Analysis of Climate Metrics for Aviation

W.A. (Liam) Megill

TU Delft

This page has intentionally been left blank. The cover image is the author's own work.

Analysis of Climate Metrics for Aviation

by

W.A. (Liam) Megill

to obtain the degree of Master of Science
at Delft University of Technology,
to be defended publicly on July 5th, 2022 at 14:00.

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TUDelft

Abstract

The aviation industry continues to contribute to anthropogenic climate change. Recent studies have shown that the total radiative forcing from aviation is around three times higher than that from CO_2 alone. To account for the full effect of aviation, climate metrics are needed, which equate the environmental impact of various emissions and effects. However, there is currently no consensus on which climate metric should be used in aviation policy. This thesis systematically analyses existing climate metrics by: 1) comparing the responses to simple emission profiles; 2) investigating the sensitivity to changes in high-level aircraft design variables; 3) using a Monte Carlo simulation to determine inherent biases in metric calculation methods; and 4) evaluating the ability of climate metrics to estimate CO_2-equivalent emissions. It is concluded that the Average Temperature Response (ATR) is the most appropriate climate metric for aviation climate policy. However, it is found that the time horizon remains a subjective choice that must be chosen carefully depending on the climate objective. The GWP*, a newly proposed climate metric, is concluded not to be suitable as a climate metric because of its high variability and secondary time horizon. Nevertheless, it is recommended to investigate the potential of using the GWP* as a Micro Climate Model to further climatic understanding.

Preface

I was interviewed for a podcast last year and was asked "Who inspired you the most, and why did you choose to study aerospace engineering?" My answers: David Attenborough, and because I thought becoming a climate scientist would be too depressing. Oh, and because I loved flying. In reflection, I find this very interesting: After focusing on the design of sustainable aircraft in my studies, at AeroDelft and at Avy, I started this thesis to gain a better understanding of aviation emissions themselves and their effects on the environment. It turns out that this particularly interests me, so much so that I will start a PhD at the Institute for Atmospheric Physics at the DLR. Maybe becoming a climate researcher is something for me after all. Depressing? Perhaps. But extremely motivating, also.

This thesis essentially marks the end of my MSc degree and, therefore, my student journey here at the TU Delft. It is safe to say that the last two and a bit years have not been how I had envisioned the MSc degree. However, I am very happy that I can make the finishing touches to this thesis while sitting in the faculty. Seeing more students back on campus has been heartening, and being able to meet regularly with friends again was key to making this thesis successful.

Pandemic or not, I am extremely grateful for the experiences I have had here in Delft. I still remember being thrown head-first into the culture in my first week in the Netherlands, as the only non-Dutch member of WASUB at our introduction weekend. Others may disagree, but I still believe our turtle-esque propulsion system design was fantastic, and I will hang up my wing as soon as I own a wall that is big enough! My tradition of spending more time at the D:DREAM Hall than anywhere else continued the following year when Thomas and I set up AeroDelft. Those first years may have been tough, but my proudest moments were undoubtedly in watching AeroDelft flourish into a fantastic student team and inspire so many students. AeroDelft will remain a force to be reckoned with in sustainable aviation, mark my words!

This thesis would not have been possible without the continuous support from my friends, here in Delft and elsewhere. Thank you for making the last six (wow!) years so memorable. A big thank you especially to those who helped and supported me during my full-time AeroDelft year - the team would not exist without you! Kiitos Roosa, for listening to my ramblings and for making sure what I was writing was understandable. I promise I will use more full stops in the future. I would also like to thank my supervisors Volker and Kathrin for their help and dedication, and for asking the critical questions that helped bring me further. Lastly, I thank my parents, my sister and the rest of my family for their unwavering support and encouragement. I am very happy that I was guided towards engineering, although I do have to say that I am glad that aerospace triumphed over maritime in the end. I hope there wasn't any money involved in that bet.

Liam Megill
Delft, June 2022

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

AGTP & GTP - Absolute & Relative Global Temperature Change Potential

AGWP & GWP - Absolute & Relative Global Warming Potential

AIC -   Aircraft Induced Cloudiness

ATR & ATR-rel - Absolute & Relative Average Temperature Response (in literature, ATR is used for both interchangeably)

CORSIA - Carbon Offsetting and Reduction Scheme for International Aviation

EGWP* - Effective GWP*, i.e. GWP* using the ERF

ERF - Effective Radiative Forcing

ETS - European Union Emissions Trading System

GWP* - Flow-based metric developed from the GWP

IPCC - Intergovernmental Panel on Climate Change

LLCP - Long-Lived Climate Pollutant

RF -   Radiative Forcing

RFI -   Radiative Forcing Index or relative RF, i.e. RF / RF(CO_2)

SLCP - Short-Lived Climate Pollutant

SSP - Shared Socio-Economic Pathway

Definitions

Background Emissions Scenario - Temporal trajectory of global CO_2 and CH_4 concentrations, e.g. SSP2-4.5 (see Figure 3.1)

Emission Profile - Temporal trajectory of aircraft/fleet fuel use, e.g. P2020 (see Figure 3.2)

Fuel Scenario - Temporal trajectory of total aviation fuel use, e.g. Fa1 (see Figure 3.4)

Stakeholder - End-users of climate metrics, including policymakers, company decision-makers and scientists

1

Introduction

Since the First Assessment Report in 1990 of the IPCC, a large amount of evidence has been gathered linking anthropogenic emissions to a global warming of the climate. However, only in the Fourth Assessment Report (AR4) in 2007 did the IPCC unequivocally announce that the climate was warming due to human activity. In 2015, the Paris Agreement sought to set new global policies to reduce greenhouse gas emissions to limit global warming to well below 2°C above pre-industrial levels, and to pursue efforts to limit warming further to 1.5°C. Key to the Paris Agreement was the creation of Nationally Determined Contributions (NDCs), which were submitted by contributing member states to reflect their mitigation and adaption strategies. Unfortunately, the Sixth Assessment Report (AR6) finds that current NDCs will fail to meet the greenhouse gas emission reductions required to limit global warming to 2°C above pre-industrial levels (Intergovernmental Panel on Climate Change, 2021).

The aviation industry by its international nature does not fit into the concept of NDCs and is, therefore, not included in the Paris Agreement. Although some member states have placed restrictions and requirements on domestic aviation, aviation in total continues to increase its contribution to global warming. In 2005, aviation emissions accounted for 5% of net anthropogenic radiative forcing, which increased markedly before the COVID-19 pandemic started (Lee et al., 2021). Figure 1.1a) shows the trajectory of revenue passenger kilometres (RPK) compared to exponential growth (Grewe et al., 2021). Drops in global aviation in response to various crises has had little to no impact on the aviation industry's growth. It is also clear to see that the increases in efficiency due to higher passenger load factors and improvements in engine fuel efficiency and aerodynamic performance are not enough to offset the climate effects of the industry's growth. Figure 1.1b) shows potential future emissions of CO_2 from the industry (Grewe et al., 2021).
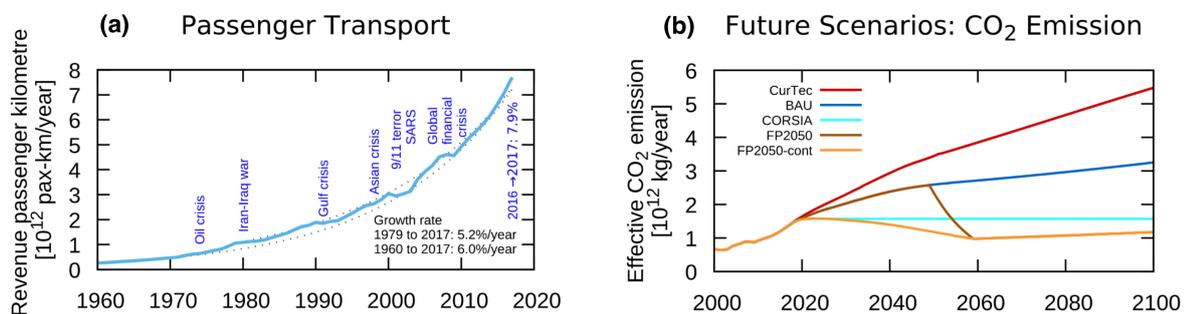
Figure 1.1: Trajectory of global aviation emissions, taken from Grewe et al. (2021). **(a)** shows the trajectory of the revenue passenger kilometre (RPK) since the year 1960, compared to an exponential trajectory (dotted line); **(b)** shows the expected trajectory of CO_2 emissions for five future scenarios

The impacts of aviation on the climate include much more than simply CO_2 from the combustion of hydrocarbons. Recent studies have shown that the total radiative forcing from aviation is around three times higher than that from CO_2 alone (Lee et al., 2021). Especially important are NO_x emissions, aircraft induced cloudiness and aerosol emissions. However, current aircraft have been designed to minimise their direct operational costs, and not to minimise their global warming impact (Proesmans and Vos, 2021). Furthermore, existing international schemes such as CORSIA (Carbon Offsetting and Reduction Scheme for International Aviation) and the European Emissions Trading System (ETS) only consider CO_2 emissions.

Climate metrics are used to equate the impact of various emissions and effects such that their impact on the environment can be analysed. Climate metrics can aid policymakers and company decision-makers place their focus and develop effective emissions mitigation strategies, thus requiring climate metrics that are robust, transparent, consistent and understandable. The standard climate metric used in international climate policy is the Global Warming Potential (GWP), which compares the time-integrated change in the global energy balance (radiative forcing, RF) of a non-CO_2 emission to that of a CO_2 pulse (Fuglestvedt et al., 2003). However, the GWP has been heavily criticised, primarily due to its dependence on the time horizon over which the integration is performed (e.g. Allen et al., 2016), for which there is no obvious choice. Numerous alternatives to the GWP have been proposed, including the Global Temperature Change Potential (GTP, Shine et al., 2005), Average Temperature Response (ATR, Dallara, 2011) and the GWP* (Allen et al., 2018), a new method of using the GWP. There is currently no consensus on which climate metric is best suited for aviation policy.

The aim of this work is to systematically analyse existing climate metrics for the different aviation climate objectives. This is done by first analysing their responses for pulse, sustained and increasing emissions to gain an understanding of their dependencies on the time horizon and background emissions scenarios, and to provide a comparison with existing literature. A second analysis involves the development of representative future fleets to investigate the sensitivity of the climate metrics to changes in aviation-specific emission species and top-level aircraft design parameters. A subsequent analysis determines the inherent bias of the climate metrics to these parameters by comparing the change in climate metric value and the change in temperature for all pairs of fleets. Finally, for comparison of the aviation industry to other industries or for the introduction of climate metrics into market-based schemes such as the ETS, the ability of climate metrics to estimate CO_2-equivalent emissions is evaluated.

This research adds to the scientific discussion on climate metrics by quantifying the suitability of climate metrics for use in aviation. Using the methods laid out in this research, stakeholders (policy-/decision-makers and scientists) should be able to determine which climate metric to use. The main research objective is therefore formulated as:

Recommend the best-suited climate metrics for existing and proposed aviation climate objectives by systematically analysing the response of existing physical climate metrics using the climate-chemistry response model AirClim.

The relevant literature is summarised in Chapter 2, with the focus on the climate impacts of aviation and the design of climate metrics. Chapter 3 develops the list of climate objectives and requirements, and describes the methodology used. The results are shown in Chapter 4 and discussed in Chapter 5. Finally, Chapter 6 concludes this work and recommends areas for further research.

2

Literature Review

This chapter provides an overview of the most relevant literature, used to guide the methodology. Portions of the following text have already appeared in a more detailed literature study, performed in conjunction with this work. The chapter is split into four main sections. Section 2.1 explores aviation emissions and their impacts on the environment. In this work, these impacts will be determined using the response model AirClim, which is described in Section 2.2. The purpose and design of climate metrics is explained in Section 2.3, followed by a description of existing climate metrics in Section 2.4.

2.1. Climate Impacts of Aviation

The vast majority of aircraft are currently propelled by the combustion of hydrocarbon fuels. The most common fuels for commercial aviation are Jet A and Jet A-1, which are both forms of kerosene that have been adapted to have different freezing points. Jet A-1 has an average ratio of carbon to hydrogen of 12:23 and is thus often written as $C_{12}H_{23}$ (Lee et al., 2010). It also contains approximately 400 ppm (mass) of sulphur. Ideal combustion of aviation fuel thus produces carbon dioxide (CO_2) and water vapour (H_2O). Table 2.1 provides the average emission indices for Jet A-1 for reference (Lee et al., 2021).

In real combustion, carbon monoxide (CO), unburned hydrocarbons (C_xH_y, often also denoted UHC), soot, nitrogen oxides (NO_x) and sulphur oxides (SO_x) are also emitted (Lee et al., 2021). The emission rate of these products is influenced by a number of factors, as shown in Table 2.1. The impact of these emissions on the environment also varies with geographic location, altitude and the current weather (Grewe et al., 2017; Frömming et al., 2021), as is described later in this section. Other sources of emissions are from (lubricant) oils (Dakhel et al., 2007) and metals within the fuel such as iron, copper and zinc (Intergovernmental Panel on Climate Change, 1999), but these have not been considered further in this research.

Table 2.1: Average emission indices of aviation emissions per kg of Jet A-1, adapted from Lee et al. (2021), with their main dependencies.

Emission	Emission Index	Dependence
CO_2	3.16 kg/kg	Thrust setting
NO_x	15.14 g/kg	Thrust setting (equivalence ratio, temperature), combustion technique
H_2O	1.231 kg/kg	Thrust setting
Soot	0.03 g/kg	Thrust setting (equivalence ratio, pressure, temperature), quality of fuel injection
SO_2	1.2 g/kg	Fuel sulphur content

The following sections describe the mechanisms leading to changes in radiative forcing of the most important aviation emissions. Section 2.1.1 describes the impact of CO_2 and Section 2.1.2 the effect of NO_x on ozone, methane and water vapour. Section 2.1.3 considers water vapour emissions and aircraft induced cloudiness. Finally, Section 2.1.4 describes the direct and indirect impact of aviation sulphate and soot aerosols.

2.1.1. Carbon Dioxide

Carbon dioxide (CO_2) is a product of complete combustion and is the most well-known and dominant anthropogenic greenhouse gas. It is also the most understood aviation emission (Lee et al., 2021). Carbon dioxide in the atmosphere is part of the global carbon cycle, which is a series of reservoirs connected by carbon fluxes. The main reservoirs are the atmosphere, biosphere, ocean, rocks and sediments. Burning fossil fuels contributes to the flux between the biosphere and the atmosphere, which has resulted in a 48% increase in atmospheric CO_2 compared to its natural state before industrialisation in 1750 (Intergovernmental Panel on Climate Change, 2021, cf. Fig. 5.12), as well as larger carbon fluxes into the oceans. The presence of CO_2 in the atmosphere results in larger absorption of outgoing (black-body) radiation from Earth, thus trapping more energy and warming the atmosphere and ocean.

Due to the complex inter-connectivity between reservoirs, CO_2 emissions cannot be modelled by a single exponential decay function like other pollutants (Intergovernmental Panel on Climate Change, 2001; Shine et al., 2007). The transfer of carbon between reservoirs means that no single timescale can be used and various modes must be used for modelling. A well-known linear response model was developed by Sausen and Schumann (2000) based on the work of Hasselmann et al. (1993), which also became the basis for AirClim as described in Section 2.2. It uses the convolution of a radiative forcing function and an impulse response function to determine the change in atmospheric CO_2 concentration (see Eq. (2.1)).

2.1.2. Nitrogen Oxides

In the context of aviation, nitrogen oxides (NO_x) refers to the combination of nitric oxide (NO) and nitrogen dioxide (NO_2). The emission index for NO_x ($EINO_x$) is usually given in g NO_2 / kg fuel, where NO is converted to NO_2 by atomic weight.

NO_x emissions change the atmospheric chemistry and shift existing balances between species. NO_x emissions in the stratosphere, for example from supersonic aircraft, result in negative RF by reducing the concentration of O_3, in turn allowing more solar radiation to enter the troposphere (Grewe et al., 2010; Lee et al., 2010). Tropospheric chemistry is more complex. Figure 2.1 shows the interaction of the NO_x and HO_x (reactive hydrogen) cycles in the troposphere. The main positive RF due to NO_x emissions can be attributed to the increase in tropospheric ozone (O_3) due to the forward direction of the reaction $HO_2 + NO \longrightarrow NO_2 + OH$ being favoured, as well as the reduction in O_3 loss by $O_3 + HO_2 \longrightarrow 2 O_2 + OH$ since less HO_2 is present (Grewe et al., 2002; Lee et al., 2010). The main negative RF is due to the destruction of methane (CH_4) by, for example, $CH_4 + OH \longrightarrow CH_3 + H_2O$. Due to the long lifetime of CH_4 (Stevenson, 2004) and because CH_4 lifetime reduction is on a global scale (Myhre et al., 2011), the effect of the methane reduction persists much longer than the ozone increase, the latter of which is negligible after 6 months (Stevenson, 2004). A secondary negative RF is due to methane-induced long-term decreases in O_3, also called the Primary Mode Ozone effect (PMO, Stevenson, 2004).
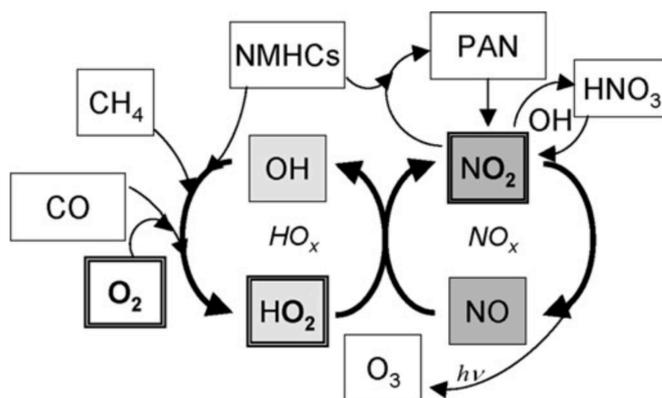
Figure 2.1: The interaction between the HO_x and NO_x cycles in the troposphere, showing the main ozone related chemistry, taken from Grewe (2009). NO_x emissions increase the reaction of HO_2 and NO to OH and NO_2, thus resulting in higher ozone concentrations.

The effectiveness of tropospheric ozone production is dependent on a number of different factors, including the ratios of NO to NO_2 and OH to HO_2, but also background NO_x levels (Stevenson and Derwent, 2009), the rate of atmospheric mixing (Berntsen et al., 2005), and importantly also the location, altitude and meteorological conditions of the NO_x emission (Stevenson, 2004; Frömming et al., 2012; Köhler et al., 2013; Frömming et al., 2021). The impacts of other NO_x-induced chemistry, such as the longer-term reductions in O_3 (Stevenson, 2004) and stratospheric water vapour (Lee et al., 2021), the direct formation of nitrate aerosols and the indirect enhancement of sulphate aerosols are still highly uncertain (Lee et al., 2021). The sign of NO_x radiative forcing is generally accepted to be positive within 90% certainty (Lee et al., 2021).

2.1.3. Water Vapour and Aircraft Induced Cloudiness

Although water vapour is a greenhouse gas and a direct product of complete combustion, it has a short 1-2 week lifetime in the troposphere (Intergovernmental Panel on Climate Change, 1999) and low climate impact when released there. Frömming et al. (2012), for example, find the accumulated water vapour emissions to be less than 1% of the background water vapour. In the stratosphere, there are fewer loss processes for water vapour (Frömming et al., 2012), making water vapour emissions at higher altitudes the dominant component of radiative forcing from supersonic aircraft (Lee et al., 2010).

Water vapour emissions are a precursor for aircraft induced cloudiness (AIC). Broadly, AIC consists of three different cloud formations: linear contrails, contrail-cirrus and soot cirrus (Lee et al., 2010). For a contrail to form, three conditions must be met. First, during mixing of the exhaust with the ambient air, the mixture must become supersaturated with respect to water. This allows water vapour to condense, form droplets and then freeze, becoming visible. Secondly and thirdly, the ambient temperature and humidity respectively must be within certain boundaries. The threshold conditions at which contrails begin to form is calculated using the Schmidt-Appleman criterion (SAC, Schmidt, 1941; Appleman, 1953).

Various aerosols, either from the engine exhaust or from the environment, can act as cloud condensation nuclei for the water droplets to form. For example, soot cirrus are defined as

cirrus clouds that have been created or altered by the presence of heterogeneous ice nuclei from aircraft (Lee et al., 2010). If the ambient air is also supersaturated with respect to ice, the contrail persists (Lee et al., 2010). Due to the atmospheric mixing and shear, the contrails take on a much larger area and are then known as contrail-cirrus (Lee et al., 2010).

In general, two main processes can be identified. Terrestrial (black-body) radiation trapped and absorbed by contrail-cirrus or cirrus is emitted at a lower temperature from the top of the cloud, resulting in a positive radiative forcing (Bock and Burkhardt, 2016). A negative radiative forcing due to the scattering of incoming shortwave solar radiation by ice crystals in the contrail, contrail-cirrus or cirrus also exists (Schumann et al., 2012; Bock and Burkhardt, 2016). However, based on current knowledge, the radiation trapping mechanism dominates and AIC contributes the largest amount of positive radiative forcing of all aviation emissions (Lee et al., 2021). Scientific understanding and modelling has improved substantially in recent years, however, the processes and effects remain highly uncertain. A number of authors have, for example, shown low efficacies for global mean surface temperature increases as a result of AIC (Ponater et al., 2005; Bickel et al., 2020; Ponater et al., 2020). Adding to this is the variation of the climate impact with geographical position, altitude and time (Irvine et al., 2014; Bier et al., 2017) - Stuber and Forster (2007), for example, attribute 60% of contrail radiative forcing to night-time flights, when solar radiation scattering does not occur.

2.1.4. Aerosols

Aviation aerosols are small compounds of carbon, sulphur and nitrogen. The impact of aerosols on the climate has been researched extensively in recent years, but especially the interaction between the particles and clouds remains uncertain (Lee et al., 2021). Sulphur oxides and soot have the largest impact for aviation.

Sulphur aerosols scatter incoming, solar radiation and absorb very little outgoing, terrestrial radiation, thus resulting in a net cooling (Intergovernmental Panel on Climate Change, 1999). A number of studies have analysed the radiative forcing of sulphate aerosol-cloud interactions and have all found a net cooling effect, although some uncertainty remains in the exact value (Gettelman and Chen, 2013; Kapadia et al., 2016).

Soot aerosols, from either black (BC) or organic (OC) carbon, form as a product of incomplete combustion. In comparison to sulphate aerosols, soot aerosls absorb solar radiation effectively and have a net positive direct radiative forcing (Fuglestvedt et al., 2010; Lee et al., 2021). The magnitude and sign of soot aerosol-cloud interaction is highly uncertainty: Penner et al. (2018) shows a substantial cooling effect, but this could well be less negative if other, more efficient aerosols are already present (Penner et al., 2018; Gettelman and Chen, 2013). Other studies, such as Zhou and Penner (2014), find both positive and negative radiative forcing. Lee et al. (2021) provides a more detailed overview of the uncertainties in aerosol impacts.

2.2. Description of the Response Model AirClim

AirClim (Grewe and Stenke, 2008) is a non-linear climate-chemistry response model that estimates the atmospheric response and near surface temperature change to emissions of CO_2, H_2O, contrails and NO_x, which leads to changes in CH_4, O_3 and PMO. Initially written primarily for the analysis of supersonic aircraft by Grewe and Stenke (2008), it has been since updated

to include higher resolutions in subsonic levels (Fichter, 2009), the transition of contrails to cirrus (Dahlmann, 2011), and other functionality such as an internal Monte Carlo simulation (Dahlmann et al., 2016b). AirClim has been used in multidisciplinary optimisation schemes (Grewe et al., 2010) as well as for air traffic trajectory analyses (Niklaß et al., 2021; Matthes et al., 2021). Since AirClim makes use of the impacts of yearly-averaged weather patterns, it is, however, not possible to use AirClim to calculate the impact of different weather patterns.
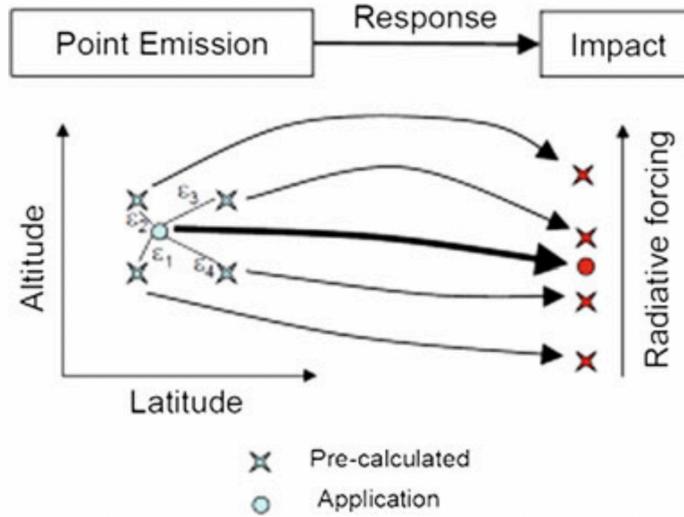
Figure 2.2: Graphical overview of the calculation method used by AirClim, taken from Grewe and Dahlmann (2012). An emission at a given altitude and latitude, shown with a circle, can be described using a linear combination of pre-calculated emissions (different ϵ). AirClim uses the same linear combination to estimate the response based on the pre-calculated responses.

AirClim is an extension to the linear response model for CO_2 developed by Sausen and Schumann (2000). AirClim combines emission data with pre-calculated, altitude- and latitude-dependent data obtained from 85 steady-state simulations performed with the E39/CA climate-chemistry model (Stenke et al., 2008), and from the ECHAM4 climate-chemistry model for contrail cirrus (Burkhardt and Kärcher, 2009; Dahlmann et al., 2016a). The pre-calculated data is created with normalised emissions in pre-defined regions, effectively giving an altitude- and latitude-dependent sensitivity. From this data, the change in atmospheric composition, radiative forcing and the temporal temperature response can be estimated. Figure 2.2 provides a graphical overview of the calculation method.

$$\Delta T = \int_{t_0}^{T} G\left(t - t'\right) \text{RF*}\left(t'\right) \mathrm{d}t' \quad \text{where} \quad G\left(t - t'\right) = \alpha \, e^{-\frac{t - t'}{\tau}} \tag{2.1}$$

For NO_x, H_2O and contrails, the temperature change is estimated using a convolution of the Green's function and the radiative forcing of a species normalised to a doubling of CO_2 (RF*), as shown in Equation (2.1) (Grewe and Stenke, 2008). For species that have lifetimes longer than one year but much shorter than a decade, such as stratospheric NO_y and water vapour, the lifetime is also taken into account using a simple linear differential equation (Grewe and Dahlmann, 2012). The calculation of methane depletion is done differently due to its lifetime of close to a decade. The change in the atmospheric lifetime of methane is calculated

using a differential equation, being the difference between the unperturbed and the perturbed situations (Grewe and Stenke, 2008; Grewe and Dahlmann, 2012). A more detailed overview of the calculation method can be found in Grewe and Stenke (2008).

AirClim is used for this analysis because it was designed specifically for aviation emissions, but also due to its low computational cost and its high usage in previous years for similar studies on the climate impacts of aviation (Grewe et al., 2010; Niklaß et al., 2021; van der Maten, 2021, for example). Since AirClim is a response model that makes use of pre-calculated data, it is not as accurate as more complex climate models. AirClim has been validated against E39/C, where it is shown to reproduce radiative forcing within 15% (Grewe and Stenke, 2008), and other models. Use of a more complex climate chemistry or circulation model is beyond the scope of this research, however the work presented here should not be dependent on the model, as is discussed further in Chapter 5.

2.3. Purpose & Design of Climate Metrics

The purpose of climate metrics is to relate the emissions of gases and aerosols to their consequences on the climate and/or society (Fuglestvedt et al., 2003). Plattner et al. (2009) identify the main uses of climate metrics as exchange rates to be implemented, for example, in international emissions trading schemes, and more generally as tools that serve to help stakeholders better understand the relative impacts of different emissions on society and the environment. It is not practical for decision-makers both in government and in industry to use complex climate models since these are slow and require expertise to use. Climate metrics must be designed such that they can be implemented without in-depth knowledge about the underlying climatic processes. However, even with the required simplifications, the climate metrics must still accurately reflect the climatic response to an emission to ensure that decisions are based on current scientific understanding (Niklaß et al., 2019).

Climate metrics must consider the impact of various different aerosol and gas emissions, which can lead to several issues, as a number of authors have noted (Tol et al., 2008). Generally, a differentiation is made between long-lived and short-lived climate pollutants (LLCPs and SLCPs respectively). CO_2 is an example of an LLCP, also known as a cumulative climate pollutant, whereas methane (CH_4) and NO_x are SLCPs. The design of a climate metric can have a major impact on the emphasis placed on SLCPs compared to LLCPs (Allen et al., 2016), which, in turn, could influence which policies are put into place. This can lead to a perceived ambiguity in the results obtained since each climate metric provides an answer to a different climate question (Grewe and Dahlmann, 2015). This is particularly important for aviation, where around two-thirds of radiative forcing originates from regionally dependent SLCP emissions (Lee et al., 2021).

To identify different climate metric options, Fuglestvedt et al. (2010) provide a cause-and-effect chain from emissions to damages, as shown in Figure 2.3 (originally in Fuglestvedt et al. (2003)). Fuglestvedt et al. (2003) argue that an 'ideal' climate metric would be at the bottom of the chain to directly compare the mitigation costs to the damage an emission causes on the environment and on society. Such a climate metric is more tangible and thus relevant to society. However, as many authors note (Fuglestvedt et al., 2003; Shine et al., 2007; Fuglestvedt et al., 2010; Dallara et al., 2011), the further down the chain a climate metric is, the more difficult its calculation is, the more assumptions need to be made and thus the larger the uncertainties and inaccuracies in our knowledge of the climatic processes and models.

The development of a climate metric is always a compromise between its relevance to society and the level of uncertainty. The most common climate metrics used by (inter-)governmental policymakers (GWP, GTP, as described in subsequent sections) are only on the third or fourth rung of the ladder in Figure 2.3.
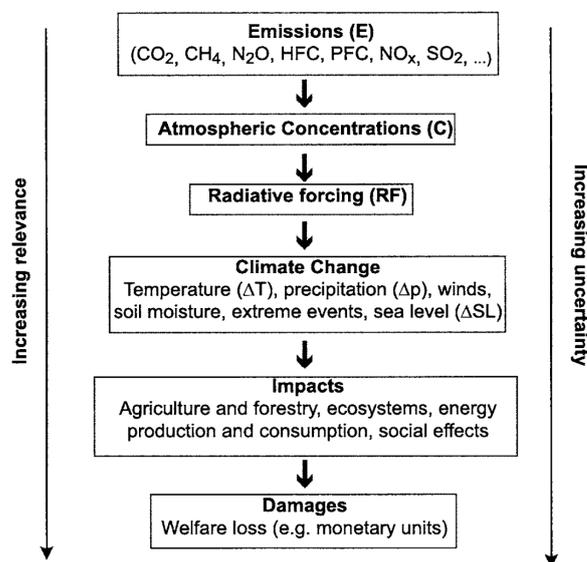
Figure 2.3: Cause-and-effect chain from emissions to damages, from Fuglestvedt et al. (2003). Climate metrics further down the chain are more relevant to society, but are also more uncertain given the larger number of assumptions that need to be made and the large number of inaccuracies in our knowledge of climatic processes and economic consequence models.

Climate metrics can be categorised (Fuglestvedt et al., 2003) 1) by their behaviour on the temporal axis - whether a climate metric considers the *instantaneous* change at a certain time (e.g. RF, GTP) or a *temporal integration* over a certain time horizon (e.g. GWP); 2) by whether the impact is related to the *rate* or *level* of change; 3) by whether the climate metric is *global* or *regional*; and 4) by the type of climatic and/or societal consequence considered. The latter categorisation considers *physical climate metrics*, which simply consider the impact an emission has on the chemistry and composition of the climate; *cost-effective climate metrics*, which are used to determine how to achieve goals at the lowest cost; and *cost-benefit climate metrics*, which aim to minimise the sum of the costs of climate change and of mitigation. Economists have developed climate metrics such as the Global Damage Potential (GDamP, Kandlikar, 1996) and the Global Cost Potential (CGP, based on Manne and Richels, 2001), but since impacts must be given a monetary value for intercomparison that can be subjective (Fuglestvedt et al., 2010), only physical climate metrics are considered in this study.

2.4. Existing Climate Metrics

This section provides an overview of the climate metrics used in this study. It is important to note that this list is not exhaustive and that a large number of other climate metrics exist, mostly as tools to answer specific climate questions. Climate metrics based on the emission mass or resulting atmospheric concentrations are not meaningful for aviation, therefore all climate metrics are on the third or fourth rung of the ladder in Figure 2.3.

2.4.1. Radiative Forcing

Radiative forcing is a net flux imbalance at a location in the atmosphere, caused by a change in atmospheric composition. This definition can be further refined, as also shown in Figure 2.4. If the temperature gradient of the atmosphere is not allowed to adapt to the changes at all, the calculated flux imbalance at the tropopause is defined as the *instantaneous radiative forcing* (Figure 2.4a) (Intergovernmental Panel on Climate Change, 1995; Fuglestvedt et al., 2003). Since the timescale for the stratosphere to adapt to a perturbation is short - of the order of months - the *stratospheric-adjusted radiative forcing* (Figure 2.4b) is often used, for which the troposphere and surface temperatures are kept constant (Fuglestvedt et al., 2003; Hansen et al., 2005) whilst the stratosphere is allowed to adapt. The *zero-surface-temperature-change radiative forcing* allows the whole atmosphere to adjust, but keeps either the global mean surface temperature or ocean temperature fixed (Figure 2.4c and d respectively) and calculates the flux imbalance at the top of the atmosphere (TOA). This is also the basis of the *effective radiative forcing* (ERF).
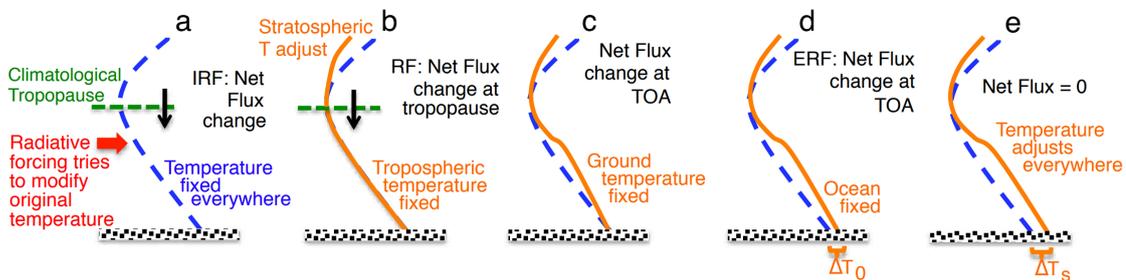
Figure 2.4: Main definitions of radiative forcing, from Intergovernmental Panel on Climate Change (2013), originally from Hansen et al. (2005). The definitions are: **(a)** instantaneous forcing; **(b)** stratospheric-adjusted forcing; **(c)** zero-surface-temperature-change forcing; **(d)** fixed sea surface temperature forcing, allowing atmospheric temperature and land temperature to adjust; **(e)** full feedback response, calculated using a climate model and allowing the temperature to adjust everywhere.

Radiative forcing can be seen as a shortcut for the global-mean surface temperature change at equilibrium $\Delta T_s = \lambda \, \text{RF}$, where $\lambda \, [\text{K/(Wm}^2)]$ is a climate sensitivity parameter and RF is the radiative forcing. The climate sensitivity parameter includes a 'no-feedback' climate response, which is well known, and a feedback and coupling response that is less well understood (Fuglestvedt et al., 2003). Since a large number of potential couplings exist, it is only possible to compare radiative forcings on a global-mean scale effectively, thus ignoring local effects (Wuebbles et al., 2010). Not all radiative forcings results in the same global-mean temperature change at equilibrium: The effectiveness of a climate agent in causing a temperature change, compared to the baseline of CO_2, is called its *efficacy* (Hansen et al., 2005; Wuebbles et al., 2010).

As a climate metric, radiative forcing in the form of RF, ERF and RFI (radiative forcing index, the ratio of a sector's total RF to the RF of its CO_2 emissions) is commonly used for aviation emission analysis (for example Lee et al., 2010, 2021). The efficacies of, for example, AIC and aerosols are hard to quantify and are a significant source of uncertainty. For the purposes of this study, the efficacy calculated by Ponater et al. (2006, cf. Table 1) using ECHAM4 will generally be used. It should also be noted that radiative forcing is a backward-looking climate metric and may not be appropriate for the analysis of future aviation emissions and policy-

making (Lee et al., 2009). However, the temporal evolution of the radiative forcing serves as the basis for many other climate metrics.

2.4.2. Global Warming Potential (GWP)

The Global Warming Potential (GWP) is the ratio between the temporal integrations of the radiative forcing caused by a unit pulse emission of a climate agent and that of the reference gas CO_2 (Fuglestvedt et al., 2003). The GWP was introduced by Rodhe (1990) and Derwent (1990) and was directly included in the first assessment of the IPCC (Intergovernmental Panel on Climate Change, 1990), becoming the most commonly used climate metric in international and national policy-making. It can be calculated as follows (Fuglestvedt et al., 2003),

$$\text{GWP}(H)_i = \frac{\int_0^H \text{RF}_i(t)\,\mathrm{d}t}{\int_0^H \text{RF}_{CO_2}(t)\,\mathrm{d}t} = \frac{\text{AGWP}(H)_i}{\text{AGWP}(H)_{CO_2}} \tag{2.2}$$

where i is the emission gas, H the time horizon and t the time. RF_i and RF_{CO_2} are the radiative forcings of the considered gas and of CO_2, as described in the previous section. The AGWP is the Absolute GWP for a single emission species.

The AGWP was shown by Irvine et al. (2014) to be equivalent to the Energy Forcing (EF) climate metric developed by Schumann et al. (2011). The EF calculates the energy per unit flight distance for a single flight and is a useful climate metric for the analysis of contrails. Conversion from the AGWP to the EF [GJ/km] for contrails follows the equation,

$$\text{EF} = \frac{\text{AGWP}}{A_{\text{cont}}}\, A_{\text{Earth}}\, t_{\text{year}}\, W_{\text{cont}} \tag{2.3}$$

where $\text{AGWP}/A_{\text{cont}}$ [$\text{Wm}^{-2}\text{yr km}^{-2}$] is the specific AGWP for a contrail coverage area A_{cont} [km^2], A_{Earth} [m^2] is the area of the Earth, t_{year} [s] is the number of seconds in a year and W_{cont} [km] is the contrail width. For other emission species, the specific AGWP per kg of emission species can be used with an estimation of the emission rate per flown kilometre. Since the AGWP and EF are equivalent, only the AGWP is used in this research.

Since its inception in 1990, the GWP has been heavily criticised, primarily due to its dependence on the time horizon H over which the temporal integration is performed. While it can be useful for a stakeholder interested in a certain effect to choose the time horizon depending on the goal (Shine et al., 2007), this can also lead to exploitable ambiguity. It is possible for two opposing groups to argue completely contradicting statements but receive support from the GWP simply by using different time horizons. Ocko et al. (2017) mention, for example, studies comparing gas and coal, or vegan and non-vegan diets.

Perhaps the most important criticism of the GWP is, therefore, that there is no obvious choice for the time horizon. The GWP is most commonly expressed with a time horizon of 100 years (GWP_{100}), and less commonly also with 20 and 500 years. Rodhe (1990), compromising between short-term and long-term climate impacts, "somewhat arbitrarily" chose a time horizon of 100 years. Numerous authors, for example Allen et al. (2016), have criticised the dominant role of the GWP_{100} for this reason, arguing that there is "no particular justification" for the 100-year time horizon to be used to inform policy decisions, especially since most climate goals focus on keeping global average temperatures below a certain limit value.

The choice of time horizon effectively hides inherent trade-offs between policy goals which aim to limit either short- or long-term impacts (Ocko et al., 2017; Fuglestvedt et al., 2010). An example of this trade-off is clear from the work of Allen et al. (2016), who show that GWP_{100} effectively measures the warming 20-40 years after the time of emission. As a result, current SLCP emissions are given a higher priority when using the GWP_{100}.

That a global *warming* potential does not actually show warming is also highly problematic. The GWP, a measure of temporally integrated radiative forcing, is an abstract concept. Since policymakers would ideally need to thus understand what the GWP is showing before making decisions (cf. requirements for climate metrics in Section 3.3), it could be argued that discussing temperature change directly would be a much more suitable climate metric for policy-making (Fuglestvedt et al., 2010). That the time horizon, for example 100 years, does not represent a change in energy balance or temperature at 100 years after the emission but rather 20-40 years as demonstrated by Allen et al. (2016), is further cause for concern.

To economists, the (monetary) values of SLCP emissions, with atmospheric lifetimes of up to a few years, are initially low and grow with time as the deadline approaches. This mechanism is not captured by a time-invariant climate metric such as the GWP_{100} (Michaelis, 1992; Bradford, 2001; Shine et al., 2007). Manne and Richels (2001) further suggest that using a purely physical climate metric such as the GWP to determine which policies should be implemented will result in unnecessarily costly mitigation. They compare the GWP_{100} to a climate metric that analyses the price of emission for CH_4 and show that use of the GWP inherently assumes that the trade-offs between which gases should be mitigated are constant with time and, importantly, independent of the mitigation goal.

Even with the large amount of criticism, the GWP has remained the dominant climate metric at a national and international level. The high level of acceptability amongst stakeholders is likely due to the perceived transparency and ease of application (Shine et al., 2005). For example, the IPCC meeting (Plattner et al., 2009) came to the conclusion that no alternative climate metric has achieved the same standard, an argument which is still commonly made today. Ocko et al. (2017) argue that continuing to debate which climate metric is best or trying to find better climate metrics could lead to a slower uptake in policies that aim to combat climate change.

Since the majority of radiative forcing is attributed to non-CO_2 emissions, the GWP is particularly problematic for aviation emissions (Fuglestvedt et al., 2010). The Intergovernmental Panel on Climate Change (1999) and Wit et al. (2005), for example, argue that defining a suitable GWP for aircraft is questionable at best. As the main aviation emission reviews such as Lee et al. (2021) show, the GWP is highly dependent on the time horizon and does not reflect the total impact of aviation. Nevertheless, given that the GWP has been the climate metric of choice for the IPCC, the GWP has been used for aviation policy-making and aircraft design, for example Svensson et al. (2004). With the creation of the GWP*, as discussed in the next section, and the more varied use of the ATR (cf. Section 2.4.5), it is possible that the GWP will no longer be used as frequently aviation policy.

2.4.3. GWP*

Radiative forcing due to cumulative pollutants with long atmospheric lifetimes, such as CO_2, scales with the cumulative integral, i.e. the total amount of emissions released, making a climate metric such as the GWP ideal. However, given their short atmospheric lifetimes, ra-

diative forcing due to SLCPs depends on the SLCP emission rate and lifetime (Allen et al., 2016, 2018). A decrease in the SLCP emission rate will result in a decrease in temperature, but this is not reflected in conventional climate metrics such as the GWP. Instead, a decrease in SLCP emissions will show a further warming since the temporal integral of the radiative forcing is used. Using the GWP to achieve temperature stabilisation by, for example, aiming for net zero CO_2-equivalent (CO_2-e) emissions would, therefore, erroneously require an indefinite cooling trend (Allen et al., 2018). This is demonstrated by Tanaka and O'Neill (2018), who show that net-zero CO_2-e emissions are not required to achieve the Paris Agreement warming limit of 1.5°C.

Allen et al. (2018) compare cumulative CO_2-e emissions and CO_2-e emission rates calculated using GWP_{100} to peak warming from the IPCC AR5 scenario database and note that there are a number of non-linearities. Since the relationship between CO_2 emissions and temperature response is approximately linear, a truly equivalent emission should show the same linear behaviour. However, for constant or decreasing SLCP emissions, such a linear trend between CO_2-e emissions and warming does not exist (Cain et al., 2019). Given this misrepresentation of the SLCPs on global temperature when converted to CO_2-e emissions, Allen et al. (2018) propose a new use of the GWP, denoted GWP*.

The GWP* is designed to approximate CO_2 forcing-equivalent (CO_2-fe) emissions (Zickfeld et al., 2009), that do result in temperature stabilisation when reduced to a net zero rate, but that require a carbon cycle model to evaluate. This is done by equating a sustained increase in an SLCP emission rate with a one-off pulse emission of CO_2, averaged over a time period following the increase (Allen et al., 2018). Specifically, a sustained increase in an SLCP emission rate ΔE_{SLCP} is equal to a pulse emission of $\Delta E_{SLCP} \times GWP_H \times H$, and is distributed over Δt = 20 years following the increase. The pulse emission quantity $GWP_H \times H$ is denoted as CO_2-e*. Here, GWP_H is the standard GWP value for the SLCP for the time horizon H. A similar equivalency can be made for radiative forcing, where an increase in the radiative forcing ΔF_{SLCP} is equal to a pulse emission of $\Delta F_{SLCP} \times H/AGWP_{H(CO_2)}$, also spread over Δt = 20 years. Here, $AGWP_{H(CO_2)}$ is the absolute global warming potential of CO_2 over a time horizon H.

The choices of Δt = 20 years and H = 100 years are chosen by Allen et al. (2018) to correspond with common emission mitigation goals and norms set by the IPCC. A major advantage of the GWP* is that it is insensitive to the time horizon as long as it is chosen to be much greater than the lifetime of the SLCPs (cf. Allen et al. (2018) and Section 4.1). In contrast, as described in the previous section, the GWP is very sensitive to the time horizon for SLCP emissions. Smith et al. (2021) show that Δt = 20 years works for SLCPs that have a lifetime of about one decade or less. The choice of averaging time does not, however, affect the cumulative emissions, and thus even though a shorter time would be arguably better for SLCPs with shorter lifetimes, Smith et al. (2021) propose to continue using 20 years for all SLCPs for simplicity.

Where conventional CO_2-e emission rates are calculated as $E_{CO_2\text{-e}} = E_{SLCP} \times GWP_H$, the emission rates of CO_2-e* can be calculated by (Allen et al., 2018),

$$E_{CO_2\text{-e}^*}(t) = \frac{\Delta E_{SLCP}(t)}{\Delta t} \times GWP_H \times H \quad (2.4) \qquad E_{CO_2\text{-e}^*}(t) = \frac{\Delta F_{SLCP}(t)}{\Delta t} \times \frac{H}{AGWP_{H(CO_2)}} \quad (2.5)$$

Improvements have been made to the GWP* by Cain et al. (2019) and Smith et al. (2021). Cain et al. (2019) adapt the GWP* method to include the long-term response of the climate system to past increases in SLCP emissions, which is notable in environments where SLCP emissions have substantially increased in recent years. This is done by including a stock term $s \times E_{\text{SLCP}} \times \text{GWP}_H$ alongside the original rate term. The two terms are constrained by $r + s = 1$ to be consistent with the initial definition of GWP*, and are found to provide good results for methane at values of $r = 0.75$ and $s = 0.25$. Using this definition, the ratio $s/(rH) = s/((1 - s)H)$ defines the required decline in SLCP emissions such that net-zero CO_2-warming equivalent (CO_2-we) is achieved, preventing further warming.

Notably in their derivation, Cain et al. (2019) use typical values of the Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR) to justify their choice of $s = 0.25$ for a time horizon H of 100 years. Smith et al. (2021) argue that, by construction, the relationship between CO_2-we and RF should directly replicate that between CO_2 emissions and RF, thus not requiring the warming response at all. By approximating the linear impulse-response model used by AR5 at the mid-range (30-200 years) using a first-order equation, Smith et al. (2021) are able to express Eq. (2.8) without reference to the TCR using a new parameter g, shown below. They argue that since g is a function of s and is thus not a tuneable parameter, it is more consistent with the linear models used for climate metric calculations.

$$E_{CO_2\text{-we}}(t) = \text{GWP}_H \times g \times \left[\frac{(1 - s)H \Delta E_{\text{SLCP}}(t)}{\Delta t} + s E_{\text{SLCP}}(t) \right] \qquad (2.6)$$

where

$$g(s) = \frac{1 - \exp\left(-s/(1 - s)\right)}{s} \qquad (2.7)$$

The peak temperature change ΔT can be calculated using the summation of the cumulative SLCP emissions aggregated with GWP* ($\sum CO_2$-we) and the cumulative LLCP emissions aggregated with GWP_{100} ($\sum CO_2$-e), as (Cain et al., 2019),

$$\Delta T(t) = \text{TCRE} \times \left(\sum_{\tau=t_0}^{t_0+t} E_{CO_2\text{-we}}(\tau) + \sum_{\tau=t_0}^{t_0+t} E_{CO_2\text{-e}}(\tau) \right) \qquad (2.8)$$

The GWP* as defined by Cain et al. (2019) has been used by Lee et al. (2021) for aviation emissions, in comparison with the more conventional GWP and GTP over 20, 50 and 100 years. Lee et al. (2021) use the GWP* to show that aviation emissions are responsible for a warming rate three times higher than that from aviation CO_2 emissions only. Since the GWP* is mentioned in the Sixth Assessment Report (Intergovernmental Panel on Climate Change, 2021) and shown to more closely follow the climate response of emission mitigation scenarios, it can be expected that more studies analysing aviation emissions using GWP* will be conducted.

It is, however, important to note that unlike other climate metrics, the GWP* does not provide a single value that can easily be compared. Rather, the GWP* is a method for calculating CO_2-equivalent emissions as a function of time and thus provides a shortcut to the temperature trajectory. This can make the GWP* less suitable for analysis purposes, as is discussed further in Section 3.4.1 and shown in Figure 3.3.

2.4.4. Global Temperature-Change Potential (GTP)

The global temperature-change potential (GTP) was first proposed by Shine et al. (2005). In contrast with the integrative GWP, the GTP is an endpoint climate metric. The absolute global temperature-change potential of gas x (AGTP$_x$) gives the temperature change ΔT [$\mathrm{K\,kg}^{-1}$] at the time horizon $t = H$. Just as in the case of GWP, the GTP is the ratio of the AGTPs of the gas in question and the reference gas CO_2.

Since the AGTP is equivalent to the temperature at the time horizon H, it can be calculated directly using a climate model such as AirClim, as is done in this study. Alternatively, Shine et al. (2005) uses a temporal exponential decay emission profile of the form $A \exp(-t/\alpha)$, where A is the initial radiative forcing and α is the time constant. Implicitly, it is assumed that both A and α are independent of the atmospheric composition and constant. In reality, both depend on the evolution of the concentrations of the gas in question in the atmosphere, as well as other gases with which it reacts, for example hydroxyl molecules (OH). For example, for a pulse emission for $\tau \neq \alpha_x$, the AGTP can be calculated as (Shine et al., 2005),

$$\mathrm{AGTP}_x^P(t) = \frac{A_x}{C\left(\tau^{-1} - \alpha_x^{-1}\right)} \left[\exp\left(-\frac{t}{\alpha_x}\right) - \exp\left(-\frac{t}{\tau}\right)\right] \tag{2.9}$$

where C is the heat capacity of the climate system and $\tau = \lambda C$ is a time constant, where λ is the climate sensitivity parameter as defined in Section 2.4.1. Other calculation methods, for example for sustained emissions, are not shown here since they are not used in this study, but can be found in Shine et al. (2005).

The GTP is designed with the criticisms of Manne and Richels (2001) in mind, namely that a climate metric should vary with time such that it is dependent not only on the climate goal, but also how close the goal is. This shows the potential of GTP to be used also in economics (Shine et al., 2007). This does, however, make its use more complex, especially since the values change over time.

Given that temperature changes are further down the cause-and-effect chain shown in Figure 2.3, Shine et al. (2007) argues that the GTP is more relevant to society than the GWP. The GTP also includes more physical processes than the GWP. At least one party to the UNFCCC argues that using the GTP "would be more consistent with the UNFCCC goal of limiting future warming" (Allen et al., 2016). This is especially true for temperature-based climate goals such as the Paris Agreements (Collins et al., 2020).

On the other hand, Godal (Plattner et al., 2009) argues that the GTP is not any better than the GWP for economic usage. Only the people alive at the endpoint are taken into account since, like the GWP, the GTP has no provision for measuring the climate effects beyond the period in which the temperature goal is reached. This latter shortcoming of the GTP is mentioned as a useful topic of further research by the IPCC Expert Meeting (Plattner et al., 2009). The GTP is also still highly dependent on the time horizon (Shine, 2009; Grewe and Dahlmann, 2012), as is demonstrated in Section 4.1.

The GTP has been used for aircraft design optimisation, for example Egelhofer et al. (2007), and optimising aircraft cruise altitudes, for example Schumann et al. (2011). The GTP is also included in the aviation emission reviews, such as Lee et al. (2021), in comparison with the GWP and GWP*. The GTP does not, however, have as widespread use as the GWP in aviation.

2.4.5. Average Temperature Response (ATR)

The average temperature response (ATR) was initially developed by Dallara (2011). It is a climate metric specifically tailored for aircraft design rather than for policy decision-making. Temperature change is chosen because it is commonly used but also because it can commonly be understood by non-experts, in comparison with radiative forcing and its temporal integration (Dallara, 2011; Dallara et al., 2011).

The formulation of the climate metric takes the form of other common climate metrics and is given by Eq. (2.10) (Dallara et al., 2011), where $\Delta T_{sust,H}$ is the time-varying global-mean temperature change, H is the aircraft lifetime, and $w(t)$ is the weighting function. A large difference between the ATR and for example the GWP and GTP is that the ATR considers average, sustained emissions over the aircraft lifetime H, rather than a pulse. Dallara et al. (2011) generally use an operational lifetime of 30 years. It is important to note that subsequent studies using the ATR generally use H as the time horizon in line with other conventional climate metrics, as is also done in this study. Common values of H are 20, 50 and 100 years.

$$\text{ATR}_H = \frac{1}{H} \int_0^\infty \Delta T_{sust,H}(t) w(t) \, \mathrm{d}t \tag{2.10}$$

The weighting function $w(t)$ can be designed as necessary to balance short-term and long-term mitigation costs and economic risks. An aircraft designed for a lower climate impact in 100 years is riskier than one designed to limit its impact within the next 30 years, thus requiring a form of discounting (Dallara et al., 2011). The discounting rate and the form of the weighting function is a value judgement based on the trade-off of short- and long-term impacts, and thus must be properly justified, as with other climate metrics discussed. This study does not use a weighting function.

Since the ATR was designed specifically for aircraft design, it has not found widespread use outside of the industry or even research. It has, however, been used in a number of aircraft design optimisation schemes and has been used extensively by the DLR (Scheelhaase et al., 2015; Dahlmann et al., 2016b; Grewe et al., 2017) and other European institutions (Koch, 2013; Proesmans and Vos, 2021). Koch (2013) argue that it is more appropriate for aviation than the GWP and GTP because it includes the thermal inertia of the climate system, considers the change in temperature rather than radiative forcing, and is an integration over a time horizon rather than a snapshot. This argument will be tested in this study.

2.4.6. Other Climate metrics

Although not used in this work for the reasons explained below, a brief overview of other existing climate metrics and their corresponding literature are provided here for general reference. The climate metrics described in this section are the CGWP, CGTP and the ARTP.

First, building on the equivalence between a step SLCP emission rate reduction and a CO_2 pulse emission used in the development of the GWP* (Allen et al., 2016, 2018), Collins et al. (2020) suggest the use of 'combined' climate metrics, specifically the combined GWP (CGWP) and combined GTP (CGTP). Both relate the step response S of a climate agent to the pulse response P of CO_2: $\text{CGWP} = \text{AGFP}^S/\text{AGFP}^P_{CO_2}$ and $\text{CGTP} = \text{AGTP}^S/\text{AGTP}^P_{CO_2}$, where AGFP($H$) is the absolute global forcing potential - an endpoint climate metric describing the radiative forcing at a time horizon H. Collins et al. (2020) further show that the GWP* is an

approximation of, but tends to underestimate, the CGTP. Neither the CGWP and CGTP are considered in this work due to the difficulty of defining the equivalent of a kilogram emission for contrails, which is required to obtain the relative climate metric, and because the climate metrics are not unitless, making their implementation more complex.

The Absolute Regional Temperature Change Potential (ARTP) is an emission climate metric developed by Shindell and Faluvegi (2009), Shindell and Faluvegi (2010) and Shindell (2012) to determine the regional time-varying surface temperature response from heterogeneous emission and radiative forcing patterns. Collins et al. (2013) and Lund et al. (2017) show that the temperature pattern can differ significantly from both the global average as well as from the radiative forcing pattern, especially for species with shorter atmospheric lifetimes. Lund et al. (2012) find that the highest loss of information by global averaging aviation emissions occurs for NO_x and aerosols. The ARTP is more complex than other, globally-averaged climate metrics and would require a large re-structuring of AirClim, which is beyond the scope of this work. Therefore, although the ARTP has potential especially for aviation non-CO_2 emissions, it is not further analysed.

3

Methodology

This chapter describes the methods used to analyse the climate metrics described in the previous chapter. The aim is to develop a framework with which the best-suited climate metric can be chosen, depending on the climate objective.

The chapter begins in Section 3.1 with an overview of the analysis procedure, research scope and limitations. The aviation climate objectives are described in Section 3.2, followed by the development of requirements in Section 3.3. This section also provides a detailed description of the selection process and which elements of the methodology correspond to the analysis of which requirement. The general response of climate metrics to aviation emissions is described in Section 3.4, in which also the GWP* and EGWP* calculation methods are explained. Section 3.5 provides an overview of the expected aircraft emission pathways, with which future fleets can be developed and analysed in Section 3.6. The sensitivity analysis is shown in Section 3.7 and, finally, the ability of climate metrics to calculate CO_2-eq emissions and estimate the temperature response is analysed in Section 3.8.

3.1. Method Overview, Research Scope and Limitations

The research objective of this work is formulated as:

> *Recommend the best-suited climate metrics for existing and proposed aviation climate objectives by systematically analysing the response of existing physical climate metrics using the climate-chemistry response model AirClim.*

The main outcome of this research is thus a framework with which the best-suited climate metric can be chosen. In Chapter 5, this framework is used to recommend a climate metric for each climate objective as defined in the following section. The intention is that the results of this research, including the best-suited climate metric, can directly be used by a stakeholder. However, sufficient information is also provided such that a stakeholder can repeat the process and implement other quantitative or qualitative requirements in their decision-making process.

The methodology presented here is informed by the aviation climate objectives and climate metric requirements generated in Sections 3.2 and 3.3. The climate metric requirements are developed specifically for aviation and thus the climate metrics used in this analysis are subject to prerequisites as described in that section. The methodology is split into four main parts: a general response analysis (Section 3.4), multivariate fleet analysis (Section 3.6), sensitivity analysis (Section 3.7) and CO_2-eq emission analysis (Section 3.8). Each aim and method of each section is described briefly here to provide an overview. It should be noted that in this research, the exact climate metric or temperature value is of lesser importance than the

comparison of the value for different climate metrics. For this reason, the analysis will generally not describe any climate metric or temperature values in detail.

The general response analysis (Section 3.4) aims to identify any systematic problems with climate metrics that would prevent their use in aviation. This is done by performing pulse, sustained and increasing emission simulations with AirClim and, using the radiative forcing and temperature profiles, by calculating the response of each climate metric over time horizons between 1 and 100 years. Using these responses, the dependency of each climate metric to the background emissions scenario and time horizon can be calculated.

The multivariate fleet analysis (Section 3.6) aims to identify the inherent biases within the climate metric calculations for different aviation-specific emission species or changes in aircraft/trajectory design. For this purpose, a set of expected aircraft emission pathways are required (Section 3.5), which provide estimations of future growth rates and changes in various parameters, for example cruise altitude, efficiency and NO_x emissions. Using these pathways, a set of 10000 fleets are generated using a Monte Carlo simulation. The analysis is done by comparing the change in climate metric value and the change in temperature for all pairs of fleets. A climate metric without inherent bias would show the same sign of change for both and thus have a linear trend. This is illustrated by Figure 3.5. The number of incorrect fleet pairings is then the result of the analysis.

The aim of the sensitivity analysis (Section 3.7) is to calculate the sensitivity of climate metrics to changes in the fleet generation variables in the multivariate fleet analysis. This helps to explain the trends seen in the other analysis steps, and identify any unexpected trends.

Finally, the CO_2-eq emission analysis (Section 3.8) evaluates the ability of climate metrics to estimate CO_2-eq emissions. This is useful for comparison between aviation and other industries, and within the industry for market-based schemes such as CORSIA. Ideally, CO_2-eq emissions would behave similarly to CO_2 emissions, which, due to their long atmospheric lifetime, are approximately linear with temperature. The analysis performs AirClim simulations for full aviation scenarios and for the fleets generated in the multivariate fleet analysis.

A number of limitations of the study are important to note, which are also discussed in more detail in Section 5.2. As the literature study attests, a large number of climate metrics exist, each serving a slightly different purpose and each answering a different climate question (cf. Grewe and Dahlmann, 2015). In this study, only physical climate metrics are considered due to the high uncertainties of damage- and impact-based climate metrics and their low acceptance among stakeholders. The latter reason is important since the intention of this work is to inform policy decisions. The climate metrics analysed are the Radiative Forcing (RF), Effective Radiative Forcing (ERF), the Global Warming Potential (GWP), the GWP* and a derived climate metric developed in this study denoted EGWP* (Effective GWP*), the Global Temperature-Change Potential (GTP) and the Average Temperature Response (ATR). The CGWP and CGTP are not analysed for the reasons given in Section 2.4.6.

Since the objective is to develop a framework that is widely applicable, the aircraft designs and fleets used are representative and do not correspond to any existing or proposed aircraft design. Because an aircraft design tool is beyond the scope of this research, assumptions on the impact of aerodynamic and propulsive efficiency, as well as novel fuels such as synfuels/biofuels, hydrogen and SAF are made. This means that, for example, the efficiencies of different propulsion systems are not taken into account and the generated fleets may not be physically feasible. The expected impacts of these assumptions are analysed in Section 5.2.

3.2. Development of Aviation Climate Objectives

Three main climate objectives are identified, against which the climate metrics are analysed. These relate to aircraft design, aircraft trajectories and finally to market-based schemes. Although more climate objectives could feasibly exist, it is assumed that the three objectives shown are broad enough to encompass these and that the results are transferable.

    1. *Optimise aircraft design for minimum climate impact.*

This objective aims to design aircraft with the climate impact in mind. Using a climate metric, a stakeholder can also easily compare proposed aircraft designs. In an aircraft design company environment, this allows the climate impact to be part of the trade-off process. Regulators would also be able to set limits or penalties depending on the climate metric value to guide the market towards more sustainable aircraft designs. Both the near-future climate impact reduction, which would aim to avoid tipping points due to overshooting, and a long-term reduction, which would aim for climate stabilisation in the spirit of the Paris Agreement, can be considered.

    2. *Optimise aircraft trajectories for minimum climate impact.*

This objective considers the potential climate impact reduction achieved by optimising trajectories or re-routing aircraft to avoid climate sensitive regions. Finding the best suited climate metric for this purpose will predominantly help policymakers at organisations such as EURO-CONTROL to make decisions easily, and will ensure that the options available are comparable.

    3. *Reduce aircraft emissions in a cost-effective manner using a market-based scheme.*

Market-based schemes such as CORSIA and the EU ETS aim to incentivise airlines to buy and operate aircraft with a lower climate impact, thus also incentivising manufacturers to develop these aircraft. This is done using a tax or requiring airlines to buy offsetting credits on released CO_2. However, as for example Frömming et al. (2012) and Niklaß et al. (2021) describe, only considering the impact of CO_2, as is currently done with both CORSIA and the EU ETS, can have the opposite effect than desired. For example, aircraft could be designed to fly higher and thus burn less fuel, but by doing so cause a higher total climate impact (Matthes et al., 2021). The aim of this objective is to determine the best suited climate metric to use to incorporate non-CO_2 impacts into CORSIA and the EU ETS.

3.3. Development of Requirements and Selection Procedure

To determine which climate metric is considered *best-suited,* as specified by the research objective, a clear set of requirements is needed. The requirements presented in this section are simultaneously also the criteria for the selection framework. The requirements originate from the climate objectives described in the previous section, from literature (Forster et al., 2006; Wuebbles et al., 2010; Dahlmann, 2011) and from discussions with experts and stakeholders. A few prerequisites are made for the climate metrics to be analysed, namely that they,

    • Shall include information on the impact an emission has on the environment;

- Shall be updateable following advancements in climate understanding; and
- Shall be flexible in their use across different industries

The first prerequisite prevents, for example, the use of the emission mass since this climate metric does not provide any meaningful information on its own about the environmental impact of all aviation emissions and effects. Making use of such a climate metric would be inappropriate for the climate objectives established. The second prerequisite is not used as a requirement in this work because all climate metrics considered are updateable following advancements in climate understanding. A similar statement can be made about the third prerequisite. Using the same climate metrics across industries is beneficial for stakeholders seeking to compare strategies and progress, however using climate metrics for industries other than aviation is outside the scope of this work. Based on the literature research performed, it can be assumed that all climate metrics considered in this work also meet the third prerequisite.

The requirements used in the framework are presented in the following. Although this list is not exhaustive, these requirements have been identified as the most important for the climate objectives. The framework aims to cover all scenarios in which a climate metric must be chosen, however there are certain limitations to the study as discussed in Chapter 5. Therefore, if deemed necessary by a stakeholder, the framework can also be further extended with other requirements. A stakeholder can also choose the weight or importance of each requirement depending on the use case. Each requirement is shown in italics, followed by a justification of the requirement and references to the methods used in the analysis. An identifier is also provided for ease of reference.

REQ 1 *Shall correctly represent the temperature change.*

The temperature change is selected as the impact against which the climate climate metrics are analysed. Temperature change is further down the cause-and-effect chain shown in Figure 2.3 than, for example, radiative forcing, and is in the category of "Climate Change". Selecting from the "Impacts" or "Damages" categories would require insights into local ecosystems, social effects and the mitigation/adaptation costs of impacts - subjects that are beyond the scope of this work. The uncertainties involved in the calculation of impacts and damages are also higher and harder to quantify. Judging potential climate metrics on their ability to estimate impacts or damages is, therefore, not appropriate for aviation policy and this study.

It can be argued that the effective radiative forcing (ERF) would be equivalent to the temperature change, since $\Delta T = \lambda \cdot \text{RF}$ where λ is the efficacy. However, for aviation emissions, only a single comprehensive study, Ponater et al. (2006), analyses the efficacy parameters. Furthermore, temperature change is easier to understand than ERF, which links to the following requirement. For these reasons, it is decided to focus on temperature change in this research. The ERF is, however, used for the EGWP* climate metric, as described later in this chapter.

This requirement will influence the choice of best-suited climate metric by ensuring all analyses judge temperature change, rather than radiative forcing. Furthermore, how well climate metrics can be used to estimate the temperature response will be judged quantitatively by the linearity of the relationship between the calculated temperature and the estimated CO_2-eq emissions, as described further in Section 3.8.

REQ 2 *Shall be easy to understand and implement.*

Since the intention is that climate metrics will play an integral part in climate policy, it is important that stakeholders are able to understand how a climate metric is calculated and what its results show. Only in this way will the climate metric be implemented correctly and appropriately, resulting in aircraft that are generally better for the environment. This requirement also covers the computational cost of a climate metric to ensure that it can be used in design optimisation schemes.

This requirement is analysed qualitatively. The ease of understanding considers the units, the calculation method and the outcome; the implementation considers whether pre-calculated data is required and how laborious it is to calculate, as well as the calculation method.

REQ 3 *Shall be largely independent of the time horizon.*

In this research, it is assumed that a low dependence on the time horizon is desired. As discussed in detail in Section 2.4.2, while it can be useful for stakeholders to have direct control over the time horizon to analyse a certain effect, this can lead to exploitable ambiguity, especially since there is often no obvious choice for the time horizon (Shine et al., 2007; Allen et al., 2016, for example). If a climate metric is perceived to be exploitable or ambiguous, even if not the case, less acceptance amongst stakeholders could be the result. The dependence of the time horizon is explored in the general response of climate metrics (Section 3.4). The dependence is shown graphically and judged quantitatively by calculating the percentage difference for the most common 20, 50 and 100 year time horizons.

REQ 4 *Shall be largely independent of background emissions.*

Low dependence on background emissions is also assumed to be desired. To understand why, consider a future where the world reduces emissions quickly, such as in SSP1-1.9 (IPCC "Sustainability" emission scenario, cf. Section 3.4). In this scenario, an emission from aviation is valued more, which is reflected also in the climate metric value being higher. Assuming aircraft operators would have to pay according to the climate metric value, then the aviation industry would in this case be incentivised to reduce its emissions to pay less. However, in a scenario where global emissions continue to rise quickly, such as in SSP5-8.5 (IPCC "Fossil-Fueled Development" emission scenario), the opposite effect occurs: the value of an emission from aviation is lower, hence there is a lower incentive to reduce emissions.

A low dependence on background emissions means that the climate metric value does not depend to a large extent on emissions of other industries. A high dependence, on the other hand, means that updates to the expected global emissions scenario can change the climate metric value substantially. Whilst no dependence would be questionable from a scientific standpoint, in this research it is argued that low dependence should be desired such that the climate metric values remain fairly consistent and policymakers, aircraft operators and other stakeholders can accurately estimate required payments from year to year.

The dependence of background emissions is explored in the general response of climate metrics (Section 3.4). As with the previous requirement, the dependence is analysed graphically and is numerically calculated. The effects described in the previous paragraphs are also demonstrated in the results (cf. Section 4.1).

REQ 5 *Shall have low inherent biases towards changes in aircraft design or trajectory.*

Depending on how a climate metric is calculated - including, among other things, whether a climate metric considers the instantaneous change at a certain time or a temporal integration over a time horizon - the impact of certain emission species can be calculated to be higher or lower than they are in reality. In other words, trade-offs between different species can be inherently built into climate metrics (cf. Ocko et al., 2017). The aim of this requirement is to ensure that the impacts of different technologies are valued appropriately, such that an aircraft design with a lower climate metric value also has a lower environmental impact.

Section 3.6 provides a more detailed explanation of the necessity of this requirement and describes how this requirement is analysed. A Monte Carlo simulation is performed for many different future fleets, using different fuels, trajectories and technologies. An error analysis calculates the number of incorrect pairings, i.e. when an aircraft with a lower climate metric value actually causes a larger temperature change, or vice versa. The number of incorrect pairings is used to establish which climate metric scores best for this requirement.

REQ 6 *Shall be appropriate for different emission profiles.*

Within the aviation industry, a climate metric should be useful for many different use cases. For example, different emission profiles are used for the climate objectives as described in the previous section: Minimum climate impact due to a single flight can be represented by a pulse emission, whereas a climate impact reduction strategy can be represented by a sustained or fleet emission. A further analysis element is that a series of pulse emissions should at least qualitatively show the same result as a sustained emission. Depending on the calculation method, however, this is not always the case and can lead to misleading results. This requirement is analysed qualitatively in the general response of climate metrics (Section 3.4).

3.4. General Response of Climate Metrics

This section describes the methodology used to determine the general response of climate metrics to aviation emissions using AirClim. In Subsections 3.4.1 and 3.4.2 it also describes the calculation method for the GWP* and EGWP* climate metrics used in this research in more detail.

The main aim of the general response analysis is to identify any potential systematic issues with the existing climate metrics described in Section 2.4 which would make them potentially inappropriate for use in aircraft climate assessments. It also aims to analyse the general impact of the background emission scenario on the results obtained using climate metrics. An ideal climate metric would show the same behaviour for a pulse as well as a sustained emission (REQ 6), and not be largely dependent on the time horizon (REQ 3) or background emission scenario (REQ 4).

The background emission scenarios used in this study are the Shared Socioeconomic Pathways (SSPs), as used in the sixth Assessment Report (AR6, Intergovernmental Panel on Climate Change, 2021). The main socioeconomic drivers of the SSPs are population, education, urbanisation and economic development (Riahi et al., 2017). Figure 3.1 shows the trajectory of global CO_2 and CH_4 concentrations from Meinshausen et al. (2020), which are inputs into AirClim. One temperature profile from each SSP is taken to provide a broad spectrum of background emission scenarios.
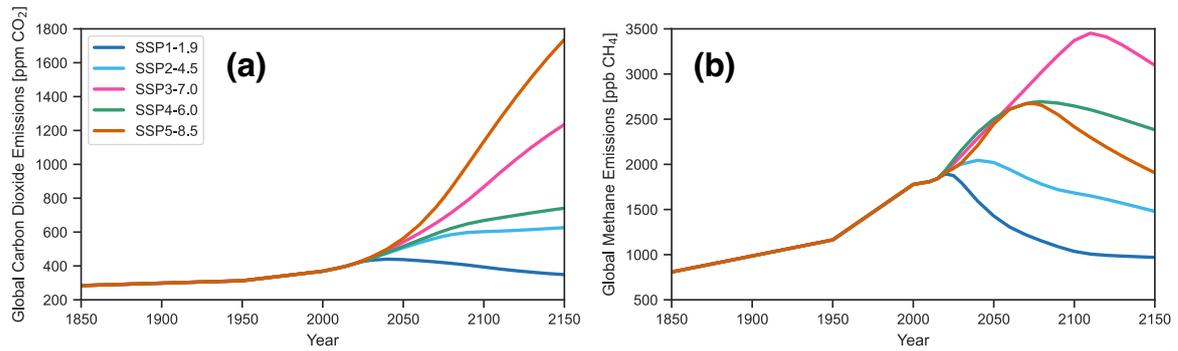
Figure 3.1: Trajectories of **(a)** global carbon dioxide and **(b)** global methane concentrations from the Shared Socioeconomic Pathways (SSPs, Meinshausen et al., 2020). "SSP1-1.9", for example, represents a temperature increase of 1.9°C under SSP1 (cf. Meinshausen et al. (2020) and supplementary data[1])

To obtain the general response of each climate metric, three emission scenarios are used: a pulse emission (P2020), a sustained emission (C2020) and a 1% increasing emission (INC1). These emission scenarios correspond with the global fleet of aircraft with routings as described by WeCare 2050 (Grewe et al., 2017) being used for one year (pulse in the year 2020), each year following the year 2020 (sustained), or increasing by 1% per year. The emission scenarios are shown graphically in Figure 3.2. The yearly fuel use of 716.2 Tg is also taken from WeCare 2050. It should be noted that the exact fuel use value is not important for the analysis since only the general response is sought.
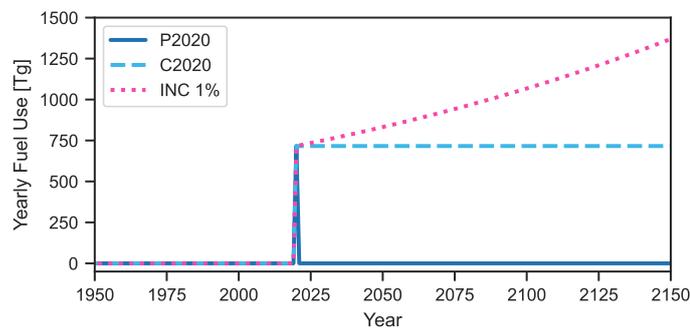
Figure 3.2: Yearly fuel use for the pulse, sustained and 1% increasing emission profiles used to analyse the general response of each climate metric

The climate metric responses, as well as the trajectories of radiative forcing and temperature, are calculated using AirClim for time horizons between one and 100 years. This is done for each background emission scenario as well as for both the pulse, sustained and increasing emission scenarios. Since the RF, GWP, GTP and ATR are already integrated into AirClim, the outputs can be used directly. The GWP*/EGWP* methodology as described in Section 2.4.3 must, however, first be integrated, described in the following sections.

Similarly to Dahlmann (2011), the response of climate metrics to a NO_x emission will be analysed separately. The NO_x emission is used to ensure that the climate metrics show similar

---

[1]Supplementary data for Meinshausen et al. (2020) available at: `https://greenhousegases.science.unimelb.edu.au/`

behaviour for a complex emission in both a pulse and sustained emission scenario, which as mentioned in the previous section is an important requirement (REQ 6). This analysis also serves as verification of the climate metric calculations.

3.4.1. GWP* Calculation Method

The definition of the GWP* as shown in Eq. (2.6) is a function of the yearly SLCP emission rate E_{SLCP} and its time-derivative ΔE_{SLCP}. These parameters are meaningless for NO_x-induced aviation effects (O_3, long-term CH_4 reduction and the PMO effect, see Section 2.1.2), and especially also for contrails. For aviation emissions, therefore, the GWP* methodology must be converted to analyse radiative forcing. This equivalent calculation is proposed in the initial development of the GWP* method by Allen et al. (2018) (cf. also Eq. (2.5)), and is modified using the improvements suggested by Cain et al. (2019) and Smith et al. (2021) to obtain the following equation,

$$E_{CO_2\text{-we}}(t) = g(s) \times \left[(1 - s) \times \frac{\Delta F_{SLCP}(t)}{\Delta t} \times \frac{H}{AGWP_{H(CO_2)}} + s \times \frac{\overline{F}_{SLCP}(t)}{AGWP_{H(CO_2)}} \right] \qquad (3.1)$$

where $\overline{F}$ is the average radiative forcing between the time $t - \Delta t$ and t, and all other parameters are already described in Section 2.4.3. It is interesting to note that $AGWP_{H(CO_2)}$ is actually a function of time and background emission scenario, as described in Section 4.1. To obtain the most accurate results, the $AGWP_{H(CO_2)}$ is thus pre-calculated. The above equation differs to the one used by Lee et al. (2021) only by the multiplication by g (see Section 2.4.3), which was introduced in the same year by Smith et al. (2021) to improve consistency with the linear models used for climate metric calculations.

The GWP* differs from other climate metrics considered in this study in that it is a flow-based climate metric. The GWP* does not provide a single climate metric value over a specific time horizon. Instead, it provides a CO_2-equivalent value as a function of time. As an example, Figure 3.3 shows the GWP* response of a fleet emission scenario. The GWP* value for each species in sub-figure (b) is clearly not constant but changes over time.

To estimate the impact of a fleet, a certain point along the temporal trajectory must nevertheless 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 inter-comparable. Using the peak CO_2-eq value can be interpreted as being equivalent to using the peak temperature directly, for example, as is done in this work. Choosing this point, however, is essentially already a trade-off between different emission species.

Determining a GWP* value corresponding to an average temperature is more difficult. As the example shows, once the emission has ended, the GWP* shows a negative rate of CO_2-eq. As described in Section 2.4.3, this negative rate is not physical but is used to ensure that the GWP* more accurately follows the temperature response. It is also important to note that the time at which the emission rate changes sign differs per emission species. Taking an average over a certain time horizon, e.g. 100 years, does not, therefore, have a physical meaning.
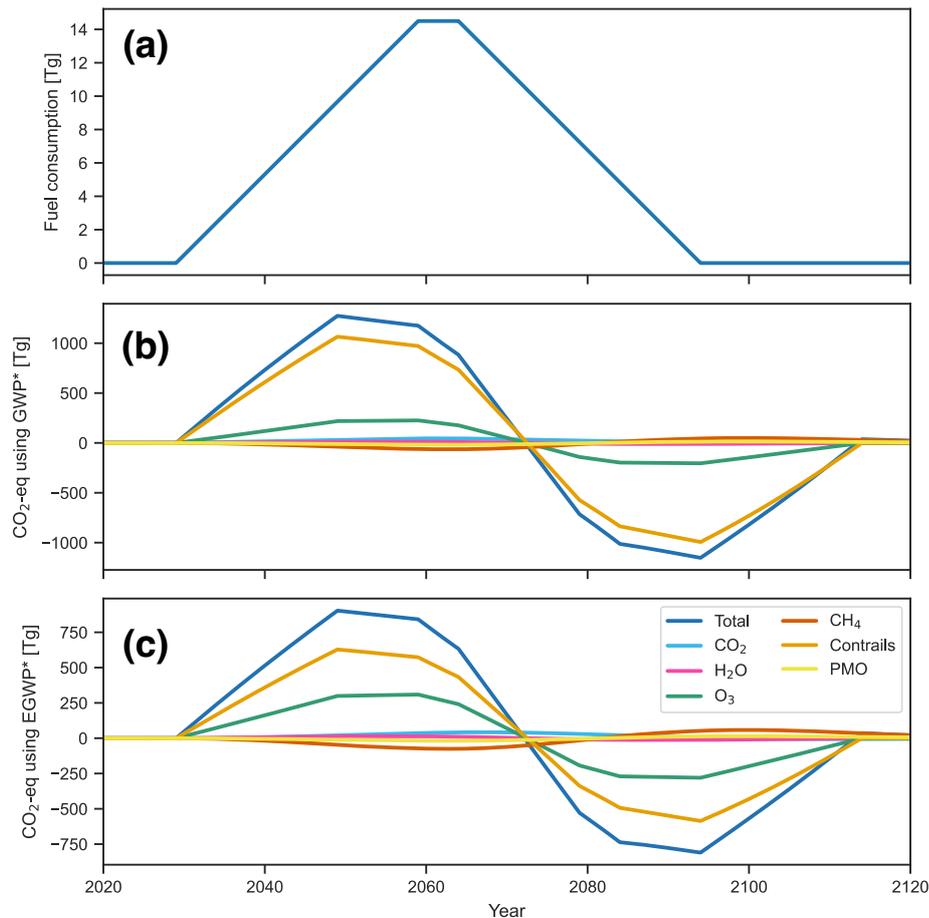
Figure 3.3: Example of GWP*/EGWP* calculation method for an example fleet emission profile demonstrating the flow-based nature of the climate metric. **(a)** shows the fleet emission profile, **(b)** the GWP* and **(c)** the EGWP* responses.

In this research, the maximum total GWP* is used for all analyses, both compared to the peak and average temperature change. The same approach is also used for the EGWP*, described in the following section. The impact of this assumption is addressed in Section 5.2.

3.4.2. EGWP* Calculation Method

The EGWP* is a climate metric developed as part of this research as a derivative of the GWP*. It is based on the recognition that the GWP* does not consider the efficacy (see Section 2.4.1) and thus overestimates the temperature induced by contrails, and underestimates the temperature induced by ozone reduction due to a NO_x emission. This effect can be seen in the general climate metric results (cf. Section 4.1). To account for this shortcoming when using temperature as the selected impact, the GWP* method is applied using Effective Radiative Forcing (ERF, cf. Section 2.4.1). The ERF for a species i is calculated simply by $ERF_i = RF_i \times r_i$ where r is the efficacy taken from Table 3.1 (Ponater et al., 2006). The difference between the GWP* and EGWP* values are shown in Figure 3.3 for an example fleet.

The use of efficacy parameters allows the EGWP* to more accurately represent the tem-

Table 3.1: Climate sensitivity parameters and efficacies (λ in $\mathrm{K/(Wm^{-2})}$ and r (unitless) respectively) used in AirClim and for the calculation of the EGWP*, as determined from equilibrium climate change simulations using ECHAM4 by Ponater et al. (2006)

	CO_2	O_3	CH_4	H_2O	Contrails
λ	0.73	1.00	0.86	0.83	0.43
r	1.00	1.37	1.18	1.14	0.59

perature, as is shown in the results section of this work, but leads to higher uncertainty. These parameters can be model- and scenario-dependent and only one reference, Ponater et al. (2006), could be found for aviation species and secondary effects. The implications of this are discussed in further detail in Section 5.2.

3.5. Overview of Expected Aircraft Emission Pathways & Technological Advances

This research will predominantly make use of the Fa1 fuel scenario developed by the IPCC (Intergovernmental Panel on Climate Change, 1999) and a new set of scenarios developed by Grewe et al. (2021), specifically the scenarios current technology (CurTec), CORSIA, Flight-path 2050 (FP2050, European Commission, 2011) and a COVID scenario with a 15 year recovery and sustained impact (COVID-15s). These scenarios are chosen such that this research builds on existing work. Figure 3.4 shows the trajectories graphically.

Since climate metrics with time horizons of up to 100 years will be analysed, the scenarios used in this research must extend quite far into the future. For this research, they have been extended to the year 2200, assuming an 0.8% annual growth rate past the year 2100 for CurTec and Fa1, and a 0.5% annual growth rate for CORSIA, FP2050 and COVID-15s. This assumption is based on Grewe et al. (2021), who estimate a 0.8% annual growth rate of Revenue Passenger Kilometres (RPK) in the year 2100, and the approximate growth rate of 0.5% measured in the year 2100 from the other scenarios.

Unfortunately, as the impact of the COVID-19 pandemic on the aviation industry has shown, estimating the future is difficult and is dependent on a large number of uncertainties. In recent years, disregarding the effects of the COVID-19 pandemic, the aviation industry has seen an annual growth of 6%, but this is expected to reduce (Grewe et al., 2021; Clean Sky 2 Joint Undertaking, 2021). Fuel use is more complex to model since it also depends on the development of new technologies, efficiency increases and the introduction of new fuels. It is likely that a constant 0.5 or 0.8% annual growth rate is an overestimation. However, a more detailed analysis is beyond the scope of this project and would likely not result in accurate data given the uncertainties. For this reason, the dependence on the fuel scenario will be considered to ensure that the results obtained from this research are nevertheless representative.

The development pathways of new aviation technologies is also important for this research since the climate impact of expected future fleets will be analysed. A range of possible improvements on existing aircraft is considered rather than specific designs. The CleanSky 2 Joint Undertaking (JU) Technology Evaluator (TE, Clean Sky 2 Joint Undertaking, 2021) expects that *mainliners*, consisting of short-/medium range and long-range passenger aircraft, released in 2030 will show up to a 20% reduction in CO_2 and NO_x emissions, improving further
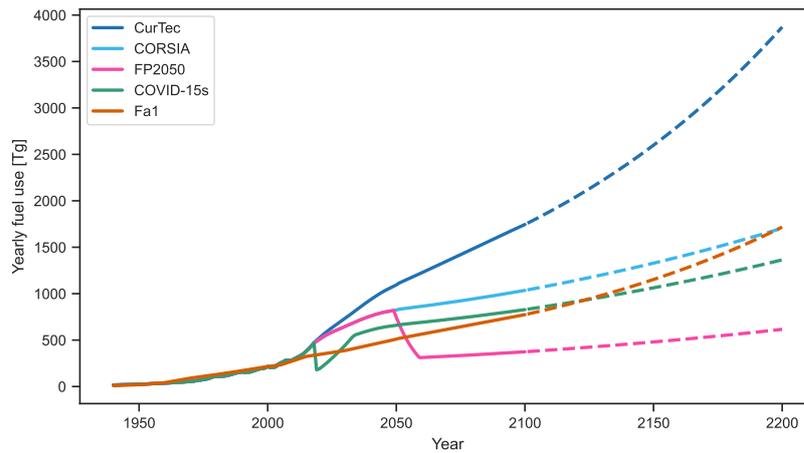
Figure 3.4: Assumed fuel scenarios until the year 2200. Solid lines indicate scenarios estimated by Grewe et al. (2021, CurTec, CORSIA, FP2050, COVID-15s) and Intergovernmental Panel on Climate Change (1999, Fa1); dashed lines indicate annual growth post-2100 of 0.5% (CORSIA, FP2050, COVID-15s) or 0.8% (CurTec, Fa1).

to 30% beyond 2035. Some short-/medium-range concepts are expected to have up to a 40% NO_x reduction. These reductions are due primarily to improvements in engine and airframe efficiency.

Research conducted by Grewe et al. (2021) is slightly less optimistic, estimating fuel burn reductions in 2035 of 22% and 18%, and NO_x reductions of up to 26% and 22% for single- and twin-aisle aircraft respectively. By 2050, single-aisle aircraft are estimated to have reduced their NO_x emissions and fuel consumption by 38% and twin-aisle aircraft by 34-44% and 25-44% respectively, depending on the configuration.

The use of new fuels such as sustainable aviation fuel (SAF) and hydrogen is also important, and part of the reason for the reductions. McKinsey & Company (2020) assumes that the next window of opportunity for short-range aircraft will be around 2030-2035, in line with the previous studies mentioned. This is corroborated by Airbus, which expects its ZEROe project to deliver hydrogen-powered airliners in time to start regular service in 2035[2].

Table 3.2: Change of in-flight emissions and emission-related effects for Sustainable Aviation Fuel (SAF) and hydrogen from McKinsey & Company (2020), used in this research. Note that the McKinsey & Company (2020) report does not provide a detailed methodology and also does not mention which climate metric was used to calculate these values. These values are thus only used to approximate future fleets, and the total values are not used.

Fuel	CO_2	NO_x	H_2O	Contrails	Total
SAF	-65-80%[3]	-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%

---

[2]Airbus Press Release ZEROe 21/09/2020 `https://www.airbus.com/en/newsroom/press-releases/2020-09-airbus-reveals-new-zero-emission-concept-aircraft` [Accessed 03/03/2022]

[3]It should be noted that 100% reduction in CO_2 emissions for SAF as specified in the McKinsey report assumes that all CO_2 is captured from the air. The EU Renewable Energy Directive (RED-II) stipulates at least 65% net-CO_2 reduction for biofuels and at least 70% for synthetic fuels after the year 2021. Using different fuel mixes

The change of in-flight emissions and emission-related effects for SAF and hydrogen is assumed to be as reported by McKinsey & Company (2020, cf. Exhibit 4), as shown in Table 3.2. This report is chosen as it provides the most comprehensive data for different fuel types. However, it should be noted that the report does not provide a methodology or list of assumptions for these values, thus preventing them from being effectively compared against other studies. More research is also needed from other authors for corroboration. Nevertheless, since the values are only used to approximate future fleets and the results of this research do not depend on the exact values, the values shown in Table 3.2 are assumed to be sufficiently accurate.

3.6. Multivariate Fleet Analysis

This section describes the development and analysis of representative, future aircraft designs and fleets. Section 3.6.1 describes the representative fleets based on the technology trajectories in the previous section. Section 3.6.2 describes the multivariate fleet analysis itself.

3.6.1. Development of Assumed Future Designs and Fleets

The development of representative fleets is loosely based on the work performed by Proesmans and Vos (2021). A constant production rate is assumed, expected to last 30 years. Each aircraft is further assumed to have a lifetime of 35 years assuming no hull losses. An example of the scenario can be seen in Figure 3.3. Production is assumed to begin in the year 2030 for the main reference fleet, approximately on par with the expected introduction of the next generation of single- and twin-aisle aircraft and new fuels such as hydrogen according to the analyses of Grewe et al. (2021) and McKinsey & Company (2020). The exact year of introduction of a new fleet is not relevant to the outcomes of this study and will thus be part of the Monte Carlo uncertainty analysis described later in this section.

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 as established by the DLR as part of the *WeCare* project (Grewe et al., 2017). This category of aircraft is classified by a seat number between 152 and 201. A twin-aisle reference aircraft is chosen to be 40% of Category 7, approximating the expected Airbus A350 market share (Cooper et al., 2018). Category 7 includes aircraft with seat numbers greater than 302. For an analysis of the trajectories and temperature responses of each category, the reader is directed to van der Maten (2021).

The fleets used in the multivariate fleet analysis are developed using Monte Carlo simulations with the parameters shown in Table 3.3. The ranges used are based on the expected technological pathways, as described in the previous section, and include various improvements in aerodynamic and propulsive efficiency as well as changes in cruising altitude (cf. Matthes et al., 2021). A uniform probability is assumed over the range of each parameter. A total of 10 000 simulations are performed. The main reference fleet, Fleet 0, uses Jet-A1, is introduced in the year 2030 in the SSP2-4.5 scenario, and has a value of 100% for all other parameters in the table. Similarly to the 6th Assessment Report (AR6, Intergovernmental Panel on Climate Change, 2021), this research makes no assumptions about the most likely background emission scenario, however, for consistency, the SSP2-4.5 "Middle of the

---

and production methods, Grewe et al. (2021) assumes that an 80% reduction can be achieved by 2100. A 100% reduction is, therefore, likely optimistic in the short-term. Therefore, a 65-80% reduction is used.

Road" scenario is used as a reference in all parts of the methodology. Based on the results of the sensitivity analysis, described in Section 3.7, it is clear which combination of parameters leads to the maximum and minimum temperature for the same year of introduction and background emissions scenario, which are also shown in Table 3.3. These fleet combinations will be included separately to the Monte Carlo simulations to ensure that the full breadth of combinations are considered.

Table 3.3: Ranges of fleet design parameters for Monte Carlo simulations, including the reference fleet (Fleet 0) and the fleets which produce the maximum and minimum temperature responses. See the accompanying text and footnote for a more detailed description of the contrail distance modifier, and Table 3.2 for the impact of different fuels.

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][4]	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

For fuels other than Jet-A1, the emissions parameters are modified according to Table 3.2. The only exception is the CO_2 emission reduction from SAF, which is assumed to be between 65 and 80% as described by the accompanying footnote. A uniform distribution is again used within the ranges shown in the table.

The contrail modifier mentioned in Table 3.2 is a direct reduction of the cruise distance for which contrails form, which is an input into AirClim. This reduction can be seen as due to the modification of the exhaust composition due to the different fuel. This is different to the contrail distance modifier shown in Table 3.3, which can be seen as due to aircraft flying further distances to avoid climate sensitive regions. As a result, a reduction in contrail distance leads to an increase in fuel burn. From Yin et al. (2018), this is estimated to be of the ratio -15%:1% contrail distance to fuel burn up to a contrail distance reduction of 60%, which is approximately the end of the quasi-linear region of the Pareto fronts calculated. At the maximum contrail distance reduction, therefore, a 4% higher fuel burn is expected.

A large number of assumptions have been made in the development of these fleets and a number of the parameters are linked. For example, hydrogen fuel cell propulsion systems are expected to be around twice as efficient as existing turbofan engines. Furthermore, aircraft flying with hydrogen fuel cells are powered by propellers and will thus fly slower and at lower altitude. These aspects have not been taken into account in the Monte Carlo simulations. However, since it is the difference between concepts that is important in this analysis, rather than the exact value, the results are still valid. The impact of the assumptions on the results is discussed further in Chapter 5.

---

[4]The contrail distance modifier is coupled with the fuel burn in the ratio of -15%:1% on top of the range included in the table.

3.6.2. Fleet Analysis Method

As discussed in the literature review, the choice of climate metric can have a large influence on what technology or design is deemed to be *better* or *optimal*. Depending on how the climate metric is calculated and the choice of time horizon, for example, the impact of certain species can be overestimated. This means that trade-offs between different species are inherently built into climate metrics such as the GWP (cf. Ocko et al., 2017).

These inherent biases can be problematic. Consider, for example, the case that two proposed aircraft designs B and C are compared to a reference aircraft A. Aircraft B shows a larger temperature change reduction than aircraft C, when compared to aircraft A. For the sake of fairness, it would thus be desirable that aircraft B receives a lower climate metric value than aircraft C. However, if a climate metric was to weigh the impact of a certain species more highly, and the reduction of this species resulted in the temperature reduction of aircraft design C, it is possible that aircraft C could receive the same or even a lower climate metric value than aircraft B. Just as only considering the impact of CO_2 emissions can result in aircraft designs that actually increase the total temperature (cf. Niklaß et al., 2021), using a climate metric with inherent bias towards a certain species runs the risk that new aircraft will be optimised for what is an essentially arbitrary calculation method, rather than for the actual climate impact.

The best suited climate metric for aircraft design is thus one whose value has a linear dependence on the temperature. The peak temperature as well as the average temperature over a 20-, 50- or 100-year period is used. It could be argued that the economic damage would be more appropriate than temperature, a similar argument to using damage-based climate metrics as discussed in Section 2.3, however both are outside the scope of this study.

The analysis is performed by determining the climate metric value of each fleet using the radiative forcing and temperature profiles calculated by AirClim. The change in climate metric value compared to a reference fleet (Fleet 0) is plotted against the change in peak or average temperature. The temperature results from AirClim are taken to be the truth for the purposes of this research, however more accurate models could also be used, as discussed further in the recommendations of this work (Chapter 6. A least squares regression is performed for each climate metric and the R^2 value is calculated. A linear trend, and thus a higher R^2 value is desired since this suggests that the climate metric does not have any inherent biases towards a certain species of effect.

This is further extended by repeating the analysis using each fleet as the reference against which all other fleets are compared. This method is adapted from Grewe et al. (2010). The climate metric gives an incorrect result when the signs of the temperature change and climate metric do not match for two arbitrary fleets 1 and 2, i.e. when the following is true (Grewe et al., 2010),

$$M_{err} = (\text{CM}_2 - \text{CM}_1) \times (\Delta T_2 - \Delta T_1) < 0 \qquad (3.2)$$

where CM is the climate metric in question. Graphically, this is represented by points in the top-left and bottom-right quadrants, since in these areas the signs do not match. Figure 3.5 shows these areas in red hatching for illustration. For each climate metric, it is thus possible to determine the number of fleet pairings for which the difference in the climate metric value does not correspond to the difference in temperature. The total number of incorrect pairings can be compared to the total number of pairings, which provides an indication of the quality of a

climate metric. To obtain a more detailed understanding of the origin of the incorrect pairings, a comparison can be made between the incorrect pairings and various other parameters, such as the parameters used to create the fleets (cf. Table 3.3) or the specific difference D_{err} between the climate metric values as calculated by (Grewe et al., 2010),

$$D_{err} = 2 \frac{CM_2 - CM_1}{|CM_2| + |CM_1|} \tag{3.3}$$
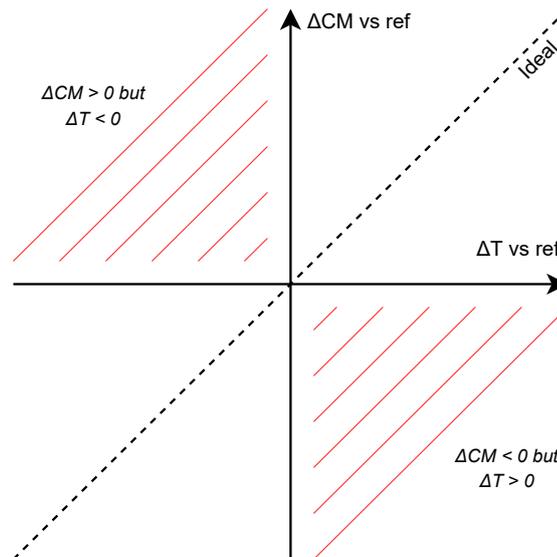
Figure 3.5: Illustration of the multivariate fleet analysis pairing method. The origin represents the current reference fleet and each fleet is compared to all others. Fleet pairings in the red, hatched quadrants are undesired since here the sign of the climate metric change and temperature change do not match. An ideal climate metric would have fleet pairings that correspond to the dashed line.

The frequency of incorrect pairings is used as a quantitative argument in the climate metric suitability framework. Furthermore, the spread and shape of the general climate metric response are analysed by plotting all fleets using each fleet as the reference.

3.7. Sensitivity Analysis

The aim of the sensitivity analysis is to quantify and visualise the dependence of climate metrics on various fleet input values and emission species and effects. This is helpful for identifying whether trends are linear in nature or more complex, and provides an opportunity for comparison with other studies for verification purposes, such as with Matthes et al. (2021) for the effect of cruise altitude.

The variables considered in the sensitivity analysis are shown in Table 3.4. The reference fleet, Fleet 0, is used and each parameter is varied separately to see its effect. A climate metric should be sensitive to all factors. The range is generally slightly higher than the values used in the Monte Carlo simulations described in the previous section.

Table 3.4: The range of parameters used in the sensitivity analysis to obtain an understanding of the dependence of each climate metric on different emission species and scenarios.

Parameter	Range
Fuel burn (40% of category)	50 - 120%
CO_2 factor	0 - 100%
NO_x factor	0 - 150%
H_2O factor	50 - 150%
Cruise pressure factor	70 - 130%
Contrail distance modifier	20 - 110%
Background emissions	SSP1 - SSP5
Fuel scenarios	Per Figure 3.4

3.8. Temporal Trajectories of CO_2-eq using Climate Metrics

Whilst a comparison between different industries can be done using climate metrics themselves, often it is easier to use the equivalent CO_2 emissions, CO_2-eq. For example, Lee et al. (2021) use the GWP, GTP (both with 20, 50 and 100 year time horizons, see their Table 5) and GWP*$_{100}$ to calculate the CO_2-eq emissions from the aviation industry in the year 2018 and are able to conclude using the GWP* that the industry is currently warming the climate at three times the rate of that from CO_2 alone.

Another use for CO_2-eq emissions is in market-based measures such as the EU Emissions Trading System (ETS). There is debate on how exactly the climate charge for aircraft operators should be created, for example as a constant, average emissions multiplier or as a more complex multiplier established on a per-flight basis (European Union Aviation Safety Agency, 2020; Niklaß et al., 2019). However, all methods of establishing a climate charge require the use of a climate metric and the underlying principle is the same: An operator must, once the climate charge is implemented, pay for or offset each tonne of CO_2-eq emitted, not just CO_2 which is currently the case. The benefit of such a system, referred to as a single-basket approach, is that only the cost per tonne of CO_2-eq must be negotiated, rather than the cost of each emission species individually. For conventional climate metrics such as the GWP, the calculation method per emission species is very simple, namely,

$$E_{CO_2\text{-eq}}(t) = \text{GWP}(t) \times E_{CO_2}(t) \tag{3.4}$$

where E is the yearly rate of an emission in the year t. In this case, the GWP is also a function of time since the value changes over time and with the emission profile. The same method can also be used for the GTP and ATR. The GWP*, in comparison, calculates the warming-equivalent CO_2 emissions, denoted CO_2-we, which are explained in Section 3.4.1.

To provide an indication as to how well climate metrics can be used for the purposes described above, the temporal trajectories of CO_2-eq emissions are calculated for each fuel scenario shown in Figure 3.4, namely CurTec, CORSIA, FP2050, COVID-15s and Fa1. The fuel scenarios are used as inputs into AirClim simulations with the SSP2-4.5 "Middle of the Road" background emissions scenario. Using the resulting trajectories of radiative forcing and temperature, the climate metric values are calculated and used to determine the CO_2-eq emissions. Each climate metric will use a time horizon of 100 years.

Depending on the objective, it can also be useful to estimate the temperature response using a climate metric. This can, for example, be used to estimate whether current strategies are set to meet climate targets. For the time horizons analysed in this study, the temperature change can be seen to vary linearly with the CO_2 emission. By extension, a good climate metric should calculate a CO_2-eq emission rate that also varies linearly with the temperature response. The temperature estimation can be written as,

$$\Delta T(t) = \text{TCRE} \times \sum_{\tau=t_0}^{t_0+t} E_{CO_2\text{-eq}}(\tau) \tag{3.5}$$

where TCRE is the transient climate response, defined as 0.49 K/TtCO$_2$ (Cain et al., 2019) in this study for consistency. The summation is cumulative in that the emissions from the first emission until the time in question are summed. For the GWP*, an extra term must be added to obtain,

$$\Delta T(t) = \text{TCRE} \times \left(\sum_{\tau=t_0}^{t_0+t} E_{CO_2\text{-we}}(\tau) + \sum_{\tau=t_0}^{t_0+t} E_{CO_2\text{-e}}(\tau) \right) \tag{3.6}$$

as described in more detail in Section 2.4.3. Here, $E_{CO_2\text{-we}}$ are emissions calculated using the GWP*, and $E_{CO_2\text{-e}}$ are emissions calculated using the GWP. For aviation emissions, these latter emissions originate only from CO_2 itself. Therefore, in practice, the CO_2 emissions can be used directly.

Finally, the above temperature estimation is repeated for the fleets generated for the multivariate fleet analysis. The estimated peak temperature is then compared to the peak temperature calculated using AirClim, with the aim of identifying the influence of the TCRE and the background emissions scenario. An ideal climate metric would show a linear trend for a given background emissions scenario.

$$4$$

Results

This chapter presents the results of the analysis methods described in the previous chapter. Only the results are shown in this chapter - the interpretation of the results to inform the choice of best climate metric is performed in Chapter 5.

The general response of climate metrics is shown in Section 4.1, in which the response of climate metrics to pulse, sustained and increasing emissions and their dependence on the background emissions scenario and time horizon are analysed. Section 4.2 shows the sensitivity of climate metrics to changes in emissions and flight conditions, which is used to inform the multivariate fleet analysis. This analysis is shown in Section 4.3 with single and pairwise fleet comparisons and an investigation into incorrect fleet pairings. Finally, in Section 4.4, the trajectory of CO_2-eq emissions and an estimation of the temperature response using climate metrics are shown.

4.1. General Response of Climate Metrics to Aviation Emissions

This section provides an overview of the general responses of climate metrics to aviation emissions. The aim of this section is to identify any potential systematic issues with existing climate metrics that would make them inappropriate for aircraft climate assessments. As described in Section 3.4, the climate metrics are calculated for pulse, sustained and increasing emissions starting in the year 2020 for different background emission scenarios. First, the temperature response is determined (Section 4.1.1), followed by the responses of each climate metric in Sections 4.1.2 to 4.1.7. The final section provides a summary of the results (Section 4.1.8), including the numerical dependence of the climate metrics on the background emissions scenario in Table 4.2.

4.1.1. Temperature Response

The temperature response of a pulse, sustained and increasing emission as described in Section 3.4 are calculated using AirClim and shown in Figure 4.1. These results are provided here for reference since they allow for a means of comparison between the climate metrics. By nature of the climate metric, the results are also equivalent to the AGTP. Similarly to the AR6 (Intergovernmental Panel on Climate Change, 2021), this research makes no assumptions as to the most likely scenario. For consistency, however, SSP2-4.5 "Middle of the Road" is plotted with a solid line, the other scenarios are plotted as uncertainty margins.

As Figure 3.1 shows, the SSP trajectories start to diverge from one another in the year 2020 for CH_4 and about 2030 for CO_2. Divergence between scenarios can be seen from about the year 2040 in the temperature response, which reflects the temporal delay from emission to temperature response. As is described in the following section on radiative forcing in more

Figure 4.1: Temperature response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5 (cf. Section 3.4)

detail, for scenarios with rapidly increasing emissions, the absolute impact and thus temperature response of an emission is lower. If taken at face value, this could lead to a devaluation of emissions in these scenarios, which may not be the aim of stakeholders.

4.1.2. Radiative Forcing

Figure 4.2 shows the radiative forcing (RF) response to a pulse, sustained and increasing emission (cf. Figure 3.2) for SSP1 to SSP5 (see Section 3.4, Figure 3.1). The pulse response demonstrates the very short lifetimes of contrails, ozone and water vapour, which decline rapidly to zero immediately following the pulse. Since the smallest time step in AirClim is one year, a higher level of detail is not possible. For readability, the initial values of ozone and contrails have been cut from the figure because they do not fit in the scale. The values of each are 108.6 and 149.5 mW/m^2 respectively, resulting in a total, initial radiative forcing of 264.4 mW/m^2. The longer-term cooling impacts of methane (CH_4) reduction and the primary mode ozone (PMO) effect are also clear. However, the positive contribution from CO_2 begins to dominate already after three years.

Figure 4.2: Radiative forcing response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

As a result of their very short lifetimes, the RF from ozone, contrails and water vapour stabilise within the first two years following the start of the sustained emission, as shown by

the horizontal nature of their responses in Figure 4.2b). In comparison, the RF from CO_2 does not stabilise, i.e. approach the horizontal, within the 100 year time horizon due to its large atmospheric lifetime. However, the RF from CO_2 shows a high dependence on the background emission scenario, as can be seen by the size of the shaded region. In scenarios where background CO_2 emissions are increasing rapidly, such as for SSP5, the RF due to CO_2 begins to decrease. In other words, the relative importance of long-lived climate pollutants decreases for background scenarios with increasing emissions. The opposite is also true: For scenarios such as SSP1 with a high reduction in CO_2 emissions, the importance of CO_2 increases. This demonstrates the necessity of REQ 4 (Section 3.3).

The changes in RF due to CH_4 reduction and the PMO effect are also dependent on the background emissions scenario, but, as short-lived effects, show the opposite behaviour. As shown in Figure 3.1, SSP1 shows the highest reduction in CH_4 emissions and SSP3 the lowest and latest. As a result, for SSP1 the RF from CH_4 reduction will become less negative over time; for SSP3 it will become more negative.

The results for an increasing emission in Figure 4.2c) are visually quite similar to those of the sustained emission. One difference is that the ozone and contrail responses no longer stabilise, but increase with the time horizon due to the increasing fuel use. The ozone response also becomes larger than the contrail response at a time horizon of around 60 years. This is because the lifetime of ozone is larger than one year, meaning that with a yearly time-step the radiative forcing will continue to build up. Contrails have a lifetime much less than one year and thus do not display the same behaviour, growing less quickly. A second noticeable difference is the larger dependence on the time horizon for both CO_2 and CH_4 reduction. This is because a larger amount of emissions are released at a later time, when background emissions have increased, or decreased in the case of SSP1, compared to the year 2020.

For a sustained or increasing emission, therefore, the total RF can show a wide variety of results after 40 years, predominantly due to the dependence of the RF from CO_2 on the background CO_2 concentration. This shows the challenge of using RF as a climate metric: as long as global CO_2 emissions continue to increase rapidly, the RF can give the impression that constant or slowly increasing emissions do not lead to further warming. This is an undesired quality of a climate metric that is to be used by stakeholders to make decisions to limit warming.

This is compounded by the NO_x results shown in Figure 4.3. For H = 1 year, the pulse emission shows warming; however, for H > 1 year, the pulse emission shows cooling and approaches zero from below. This is in contrast to the sustained emission, for which the radiative forcing remains positive for all H. Ideally, a climate metric should show the same results regardless of whether a pulse or sustained emission is used, since a sustained emission is a summation of pulse emissions (cf. REQ 6). This effect is not seen for aviation specifically because the initially largest contributors to radiative forcing, namely contrails and ozone, decline to near-zero within the first year. This is in line with the analysis performed by Dahlmann (2011, see her Figure 3.3).

The RF has been used as a relative climate metric, denoted the Radiative Forcing Index (RFI). The RFI response for the same emission profiles is shown in Figure 4.4. Where, like the RF responses as shown above, the results of the sustained and increasing scenario are similar, the RFI response to a pulse emission is not. Similarly to the NO_x results, the total RFI can be negative at low time horizons and then approaches the CO_2 response from below. It is clear that this difference means that the RF and RFI do not meet the requirement that the
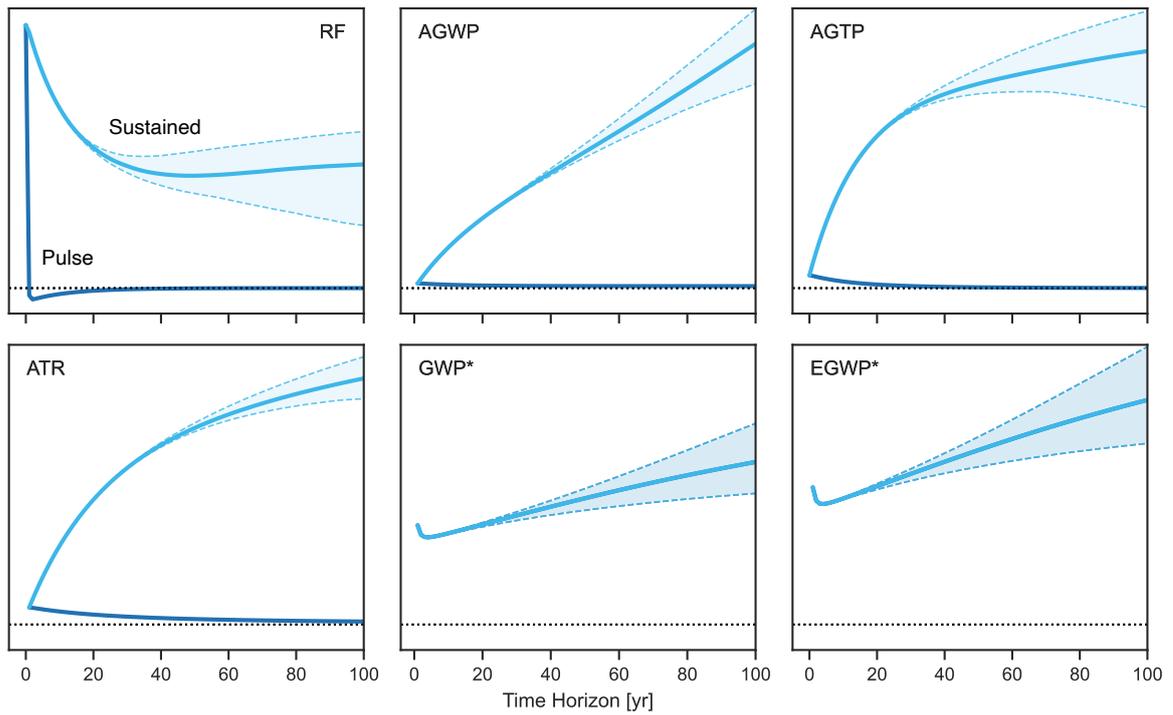
Figure 4.3: Analysis of the climate metric response to a pulse emission (dark blue) and a sustained emission (light blue) of aviation NO_x. The y-axis is not labelled for clarity since the aim is to identify the general trend of each climate metric. The dotted black line is y = 0. Note: for the GWP* and EGWP*, the pulse and sustained emissions overlap.

response from a sustained emission should be equal to that from a sum of pulse emissions (REQ 6).
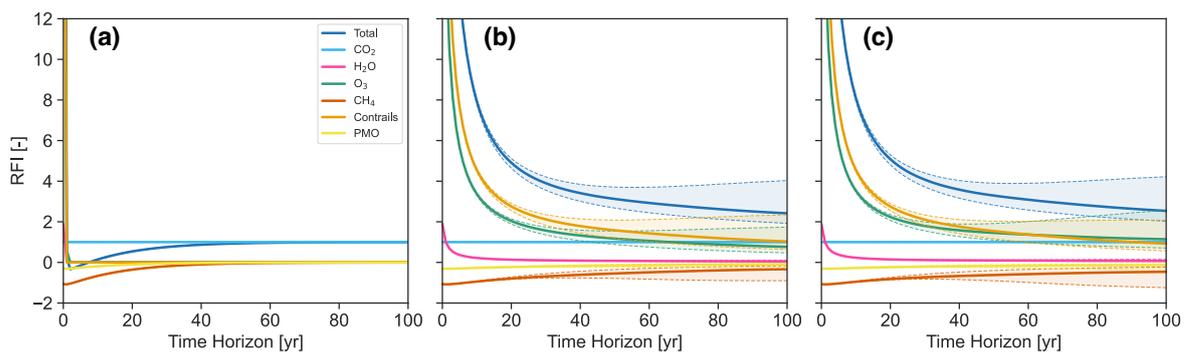
Figure 4.4: Radiative Forcing Index (RFI) response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

4.1.3. Global Warming Potential

The response of the Absolute Global Warming Potential (AGWP) to a pulse, sustained and increasing emission is shown in Figure 4.5. As is evident when comparing Figures 4.5a) and 4.2b), the temporal integration performed in the calculation of the AGWP substantially reduces the dependence on the background emission scenario. Another benefit of the AGWP compared to the RF is that the impact of short-lived climate pollutants such as contrails and ozone for a pulse emission remains non-zero with increasing time horizon, rather than decaying to zero after the first year as in Figure 4.2a).

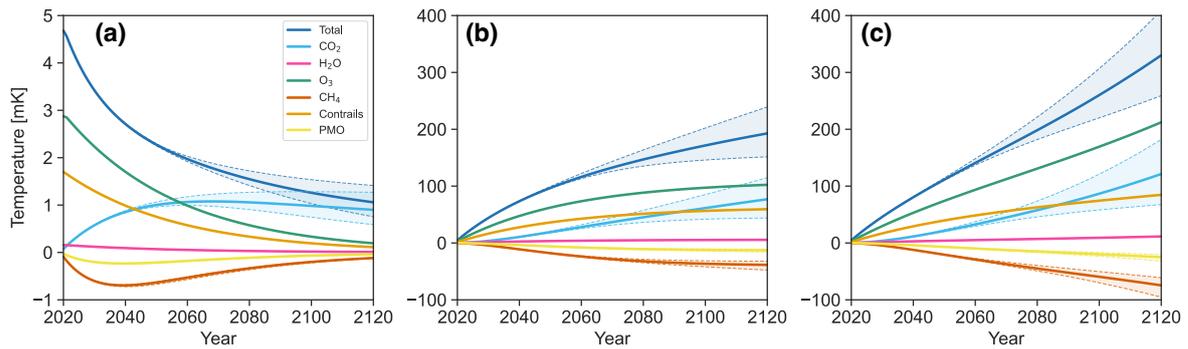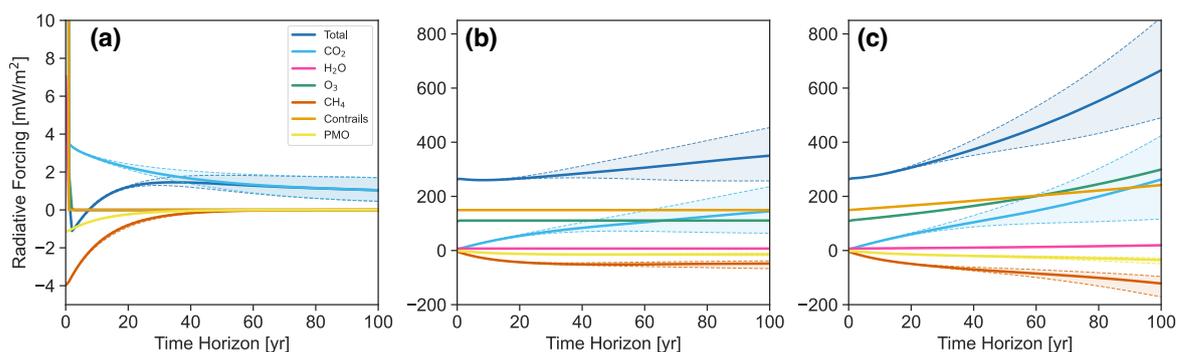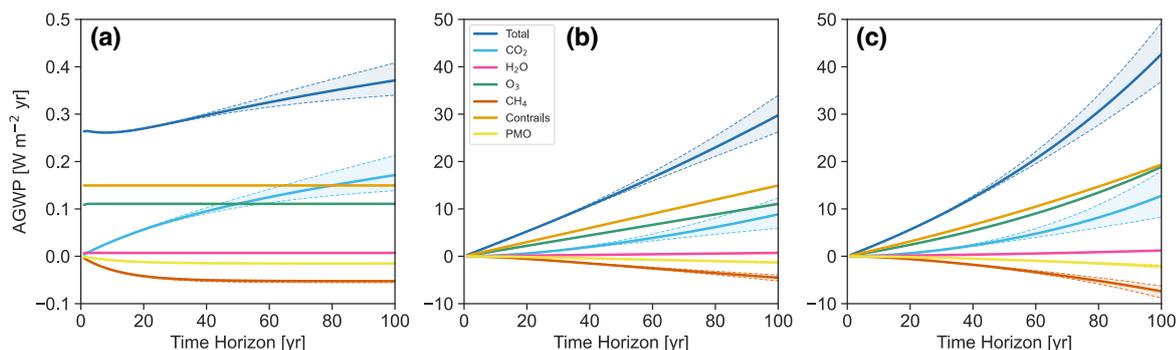Figure 4.5: Absolute Global Warming Potential (AGWP) response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

The sustained emission response is shown in Figure 4.5b). Since all aviation emissions except CO_2 are short-lived, the response of the AGWP to a sustained emission is mostly linear. The CO_2 and CH_4 responses are the outliers, and are the only two effects to also have a noticeable dependence on the background emission scenario. As before with the RF, this is due to changes in global CO_2 and CH_4 emissions: for scenarios with rapidly increasing emissions, the absolute value of the emission decreases with increasing time horizon.

The AGWP response due to an increasing emission is non-linear, as shown in Figure 4.5c). The ozone and contrail responses are noticeably non-linear too. The ozone response dominates over the contrail response because ozone has a lifetime larger than one year. This effect is the direct result of the RF response shown in Figure 4.2c) and is described in the accompanying text. The higher dependence of CO_2 and CH_4 on the background emission scenario for increasing emissions compared to sustained emissions, as also identified in the previous section on RF, is also to be seen in Figure 4.5c).

The AGWP generally shows the same sign for a pulse and sustained emission regardless of time horizon or background emission scenario, as demonstrated by Figure 4.3 for NO_x. Although it is possible that, with a large, initially negative RF, the pulse emission of the AGWP switches to a negative value, this is unlikely to occur for total aviation emissions due to the large positive RF from contrails.

The response of the GWP to a pulse, sustained and increasing emission is shown in Figure 4.6. Although the values differ between the three emissions, it is clear that they all show the same overall trend, which is desirable for a climate metric. However, the criticism that there is no obvious choice for the time horizon is clear. For short-lived pollutants such as contrails and ozone, the GWP varies by more than an order of magnitude: For a pulse emission of contrails,

for example, the GWP varies from 35.7 at H = 1 to 0.78 at H = 100. Since CO_2 is the only long-lived aviation emission, with time horizons above 100 years the GWP values of all other species continue to tend to zero.
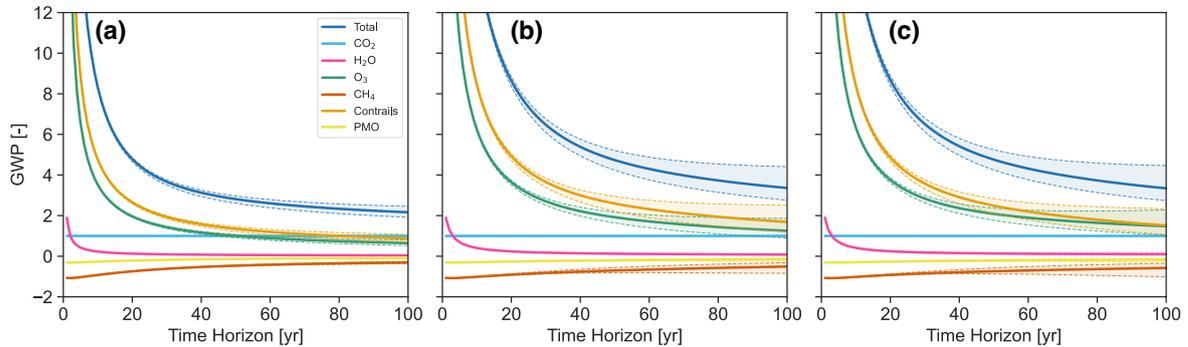
Figure 4.6: Global Warming Potential (GWP) response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

An important note should be made here about the dependence on the background emissions scenario. Where for the AGWP higher decarbonisation in for example SSP1-1.9 leads to higher climate metric values and vice versa, for relative climate metrics this effect reverses. For the GWP as shown in Figure 4.6, the higher trend of the total emissions corresponds to scenarios with quickly increasing emissions (SSP5-8.5) and the lower trend to SSP1-1.9. In other words, the value of an emission in a scenario with rapidly increasing global emissions is valued more than in a scenario with decreasing emissions. This switch is seen for the GWP, GTP and ATR, as is shown in the following figures, and will be further discussed in Section 5.1. The numerical dependence of the GWP on the background emission scenario is given in Table 4.2.

4.1.4. Global Temperature Change Potential

Since AirClim provides a direct way of calculating the temperature response, the AGTP is equivalent to the temperature profile shown in Figure 4.1 where the time horizon H is the time after the year 2020. As with the previous climate metrics, the impacts from CO_2, CH_4 reduction and PMO are dependent on the background emission scenario. However, also from Figure 4.3 it is clear that the AGTP is more dependent on the background emission scenario than the AGWP, as numerous authors have previously identified (Shine, 2009; Dahlmann, 2011).

The behaviour of the GTP, shown in Figure 4.7, is visually similar to the GWP. An interesting observation is that the GTP shows a lower impact from contrails than from ozone, which is opposite to the results obtained using the RF, GWP and GWP*, and more accurately reflects the temperature profile. This observation is important since it highlights the influence of the efficacy. As discussed in Section 2.3 in reference to the cause-and-effect chain (Figure 2.3), this is an advantage of temperature-based climate metrics, which are one step lower on the chain: a temperature change is more relevant to society than a change in RF.

The disadvantages of the GTP are, however, also clear. Along with the high dependence on the background emissions scenario, the nature of the GTP as an endpoint climate metric means that the responses of short-lived species and effects are generally highly dependent
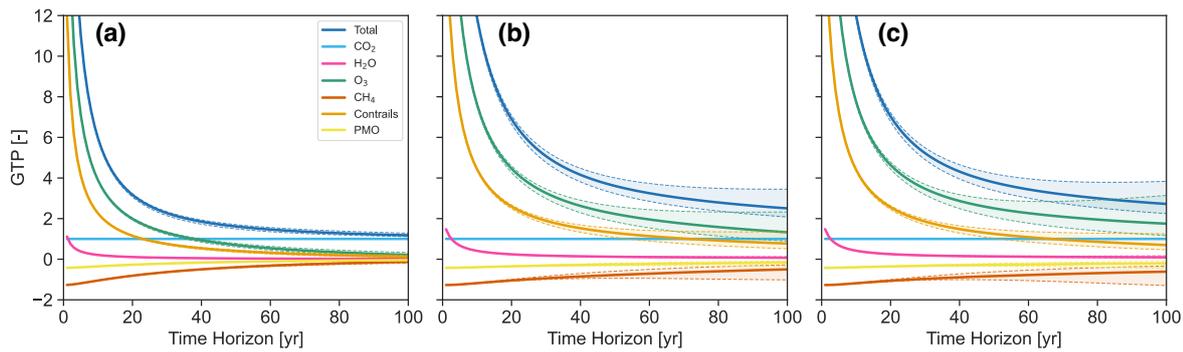
Figure 4.7: Global Temperature Change Potential (GTP) response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

on the time horizon. This is clear from the steep gradient and lower values of the short-lived effects in Figure 4.7 and is due to the lack of memory of previous temperature changes by the GTP.

4.1.5. Average Temperature Response

The Average Temperature Response (ATR) combines the advantages of the GWP and GTP and its response to a pulse, sustained and increasing emission is shown in Figure 4.8 in its absolute form. It calculates a temperature response, shows a smaller dependence on the background emissions scenario than the GTP and, like the GWP, has a temporal memory. The latter two elements are due to the temporal averaging performed in the calculation of the ATR. For a pulse emission, this results in shallower gradients than the GTP, but also higher temperature change and a higher dependence on the time horizon. A general stabilisation of the CO_2 response can also be seen within a time horizon of 100 years.
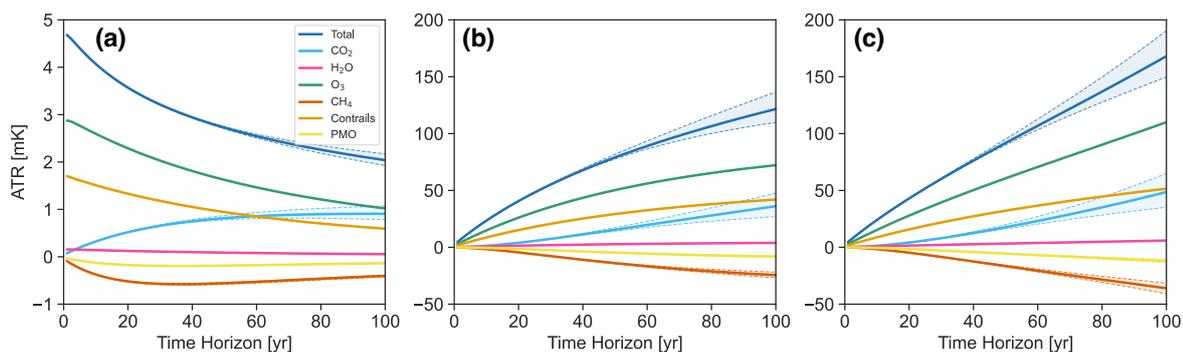
Figure 4.8: Absolute Average Temperature Response (ATR) of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

For a sustained or increasing emission, the ATR can once again be seen to vary less with the background emission scenario. This is also clear from Figure 4.3 for a NO_x sustained emission. However, the ATR, compared to the AGTP showing the actual temperature response,

does generally show lower temperature changes.

The response of the relative ATR (ATR-rel) to the three emissions is shown in Figure 4.9. The effect of averaging is clear from the shallower gradients especially of the contrail and ozone responses. The shallower gradient, however, means that the total ATR-rel is more dependent on the time horizon than other climate metrics. The ATR-rel value at the lowest commonly used time horizon of 20 years is much higher than the value at 100 years, compared to the other climate metrics analysed.
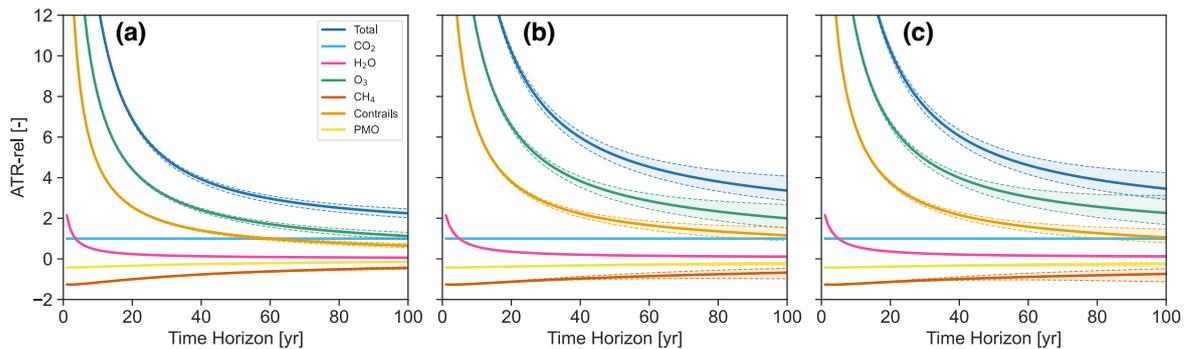
Figure 4.9: Relative Average Temperature Response (ATR-rel) of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5.

4.1.6. GWP*

As described in Section 3.4, the GWP* method requires the pulse response of the CO_2 AGWP, which differs depending on the background emission scenario, as shown in Figure 4.5. Figure 4.10 shows the AGWP response for different background emissions scenarios compared to the response used by the IPCC in AR5 (Intergovernmental Panel on Climate Change, 2013, Figure 8.29). In this work, the response from SSP2-4.5 is used for consistency, unless otherwise stated.
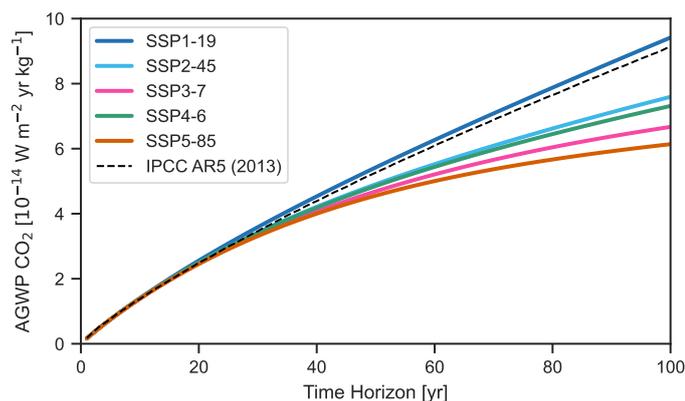
Figure 4.10: Response of the CO_2 AGWP to a 1 kg pulse emission in the year 2020 for different background emission scenarios calculated using AirClim, compared to an AGWP response used by the IPCC in AR5 (Intergovernmental Panel on Climate Change, 2013, Figure 8.29)

The response of the GWP* to a pulse, sustained or increasing emission is shown in Figure 4.11, where the value is taken at the year of the pulse or start of the sustained emission, rather than at a peak such as in the multivariate fleet analysis (see Section 3.4.1). Unlike all previous climate metrics, the GWP* has identically the same response for the three emissions. This is only the case when an analysis is done for the starting year since in this case $\Delta F / \Delta t$ and $\overline{F}$ in Eq. (3.1) are identical for both emission scenarios.
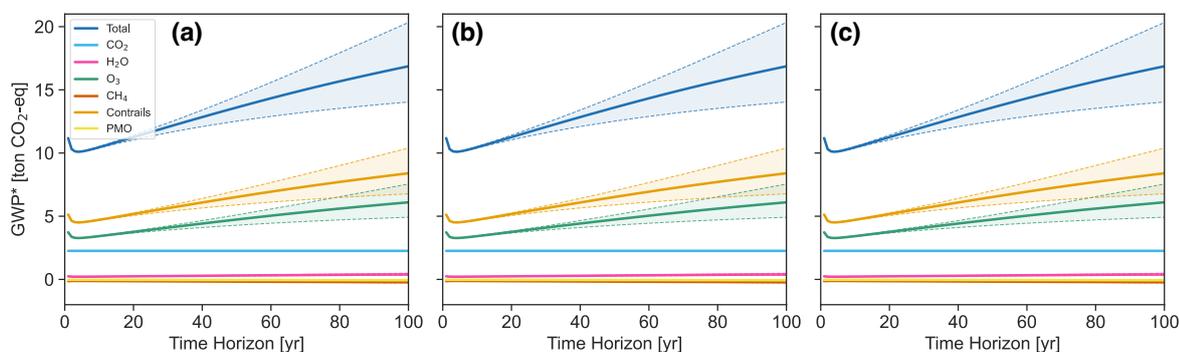
Figure 4.11: GWP* (s = 0.25) response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5. Both responses are identical.

As Figure 4.11 shows, the GWP* is dependent on the time horizon and the background emissions scenario. Both dependencies are due to the differences in the CO_2 AGWP trajectory as shown in Figure 4.10. Unlike the GWP, which as a temporal integration considers the trajectory of the RF up to the time horizon, the GWP* only considers the change in RF and average RF in the previous Δt years, which is constant with increasing time horizon. Only the time horizon H and the CO_2 AGWP change with increasing time horizon, giving the GWP* a dependence on the time horizon. It is important to note the scale of the plot: although visually the GWP* shows a large dependence on the background emissions scenario, due to the high values obtained compared to the other climate metrics the dependency is fairly low. As with the relative climate metrics described above, the upper values in Figure 4.11 correspond to SSP5-8.5 and the lower values to SSP1-1.9.

The initial drop of the GWP* value at low time horizons is a result of the value of the stock term s, which is generally the result of a regression (Cain et al., 2019; Smith et al., 2021). A value of 0.25, as established by Cain et al. (2019), is used in this work for consistency. However, it should be noted that this value was obtained primarily for methane emissions and it is thus questionable whether the same value should be used for all aviation emissions. An investigation into the optimal value of s for aviation did not provide a more optimal value of s due to the high uncertainties involved and the large influence of the transient climate response (TCRE), as is demonstrated in Section 4.4.1. The investigation is thus not presented in this work and its further development is left as a recommendation. The influence of the parameter s on the climate metric results is discussed in more detail in Section 5.2.

4.1.7. EGWP*

Since the EGWP* differs to the GWP* only by multiplication of the efficacy, the results obtained are qualitatively very similar. As the summary table shows, the dependence on the background

emissions scenario is almost identical. The main difference is the influence of ozone and contrails, which now more closely resemble the temperature change as shown in Figure 4.1. Otherwise, the pulse, sustained and increasing emission responses are identical for the same reason as described above.
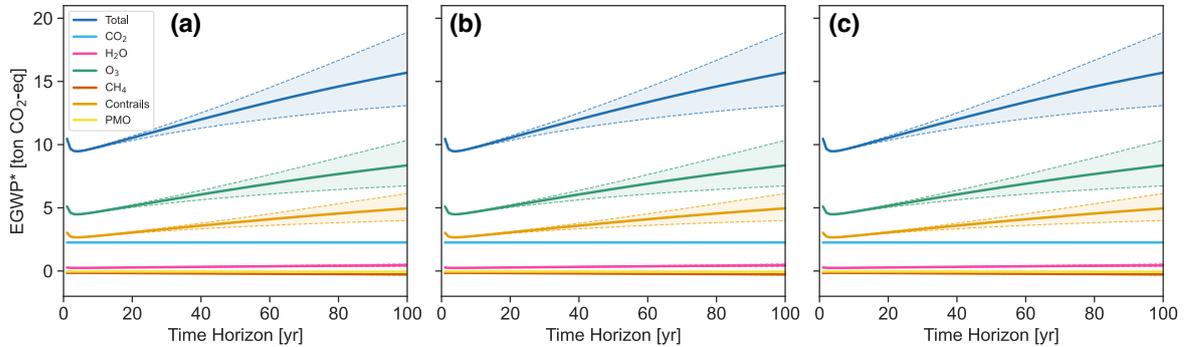
Figure 4.12: EGWP* (s = 0.25) response of **(a)** a pulse emission, **(b)** a sustained emission and **(c)** a 1% increasing emission for background emission scenario SSP2-4.5 with uncertainty margins for scenarios SSP1 to SSP5. Both responses are identical.

4.1.8. Quantification of Time Horizon and Background Emissions Scenario Dependency

This section provides a quantification of the dependence of climate metrics on the time horizon and background emissions scenario for the analysis of REQ 3 and REQ 4. Table 4.1 shows the dependence on the time horizon for the total RFI, GWP, GTP, ATR-rel, GWP* and EGWP* for the background emission scenario SSP2-4.5.

Table 4.1: Summary of the dependence of the total climate metric values on the time horizon. The darker and the bluer the colour, the less dependent; the darker and the redder the colour, the more dependent. The values are for total emissions and use the 100-year time horizon with background emissions scenario SSP2-4.5 as the basis.

		RFI	GWP	GTP	ATR-rel	GWP*	EGWP*
Pulse	100	**1**	**2.16**	**1.18**	**2.25**	**16865.5**	**15694.8**
	20	-45.8 %	+120.9 %	+169.2 %	+206.4 %	-33.2 %	-32.9 %
	50	-6.5 %	+30.0 %	+37.1 %	+48.8 %	-19.4 %	-19.1 %
Sustained	100	**2.42**	**3.36**	**2.5**	**3.36**	**16865.5**	**15694.8**
	20	+103.0 %	+156.8 %	+175.6 %	+203.3 %	-33.2 %	-32.9 %
	50	+29.3 %	+40.9 %	+44.1 %	+51.9 %	-19.4 %	-19.1 %
Increasing	100	**2.54**	**3.34**	**2.72**	**3.46**	**16865.5**	**15694.8**
	20	+100.3 %	+159.4 %	+158.3 %	+196.9 %	-33.2 %	-32.9 %
	50	+29.7 %	+42.3 %	+39.0 %	+50.1 %	-19.4 %	-19.1 %

As noted before, the RFI for a pulse emission approaches the CO_2 response from below with increasing time horizon. The slope of the RFI is, however, lower than for the GWP, GTP and ATR-rel for all emissions. The high dependence on the time horizon from the GWP and

GTP is clear, especially for low time horizons such as 20 years. If used at higher time horizons, the dependence could be seen as acceptable but must still be accounted for. The ATR-rel shows the highest dependence on the time horizon. This is due to the shallower gradients obtained by averaging, as is clear from Figure 4.9. The choice of time horizon is thus an important consideration also for the ATR, especially since the ATR_{20} is often used for example for route optimisation (e.g. Frömming et al., 2021). An interesting observation is that the ATR-rel and GWP show quite similar values for H = 100 years. Niklaß et al. (2019), for example, use this behaviour to describe a conversion factor between the two climate metrics which would help in the introduction of the ATR into policy. The CO_2-eq emission profiles in Section 4.4.1 also show similar behaviour.

As is clear from the figures above, the value of the GWP* and EGWP* increases with time horizon, opposite to the relative climate metrics in this list. However, both have the lowest dependence on the time horizon. This must be understood in context, however, because unlike the relative climate metrics, the GWP*/EGWP* values do not act asymptotically within the 100-year time horizon for all background emission scenarios. The GWP* thus also falls foul of the criticism that the 100 year time horizon is arbitrary (cf. Section 2.4.2), although the dependence is lower than for the GWP. The claim that results under the GWP* are insensitive to the time horizon provided that the time horizon is much larger than the lifetime of a species (Allen et al., 2018) should thus be called into question when the GWP* is applied to aviation emissions. As the results for contrails and ozone in Figure 4.11 attest, the GWP* results vary with the time horizon and do not stabilise within a 100-year time horizon for all background emission scenarios, even though their atmospheric lifetimes are notably short.

Figure 4.13 extends the analysis in Table 4.1 to include the dependency of individual species and effects. The figures shows the percentage of each species or effect to the total climate metric value for the three most common time horizons of 20, 50 and 100 years. Here, only the increasing emission response is shown, the pulse and sustained emission responses are included in Section A.1.

The climate metrics on the top row all make use of the RF, whereas those on the bottom row use the effective RF (EGWP*) or temperature. The main difference is that the contrail impact dominates the climate metrics using RF, whereas the ozone impact dominates those using ERF and temperature. Across all climate metrics, the ozone impact is very consistent and hardly depends on the time horizon. CO_2 and CH_4 generally have high dependencies; for RF, AGWP and AGTP contrails also have a clear dependence.

Although the total ATR-rel shows a high dependence on the time horizon in Table 4.1, except for CO_2 and CH_4 the species and effects in Figure 4.13 have low dependence. The GWP* and EGWP* have low dependence for all emission species and effects, but demonstrate an opposite trend for CO_2 as explained above. The figure also shows the very low influence of the CH_4 reduction to the total value for the starred climate metrics, compared to the others shown.

Table 4.2 provides an overview of the dependence of the temperature response and the climate metrics on the background emissions scenario. The dependence has two potential origins. First, changes in background emissions, specifically CO_2 and CH_4 which are inputs to the AirClim simulation, have a direct impact on the response of CO_2, CH_4 and PMO. Second, the division by the CO_2 response to obtain a relative climate metric causes further dependence on the background emissions scenario. Since within AirClim the impact and lifetimes
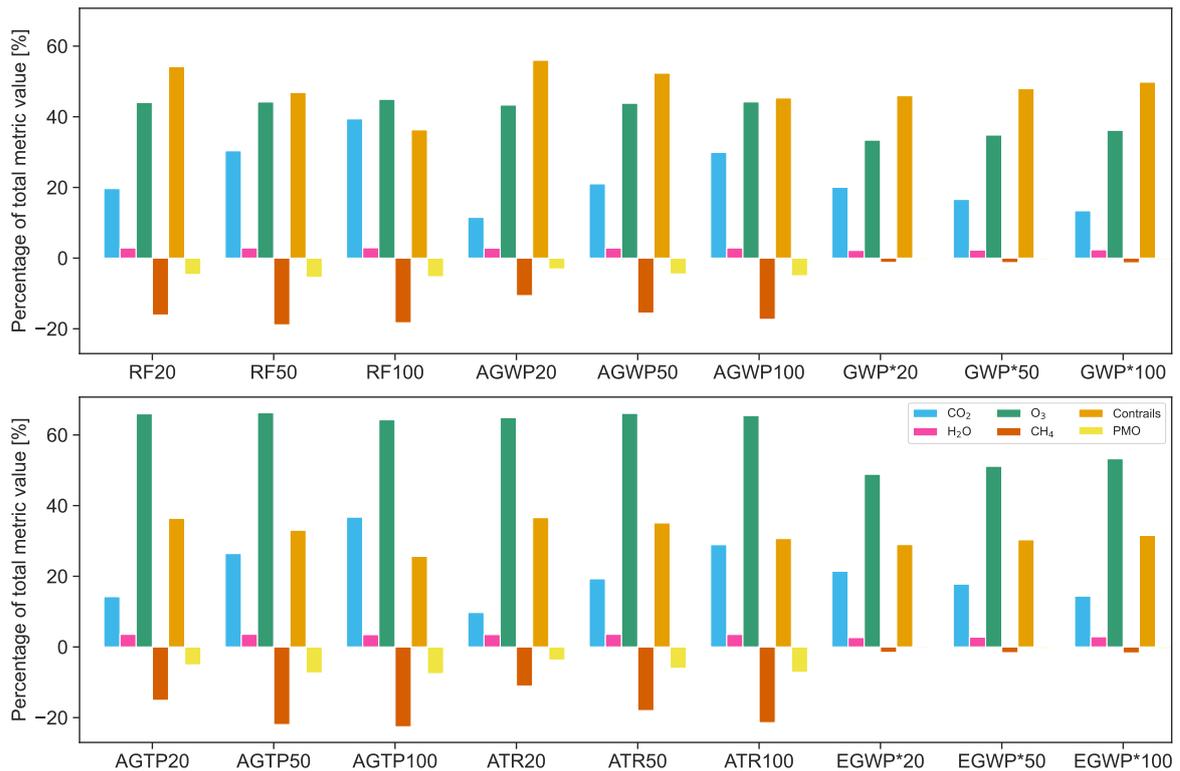
Figure 4.13: Percentage contribution of each species or effect to the total climate metric value for time horizons 20, 50 and 100 years for the 1% increasing emission scenario. The results of the SSP2-4.5 scenario are used. All metrics are shown in their absolute form.

of H_2O, O_3 and contrails are not dependent on the background emissions of CO_2 and CH_4, the dependence of these species is identical per climate metric and fuel scenario. A similar statement can be made for the impact of CH_4 reduction and PMO, which are dependent only on the background CH_4 emissions.

In general, the sustained and increasing emissions have higher dependencies than the pulse emission. This is due to emissions being released in the future being subjected to a larger divergence in background emissions, as shown in Figure 3.1. A similar reasoning likely explains why the largest dependencies originate from the end-point climate metrics RF and GTP: Since end-point climate metrics only consider the value at one time horizon in the future, the result is that end-point climate metrics are subjected to a larger divergence in background emissions than climate metrics that perform integration or averaging.

An important observation from the results is that the influence of the CO_2 background emissions is larger than that of the CH_4 background emissions. The CO_2 emission profile is thus more important for the temperature change than the CH_4 profile. This can be seen for the results of the temperature in the first row of Table 4.2. For the relative climate metrics in the table that divide by the CO_2 response, this, therefore, means that the CO_2 dependence will dominate over the CH_4 dependence.

Table 4.2: Summary of the dependence of climate metrics on the background emissions scenario. The darker and the bluer the colour, the less dependent; the darker and the redder the colour, the more dependent. All values are shown in their absolute form and all climate metrics are shown in their relative form.

Metric		Pulse (% diff)							Sustained (% diff)							Increasing (% diff)						
		Total	CO_2	H_2O	O_3	CH_4	Cont	PMO	Total	CO_2	H_2O	O_3	CH_4	Cont	PMO	Total	CO_2	H_2O	O_3	CH_4	Cont	PMO
Temp	Upper	33.5	41	0	0	5.5	0	5.5	24.1	49.3	0	0	16.7	0	16.7	23.8	50.2	0	0	18	0	18
	Lower	28.8	34.6	0	0	9.1	0	9.1	21.3	42.9	0	0	23.5	0	23.5	21.5	44.3	0	0	27.7	0	27.7
	Total	**62.3**	**75.6**	**0**	**0**	**14.6**	**0**	**14.6**	**45.4**	**92.2**	**0**	**0**	**40.2**	**0**	**40.2**	**45.3**	**94.5**	**0**	**0**	**45.7**	**0**	**45.7**
RFI	Upper	73.8	0	7.5	7.5	22.4	0	22.4	66.9	0	127.3	127.3	51.3	127.3	51.3	66.3	0	125.9	125.9	51.1	125.9	51.1
	Lower	60.9	0	10	10	95.2	0	95.2	20.3	0	38.5	38.5	171.1	38.5	171.1	20	0	38.1	38.1	168.4	38.1	168.4
	Total	**134.7**	**0**	**17.5**	**17.5**	**117.6**	**0**	**117.6**	**87.2**	**0**	**165.8**	**165.8**	**222.4**	**165.8**	**222.4**	**86.3**	**0**	**164**	**164**	**219.5**	**164**	**219.5**
GWP	Upper	13.3	0	23.7	23.7	14.4	23.7	14.4	31.2	0	48.9	48.9	37.3	48.9	37.3	33.5	0	54.4	54.4	39.7	54.4	39.7
	Lower	11.3	0	19.4	19.4	20.6	19.4	20.6	17.9	0	28.1	28.1	64.8	28.1	64.8	18.1	0	29.2	29.2	74.8	29.2	74.8
	Total	**24.6**	**0**	**43.1**	**43.1**	**35**	**43.1**	**35**	**49.1**	**0**	**77**	**77**	**102.1**	**77**	**102.1**	**51.6**	**0**	**83.6**	**83.6**	**114.5**	**83.6**	**114.5**
GTP	Upper	9	0	53	53	22.6	53	22.6	37.7	0	75.1	75.1	44.2	75.1	44.2	40.9	0	79.5	79.5	45.4	79.5	45.4
	Lower	7.3	0	29.1	29.1	46.1	29.1	46.1	16.8	0	33	33	103	33	103	17.5	0	33.4	33.4	110.7	33.4	110.7
	Total	**16.3**	**0**	**82.1**	**82.1**	**68.7**	**82.1**	**68.7**	**54.5**	**0**	**108.1**	**108.1**	**147.2**	**108.1**	**147.2**	**58.4**	**0**	**112.9**	**112.9**	**156.1**	**112.9**	**156.1**
ATR-rel	Upper	9.6	0	15.9	15.9	10.3	15.9	10.3	21.2	0	34.2	34.2	32.1	34.2	32.1	23	0	38.1	38.1	34.6	38.1	34.6
	Lower	9.8	0	15.3	15.3	13.3	15.3	13.3	14.7	0	24.1	24.1	44.7	24.1	44.7	15.3	0	25.2	25.2	52.2	25.2	52.2
	Total	**19.4**	**0**	**31.2**	**31.2**	**23.6**	**31.2**	**23.6**	**35.9**	**0**	**58.3**	**58.3**	**76.8**	**58.3**	**76.8**	**38.3**	**0**	**63.3**	**63.3**	**86.8**	**63.3**	**86.8**
GWP*	Upper	20.5	0	23.7	23.7	19.7	23.7	19.7	20.5	0	23.7	23.7	19.7	23.7	19.7	20.5	0	23.7	23.7	19.7	23.7	19.7
	Lower	16.8	0	19.4	19.4	23.6	19.4	23.6	16.8	0	19.4	19.4	23.6	19.4	23.6	16.8	0	19.4	19.4	23.6	19.4	23.6
	Total	**37.3**	**0**	**43.1**	**43.1**	**43.3**	**43.1**	**43.3**	**37.3**	**0**	**43.1**	**43.1**	**43.3**	**43.1**	**43.3**	**37.3**	**0**	**43.1**	**43.1**	**43.3**	**43.1**	**43.3**
EGWP*	Upper	20.3	0	23.7	23.7	19.7	23.7	19.7	20.3	0	23.7	23.7	19.7	23.7	19.7	20.3	0	23.7	23.7	19.7	23.7	19.7
	Lower	16.6	0	19.4	19.4	23.6	19.4	23.6	16.6	0	19.4	19.4	23.6	19.4	23.6	16.6	0	19.4	19.4	23.6	19.4	23.6
	Total	**36.9**	**0**	**43.1**	**43.1**	**43.3**	**43.1**	**43.3**	**36.9**	**0**	**43.1**	**43.1**	**43.3**	**43.1**	**43.3**	**36.9**	**0**	**43.1**	**43.1**	**43.3**	**43.1**	**43.3**

Generally, a pattern in the individual species can be seen for the relative climate metrics and the GWP*/EGWP*, namely that the upper values cause a larger absolute change for H_2O and O_3. For relative climate metrics and emissions causing a warming, the upper values correspond to SSP5-8.5 and the lower values correspond to SSP1-1.9. Since SSP2-4.5 is used as the reference, which is closer to the SSP1-1.9 than SSP5-8.5, this can be expected. For the CH_4 reduction and the PMO effect, the lower values correspond to the SSP5-8.5 scenario. This is because a larger background concentration of CH_4 results in a larger influence of each NO_x emission, which in turn results with a larger negative radiative forcing.

The GWP generally improves upon the dependency shown by the GTP for both fuel scenarios. The ATR-rel shows a further improvement on an individual emission species level, but nevertheless still has a substantial dependence in the sustained and increasing emission scenarios. The GWP* and EGWP* show very similar results and at the individual emission species level show the lowest dependency in the sustained emission scenario. In total emissions, however, the ATR-rel shows lower dependency. The ATR-rel, GWP* and EGWP* thus score well for dependence on the background emissions scenario, whereas the RF and GTP do not.

4.2. Sensitivity of Climate Metrics to Changes in Emissions and Flight Conditions

The aim of this section is to identify which variables have the largest difference on the climate metric values to aid the development of the Monte Carlo simulation required for the multivariate fleet analysis. In this section, all climate metrics are calculated with a 100-year time horizon because it is commonly used and is long enough to incorporate the full fleet. To aid the understanding of the results shown in this section, Figure 4.14 shows the radiative forcing and temperature responses of the reference fleet (see Table 3.3).
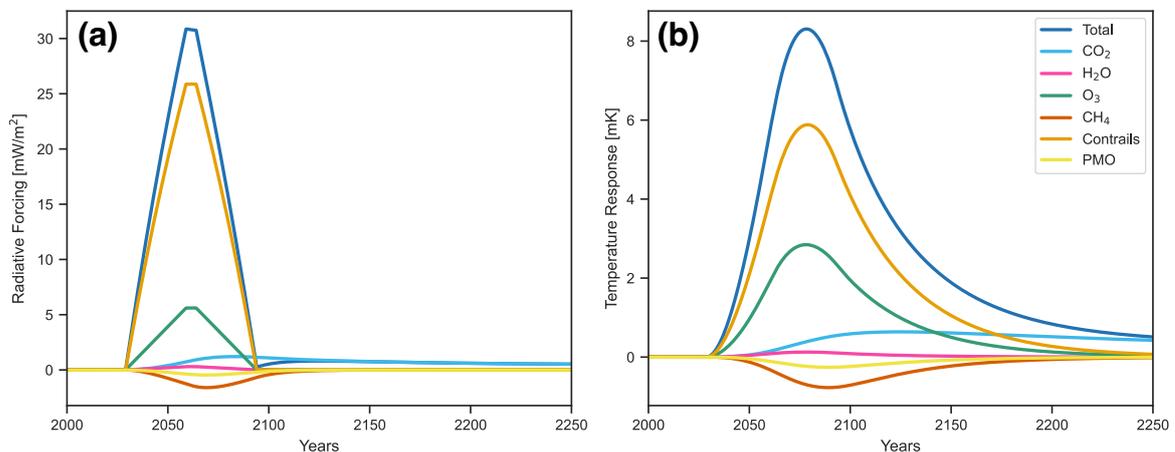
Figure 4.14: Responses of **(a)** radiative forcing and **(b)** temperature (right) of the reference fleet described in Table 3.3 for the background emission scenario SSP2-4.5.

The sensitivity of the climate metrics to changes in fuel burn and CO_2 emission are shown in Figures 4.15a) and b) respectively. As expected given the long atmospheric lifetime of CO_2, the climate metrics show a linear trend in both cases. The gradient, however, differs

significantly between the different climate metrics.
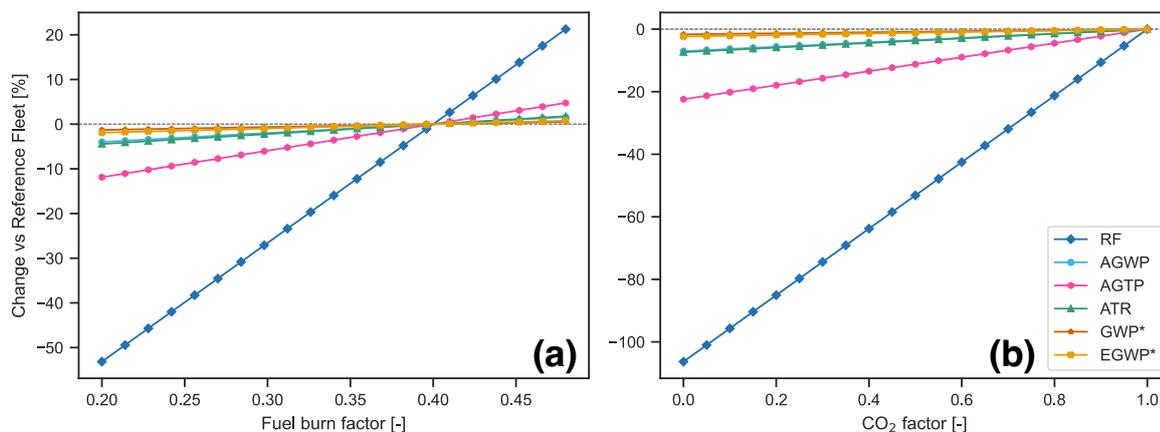
Figure 4.15: Sensitivity of climate metrics to **(a)** fuel burn and **(b)** CO_2 emission factors. The reference fleet is established in Table 3.3.

The radiative forcing shows a large sensitivity to both the fuel burn and the CO_2 emission factors. The reason for this is evident when the radiative forcing response in Figure 4.14a) is analysed. The radiative forcing is used in this analysis as an endpoint climate metric, with the climate metric value taken to be that at a time horizon of H = 100 years after the introduction of the fleet. Figure 4.14a) shows that in the year 2130 the CO_2 response dominates, with only a minor negative contribution from long-term CH_4 reduction and the PMO effect. Since the radiative forcing itself does not have a memory of previous emissions, the effects of the other emission species, notably contrails and ozone, do not remain. Therefore, the sensitivity to the fuel burn and CO_2 emission factors can be expected to be high.

In the case that no CO_2 is emitted, the negative contribution from CH_4 and PMO results in a more than 100% reduction compared to the reference fleet. At face value, this would suggest that an aircraft flying without CO_2 emissions, as would be the case for hydrogen-powered aircraft, would cause a net cooling of the climate. This is, however, misleading, as Figure 4.14b) demonstrates: the temperature caused by contrails and ozone remains. It is thus clear that the radiative forcing is not applicable for market-based measures, for example, because airlines could be owed money for these flights.

The other climate metrics show only a small sensitivity to the fuel burn and the CO_2 emission factors. For the endpoint climate metric AGTP this is due to the remaining temperature from emission species in the year 2130 as discussed before. For the AGWP, the integral of the contrail and ozone radiative forcing over the 100 year time horizon continues to dominate the integral of CO_2. As described in Section 3.4.1, in this research the peak GWP* and EGWP* values are used for which CO_2 plays only a minor role compared to contrails and ozone, as Figure 3.3 demonstrates. As a result, the GWP* and EGWP* have a very low sensitivity to fuel burn and CO_2.

Figure 4.16a) shows the sensitivity of climate metrics to the NO_x emission factor. The radiative forcing has a negative gradient, demonstrating a lower value for higher NO_x emissions. As before, this is due to nature of the radiative forcing as an endpoint climate metric and the fact that the atmospheric lifetime of ozone is of the order of one year. The positive radiative

forcing of ozone is thus no longer present at the time horizon, whereas the negative radiative forcing from CH_4 reduction and the PMO effect are. Performing the temporal integral or using the temperature response, however, produce positive gradients. The GWP* method also produces a positive gradient, recognisable due to the large influence of ozone in the peak CO_2-eq value in Figure 3.3. The EGWP* shows the largest sensitivity, achieving a 27% reduction in climate metric with no NO_x. The NO_x factor is thus an important variable in the multivariate fleet analysis.
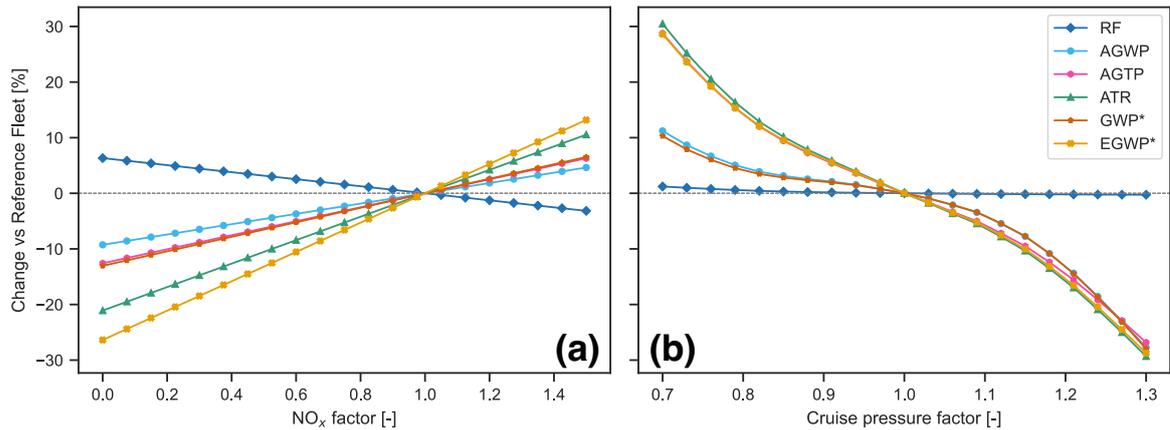
Figure 4.16: Sensitivity of climate metrics to **(a)** NO_x emission and **(b)** cruise pressure factors. The reference fleet is established in Table 3.3.

The sensitivity of the climate metrics to the cruise pressure factor is shown in Figure 4.16b). It should be noted that a shift to higher cruise altitudes generally also comes with a reduction in fuel burn due to the lower pressure and thus drag (cf. Matthes et al., 2021), however, this is not implemented here. It is likely this difference that explains the difference between these non-linear results to the linear results obtained by (Matthes et al., 2021, , see their Figure 8). When comparing the figures, it should also be noted that a lower cruise pressure factor is equivalent to a higher cruise altitude and vice versa.

The climate metrics show a non-linear sensitivity to the cruise pressure. A clear difference can be seen at higher altitudes between climate metrics using radiative forcing as their base, thereby not including the efficacy (AGWP and GWP*), and those using temperature or effective radiative forcing (ATR, AGTP, EGWP*). At lower altitudes, the effect of the efficacy is not as prevalent as at higher altitudes. The radiative forcing shows a very low sensitivity, as before due to only the effects of CO_2, CH_4 and PMO being present at the time horizon. The small increase seen at higher altitudes is due to the release of NO_x at higher altitudes. The cruise pressure factor is also an important variable in the multivariate fleet analysis.

The largest sensitivity for climate metrics other than the radiative forcing is due to the contrail distance factor, as demonstrated by Figure 4.17a). The reason for this is evident when analysing the radiative forcing and temperature profiles in Figure 4.14. In contrast, the sensitivity of climate metrics to changes in water vapour emissions is very low, as shown by Figure 4.17b). In both cases, for RF, the short atmospheric lifetimes of both emission species result in no sensitivity for the same reasons as described previously. Also noticeable is that the ATR and EGWP* show very similar responses, as do the AGWP and GWP*.
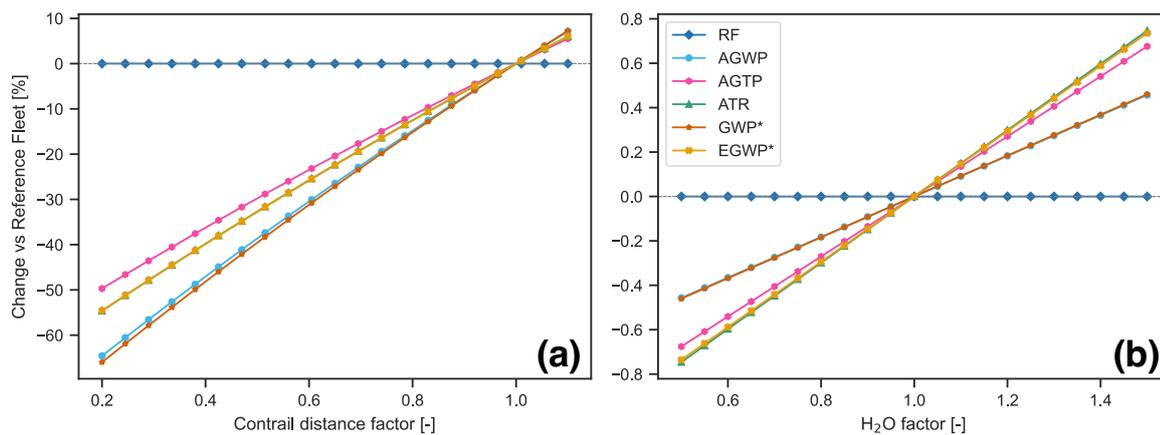
Figure 4.17: Sensitivity of climate metrics to **(a)** contrail distance and **(b)** H_2O emission factors. The reference fleet is established in Table 3.3.

The influence of the background emissions scenario is shown in Figure 4.18 and can be compared to the results of the general response of climate metrics analysis in Table 4.2 and the corresponding figures. The two endpoint climate metrics, but specifically the RF, demonstrate very high sensitivities to the background emissions scenario. The reason for this follows from the results of Section 4.1: Endpoint climate metrics only consider the value at one time horizon in the future, where the background emissions scenarios can differ substantially, whereas integrated or averaged climate metrics consider all years leading up to the endpoint and are thus generally less sensitive. The AGWP and ATR show similar sensitivity, within $\pm$ 5% of SSP2-4.5. In comparison with the results in Section 4.1, however, the GWP* and EGWP* show a notably smaller sensitivity, within $\pm$ 1% of SSP2-4.5.
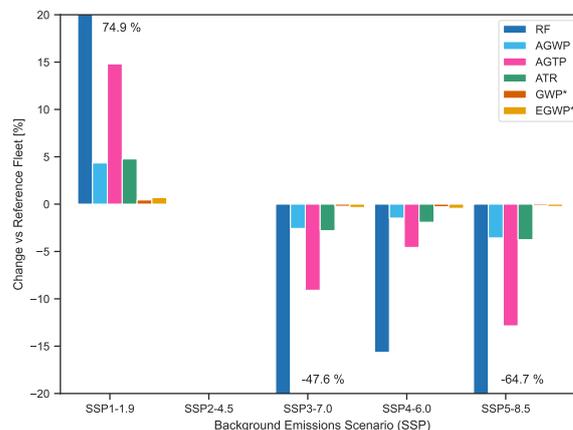
Figure 4.18: Sensitivity of climate metrics to the background emission scenario. SSP2-4.5 is used in the reference fleet (see Table 3.3) and thus shows no change for all climate metrics.

The sensitivity to the year of fleet introduction is shown in Figure 4.19. This sensitivity is caused by the trajectory of background CO_2 and CH_4 concentrations in the background emissions scenario and is low for all climate metrics. For the AGWP, AGTP and ATR, the year

2030 is a minimum value, likely since this corresponds with the peak in global methane concentrations in SSP2-4.5 (see Figure 3.1). In the SSP2-4.5 scenario, the rate of CO_2 emissions also starts to decrease. As with previous parameters, the RF is generally not affected by the change in global methane concentrations due to the 100 year time horizon. Its response is likely only due to the change in CO_2 concentration and thus does not show a minimum.
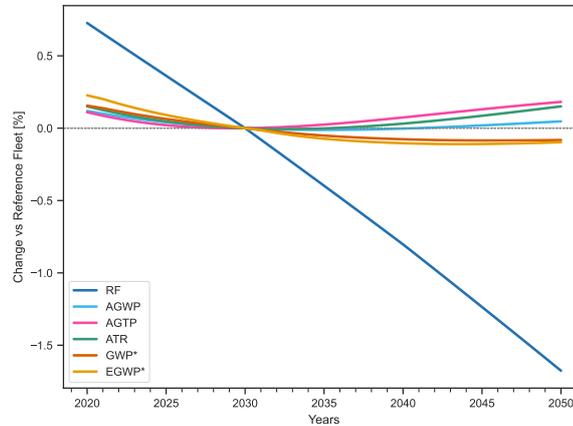
Figure 4.19: Sensitivity of climate metrics to the year of fleet introduction. The reference fleet is shown in Table 3.3.

4.3. Multivariate Fleet Analysis

This section presents the results of Monte Carlo simulation in detail. The linear regression for a single reference fleet is described in Section 4.3.1, followed by the pairwise analysis for all fleets in Section 4.3.2. Here, the inherent biases built into climate metrics, the basis for REQ 5, are investigated. Finally, in Section 4.3.3, the errors identified in the pairwise analysis are quantified to provide a better understanding of which parameters should be monitored carefully when making use of climate metrics.

The temperature profiles of all fleets are shown in Figure 4.20a) along with the reference, low and high fleets as established in Table 3.3. The influence of the year of fleet introduction can be seen, however the general shape of the response remains the same for all fleets. To see the influence of the variables to do with efficiency and trajectory changes, Figure 4.20b) shows the temperature profiles of only the fleets powered by Jet-A1. This figure shows that a substantial reduction in temperature can be achieved even without changing the fuel used.

4.3.1. Linear Regression of Fleets compared to Single Reference Fleet

The comparison of all fleets with the reference fleet for the peak temperature is shown in Figure 4.21. As described in Section 3.6, an ideal climate metric would show a linear trend since then the sign of the temperature change would match the sign of the climate metric change between each fleet. As a result, the climate metric would not demonstrate any inherent biases towards certain emission species or effects. A linear regression is performed and plotted for each climate metric, shown by the dashed line, and the gradient b is given in the legend. It is important to note that the gradient is given as a function of the percentage change in the climate metric value and not as a function of the absolute climate metric value itself, which
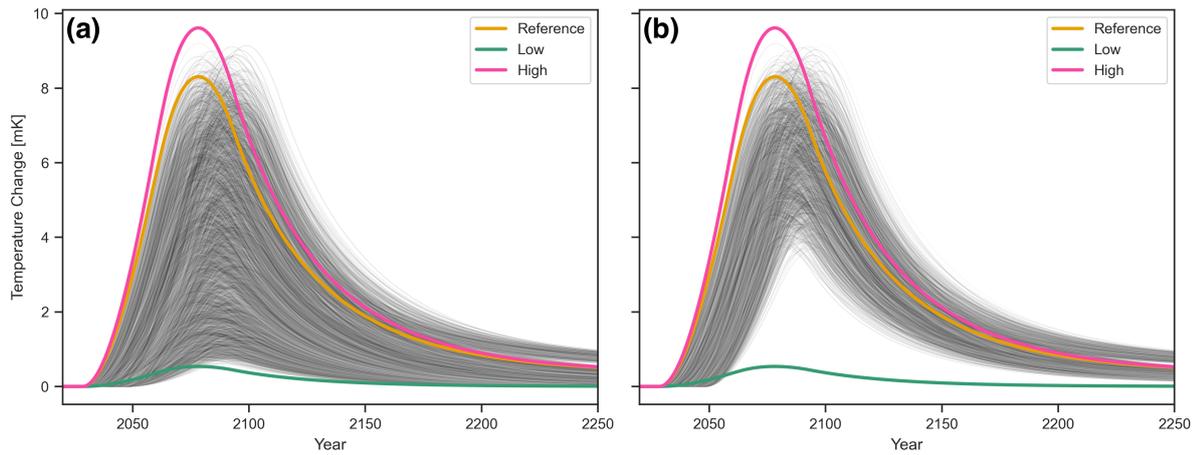
Figure 4.20: Temperature profiles of **(a)** all fleets and **(b)** only fleets using Jet-A1 in the Monte Carlo simulation, as defined by Table 3.3. The reference fleet, approximately equivalent to the Airbus A320, as well as the high and low references are also shown.

allows for direct comparison.
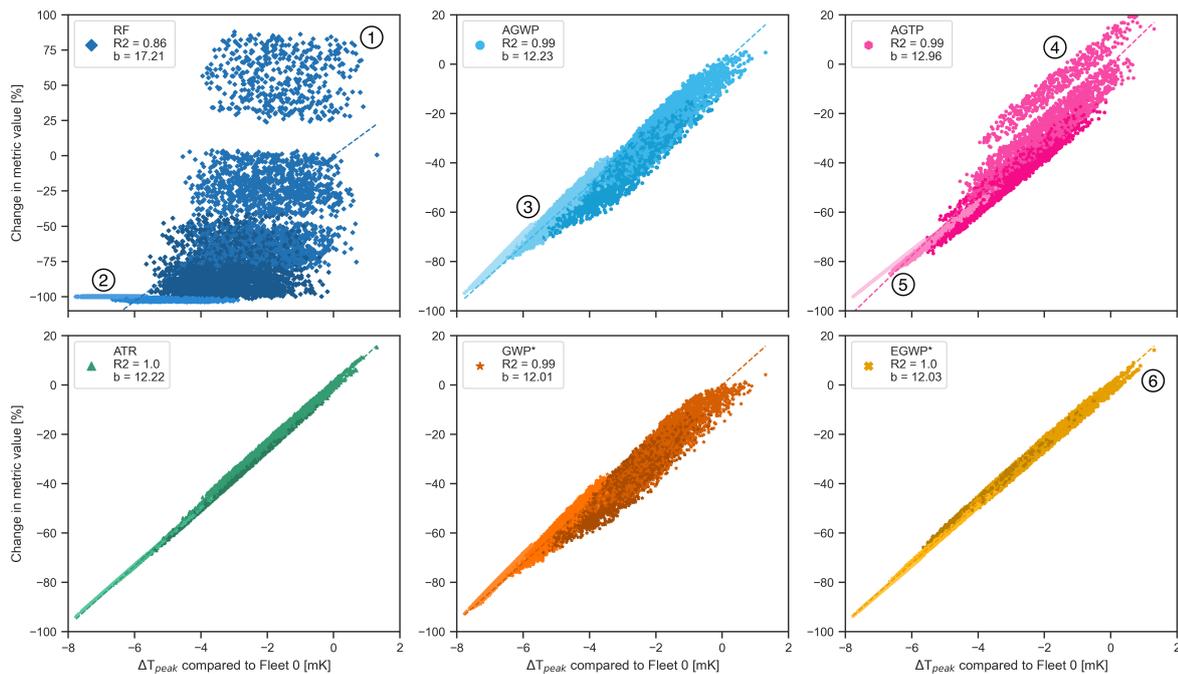
Figure 4.21: Comparison of all fleets with the reference fleet (Fleet 0) for the peak temperature. Note the different vertical axis limits for the radiative forcing results. Within each plot, the colours in ascending brightness show the fleets powered by SAF, Jet-A1, H_2 combustion and H_2 fuel cell respectively. The numbering corresponds to effects described in the text.

It is immediately clear that a number of variables used to create the fleets have a large effect on the radiative forcing results. Two main effects are identified here since they will be useful in the analysis of subsequent sections.

First, the cloud of fleets showing higher RF than the reference for a lower peak temperature is identified as consisting of kerosene-powered designs within the SSP1-1.9 scenario $\textcircled{1}$. A closer analysis of the RF results reveals five clusters of fleets, each corresponding to different background emissions scenarios. However, the difference between SSP1-1.9 and SSP2-4.5 is large enough to cause a clear separation. As is demonstrated by Figure 4.18, the RF is very sensitive to the background emissions scenario and the difference between the SSP1-1.9 and SSP2-4.5 scenarios is around 75% for the reference fleet.

The second effect is the largely horizontal cluster of fleets in the bottom-left of the plot $\textcircled{2}$. The cluster consists of two linear sections, the left-most section corresponding to hydrogen fuel cell fleets and the other to hydrogen combustion. The reason for this clustering around a constant RF value is due to the lack of CO_2 emissions by these fleets. Using Figure 4.14a) as a reference, it is clear that 100 years after the fleet is introduced, the CO_2 response has the largest effect on the RF value. Without any CO_2 emissions, the only effect left is the smaller long-term CH_4 reduction. As identified in the sensitivity analysis, the RF reduces with increasing NO_x emissions due to this long-term reduction (see Figure 4.16 and accompanying text). As a result, the RF value for all hydrogen fuel cell fleets, which do not produce any NO_x emissions, is 0, whereas for hydrogen combustion fleets, which do emit NO_x, the value is slightly negative. This effect is visible in Figure 4.21 by the second-brightest colours in the upper left plot.

The AGWP and GWP* show similar results. Both have a clear linear trend with a high R^2 value as well as an increasing spread with increasing peak temperature. This is primarily the result of the cruise altitude: fleets flying a higher altitudes, i.e. have lower cruise pressure factors, are clustered to the right edge of the trend. This indicates either a higher peak temperature for the same climate metric value, a lower climate metric value for the same peak temperature, or a combination of both. The contrail distance factor has a similar effect and is responsible for creating a spread in the values. The hydrogen fleets, which, like the RF, are clustered in two rows, are responsible for the two gradient changes seen on the left edge of the overall trend $\textcircled{3}$. The spread for the hydrogen fleets is lower since the contrail distance factor varies less.

The AGTP has a clear separation of a cluster of points, which, like the RF, are identified as kerosene-powered fleets within the SSP1-1.9 scenario $\textcircled{4}$. The spread is higher than for the AGWP and GWP* and primarily dependent on the background emissions scenario, rather than on the contrail distance and cruise pressure factors. This is due to the high dependence of the AGTP on the background emissions scenario, identified in the sensitivity analysis (see Figure 4.18). An interesting observation for the AGTP results is that the hydrogen fleets, in the same general location as the AGWP and GWP*, have a different gradient to the rest of the results $\textcircled{5}$. The reason for this could not be established.

The ATR and EGWP* show promising results. Both have a lower spread to the other climate metrics and have a well-defined linear trend. The spread is primarily attributed to the difference in background emissions scenario. However, it is important to note that for the ATR and the other absolute climate metrics, the fleets corresponding to SSP1-1.9 are on the left edge, whereas for the EGWP* and GWP* they are on the right edge and are responsible for the right "prong" $\textcircled{6}$ in the EGWP* results at higher peak temperatures. As discussed in Section 4.1, this is because absolute and relative climate metrics have opposite trends with respect to the background emissions scenario.

4.3.2. Pairwise Analysis of all Fleets

The pairwise analysis of the peak temperature is shown in Figure 4.22. Since each fleet is compared to each other fleet, the results are symmetric abound the origin. This section visually analyses the response of each climate metric; the error is quantified in the following section. The analyses of the average temperature are shown in Appendix A.
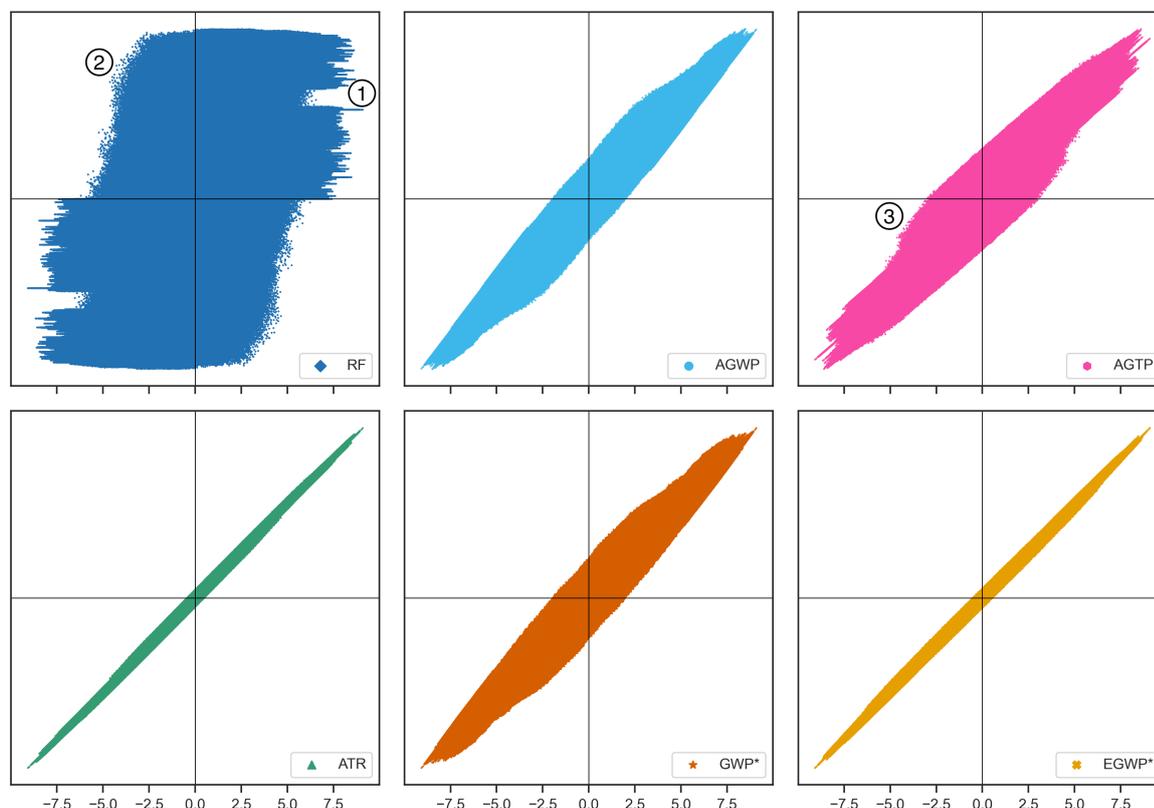
Figure 4.22: Peak temperature pairwise analysis of all fleets developed with the Monte Carlo analysis. The vertical axis is the change in absolute climate metric value, with zero in the middle. The numbering corresponds to effects described in the text.

The radiative forcing shows a very large spread of values, demonstrating a large bias towards certain technologies. Interesting to note are the horizontal lines visible in the top-right and bottom-left quadrants, which are caused by two main effects. The first is the low sensitivity of the RF to changes in the contrail distance factor, cruise pressure factor and, to a lesser extent, the water vapour factor (see Figures 4.17 and 4.16). The low sensitivity, or in the case of the contrail distance and water vapour factors lack of sensitivity, results in fleets obtaining the same RF value if only these variables are changed, regardless of the change in peak or average temperature. The second effect can be seen in the linear regression in Figure 4.21. For hydrogen-powered fleets, both using combustion and fuel cells, the radiative forcing shows largely a constant RF value. Analysing the sensitivity response of the CO_2 factor in Figure 4.15, it is likely that the high sensitivity of the RF to CO_2 emissions is much higher than the effect of any other variable. These effects make the RF far from ideal as a climate metric for aircraft design optimisation per REQ 5.

Other interesting artefacts seen in the RF results include an apparent gap in the top-right and bottom-left quadrants ①. The origin for this gap is determined to the large gap between results obtained from the background emissions scenarios SSP1-1.9 and SSP2-4.5, as is demonstrated in Figure 4.21 and explained in the accompanying text. The less well defined edges in the top-left and bottom-right quadrants ② are the result of comparison between fleets using conventional kerosene, of which there are only 2448.

The AGWP and GWP* show very similar results, both showing a clear linear trend with approximately the same gradient. Looking at the sensitivity analysis, this is likely because for a number of high-sensitivity parameters, such as the cruise pressure and contrail distance factor, the response from both climate metrics is very similar. For other factors, the difference between the responses is also not high. The overall shape remains the same as for the analysis against a single reference fleet, as in Figure 4.21. As is quantified numerically in the next section, the GWP* results are slightly more condensed and have slightly fewer errors.

Like the AGWP and GWP*, overall the AGTP shows a linear trend, but has a larger spread and a lower gradient. The separation of the kerosene powered fleets in the SSP1-1.9 scenario as described in the previous section are also visible in the AGTP results as a large increase in the spread near the origin ③. The contrail distance factor also increases the spread, shifting climate metric values higher for lower contrail distance and the same temperature change.

The ATR and EGWP* have the lowest spread and thus show the greatest promise for REQ 5. The difference between the GWP* and EGWP* is noticeable and demonstrates the importance of including the efficacy, even in a simplified manner. The EGWP* also performs well for the average temperature, which is unexpected considering that the peak value is used for the climate metric. Looking at the temperature response of a single fleet (Figure 4.14), it is likely that this is due to the average temperature response being similar until over 100 years after the fleet introduction. For example, the CO_2 temperature response is higher than the ozone response only around 110 years after fleet introduction. For a fleet as used in this work, the peak temperature is thus also a good indicator of the average temperature at least until a time horizon of 100 years.

4.3.3. Error Analysis

The results from the pairwise analysis can be quantified by calculating the number of pairings for which the climate metric produces the incorrect sign. Graphically, this corresponds to the points in the top-left and bottom-right quadrants in Figure 4.22. The frequency with which this occurs for the peak and average temperatures is plotted against the percentage distance between the climate metric values of each fleet in Figure 4.23. Each plotted value corresponds to the value within the 1% bracket.

From Figure 4.23, it can be concluded that the ATR and EGWP* generally perform best and the RF performs worst. Even for a 100% difference between the fleets, the RF provides an incorrect climate metric value compared to the temperature 35 - 40% of the time, depending on whether the peak or average temperature is used. In comparison, the ATR and EGWP* both always show the correct pairing above a difference of 20%.

The AGWP and GWP* show very similar results in all cases. In comparison, the AGTP generally takes longer to reach 0, but at lower differences between fleets is approximately equivalent to the AGWP and GWP*. The EGWP* performs best for the 20-year average temperature and the ATR, unsurprisingly, performs best for the 100-year average temperature. In
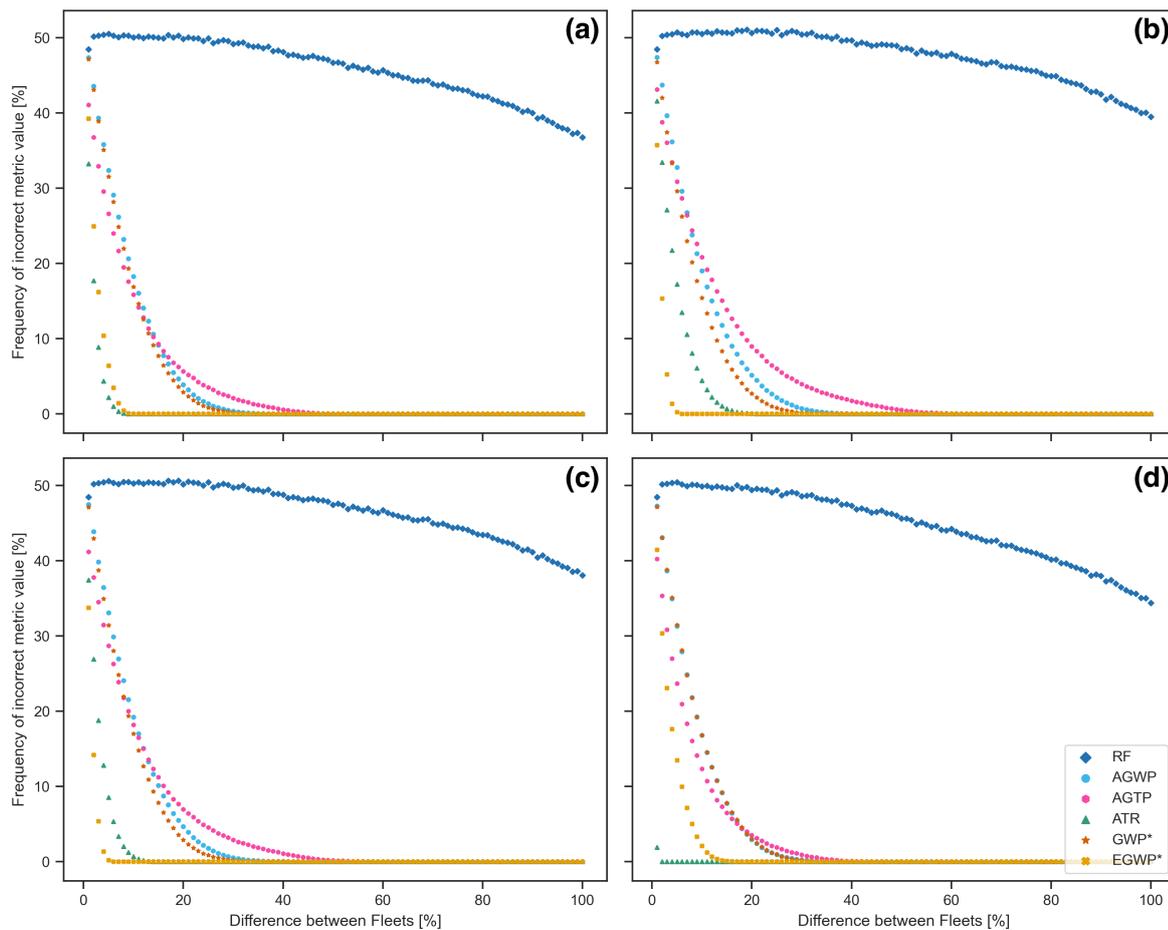
Figure 4.23: Frequency of an incorrect climate metric value as a function of the percentage difference between each fleet pair for **(a)** the peak temperature, **(b)** the 20-year average temperature, **(c)** the 50-year average temperature and **(d)** the 100-year average temperature.

general, the results do not differ substantially between the different temperature objectives.

4.4. Temporal Trajectories of CO_2-eq using Climate Metrics

The final analysis method for climate metrics concerns the development of CO_2-eq trajectories and how these can be used to estimate the temperature response. The calculation of CO_2-eq emissions are important since they are often used to compare industries to one another and to ensure that emissions are being reduced. Section 4.4.1 analyses CO_2-eq emissions for the fuel scenarios, which would be used by a stakeholder to track emissions over time. Section 4.4.2 then analyses the difference between the temperature calculated using AirClim and that estimated using climate metrics for the fleets in the Monte Carlo simulation.

4.4.1. CO_2-Equivalent Responses of Emission Profiles using Climate Metrics

The responses of the CurTec and FP2050 scenarios are shown and described in this section. The results of the other scenarios (CORSIA, COVID-15s, Fa1) are shown in Appendix A but

are not included here since the scenarios shown here can be used to explain all effects.

The response of the CurTec scenario is shown in Figure 4.24a) with the fuel use and temperature shown in b) for reference. Two elements of the response are highlighted here. First, the difference between the total values calculated using the GWP and ATR-rel are very similar, potentially reducing the political capital required to change from the standard GWP_{100} to the ATR_{100}-rel. It is interesting to note that the total values are similar, although the values for individual species may not be. This effect can be seen especially for the responses of ozone, contrails and methane reduction. A similar observation can be made with the endpoint RF and GTP climate metric results.
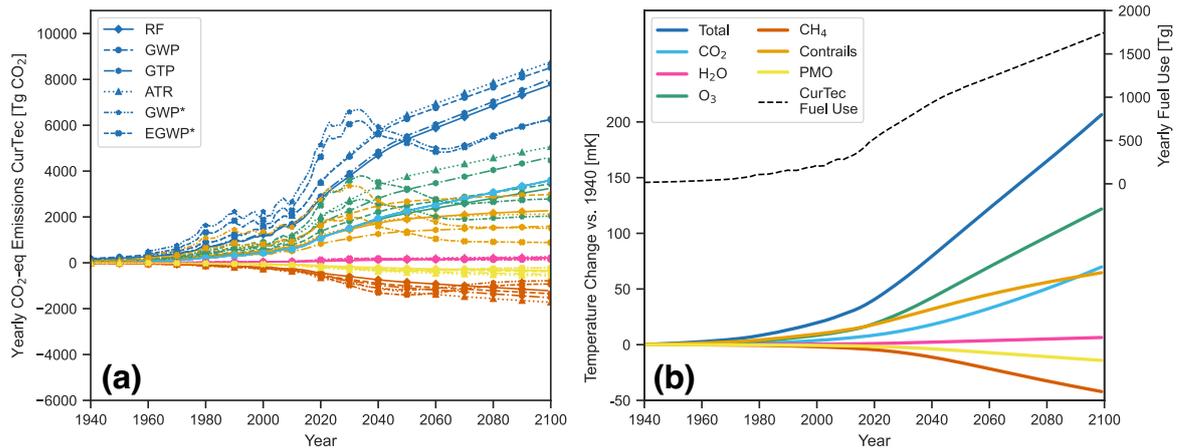
Figure 4.24: CurTec fuel scenario: **(a)** CO_2-eq emissions calculated using climate metrics; **(b)** fuel use and temperature response from each emission species. The colours in the left figure correspond to the emission species in the right figure.

The second noticeable element is the rapid departure of the GWP* and EGWP* from the response shown by all other climate metrics around the year 2025. The CO_2-eq calculated using GWP* and EGWP* reach a peak, then start to reduce until the year 2060, whereupon they start to increase again. The reason for this is that the GWP* calculation method uses an average of the previous 20 years of radiative forcing to calculate its value. Looking closely at the fuel scenario, it is clear that the rate of emission increase reduces around the year 2020 and again around 2045. By mathematical definition of the GWP*, this change in rate results in a lower GWP* climate metric value.

The temperature response calculated using the climate metrics is compared to the response calculated with AirClim in Figure 4.25. The effect of the plateau and decrease in CO_2-eq emissions can be seen in the GWP* and EGWP* results by a noticeable decrease in the slope for total emissions, especially when compared to the other climate metrics. The benefit of the GWP* method can be seen in the estimation of the contrail response: The temperature response shows the beginning of asymptotic behaviour, which is picked up by the GWP* climate metrics but not by the others due to the reducing CO_2-eq values. Similar behaviour is also seen for the ozone response.

It is interesting to note that all climate metrics overestimate the temperature impact of CO_2, which shows the limitations of using climate metrics in this manner: How well the results match the temperature is primarily dependent on the value for the transient climate response (TCRE),
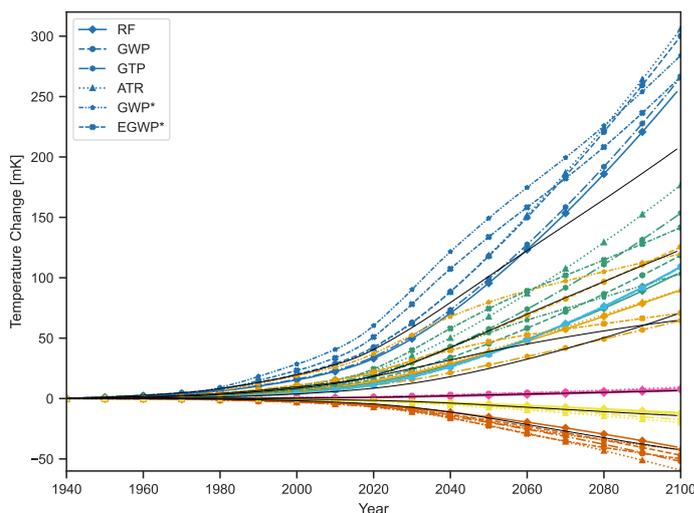
Figure 4.25: CurTec temperature response calculated using AirClim (thin black lines) compared to the response estimated using climate metrics. The colours correspond to the emission species in Figure 4.24.

defined as 0.49 K/TtCO$_2$ in this study (Cain et al., 2019). The trends are, however, promising from all climate metrics.

The FP2050 scenario, shown in Figure 4.26a), demonstrates the erratic nature of the GWP* method. Between the years of around 2050 and 2080, the GWP* and EGWP* show negative CO_2-eq values due to reducing emissions in the fuel scenario. The sign of the CO_2-eq emissions from all emission species except CO_2 is reversed. For the CH_4 reduction and PMO effect, the reversal lasts until after the year 2100 due to their longer atmospheric lifetimes. The negative values, when the CO_2-eq emissions are used to estimate the temperature change, correspond to the peak and slight drop in total temperature shown in Figure 4.26b).
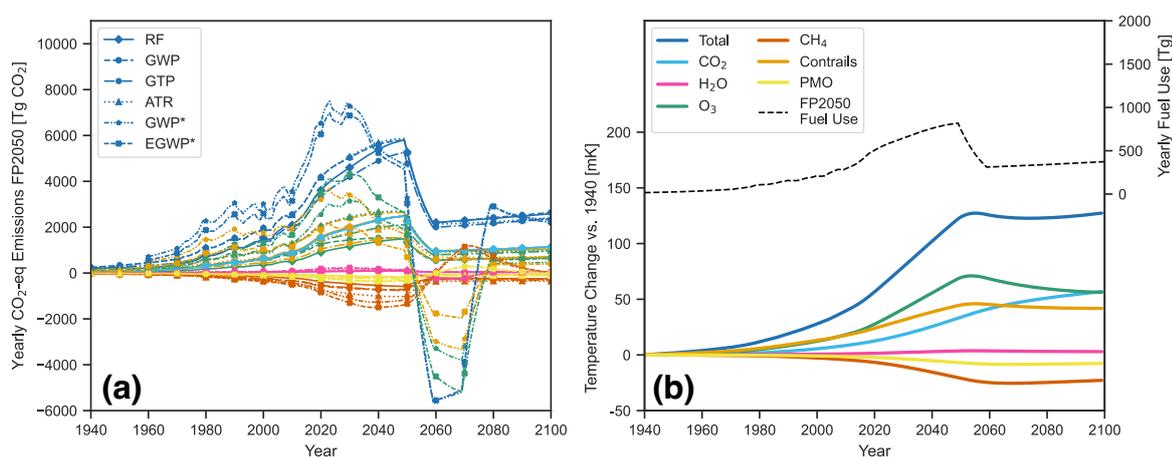
Figure 4.26: FP2050 fuel scenario: **(a)** CO_2-eq emissions calculated using climate metrics; **(b)** fuel use and temperature response from each emission species. The colours in the left figure correspond to the emission species in the right figure.

The other climate metrics show a clear reduction in the year 2050, but remain positive.

Estimating the temperature response by the CO_2-eq emissions from these climate metrics would, therefore, not show the peak and slight drop. This effect can be seen in Figure 4.27. Overall, however, the CO_2-eq emissions can be seen to be more stable than those calculated using the GWP*/EGWP*. The similarity of the GWP and ATR-rel, and the RF and GTP, is also recognisable for total emissions. For the individual species, the differences caused by the inclusion of the efficacy are nevertheless present.

Interesting to note from the temperature response calculated using the GWP* and EGWP* in Figure 4.27 is that the peak and subsequent reduction in temperature is overestimated. This originates primarily from the ozone and contrail responses, which also continue to rise rather than continue to reduce towards the end of the century. It is likely that a more optimal value of the stock term s in Eq. (2.6) would help alleviate this problem and that the value of 0.25 applies primarily to the CH_4 emissions used in the development of the GWP* (Cain et al., 2019; Smith et al., 2021). This is discussed further in Section 5.1, however, a full analysis is beyond the scope of this research.
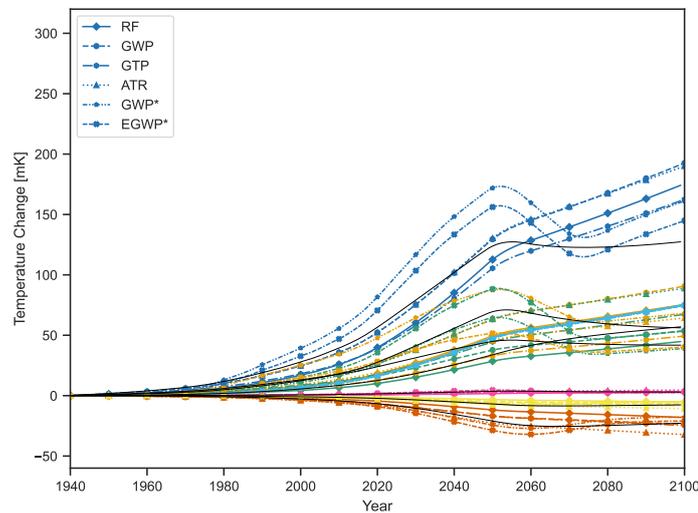
Figure 4.27: FP2050 temperature response calculated using AirClim (thin black lines) compared to the response estimated using climate metrics. The colours correspond to the emission species in Figure 4.26.

4.4.2. Temperature Error Analysis of Pairwise Fleets

The analysis in Section 4.3 compares the climate metric values for all fleets and aims to determine which climate metric most accurately shows a linear trend, thereby showing a low bias towards different emission species. Not considered in that analysis is whether the climate metric value itself has any relevance to the temperature estimation that can be done using the temporal response of each climate metric. This section uses the CO_2-eq emission trajectory to estimate the peak temperature using a climate metric, to compare this to the temperature calculated using AirClim.

Figure 4.28 shows the results of this analysis. Ideally, all fleets would be on the black, linear trend shown, since this would mean that the peak temperature estimated using the climate metric matches the peak temperature calculated using AirClim. At first glance, it is thus clear that the GWP* climate metrics, specifically the EGWP*, follow the desired trend.
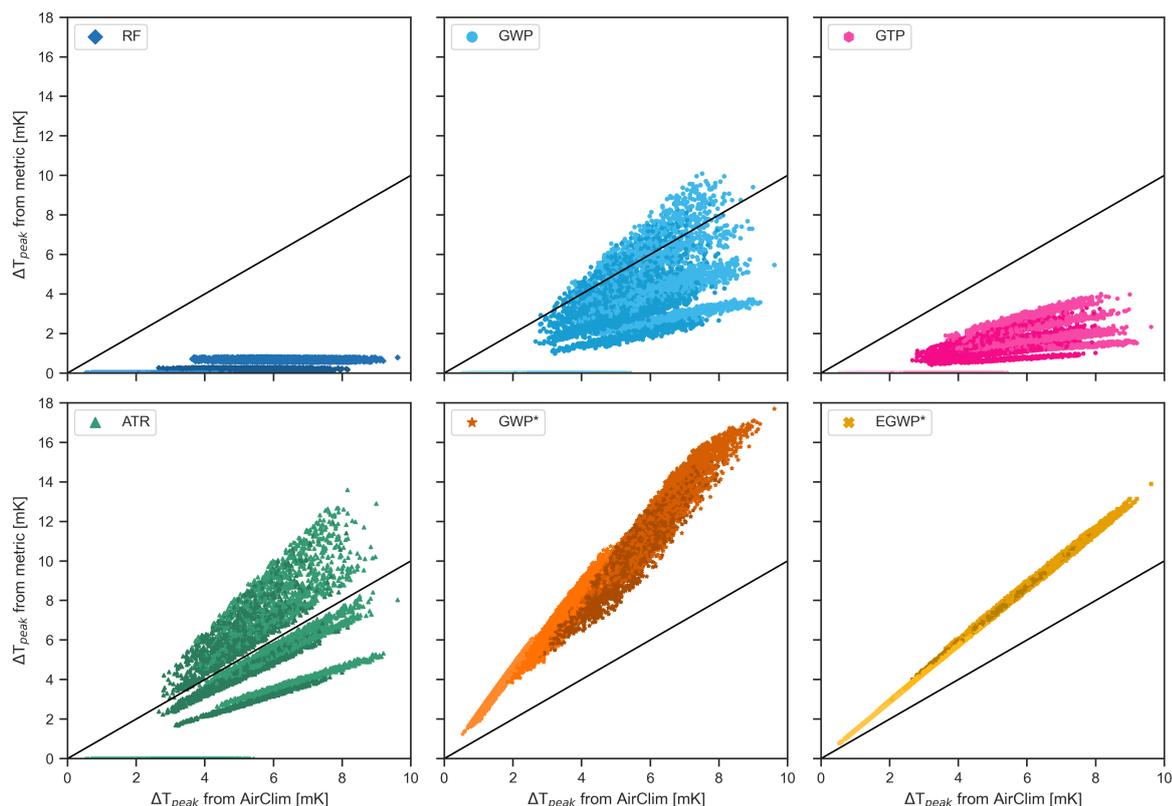
Figure 4.28: Change in peak temperature calculated with AirClim compared to the peak temperature estimated using climate metrics. Within each plot, the colours in ascending brightness show the fleets powered by SAF, Jet-A1, H_2 combustion and H_2 fuel cell respectively.

The conventionally calculated climate metrics GWP, GTP and ATR-rel show a large dependence on the background emissions scenario, which influences the slope. The lowest slope corresponds to SSP1-1.9, whereas the highest slope corresponds to SSP5-8.5. The conventional climate metrics also do not calculate a temperature change from hydrogen-powered fleets. This is due to the calculation method, which requires division by the absolute climate metric value of the fleet CO_2 emission (e.g. $AGWP_{CO_2}$), which is zero for hydrogen-powered fleets. A potential solution to this is to directly calculate the CO_2-eq using the absolute climate metric response divided by the response due to a pulse of CO_2. In this scenario, a decision would have to be made what pulse to use: a pulse that depends on the background emissions scenario and year of emission, or a standard pulse. In the latter scenario, accuracy could be lost.

The GWP* and EGWP* do not show a difference in slope due to the background emissions scenario. This is because the $AGWP_{CO_2}$ values for the SSP2-4.5 scenario are used throughout. If the $AGWP_{CO_2}$ trajectory corresponding to the scenario used by each fleet is used, then a slope change is visible, as Figure 4.29 demonstrates. The effect of using the efficacy is visible in the comparison of the responses: the EGWP* response shows much lower spread than the GWP* response.

Both figures demonstrate the large influence of the transient climate response (TCRE),
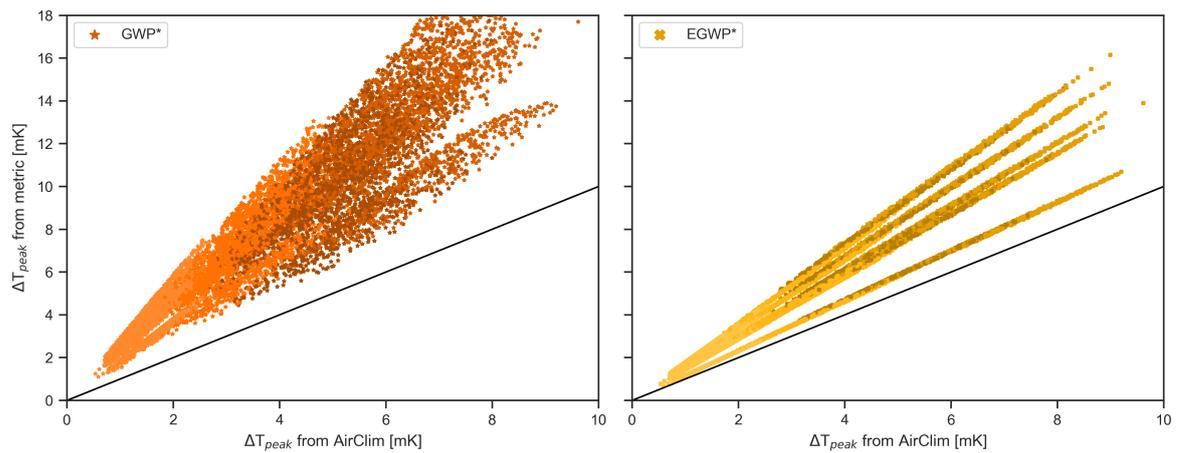
Figure 4.29: Change in peak temperature calculated within AirClim compared to the peak temperature estimated using the GWP* and EGWP*. The AGWP$_{CO_2}$ value corresponding to the background emissions scenario has been used in the calculation of the climate metric, resulting in the varying slopes. Within each plot, the colours in ascending brightness show the fleets powered by SAF, Jet-A1, H_2 combustion and H_2 fuel cell respectively.

as identified in the previous section. Each background emissions scenario is linked with a different transient climate response, which was not calculated in this research. The accuracy of the temperature estimation using climate metrics, therefore, highly depends on the accuracy of the TCRE. However, as the comparison of the GWP, ATR-rel and EGWP* responses show, the required TCRE to obtain the correct slope is not the same across climate metrics: the slope shown by the black line corresponds approximately to SSP4 for the GWP, SSP2 for the ATR-rel and a higher decarbonisation than SSP1 for the EGWP*. This shows the difficulty of using climate metrics for the purpose of temperature estimation.

Discussion

This chapter discusses the results presented in the previous chapter. An analysis of each requirement is given in Section 5.1, followed by a choice of climate metric for each use case. The uncertainties and assumptions are discussed in Section 5.2.

5.1. Choice of Best-Suited Climate Metric for Aviation

This section discusses the results and how well each climate metric meets the requirements set in Section 3.3. Table 5.1 provides a summary of the discussion, which is also shown visually in Figure 5.1. The table and figure suggest that the ATR and EGWP* are contenders for best-suited climate metric. In the following, the reasoning for the shown values is shown and a best-suited climate metric is chosen for each use case. A detailed analysis of the GWP* method is also provided.
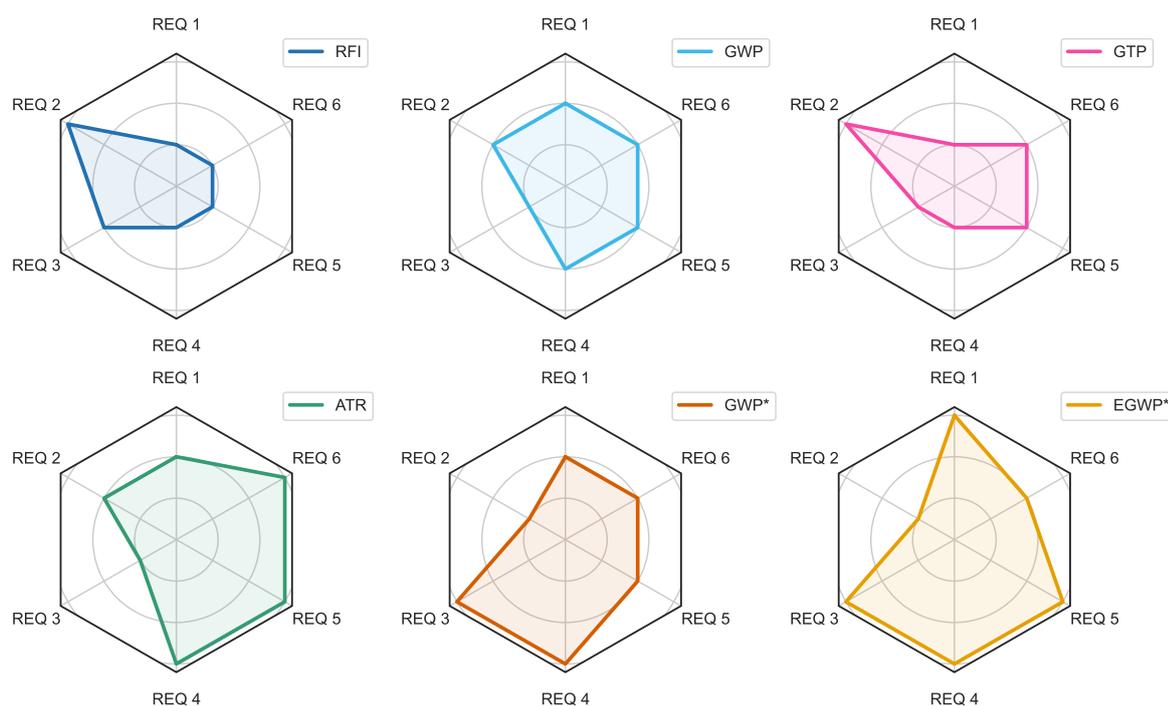
Figure 5.1: Visualisation of the summary table. The better a climate metric performs for a given requirement, the further outwards its value is. Note that this is a simplification of the results and that some values depend on the emission species or effect in question. More detail about how the climate metrics perform for each requirement is found in this section.

Table 5.1: Summary of the behaviour of climate metrics with respect to the requirements. Note that this is a simplification of the results and that some values depend on the emission species or effect in question. More detail can be found in this section.

Requirement		RFI	GWP	GTP	ATR	GWP*	EGWP*
REQ 1	Representation of the temperature change	Low spread, low accuracy	Large spread, medium accuracy	Low spread, low accuracy	Large spread, medium accuracy	Medium spread, medium accuracy	Low spread, high accuracy
REQ 2	Ease of understanding and implementation	Low complexity	Less complex	Low complexity	Less complex	Complex	Complex
REQ 3	Dependence on the time horizon	Medium	High	High	High	Low	Low
REQ 4	Dependence on background emissions	High	Medium	High	Low	Low	Low
REQ 5	Inherent biases	High bias, high incorrect fleet pairings	Medium bias, medium incorrect fleet pairings	Medium bias, lower quality linear trend than AGWP/GWP*, medium incorrect fleet pairings	Low bias, low incorrect fleet pairings	Medium bias, medium incorrect fleet pairings	Low bias, low incorrect fleet pairings
REQ 6	Appropriate for different emission profiles	Shows differences in qualitative responses of pulse and sustained emissions	Might show qualitative differences between pulse and sustained emissions	Might show qualitative differences between pulse and sustained emissions	Shows qualitatively similar behaviour	Identical behaviour for pulse and sustained, potentially misleading	Identical behaviour for pulse and sustained, potentially misleading

REQ 1: Shall correctly represent the temperature change

The first requirement has a larger effect than just the estimation of the temperature using climate metrics. All analyses are done with respect to the temperature change, rather than, for example, the radiative forcing. Therefore, this requirement informs the results of the other requirements as well.

The value of the transient climate response (TCRE) has a large impact on the results for all climate metrics since it defines how much warming each CO_2-eq emission produces. Whether a climate metric over- or underestimates the temperature should thus always be analysed in the context of comparison with other climate metrics. The GWP tends to underestimate the temperature of fleet emissions using the standard TCRE. The ATR has a larger spread than the GWP, but generally seems to match the temperature more accurately. This could be expected since the ATR makes use of the temperature response in its calculation. For the full aviation emissions scenario, however, both climate metrics tend to overestimate the total temperature, likely since conventional climate metrics cannot show decreasing temperature or peaks, as described below.

Endpoint climate metrics, in this research the RF and GTP, have difficulty estimating the temperature change for short emission profiles such as the pulse and fleet emissions used in this work. This is because endpoint climate metrics have no memory of previous emissions and only consider the impact at the time horizon. For longer-term profiles such as the full aviation emissions scenarios, these climate metrics are nevertheless capable of estimating CO_2-eq emissions and thus temperature change.

A note should be made here about the calculation method of conventional climate metrics, in this research the RF, GWP, GTP and ATR. As described in Section 2.4.2, the original definition of the GWP, for example, compared the response of an emission species to a pulse of CO_2. In this research, the conventional climate metrics are used to compare the response of an emission species to the full response of CO_2. Since CO_2 has a very long atmospheric lifetime, over the time horizons being analysed in this research, the quantity of CO_2 and its effect on the radiative forcing and temperature can be approximated as simply being linear - i.e. the full CO_2 response is equivalent to the sum of CO_2 pulses. The advantage of this method is that the relative climate metric value directly provides a multiplier for the conversion

between CO_2 and CO_2-eq emissions.

This method does not, however, work for fleets with reduced CO_2 emissions such as those powered by SAF and, especially, H_2. Since the relative climate metric values have the CO_2 response in the denominator, decreasing CO_2 emissions result in higher climate metric values, with the limit case of H_2 fleets being an infinite value. In this research, these fleets are shown to have 0 K temperature change, but technically they are undefined. Since it is not possible to compare contrail emissions per kg, i.e. to calculate the response of 1 kg of an emission to a 1 kg emission of CO_2, comparing the mass of released emissions is not a viable alternative. Instead, it is suggested that the conventional climate metrics should be used as the absolute climate metric value of an emission species compared to a pulse of CO_2. This does, however, require a choice to be made for the CO_2 response: either, as has been done in this research, the CO_2 pulse emission takes place in the same year as the non-CO_2 emission, or a standard CO_2 pulse emission should be defined. The latter option is easier to implement, but the former upholds the dependency on the background emissions scenario.

REQ 2: Shall be easy to understand and implement

Endpoint climate metrics are the easiest to understand: GTP_{100} and RF_{100} are simply the temperature or radiative forcing 100 years in the future. As a result, it is straightforward to determine how these climate metrics behave for different time horizons, background emissions scenarios and fuel scenarios. Integrated climate metrics such as the GWP are more complex, and it is more difficult to ascertain the dependence of various effects and emission species. A similar argument can be made for averaged climate metrics such as the ATR.

The starred climate metrics are the most difficult to understand. To comprehend and analyse the results obtained using the GWP*, a stakeholder must first have researched and understood the equivalency between a change in emission rate of SLCPs and a pulse of CO_2. Even the responses to simple emission profiles are difficult to understand and require more in-depth knowledge of the climate metric than should be expected of a policymaker or company decision-maker. For more complex fuel scenarios, or for the analysis of dependencies, the behaviour of the climate metric is hard to estimate in advance and can show at first counter-intuitive results (see, for example, the calculation of CO_2-eq emissions from full aviation fuel scenarios in Section 4.4.1). Stakeholders must also be aware of a number of assumptions that are used in the calculation method, as are discussed later in this chapter.

In terms of implementation, the climate metrics based on radiative forcing (RF, GWP) are easier to implement than temperature-based climate metrics (GTP, ATR) since they do not require a full climate or carbon model. The same argument can be made for the GWP*, although it does require the $AGWP_{CO_2}$ for the time horizon in question, which should either be a standard value or should be calculated for the specific scenario. Additionally, the GWP* requires the profile of emissions from twenty years prior since it runs on a 20-year average. On top of this, the EGWP* requires an estimation of the efficacies of the emission species. Implementation of the starred climate metrics is thus more complex.

REQ 3: Shall be largely independent of the time horizon

A high dependence on the time horizon is essentially a trade-off between long-lived and short-lived climate pollutants. Ideally, this trade-off is done by the stakeholder rather than it being built inherently into the climate metric, making a low dependence desirable. However, given

that long-lived and short-lived climate pollutants have fundamentally different impacts on the environment, full independence of a climate metric to the time horizon is not possible.

The general response analysis in Section 4.1 shows that it is not possible to mathematically determine an optimal time horizon and that all climate metrics depend on it. The choice of time horizon, therefore, remains subjective. The starred climate metrics generally show the lowest dependence on the time horizon between 20 and 100 years for total emissions, whereas the ATR shows the highest. This dependence will vary depending on the emission profile and the emission species or effect, however, so care must nevertheless be taken when analysing a certain profile.

A further point of consideration is that for non-endpoint climate metrics, the value of the time horizon can be misleading. For example, as Allen et al. (2016) show, the GWP_{100} effectively estimates the temperature change of SLCPs 20-40 years in the future. A similar analysis could be done for the ATR and other conventional, non-endpoint climate metrics. In other words, while for an endpoint climate metric, the conditions at the time horizon are being used, for integrated or averaged climate metrics it is often not possible to know a priori which point in time is being estimated. This further complicates the choice of time horizon.

REQ 4: Shall be largely independent of the background emissions scenario

The endpoint climate metrics RF and GTP perform poorly for this requirement because they only consider the change in radiative forcing or temperature change at a certain time in the future. The difference between the background emissions scenarios at a certain time in the future is larger than the integrated or average difference until that time, meaning that endpoint climate metrics will always show larger dependence as long as background emissions scenarios diverge. The ATR performs better than the GWP for pulse, sustained and increasing emissions, but the climate metrics give similar results for fleet scenarios. The GWP* and EGWP* perform well for the three emissions and show almost no dependence at all for the fleet scenarios. As shown in Section 4.4.2, specifically the difference between the responses in Figures 4.28 and 4.29, this is likely due to a single, reference $AGWP_{CO_2}$ for SSP2-4.5 being used.

It is important to note the differences between relative and absolute climate metrics and their use. When using absolute climate metrics, the value of an emission is higher for scenarios with high decarbonisation and lower when emissions are increasing quickly. An argument for why this might be desirable is that when other industries are decarbonising quickly, a higher climate metric value should incentivise the aviation industry to also decarbonise. However, from the results it is clear that relative climate metrics show the opposite effect: for scenarios with high decarbonisation, the value of an emission - if introduced into an emission trading scheme then also the price of an emission - is lower, and for scenarios with rapidly increasing emissions, the value is higher. It could be argued that this is also beneficial, in that it stimulates the industry to decarbonise when emissions overall are continuing to rise. However, the lower value for scenarios with high decarbonisation does not make sense.

The reason for the opposing trends likely lies in the relative importance of long-lived and short-lived climate pollutants. The importance of LLCPs, in aviation only CO_2, decreases for scenarios with rapidly increasing emissions such as SSP5-8.5, and increases for high decarbonisation scenarios such as SSP1-1.9. It is likely that the division of the absolute CO_2 climate metric value, which is more highly dependent on the background CO_2 emissions than

CH_4 reduction and PMO are on background CH_4 emissions (see Table 4.2), causes the trend to reverse.

It is clear that regardless of how it is argued, the trend reversal can lead to problematic conclusions. Stakeholders must be aware of this behaviour of climate metrics. A potential solution to the problem described above would be to use a consistent CO_2 response in all calculations. However, doing so would reduce accuracy since effectively that would remove dependence on the CO_2 background emissions scenario. An analysis of the environmental and economic ramifications of using a variable or fixed CO_2 response for example in a trading scheme is beyond the scope of this work, but is recommended for future research.

REQ 5: Shall have low inherent biases towards changes in aircraft design or trajectory

It is clear from the results of the multivariate fleet analysis that the RF with a 100 year time horizon has large biases towards different emission species and effects. The RF does not change linearly with the peak temperature and displays large clustering, owing to a lack of sensitivity to changes in various fleet parameters. The choice of the 100 year time horizon was made to enable direct comparison with the other climate metrics, however, it could be argued that a better choice would be the peak RF or the RF at an earlier time horizon. The potential issues with choosing a peak value is discussed in more detail later in this chapter in reference to the GWP*. A lower time horizon is possible, however the choice must nevertheless be properly argued.

The ATR and EGWP* clearly show the lowest bias of all climate metrics considered. These climate metrics have the lowest number of errors, regardless of the form of the temperature used for comparison (peak or average). The AGWP and GWP* show a sizeable spread due to the cruise altitude and contrail distance factor, but are nevertheless usable. The AGTP shows a somewhat larger spread due to the influence of the background emissions scenario. All climate metrics except the RF would be appropriate for use according to this requirement, provided their limitations are kept in mind.

REQ 6: Shall be appropriate for different emission profiles

In this study, the pulse, sustained and increasing emissions are compared. Since the sustained emission can be seen as a series of pulse emissions, in the case of AirClim one pulse per year, climate metrics should show at least qualitatively the same result for both emission profiles. The RF does not meet this requirement since it shows fundamentally different results for a NO_x pulse and sustained emission: for a pulse emission, the RF shows an initial positive peak, but negative values after the first year. In comparison, for a sustained emission, the RF shows continuous positive values.

The other climate metrics show qualitatively the same results. It is, however, possible for the AGWP to show negative values for a pulse emission if the initial radiative forcing is negative enough. Since contrails cause a large positive radiative forcing, it is unlikely to occur for total aviation emissions. With the introduction of hydrogen as a future fuel, however, this should not be ruled out. The AGTP is even more susceptible to changes in sign. Although the change in sign is not seen in this analysis, this has been identified by previous authors (e.g. Dahlmann, 2011).

The results of the GWP* and EGWP* are more complex and potentially confusing. Both climate metrics show the same amount of CO_2-eq for all three emissions. However, the sustained emission scenario results in more emissions than, for example, the pulse emission scenario and should, therefore, be responsible for more CO_2-eq. The reason for the identical value is due to the mathematical description of the climate metric. The GWP* uses the change in the 20-year running average to calculate a value and has no forward-looking time horizon. If there is no change in the previous 20 years, then the GWP* will not identify any CO_2-eq emissions. Furthermore, any changes in the emission profile in the years following the introduction of the emission are not considered. For the three emission scenarios, there are no emissions in the 20 years leading up to the introduction of the emission, thereby making the results equal. This is odd behaviour when one is used to the GWP or similar climate metrics.

Further discussion of the GWP*

Two elements of the GWP* method are discussed in more detail here. The first is that the GWP* can show reducing or even negative CO_2-eq emissions, as demonstrated by Figures 4.24 and 4.26. This effect is problematic. For the CurTec scenario, if the GWP* were to be used to account for all CO_2-eq emissions from aviation on a yearly basis, then a stakeholder looking only at the values between the years of 2025 and 2060 could be forgiven for assuming that emissions from aviation were reducing. However, it is clear that this is not actually the case: emissions continue to rise, just at a lower rate. This is compounded by the fact that stakeholders are likely used to the response of conventional climate metrics such as the GWP, which do not behave in a similar manner. For the FP2050 scenario (see Figure 4.26), a stakeholder could come to the conclusion that aviation is causing a cooling effect on the climate due to the negative CO_2-eq emissions for the same reason.

The figures mentioned provide a visual description of the discussion provided by Lee et al. (2021). They mention that the GWP* multiplier, i.e. how much total aviation emissions are warming the climate compared to how much only CO_2 emissions are, should not be applied "to future scenarios that deviate substantially from the current trend of increasing aviation-related emissions". The results of this research show that even a small deviation can cause CO_2-eq emissions calculated using the GWP* to reduce and therefore provide an incorrect impression of the multiplier trajectory over time. This research suggests that the GWP* should not be used to calculate CO_2-eq emissions from the industry in the manner done by Lee et al. (2021).

A second issue with the GWP* identified in this research is that it essentially has two time horizons. Meinshausen and Nicholls (2022) argue that the GWP* is a model rather than a climate metric. This can be seen in Figure 3.3 - rather than providing a single value, as conventional climate metrics do, the GWP* provides a temporal trajectory for each emission species. Just as with the time horizon, as discussed above for REQ 3, there is no obvious point to choose. In this research, the time of the peak total value is chosen, which is almost entirely caused by the impact of contrails and ozone for the fleets analysed. The location of the peak is mathematically at least 20 years in the future, but the exact location can differ depending on the emission profile. For example, peak CO_2 emissions occur much further along in time. Picking a different point could be possible, but this needs to be soundly reasoned and should not be chosen at a location where emissions have reversed due to a reduction in emissions since doing so would give the wrong outcome. Adding what amounts to a second time horizon that changes depending on the emission profile is problematic.

Nevertheless, the GWP* should not be discounted entirely. As the results show, in certain circumstances, the GWP* or its derivative EGWP* can act like a conventional climate metric and provide very usable results. Provided the peak CO_2-eq value is used, the starred climate metrics are able to estimate the temperature response accurately and can be a useful tool. However, the aim of the GWP* should not be to replace the GWP in all areas of use since this can lead to inappropriate results.

Use Case: Aircraft Design Optimisation

For aircraft design optimisation, the low inherent bias requirement (REQ 5) is most important since it ensures that an aircraft design with a lower climate metric value also has a lower environmental impact. To properly account for the climate impact of all aviation emissions, the temperature should be used as an input (REQ 1). It is assumed that aircraft designers do not mind using climate metrics as a "black box" in their design work, therefore REQ 2 is given less weight. Since the introduction of a new fleet will take some time and the fleet will be in service for many years, a high time horizon should be used. As the results show, at higher time horizons, the dependency is lower, therefore reducing the importance of REQ 3. However, for high time horizons the dependency of background emissions (REQ 4) is important. Finally, since the results show that all climate metrics except the RF are appropriate for fleet emission profiles, REQ 6 is of lower importance in this discussion.

Based on these requirements, the absolute ATR_{100} is recommended. It performs excellently in the pairwise fleet analysis for both the peak and average temperature, meaning that the likelihood of the climate metric suggesting incorrectly that one fleet is better than another is low. It also has a generally low dependence on the background emissions scenario and accounts for a change in temperature rather than in radiative forcing. The 100-year time horizon is kept to be in line with the standard GWP_{100} used in international climate policy. As the results of the CO_2-eq emission profiles show, the total ATR_{100} and GWP_{100} values do not differ by much, making it possible to convert from one to the other for comparison with other studies. Niklaß et al. (2019) provide an example of a conversion between the two climate metrics. With the introduction of different fuels such as hydrogen, it is, however, possible that the GWP and ATR will start showing different results since the main differences are for contrails and ozone.

In general, relative climate metrics using the full CO_2 response are not recommended for aircraft design optimisation. Relative climate metrics using a CO_2 pulse emission are more appropriate. This is because a reduction in CO_2 would erroneously lead to an increase in the climate metric value. Amongst the absolute climate metrics, the the EGWP* shows similarly good results as the ATR_{100} for the pairwise fleet analysis, but it is more uncertain and thus also not recommended. It relies on estimates of efficacies, for which not much data is available. While it is less dependent on the background emissions scenario, it has the secondary time horizon issue as described above and can be less easily compared to existing studies.

Use Case: Trajectory Optimisation

Since trajectory optimisation focuses mainly on reducing the effect of contrails and other SLCPs, it is beneficial for the climate metric to be temperature-based (REQ 1). Similarly to the aircraft design optimisation, it is assumed that a "black box" is sufficient for this case, thereby giving REQ 2 a low weight. Trajectory optimisation primarily focuses on short-term effects, requiring a low time horizon. Since the gradient of the time horizon dependency is high at low

time horizons, REQ 3 is an important consideration. On the other hand, the low time horizon reduces the importance of the background emissions scenario (REQ 4). REQ 5 is not relevant for this use case. How appropriate a climate metric is for a pulse emission (REQ 6), the likely emission profile for trajectory optimisation, is also relevant for the discussion.

Based on these requirements, the ATR_{20} is recommended. The time horizon of 20 years is chosen since it is longer than the atmospheric lifetime of aviation SLCPs, and because it is commonly found in literature. Other time horizons could chosen depending on the exact outcome desired. The ATR is chosen as a climate metric because it uses the temperature change, has a low dependence on the background emissions scenario and is appropriate for pulse emissions. It could be argued that the GTP would be a more appropriate climate metric for this purpose since it has a lower dependence on the time horizon, a key requirement for this use case. However, for future fuels such as hydrogen which have a substantial reduction in both contrail and ozone impact, it is possible that the GTP becomes inappropriate for pulse emissions. The averaging performed by the ATR removes this issue. It is also very likely that the integrated GTP ($iGTP_{20}$) would be a useful climate metric for this purpose, but more work will have to be conducted since this climate metric is not part of this research. If the ATR_{20} is chosen, it is important that the influence of the time horizon is kept in mind when the climate metric is applied.

Use Case: Market-Based Schemes

For market-based schemes, it is important that policymakers understand the workings of the climate metric to be used (REQ 2) because they must decide which climate metric to use and successfully face scrutiny. It is beneficial also for this reason that a climate metric is temperature-based (REQ 1), although it could also be argued that existing climate metrics based on radiative forcing would be beneficial since the GWP is currently the most common climate metric in climate policy. The time horizon chosen depends on whether the objective of the market-based measure is targeting short-term, medium-term or long-term impacts. It is assumed here that long-term impacts should be reduced, therefore making the dependence on the time horizon (REQ 3) less important and the dependence on the background emissions scenario (REQ 4) more important. REQ 5 is not relevant for this use case, and a similar analysis is made for trajectory optimisation for REQ 6, assuming that a series of pulse emissions are most relevant for market-based schemes.

Based on these requirements, the ATR_{100} is recommended. The ATR is easy to understand for non-specialists, being simply the average temperature over the 100-year time horizon. The ATR also displays low dependence on the background emissions scenario. The time horizon of 100 years is chosen since it is assumed the long-term impacts are considered.

If consistency with existing climate policy is deemed an important requirement, the GWP_{100} could also be continued to be used. In that case, the lack of efficacy must be taken into account. This would primarily impact the value of contrail and NO_x emissions: the GWP generally overstates the value of contrails and understates the value of O_3 and CH_4 reduction. It should also be noted that, as with the aircraft design optimisation, it is possible that with the introduction of hydrogen as a fuel, the GWP and ATR will start showing substantially different values. This would require a re-evaluation of the conversion methods, such as the one presented by Niklaß et al. (2019).

Although a powerful tool, the GWP* is not recommended for use in market-based schemes.

As demonstrated in this research and more generally by Meinshausen and Nicholls (2022), the GWP* shows a high variability year-to-year, even with the 20-year rolling average. The chance of misunderstanding stagnating, reducing or even negative CO_2-eq emissions, especially when combined with the additional time horizon, is high. The complexity of the GWP* and its derivative EGWP* make them unsuitable for policy, although it is possible that these climate metrics will be used as Micro Climate Models (as suggested by Meinshausen and Nicholls, 2022) and thereby inform decision-making.

5.2. Estimation of Impacts of Assumptions and Uncertainties

The results of this research are subject to a number of assumptions and uncertainties. It is important to recognise these and estimate their impacts to ensure that the results are valid. This section provides an overview of the main uncertainties and assumptions to determine remaining research gaps.

Climate Model Uncertainties

The main model-related uncertainties relate to the use of AirClim. AirClim has been used in a number of previous studies and has been validated against the E39/C climate-chemistry model (Stenke et al., 2008). However, as a response model, AirClim makes use of pre-calculated data and is not as accurate as more complex climate models. A Monte Carlo analysis is built into AirClim to consider the uncertainties that arise from this, specifically uncertainties in the atmospheric lifetime, radiative forcing and temperature response of each species or effect (Dahlmann et al., 2016a; Grewe et al., 2021). Using this analysis, it is possible to determine likelihood ranges of, for example, a climate target being met for a given scenario.

The largest uncertainties due to changes in radiative forcing originate from contrails, PMO, ozone and water vapour. As the results show, the impacts of water vapour and PMO are low and are, therefore, of lower importance. On the other hand, contrails and ozone make up the majority of the initial positive radiative forcing. The effect of under- or over-estimation of the RF is thus two-fold: First, the relative importance of contrails and ozone has an influence on the value of NO_x emissions compared to contrail emissions. In other words, a pairwise analysis may find that fleets that produce less NO_x are favoured over those that reduce contrail emissions, a key element in the discussion on inherent biases (REQ 5). Secondly, the relative importance of contrails and ozone compared to CO_2 has an impact on the relative climate metric values. An overestimation of the response of various species means that these species would have a larger dependence on the background emissions scenario since the relative impact of CO_2 is higher. The opposite is true for underestimation. A similar argumentation can be devised for the uncertainty related to temperature, for which the largest uncertainties are due to ozone, PMO and H_2O.

Since the main objective of this research is to analyse the design of climate metrics, the accuracy of the radiative forcing or temperature responses are of lesser importance, so long as all responses are relative to one another. Therefore, the three uncertainties mentioned above are not considered directly. Instead, it is reasoned that the Monte Carlo analysis used in this research already covers a wide part of the uncertainty range by changing the emissions of different species without optimising the full aircraft. For example, a 20% reduction in contrail RF can also be obtained by reducing the contrail distance by 20%, which is possible in the

Monte Carlo simulation. Furthermore, the sensitivity analysis can be used to demonstrate that climate metrics are sensitive to the changes of individual emission species - since all climate metrics show sensitivity to each species[1], the exact value obtained should not influence the pairwise distribution of the climate metrics. The general trend of climate metrics shown in Section 4.1 would also not be affected. The CO_2-eq and temperature responses calculated using climate metrics could be different, however, since the temperature calculated using AirClim is taken as a reference, it is unlikely that the general trend of the climate metrics would change.

To alleviate this uncertainty, the methods shown in this research could be re-done with a more accurate climate model. However, given the large number of simulations required, especially for the multivariate fleet analysis, that may not be feasible.

Future Estimation Uncertainties

Another source of uncertainty relates to the difficulty of estimating the future. This is important firstly for the location- and altitude-dependent emission data set, and secondly for the expected changes in growth rate, efficiency and emission rate of various species. In this research, the WeCare 2050 data set (Grewe et al., 2017) and the Category 4 sub-set (van der Maten, 2021) are used. These data sets are important because they provide the underlying location and altitude data for all simulations. The analysis by van der Maten (2021) shows that the Category 4 sub-set covers a wide range of altitudes, distances and locations around the world. Using a different category would have consequences on the results, for example by raising or lowering the mean cruise altitude, by changing the contrail distance or by changing the total fuel used. The location of Category 4 emissions, mostly in the northern hemisphere, likely results in a higher impact calculated for NO_x since aviation ozone is most effective in the northern polar and mid-latitudes (Grewe et al., 2002; Ponater et al., 2006).

As before, both sources of uncertainty have an impact on the radiative forcing and temperature responses. If these responses themselves were the results of this research, then more importance would have to be placed on ensuring that they are as accurate and representative as possible, and the uncertainties would have to be considered. However, since either the qualitative response of climate metrics or the relative difference between the climate metrics are considered, the uncertainties influence all climate metrics and should not affect the results.

GWP* Calculation Method

The stock parameter s has a large influence on the results obtained using the GWP* method. As discussed by Cain et al. (2019) in the introduction of the parameter to the GWP* method, s is estimated using multiple linear regression of the GWP* response onto the temperature response of methane between the years of 1900-2100, and an optimal value could differ per short-lived climate pollutant. For methane, an optimal s of 0.25 could be found; for other SLCPs, the value will depend on the trajectory of past emissions.

As part of this research, a number of linear regressions were performed with the aim of determining the optimal s per emission species or effect. However, the optimisation did not provide useful results for a number of species, notably ozone, contrails and methane reduction for the standard TCRE as described by Cain et al. (2019). It was also found that an optimal s

---

[1]A notable exception is the RF_{100}. However, the reason for the lack of or incorrect sensitivity is because a time horizon of 100 years is being used, as described in more detail in Section 4.2.

would depend on both the value of the TCRE and the background emissions scenario. As the results in Section 4.4.1 show, the same closeness of fit achieved for methane in Cain et al. (2019) could not be duplicated to other species for the full fuel scenarios. In this research, the standard 0.25 value is used, and it is recommended to perform a more vigorous analysis should the GWP* be used for aviation emissions.

The EGWP* has the potential to potentially solve this issue, although the same regression was not performed due to time constraints. The EGWP* has an additional source of uncertainty, however, which is its use of the efficacy. The efficacy parameters used in this study stem from Ponater et al. (2006), the only reference that could be found that calculates the parameters specifically for aviation, although the authors do mention that the maximum deviations match those of Hansen et al. (2005) for non-aviation radiative forcings. The aviation-specific efficacies are determined using 30-year simulations of ECHAM4, coupled with a mixed layer ocean model. Although the results of this research show the EGWP* to be highly effective, especially with regards to the multivariate fleet analysis, the introduction of the efficacy as a shortcut for a more complex temperature estimation should be kept in mind.

Multivariate Fleet Analysis Assumptions

To perform the multivariate fleet analysis, a number of simplifying assumptions were made. As an example, hydrogen-powered aircraft require heavier and, especially in the case of compressed hydrogen, more voluminous fuel tanks. Fuel cells or hydrogen combustion systems are also required. These changes will have an affect on aircraft weight and fuel efficiency and likely on the cruising speed and altitude. However, developing an aircraft design tool capable of comparing fleets using different fuels is beyond the scope of this research. Instead, a number of parameters were decoupled and changes to the aircraft weight, fuel efficiency, speed and altitude were neglected except where stated in the methods. The result is that the fleets are theoretical and, therefore, some combinations of parameters result in fleets that are potentially not physically feasible.

Assumptions were also made for the changes in emissions and emission-related effects due to changes in the fuel. There is a lack of research available about the expected environmental impact of these changes for new fuels such as SAF and hydrogen. In this research, the data reported by McKinsey & Company (2020) is applied, but, as discussed around Table 3.2, it is unclear which climate metrics and methods were used to obtain the results in that report. Research into the exhaust consistency, which above all will affect the formation and longevity of contrails, is recommended.

These shortcomings are acceptable as long as the fleets shown in this research are used only in comparison to one another in a theoretical assessment framework. Since the objective of the multivariate fleet analysis is to ensure that a climate metric does not have an inherent bias towards a certain emission species, the fact that a fleet may not be physically feasible is not as relevant and, as discussed earlier in this section, helps cover the lack of uncertainty analysis for radiative forcing and efficacy.

Reliability of Best Climate Metric Recommendation

The final discussion topic relates to the reliability of the recommendation of the best-suited climate metric for each objective. The first important note is that only a limited number of climate metrics could be analysed, given the time constraints of this research. Potentially interesting

climate metrics for use in aviation include the integrated global temperature change (iGTP) and the combined climate metrics CGWP and CGTP, although the latter climate metrics could be potentially problematic for contrail emissions which cannot be described on a per kilogram basis. Furthermore, only a few different emission profiles, notably the pulse, sustained, increasing and fleet profiles, were considered.

The tests conducted depend primarily on the requirements, which are specific for the aviation industry. It is possible that other requirements may be important for stakeholders, which would change which tests should be performed or the weight of each result on the final decision. An important example is how well the climate metric fits into existing climate policy: The relevance of this requirement might be high to ensure that the measures being put into place in the aviation industry matches those in other industries. Were this to be the case, the value of the GWP may increase and become the climate metric of choice. Another potential requirement could consider how flexible a climate metric is to be used in other industries.

This research provides stakeholders at various levels the means to determine the optimal climate metric for their purposes and by no means ends the discussion over climate metrics for aviation. The choice of climate metric will always depend on the objective, the requirements and the weight of those requirements. The results indicate that the ATR should be considered as a replacement for the GWP for aviation. A significant benefit is to be gained if the same climate metric is used across the industry.

6

Conclusions & Recommendations

This research aims to recommend the best-suited climate metrics for aviation climate objectives. Based on the systematic analysis of existing, physical climate metrics for pulse, sustained, increasing and aviation industry emission profiles, it can be concluded that the Average Temperature Response (ATR) is the most appropriate climate metric for implementation in aircraft and trajectory optimisation as well as in market-based schemes. The results do, however, indicate that the choice of time horizon remains subjective and must be chosen carefully.

Four main analyses are conducted to identify how well the climate metrics perform for each requirement. First, the general response analysis aims to identify systematic problems with climate metrics and calculate the dependence on the background emissions scenario and time horizon. The multivariate fleet analysis explores any inherent biases within the climate metric calculations for different aviation-specific emission species or changes in aircraft or trajectory design. The discussion of this section is aided by the sensitivity analysis, which calculates the sensitivity of climate metrics to changes in the fleet generation variables. Finally, the CO_2-eq emission analysis evaluates the ability of climate metrics to estimate CO_2-eq emissions and the temperature, which are useful assets for use in market-based schemes such as CORSIA or the EU ETS.

The analysis of pulse, sustained and increasing emissions show that it is not possible to mathematically determine an optimal time horizon and that all climate metrics have a dependence on the time horizon. This also includes the GWP*, for which proponents had previously suggested that a dependence was not to be expected as long as the time horizon was much greater than the lifetime of all short-lived climate pollutants. The results further indicate that absolute and relative climate metrics can demonstrate the opposite dependence to the background emissions scenario. For absolute climate metrics, scenarios with high decarbonisation result in higher values and scenarios with rapidly increasing emissions result in lower values. For relative climate metrics, this is reversed. This could potentially be problematic when climate metrics are introduced into emission trading schemes since the wrong incentive could be given to the market.

The CO_2-eq emission analysis shows the similarity between the GWP_{100} and relative ATR_{100} for total emissions. This allows the use of simple conversion formulae between the two climate metrics and could make the introduction of the ATR_{100} in place of the GWP_{100} easier for stakeholders. It is, however, possible that with the introduction of new fuels such as hydrogen that have a large influence on contrail production and NO_x emission, larger differences between the results will be seen. This is because the main differences between the two climate metrics originate in the responses of contrails and ozone, a result of NO_x emissions. The results of the sensitivity analysis indicate that the contrail distance and NO_x factors, with the cruise pressure factor, have the largest influence on the climate metric values. These fac-

tors could also be identified as the main causes of data spread and errors in fleet pairings in the multivariate fleet analysis, along with the influence of the background emissions scenario.

The CO_2-eq emission analysis also builds on the discussion about the use of relative and absolute climate metrics. In literature, it has been proposed to use simple multipliers, i.e. the ratio between the total emission and CO_2 emission, to convert CO_2 emissions to CO_2-eq emissions in market-based schemes. This research shows that simple multipliers do not work when CO_2 emissions are reducing, such as when biofuels, synthetic fuels or hydrogen are used, because the multipliers increase and tend to infinity. Instead, the absolute climate metric values must be compared to a pulse of CO_2 to determine the amount of CO_2-eq, which can be assigned a price. This method retains the advantage of the simple multiplier that only a single emission price, namely that of CO_2-eq, must be negotiated. The CO_2 pulse must be chosen either as a standard value or should be calculated for the conditions in which an aircraft is flying. For aircraft and trajectory optimisation, the absolute climate metric is also a valid form of comparison.

Based on the results and discussion presented in this work, it is recommended that policymakers consider introducing the ATR as the standard climate metric for aviation policy. For market-based schemes such as the EU ETS and CORSIA as well as for aircraft design optimisation, a time horizon of 100 years is best-suited to match existing policy and to ensure that long-lived emission species such as CO_2 are properly accounted for. For trajectory optimisation schemes through entities such as EUROCONTROL, the ATR_{20} is recommended to focus on the short-term impact but still incorporate the full lifetime of short-lived climate impacts.

Future studies could use different climate models and aircraft design optimisation tools to extend and verify this work. A number of uncertainties related to the climate model AirClim could be alleviated by performing the same or similar analyses with different climate models. These include the uncertainties that arise from atmospheric lifetimes of the emission species and their radiative forcing and temperature responses, as well as the impact of other species such as aerosols. Furthermore, uncertainties and assumptions related to the fleet design process used in this work can be validated and improved upon using tools that incorporate more facets of aircraft design, such as changes in maximum takeoff weight, cruise altitude and cruise speed.

Finally, further research is needed to determine whether the GWP* could be used as a tool in the aviation industry, and whether this is desired. Based on the results of this work, it is not recommended to use the GWP* to account for aviation non-CO_2 emissions due to its propensity to show declining or negative CO_2-eq emissions, which could be misinterpreted and hinder vital emission reduction. Since the GWP* provides a temporal trajectory for each emission species, it also essentially has two time horizons. The results indicate that the GWP* behaves very differently to conventional climate metrics, and it should thus not be seen as a replacement to the standard GWP. Nevertheless, the GWP* and its derivative EGWP* can act as a conventional climate metric and provide useful results. It is possible that in its capacity as a Micro Climate Model, as suggested by Meinshausen and Nicholls (2022), the GWP* method could be used effectively as a shortcut for temperature estimations. Further research could, therefore, focus on investigating optimal values of the stock parameter s, the influence of the Transient Climate Response (TCRE) and efficacies of aviation-specific forcings, all of which are required to ensure the correct implementation of the GWP* and EGWP* for aviation.

A

Appendix

A.1. Dependence of Individual Species on Time Horizon for P2020 and C2020 Scenarios

This section shows the dependence of individual species on the time horizon for the pulse and sustained emissions. The results of the increasing emission are shown in Figure 4.13 in Section 4.1.8.

Figure A.1: Percentage contribution of each species or effect to the total climate metric value for time horizons 20, 50 and 100 years for the pulse emission scenario. The SSP2-4.5 background scenario is used.

The results of the pulse emission are shown in Figure A.1. Since all emissions and effects apart from CO_2 are short-lived, the RF is dominated by CO_2 as the time horizon increases. A similar effect can be seen for the AGWP, AGTP and ATR, where the influence of CO_2 increases with time horizon. The GWP* and EGWP* show the opposite behaviour: the influence of CO_2 decreases with increasing time horizon. The dependence of the GWP* and EGWP* to the

time horizon is generally much lower than the other climate metrics shown. As described in the main text body, the GWP* and EGWP* show the same response for all scenarios.

The results of the sustained emission are shown in Figure A.2. The results are visually very similar to those of the increasing emission shown in the main text body. The contrail response dominates the climate metrics based on radiative forcing (RF, AGWP and GWP*), whereas the ozone response dominates those based on effective radiative forcing (EGWP*) or temperature change (AGTP and ATR). CO_2 generally shows the highest dependence on time horizon, followed by CH_4 reduction. Except for the RF and AGTP, ozone and contrails show a low dependence on the time horizon.
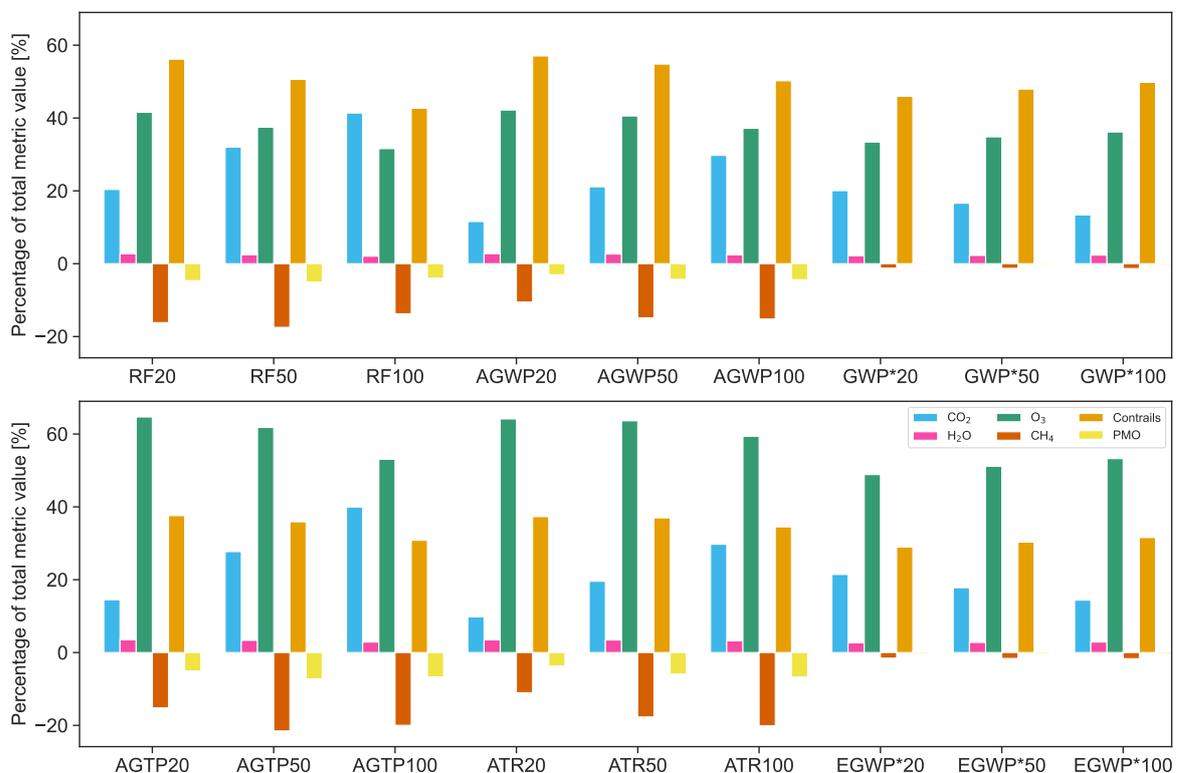
Figure A.2: Percentage contribution of each species or effect to the total climate metric value for time horizons 20, 50 and 100 years for the sustained emission scenario. The SSP2-4.5 background scenario is used.

A.2. Average Temperature Pairwise Fleet Analysis

This section shows the pairwise analysis of all fleets in the multivariate fleet analysis for the 20-, 50- and 100-year average temperature. The pairwise analysis of only the Jet-A1 fleets is also shown. The peak temperature pairwise analysis is described in Section 4.3.2.
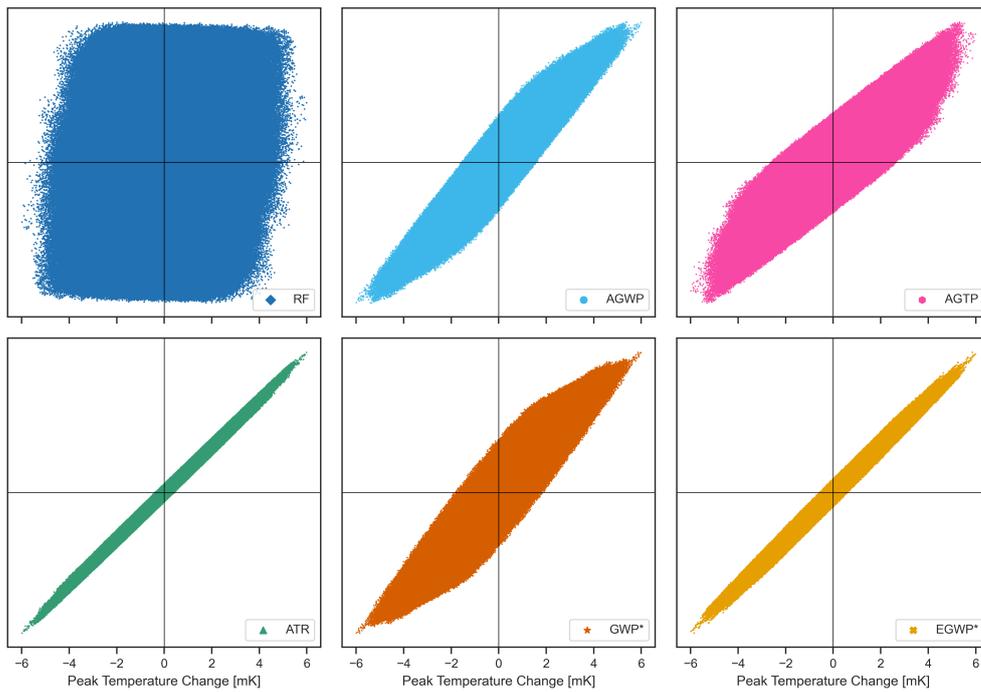
Figure A.3: Peak temperature pairwise analysis of fleets powered by Jet-A1. The vertical axis is the change in absolute climate metric value, with zero in the middle. Note the almost rectangular response from the RF and lack of extra extrusions caused by the SAF and hydrogen fleets in comparison to Figure 4.22.
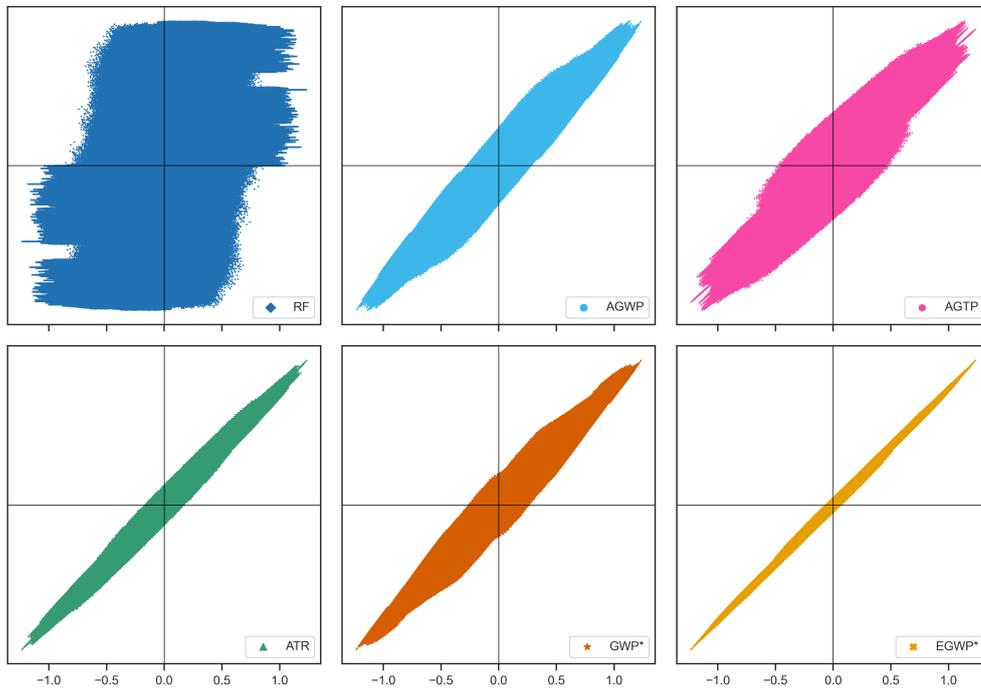
Figure A.4: 20-year average temperature pairwise analysis of all fleets developed with the Monte Carlo analysis. The vertical axis is the change in absolute climate metric value, with zero in the middle.
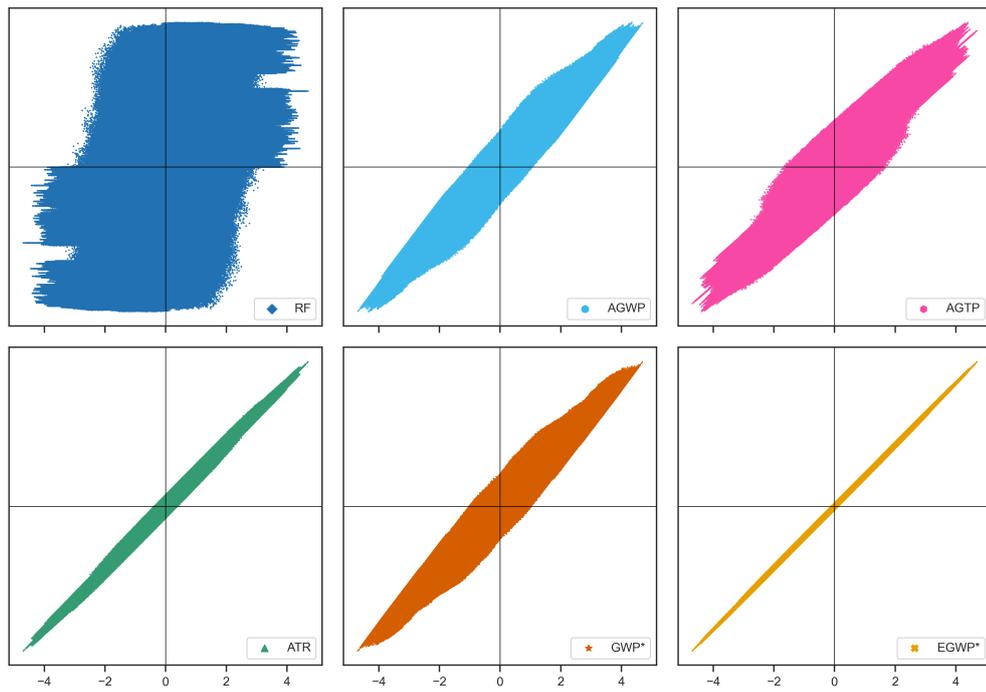
Figure A.5: 50-year average temperature pairwise analysis of all fleets developed with the Monte Carlo analysis. The vertical axis is the change in absolute climate metric value, with zero in the middle.
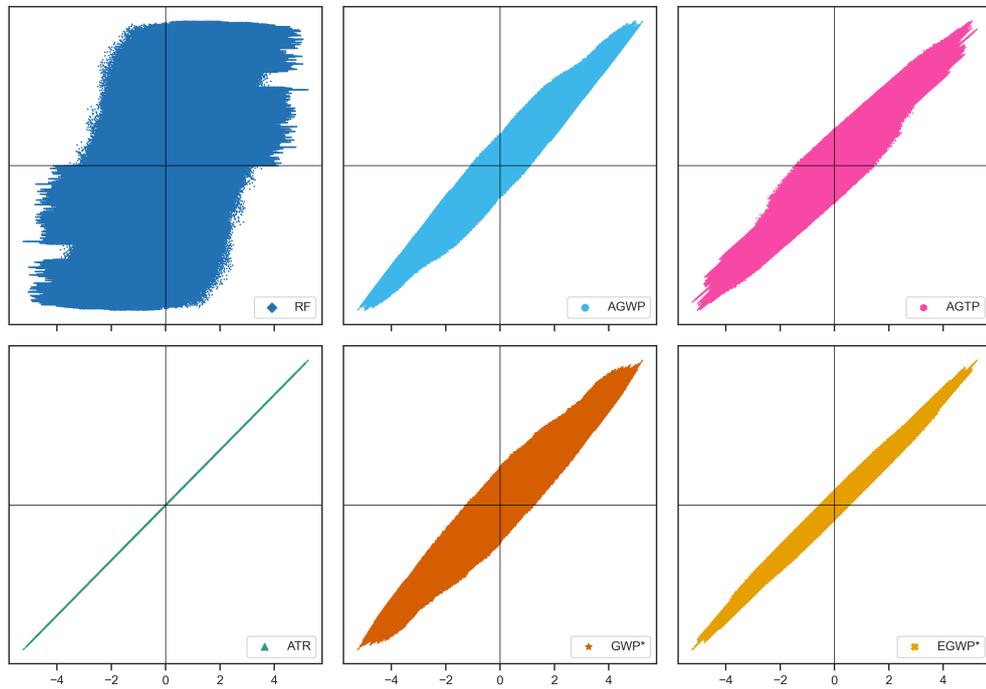
Figure A.6: 100-year average temperature pairwise analysis of all fleets developed with the Monte Carlo analysis. The vertical axis is the change in absolute climate metric value, with zero in the middle.

A.3. CO$_2$-Equivalent Responses of Emission Profiles using Climate Metrics

This section shows the CO$_2$-eq emissions calculated using climate metrics for the CORSIA, COVID-15s and Fa1 fuel scenarios. These are not included in the main text, specifically Section 4.4.1, because the CurTec and FP2050 scenarios are able to explain all effects.
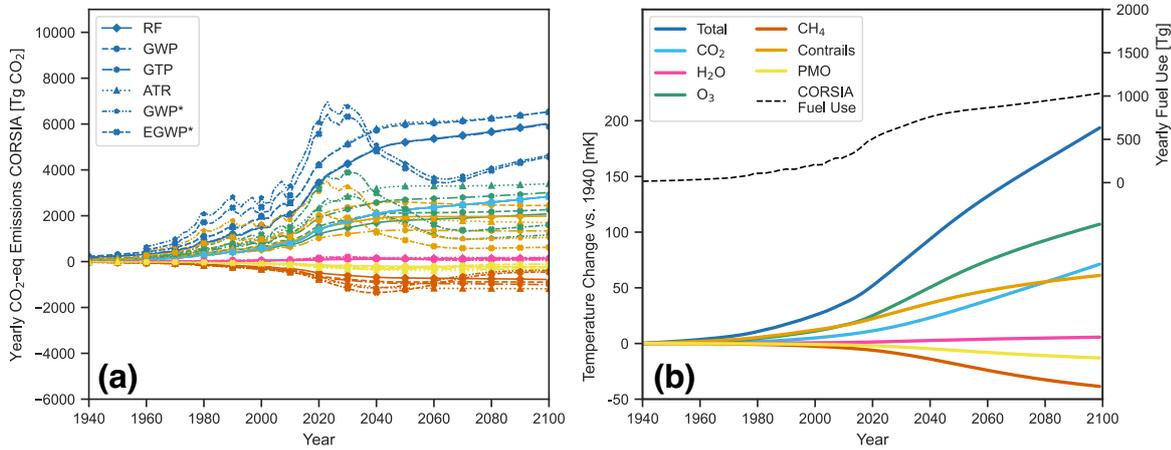
Figure A.7: CORSIA fuel scenario: **(a)** CO$_2$-eq emissions calculated using climate metrics; **(b)** fuel use and temperature response from each emission species. The colours in the left figure correspond to the emission species in the right figure.
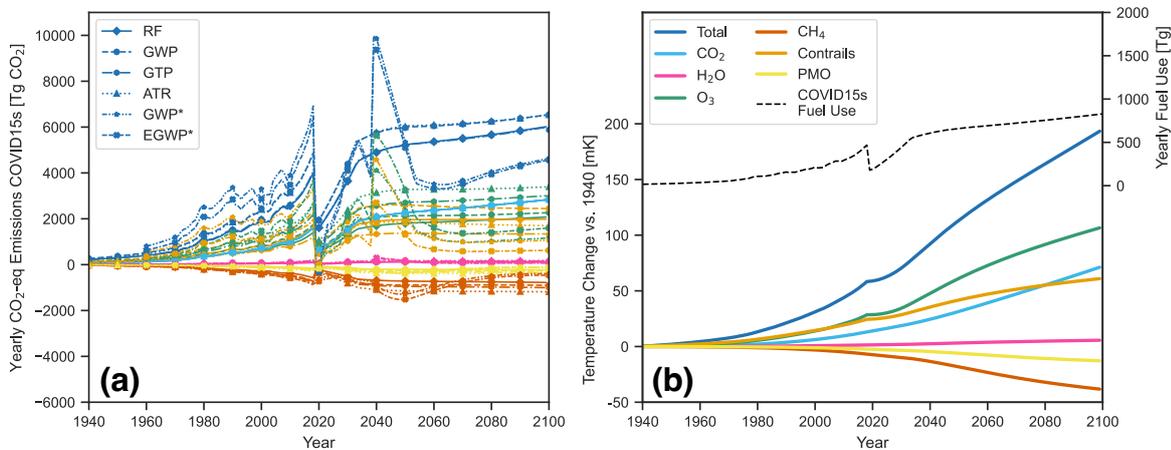
Figure A.8: COVID-15s fuel scenario: **(a)** CO$_2$-eq emissions calculated using climate metrics; **(b)** fuel use and temperature response from each emission species. The colours in the left figure correspond to the emission species in the right figure.
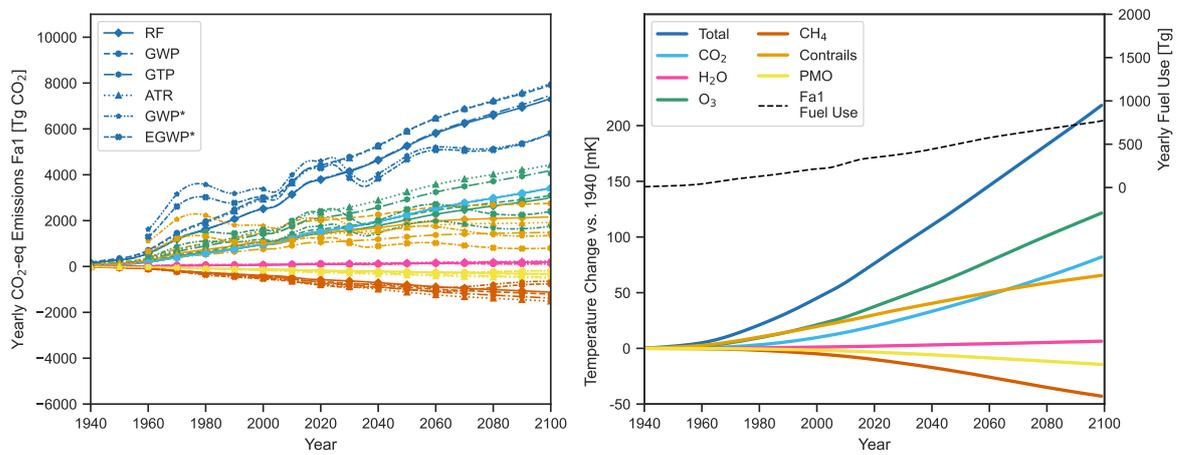
Figure A.9: Fa1 fuel scenario: **(a)** CO$_2$-eq emissions calculated using climate metrics; **(b)** fuel use and temperature response from each emission species. The colours in the left figure correspond to the emission species in the right figure. Note the erratic behaviour of the GWP* and EGWP* compared to the other climate metrics.

Bibliography

Allen, M. R., Fuglestvedt, J. S., Shine, K. P., Reisinger, A., Pierrehumbert, R. T., and Forster, P. M. (2016). New use of global warming potentials to compare cumulative and short-lived climate pollutants. *Nature Climate Change*, **6**, 773-776.

Allen, M. R., Shine, K. P., Fuglestvedt, J. S., Millar, R. J., Cain, M., Frame, D. J., and Macey, A. H. (2018). A solution to the misrepresentations of CO_2-equivalent emissions of short-lived climate pollutants under ambitious mitigation. *npj Climate and Atmospheric Science*, **1**, 1-8.

Appleman, H. (1953). The formation of exhaust condensation trails by jet aircraft. *Bulletin of the American Meteorological Society*, **34**, 14–20.

Berntsen, T. K., Fuglestvedt, J. S., Joshi, M. M., Shine, K. P., Stuber, N., Ponater, M., Sausen, R., Hauglustaine, D. A., and Li, L. (2005). Response of climate to regional emissions of ozone precursors: sensitivities and warming potentials. *Tellus B: Chemical and Physical Meteorology*, **57**, 283–304.

Bickel, M., Ponater, M., Bock, L., Burkhardt, U., and Reineke, S. (2020). Estimating the effective radiative forcing of contrail cirrus. *Journal of Climate*, **33**, 1991–2005.

Bier, A., Burkhardt, U., and Bock, L. (2017). Synoptic control of contrail cirrus life cycles and their modification due to reduced soot number emissions. *Journal of Geophysical Research: Atmospheres*, **122**, 11584–11603.

Bock, L. and Burkhardt, U. (2016). Reassessing properties and radiative forcing of contrail cirrus using a climate model. *Journal of Geophysical Research: Atmospheres*, **121**, 9717–9736.

Bradford, D. F. (2001). Time, money and tradeoffs. *Nature*, **410**, 649–650.

Burkhardt, U. and Kärcher, B. (2009). Process-based simulation of contrail cirrus in a global climate model. *Journal of Geophysical Research*, **114**, 1–13.

Cain, M., Lynch, J., Allen, M. R., Fuglestvedt, J. S., Frame, D. J., and Macey, A. H. (2019). Improved calculation of warming-equivalent emissions for short-lived climate pollutants. *npj Climate and Atmospheric Science*, **2**, 1–7.

Clean Sky 2 Joint Undertaking (2021). Technology evaluator. first global assessment 2020. Technical report, European Commission.

Collins, W. J., Frame, D. J., Fuglestvedt, J. S., and Shine, K. P. (2020). Stable climate metrics for emissions of short and long-lived species—combining steps and pulses. *Environmental Research Letters*, **15**, 1–9.

Collins, W. J., Fry, M. M., Yu, H., Fuglestvedt, J. S., Shindell, D. T., and West, J. J. (2013). Global and regional temperature-change potentials for near-term climate forcers. *Atmospheric Chemistry and Physics*, **13**, 2471–2485.

Cooper, T., Smiley, J., Porter, C., and Precourt, C. (2018). Global fleet & MRO market forecast commentary 2018–2028. Technical report, Oliver Wyman.

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

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

Dahlmann, K., Koch, A., Linke, F., Lührs, B., Grewe, V., Otten, T., Seider, D., Gollnick, V., and Schumann, U. (2016b). Climate-compatible air transport system-climate impact mitigation potential for actual and future aircraft. *Aerospace*, **3**, 1–25.

Dakhel, P. M., Lukachko, S. P., Waitz, I. A., Miake-Lye, R. V., and Brown, R. C. (2007). Postcombustion evolution of soot properties in an aircraft

engine. *Journal of Propulsion and Power*, **23**, 942–948.

Dallara, E. S. (2011). *Aircraft Design for Reduced Climate Impact*. PhD thesis, Stanford University, Stanford, CA.

Dallara, E. S., Kroo, I. M., and Waitz, I. A. (2011). Metric for comparing lifetime average climate impact of aircraft. *AIAA Journal*, **49**, 1600-1613.

Derwent, R. G. (1990). Trace gases and their relative contribution to the greenhouse effect. *AERE R 13716*.

Egelhofer, R., Bickerstaff, C., and Bonnet, S. (2007). Minimizing impact on climate in aircraft design. In *AeroTech Congress & Exhibition, September 17 - 20, 2007, Los Angeles, California*, number 2007-01-3807 in SAE Technical Paper Series.

European Commission (2011). *Flightpath 2050: Europe's Vision for Aviation. Report of the High Level Group on Aviation Research*. European Union, Brussels.

European Union Aviation Safety Agency (2020). Updated analysis of the non-CO2 climate impacts of aviation and potential policy measures pursuant to the EU Emissions Trading System Directive Article 30(4). Full-length report accompanying the document. Technical Report 52020SC0277, European Commission.

Fichter, C. (2009). *Climate impact of air traffic emissions in dependency of the emission location*. PhD thesis, Manchester Metropolitan University.

Forster, P. M. d. F., Shine, K. P., and Stuber, N. (2006). It is premature to include non-CO_2 effects of aviation in emission trading schemes. *Atmospheric Environment*, **40**, 1117–1121.

Frömming, C., Grewe, V., Brinkop, S., Jöckel, P., Haslerud, A. S., Rosanka, S., van Manen, J., and Matthes, S. (2021). Influence of weather situation on non-CO_2 aviation climate effects: the REACT4C climate change functions. *Atmospheric Chemistry and Physics*, **21**, 9151–9172.

Frömming, C., Ponater, M., Dahlmann, K., Grewe, V., Lee, D. S., and Sausen, R. (2012). Aviation-induced radiative forcing and surface tempera-

ture change in dependency of the emission altitude. *Journal of Geophysical Research: Atmospheres*, **117**, 1–15.

Fuglestvedt, J. S., Berntsen, T. K., Godal, O., Sausen, R., Shine, K. P., and Skodvin, T. (2003). Metrics of climate change: Assessing radiative forcing and emission indices. *Climatic Change*, **58**, 267-331.

Fuglestvedt, J. S., Shine, K. P., Berntsen, T., Cook, J., Lee, D. S., Stenke, A., Skeie, R. B., Velders, G. J. M., and Waitz, I. A. (2010). Transport impacts on atmosphere and climate: Metrics. *Atmospheric Environment*, **44**, 4648-4677.

Gettelman, A. and Chen, C. (2013). The climate impact of aviation aerosols. *Geophysical Research Letters*, **40**, 2785–2789.

Grewe, V. (2009). Impact of lightning on air chemistry and climate. In Betz, H. D., Schumann, U., and Laroche, P., editors, *Lightning: Principles, Instruments and Applications*, pages 537–549. Springer, Dordrecht.

Grewe, V. and Dahlmann, K. (2012). Evaluating climate-chemistry response and mitigation options with AirClim. In Schumann, U., editor, *Atmospheric Physics, Research Topics in Aerospace*, pages 591–606. Springer, Berlin, Heidelberg.

Grewe, V. and Dahlmann, K. (2015). How ambiguous are climate metrics? And are we prepared to assess and compare the climate impact of new air traffic technologies? *Atmospheric Environment*, **106**, 373-374.

Grewe, V., Dahlmann, K., Flink, J., Frömming, C., Ghosh, R., Gierens, K., Heller, R., Hendricks, J., Jöckel, P., Kaufmann, S., Kölker, K., Linke, F., Luchkova, T., Lührs, B., van Manen, J., Matthes, S., Minikin, A., Niklaß, M., Plohr, M., Righi, M., Rosanka, S., Schmitt, A., Schumann, U., Terekhov, I., Unterstrasser, S., Vázquez-Navarro, M., Voigt, C., Wicke, K., Yamashita, H., Zahn, A., and Ziereis, H. (2017). Mitigating the climate impact from aviation: Achievements and results of the DLR WeCare project. *Aerospace*, **4**, 1-50.

Grewe, V., Dameris, M., Fichter, C., and Sausen, R. (2002). Impact of aircraft NO_x emissions. Part 1: Interactively coupled

climate-chemistry simulations and sensitivities to climate-chemistry feedback, lightning and model resolution. *Meteorologische Zeitschrift*, **11**, 177–186.

Grewe, V., Plohr, M., Cerino, G., Muzio, M. D., Deremaux, Y., Galerneau, M., de Saint Martin, P., Chaika, T., Hasselrot, A., Tengzelius, U., and Korovkin, V. D. (2010). Estimates of the climate impact of future small-scale supersonic transport aircraft – results from the HISAC EU-project. *The Aeronautical Journal*, **114**, 199-206.

Grewe, V., Rao, A. G., Grönstedt, T., Xisto, C., Linke, F., Melkert, J., Middel, J., Ohlenforst, B., Blakey, S., Christie, S., Matthes, S., and Dahlmann, K. (2021). Evaluating the climate impact of aviation emission scenarios towards the Paris agreement including COVID-19 effects. *Nature Communications*, **12**, 1–10.

Grewe, V. and Stenke, A. (2008). AirClim: an efficient tool for climate evaluation of aircraft technology. *Atmospheric Chemistry and Physics*, **8**, 4621–4639.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S. (2005). Efficacy of climate forcings. *Journal of Geophysical Research*, **110**, 1–45.

Hasselmann, K., Sausen, R., Maier-Reimer, E., and Voss, R. (1993). On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Climate Dynamics*, **9**, 53–61.

Intergovernmental Panel on Climate Change (1990). *Climate Change. The IPCC Scientific Assessment*. Cambridge University Press, Cambridge, U.K.

Intergovernmental Panel on Climate Change (1995). *Climate Change 1994. Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios*. Cambridge University Press, Cambridge, U.K.

Intergovernmental Panel on Climate Change (1999). *Aviation and the Global Atmosphere. A Special Report of IPCC Working Groups I and III*. Cambridge University Press, Cambridge, U.K.

Intergovernmental Panel on Climate Change (2001). *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, U.K.

Intergovernmental Panel on Climate Change (2013). *Climate Change 2013: The Physical Science Bases. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingfom and New York, NY, USA.

Intergovernmental Panel on Climate Change (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, U.K.

Irvine, E. A., Hoskins, B. J., and Shine, K. P. (2014). A lagrangian analysis of ice-supersaturated air over the North Atlantic. *Journal of Geophysical Research: Atmospheres*, **119**, 90–100.

Kandlikar, M. (1996). Indices for comparing greenhouse gas emissions: integrating science and economics. *Energy Economics*, **18**, 265–281.

Kapadia, Z. Z., Spracklen, D. V., Arnold, S. R., Borman, D. J., Mann, G. W., Pringle, K. J., Monks, S. A., Reddington, C. L., Benduhn, F., Rap, A., Scott, C. E., Butt, E. W., and Yoshioka, M. (2016). Impacts of aviation fuel sulfur content on climate and human health. *Atmospheric Chemistry and Physics*, **16**, 10521–10541.

Koch, A. (2013). *Climate Impact Mitigation Potential given by Flight Profile and Aircraft Optimization*. PhD thesis, Technischen Universität Hamburg-Harburg, Hamburg, Germany.

Köhler, M., Rädel, G., Shine, K., Rogers, H., and Pyle, J. (2013). Latitudinal variation of the effect of aviation NO$_x$ emissions on atmospheric ozone and methane and related climate metrics. *Atmospheric Environment*, **64**, 1–9.

Lee, D. S., Fahey, D. W., Forster, P. M., Newton, P. J., Wit, R. C., Lim, L. L., Owen, B., and Sausen, R. (2009). Aviation and global climate change in the 21st century. *Atmospheric Environment*, **43**, 3520–3537.

Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., Doherty, S. J., Freeman, S., Forster, P. M., Fuglestvedt, J., Gettelman, A., De León, R. R., Lim, L. L., Lund, M. T., Millar, R. J., Owen, B., Penner, J. E., Pitari, G., Prather, M. J., Sausen, R., and Wilcox, L. J. (2021). The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric Environment*, **244**, 1–29.

Lee, D. S., Pitari, G., Grewe, V., Gierens, K., Penner, J. E., Petzold, A., Prather, M. J., Schumann, U., Bais, A., Berntsen, T., Iachetti, D., Lim, L. L., and Sausen, R. (2010). Transport impacts on atmosphere and climate: Aviation. *Atmospheric Environment*, **44**, 4678-4734.

Lund, M. T., Aamaas, B., Berntsen, T., Bock, L., Burkhardt, U., Fuglestvedt, J. S., and Shine, K. P. (2017). Emission metrics for quantifying regional climate impacts of aviation. *Earth System Dynamics*, **8**, 547–563.

Lund, M. T., Berntsen, T., Fuglestvedt, J. S., Ponater, M., and Shine, K. P. (2012). How much information is lost by using global-mean climate metrics? An example using the transport sector. *Climatic Change*, **113**, 949–963.

Manne, A. S. and Richels, R. G. (2001). An alternative approach to establishing trade-offs among greenhouse gases. *Nature*, **410**, 675-677.

Matthes, S., Lim, L., Burkhardt, U., Dahlmann, K., Dietmüller, S., Grewe, V., Haslerud, A. S., Hendricks, J., Owen, B., Pitari, G., Righi, M., and Skowron, A. (2021). Mitigation of non-CO$_2$ aviation's climate impact by changing cruise altitudes. *Aerospace*, **8**, 1–20.

McKinsey & Company (2020). Hydrogen-powered aviation: A fact-based study of hydrogen technology, economics, and climate impact by 2050. Technical report, Clean Sky 2 JU and Fuel Cells and Hydrogen 2 JU.

Meinshausen, M. and Nicholls, Z. (2022). GWP* is a model, not a metric. *Environmental Research Letters*, **17**, 1–6.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, **13**, 3571–3605.

Michaelis, P. (1992). Global warming: Efficient policies in the case of multiple pollutants. *Environmental and Resource Economics*, **2**, 61-77.

Myhre, G., Shine, K., Rädel, G., Gauss, M., Isaksen, I., Tang, Q., Prather, M., Williams, J., van Velthoven, P., Dessens, O., Koffi, B., Szopa, S., Hoor, P., Grewe, V., Borken-Kleefeld, J., Berntsen, T., and Fuglestvedt, J. (2011). Radiative forcing due to changes in ozone and methane caused by the transport sector. *Atmospheric Environment*, **45**, 387–394.

Niklaß, M., Dahlmann, K., Grewe, V., Maertens, S., Plohr, M., Scheelhaase, J., Schwieger, J., Brodmann, U., Kurzböck, C., Repmann, M., Schweizer, N., and von Uger, M. (2019). Integration of non-CO$_2$ effects of aviation in the EU ETS and under CORSIA - final report. Technical report, Umweltbundesamt Deutschland (German Environment Agency), Dessau-Roßlau.

Niklaß, M., Grewe, V., Gollnick, V., and Dahlmann, K. (2021). Concept of climate-charged airspaces: a potential policy instrument for internalizing aviation's climate impact of non-CO$_2$ effects. *Climate Policy*, **21**, 1-20.

Ocko, I. B., Hamburg, S. P., Jacob, D. J., Keith, D. W., Keohane, N. O., Oppenheimer, M., Roy-Mayhew, J. D., Schrag, D. P., and Pacala, S. W.

(2017). Unmask temporal trade-offs in climate policy debates. *Science*, **356**, 492-493.

Penner, J. E., Zhou, C., Garnier, A., and Mitchell, D. L. (2018). Anthropogenic aerosol indirect effects in cirrus clouds. *Journal of Geophysical Research: Atmospheres*, **123**, 11652–11677.

Plattner, G.-K., Stocker, T., Midgley, P., and Tignor, M., editors (2009). *IPCC Expert Meeting on the Science of Alternative Metrics: Meeting Report*, Oslo, 18 to 20 March 2009. IPCC Working Group I, Technical Support Unit, Bern, Switzerland, 2009.

Ponater, M., Bickel, M., Bock, L., and Burkhardt, U. (2020). Towards determining the efficacy of contrail cirrus. In Matthes, S. and Blum, A., editors, *Making Aviation Environmentally Sustainable, 3rd ECATS Conference, Book of Abstracts*, volume 1, pages 53–55. ECATS International Association.

Ponater, M., Marquart, S., Sausen, R., and Schumann, U. (2005). On contrail climate sensitivity. *Geophysical Research Letters*, **32**, 1–5.

Ponater, M., Pechtl, S., Sausen, R., Schumann, U., and Hüttig, G. (2006). Potential of the cryoplane technology to reduce aircraft climate impact: A state-of-the-art assessment. *Atmospheric Environment*, **40**, 6928–6944.

Proesmans, P.-J. and Vos, R. (2021). Airplane design optimization for minimal global warming impact. In *AIAA Scitech 2021 Forum: 11-15 & 19-21 January 2021, Virtual Event [AIAA 2021-1297]*. American Institute of Aeronautics and Astronautics (AIAA).

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Silva, L. A. D., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, **42**, 153–168.

Rodhe, H. (1990). A comparison of the contribution of various gases to the greenhouse effect. *Science*, **248**, 1217–1219.

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

Scheelhaase, J., Dahlmann, K., Jung, M., Keimel, H., Murphy, M., Nieße, H., Sausen, R., Schaefer, M., and Wolters, F. (2015). Die einbeziehung des luftverkehrs in internationale klimaschutzprotokolle (AviClim). Abschlussbericht. Technical report, Deutsches Zentrum für Luft- und Raumfahrt e.V.

Schmidt, E. (1941). Die Entstehung von Eisnebel aus den Auspuffgasesn von Flugmotoren. *Schriften der Deutschen Akademie der Luftfahrtforschung*, **5**, 1–22.

Schumann, U., Graf, K., and Mannstein, H. (2011). Potential to reduce the climate impact of aviation by flight level changes. In *3rd AIAA Atmospheric Space Environments Conference, 27 - 30 June 2011, Honolulu, Hawaii*. American Institute of Aeronautics and Astronautics.

Schumann, U., Mayer, B., Graf, K., and Mannstein, H. (2012). A parametric radiative forcing model for contrail cirrus. *Journal of Applied Meteorology and Climatology*, **51**, 1391–1406.

Shindell, D. and Faluvegi, G. (2009). Climate response to regional radiative forcing during the twentieth century. *Nature Geoscience*, **2**, 294–300.

Shindell, D. and Faluvegi, G. (2010). The net climate impact of coal-fired power plant emissions. *Atmospheric Chemistry and Physics*, **10**, 3247–3260.

Shindell, D. T. (2012). Evaluation of the absolute regional temperature potential. *Atmospheric Chemistry and Physics*, **12**, 7955–7960.

Shine, K. P. (2009). The global warming potential - the need for an interdisciplinary retrial. *Climatic Change*, **96**, 467–472.

Shine, K. P., Berntsen, T. K., Fuglestvedt, J. S., Skeie, R. B., and Stuber, N. (2007). Comparing the climate effect of emissions of short- and long-lived climate agents. *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, **365**, 1903-1914.

Shine, K. P., Fuglestvedt, J. S., Hailemariam, K., and Stuber, N. (2005). Alternatives to the global warming potential for comparing climate impacts of emissions of greenhouse gases. *Climatic Change*, **68**, 281-302.

Smith, M. A., Cain, M., and Allen, M. R. (2021). Further improvement of warming-equivalent emissions calculation. *npj Climate and Atmospheric Science*, **4**, 1–3.

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

Stevenson, D. S. (2004). Radiative forcing from aircraft NO_x emissions: Mechanisms and seasonal dependence. *Journal of Geophysical Research*, **109**, 1–13.

Stevenson, D. S. and Derwent, R. G. (2009). Does the location of aircraft nitrogen oxide emissions affect their climate impact? *Geophysical Research Letters*, **36**, 1–5.

Stuber, N. and Forster, P. (2007). The impact of diurnal variations of air traffic on contrail radiative forcing. *Atmospheric Chemistry and Physics*, **7**, 3153–3162.

Svensson, F., Hasselrot, A., and Moldanova, J. (2004). Reduced environmental impact by lowered cruise altitude for liquid hydrogen-fuelled aircraft. *Aerospace Science and Technology*, **8**, 307–320.

Tanaka, K. and O'Neill, B. C. (2018). The Paris Agreement zero-emissions goal is not always consistent with the 1.5 °C and 2 °C temperature targets. *Nature Climate Change*, **8**, 319–324.

Tol, R. S. J., Berntsen, T. K., O'Neill, B. C., Fuglestvedt, J. S., Shine, K. P., Balkanski, Y., and Makra, L. (2008). Metrics for aggregating the climate effect of different emissions: A unifying framework. Technical report, The Economic and Social Research Institute (ESRI).

van der Maten, J. T. (2021). An inter-comparison of the climate impact of short-, medium- and long-range aircraft. Master's thesis, Delft University of Technology, Delft, NL.

Wit, R. C. N., Boon, B. H., van Velzen, A., Cames, M., Deuber, O., and Lee, D. S. (2005). Giving wings to emission trading. inclusion of aviation under the European emission trading system (ETS): design and impacts. Report for the European Commission. Technical Report ENV.C.2/ETU/2004/0074r, CE Delft, Delft, The Netherlands.

Wuebbles, D., Forster, P., Rogers, H., and Herman, R. (2010). Issues and uncertainties affecting metrics for aviation impacts on climate. *Bulletin of the American Meteorological Society*, **91**, 491-496.

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

Zhou, C. and Penner, J. E. (2014). Aircraft soot indirect effect on large-scale cirrus clouds: Is the indirect forcing by aircraft soot positive or negative? *Journal of Geophysical Research: Atmospheres*, **119**, 11303–11320.

Zickfeld, K., Eby, M., Matthews, H. D., and Weaver, A. J. (2009). Setting cumulative emissions targets to reduce the risk of dangerous climate change. *Proceedings of the National Academy of Sciences*, **106**, 16129–16134.