carbon cycle dynamics</td>
<td>- Easy to apply; No need to define a $\Delta \alpha$ scenario a priori</td>
</tr>
</tbody>
</table>

https://doi.org/10.5194/acp-2020-1109

Preprint. Discussion started: 10 December 2020
© Author(s) 2020. CC BY 4.0 License.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
<th>Uniformity</th>
<th>Sensitivity to TH</th>
<th>Sensitivity to CO₂</th>
<th>Time Dependence</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EESF/TH</strong></td>
<td>Series of uniform pulses</td>
<td>No</td>
<td>Same as above</td>
<td>CO₂-eq. series does not reproduce $RF_{\Delta \alpha}(t)$</td>
<td>Easy to apply</td>
<td></td>
</tr>
<tr>
<td><strong>TDEE</strong></td>
<td>Series of non-uniform pulses</td>
<td>Yes</td>
<td>Not scalar</td>
<td>CO₂-eq. series reproduces $RF_{\Delta \alpha}(t)$</td>
<td>Can be compared to an emission scenario</td>
<td>Insensitive to TH</td>
</tr>
<tr>
<td><strong>ΣTDEE</strong></td>
<td>Accumulation of a series of non-uniform pulses</td>
<td>Yes</td>
<td>Cannot be compared to a CO₂ pulse of the present</td>
<td>Compatible with policy targets based on cumulative emissions</td>
<td>Insensitive to TH</td>
<td></td>
</tr>
<tr>
<td><strong>GWP</strong></td>
<td>Single pulse</td>
<td>Yes</td>
<td>Sensitive to TH</td>
<td>Well-known; IPCC conformity</td>
<td>Compatible with IPCC assessments and UNFCCC accounting conventions</td>
<td></td>
</tr>
<tr>
<td><strong>GWP(TH)</strong></td>
<td>Series of uniform pulses</td>
<td>Yes</td>
<td>Sensitive to TH</td>
<td>CO₂-eq. series does not reproduce $RF_{\Delta \alpha}(t)$ except at $t = TH$</td>
<td>GWP method is well-known</td>
<td></td>
</tr>
</tbody>
</table>

For cases when $\Delta \alpha$’s time-dependency is not known or defined a priori, the EESF measure is the only applicable measure, although it was shown here to be highly sensitive to the value chosen to represent CO₂’s airborne fraction (AF; Fig. 5) – a key input variable taking on a wide range of values depending on how it was defined. In general, when AF is defined according to historical accounts of global carbon cycling, its value is prone to large fluctuations across short time scales (Fig. 1) due to natural variability in the global carbon cycle (Ciais et al., 2013). When defined as the fraction of CO₂ remaining in the atmosphere following a pulse emission – as would be obtained from a simple carbon cycle model (i.e., a CO₂ impulse-response function) -- its value depends on the time horizon chosen and underlying model representation of atmospheric removal processes (i.e., time-constants). Use of the latter definition of AF affixes a forward-looking time-dependency to the EESF measure which is inconsistent with the definition of $\Delta \alpha$ and adds subjectivity (i.e., the choice in TH). Basing the AF on global carbon budget reconstructions would at least preserve some element of objectivity, although given the measure’s sensitivity to

---

\[AF = TH^{-1} \int_{t=0}^{t=TH} y_{CO_2}(t) \, dt\]
AF it would be prudent to compute the measure for a range of AFs (i.e., as constrained by the observational record) in effort to boost transparency. Forgoing the use of an AF altogether would eliminate all subjectivity, as has been suggested elsewhere (Bright et al., 2016).

For cases involving a time-dependent $\Delta \alpha$ scenario that is defined *a priori*, forward-looking measures are identified whose methodological differences give rise to different interpretations of “CO$_2$-equivalence” (Table 4). For example, the $GWP$ measure can be interpreted as a unit pulse of CO$_2$ emitted today yielding the accumulated radiative forcing of the $\Delta \alpha$ scenario at $TH$ years into the future. The $TDEE$ measure, on the other hand, can be interpreted as a complete time series of emission pulses (i.e., a complete emission scenario) yielding the instantaneous radiative forcing of the $\Delta \alpha$ scenario. When summed to $TH$, the latter (as $\Sigma TDEE$) also informs about the instantaneous radiative effect of the $\Delta \alpha$ scenario, albeit in terms of an accumulated CO$_2$ emission of the *future* as opposed to an emission pulse at the present (i.e., $GWP$). Both $GWP$ and $\Sigma TDEE$ have merit, and it would be unfair to claim one is superior to the other. For instance, if the objective is to weigh the effect of a $\Delta \alpha$ scenario against a basket of GHG emissions evaluated at the present time following some land use change (as is done when compiling national GHG emission inventories), the $GWP$ measure is the most suitable measure given its conformity to established reporting methods, accounting standards, or decision-support tools such as Life Cycle Assessment (e.g., Cherubini et al. (2012); Sieber et al. (2020)). On the other hand, if the objective is to weigh the effect of a $\Delta \alpha$ scenario against cumulative CO$_2$ emissions in the future – as would be required to evaluate the mitigation potential of land use policies in the context of emission budgets or policy targets based on cumulative emissions (e.g. Bright et al. (2020)) – the $\Sigma TDEE$ is the more suitable measure.

The permutations of $GWP$ and $EESF$ applied to arrive at a time series of CO$_2$-eq. pulses -- $GWP(TH)/TH$ and $EESF/TH$ -- have little merit on the grounds that the resulting series does not reproduce $RF_{\Delta \alpha}(t)$ (Fig. 3). The $TDEE$ approach was proposed to overcome this limitation, although it should be stressed that – like $GWP(TH)/TH$ – its derivation requires that a time-dependent $\Delta \alpha$ scenario be defined *a priori*, which adds uncertainty and may not always be possible.

## 6 Spatial disparity in climate response between CO$_2$ emissions and $\Delta \alpha$ perturbations

The climate (i.e., temperature) response to a $\Delta \alpha$ perturbation either isolated (e.g., Jacobson and Ten Hoeve, (2012)) or as part of LULCC (e.g., (Pongratz et al., 2010; Betts, 2001)) is highly heterogeneous in space, the magnitude and extent of which depends on its location (Brovkin et al., 2013; de Noblet-Ducoudré et al., 2012). This is because the response pattern of climate feedbacks has a strong spatial dependency – feedbacks are generally larger at higher latitudes due to higher energy budget sensitivity to clouds, water vapor, and surface albedo, which generally increases the effectiveness of RF in those regions (Shindell et al., 2015). This is in contrast to CO$_2$ emissions where both RF and the temperature response are more homogeneous in space (Hansen and Nazarenko, 2004; Hansen et al., 2005; Myhre et al., 2013). This has caused some researchers to question
the utility of a CO2-eq. measure for $\Delta \alpha$ (Jones et al., 2013) or encouraged others to look for solutions or further methodological refinements. For instance, some researchers (e.g., Cherubini et al. (2012); Zhao & Jackson (2014)) have applied climate “efficacies” – or the climate sensitivity of a forcing agent relative to CO2 (Joshi et al., 2003; Hansen et al., 2005) – to adjust $RF_{\Delta \alpha}$ prior to the CO2-eq. calculation. Such adjustments recognize that the temperature response to $RF$ depends on the geographic location, extent, and type of underlying forcing associated with the $\Delta \alpha$ (e.g., land use/land cover change (LULCC), white-roofing, etc.) which can be co-associated with other perturbations (Table 5) like those arising from changes to vegetative physical properties (for the LULCC case) which can modify the partitioning of turbulent heat fluxes above and beyond the purely radiatively-driven change (Davin et al., 2007; Bright et al., 2017).

Using a climate efficacy to adjust $RF_{\Delta \alpha}$, however, is not without its drawbacks. A first and obvious drawback is that efficacies are climate model dependent (Hansen et al., 2005; Smith et al., 2020; Richardson et al., 2019). Climate models vary in their underlying physics, which is evidenced by the large spread in CO2’s climate sensitivity across CMIP6 models (Meehl et al., 2020; Zelinka et al., 2020). A second drawback is that climate sensitivities for certain forcing agents like $\Delta \alpha$ are tied to experiments that differ largely in the way forcings have been imposed in time and space. Both drawbacks contribute to large uncertainties in the choice of efficacy for $\Delta \alpha$. The latter drawback is especially problematic since the $\Delta \alpha$ perturbation is often accompanied by perturbations to other surface properties and fluxes (Table 5) having large spatial- and temporal dependencies. The turbulent heat flux perturbations that accompany a net radiative flux change at the surface affect atmospheric temperature and humidity profiles (Bala et al., 2008; Modak et al., 2016; Schmidt et al., 2012; Kravitz et al., 2013), causing the atmosphere to adjust to a new state resulting in a net radiative flux change at TOA that extends beyond the instantaneous shortwave radiative flux change (i.e., $RF_{\Delta \alpha}$).

Table 5. Differences in surface property and flux perturbations between geoengineering-type forcings involving non-vegetative solar radiation management (SRM) and forcings from LULCC, land management change (LMC), or forest management change (FMC). “$\Delta r_a$” = change to bulk aerodynamic resistance; “$\Delta r_s$” = change to bulk surface resistance; “$\Delta \lambda(E)$” = latent heat flux change from a change to evaporation; “$\Delta \lambda(E+T)$” = latent heat flux change from a change to both evaporation and transpiration; “$\Delta H$” = sensible heat flux change.

<table>
<thead>
<tr>
<th>Forcing type</th>
<th>Surface property perturbation</th>
<th>Surface flux perturbation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geoengineering (non-veg. SRM)</td>
<td>$\Delta \alpha$</td>
<td>$\Delta \lambda(E)$, $\Delta H$</td>
</tr>
<tr>
<td>LULCC; LMC; FMC</td>
<td>$\Delta \alpha$, $\Delta r_a$, $\Delta r_s$</td>
<td>$\Delta \lambda(E+T)$, $\Delta H$</td>
</tr>
</tbody>
</table>

For example, the efficacy of LULCC forcing across the six studies reviewed by Bright et al. (2015) ranged from 0.5 to 1.02 owed to differences in model set-up (e.g., fixed SST vs. slab vs. dynamic ocean), differences in the spatial extent and magnitude of the imposed LULCC forcing (e.g., historical transient vs. idealized time slice), as well as the LULCC definition (i.e., the type of LULCC that was included in the study such as only aff/deforestation vs. all LULCC). Even when controlling for
differences in experimental design (e.g., CMIP protocols), the climate efficacy of historical LULCC has been found to vary considerably in both sign and magnitude (c.f. Figure 8, Richardson et al. 2019), which is more likely attributed to the larger spread in effective radiative forcing (ERF) for LULCC than for CO₂. For instance, Smith et al. (2020) report a standard deviation of 6% in the ERF of CO₂ (4xabrupt) across 13 GCMs/ESMs participating in RFMIP in contrast to 175% for LULCC, although it should be kept in mind that the ERF is weak for LULCC thus relative differences become large.

An additional drawback and source of uncertainty underlying efficacies is related to differences in their definition. Differences in definition can stem from either different definitions of RF itself or differences in the definition of the temperature response per unit RF (Richardson et al., 2019; Hansen et al., 2005). Regarding the latter, most base the temperature response for CO₂ on the equilibrium climate sensitivity (ECS) for a CO₂ doubling, although good arguments have been made for using the transient climate response (TCR) instead, particularly for short-lived forcing agents (Marvel et al., 2016; Shindell, 2014). The temperature response for the forcing agent of interest is rarely taken as the equilibrium response although there are some exceptions (e.g. “Eₐ” in Richardson et al. 2019 which is based on climate feedback parameters obtained from ordinary least square regressions). Efficacies are also sensitive to the definition of RF (Richardson et al., 2019; Hansen et al., 2005). For example, the efficacy of sulfate forcing (5×SO₄) has recently been shown to vary from 0.94 to 2.97 depending on whether RF is based on the net radiative flux change at TOA from fixed SST experiments or the instantaneous shortwave flux change at tropopause (Richardson et al., 2019).

Ideally, CO₂-eq. metrics based on the RF concept should be based on an RF definition yielding efficacies approaching unity for a broad range of forcing types. Although there is currently no consensus here, strong arguments have been made for RF definitions based on the net radiative flux change at TOA resulting from fixed SST experiments with GCMs/ESMs (i.e., “Fₛ” in Hansen et al. 2005; “ERFₛ” in Richardson et al. 2019), since such definitions yield efficacies approaching unity for a broad range of forcing types. However, for most Δα metric practitioners it is not feasible to quantify atmospheric adjustments and hence the ERF. Efficacies compatible with RFΔα (instantaneous ΔSW at TOA) could be the more feasible option for metric calculations, but broad consensus would need to be established first surrounding appropriate efficacy values for different forcing types associated with the Δα perturbation (Table 5). This is especially true for forcings involving changes to the biophysical properties of vegetation – such as LULCC, forestry, etc. – since these are constructs representing a seemingly myriad combination of biophysical perturbations acting on non-radiative controls (i.e., Δrₛ and Δrₛ) of the surface energy balance. Building consensus for efficacies applicable to geoengineering-type forcings where the only physical property perturbed is the surface albedo (e.g., white roofing, sea ice brightening, etc.) would be less challenging since the confounding perturbations to Δrₛ and Δrₛ and hence to the turbulent heat fluxes are removed. Nevertheless, irrespective of whether broad scientific consensus can be reached surrounding efficacies suitable for Δα metrics, additional responsibility would always be imposed on the metric practitioner to ensure that the chosen efficacy aligns with the forcing type underlying the RFΔα.
7 Discussion

In this review, we quantitatively and qualitatively reviewed metrics (methods) to characterize \( RF_{\Delta a} \) in terms of a \( CO_2 \)-equivalent effect. Differences among metrics could be attributed to the different ways of dealing with the time-dependency of \( RF_{CO_2} \), which to a large extent was determined by whether a time-dependency for \( \Delta a \) was prescribed following the initial \( \Delta a \) perturbation. When \( \Delta a \) was assumed to have no time-dependency, as was the case for the EESF metric, uncertainties arose from the choice of AF giving a mere snapshot in time of the \( CO_2 \) perturbation. When a time-dependency for \( \Delta a \) was prescribed or defined \textit{a priori}, differences stemmed from whether the goal of the analysis was to derive a single (e.g., GWP) or a time-series of \( CO_2 \)-eq. emission (or removal) pulses (e.g., TDEE). As a single pulse or scalar metric, the GWP has merit when conformity with IPCC emission metrics is desired, although in general the need to define a time-dependency for \( \Delta a \) (also for TDEE) introduces significant uncertainty owed to the reversible nature of \( \Delta a \). For TDEE and GWP, reducing uncertainty by way of accounting for the time-dependency of \( RF_{CO_2} \) comes at the expense of increasing uncertainty by way of adding a time-dependency to \( RF_{\Delta a} \).

Our review leads us to conclude that no single metric is necessarily superior to another; rather, their relative merits (Table 4) are context-dependent. For instance, application of EESF to \( \Delta a \) perturbations in dynamic systems (i.e., systems in which \( \Delta a \) is time-varying over long time scales) opens the risk for grossly mis-characterizing the system, particularly when the chosen \( \Delta a \) is not representative of the mean \( \Delta a \) of the system (Fig. 6 B). For such systems, time series metrics accounting for the time-dependency of \( \Delta a \) (i.e., TDEE) have greatest merit, but GWP or \( \Sigma TDEE \) have merit when scalar metrics are desired, which is often the case in many decision-support contexts. The choice between GWP and \( \Sigma TDEE \) when scalar metrics are desired then depends either on one of two criteria (Fig. 2): i) on the decision support time-frame; that is, on whether the physical interpretation of \( CO_2 \) “equivalence” corresponds to the present (i.e, GWP) or the future (i.e, \( \Sigma TDEE \)); or, ii) on whether compatibility with IPCC emission metrics is desired.

Although this review has provided needed guidance for choosing appropriate \( \Delta a \) metrics according to the context in which they have merit, practitioners should always be mindful that \( RF_{CO_2} \) and \( RF_{\Delta a} \) are not necessarily additive. The global mean temperature may respond differently to identical RFs, and, although there are ways to deal with this discrepancy – either by using ERFs directly in the metric calculation or by adjusting RFs with appropriate efficacy factors – such approaches require additional modeling tools which introduces notable additional uncertainties (section 6). Efficacies for inhomogeneous forcings like \( RF_{\Delta a} \) are spatial pattern- and scale-dependent (Shindell et al., 2015), and are sensitive to the climate model set-up and experimental conditions (i.e., how, where, and when \( \Delta a \) is imposed in the model). Moreover, efficacies are forcing-type dependent; that is, the forcing signal driving the underlying temperature response may depend on multiple additional perturbations at the surface that are co-associated with \( \Delta a \). A good example is LULCC, which perturbs a suite of additional biogeophysical properties affecting surface fluxes (Table 5), some of which resulting in atmospheric feedbacks that can
counteract the $\Delta a$-driven signal (Laguë et al., 2019). Since LULCC represents a broad range of land-based forcings, each of which in turn representing a myriad combination of surface biogeophysical property perturbations, the risk of misapplication of efficacies derived from climate modeling simulations of LULCC is inherently large.

7.1 Research Roadmap

Research efforts directed towards building a scientific consensus surrounding the most appropriate $RF_{\Delta a}$ estimation method (or model) for use in metric computation would serve to enhance metric transparency and facilitate comparability across studies. Given the ease and efficiency of applying radiative kernels for $RF_{\Delta a}$ calculations, such efforts might entail systematic evaluations and benchmarking of radiative kernels (e.g., as in Kramer et al. (2019)) for $\Delta a$.

Reducing uncertainty surrounding the efficacy of $RF_{\Delta a}$ associated with a variety of underlying surface forcing types (i.e., specific LULCC conversions, geoengineering methods, etc.) is paramount to reducing the “additivity” uncertainty (Jones et al., 2013) of RF-based metrics for $\Delta a$. This can be achieved through extending existing climate modeling experimental protocols (e.g., LUMIP, GeoMIP) or by creating new protocols that seek to systematically quantify the sensitivity of the global mean temperature response to variations in the spatial pattern, extent, and magnitude of surface and TOA radiative forcings associated with $\Delta a$.

Research is also needed to examine the relevance of accounting for the climate-carbon feedback in $\Delta a$ metrics, given that such feedback is implicitly included in CO$_2$’s impulse-response function (Gasser et al., 2017). Such research should be mindful of the regional climate response patterns of the various surface forcing types associated with $\Delta a$, and how regional CO$_2$ sinks are affected in turn by the regional response patterns.

Finally, while not a research need per se, a discussion between metric scientists and users/policy makers is needed surrounding three topics (Myhre et al., 2013): i) useful applications; ii) comprehensiveness; and iii) the value of simplicity and transparency. The first involves identifying which application(s) a particular $\Delta a$ metric is meant to serve. We have already shown for instance that the EESF metric is not ideal for characterizing dynamic systems. As for comprehensiveness, from a scientific point of view we would ideally wish to be informed about the totality of climate impacts of a $\Delta a$ perturbation at multiple scales (i.e., at both the local and global levels). But a user may often need to aggregate this information, which necessitates trade-offs between impacts at different points in space, between impacts at different points in time, and even between the choice of metric indicator (e.g., $RF$ vs. $\Delta T$). Related to comprehensiveness is the value of having simple and transparent metrics versus more complex model-based metrics (e.g., those based on ERF). The discussion here should weigh their trade-offs: the former may convey incomplete information, whereas the latter are more uncertain.
7.2 Concluding remarks

For the past several decades, normalized emission metrics have served useful in enabling users or decision makers to quickly perform calculations of the climate impact of GHG emissions. Their common CO\textsubscript{2}-equivalent scale has provided flexibility in emissions trading schemes and international climate policy agreements like The Kyoto Protocol. With the advent of the Paris Agreement and a broadened emphasis (Article 4) to include both emissions and removals, more attention to land-based mitigation seems likely, and the need for a way to compare albedo and CO\textsubscript{2} on an equivalent scale may increase. This obliges the scientific community to provide users with better tools to do so.

This review has highlighted many challenges associated with quantifying and interpreting CO\textsubscript{2}-equivulant metrics for $\Delta a$ based on the RF concept. A variety of metric alternatives exist, each with their own set of merits and uncertainties depending on the context in which they are applied. This review has provided guidance to practitioners for choosing a metric with maximum merit and minimum uncertainty according to the specific application context. Going forward, practitioners should always be mindful of the inherent limitations of RF-based measures for $\Delta a$, carefully weighing these against the uncertainties of metrics based on impacts further down the cause-effect chain – such as a change in temperature.

Acknowledgements

We acknowledge funding from the Research Council of Norway, grant number 254966. We are grateful for comments and feedback provided by Dr. Gunnar Myhre.

References


Zhao, K., and Jackson, R. B.: Biophysical forcings of land-use changes from potential forestry activities in North America, Ecological Monographs, 84, 329-353, 10.1890/12-1705.1, 2014.