Alternative climate metrics to the Global Warming Potential are more suitable for assessing aviation non-CO$_2$ effects

Check for updates

Liam Megill [1,2] ✉, Kathrin Deck[2] & Volker Grewe [1,2]

A growing body of research has highlighted the major contribution of aviation non-CO$_2$ emissions and effects to anthropogenic climate change. Regulation of these emissions, for example in the EU Emissions Trading System, requires the use of a climate metric. However, choosing a suitable climate metric is challenging due to the high uncertainties of aviation non-CO$_2$ climate impacts, their variability in atmospheric lifetimes and their dependence on emission location and altitude. Here we use AirClim to explore alternatives to the conventional Global Warming Potential (GWP) by analysing the neutrality, temporal stability, compatibility and simplicity of existing climate metrics and perform a trade-off. We find that using the temperature-based Average Temperature Response (ATR) or using an Efficacy-weighted GWP (EGWP) would enable a more accurate assessment of existing as well as future aircraft powered by novel aviation fuels.

The aviation industry contributes to anthropogenic climate change through CO$_2$ and non-CO$_2$ emissions. Recent studies have underscored the significance of aviation non-CO$_2$ emissions, which are now thought to be responsible for around two-thirds of the total warming from aviation[1]. Of primary importance is the release of nitrogen oxides (NO$_x$)[2–6], water vapour[7,8] and aerosols[9–12], and also the formation of contrails[13–18]. The EU Parliament recently adopted legislation (Directive 2023/958 of May 10, 2023 amending Directive 2003/87/EC) that aims to revise the EU Emissions Trading System (ETS) for aviation, inter alia requiring the European Commission to include aviation non-CO$_2$ effects in a monitoring, reporting and verification (MRV) framework and, if deemed appropriate, expand the scope of the ETS to include aviation non-CO$_2$ effects by the end of 2027. Such implementation requires the use of a climate metric, which relates non-CO$_2$ emissions and effects to their consequences on the climate and/or on society[19–22]. To maintain compatibility with market-based or offsetting schemes such as the ETS, climate metrics are often used as exchange rates, expressing non-CO$_2$ emissions on a common scale with CO$_2$ emissions. This single-basket approach can simplify climate negotiations and the implementation of climate policies[22].

However, establishing an adequate equivalence is not trivial and there is currently no consensus on which climate metric is most appropriate for aviation. In international climate policy, the most commonly used climate metric is the Global Warming Potential (GWP)[19,23], although it has been heavily criticised, primarily due to its dependence on the time horizon[20,24–27]. The choice of climate metric for aviation climate policy is further complicated because aviation non-CO$_2$ emissions and effects have highly varying atmospheric lifetimes and efficacies[28], are dependent on the emission time, altitude and location[29,30] and their impacts on the climate have a high degree of uncertainty[1,31]. Furthermore, since each climate metric uses a different climate indicator (e.g., stratospheric-adjusted radiative forcing or global mean near-surface temperature change)[32] and calculation method, a climate metric can inherently and inadvertently place emphasis on certain aircraft design choices, emission species or effects. The choice of climate metric is thus an important consideration for all stakeholders to ensure that the implementation of climate policy results in the desired reduction of the aviation industry's impact on climate and on society[22,32].

In this paper, we explore the applicability of existing, physical climate metrics to the aviation industry. Specifically, we analyse the compatibility to aircraft design and aviation policy, methodological simplicity, neutrality and stability of the following conventional climate metrics: Radiative Forcing (RF) and the Radiative Forcing Index (RFI—relative RF)[33,34], GWP[19,23], Global Temperature Change Potential (GTP)[35], Integrated GTP (iGTP)[36] and Average Temperature Response (ATR)[37]. The performance of a recently proposed, unconventional method that relates the changes in emission rates of short-lived species to pulses of CO$_2$, denoted GWP*[38–40], is also evaluated. We further analyse derivatives of the GWP and GWP* that are weighted by the efficacy, which we denote the Efficacy-weighted Global Warming Potential (EGWP) and EGWP* respectively. We find that, compared to the dominant GWP, a more accurate assessment of existing as well as future aircraft powered by novel

[1]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany. [2]Delft University of Technology, Faculty of Aerospace Engineering, Section Aircraft Noise and Climate Effects, Delft, The Netherlands. ✉e-mail: liam.megill@dlr.de

aviation fuels would be enabled through the introduction of the ATR or EGWP into climate policy. We recommend further research into the potential use of the EGWP and into new efficacy estimates for aviation non-CO_2 emissions.

Results

Development of climate metric requirements

Climate metrics are used in both their absolute form and relative to CO_2. In aviation, absolute climate metrics have two primary use cases: in trajectory optimisation, where aircraft are re-routed to avoid climate-sensitive regions[41,42]; and in aircraft design, where the climate metric can be part of the design trade-off process[43]. Relative metrics are primarily used at a policy level, notably to calculate CO_2-equivalent emissions and multipliers in single-basket emissions trading schemes such as the ETS[19,22].

From these use cases, we identify the following main requirements for climate metrics used for aviation. Aviation climate metrics shall:

1. **Neutrally represent the chosen climate indicator (REQ 1)**[32,44]. Value judgements should be left to policymakers and should not be built into climate metrics. Therefore, a climate metric should not exhibit any inherent bias towards specific aircraft design changes.
2. **Be temporally stable (REQ 2)**[45]. For aviation policy, it should be possible to use climate metrics to monitor how well the industry is performing through annual and quarterly reports. The results shown by a climate metric, and thus the emission offsetting cost, should not vary to the extent that policymakers cannot gauge the effectiveness of their policies, and airlines and other stakeholders cannot estimate their offsetting cost.
3. **Be compatible with existing climate policy (REQ 3)**[45]. A new climate metric must still be able to perform the same functions in the current climate policy context.
4. **Be simple to understand and implement (REQ 4)**[20,21,45]. Non-specialists should be able to understand how a climate metric is calculated and be able to correctly interpret what its results show.

In the following, we analyse each requirement individually and recommend the best-suited climate metric based on the results. We also analyse the impact of the time horizon on the results. We perform our analyses using the climate-chemistry response model AirClim[46,47], which provides yearly global mean radiative forcing and temperature change values from spatially resolved aviation scenarios for CO_2, water vapour, contrails and NO_x-induced changes in ozone (short and long-term) and methane. For the purposes of analysing climate metrics with time horizons in the order of years, the responses of other very short-lived species such as aerosols are expected and assumed to be qualitatively the same as for contrails.

Note that in this paper, we add an A to denote an absolute climate metric (i.e., AGWP, AEGWP, AGTP, iAGTP) and use rATR to denote the relative ATR (to CO_2). The GWP* and EGWP* do not have absolute and relative forms. Where necessary for clarification, we also use P-, F- and S- to denote climate metrics calculated using a pulse, fleet or total aviation industry emission scenario, respectively.

REQ 1: climate metric neutrality with respect to aviation emissions

To assess the neutrality of climate metrics for aircraft design, the peak and average total temperatures and climate metric values of potential future fleets are compared. A wide range of narrowbody fleets are generated using a Monte Carlo simulation of various high-level aircraft parameters, including the use of conventional as well as novel aviation fuels such as SAF and hydrogen. The fleets are analysed using the climate-chemistry response model AirClim[46,47], as described in "Methods" (cf. Table 3). The neutrality of each climate metric is gauged by the frequency f of incorrect fleet pairs— defined here as when the signs of the differences in peak/average temperature (ΔT) and total climate metric value (CM) between any two fleets i and j do not match—compared to the total number of fleet pairs (cf. the

method used by Grewe et al.[48]):

$$f(H) = \frac{1}{C(N,2)} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \begin{cases} 1 & \text{if } (CM_{H,j} - CM_{H,i}) \times (\Delta T_j - \Delta T_i) < 0 \\ 0 & \text{otherwise} \end{cases}$$

(1)

where $N = 10{,}000$ is the total number of fleets. Figure 1 shows the frequency of incorrect fleet pairs as a function of the time horizon H; example results for a single time horizon $H = 100$ are shown in Supplementary Fig. 1.

In general, the similarity shown in Fig. 1 between the results for all climate metrics and climate objectives suggest that the peak temperature is also a good indicator for the average temperature and vice versa over a wide range of time horizons. The endpoint climate metrics F-RF and F-AGTP show a clear dependence on the time horizon and hence shape of the temporal emission profile (temporal evolution of yearly emissions per species). This demonstrates that a single value of radiative forcing or temperature at a time in the future is not a good indicator of the peak or average temperature. Since the radiative forcing from a number of aviation non-CO_2 effects, such as the warming effect from the NO_x-induced short-term increase in ozone (e.g., ref. 1), has dissipated at large time horizons, the F-RF in particular can have low or altogether no sensitivity to a number of aircraft and engine design changes, leading to a high rate of incorrect fleet pairs. The integrated climate metrics F-AGWP, F-AEGWP and F-iAGTP/F-ATR, in comparison, have a memory of these previous emissions and are less dependent on the temporal emission profile.

The F-AGWP and F-GWP* show largely linear responses for peak and average temperature, particularly for time horizons above 60 years, but in general have higher frequencies of incorrect fleet pairs than climate metrics based on temperature or using efficacy. The F-EGWP* has a similarly low dependence on the time horizon with a lower frequency of incorrect fleet pairs, demonstrating almost ideal behaviour in this context. Whilst the F-iAGTP/F-ATR has a clear minimum at 70 years for peak temperature and at 20, 50 and 100 years for the corresponding average temperatures, it generally has a frequency of incorrect fleet pairs of less than 2%. The F-AEGWP performs very similarly, surpassing 2% error frequency only for the 20-year average temperature, and has clear minima at slightly lower time horizons: 55, 15, 45 and 80 years, respectively.

REQ 2: temporal stability

The temporal stability of climate metrics is judged using CO_2-eq trajectories for the full aviation industry. In this work, we use the CORSIA and FP2050 scenarios developed by Grewe et al.[49] as examples. The CORSIA scenario assumes business as usual, but that CO_2 emissions are offset beyond 2020; whereas the FP2050 scenario makes use of the Flightpath 2050 targets: 75% CO_2 and 90% NO_x reduction by 2050. Figure 2 shows the CO_2-eq emissions calculated for both scenarios using each climate metric with a 100-year time horizon. Two elements of the responses are highlighted here.

First, the total CO_2-eq values calculated using the endpoint S-RFI and S-GTP climate metrics are very similar, although as the emission rate (rate of change of yearly emissions over time) reduces around the year 2020 in the FP2050 scenario the results begin to drift apart. The RFI and GTP can thus be seen to be stable for the analysed full aviation emissions. However, both climate metrics can struggle to show qualitatively the same response for pulse and constant emissions (see e.g., ref. 50, their Fig. 3.3), depending on the chosen time horizon.

The S-GWP and S-iGTP/S-rATR show very similar responses; the S-EGWP also produces similar, albeit generally lower results. The similarity between the S-GWP and S-iGTP/S-rATR potentially allows for species-dependent conversion factors and reduces the political capital required to switch from the standard GWP to either the iGTP or rATR in climate policy. This is a somewhat surprisingly result, since the bases for the climate metric calculations differ, affecting the contributions of individual species to the total CO_2-eq: The GWP is RF-based, whereas the ATR is temperature-
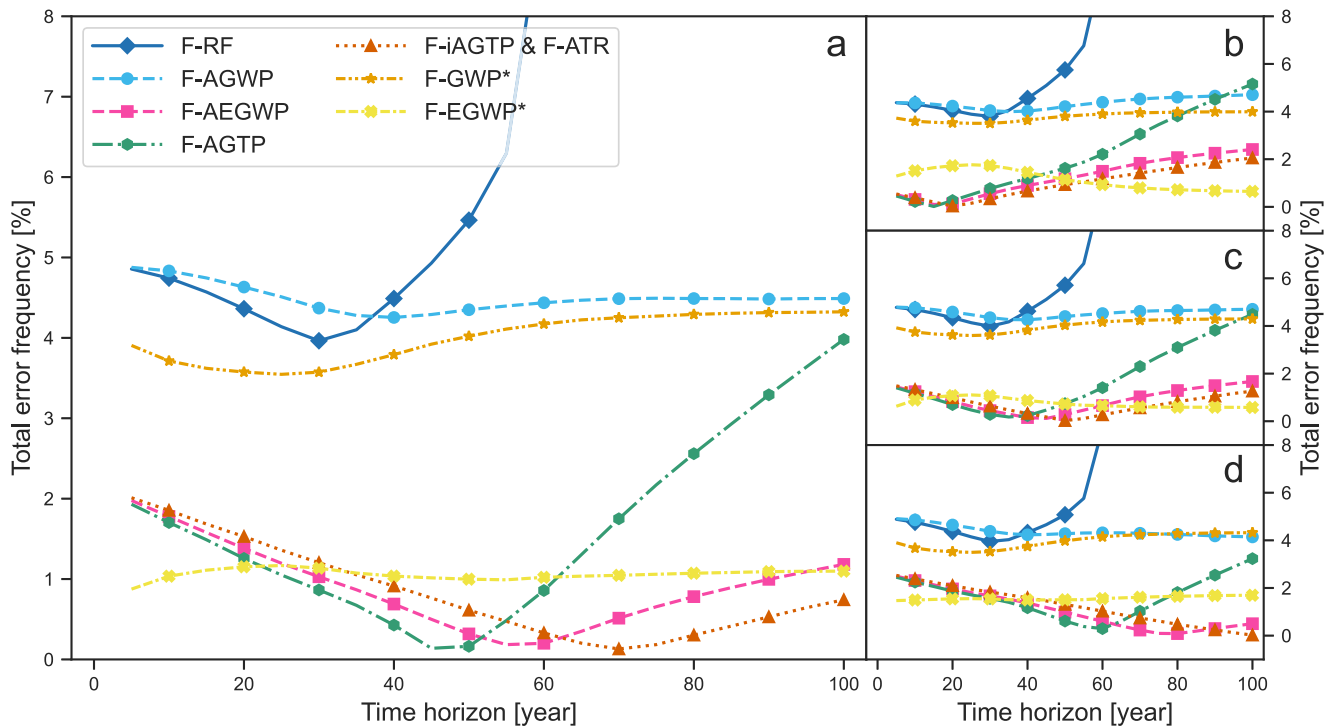
Fig. 1 | Comparison of the neutrality of different climate metrics to changes in aircraft design. Shown is the frequency of incorrect fleet pairs, corresponding to those where the signs of the peak/average temperature change and climate metric change do not match, as a function of the time horizon for the peak temperature (**a**) and 20-, 50- and 100-year average temperature (**b–d**) climate objectives. Each climate metric is represented by a different combination of colour, line style and marker. Values are available between 5 and 100 years with a time horizon step of 5 years; however, markers are shown every 10 years for clarity.
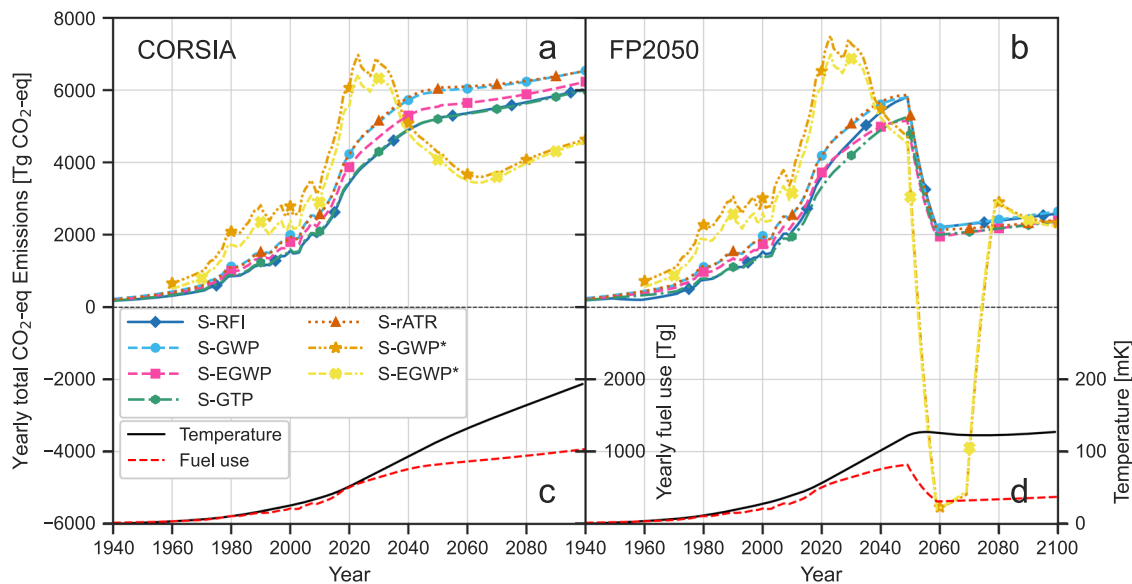

Fig. 2 | Comparison of the climate metric responses for full aviation emission scenarios. Shown are the CO_2-eq emissions calculated using each climate metric with a 100-year time horizon for the CORSIA (**a**) and FP2050 (**b**) scenarios. Each climate metric is represented by a different combination of colour, line style and marker for clarity. Also shown is the fuel use (red dashed line) and temperature (solid black line) response for each emission species calculated using AirClim for the CORSIA (**c**) and the FP2050 (**d**) scenarios. All values are calculated on a yearly basis —the markers for each species are differently spaced such that overlapping lines can more easily be identified.

based. As a result, the GWP emphasises contrail cirrus, and the ATR the warming effect of NO_x-induced ozone (see Supplementary Fig. 2). Nevertheless, for full aviation scenarios assuming Jet-A1 fuel, the differing contributions seem to balance out. It is, however, likely that the introduction of novel propulsion technologies and fuels, which change the emission indices relative to one another, will result in a divergence of the total CO_2-eq

emissions calculated by both metrics. The rATR would then likely more closely match the EGWP than the GWP. Further research could analyse the response from different models and emission inventories to check the validity of conversion factors, in particular for novel aviation fuels.

A second noticeable element is the rapid deviation of the S-GWP* and S-EGWP* from the response shown by all other climate metrics in both

scenarios. The GWP* calculation method uses an average of the previous 20 years of radiative forcing and is closely tied to the emission rate given by the scenario. Therefore, small changes in the emission rate, in the years 2020 and 2050 in the CORSIA scenario, result in large changes in the CO_2-eq trajectory. A policymaker using the GWP* or EGWP* method to monitor CO_2-eq emissions between 2030 and 2050 could incorrectly assume that the impact of aviation is reducing, when in actuality only the emission rate has decreased.

This instability is particularly problematic for the FP2050 scenario. Whilst the values from all other climate metrics largely correspond to the reducing fuel use, the S-GWP* and S-EGWP* show negative CO_2-eq emissions between 2050 and 2080. Whilst this behaviour is useful for representing the temperature using a cumulative integral, negative CO_2-eq values could easily be misinterpreted as a sign that aviation is causing an active cooling. The magnitude of the negative CO_2-eq values is also disproportionately large compared to the shallow peak shown in the temperature.

REQ 3: compatibility with existing climate policy

To be compatible with existing climate policy means that a climate metric can be used in current climate frameworks and methods. These have generally been established on the basis of the GWP, which has become the most commonly used climate metric. There is thus a natural bias towards climate metrics that behave in a similar manner to the GWP. For aviation, this functionally means that any alternative to the GWP must be able to (1) calculate the temporal trajectories of CO_2-eq emissions; and (2) calculate single values for fleets and individual flights, the latter of which is necessary for the introduction of aviation non-CO_2 emissions into the ETS. The climate metrics RF, GWP, EGWP, GTP, iGTP and ATR are able to perform these functions. However, the GWP* and EGWP* struggle to provide a single value for an individual fleet or flight.

Rather than providing a single value for a given time horizon, the GWP* method provides a temporal trajectory for each emission species, as shown in Fig. 3 for a simple fleet temporal emission profile. If used as a climate metric, for example, to compare this fleet to another, there is no obvious point along the temporal trajectory to choose. Indeed, the choice of which point to use is itself a trade-off between different emission species. Therefore, whilst the GWP* method is useful for certain technical discussions, it should be seen as a model rather than a metric, as previously argued by Meinshausen and Nicholls[45], and should not be viewed as a potential replacement for the GWP in aviation policy.

REQ 4: simplicity

The endpoint climate metrics RF and GTP are clearly the easiest to understand. It is straightforward to determine how these climate metrics behave for different time horizons, background emissions scenarios and fuel scenarios. Integrated climate metrics—GWP, EGWP, iGTP and ATR—are more complex and it can be difficult to ascertain the impacts of individual effects and species on the results. The least simple to understand and

implement are the GWP* and EGWP*: Their behaviour can be puzzling even for simple temporal emission profiles and can show initially counter-intuitive results, such as the negative emissions in Fig. 2.

In comparison to temperature-based climate metrics (GTP, iGTP, ATR), climate metrics based on radiative forcing (RF, GWP) are easier to implement since they do not need a full climate or carbon-cycle model. The EGWP requires efficacy values and is thus more complex. However, the demand on computational time depends on which model is used to calculate the climate metric values. The GWP* and EGWP* are RF-based climate metrics, but they do use the AGWP for CO_2, which would need to be defined as a standard value or calculated for a given scenario. In addition, these metrics require the temporal emission profile twenty years prior to any value, since the method uses a 20-year running average. This could potentially complicate the implementation of the GWP* method.

Choice of climate metric and time horizon

An overview of the performance of all analysed climate metrics is shown in Table 1. It is clear that the choice of climate metric must be the result of a trade-off. Based on our analysis and definition of requirements, the ATR and EGWP can be seen to perform best. Here, we inspect in more detail the advantages and disadvantages of these metrics in comparison to the GWP and investigate their dependence on the time horizon. We note that the ATR and iAGTP differ only in the division of the time horizon; in their relative forms, the rATR and iGTP are identical. However, the ATR is chosen rather than the iAGTP because the division by the time horizon improves the stability of the absolute climate metric responses.

The ATR and EGWP perform similarly well in the pairwise fleet analysis (REQ 1) for the peak and average temperature climate indicators, as well as in analysis on temporal stability (REQ 2). The rATR$_{100}$ produces CO_2-eq emissions that very closely match those calculated using the GWP$_{100}$, potentially easing the introduction of the ATR in climate policy (REQ 3); introduction of the EGWP is even simpler since it is only a derivative of the GWP. Finally, although the concept of an average temperature change is simple to understand for non-specialists, it can be difficult to identify the impacts of specific effects on results. In comparison to the GWP, the EGWP and in particular the ATR as a temperature-based metric, include more climatic processes, but thus also more assumptions and uncertainties. Implementation of both climate metrics can thus be seen as complex (REQ 4). Further work is required to better understand the benefits and potential downsides of the EGWP in other contexts. Further research into best estimates of the efficacy would also be beneficial.

The dependence of the GWP, EGWP and ATR on the time horizon is shown in Fig. 4 for three different emission scenarios. Individual overviews for each of these metrics and the GTP are provided in Supplementary Figs. 3–6. The inclusion of the efficacy in the EGWP is evident from the lower panel of the figure, affecting the relative importance of ozone and contrails in particular. The results of the RF-based EGWP now much more

Fig. 3 | CO_2-eq emissions calculated using the GWP*$_{100}$ method for an example fleet demonstrating the flow-based nature of the GWP*. The inset figure shows the temporal emission profile; the main figure the CO_2-eq response for each species (colours). For the comparison of fleets in this study (F-GWP*), the peak total value is used, in this example occurring in the year 2050. Note that the total value is dominated by the contrail impact and that different species have their peaks at later times. Therefore, the choice of which value to use is itself a trade-off between different emission species.
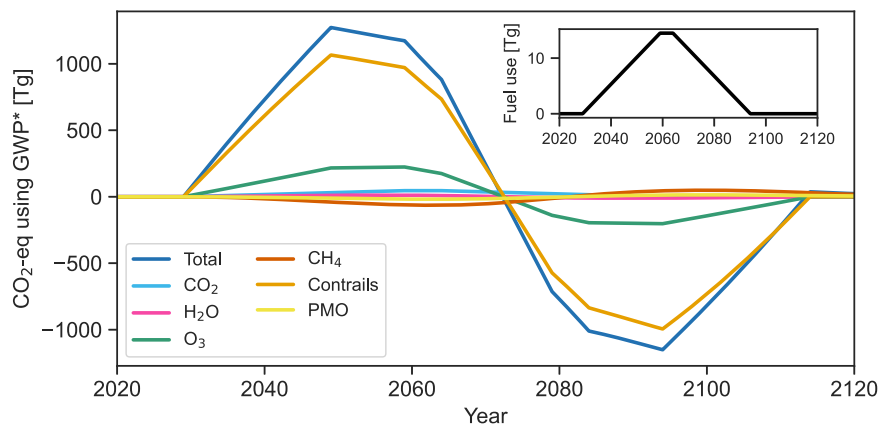
Table 1 | Overview of the performance of the analysed climate metrics with respect to each requirement

Requirement	Neutrality (REQ 1)	Stability (REQ 2)	Compatibility (REQ 3)	Simplicity (REQ 4)
RF	Very low neutrality	Generally stable	Compatible	Simple to understand and implement
GWP	Low, but consistent neutrality	Stable	Compatible (standard climate metric)	Complex to understand, simple to implement
EGWP	High neutrality	Stable	Compatible	Complex to understand and implement
GTP	Low, but inconsistent neutrality	Generally stable	Compatible	Simple to understand, complex to implement
ATR & iGTP	High neutrality	Stable	Compatible	Simple/complex to understand, complex to implement
GWP*	Low, but consistent neutrality	Highly unstable	Not compatible	Highly complex to understand and implement
EGWP*	High and consistent neutrality	Highly unstable	Not compatible	Highly complex to understand and implement

"Generally stable" refers to the finding that the RF and GTP are stable for full aviation emission scenarios, but can struggle to show qualitatively the same response for pulse and constant emissions (see e.g., ref. 50, their Fig. 3.3).
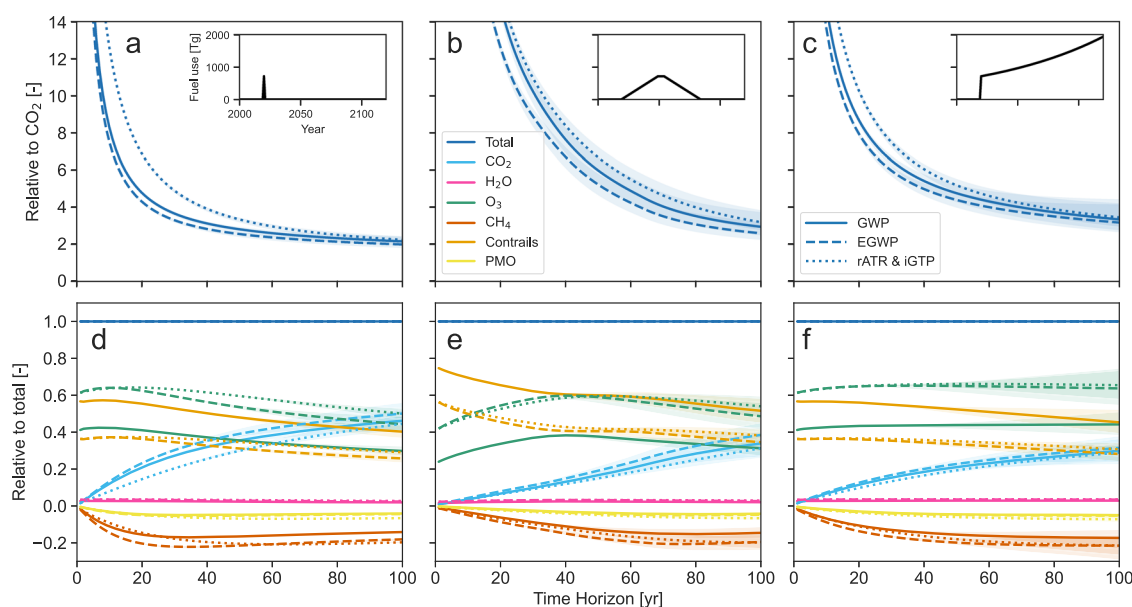

Fig. 4 | Comparison of the GWP, EGWP and ATR responses for time horizons between 0 and 100 years. Shown are the responses from the GWP (solid line), EGWP (dashed line) and rATR (dotted line), which in its relative form is equivalent to the iGTP. The top row shows the total metric value relative to CO_2; the bottom row the responses calculated for each species (colours) relative to the total. Three temporal emission profiles are used, for which the fuel usage profiles are shown in the inset plots: a pulse emission (P-) in (**a**, **d**); a fleet emission (F-) in (**b**, **e**) and a 1% increasing emission (I-) in (**c**, **f**). Each response is shown for the Shared Socio-economic Pathway SSP2-4.5 with margins (shading) for scenarios SSP1 to SSP5, which are used as the background emissions scenarios in AirClim.

closely match those of the temperature-based ATR. Nevertheless, the total relative metric values and thus calculated CO_2-eq emissions of the GWP, EGWP and ATR are very similar, especially for large time horizons. In general, the sensitivity of all three metrics to the time horizon, represented by the gradient of the relative metric values, decreases with increasing time horizon. Using a low time horizon, for example, 20 years, would require particular justification: Why was 20 years chosen rather than say 15, 25 or even 19 years? Instead, the responses suggest that larger time horizons are most suitable for integrated climate metrics, greater than around 70 years. This is particularly true for the ATR, which requires larger time horizons to properly account for the delay in the temperature response of the atmosphere.

Discussion

Our analyses demonstrate that the selection of a climate metric plays a crucial role in ensuring that implemented climate policies effectively reduce the aviation industry's impact on the climate. In the fleet pairing analysis, we illustrate that climate metrics can have inherent trade-offs and favour certain aircraft designs over others. These inherent biases are undesirable since value judgements should be left to policy decision-making and not embedded into climate metrics.

The choice of climate metric is always the result of a trade-off. Due to the historical dominance of the GWP, there is a natural bias towards climate metrics that behave in a similar manner. However, our research clearly suggests that there are derivatives and alternatives that outperform the GWP for aviation. We require that a suitable climate metric displays neutrality with respect to different emission species; exhibits temporal stability; is compatible with existing climate policy; and is simple to understand and implement. These requirements are in line with those stated by others[20,45]. Based on these requirements, we identify the Efficacy-weighted Global Warming Potential (EGWP) and Average Temperature Response (ATR) as the most appropriate climate metrics for aircraft design and aviation policy. Both metrics are stable and can monitor the impact of the aviation industry using CO_2-equivalents effectively. They also do not favour specific emission species for both peak and average temperature climate indicators across a wide range of time horizons and emission scenarios.

Whilst the ATR as a temperature-based climate metric has the potential to include more climatic processes and be more relevant for temperature-based targets than the GWP, the larger number of assumptions and uncertainties must also be considered. The EGWP may, therefore, be a useful compromise for policymakers, in that it can more accurately represent the climate impact of aviation whilst still using the GWP methodology.

Table 2 | Calculation methods for all climate metrics used in this research

Climate metric	Calculation method
Radiative forcing (RF)	$RF_H = RF(t_0 + H)$
Global warming potential (GWP)	$AGWP_H = \int_{t_0}^{t_0+H} RF(t)dt$
Efficacy-weighted GWP (EGWP)	$AEGWP_{i,H} = r_i \int_{t_0}^{t_0+H} RF_i(t)dt$
Global temperature change potential (GTP)	$AGTP_H = \Delta T(t_0 + H)$
Integrated GTP (iGTP)	$iAGTP_H = \int_{t_0}^{t_0+H} \Delta T(t)dt$
Average temperature response (ATR)	$ATR_H = \frac{1}{H} \int_{t_0}^{t_0+H} \Delta T(t)dt = \frac{1}{H} iAGTP_H$
GWP*	$E_{CO_2-we}(t) = g(s) \times \left[(1-s) \frac{\Delta RF(t)}{\Delta t} \frac{H}{AGWP_{H(CO_2)}} + s \times \frac{\overline{RF(t)}}{AGWP_{H(CO_2)}} \right]$
EGWP*	Same as GWP* except replacing RF_i with $RF_i \times r_i$

t_0 is the start of an emission series and, therefore, also the year for which a climate metric is to be calculated; H is the time horizon, r the efficacy and i a single species. E_{CO_2-we} is explained in the calculation description of the GWP*. Note that these methods all require the time series of radiative forcing (RF) or temperature change ΔT at least from t_0 until $t_0 + H$, which requires the use of a separate climate model. The GWP* and EGWP* require the time series of RF at least from $t_0 - 20$ until $t_0 + H$.

Further research is recommended into the advantages and potential disadvantages of using the EGWP. If the ATR were to be chosen, it would benefit from the close match of the total CO_2-eq emissions calculated by the S-rATR$_{100}$ and S-GWP$_{100}$, despite the differences in contributions of individual species. The S-EGWP$_{100}$, for its part, also produces similar results. However, it is likely that the total emissions calculated by the ATR and GWP will diverge with the introduction of novel propulsion technologies and fuels such as hydrogen since the relative contributions of the non-CO_2 emissions will change.

Determining an appropriate time horizon for both the EGWP and ATR remains a challenge. The time horizon is a trade-off between incorporating the long-term response to an emission and ensuring the predictability and accuracy of a future emission scenario. We find that integrated climate metrics generally require larger time horizons to account for the atmospheric radiative forcing and temperature adjustment. If a short time horizon is chosen, policymakers must provide sufficient justification for the choice. Alternatively, values for different time horizons could be provided together, as proposed in ref. 27, although this complicates the calculation of CO_2-eq emissions, for example in the upcoming ETS revision.

The accuracy of our results could be improved by using real aircraft designs: Since design parameters are chosen randomly within a given range, some fleets may not be physically feasible. However, since we used the same method of randomly choosing parameters, any additional incorrect fleet pairings caused by this limitation are assumed to cancel out and not impact our conclusions. Similarly, given the wide range of potential aircraft designs analysed in this study, it is unlikely that the choice of climate model, Air-Clim, has influenced the results significantly. Verification with another climate model may enhance our understanding of the results.

Ultimately, the most suitable climate metric and corresponding time horizon must be determined by policymakers depending on the policy and climate objective, the emission scenario, and whether a relative or absolute climate metric is required. Based on a general set of requirements suitable for policymaking, our findings endorse the use of the ATR and EGWP with a time horizon greater than 70 years for aircraft design and aviation policy to assess the long-term climate impact of aviation. However, the choice of climate metric does not have to be contentious or controversial: As our analysis and the numerous previous studies have demonstrated, tools exist with which the performance of any climate metric can be analysed and potential shortcomings and pitfalls identified, such that these can be addressed in climate policy.

Methods
Climate metric calculation methods
The calculation methods for all climate metrics are given in Table 2. These methods require a time series of radiative forcing (RF) and resulting temperature change ΔT. The calculation methods of the EGWP, ATR/iGTP and GWP*/EGWP* are described in more detail below. In this research, we use the climate-chemistry response model AirClim[46,47] to calculate the RF and

ΔT responses of individual aircraft fleets using data from the DLR WeCare project[51], and of the full global fleet using scenarios developed by Grewe et al.[49]. These are described in more detail in subsequent sections. For ease of comparison, we use the Shared Socioeconomic Pathway SSP2-4.5[52] as the default background emissions scenario for our analyses, but vary between SSP1 to SSP5 in the multivariate fleet analysis. AirClim is an extension to the linear response model for CO_2 developed by Sausen and Schumann[53] and combines emission data with pre-calculated altitude- and latitude-dependent data obtained from steady-state simulations with the E39/CA[54] climate-chemistry model (for ozone, methane, water vapour and contrails) and ECHAM4-CCMod[55] (for contrail cirrus). It was chosen for this research due to its low computational cost and flexibility.

EGWP—The Efficacy-weighted Global Warming Potential (EGWP) was developed as a derivative of the GWP. It aims to introduce the efficacy of non-CO_2 emissions into the GWP method, such that the results obtained by the GWP more closely match those of temperature-based climate metrics. The EGWP for a single species i is then the GWP of that species multiplied by its efficacy r_i, taken from ref. 28 (their Table 1). We note that this calculation method is still quite uncertain, in particular for contrail cirrus[56], although it is not expected to affect the results of this study. Another potential approach to calculate the EGWP would be to use the Effective Radiative Forcing (ERF) and corresponding efficacies r'_i. Further work is required to analyse the performance of these two climate metrics and to develop better estimates of the efficacies for aviation emissions.

ATR—The Average Temperature Response (ATR) was initially developed by Dallara et al.[37] specifically for aircraft design. Initially, it included a weighting function and used an infinite time horizon H. However, the infinite time horizon in particular made it inappropriate for global fuel scenarios, for example. Since its inception, therefore, the ATR has been repurposed and is now generally used as the average temperature change over a given time horizon, as shown in Table 2. The weighting function is also no longer used. The relative ATR is denoted rATR in this research for clarity. Note that this definition of the ATR is related to the iAGTP by: $ATR_H = iAGTP_H / H$.

GWP*—Since emission rates are meaningless for NO_x-induced aviation effects (O_3, long-term CH_4 reduction and the Primary Mode Ozone (PMO) effect) and for contrails, the GWP* methodology must be adapted to use radiative forcing. This equivalent calculation is proposed in the initial development of the GWP* by Allen et al.[38] and is modified using the improvements suggested by Cain et al.[39] and Smith et al.[40] to obtain:

$$E_{CO_2-we}(t) = g(s) \times \left[(1-s) \times \frac{\Delta RF(t)}{\Delta t} \times \frac{H}{AGWP_{H(CO_2)}} + s \times \frac{\overline{RF(t)}}{AGWP_{H(CO_2)}} \right]$$

(2)

where $E_{CO_2-we}(t)$ are CO_2-warming equivalent emissions as a function of time, ΔRF the change in radiative forcing over the previous $\Delta t = 20$ years, $\overline{RF}$

Table 3 | Ranges of fleet design parameters for the fleet pairing analysis simulations

Parameter	Range	Reference	Minimum	Maximum
Fuel burn (40% of category) [Tg]	70–100%	100%	70%	100%
NO_x emission [Tg]	70–100%	100%	70%	100%
Cruise pressure [hPa]	80–120%	100%	120%	80%
Contrail distance modifier [km]	40–100%	100%	40%	100%
Fuel used [—]	Jet-A1, SAF, H_2	Jet-A1	H_2(FC)	Jet-A1
Year of fleet introduction [yr]	2030–2050	2030	2030	2030
Background emissions [—]	SSP1–SSP5	SSP2-4.5	SSP2-4.5	SSP2-4.5

See the accompanying text for a more detailed description of the contrail distance modifier, and Table 4 for the impact of different fuels.

the running average of RF and $AGWP_{H(CO_2)}$ the AGWP of a CO_2 pulse at a time horizon of H years. The above equation differs to the one used by Lee et al.[1] only by the multiplication by $g(s)$, which was introduced in the same year by Smith et al.[40] to improve consistency with the linear models used for climate metric calculations. In this research, we use $s = 0.75$ to be consistent with Smith et al.[40]. However, we note that this value was calculated for methane (CH_4) and thus may not be optimal for other aviation non-CO_2 emissions and effects.

The EGWP* is a climate metric developed as part of this research as a derivative of the GWP*. Similarly to the EGWP, it makes use of the efficacy r_i, also taken from ref. [28], to more closely match the results obtained by temperature-based climate metrics. The GWP* methodology is adapted by replacing RF_i with $RF_i \times r_i$.

The GWP* and EGWP* differ from the other climate metrics considered in this study in that they are flow-based climate metrics: The GWP* method does not provide a single value over a specific time horizon. Instead, it provides a CO_2-eq value as a function of time, as shown in Fig. 3 (main text). To estimate the impact of a fleet or flight, a certain point along the temporal trajectory must be chosen. It can be argued that for the analysis of the peak temperature, the peak CO_2-eq value should be chosen. However, the time at which the peak occurs differs per species, and can also differ per fleet, thereby raising the question whether the climate metric values of each fleet are showing the same thing and are thus intercomparable. In the example shown in the Figure, the peak total CO_2-eq value is dominated by the contrail impact—all other emissions have their peaks at a later time. However, since no other point could be identified as appropriate, in this research we use time of the peak total CO_2-eq value for fleet comparisons—in Fig. 3 thus the values in 2050.

Development of fuel scenarios

This research is based on the CORSIA and Flightpath 2050 (FP2050) fuel scenarios developed by Grewe et al.[49]. Since time horizons of up to 100 years are analysed, the scenarios needed to be extended. For this research, they have been extended until the year 2200, assuming a 0.5% annual growth rate after the year 2100. The scenarios are developed to test climate metrics and have not been evaluated for reliability and accuracy.

Fleet pairing analysis

The fleets used in this research are theoretical and characterised with a set of input parameters in AirClim, chosen uniformly from ranges shown in Table 3. The parameter ranges are based on expected technological pathways developed by Grewe et al.[49], within the Clean Sky 2 Technology Evaluator[57] and by the "Hydrogen-powered aviation" report by the Clean Hydrogen Joint Undertaking (2020, https://doi.org/10.2843/766989). The contrail distance modifier mentioned in the Table is a multiplier for the total cruise distance for which contrails form, which is an AirClim input. In this context, a factor below unity corresponds to aircraft flying further to avoid climate-sensitive regions and, therefore, contrail formation. As a result, the reduction in contrail distance is coupled with an increase in fuel burn, estimated from ref. [58] to be of the ratio −15%:1% contrail distance to fuel burn up to a contrail distance reduction of 60% (contrail distance modifier

Table 4 | Assumed change of in-flight emissions and emission-related effects for Sustainable Aviation Fuel (SAF) and hydrogen

Fuel	CO_2	NO_x	H_2O	Contrails	Total
SAF	−65 − 80%	−0%	−0%	−10 − 40%	−30 − 60%
Hydrogen combustion	−100%	−50 − 80%	+ 150%	−30 − 50%	−50 − 75%
Hydrogen fuel cell	−100%	−100%	+ 150%	−60 − 80%	−75 − 90%

The values are taken from the "Hydrogen-powered aviation" report (Clean Hydrogen Joint Undertaking, 2020, https://doi.org/10.2843/766989). Note that these values cannot currently be corroborated through other studies and are, therefore, only used to provide a wide spectrum of potential future fleets.

of 40%), which is approximately the end of the quasi-linear region of the Pareto fronts calculated. For fleets using fuels other than Jet-A1, the emissions parameters are further modified according to Table 4. Here, the contrail reduction is assumed to correspond to changes in the exhaust composition due to the use of different fuels. We note that the data in both tables are a simplification and that comprehensive data is not yet available for different fuel types. However, since our objective is to provide a wide range of potential future fleets, this simplification is deemed appropriate and should not affect the results of this research.

For each fleet, a constant production rate is assumed, expected to last 30 years. Production is assumed to begin after 2030, approximately on par with the expected introduction of the next generation of single-aisle aircraft and new fuels such as hydrogen according to the analyses of Grewe et al.[49]. The exact year of introduction of new fleets is, however, not relevant to the outcome of this study and is thus varied. Each aircraft is further assumed to have a lifetime of 35 years with no hull losses. A single-aisle aircraft about the size of the Airbus A320 is chosen for reference. For simplicity, the fuel use of this fleet is taken to be 40% of Category 4 of the DLR WeCare project[51], characterised by aircraft with seat numbers between 152 and 201. A total of 10,000 fleets are simulated using AirClim.

Data availability

All scenario data and AirClim simulation results used in this study are available in the 4TU.ResearchData repository https://doi.org/10.4121/344e24ad-b2f5-4ed9-8d49-6efa2081d30c.

Code availability

The Jupyter notebooks (Python) used to create the scenarios, generate the multivariate fleets and analyse the data in this study are provided in the 4TU.ResearchData repository https://doi.org/10.4121/344e24ad-b2f5-4ed9-8d49-6efa2081d30c. The software code AirClim is confidential proprietary information of the DLR and cannot be made available to the public or readers without restrictions. Licensing of the code to third parties is conditioned upon the prior conclusion of a licensing agreement with the DLR. Qualified researchers can request an agreement on reasonable request from the corresponding author.

References

1. Lee, D. et al. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmos. Environ.* **244**, 117834 (2021).
2. Fuglestvedt, J. et al. Climatic forcing of nitrogen oxides through changes in tropospheric ozone and methane; global 3D model studies. *Atmos. Environ.* **33**, 961–977 (1999).
3. Grewe, V., Dameris, M., Fichter, C. & Sausen, R. Impact of aircraft NOx emissions. Part 1: interactively coupled climate-chemistry simulations and sensitivities to climate-chemistry feedback, lightning and model resolution. *Meteorologische Zeitschrift* **11**, 177–186 (2002).
4. Stevenson, D. S. et al. Radiative forcing from aircraft NOx emissions: mechanisms and seasonal dependence. *J. Geophys. Res.* **109**, D17307 (2004).
5. Myhre, G. et al. Radiative forcing due to changes in ozone and methane caused by the transport sector. *Atmos. Environ.* **45**, 387–394 (2011).
6. Köhler, M., Rädel, G., Shine, K., Rogers, H. & Pyle, J. Latitudinal variation of the effect of aviation NOx emissions on atmospheric ozone and methane and related climate metrics. *Atmos. Environ.* **64**, 1–9 (2013).
7. Frömming, C. et al. Aviation-induced radiative forcing and surface temperature change in dependency of the emission altitude. *J. Geophys. Res. Atmos.* **117**, D19 (2012).
8. Wilcox, L., Shine, K. & Hoskins, B. Radiative forcing due to aviation water vapour emissions. *Atmos. Environ.* **63**, 1–13 (2012).
9. Righi, M., Hendricks, J. & Brinkop, S. The global impact of the transport sectors on aerosol and climate under the Shared Socioeconomic Pathways (SSPs). *Earth Syst. Dyn. Discuss.* **2023**, 1–34 (2023).
10. Penner, J. E., Zhou, C., Garnier, A. & Mitchell, D. L. Anthropogenic aerosol indirect effects in cirrus clouds. *J. Geophys. Res. Atmos.* **123**, 11–652 (2018).
11. Gettelman, A. & Chen, C. The climate impact of aviation aerosols. *Geophys. Res. Lett.* **40**, 2785–2789 (2013).
12. Righi, M., Hendricks, J. & Sausen, R. The global impact of the transport sectors on atmospheric aerosol: Simulations for year 2000 emissions. *Atmos. Chem. Phys.* **13**, 9939–9970 (2013).
13. Bickel, M., Ponater, M., Bock, L., Burkhardt, U. & Reineke, S. Estimating the effective radiative forcing of contrail cirrus. *J. Clim.* **33**, 1991–2005 (2020).
14. Burkhardt, U., Bock, L. & Bier, A. Mitigating the contrail cirrus climate impact by reducing aircraft soot number emissions. *NPJ Clim. Atmos. Sci.* **1**, 37 (2018).
15. Burkhardt, U. & Kärcher, B. Global radiative forcing from contrail cirrus. *Nat. Clim. Change* **1**, 54–58 (2011).
16. Schumann, U. et al. Properties of individual contrails: a compilation of observations and some comparisons. *Atmos. Chem. Phys.* **17**, 403–438 (2017).
17. Bock, L. & Burkhardt, U. Contrail cirrus radiative forcing for future air traffic. *Atmos. Chem. Phys.* **19**, 8163–8174 (2019).
18. Stuber, N. & Forster, P. The impact of diurnal variations of air traffic on contrail radiative forcing. *Atmos. Chem. Phys.* **7**, 3153–3162 (2007).
19. Fuglestvedt, J. S. et al. Metrics of climate change:: assessing radiative forcing and emission indices. *Clim. Change* **58**, 267–331 (2003).
20. Fuglestvedt, J. et al. Transport impacts on atmosphere and climate: metrics. *Atmos. Environ.* **44**, 4648–4677 (2010).
21. Wuebbles, D., Forster, P., Rogers, H. & Herman, R. Issues and uncertainties affecting metrics for aviation impacts on climate. *Bull. Am. Meteorol. Soc.* **91**, 491–496 (2010).
22. Tanaka, K., Peters, G. P. & Fuglestvedt, J. S. Policy update: multicomponent climate policy: why do emission metrics matter? *Carbon Manag.* **1**, 191–197 (2010).
23. Rodhe, H. A comparison of the contribution of various gases to the greenhouse effect. *Science* **248**, 1217–1219 (1990).
24. Manne, A. S. & Richels, R. G. An alternative approach to establishing trade-offs among greenhouse gases. *Nature* **410**, 675–677 (2001).
25. Shine, K. P., Berntsen, T. K., Fuglestvedt, J. S., Skeie, R. B. & Stuber, N. Comparing the climate effect of emissions of short- and long-lived climate agents. *Philos. Trans. Royal Soc. A: Math. Phys. Eng. Sci.* **365**, 1903–1914 (2007).
26. Allen, M. R. et al. New use of global warming potentials to compare cumulative and short-lived climate pollutants. *Nat. Clim. Change* **6**, 773–776 (2016).
27. Ocko, I. B. et al. Unmask temporal trade-offs in climate policy debates. *Science* **356**, 492–493 (2017).
28. Ponater, M., Pechtl, S., Sausen, R., Schumann, U. & Hüttig, G. Potential of the cryoplane technology to reduce aircraft climate impact: a state-of-the-art assessment. *Atmos. Environ.* **40**, 6928–6944 (2006).
29. Frömming, C. et al. Influence of weather situation on non-CO_2 aviation climate effects: the REACT4C climate change functions. *Atmos. Chem. Phys.* **21**, 9151–9172 (2021).
30. Lund, M. T. et al. Emission metrics for quantifying regional climate impacts of aviation. *Earth Syst. Dyn.* **8**, 547–563 (2017).
31. Lee, D. et al. Transport impacts on atmosphere and climate: aviation. *Atmos. Environ.* **44**, 4678–4734 (2010).
32. Grewe, V. & Dahlmann, K. How ambiguous are climate metrics? And are we prepared to assess and compare the climate impact of new air traffic technologies? *Atmos. Environ.* **106**, 373–374 (2015).
33. Penner, J., Lister, D., Griggs, D., Dokken, D. & McFarland, M. (eds) *Aviation and the Global Atmosphere* (Cambridge University Press, Cambridge, 1999).
34. Forster, P. M., Shine, K. P. & Stuber, N. It is premature to include non-CO_2 effects of aviation in emission trading schemes. *Atmos. Environ.* **40**, 1117–1121 (2006).
35. Shine, K. P., Fuglestvedt, J. S., Hailemariam, K. & Stuber, N. Alternatives to the Global Warming Potential for comparing climate impacts of emissions of greenhouse gases. *Clim. Change* **68**, 281–302 (2005).
36. Peters, G. P., Aamaas, B., Berntsen, T. & Fuglestvedt, J. S. The integrated global temperature change potential (iGTP) and relationships between emission metrics. *Environ. Res. Lett.* **6**, 044021 (2011).
37. Dallara, E. S., Kroo, I. M. & Waitz, I. A. Metric for comparing lifetime average climate impact of aircraft. *AIAA J.* **49**, 1600–1613 (2011).
38. Allen, M. R. et al. A solution to the misrepresentations of CO_2-equivalent emissions of short-lived climate pollutants under ambitious mitigation. *NPJ Clim. Atmos. Sci.* **1**, 16 (2018).
39. Cain, M. et al. Improved calculation of warming-equivalent emissions for short-lived climate pollutants. *NPJ Clim. Atmos. Sci.* **2**, 29 (2019).
40. Smith, M. A., Cain, M. & Allen, M. R. Further improvement of warming-equivalent emissions calculation. *NPJ Clim. Atmos. Sci.* **4**, 19 (2021).
41. Niklaß, M., Grewe, V., Gollnick, V. & Dahlmann, K. Concept of climate-charged airspaces: a potential policy instrument for internalizing aviation's climate impact of non-CO_2 effects. *Clim. Policy* **21**, 1066–1085 (2021).
42. Grewe, V. et al. Aircraft routing with minimal climate impact: the REACT4C climate cost function modelling approach (V1.0). *Geosci. Model Dev.* **7**, 175–201 (2014).
43. Proesmans, P.-J. & Vos, R. Airplane design optimization for minimal global warming impact. *J. Aircraft* **59**, 1363–1381 (2022).
44. Rogelj, J. & Schleussner, C.-F. Unintentional unfairness when applying new greenhouse gas emissions metrics at country level. *Environ. Res. Lett.* **14**, 114039 (2019).
45. Meinshausen, M. & Nicholls, Z. GWP* is a model, not a metric. *Environ. Res. Lett.* **17**, 041002 (2022).

46. Grewe, V. & Stenke, A. AirClim: an efficient tool for climate evaluation of aircraft technology. *Atmos. Chem. Phys.* **8**, 4621–4639 (2008).

47. Dahlmann, K., Grewe, V., Frömming, C. & Burkhardt, U. Can we reliably assess climate mitigation options for air traffic scenarios despite large uncertainties in atmospheric processes? *Transport. Res. Part D: Transport Environ.* **46**, 40–55 (2016).

48. Grewe, V., Stenke, A., Plohr, M. & Korovkin, V. D. Climate functions for the use in multi-disciplinary optimisation in the pre-design of supersonic business jet. *Aeronautical J.* **114**, 259–269 (2010).

49. Grewe, V. et al. Evaluating the climate impact of aviation emission scenarios towards the Paris Agreement including COVID-19 effects. *Nat. Commun.* **12**, 3841 (2021).

50. Dahlmann, K. *Eine Methode zur effizienten Bewertung von Maßnahmen zur Klimaoptimierung des Luftverkehrs*. Ph.D. thesis, LMU München (2011).

51. Grewe, V. et al. Mitigating the climate impact from aviation: achievements and results of the DLR WeCare project. *Aerospace* **4**, 34 (2017).

52. Meinshausen, M. et al. The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geosci. Model Dev.* **13**, 3571–3605 (2020).

53. Sausen, R. & Schumann, U. Estimates of the climate response to aircraft CO_2 and NOx emissions scenarios. *Clim. Change* **44**, 27–58 (2000).

54. Stenke, A., Grewe, V. & Ponater, M. Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias. *Clim. Dyn.* **31**, 491–506 (2008).

55. Burkhardt, U. & Kärcher, B. Process-based simulation of contrail cirrus in a global climate model. *J. Geophys. Res.* **114**, D16 (2009).

56. Ponater, M., Bickel, M., Bock, L. & Burkhardt, U. Towards determining the contrail cirrus efficacy. *Aerospace* **8**, 42 (2021).

57. Gelhausen, M. C., Grimme, W., Junior, A., Lois, C. & Berster, P. Clean Sky 2 Technology Evaluator—results of the first air transport system level assessments. *Aerospace* **9**, 204 (2022).

58. Yin, F., Grewe, V., Frömming, C. & Yamashita, H. Impact on flight trajectory characteristics when avoiding the formation of persistent contrails for transatlantic flights. *Transport. Res. Part D: Transport Environ.* **65**, 466–484 (2018).

Acknowledgements

Author contributions

Conceptualisation: L.M. and V.G. Methodology, investigation and visualisation: L.M. Supervision: K.D. and V.G. Original draft: L.M. Review and editing: L.M., K.D. and V.G. All authors read and approved the final manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01423-6.

Correspondence and requests for materials should be addressed to Liam Megill.

Peer review information *Communications Earth and Environment* thanks Takashi Sekiya and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Prabir Patra, Heike Langenberg. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

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