AerChemMIP: Quantifying the effects of chemistry and aerosols in CMIP6

William J. Collins[1], Jean-François Lamarque[2], Michael Schulz[3], Olivier Boucher[4], Veronika Eyring[5], Michaela I. Hegglin[1], Amanda Maycock[6], Gunnar Myhre[7], Michael Prather[8], Drew Shindell[9], Steven J Smith[10]

[1]Department of Meteorology, University of Reading, Reading, RG6 6BB, UK
[2]National Center for Atmospheric Research, Boulder, Colorado, USA
[3]Norwegian Meteorological Institute, Oslo, Norway
[4]Laboratoire de Météorologie Dynamique, IPSL, Université Pierre et Marie Curie/CNRS, Paris, France
[5]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[6]School of Earth and Environment, University of Leeds, Leeds, UK
[7]CICERO – Center for International Climate and Environmental Research Oslo, Oslo, Norway
[8]University of California, Irvine, CA, USA
[9]Nicholas School of the Environment, Duke University, Durham, North Carolina, 27708, USA
[10]Joint Global Change Research Institute, Pacific Northwest National Laboratory, 5825 University Research Court, Suite 3500, College Park, MD 20740, USA

Correspondence to: William Collins (w.collins@reading.ac.uk)

Abstract. The Aerosol Chemistry Model Intercomparison Project (AerChemMIP) is endorsed by the Coupled-Model Intercomparison Project 6 (CMIP6) and is designed to quantify the climate and air quality impacts of aerosols and chemically-reactive gases. These are specifically near-term climate forcers (NTCFs: methane, tropospheric ozone and aerosols, and their precursors), nitrous oxide and ozone-depleting halocarbons. The aim of AerChemMIP is to answer four scientific questions:

1. How have anthropogenic emissions contributed to global radiative forcing and affected regional climate over the historical period?
2. How might future policies (on climate, air quality and land use) affect the abundances of NTCFs and their climate impacts?
3. How do uncertainties in historical NTCF emissions affect radiative forcing estimates?
4. How important are climate feedbacks to natural NTCF emissions, atmospheric composition, and radiative effects?

These questions will be addressed through targeted simulations with CMIP6 climate models that include an interactive representation of tropospheric aerosols and atmospheric chemistry. These simulations build on the CMIP6 Diagnostic, Evaluation and Characterization of Klima (DECK) experiments, the CMIP6 historical simulations, and future projections performed elsewhere in CMIP6, allowing the contributions from aerosols and/or chemistry to be quantified. Specific diagnostics are requested as part of the CMIP6 data request to highlight the chemical composition of the atmosphere, to evaluate the performance of the models, and to understand differences in behaviour between them.

1 Introduction

1.1 Motivation for AerChemMIP

Aerosols and chemically reactive gases in the atmosphere can exert important influences on global and regional air quality and climate. Scientific questions and uncertainties regarding chemistry-climate interactions are relevant to regional scale climate change (e.g., tropospheric ozone and aerosols interacting with regional meteorology), to long-range connections (e.g., hemispheric transport of air pollution, the impacts of lower stratospheric ozone and temperatures on surface climate) and globally integrated effects (e.g., the lifetimes of methane (CH_4) and nitrous oxide (N_2O)).

Past climate change has been forced by a wide range of chemically reactive gases, aerosols and well-mixed greenhouse gases (WMGHGs), in addition to CO_2. More specifically, anthropogenic effects on methane, aerosol and ozone abundances (also known as near-term climate forcers, NTCFs) are estimated to have been responsible for a climate forcing that is presently nearly equal in magnitude to that of CO_2 (Shindell et al., 2013a; Myhre et al., 2013a). These emissions are thought to have led to a variety of global climate impacts including changes in regional patterns of temperature and precipitation (Rotstayn et al., 2015). In addition, NTCF forcing is inherently spatially inhomogeneous (Shindell et al., 2013a), which leads to regional responses, particularly for aerosols, and there is some evidence that the global climate response to a regional scale NTCF differs from that of an equivalent globally homogeneous radiative forcing (Shindell et al., 2012b; Shindell et al., 2015). Changes in the abundance of NTCFs can also induce rapid adjustments in meteorological quantities (such as atmospheric temperature, water vapour, clouds) through radiative heating/cooling and/or effects on precipitation and cloud lifetime (Sherwood et al., 2015). These adjustments are in principle independent of surface temperature changes (although in practical model configurations the land surface temperature will also respond). The contribution of such rapid adjustments to the change in Earth's energy budget following a perturbation of a radiatively active species can be incorporated into an effective radiative forcing (ERF) (Myhre et al. 2013a; Boucher et al. 2013; Sherwood et al., 2015), which has been shown to be a better predictor of the eventual surface temperature change than the traditional instantaneous or stratosphere-adjusted definitions of radiative forcing.

NTCFs were also identified in IPCC AR5 (Myhre et al., 2013a) as the main source of uncertainty in the total anthropogenic ERF since pre-industrial times. In particular, natural aerosols originating from biogenic sources, dust or sea-salt are a primary contributor to the uncertainty in present day aerosol forcing (Carslaw et al., 2013). This is because, the response of the climate system to human-induced aerosol perturbations depends critically on the natural aerosol background (Carlton et al., 2010, Gordon et al., 2016), due in part to the nonlinear response of aerosol-cloud interactions.

The forcing of climate by ozone changes has resulted from increases in ozone driven by changes in NO_X, CO and VOC emissions and methane abundance (mostly affecting the troposphere) and decreases in ozone driven by ozone depleting substances (ODSs), such as chlorofluorocarbons and hydrochlorofluorocarbons (mostly affecting the stratosphere) (Shindell et al., 2013b). The net ERF due to ozone is primarily the result of multiple emission changes. For example, one of the largest components of the ERF due to past methane emissions comes from the associated increase in tropospheric ozone (Prather et al, 2001; Stevenson et al. 2013). In addition, stratospheric ozone losses due to ODSs since the 1970s have led to a significant cooling of the stratosphere (Shine et al., 2003; McLandress et al., 2015), and the Antarctic ozone hole is linked to changes in tropospheric circulation and rainfall patterns in the southern hemisphere, especially during austral summer (WMO, 2014). In the Southern Hemisphere, future changes in summertime tropospheric circulation are expected to be controlled by both the rates of ozone recovery and WMGHG increases (McLandress et al., 2011; Polvani et al., 2011), indicating the need to account for ozone changes in future climate projections.

IPCC AR5 (Kirtman et al., 2013) found large uncertainties in projecting the future chemical composition of the atmosphere and climate insofar as it affects climate and air quality. Natural and managed ecosystems provide a large fraction of the methane and nitrous oxide emissions, and also emit aerosol and ozone precursors (e.g., through emissions of soil nitrogen oxides, biogenic volatile organic compounds, and wildfires). These sources are likely to be affected by climate change, leading to a variety of feedbacks (Arneth et al., 2010) that to date have only been quantified from a limited number of studies (and models) and thus the need for a coordinated set of simulations that allows for a consistent and clean comparison between models. For example, the CMIP5 ACCMIP exercise focusing on chemistry had only three model results that could be used to assess climate-air quality links (Schnell et al., 2016).

NTCF precursor emissions are also responsible for driving regional and local air quality (Fiore et al., 2012). This has led to the recognition that a combined mitigation policy for climate change and air pollution has clear economic benefits compared to separate mitigation (Clarke et al., 2014). Most, if not all, scenarios for the future actions of societies lead to changes in the emissions and meteorology that determine air quality and create pollution episodes. The exposure risks of human health and assets (agriculture, built environment, ecosystems) will be driven by daily variations in surface ozone and particulate matter in addition to deposition of nitrate and sulphate and any interactions of atmospheric and land-use changes. CMIP6 will provide comprehensive information on the future large-scale evolution of atmospheric composition thus updating the knowledge base used to manage air pollution.

The Aerosol and Chemistry Model Intercomparison Project, AerChemMIP, contributes to CMIP6 by diagnosing climate forcings and feedbacks involving NTCFs and chemically reactive WMGHGs (collectively, tropospheric aerosols and ozone, their precursors, methane, nitrous oxide, some halocarbons and impacts on stratospheric ozone), documenting and understanding past and future changes in the chemical composition of the atmosphere, and estimating the global-to-regional climate response from these changes.

Improving our insight and understanding of the inner functioning of climate models and climate itself is a primary motivation for AerChemMIP. To characterise the overall ERF from all species, including NTCFs, requires extra efforts and has not been properly done in earlier CMIP experiments (Forster et al., 2016). Climate feedback processes that involve changes to the atmospheric composition of reactive gases and aerosols may affect the temperature response to a given WMGHG concentration level. Better exploration and documentation of the changes in atmospheric composition in a changing climate is of interest for many other purposes such as choosing among different air quality mitigation options or understanding perturbations to ecosystems and biogeochemical cycles. Finally, uniform evaluation of the models will expose systematic biases and better constrain our overall goal of quantifying the role of aerosols and reactive gases on climate forcing.

1.2 Previous work

The contribution of tropospheric ozone precursors to radiative forcing (through changes in ozone and methane) has been considered in successive IPCC assessments since IPCC (1994) and the Second Assessment Report (IPCC 1996) where a combination of 2D and 3D chemistry models were used (PhotoComp in Olson et al. 1997). A more rigorous intercomparison of 3D chemistry transport models (OxComp in Prather et al. 2001; Gauss et al. 2003) provided information on the geographical distribution of ozone forcing for the IPCC Third Assessment Report (Ramaswamy et al. 2001). The IPCC Fourth Assessment report (AR4) (Forster et al. 2007) again used a multi model framework (Atmospheric Composition Change European Network – ACCENT) to calculate maps of ozone radiative forcing (Gauss et al. 2006). Here the models were still nearly all offline chemistry transport models, and none of the climate models used in AR4 (those participating in the CMIP3 project) included tropospheric ozone chemistry. The radiative forcing of ozone in all cases was calculated using offline radiative transfer models, usually for 'pre-industrial', 'present' and one or two future timeslices. It was not until the CMIP5 project that a few of the climate models included interactive tropospheric chemistry. The aim of the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) (Lamarque et al. 2013) was to quantify the contribution of ozone and aerosols to the radiative forcing in the CMIP5 models that included these components. In practice, the model setups for CMIP5 and ACCMIP tended to be different (in terms of resolution and complexity) so that ACCMIP was not able to fully characterise the forcings of most simulations submitted to the CMIP5 archive. ACCMIP combined the results from chemistry-climate models (CCMs) and offline chemistry transport models (CTMs) to quantify the central estimate and range of historical and future ozone and aerosol forcings, air quality, and the contributions of individual ozone precursor emissions. Surface ozone diagnostics in ACCMIP were used to evaluate CCM ability to match current air quality episodes and predict future ones (Schnell et al., 2015; 2016). NTCF forcings were diagnosed using a mixture of offline radiative transfer models and double call diagnostics, whereby a model radiation scheme is called twice with the second call containing one or all radiative species set to fixed values.

The historical and future climate effects of ozone depletion were first addressed in multi-model studies using CTMs in ACCENT (Gauss et al. 2006; Forster et al. 2007), focussing on changes in global radiative forcing. Son et al. (2008) highlighted the specific impact of the Antarctic ozone hole on regional surface climate by contrasting CMIP3 models with and without prescribed stratospheric ozone changes, and by comparing them to online CCMs from the SPARC Chemistry-Climate Model Validation Activity phase 1 (CCMVal-1; Eyring et al., 2007), with follow-on studies using model simulations from the CCMVal phase 2 (Son et al., 2010; SPARC 2010; WMO, 2010; 2014). Most of the CCMs included stratospheric chemistry only, while the newer generation CCMs available now encompass both tropospheric and stratospheric chemistry, with a few of these models being coupled to an ocean (John et al., 2012; Lamarque et al., 2013; Shindell et al. 2013c; Morgenstern et al., 2016). The most recent comprehensive assessment of the performance of these CCMs regarding stratospheric and tropospheric chemistry and dynamics is currently being performed within the SPARC/IGAC Chemistry-Climate Model Initiative (CCMI; Eyring et al., 2013a). In contrast to CMIP3 where half of the models prescribed a constant stratospheric ozone climatology, the CMIP5 models all considered time-varying ozone either prescribed or calculated interactively (Eyring et al., 2013b). This has led to substantial improvements in the representation of climate forcing by stratospheric ozone in climate models since the AR4 (Flato et al., 2013). The importance of "whole atmosphere" chemistry-climate coupling for the climate effects of ozone has also been recently highlighted, since changes in stratospheric ozone abundances, e.g. due to changes in ozone depleting substances, can affect tropospheric ozone through stratosphere-to-troposphere exchange (Shindell et al., 2013b; Banerjee et al., 2016).

The radiative forcing from historical aerosol emissions was quantified on the basis of one model (Langner and Rodhe 1991) in IPCC (1994) and the Second Assessment Report (IPCC 1996). The effects started to be included online in some climate models by the Third Assessment Report (Penner et al. 2001), but CTMs continued to play an important role even in the Fourth and Fifth Assessment Reports. Radiative forcing estimates for anthropogenic aerosol components were largely derived from a multimodel ensemble through the Aerosol Comparison (AeroCom) initiative (Schulz et al., 2006; Myhre et al. 2013b). AeroCom is a long-standing activity of aerosol model intercomparison (Textor et al., 2006, Kinne et al., 2006, Schulz et al., 2006), which provided estimates of radiative forcings from a large set of global aerosol (mostly offline) models for AR4 and AR5. The complex path from precursor emissions to aerosol loads, to optical and cloud-perturbing properties, to finally forcing has been simulated with well-documented diversity (Schulz et al., 2006; Myhre et al. 2013b, Ghan et al., 2016). More recently, aerosol forcing was quantified in ACCMIP, with many models using the same setups as used in CMIP5. These results were combined with additional simulations under CMIP5 to quantify the central estimate and range of historical to present-day aerosol forcing. Future forcing estimates relied solely on ACCMIP (Shindell et al., 2013a). It was problematic in CMIP5 to properly quantify the ERF by aerosols in the historic period and future scenarios, in most GCM models. This was mainly due to missing experiments, diagnostics and insufficiently characterised feedbacks involving natural aerosols.

Since the ERF calculations for ozone (tropospheric and stratospheric) and aerosols in ACCMIP were decoupled from the CMIP5 climate model simulations that informed the IPCC AR5 chapters on climate change (Bindoff et al., 2013; Kirtman et al., 2013; Collins et al. 2013), this made it difficult to relate the temperature responses to radiative forcing due to NTCFs and also to constrain the climate sensitivity. AerChemMIP is designed to fill in this information gap to inform IPCC AR6. The lessons learned in AeroCom, ACCMIP, and CCMI largely contributed to the design of AerChemMIP.

2 Science questions

This section provides the scientific justification for the four science questions that AerChemMIP will address. As a multi-model exercise, AerChemMIP will identify areas of consensus and disagreement in the answers. Owing to the strong connection between clouds and aerosols (Boucher et al., 2013), AerChemMIP will provide crucial information to support the World Climate Research Program (WCRP) Grand Challenge on “Clouds, Circulation and Climate sensitivity”. In addition, through the importance of natural sources on WMGHGs and NTCFs, the proposed research questions in AerChemMIP are well-aligned with the WCRP theme “Biogeochemical forcings and feedbacks”. The AerChemMIP proposal focuses on four broad questions, listed and discussed below.

2.1 How have anthropogenic emissions contributed to global radiative forcing and affected regional climate over the historical period?

Anthropogenic non-CO_2 emissions (e.g., NTCFs, in addition to other WMGHGs like halocarbons and nitrous oxide) have led to a climate forcing that is commensurate to the CO_2 forcing in some regions, especially over the last few decades (Myhre et al., 2013a). There are many couplings between different chemically and radiatively active species that remain to be fully understood. For example, in addition to its direct climate forcing of 0.48 ± 0.05 W m^{-2} (Myhre et al., 2013a), methane acts as a precursor to tropospheric ozone (Prather et al., 1994; Fiore et al., 2012), and is a dominant sink of the hydroxyl radical (OH), the primary tropospheric oxidising agent (Naik et al., 2013). As such, changes in methane emissions will also affect the lifetime of CH_4 and related gases (Prather, 1994) and the formation of aerosols through oxidation of anthropogenic and natural precursors (Shindell et al., 2009). Methane directly affects the chlorine chemistry of stratospheric ozone depletion (Pawson et al., 2014). Furthermore, methane is a source of water vapour in the stratosphere; this is an important contributor to stratospheric ozone depletion, especially away from the polar regions. Stratospheric water vapour is also a greenhouse gas and changes in methane thus have a further indirect radiative forcing (Myhre et al., 2013a). As the methane concentration has more than doubled since pre-industrial times (from 722 ± 25 ppb in 1750 to 1803 ± 2 ppb by 2011, it is imperative to quantify its historical forcing and the combined climate impacts associated with those changes. Furthermore the ERF due to ozone and aerosol changes since pre-industrial times is a key factor behind the large uncertainty in constraining climate sensitivity using observations over the historical record.

The ERF due to NTCFs has an inhomogeneous spatial distribution. The degree of regional temperature and precipitation responses due to such heterogeneous forcing remains an open question within the scientific community. There is also evidence that NTCFs, which are primarily located over Northern Hemisphere midlatitude land areas, have led to a larger climate response, both there and globally, relative to the more homogeneous ERF from WMGHGs (Shindell, 2014; Shindell et al., 2015; Rotstayn et al., 2015; Marvel et al., 2016). Climate response to regional forcings is not limited to the region of origin, and remote responses have been clearly demonstrated in numerical experiments (Teng et al., 2012; Levy et al., 2013; Bollasina et al., 2013; Shindell et al. 2015). In particular, the position of the ITCZ has been shown to depend on the differential rate of aerosol forcing between hemispheres (Hwang et al., 2013).

A detectable regional response to inhomogeneous climate forcing concerns the Southern hemisphere summertime surface circulation changes which have been induced by the Antarctic ozone hole as an indirect response to stratospheric ozone depletion from increasing halocarbons. These changes have been argued to lead to changes in the subtropical jet position, rainfall patterns, ocean circulation, and possibly sea-ice cover (Arblaster and Meehl, 2006; McLandress et al., 2011; Polvani et al, 2011). The relative role of these ozone-induced changes for observed Southern hemisphere summertime climate compared to other anthropogenic forcings and natural variability is not fully resolved by the scientific community with some contradictory studies in particular for the Antarctic sea-ice response (WMO, 2014). Hence there is a need for a multi-model ensemble of simulations that resolve stratospheric chemistry to isolate the role of stratospheric ozone depletion.

2.2 How might future policies (on climate, air quality and land use) affect the abundances of NTCFs and their climate impacts?

In the upcoming decades, policies that will impact atmospheric chemistry can be categorized in three broad areas: 1) climate change policies targeting mostly WMGHG emissions, 2) air quality policies targeting mostly NTCF emissions affecting tropospheric aerosols and ozone, and 3) land-use policies and practices. AerChemMIP aims to identify the patterns of chemical change at the global and regional levels, as well as the ERF associated with NTCF mitigation efforts (focusing on policy choices in areas 1 and 2 above) and their impact on climate (surface temperature and precipitation) and other environmental change (health, ecosystem, visibility etc.) between 2015 and 2055 (as the time frame over which aerosol and precursor emissions are expected to be significant; Shindell et al., 2012a; Fiore et al., 2015). Such impact analysis can be performed by contrasting two simulations: a) a reference with weak air quality policies and relatively high aerosol and ozone precursor emissions; and b) a perturbation experiment where strong air quality policies are applied, leading to much reduced NTCF emissions. These perturbations are designed in collaboration with ScenarioMIP to ensure that the NTCF perturbations are consistent with the underlying storylines (see section 3.2). A comparison of the reference and perturbation simulation will provide the background for understanding the effects of air quality policies over the next few decades. Analysis of results from these simulations will be critical to understand the interactions between NTCFs (aerosols in particular) and weather systems (Jacob and Winner, 2009; Leibensperger et al., 2012).

2.3. How do uncertainties in historical NTCF emissions affect radiative forcing estimates?

The primary focus of this question is to understand the sensitivity of present-day ERF to uncertainties in estimates of historical NTCF emissions. Indeed, while all proposed simulations rely on the usage of a central estimate, it is clear that there is a range of emission estimates (as discussed in Granier et al., 2011; Smith et al., 2011; Bond et al., 2013) that needs to be considered. While this uncertainty will clearly be region, sector and species dependent, it would be unrealistic to explore the full spectrum of variations. For that purpose, we will make use of perturbation (pre-industrial to present-day) simulations. This is likely to provide an upper bound on the impact of emission uncertainties. Results from the simulations can be directly compared to the simulations in section 3.1 and analysed for differences in radiative forcing as well as air quality and overall atmospheric composition. Inter-model differences will document their varying sensitivities to emissions.

2.4. How important are climate feedbacks to natural NTCF emissions, atmospheric composition, and radiative effects?

In a recent assessment of 28 modelled factors that could be a source of uncertainty in simulated cloud brightness, Carslaw et al. (2013) identified that, in their model, approximately 45% of the variance came from natural aerosols, especially from dimethysulfide (DMS) and volcanic SO_2 emissions. This can be compared with 34% of the variance due to anthropogenic aerosols. Additional studies have highlighted the role of marine biogenic aerosols (McCoy et al., 2015) and isoprene emissions (Archibald et al., 2010) in biogeochemical feedbacks. These are all examples of couplings and potential climate feedbacks involving diverse biogeochemical cycles, terrestrial (Isaksen et al., 2009; Arneth et al., 2010) and marine ecosystems (Cameron-Smith et al., 2011). AerChemMIP therefore proposes to quantify the climate impacts associated with specific biogeochemical cycles. To do this it will be necessary to quantify the climate response to the heterogeneous forcing patterns from naturally emitted short-lived species (the climate responses to WMGHGs are already covered in section 2.1). Six different feedbacks will be examined: 1) dust emissions, 2) sea salt emissions, 3) DMS emissions, 4) fire emissions, 5) NOx emissions from lightning, and 6) biogenic VOC emissions. Each will have a specified perturbation experiment. The comparison of each simulation with the CMIP DECK (Diagnostic, Evaluation and Characterization of Klima) pre-industrial control experiment (Eyring et al., 2016a) will enable a quantification of the importance of the considered climate-emission feedbacks.

3 Experimental design

The AerChemMIP experiments focus primarily on understanding atmospheric composition changes (from NTCFs and other chemically-active anthropogenic gases) and their impact on climate. We have devised a series of experiments that enable the forcing of various NTCFs to be contrasted with that of WMGHGs for historical and future climate change. In addition, the proposed chemistry-climate simulations will enable diagnosis of changes in regional air quality through changes in surface ozone and particulate matter. The effective radiative forcings are calculated from the net top of atmosphere flux difference between atmosphere-only simulations with identical SSTs, but differing composition (Forster et al. 2016; Pincus et al., 2016). The climate responses are calculated from the differences in climate between atmosphere-ocean simulations with differing composition.

AerChemMIP is designed to quantify the climate effects of interactive aerosols, tropospheric chemistry and stratospheric chemistry. Ideally participating models will include all three components; however, we realise that this may not always be possible or practical. Many CMIP models include emission-driven interactive aerosol schemes, but with limited or no further tropospheric chemistry. To understand their overall behaviour we encourage such models to participate in those AerChemMIP experiments that are relevant to quantifying the climate effects of the aerosols. For models with tropospheric chemistry, but lacking a stratospheric chemistry, we encourage participation in all experiments except those explicitly addressing the effects of halocarbons. For tropospheric-only chemistry, the CMIP6-specified stratospheric ozone dataset should be used (Hegglin et al., in preparation). Modelling groups with full chemistry and aerosol models are encouraged to perform all simulations they deem relevant to their objectives. Note that, for consistency, the concentrations of chemically and radiatively active species should be the same in the radiation and chemistry schemes. Tables 1 to 6 list the minimum model configurations required for each experiment. The suffix "CHEM^T" or "CHEM^S" indicates interactive tropospheric or stratospheric chemistry is the minimum needed for these experiments. The suffix "AER" indicates that interactive aerosols are needed. For models without interactive tropospheric chemistry, the Tier 1 ...*NTCF* and Tier 2 ...*Aer* simulations will be identical so only need to be run once. **Models capable of running with both interactive chemistry and aerosols should do so for all experiments.**

To participate in AerChemMIP, climate models must be run for the CMIP DECK and CMIP6 *historical* (atmosphere-ocean simulation with forcings evolving over 1850-2014) experiments with the same setup as in AerChemMIP, i.e., with the same levels of sophistication activated in the chemistry and aerosol schemes, and with the AerChemMIP diagnostics as specified in section 4. It is likely that groups will first spin up their model to pre-industrial conditions without interactive chemistry. This would then be followed with a shorter spin up with fully interactive chemistry for as long as is needed to ensure the chemistry does not introduce any additional drift. This process will be quicker if the non-chemistry spin up uses ozone and oxidants from a prior run of the interactive chemistry model (e.g., Collins et al. 2011). The length of the *piControl* (or *esm-piControl*) needs to be at least as long as the experiments. For AerChemMIP this is 205 years (164 years historical + 41 years future), although note that Eyring et al. (2016) recommend 500 years. It is necessary to have CMIP6 *historical* simulations with the same chemistry and aerosols as in AerChemMIP as this is used as the baseline. The AerChemMIP configured *abrupt-4xCO2* should be run for at least 150 years as recommended in Eyring et al. (2016), to quantify the climate-chemistry feedbacks. It is recommended that modelling groups document the aerosol and chemistry schemes in their climate model, and evaluate their performance in the DECK AMIP simulation.

The emissions of anthropogenic aerosols and reactive species are provided by Hoesly et al. (in prep) http://www.globalchange.umd.edu/ceds/ceds-cmip6-data/. Models should use their own schemes for natural emissions. The WMGHGs (methane, nitrous oxide, halocarbons) will be specified as CMIP-specified concentrations, either throughout the troposphere or at the surface.

We also realise that valuable contributions to answering the AerChemMIP scientific questions can be made by groups unable to participate in CMIP6, such as those running offline CTMs. Participation from these groups is welcomed and encouraged in the wider Chemistry-Climate Model Initiative (CCMI; Eyring et al., 2013b) and AeroCom projects, but the data will not form part of the official CMIP6 submission.

We have arranged the experiments into 3 Tiers to reflect their priority. Tier 1 experiments are those necessary to answer science questions 1 and 2 in terms of overall impacts of NTCFs and reactive well-mixed gases. Tier 2 experiments will answer question 4 and provide further detail on questions 1 and 2 by separating the effects of aerosol and ozone precursors. Tier 3 experiments will contribute to question 3 and provide additional detail and speciation. The total simulation years requested are 1265 for Tier 1, 1369 for Tier 2 and 270 for Tier 3, split between coupled ocean and fixed SST experiments. This includes 30 years for pre-industrial fixed SST control in common with RFMIP. In addition, models should have been run for the DECK experiments (501 years excluding control). Finally, modelling groups interested in studying the climate and air quality impacts of future emission reduction will need to perform a 3-member ensemble of SSP3-7.0 as described in ScenarioMIP (O'Neill et al., 2016) (41 years each member).

3.1 Historical (in support of Science question 2.1)

These experiments are designed to quantify the contributions of aerosols (tropospheric) and ozone (tropospheric and stratospheric) to climate change over the historical period. For aerosol and tropospheric ozone precursor emissions the model simulations will span 1850-2014, for halocarbons the simulations will span 1950-2014, since halocarbons only significantly increased after 1950. The latter experiments will enable the evaluation of the relative role of ozone depletion compared to other anthropogenic forcings and natural variability in determining past changes in southern hemisphere summertime climate (WMO, 2014). Methane and nitrous oxide have indirect climate effects as their chemistry affects tropospheric ozone production, stratospheric ozone chemistry, aerosol oxidation, and the lifetimes of each other.

The historical increases in aerosols and tropospheric ozone have inhomogeneous spatial distributions, and the degree of regional temperature and precipitation responses to such heterogeneous forcing remains an open question within the scientific community which these experiments and their pairwise ERF experiments described in Section 3.1.2 should help to answer. These will also enable the community to quantify whether NTCF emissions, which are primarily located over Northern Hemisphere midlatitude land areas have led to a larger climate response there, relative to forcing from WMGHGs (Shindell 2014). To distinguish between the warming effects of ozone and the net cooling effect of aerosols, further experiments separate the two groups. We choose to allocate NO_X to the ozone precursor group (as this is where it has the largest climate impact) even though it will generate both ozone and nitrate aerosol in models.

The experimental setup has been designed to pair coupled-ocean simulations with specified SST atmosphere-only experiments to calculate the ERFs due to each category of forcing agent (see section 3.1.2 for more details). Comparison between the temperature and precipitation changes in the coupled-ocean simulations with the ERFs (top of atmosphere, and surface) will provide information on the efficacies of the forcings to drive changes in climate.

3.1.1. Transient historical coupled-ocean simulations

These simulations parallel *historical* which is a simulation from 1850 to 2014 with all forcings applied (Eyring et al. 2016), and differ only by fixing the anthropogenic emissions or concentrations of a specified class of species. All other forcing agents must evolve as in *historical*. Perturbations to the total NTCF emissions (*hist-piNTCF*) or aerosol component (*hist-piAer*) start at the same point in the 1850 control as *historical* with the anthropogenic emissions fixed at that point. Perturbations to the halocarbon ODSs branch from the *historical* run at 1950 with chlorofluorocarbon and hydrochlorofluorocarbon concentrations fixed at that point (*hist-1950HC*), as specified in Table 1.

The individual climate signals from the proposed perturbations are likely to be small compared to internal climate variability, therefore we request at least three ensemble members for each experiment, using different ensemble members of *historical* as the starting points. The climate impacts of the anthropogenic emissions of NTCFs, aerosols and ODS halocarbons can then be diagnosed by subtracting the perturbed runs from the *historical* climate and evaluated against internal variability diagnosed from *piControl*. For models without interactive chemistry, *hist-piNTCF* and *hist-piAer* are identical and the same ozone climatology as *historical* should be used. Note that the climate signal from aerosols and reactive gas perturbations will include biogeochemical feedbacks via climate impacts on emissions, chemistry and transport of constituents, which requires further experiments and diagnostics describe below in section 3.4.

The total simulation years requested for this set of experiments are 684 for Tier 1 and 492 for Tier 2.

3.1.2. Transient historical prescribed SSTs simulations

In order to calculate the transient ERFs that drive the above climate changes, a set of simulations repeats the above sensitivity runs in 3.1.1, but using atmosphere-only configurations with prescribed sea-surface temperatures (SSTs) and sea-ice. The SSTs and sea ice should be specified as the monthly mean time-evolving values from one ensemble member of the historical simulations. This differs from the usual definition of ERF where the SSTs are specified to be a fixed repeating climatology throughout the simulation, usually taken from a preindustrial control experiment. Including evolving SSTs means that the underlying climate state is consistent with the *historical* simulation that is used as the reference for all these experiments. Use of historical SSTs rather than pre-industrial will eliminate any effects of using an inconsistent background climate state (such as different cloud cover and natural emissions) that could affect concentrations of aerosols and reactive species and the transient ERFs. The impact of background state on the diagnosis of ERFs is likely to be small (Forster et al., 2016). This is further discussed in section 5.1. The control simulation (*histSST*) uses prescribed historical SSTs with all other components as *historical*. This is not likely to be significantly different from the coupled-ocean historical simulation, but we request this simulation for consistency with the rest of the experiments. There are Tier 2 experiments to calculate separately the transient ERFs from aerosol and ozone precursors, as specified in Table 2.

The total simulation years requested are 556 for Tier 1 and 492 for Tier 2.

Historical changes in methane and nitrous oxide abundances have altered atmospheric chemistry and the NTCF radiative forcing. These indirect effects are complex and have previously been calculated in piecemeal ways (Myhre et al. 2013a). In AerChemMIP, the historical transient ERFs will be calculated for models that have reactive gas chemistry (at least tropospheric). The transient ERF pattern from these simulations is expected to be relatively homogeneous, although their chemical effects on ozone and secondary aerosols may be less so. Therefore, AerChemMIP does not include any experiments to derive the climate responses to methane or nitrous oxide forcing. The climate response to homogeneous forcing is quantified in Detection and Attribution Model Intercomparison Project (DAMIP) from the *hist-GHG* simulation (as *historical*, but only the WMGHG forcings evolve).

The transient ERFs for each species or group of species will be diagnosed by subtracting the top of atmosphere and surface radiative fluxes in the perturbed runs from those in histSST. For models without interactive chemistry, *histSST-piNTCF* and *histSST-piAer* are identical, and the same ozone climatology as *histSST* should be used.

3.2. Future simulations (in support of Science question 2.2)

AerChemMIP further aims to identify the patterns of change in surface temperature and precipitation at the global and regional levels associated with future NTCF mitigation efforts focusing on air pollutant species. These experiments cover on the time frame from 2015 to 2055, as this is when reductions in aerosol and ozone precursor emissions are expected to be significant, at least for some regions. The future scenarios are based on Shared Socio-economic Pathways (SSPs) as described in O'Neill et al. (2014) and van Vuuren et al. (2014). As well as socio-economic scenarios, the SSPs include representations of different levels of controls on air quality pollutants – weak, medium and strong (Table 3). The medium strength of pollution control corresponds to following current legislation (CLE) until 2030 and progressing three-quarters of the way towards maximum technically feasible reduction (MTFR) thereafter. The rate of progress is different for high, medium and low-income countries. Strong pollution control exceeds CLE and progresses ultimately towards MTFR. Weak pollution controls assume delays to the implementation of CLE and make less progress towards MTFR than the medium scenario. For more details, see Rao et al. (2016). To detect the largest signal we choose the reference scenario to be SSP3-7.0 "Regional Rivalry" without climate policy (7.0 Wm^{-2} at 2100, experiment *ssp370*), see Fujimori et al. (2016), as this has the highest levels of short-lived climate pollutants and "Weak" levels of air quality control measures (O'Neill et al. 2016; Rao et al. 2016). The *ssp370* ScenarioMIP simulation will need to have been run with the AerChemMIP setup and diagnostics, or repeated here. The data for the perturbation experiment to this within AerChemMIP will be generated by the AIM group using the same socio-economic scenario as in Fujimori et al. (2016), but with "Strong" levels of air quality control measures SSP3-7.0-lowNTCF. The energy use and levels of climate mitigation are the same in these two scenarios. Levels of WMGHGs (including methane) will be unchanged unless they are directly affected by the air quality control measures. Differences in climate, transient ERF, chemical composition and air quality between the two scenarios will be solely due to the alternative air quality control measures. SSP3-7 does run out until 2100 as part of ScenarioMIP but, to save computation expense, AerChemMIP is only requesting simulations out to 2055. This is the time period over which the divergence in air quality policies is expected to be largest.

3.2.1. Transient future coupled ocean

The two transient future coupled ocean experiments start in January 2015 from the end of *the historical* simulation and are run for 41 years (to December 2055) following the SSP3-7 and SSP3-7-lowNTCF scenarios for WMGHG and NTCF emissions (experiments *ssp370* and *ssp370-lowNTCF*, see Figure 1 and Table 4). Note the *ssp370* reference simulation is the same as in ScenarioMIP; it is therefore required that the ScenarioMIP *ssp370* is performed using the same model configuration as for AerChemMIP. The climate and air quality signals will be derived by subtracting the experiment ("clean", see Tables 4,5) from the reference. This signal is expected to be globally small (commensurate with a forcing on the order of 0.1 Wm^{-2}, although much larger locally), so at least 3 ensemble members for both SSP3-7 and SSP3-7-lowNTCF variants are requested as continuation from existing historical ensemble members. Where natural emissions are modelled interactively, these will vary with the evolving climate and will differ between ssp370 and ssp370-lowNTCF as the climate diverges.

The total simulation years requested are 123 (all Tier 1).

3.2.2. Transient future prescribed SSTs simulations

As for the historical experiments (section 3.1), the above scenarios (SSP3-7.0 and SSP3-7.0-lowNTCF) are repeated with prescribed SSTs. These SSTs (and sea ice) are taken from the monthly mean evolving values from one of the ensemble members of the coupled *ssp370* run. The differences in radiative fluxes between the reference and "clean" simulations will give the TOA and atmospheric transient ERFs. Comparison between the magnitudes and patterns of transient ERF with surface temperature and precipitation from the previous coupled model simulations will provide quantification of the efficacy of the NTCFs to affect climate.

The contributions of the different groups of NTCFs to future climate will be quantified by further simulations in which only a subset of the emissions (aerosols, ozone precursors, black carbon) follow the "clean" scenario with the rest following the reference SSP3-7.0. All these perturbation experiments are Tier 2 (see Figure 2 and Table 5). An additional scenario SSP3-7.0-lowCH4 will differ from the control SSP3-7.0 only in using lower emission factors for methane.

The *ssp370SST-ssp126Lu* experiment will study the atmospheric chemical impacts of land-use changes through natural emissions (biogenic VOCs, fire, dust) and surface uptake for models that include interactive schemes for emission and deposition. Not all models will model all these processes interactively. The simulation will be parallel to *ssp370SST* using the same WMGHGs and anthropogenic NTCF emissions, but with land use specified according to the SSP1-2.6 scenario. The transient ERFs calculated with respect to the *ssp370SST* control will include the effects of albedo changes as well as NTCFs. The pair *ssp370SST* - *ssp370SST-ssp126Lu* are the prescribed-SST equivalents of the coupled-ocean Land Use Model Intercomparison Project (LUMIP) pair *ssp370* - *ssp370-ssp126Lu* (Lawrence et al. 2016).

The total simulation years requested are 82 for Tier 1 and 205 for Tier 2,.

3.3. Timeslice historical ERF simulations (in support of Science questions 2.1 and 2.3)

The quantification of pre-industrial to present day ERFs due to different drivers (such as in Myhre et al. 2013a, figure 8.17) is used widely. The AerChemMIP timeslice experiments will provide the data to generate a consistent table of present day ERFs for the reactive gases and aerosols. The ERFs are calculated by comparing the change in net TOA radiation fluxes between two runs with the same prescribed SSTs, but with NTCF emissions or WMGHGs (methane, nitrous oxide, halocarbon) concentrations perturbed from their preindustrial to present day values. Internal variability (mainly clouds) generates considerable interannual variability in ERFs; therefore, at least 30 years of simulation are needed to characterize the present day ERF from some species (Forster et al, 2016.). These simulations differ from the transient ERF simulations in

3.1.2 in that they use pre-industrial SSTs and maintain the same emissions (or concentrations) for 30 years. They therefore give a more accurate representation of the pre-industrial to present ERF than would be obtained from using portions of the transient historical ERF simulations.

The control simulation for these experiments will use 1850 concentrations of WMGHGs and emissions of NTCFs, run for 30 years in atmosphere-only mode with SSTs and sea-ice prescribed as a (monthly-varying) climatology taken from 30 years of the pre-industrial control (experiment piClim; see Table 6) following the Radiative Forcing Model Intercomparison Project (RFMIP) specification (Pincus et al., 2016). Provided this experiment is run with the same interactive chemistry and aerosols as the model configuration contributing to AerChemMIP, this will be the same control as in RFMIP. The TOA radiative fluxes from this control are expected to be very similar to the climatology from the coupled pre-industrial control. However, this extra simulation ensures consistency with the ERF definition and with RFMIP.

The perturbation experiments are run for 30 years following the control, using the same control SST and sea-ice, but with the concentrations (for WMGHGs) or emissions (for short-lived species) of the selected species set to present day (2014) values (Table 6). The WMGHG experiments should allow as complete a representation of the chemical effects on aerosol oxidation, tropospheric and stratospheric ozone, and stratospheric water vapour as the models are capable. Note that in this setup methane concentrations are fixed and do not respond to changes in oxidation rate. The ozone ERF estimates are not broken down by their location (tropospheric or stratospheric), but whether they are driven by changes in ozone precursors or ODSs. Models without interactive chemistry should only run the aerosol specific experiments, and use the same ozone climatology as *piClim*.

The total simulation years requested are 120 for Tier 1, 120 for Tier 2 and 150 for Tier 3.

3.4. Natural emissions simulations (in support of Science question 2.4)

Climate change will affect the natural emissions of NTCFs and reactive WMGHGs. These natural emissions will have a radiative effect and so feedback on to climate change. To simplify the experimental setup, the experiments detailed here simply double the natural emissions. The radiative effects of natural WMGHGs (e.g., methane from natural sources) are not calculated as these can be obtained from experiment *piClim-CH4*. The control simulation is the 30 year 1850 fixed-SST piClim as for the timeslice ERFs. Each experiment parallels the 30 year control except the emission fluxes from an interactive parameterization are doubled (see Table 7). For models that do not interactively parameterize particular emissions, the fluxes from the 1850 climatological dataset should be doubled.

The radiative perturbation from these experiments will give ERF per Tg yr^{-1} change in emissions. When scaled by the simulated changes in emission fluxes per K temperature change from either the DECK 4xCO_2 or 1% yr^{-1} CO_2 simulations, these determine the feedback parameter given as Wm^{-2} per K in surface temperature.

The total simulation years requested are 60 for Tier 2 and 120 for Tier 3.

4 Diagnostics

The AerChemMIP specific diagnostics are designed to answer the following questions: How large are forcing, feedback and response associated with reactive gases and aerosols in the models participating in CMIP6 historical and scenario simulations? Which processes and mechanisms need to be represented in the models for a credible description of climate-chemistry-aerosol interactions? How well do models reproduce the observed spatial distribution and historical evolution of NTCF concentrations, depositions, optical properties, and observable interactions with climate?

To guide the diagnostic process, the data request is structured according to overarching analysis subjects. These are detailed in the subsections below: Climate response, Forcing, Feedbacks, Chemistry-Climate Interactions, Air Quality, and Evaluation of model performance. Considerable experience has been gained in previous model intercomparison exercises (namely CCMVal, CCMI, AeroCom, ACCMIP, Hemispheric Transport of Air Pollution (HTAP) and CMIP5), but all too often model versions were different from those used in CMIP. AerChemMIP provides a unique opportunity to generate a complete data set, requested directly from those GCMs providing climate sensitivity and scenario information to CMIP6. A specific problem may be the expected diversity in model complexity, as mentioned in section 3. Models may contain interactive aerosols, tropospheric chemistry, stratospheric chemistry and any combination of these. AerChemMIP requests all output unless unavailable from an individual model configuration with good reason.

The diagnostics requested for the AerChemMIP experiments are assembled in two Excel sheets (available at https://wiki.met.no/aerocom/aerchemmip/diagnostics), and the definitive and detailed request will be found in the CMIP6 data request (https://earthsystemcog.org/projects/wip/CMIP6DataRequest; Juckes et al., in preparation). Since the AerChemMIP model versions are requested to also perform the DECK experiments, the data request contains suggestions for output limitations for these experiments (see for details final data request). Here we provide an overview along the analysis subjects mentioned above. Suggestions for best practice of diagnosing processes and outputting variables are given in some cases, in particular where previous model intercomparison projects failed to harmonize model output. The specific AerChemMIP request is grouped in 8 tables and these are to be found in the excel file and the CMIP6 data request: aerfixed, aermonthly-3d, aermonthly-2d, aerdaily, aer-6h, aerhourly, aerzonal-vert, aerzonal. Modellers are asked to read the explanatory notes found for each CF standard name on the CF website and the specific explanatory remarks in the AerChemMIP data request and corresponding excel worksheet.

4.1 Climate response

The characterisation of the climate response to NTCF forcing requires a set of diagnostics, which are fairly standard to all CMIP experiments. They include the variables that altogether describe the state of the atmosphere, the ocean and cryosphere, land surfaces including essential biosphere and carbon cycle parameters.

Specific attention should be devoted to the COSP simulator, which AerChemMIP models are encouraged to install. To facilitate the exploitation of A Train satellite data in numerical models, the COSP system has been developed that allows simulation of the signal that CloudSat/CALIPSO would see in a model-generated world. A better understanding of cloud and aerosol interactions may be possible if models add specific diagnostic aerosol calls, which would allow analysis together with the COSP diagnostic package output. Of particular interest is the observable aerosol backscatter and extinction coefficient, which provides, since the arrival of the CALIOP satellite lidar in the A-train, a constraint for the global 3D distribution of aerosols. Therefore modellers are asked to provide 3D 6-hourly fields of aerosol extinction and backscatter coefficient for one realisation of the DECK Atmospheric Model Intercomparison Project (AMIP) experiment (1979-2014).

4.2 Forcing

For a better documentation of which forcing is actually present in a given climate model, several sets of diagnostics are needed: 1) flux parameters providing ERF from fixed-SST simulations, 2) 3D mass mixing ratios and optical thickness in transient simulations, and 3) repeated aerosol-free calls to the radiation code (with aerosol scattering and absorption set to zero) in transient climate simulations (Ghan et al., 2012). This allows characterisation of the radiative forcing of the aerosol radiation interaction and separation of it from the aerosol cloud interactions and rapid adjustments. This is strongly recommended for the reference historical simulation and ERF time slice experiments aimed at diagnosing aerosol forcing, but not essential if it is not possible to implement. Combined with additional RFMIP diagnostics, this will generate a fairly complete forcing characterisation.

ERFs of gases and aerosols will be derived in the 30 year long fixed-SST simulations (see section 3) by investigating clear-sky and all-sky shortwave and longwave fluxes at the top of the atmosphere and at the surface. In the same simulations a range of auxiliary variables are requested which characterise emissions, 3D mass or molar mixing ratios, aerosol optical properties and cloud properties to complement the radiative fluxes with actual composition diagnostics. With these, forcing efficiencies may be established to be used for the interpretation of transient simulations.

Characterising the atmospheric forcing in transient simulations of a fully coupled model poses problems, because the climate system response alters cloud cover, lapse rates, and even cryosphere and land surface properties. To first order and for most components atmospheric forcing is proportional to the amount of the species perturbing the pristine atmosphere – with the important exception of interactions involving clouds. 3D fields of mass mixing ratios and column integrated optical thickness for aerosol species allow tracking of the extent to which perturbations are present. For gaseous pollutants ozone molar mixing ratios and methane lifetime are requested in order to diagnose forcing offline.

Tracking the anthropogenic fraction of column loads would require additional tracers. For a first order analysis, we will use anthropogenic-only emissions to compute the average anthropogenic fraction in transient simulations, as compared to the preindustrial reference in 1850. Some emissions may include natural components, which may have changed along with climate change, such as NO_X from lightning and soil degradation and biogenic volatile organic compounds. We therefore request total emissions for each species for 1850 and 2014.

The 3D fields of mass mixing ratios provide also the vertical distributions, which are useful for the understanding of forcing components (semi-direct, direct, cloud-aerosol interaction) of heterogeneously distributed species - in particular black carbon and other anthropogenic aerosols. Indeed, positive forcing (warming) and the semi-direct of black carbon have received recently more attention. 3D fields of black carbon mass mixing ratios and column integrated absorption optical depth have been used widely for analysing the black carbon forcing efficiency.

A source of confusion has been the aerosol optical thickness (AOT) diagnostics (e.g., Flato et al. 2013 figure 9.29) in CMIP models. Natural (particularly dust and sea salt) and anthropogenic aerosols have been blended together in different ways in models. We request output to diagnose the different aerosol species contributions to total AOT, and to provide more insight into the reasons for differences in AOT between models. If possible AOT should be output for sulphate, organic matter, dust, sea salt, black carbon, nitrate at ambient relative humidity. In the case of internal mixed aerosol modes, total AOT shall be distributed according to the volume of the dry aerosol species present in the mixed aerosol mode. The sum of speciated AOT from all species simulated should be equal to total AOT at any given point in time and space.

Providing 3D fields of mass mixing ratios and AOT consistently will allow analysis of differences in aerosol optical property calculations, as well as changes in aerosol humidity growth in a changing climate and with changing emission patterns. In combination with emissions in transient climate simulations these diagnostics will allow complete analysis of feedback processes, see below.

Another problem is which ambient humidity in the model is picked to compute AOT. Some models compute an all-sky AOT, including AOT in cloudy fractions with high humidities, while others restrict output to clear-sky AOT. The latter is preferred here, because it may be compared to that AOT which is observed under clear-sky conditions from satellites and sun photometers. Aerosol radiation interactions are also most effective in clear-sky scenes and it is thus more relevant to base forcing efficiencies on clear-sky AOT. If models compute normally an all-sky AOT using high relative humidities in cloudy fractions of the grid box, they are asked to compute also a clear-sky AOT (od550csaer) using clear sky relative humidities.

Aerosol-cloud interactions are still among the most uncertain of forcing components. Here we have selected rather standard parameters which are also used in the Cloud Feedback Model Intercomparison Project (CFMIP) and which allow for a 3D characterisation of cloud fraction, cloud liquid water path and cloud as well as ice number concentrations. The fixed-SSTs approach can further be applied with additional radiation calls to diagnose the various aerosol-cloud effects (Ghan et al., 2012). Most models will provide all-sky and clear-sky radiative fluxes, by computing fluxes with a repeated double call to the radiation routine neglecting cloud scattering. Here we propose a repeated “aerosol-free” call invoking this all-sky / clear sky double call to radiation once more, by setting the aerosol scattering and absorption properties to zero. Fluxes for this repeated call have to be stored separately (variables rsutca, rsutcsca, rlutca, rlutcsca). To limit computational burden we propose to invoke this call during the DECK historical simulation and the *piClim…* experiments. In all cases, separate diagnostics for shortwave and longwave changes are applied.

4.3 Feedbacks

Feedback processes will change natural emissions of reactive gases and aerosols. The short-lived nature of dust, sea salt, biogenic gases and aerosols as well as reactive nitrogen components and ozone will exert a rather rapid feedback loop if triggered. A thorough documentation of natural emissions and 3D fields of reactive gases and aerosols is needed. To relate natural emission changes to forcing, specific experiments are designed in AerChemMIP, which resemble the ERF experiments (see table 7). In these feedback experiments radiation flux and cloud variables are requested as in the ERF forcing experiments. Other variables such as those characterizing aerosol and cloud optical properties, land-sea-ice distribution will help with the analysis of processes (fires, dynamics, volcanic perturbation, land cover change, sea ice change) involved in the feedback process. Deposition of nutrients such as nitrogen and dust has been suspected to be involved in feedback processes (Collins et al. 2011). Their output is thus requested too.

4.4 Chemistry-climate interactions

Chemistry-climate interactions involve impacts of composition on climate (as discussed in 4.1), but also crucially the effect of climate change on atmospheric composition, which happens through both changes in transport and chemistry. The availability of stratosphere-troposphere resolving chemistry-climate model simulations thereby provides the unique opportunity to look at these chemistry-climate interactions in a more comprehensive way than what was hitherto possible based on the models contributing to CMIP5. The DECK Control and 1%/yr CO_2 runs will be particularly valuable for this. The stratosphere has been identified as potentially important contributor to model-differences for both tropospheric ozone (Young et al., 2013) and the OH budget (Voulgarakis et al., 2013). Particular focus within AerChemMIP will be placed on the study of how physical climate parameters such as temperature, wind, clouds, and precipitation affect tropospheric composition and the oxidizing capacity of the atmosphere in addition to changing emissions. To study and disentangle key processes that lead to model differences, a comprehensive list of monthly mean 3D-output of key meteorological parameters (standard variables temperature and precipitation, convective mass fluxes (mcu), lightning NOx production (emilnox) and chemical species (also annual loss terms of methane, carbon monoxide, and nitrous oxides) is requested. In addition, two transport tracers will help to track changes in tropospheric transport between hemispheres, the artificial tracers called aoa_nh and nh_50, the first one with a uniform source [1year/year], constant in space and time, above the surface layer, 30-50N, the second one applying a uniform surface mixing ratio [100 pbbv], 30-50N, with a 50 day exponential decay (see definitions in the CCMI-1 data request at http://blogs.reading.ac.uk/ccmi/data-requests-and-formats/) . In the coupled troposphere-stratosphere system, climate will affect tropospheric composition and also its oxidation capacity through changes in the stratospheric circulation and resulting changes in the stratospheric distribution of ozone and stratosphere-troposphere transport of ozone (Collins et al. 2003; Stevenson et al., 2006; Hegglin and Shepherd, 2009). A tagged stratospheric ozone variable (o3ste) is defined to diagnose stratosphere-troposphere exchange, with the simulations *hist-1950HC* and *histSST-1950HC* designed to help disentangle the impact of ODSs and climate change on the stratospheric influence on tropospheric composition. Loss terms of CO, methane, ozone and nitrous oxide are suggested to help interpretation of their budget in a changing climate.

4.5 Air Quality

The simulations in the AerChemMIP provide the opportunity to retrieve from historical and scenario runs air quality related parameters which relate the broadly used CMIP emissions to a transient description of climate to air quality metrics. An ensemble of models can be used to establish consequences for air quality. Most interest is on particulate matter concentrations and high ozone peaks. Since air pollution standards have been defined as exceedances for a given time window, we request hourly data at surface level for some few key substances, such as ozone, $PM_{2.5}$ and NO_2. From these frequencies of daily maximum, diurnal cycles in different climate regimes, boundary layer characteristics can be obtained, which do characterise the model also in terms of chemical reactivity but also with respect to boundary layer mixing. The output is required for the fully coupled *historical* and the two coupled SSP3-7.0 scenarios.

4.6 Evaluation of model performance

In addition to the above diagnostics focussed on the science questions, some variables will be used to specifically help to evaluate model performance. Testing model behaviour against observations is critical for gaining confidence in their simulation of the historical past and predictions of the future. The output requested refers to variables that have been observed by different observational networks (based on ground-based, balloon, aircraft or satellite sensors) over the recent past. Comparison to these data will enable investigation of model bias, but may also help to rank models with respect to their ability to capture critical variability (see for example SPARC CCMVal, 2010).

The diagnostics requested represent a subset of the diagnostics requested for the AeroCom and CCMI model comparison activities. These include 2D hourly (surface level ozone, $PM_{2.5}$, and NO_2), 3D monthly mean concentrations of aerosol species, ozone and ozone precursors (including methane, CO, NO_2, OH, and VOCs), column data (ozone), AOTs at different wavelengths, and deposition rates (including wet and dry deposition of nitrates and sulphates, dust, and BC). The hourly and 6 hourly model output (contained in aerhourly and aer-6h tables) is requested specifically for the DECK AMIP simulations (1979-2014), and are not needed for other experiments. The other variables are included as being essential for forcing, feedback, chemistry climate interactions and air quality analysis. Outputting such variables will hence serve multiple purposes.

Of particular importance are variables as assembled in the framework of Global Atmospheric Watch (GAW) and observations for Model Intercomparison Projects (obs4MIP) (Teixeira et al., 2014; Ferraro et al., 2015). In addition, other datasets such as from the SPARC Data Initiative (Hegglin et al., in preparation) for the stratosphere and from the ESA CCI (Hollmann et al., 2009) for the troposphere will be valuable for comparisons. Note, both CCMI and AeroCom will contribute with their model evaluation experience and will feed selected observational data sources into the Earth System Model Evaluation Tool (ESMValTool; Eyring et al., 2016b; Righi et al., 2015). The ESMValTool will run - together with other evaluation tools such as the PCMDI metrics package (PMP, Gleckler et al., 2016) - alongside the Earth System Grid Federation (ESGF) as soon as the output is submitted to the CMIP archive so that evaluation results can be made available at a time much faster than in CMIP5 (Eyring et al., 2016c). This will include the evaluation of chemistry and aerosols in the CMIP DECK and CMIP6 historical simulations.

5 Relations with other MIPs

AerChemMIP is self-contained in so far as the questions posed can be answered by running only the experiments listed here, the DECK, and *historical*. For a full analysis of the past and future climate-composition interactions (including unreactive greenhouse gases) in the CMIP6 chemistry climate models we recommend that as many as possible of the Tier 1 simulations of RFMIP (Pincus et al., 2016), DAMIP (Gillett et al. 2016), ScenarioMIP (O'Neill et al., 2016), C4MIP (Jones et al. 2016) and LUMIP (Lawrence et al., 2016) are run with the AER CHEM model configuration and with AerChemMIP diagnostics.

5.1 Radiative Forcing MIP (RFMIP)

There are considerable synergies between AerChemMIP and RFMIP. RFMIP addresses the ERF due to all drivers for the historical and future periods. AerChemMIP specifically looks to quantify the ERFs for reactive species and aerosols, and to separate individual components of these. RFMIP also contains other components related to the assessment of model radiation code performance and simulations with prescribed aerosol distributions and aerosol optical properties for *historical* following a similar philosophy to the "Easy Aerosol" project.

For the prescribed-SST experiments to diagnose transient ERFs, RFMIP uses the pre-industrial conditions as the reference and perturbs one group of species at a time to evolve following *historical* (e.g. in *piClim-histaerO3*); consequently RFMIP specifies a pre-industrial SST and sea ice climatology. AerChemMIP uses the evolving conditions (*historical*) as the reference with one group of species perturbed back to the pre-industrial conditions (e.g. in *histSST-piNTCF*); therefore AerChemMIP specifies a time evolving monthly SST and sea ice distribution taken from a coupled *historical* experiment. The impacts of different approaches for specifying SSTs and sea ice on the total ERF over the satellite era have been estimated to be small in one climate model (Forster et al., 2016).

For models with interactive chemistry, *piClim-NTCF* and *piClim-aerO3* are identical and only need to be run once. For models without interactive chemistry, RFMIP specifies that *piClim-aerO3* uses the present day tropospheric and stratospheric ozone climatology, whereas AerChemMIP specifies that *piClim-NTCF* uses the control (pre-industrial) ozone climatology.

5.2 Detection and Attribution MIP (DAMIP)

There is some overlap between the AerChemMIP coupled model experiments and those requested in DAMIP. For example, AerChemMIP requires the extra *historical* runs from DAMIP to increase the ensemble size to at least 3 members.

The DAMIP historical experiments use the pre-industrial as the control and have one class of species evolving at a time (e.g. *hist-aer* parallels *historical*, but with only aerosol forcing evolving) with all others fixed to pre-industrial levels. In contrast, AerChemMIP uses *historical* as the control and fixes one class of species at a time to pre-industrial levels (e.g. *hist-piAer*).

The DAMIP *hist-stratO3* run has only stratospheric ozone evolving, using either the prescribed CMIP6 ozone dataset (Hegglin et al., in prep.) or ozone output from the previous historical run. In the troposphere ozone is fixed to pre-industrial levels. The equivalent experiment for the models with stratospheric chemistry in AerChemMIP fixes halocarbon concentrations at 1950s levels allowing the model chemistry to generate the difference in ozone compared to *historical*. DAMIP will use SSP2-4.5 as its future scenario, with *ssp245-ghg* and *ssp245ae*r as variants, so this does not overlap with the AerChemMIP future experiments.

5.3 Other MIPs

The future scenario SSP3-7.0 (experiment *ssp370*) is prescribed as a Tier 1 scenario in ScenarioMIP with extra ensemble members as Tier 2. A total of three members (using the AerChemMIP model configuration) are required as the baseline for the AerChemMIP future experiments. We recommend that the AER CHEM configuration with AerChemMIP diagnostics be used for as many as possible of the other ScenarioMIP experiments in order to understand the range of possible future evolution of aerosols, reactive gases and surface air quality concentrations.

The future land-use ERF calculations (*ssp370SST-ssp126Lu*) in AerChemMIP (section 3.2.2) parallel the full climate land-use perturbation (*ssp370-ssp126Lu*) in LUMIP. If the same model configurations are used for both, this will allow direct quantification of the efficacy of land-use changes.

We recommend that the AER CHEM configuration with AerChemMIP diagnostics be used for the C4MIP 1% yr^{-1} CO_2 RAD and BGC experiment in order to explore fully the biogeochemical couplings involving aerosols and reactive gases.

6 Summary

Advances in climate model development mean that for CMIP6 a larger set of climate models will include interactive simulation of aerosols than at the time of CMIP5, and many will include interactive chemistry of the troposphere and/or stratosphere. AerChemMIP has therefore been designed to quantify the effects of these NTCFs and reactive WMGHGs on climate and also on atmospheric composition and surface air quality.

A focus is on comparing the climate responses (both global and regional) to the heterogeneous forcing patterns generated by changes in emissions of NTCFs and their precursors over the historical period and in future scenarios. The future scenarios consist of a pair differing only in their levels of ambition in air quality policy. The results from these will provide information on the impacts of air quality policies on climate. The forcings are characterised by the ERFs using model simulations with fixed SSTs and sea ice; the responses are characterised by changes in surface temperature and precipitation (amongst others) using model simulations with coupled oceans.

AerChemMIP will identify the contributions to the present day climate (in terms of ERF) made by aerosol emissions, tropospheric ozone production, stratospheric ozone depletion, and changes in the reactive gases methane and nitrous oxide. To add to the forcing-response relationships AerChemMIP will also provide information on climate feedbacks by calculating the radiative effect of natural emissions of aerosols or ozone precursors in the same way as ERFs of anthropogenic species. Combining these ERFs with diagnosed changes in natural emissions from the DECK 1% yr^{-1} CO_2 (1pctco2) or 4×CO_2 (abrupt4co2) simulations would give the climate feedback parameters. AerChemMIP is therefore key to understanding the behaviours of models with aerosols and chemistry in CMIP6 and we would encourage all such models to participate.

Data Availability

The climate model output from AerChemMIP experiments described in this paper will be distributed through the Earth System Grid Federation (ESGF) with DOIs assigned. As in CMIP5, the model output will be freely accessible through data portals after registration. In order to document CMIP6's impact and enable ongoing support of CMIP, users are obligated to acknowledge CMIP6 and the participating modelling groups (see details on the CMIP Panel website at http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip). In order to run the experiments, datasets for natural and anthropogenic forcings are required. These forcing datasets are described in separate invited contributions to this Special Issue. The forcing datasets will be made available through the ESGF (https://pcmdi.llnl.gov/projects/input4mips/) with version control and DOIs assigned.

Acknowledgements. CRESCENDO project members (Bill Collins, Michael Schulz, Olivier Boucher, and Veronika Eyring) acknowledge funding received from the European Union's Horizon 2020 research and innovation programme under grant agreement No 641816 (CRESCENDO). This work benefitted also from the Norwegian research council projects #235548 (Role of SLCF in Global Climate Regime) and #229796 (AeroCom-P3).

References

Arblaster, J. M., and Meehl G. A.:, Contributions of external forcings to southern annular mode trends, J. Clim., 19, 2896–2905, doi:10.1175/JCLI3774.1, 2006

Archibald, A. T., Cooke, M. C., Utembe, S. R., Shallcross, D. E., Derwent, R. G., and Jenkin, M. E.: Impacts of mechanistic changes on HO_x formation and recycling in the oxidation of isoprene, Atmos. Chem. Phys., 10, 8097-8118, doi:10.5194/acp-10-8097-2010, 2010.

Arneth, A., Harrison, S. P., Zaehle, S., Tsigaridis, K., Menon, S., Bartlein, P. J., Feichter, J., Korhola, A., Kulmala, M., O'Donnell, D., Schurgers, G., Sorvari, S., Vesala, T: Terrestrial biogeochemical feedbacks in the climate system, Nature Geosci., 3, 525–532, 2010

Banerjee, A., Maycock A. C., Archibald A. T., Abraham N. L., Telford, P., Braesicke P., and Pyle J.A.:. Drivers of changes in stratospheric and tropospheric ozone between year 2000 and 2100, Atmos. Chem. Phys., 16, DOI: 10.5194/acp-16-2727-2016, 2016

Bindoff, N.L., et al., Detection and Attribution of Climate Change: from Global to Regional. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* edited by Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Bollasina, M. A., Ming, Y., and Ramaswamy V.: Earlier onset of the Indian monsoon in the late twentieth century: The role of anthropogenic aerosols, Geophys. Res. Lett., 40, 3715–3720, doi:10.1002/grl.50719, 2013

Bond, T. C., et al.: Bounding the role of black carbon in the climate system: A scientific assessment, J. Geophys. Res., 118, 5380–5552, doi:10.1002/jgrd.50171, 2013

Boucher, O., et al., Clouds and Aerosols, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, pp. 571-657, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. 2013

Cameron-Smith, P., Elliott S., Maltrud M., Erickson D., and Wingenter O.: Changes in dimethyl sulfide oceanic distribution due to climate change, Geophys. Res. Lett., 38, L07704, doi:10.1029/2011GL047069, 2011

Carlton, A.G., Pinder R. W., Bhave P. V., and Pouliot G. A.: To what extent can biogenic SOA be controlled?, Environ. Sci. Technol., 44, 3376–3380, doi:10.1021/es903506b, 2010

Carslaw, K. S., Lee L. A., Reddington C. L., Pringle K. J., Rap A., Forster P. M., Mann G. W., Spracklen D. V., Woodhouse M. T., Regayre L. A., Pierce J. R.: Large contribution of natural aerosols to uncertainty in indirect forcing. Nature, 503, 7474, 10.1038/nature12674, 2013

Cionni, I., Eyring, V., Lamarque, J. F., Randel, W. J, Stevenson, D. S., Wu, F., Bodeker, G. E., Shepherd, T. G., Shindell, D. T., Waugh, D. W.: Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, Atmos. Chem. Phys., 11, 11267-11292, 2011

Clarke L., K. Jiang, K. Akimoto, M. Babiker, G. Blanford, K. Fisher-Vanden, J.-C. Hourcade, V. Krey, E. Kriegler, A. Löschel, D. McCollum, S. Paltsev, S. Rose, P.R. Shukla, M. Tavoni, B.C.C. van der Zwaan, and D.P. van Vuuren,: Assessing Transformation Pathways. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Collins, W.J., Derwent, R.G., Garnier, B., Johnson, C.E., Sanderson, M.G., Stevenson, D.S.,. The effect of stratosphere–troposphere exchange on the future tropospheric ozone trend. Journal of Geophysical Research 108. doi:10.1029/2002JD002617, 2003

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geosci. Model Dev., 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Collins, M., R. et al., Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Eyring, V., et al.,: Multimodel projections of stratospheric ozone in the 21st century. *J. Geophys. Res. Atmos.*, 112, D16303, 2007

Eyring, V., et al.: Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) Community Simulations in Support of Upcoming Ozone and Climate Assessments, SPARC Newsletter, *40*, 48-66, 2013a

Eyring, V., Arblaster J. M., Cionni I., Sedlacek J., Perlwitz J., Young P. J., Bekki S., Bergmann D., Cameron-Smith P., Collins W. J., Faluvegi G., Gottschaldt K.-D., Horowitz L. W., Kinnison D. E., Lamarque J.-F., Marsh D. R., Saint-Martin D., Shindell D. T., Sudo K., Szopa S., and Watanabe S.:, Long-term ozone changes and associated climate impacts in CMIP5 simulations, J. Geophys. Res. Atmos, 118, 5029-5060, 2013b

Eyring, V., Bony, S., Meehl, G. A., Senior, C., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016a.

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K.-D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S., and Williams, K. D.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, Geosci. Model Dev., 9, 1747-1802, doi:10.5194/gmd-9-1747-2016, 2016b.

Eyring, V., Gleckler, P. J., Heinze, C., Stouffer, R. J., Taylor, K. E., Balaji, V., Guilyardi, E., Joussaume, S., Kindermann, S., Lawrence, B. N., Meehl, G. A., Righi, M., and Williams, D. N.: Towards improved and more routine Earth system model evaluation in CMIP, Earth Syst. Dynam., doi:10.5194/esd-2016-26, 2016.

Ferraro, R., Waliser, D. E., Gleckler, P., Taylor, K. E., and Eyring, V.: Evolving obs4MIPs to Support the Sixth Coupled Model Intercomparison Project (CMIP6), B Am Meteorol Soc, doi: 10.1175/BAMS-D-14-00216.1, 2015.

Fiore, A. M., Naik V., Spracklen D., Steiner A., Unger N., Prather M., Bergmann D., Cameron-Smith P. J., Collins W., Dalsøren S., Folberth G., Ginoux P., Horowitz L. W., Josse B., Lamarque J.-F., Nagashima T., O'Connor F., Rumbold S., Shindell D. T., Skeie R. B., Sudo K., Takemura T., Zeng G.: Global Air Quality and Climate, Chemical Society Reviews, 41, 6663-6683, doi: 10.1039/C2CS35095E, 2012

Fiore, A.M., Naik V., Leibensperger E.M.: Air quality and climate connections J. Air Waste Manag. Assoc., 65 pp. 645–685, doi:10.1080/10962247.2015.1040526, 2015

Flato, G., Marotzke J., Abiodun B., Braconnot P., Chou S.C., Collins W., Cox P., Driouech F., Emori S., Eyring V., Forest C., Gleckler P., Guilyardi E., Jakob C., Kattsov V., Reason C. and Rummukainen M: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013

Forster, P., et al.: Changes in Atmospheric Constituents and in Radiative Forcing. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (eds.)] Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 129-234 , 2007

Forster, P. M., T. Richardson, A. C. Maycock, C. J. Smith, B. H. Samset, G. Myhre, T. Andrews, R. Pincus, and M. Schulz: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, J. Geophys. Res. Atmos., 121, doi:10.1002/2016JD025320, 2016.

Gauss, M., Myhre, G., Pitari, G., et al.: Radiative forcing in the 21st century due to ozone changes in the troposphere and lower stratosphere, J. Geophys. Res., 108, 4292, doi:10.1029/2002JD002624, 2003

Gauss, M., et al.: Radiative forcing since preindustrial times due to ozone changes in the troposphere and the lower stratosphere, Atmos. Chem. Phys., 6, 575–599, 2006

Ghan, S. J., Liu X., Easter R. C., Zaveri R., Rasch P. J., Yoon J.-H., and Eaton B.: Toward a minimal representation of aerosols in climate models: Comparative decomposition of aerosol direct, semi-direct and indirect radiative forcing, J. Clim., 25, 6461–6476, doi: 10.1175/JCLI-D-11-00650.1, 2012

Ghan, S. J.: Technical Note: Estimating aerosol effects on cloud radiative forcing, Atmos. Chem. Phys., 13, 9971-9974, doi:10.5194/acp-13-9971-2013, 2013.

Ghan, S., Wang, M., Zhang, S., Ferrachat, S., Gettelman, A., Griesfeller, J., Kipling, Z., Lohmann, U., Morrison, H., Neubauer, D., Partridge, D. G., Stier, P., Takemura, T., Wang, H., Zhang, K.: Challenges in constraining anthropogenic aerosol effects on cloud radiative forcing using present-day spatiotemporal variability, Proceedings of the National Academy of Sciences, 2016

Gillett N. P., Shiogama H., Funke B., Hegerl G., Knutti R., Matthes K., Santer B. D., Stone D., and Tebaldi C.: Detection and Attribution Model Intercomparison Project (DAMIP). Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-74, 2016.

Gleckler, P. J., Doutriaux, C., Durack P. J., Taylor K. E. , Zhang, Y., Williams, D. N., Mason, E., and Servonnat, J.: A more powerful reality test for climate models, Eos Trans. AGU, 97, 2016

Gordon, H., Sengupta, K., Rap, A., Duplissy, J., Frege, C., Williamson, C., Heinritzi, M., Simon, M., Yan, C., Almeida, J., Tröstl, J., Nieminen, T., Ortega, I. K., Wagner, R., Dunne, E. M., Adamov, A., Amorim, A., Bernhammer, A.-K., Bianchi, F., Breitenlechner, M., Brilke, S., Chen, X., Craven, J. S., Dias, A., Ehrhart, S., Fischer, L., Flagan, R. C., Franchin, A., Fuchs, C., Guida, R., Hakala, J., Hoyle, C. R., Jokinen, T., Junninen, H., Kangasluoma, J., Kim, J., Kirkby, J., Krapf, M., Kürten, A., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Mathot, S., Molteni, U., Monks, S. A., Onnela, A., Peräkylä, O., Piel, F., Petäjä, T., Praplan, A. P., Pringle, K. J., Richards, N. a. D., Rissanen, M. P., Rondo, L., Sarnela, N., Schobesberger, S., Scott, C. E., Seinfeld, J. H., Sharma, S., Sipilä, M., Steiner, G., Stozhkov, Y., Stratmann, F., Tomé, A., Virtanen, A., Vogel, A. L., Wagner, A. C., Wagner, P. E., Weingartner, E., Wimmer, D., Winkler, P. M., Ye, P., Zhang, X., Hansel, A., Dommen, J., Donahue, N. M., Worsnop, D. R., Baltensperger, U., Kulmala, M., Curtius, J. & Carslaw, K. S. 2016. Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1602360113

Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., Denier van der Gon, H., Frost, G. J., Heil, A., Kaiser, J. W., Kinne, S.,Klimont, Z., Kloster, S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M. G., Smith, S. J., Thompson, A., van Aardenne, J., van der Werf, G. R., and van Vuuren, D. P.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period, Climatic Change, 109, 163–190, doi:10.1007/s10584-011-0154-1, 2011.

Gregory, J. M., Ingram W. J, Palmer M. A., Jones G. S., Stott P. A., Thorpe R. B., Lowe J. A., Johns T. C., and Williams K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31(3), L03205, 2004

Hegglin, M. I., and Shepherd T. G.: Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux, Nature Geosci., 2, 687–691, 2009

Hegglin, M. I., et al., SPARC Data Initiative: Comparison of stratospheric composition measurements from international satellite limb sounders, in preparation.

Hollmann, R., Merchant, C. J., Saunders, R., Downy, C., Buchwitz, M,. Cazenave, A,. Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R,. Holzer-Popp, T., Paul, F,. Sandven, S., Sathyendranath, S., van Roozendael, M,. and Wagner, W.:

The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables. *Bull. Amer. Meteor. Soc.*, **94**, 1541–1552, doi: 10.1175/BAMS-D-11-00254.1., 2013

Hwang, Y. T., Frierson D. M., and Kang S. M.: Anthropogenic sulfate aerosol and the southward shift of tropical precipitation in the late 20th century, *Geophys. Res. Lett.*, 40, 2845–2850, doi:10.1002/grl.50502, 2013

IPCC: Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios, J.T. Houghton, L.G. Meira Filho, J. Bruce, Hoesung Lee, B.A. Callander, E. Haites, N. Harris and K. Maskell (eds.). Cambridge University Press, Cambridge, UK, 339 pp., 1994

IPCC: Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 572 pp., 1996

Isaksen, I.S.A., Granier C., Myhre G., Berntsen T.K., Dalsøren S.B., Gauss M., Klimont Z., Benestad R., Bousquet P., Collins W.: Atmospheric composition change: climate–chemistry interactions, Atmos. Environ., 4, 5138–5192, 2009

Jacob, D.J., Winner D.A., Effect of climate change on air quality, *Atmos. Environ.*, 43, 51–63. doi:10.1016/j.atmosenv.2008.09.05, 2009

John, J. G., Fiore, A. M., Naik, V., Horowitz, L. W., and Dunne, J. P.: Climate versus emission drivers of methane lifetime against loss by tropospheric OH from 1860–2100, Atmos. Chem. Phys., 12, 12021-12036, doi:10.5194/acp-12-12021-2012, 2012.

Jones, C.D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J.G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J.T., and Zaehle, S.: C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6, Geosci. Model Dev., 9, 2853-2880, doi:10.5194/gmd-9-2853-2016, 2016

Kinne, S., Schulz M., Textor C., Guibert S., Balkanski Y., Bauer S.E., Berntsen T., Berglen T.F., Boucher O., Chin M., Collins W., Dentener F., Diehl T., Easter R., Feichter J., Fillmore D., Ghan S., Ginoux P., Gong S., Grini A., Hendricks J., Herzog M., Horowitz L., Isaksen I., Iversen T., Kirkevåg A., Kloster S., Koch D., Kristjansson J.E., Krol M., Lauer A., Lamarque J.F., Lesins G., Liu X., Lohmann U., Montanaro V., Myhre G., Penner J., Pitari G., Reddy S., Seland O., Stier P., Takemura T., and Tie X.: An AeroCom initial assessment optical properties in aerosol component modules of global models. Atmos. Chem. Phys., 6, 1815-1834, 2006

Kirtman, B., et al., Near-term Climate Change: Projections and Predictability. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Lamarque, J. F., et al.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): Overview and description of models, simulations and climate diagnostics. *Geosci. Model Dev.*, 6, 179–206, 2013

Langner, J. and Rodhe H.: A global three-dimensional model of the tropospheric sulphur cycle. *J. Atmos. Chem.*, 13, 255-263, 1991

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin. K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., de Noblet-Ducoudré, N., Pongratz, J., Seneviratne, S. I., Shevliakova, E.: The Land Use Model Intercomparison Project (LUMIP): Rationale and experimental design, Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-76, 2016.

Leibensperger, E. M., Mickley, L. J., Jacob, D. J., Chen, W.-T., Seinfeld, J. H., Nenes, A., Adams, P. J., Streets, D. G., Kumar, N., and Rind, D.: Climatic effects of 1950–2050 changes in US anthropogenic aerosols – Part 2: Climate response, Atmos. Chem. Phys., 12, 3349–3362, doi:10.5194/acp-12-3349-2012, 2012

Levy, H., Horowitz L. W., Schwarzkopf M. D., Ming Y., Golaz J.-C., Naik V., and Ramaswamy V.: The roles of aerosol direct and indirect effects in past and future climate change, *J. Geophys. Res. Atmos.*, 118, 4521–4532, doi:10.1002/jgrd.50192, 2013

Marvel, K., Schmidt, G.A., Miller, R.L., Nazarenko, L.S.: Implications for climate sensitivity from the response to individual forcings. *Nature Climate Change*, 6, doi:10.1038/nclimate2888, 2016

McCoy et al. Natural aerosols explain seasonal and spatial patterns of Southern Ocean cloud albedo, *Sci. Adv.*;1:e1500157, 2015

McLandress, C., Shepherd T. G., Scinocca J. F., Plummer D. A., Sigmond M., Jonsson A. I., and Reader M. C.: Separating the dynamical effects of climate change and ozone depletion. Part II Southern Hemisphere troposphere, *J. Clim.*, 24(6), 1850–1868, 2011

McLandress, C., Shepherd, T. G., Jonsson, A. I., von Clarmann, T., and Funke, B.: A method for merging nadir-sounding climate records, with an application to the global-mean stratospheric temperature data sets from SSU and AMSU, *Atmos. Chem. Phys.*, 15, 9271-9284, doi: 10.5194/acp-15-9271-2015, 2015.

Myhre, G., et al.: Anthropogenic and Natural Radiative Forcing, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, pp. 659-740, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013a

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Luo, G., Ma, X., Penner, J. E., Rasch, P. J., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J.-H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, Atmos. Chem. Phys., 13, 1853-1877, doi:10.5194/acp-13-1853-2013, 2013b.

Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J.-F., Lin, M., Prather, M. J., Young, P. J., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R., Eyring, V., Faluvegi, G., Folberth, G. A., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., van Noije, T. P. C., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R., Shindell, D. T., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., and Zeng, G.: Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 5277–5298, doi:10.5194/acp-13-5277-2013, 2013.

Olson, J., Prather M., Berntsen T., Carmichael G., Chatfield R., Connell P., Derwent R., Horowitz L., Jin S., Kanakidou M., Kasibhatla P., Kotamarthi R., Kuhn M., Law K., Penner J., Perliski L., Sillman S., Stordal F., Thompson A., and Wild O.: Results from the Intergovernmental Panel on Climatic Change Photochemical Model Intercomparison (PhotoComp) J. Geophys., Res., 102, D5, 5979-5991, 1997

O’Neill, B.C., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R., Van Vuuren, D.P., 2014. A new scenario framework for climate change research: The concept of shared socioeconomic pathways. Clim. Change 122, 387–400, 2014.

O'Neill, B.C., Tebaldi C., van Vuuren D., Eyring V., Friedlingstein P., Hurtt G., Knutti R., Kriegler E., Lamarque J.-F., Lowe, J. Meehl J., Moss R., Riahi K., and. Sanderson B. M, The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev. Discuss.*, doi:10.5194/gmd-2016-84, 2016

Pawson, S., and W. Steinbrecht (Lead Authors), A.J. Charlton-Perez, M. Fujiwara, A.Yu. Karpechko, I. Petropavlovskikh, J. Urban, and M. Weber, Update on global ozone: Past, present, and future, Chapter 2 in Scientific Assessment of Ozone Depletion: 2014, Global Ozone Research and Monitoring Project –Report No. 55, World Meteorological Organization, Geneva, Switzerland, 2014.

Penner, J.E., et al.: Aerosols, their direct and indirect effects. In: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 289–348, 2001

Pincus, R., Forster, P.M., Stevens, B.: The Radiative Forcing Model Intercomparison Project, Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-88, 2016

Polvani, L. M., Waugh D. W., Correa G. J. P., and Son S. W.: Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere, *J. Clim.*, 24(3), 795–812, 2011

Prather, M.J., Lifetimes and eigenstates in atmospheric chemistry, Geophys.Res.Lett., 21, 801-804, 1994.

Prather, M., et al. Atmospheric Chemistry and Greenhouse Gases, in Climate Change 2001: The Scientific Basis, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., et al., Cambridge University Press, Cambridge, UK, pp. 239–287, 2001

Prather, M., R. Derwent, D. Ehhalt, P. Fraser, E Sanhueza, X. Zhou, Chapter 2: Other tracer gases and atmospheric chemistry, in Climate Change 1994, Intergovernmental Panel on Climate Change, J.T. Houghton et al., eds., Cambridge U. Press, pp. 73-126. 1994.

Prather, M.J., C.D. Holmes, J. Hsu, Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, Geophys. Res. Lett., 39, L09803, 5 pp., doi:10.1029/2012GL051440, 2012

Ramaswamy, V., et al.: Radiative forcing of climate change. In: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [J.T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell and C. A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 349-416., 2001

Righi, M., V. Eyring, K.-D Gottschaldt, C. Klinger, F. Frank, P. Jöckel, and I. Cionni, Quantitative evaluation of ozone and selected climate parameters in a set of EMAC simulations, Geosci. Model Dev., 8, 733-768, doi:10.5194/gmd-8-733-2015, 2015.

Rotstayn, L., Collier M., Shindell D., and Boucher O., Why does aerosol forcing control historical global-mean surface temperature change in CMIP5 models?, J. Climate, 28, 6608-6625, 2015.

Schnell, J.L., Prather, M. J. , Josse, B., Naik, V., Horowitz, L. W., Cameron-Smith, P., Bergmann, D., Zeng, G., Plummer, D., Sudo, A. K., Nagashima, T., Shindell, D. T., Faluvegi, G. and Strode, S. A., Use of North American and European air quality networks to evaluate global chemistry-climate modeling of surface ozone, Atmos. Chem. Phys., 15, 10581-10596, 2015.

Schnell, J.L., Prather, M. J. , Josse, B., Naik, V., Horowitz, L. W., Zeng, G., Shindell, D. T., Effect of climate change on surface ozone over North America, Europe, and East Asia, Geophys. Res. Lett., 43, 3509–3518, 2016.

Schulz, M., Textor C., Kinne S., Balkanski Y., Bauer S., Berntsen T., Berglen T., Boucher O., Dentener F., Grini A., Guibert S., Iversen T., Koch D., Kirkeveg A., Liu X., Montanaro V., Myhre G., Penner J., Pitari G., Reddy S., Seland X., Stier P., and Takemura T.: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. Atmos Chem Phys, 6, 5225-5246, 2006.

Sherwood, S. C., Bony S., and Dufresne J. L.: Spread in model climate sensitivity traced to atmospheric convective mixing, *Nature*, 505, 37-42, 2014

Sherwood, S., Bony S., Boucher O., Bretherton C., Forster P., Gregory J., and Stevens B.: Adjustments to the forcing-feedback framework for understanding climate change, Bull. Amer. Meteorol. Soc., 96, 217-228, 2015.

Shindell, D. T.: Inhomogeneous forcing and transient climate sensitivity, *Nature Climate Change*, 4, 274–277, doi:10.1038/nclimate2136, 2014

Shindell, D. T., Faluvegi, G., Koch, D.M., Schmidt, G.A., Unger, N., Bauer, S.E.: Improved Attribution of Climate Forcing to Emissions, Science, 326, 716-718, 2009.

Shindell, D., et al.: Simultaneously mitigating near-term climate change and improving human health and food security, *Science*, 335, 183–189, 2012a

Shindell, D. T., Voulgarakis A., Faluvegi G., and Milly G.: Precipitation response to regional radiative forcing. Atmos. Chem. Phys., 12, 6969-6982, 2012b

Shindell, D. T., Lamarque J.-F., Schulz M., Flanner M., Jiao C., Chin M., Young P. J., Lee Y., Milly G., Faluvegi G., Balkanski Y., Collins W. J.,. Conley A. J, Dalsoren S., Easter R., Ghan S. J., Horowitz L., Liu X., Myhre G., Nagashima T., Naik V., Rotstayn L., Rumbold S., Skeie R., Sudo K., Szopa S., Takemura T., Yoon J.-H.: Radiative forcing in the ACCMIP historical and future climate simulations, Atmos. Chem. Phys., 13, 2939-2974, doi:10.5194/acp-13-2939-2013, 2013a.

Shindell, D. T., Faluvegi G.,. Nazarenko L, Bowman K., Lamarque J.-F., Voulgarakis A., Schmidt G., Pechony O., Reudy R., Attribution of historical whole atmosphere ozone forcing. Nature Climate Change, 3, 567–570, doi:10.1038/nclimate1835, 2013b.

Shindell, D.T., O. Pechony, A. Voulgarakis, G. Faluvegi, L. Nazarenko, J.-F. Lamarque, K. Bowman, G. Milly, B. Kovari, R. Ruedy, and G. Schmidt, 2013: Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations. Atmos. Chem. Phys., 13, 2653-2689, doi:10.5194/acp-13-2653-2013.

Shindell, D. T., Faluvegi, G., Rotstayn, L., Milly, G., Spatial patterns of radiative forcing and surface temperature response, *J.* Geophys. Res., 120, 5385-5403, 2015

Shine, K. P., Bourqui, M. S., de Forster, P. M., Hare, S. H. E., Langematz, U., Braesicke, P., Grewe, V., Ponater, M., Schnadt, C., Smith, C. A., Haigh, J. D., Austin, J., Butchart, N., Shindell, D. T., Randel, W. J., Nagashima, T., Portmann, R. W., Solomon, S., Seidel, D. J., Lanzante, J., Klein, S., Ramaswamy, V., and Schwarzkopf M. D.: A comparison of model-simulated trends in stratospheric temperatures, Q. J. Roy. Meteorol. Soc., 129, 1565–1588, 2003

Smith, S. J., van Aardenne, J., Klimont, Z., Andres, R. J., Volke, A., and Delgado Arias, S.: Anthropogenic sulfur dioxide emissions: 1850–2005, Atmos. Chem. Phys., 11, 1101-1116, doi:10.5194/acp-11-1101-2011, 2011.

Son, S.-W., Polvani L. M., Waugh D. W., Akiyoshi H., Garcia R., Kinnison D., Pawson S., Rozanov E., Shepherd T. G., and Shibata K.: The impact of stratospheric ozone recovery on the Southern hemisphere westerly jet, *Science*, 320, 1486-1489., 2008

Son, S.-W., Gerber E. P., Perlwitz J., et al.: Impact of Stratospheric Ozone on Southern Hemisphere Circulation Change: A Multimodel Assessment, J. Geophys. Res., 115, D00M07, doi:10.1029/2010JD014271 , 2010

SPARC CCMVal, SPARC Report on the Evaluation of Chemistry-Climate Models, V. Eyring, T. G. Shepherd, D. W. Waugh (Eds.), SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526, http://www.atmosp.physics.utoronto.ca/SPARC, 2010.

Stevenson, D. S., et al.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone. J. Geophys. Res. Atmos., 111, D08301, 2006

Stevenson, D. S., et al.: Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 3063-3085, doi:10.5194/acp-13-3063-2013, 2013

Teixeira, J., Waliser, D., Ferraro, R., Gleckler, P., Lee, T., and Potter, G.: Satellite Observations for CMIP5: The Genesis of Obs4MIPs, B Am Meteorol Soc, 95, 1329-1334, 2014.

Teng H., Washington W. M., Branstator G., Meehl G. A., and Lamarque J. F.: Potential impacts of Asian carbon aerosols on future US warming, Geophys. Res. Lett., 39, L11703, doi:10.1029/2012GL051723, 2012

Textor, C., Schulz M., Guibert S., Kinne S., Balkanski Y., Bauer S., Berntsen T., Berglen T., Boucher O., Chin M., Dentener F., Diehl T., Easter R., Feichter H., Fillmore D., Ghan S., Ginoux P., Gong S., Grini A., Hendricks J., Horowitz L., Huang P., Isaksen I., Iversen T., Kloster S., Koch D., Kirkevåg A., Kristjansson J.E., Krol M., Lauer A., Lamarque J.F., Liu X., Montanaro V., Myhre G., Penner J., Pitari G., Reddy S., Seland Ø., Stier P., Takemura T., and Tie X.: Analysis and quantification of the diversities of aerosol life cycles within AeroCom, Atmos. Chem. Phys., 6, 1777-1813, 2006

van Vuuren, D.P., Kriegler, E., O’neill, B.C., Ebi, K.L., Riahi, K., Carter, T.R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R., Winkler, H.: A new scenario framework for climate change research: scenario matrix architecture. Clim. Change 122, 373–386, 2014.

Voulgarakis, A., Naik, V., Lamarque, J.-F., Shindell, D. T., Young, P.J., Prather, M. J., Wild, O., Field, R. D., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Folberth, G. A., Horowitz, L.W., Josse, B., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Stevenson, D. S., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Analysis of present day and future OH and methane lifetime in the ACCMIP simulations, Atmos. Chem. Phys., 13, 2563–2587, doi:10.5194/acp-13-2563-2013, 2013.

Vuolo, M. R., Schulz, M., Balkanski, Y., and Takemura, T.: A new method for evaluating the impact of vertical distribution on aerosol radiative forcing in general circulation models, Atmos. Chem. Phys., 14, 877-897, doi:10.5194/acp-14-877-2014, 2014.

WMO: Scientific Assessment of Ozone Depletion 2010. Global Ozone Research and Monitoring Project Report, 52. World Meteorological Organisation, Geneva, 2010.

WMO: Scientific Assessment of Ozone Depletion 2014. Global Ozone Research and Monitoring Project Report, 55. World Meteorological Organisation, Geneva, 2014.

Young, P. J., Archibald, A. T., Bowman, K. W., Lamarque, J. F., Naik, V., Stevenson, D. S., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsoren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Horowitz, L. W., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R. B., Shindell, D. T., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 2063-2090, 10.5194/acp-13-2063-2013, 2013.

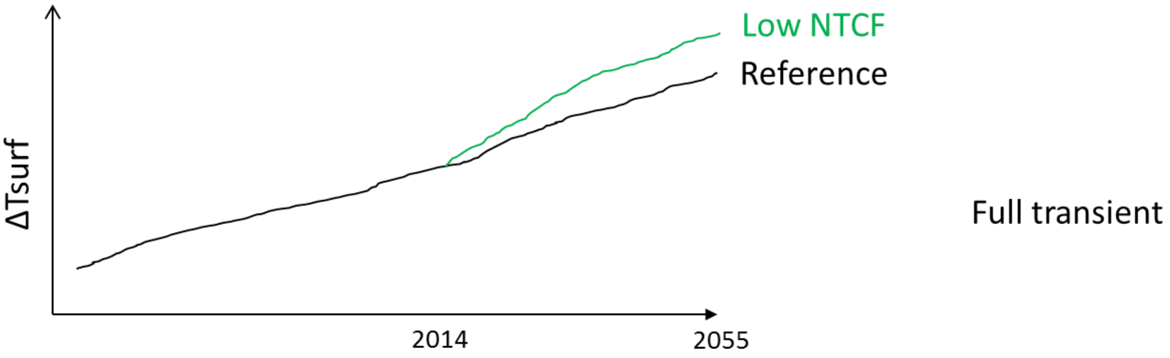

Figure 1: Schematic of future coupled-ocean simulations based on the SSP3-7.0 scenario.

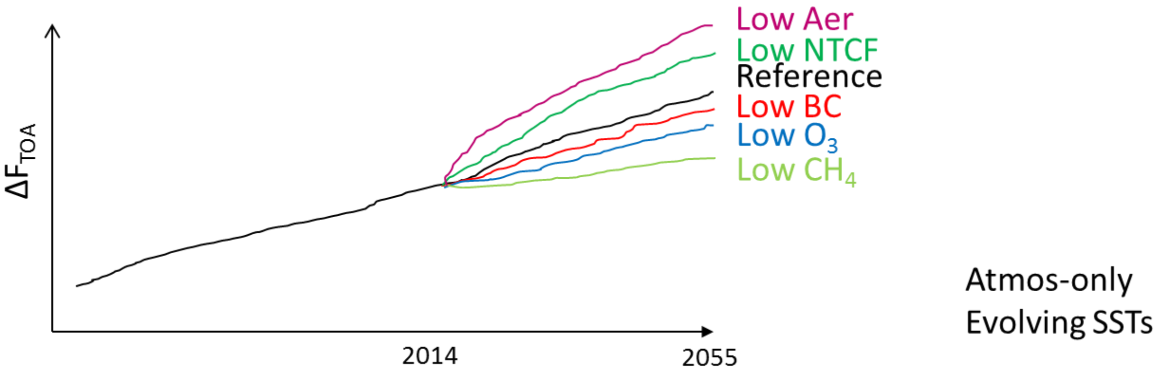

Figure 2: Schematic of future prescribed SST simulations based on the SSP3-7.0 scenario.

Experiment ID	Minimum model configuration	CH_4	N_2O	Aerosol Precursors	Ozone precursors	CFC/ HCFC	Tier
*hist-**piNTCF***	AOGCM_AER	Hist	Hist	**1850**	**1850**	Hist	1
*hist-**piAer***	AOGCM_AER	Hist	Hist	**1850**	Hist	Hist	2
*hist-**1950HC***	AOGCM_CHEM^{S}	Hist	Hist	Hist	Hist	**1950**	1

Table 1: List of historical coupled-ocean experiments. Experiments cover the period between 1850-2014, except *hist-1950HC* which starts in 1950. The "AER" suffix means models should calculate tropospheric aerosols driven by emission fluxes. The "CHEM^{S}" suffix means tropospheric or stratospheric chemistry are required. The species columns refer to the specifications for concentrations (CH_4, N_2O and CFC/HCFC) or emissions (Aerosol and Ozone precursors). "Hist" means the concentrations or emissions should evolve as for the CMIP6 historical simulation, a year means the concentrations or emissions should be fixed to that year. Three ensemble members are requested for each experiment.

Experiment ID	Minimum model configuration	CH_4	N_2O	Aerosol Precursors	Ozone precursors	CFC/ HCFC	Tier
histSST	AGCM_AER	Hist	Hist	Hist	Hist	Hist	1
*histSST-**piNTCF***	AGCM_AER	Hist	Hist	**1850**	**1850**	Hist	1
*histSST-**piAer***	AGCM_AER	Hist	Hist	**1850**	Hist	Hist	2
*histSST-**piO3***	AGCM_CHEM^{T}	Hist	Hist	Hist	**1850**	Hist	2
*histSST-**1950HC***	AGCM_CHEM^{S}	Hist	Hist	Hist	Hist	**1950**	1
*histSST-**piCH4***	AGCM_CHEM$^{T/S}$	**1850**	Hist	Hist	Hist	Hist	1
*histSST-**piN2O***	AGCM_CHEM^{S}	Hist	**1850**	Hist	Hist	Hist	2

Table 2: List of historical prescribed SST experiments. Experiments cover the period between 1850-2014, except *histSST-1950HC* which starts in 1950. The "AER" suffix means models should calculate tropospheric aerosols driven by emission fluxes. The "CHEM^{T}" or "CHEM^{S}" suffix means that tropospheric or stratospheric chemistry are required. The species columns refer to the specifications for concentrations (CH_4, N_2O and CFC/HCFC) or emissions (Aerosol and Ozone precursors). "Hist" means the concentrations or emissions should evolve as for the CMIP6 historical simulation, a year means the concentrations or emissions should be fixed to that year. Note that the N_2O and HC runs will have a small forcing, but will provide valuable information on their historical impacts on stratospheric ozone changes.

Policy strength	Policy targets		Technological innovation
	High Income countries	Medium and Low income countries	
Strong	Aim for much lower pollutant levels than current targets	Comparatively quick catch-up with the developed world	Pollution control technology costs drop substantially with control performance increasing.
Medium	Lower than current targets	Catch-up with the developed world, but not as quickly as "Strong"	Continued modest technology advances
Weak	Regionally varied policies	Trade barriers and/or institutional limitations substantially slow progress in pollution control.	Lower levels of technological advance overall.

Table 3: Qualitative framework for pollution control in the SSPs, based on Rao et al. (2016) table 2.

Experiment ID	Minimum model configuration	CH_4	Aerosol Precursors	Ozone precursors	Tier
ssp370†	AOGCM_AER	Reference	Reference	Reference	1
*ssp370-**lowNTCF***	AOGCM_AER	Reference	**Clean**	**Clean**	1

Table 4: List of future coupled ocean experiments. Experiments cover the period 2015 to 2055. The "AER" suffix models should calculate tropospheric aerosols driven by emission fluxes. The species columns refer to the specifications for concentrations (CH_4) or emissions (Aerosol and Ozone precursors). "Reference" means the concentrations or emissions should evolve as for the SSP3-7.0, "Clean" means the concentrations or emissions should evolve following a version of SSP3-7.0 with cleaner air quality policies. CFC/HCFC concentrations should follow the SSP3-7.0 scenario in both cases. Three ensemble members are requested for each experiment. †*ssp370* is also specified as Tier 1 in ScenarioMIP

Experiment ID	Minimum model configuration	CH_4	Aerosol Precursors	Ozone precursors	Tier
ssp370SST	AGCM_AER	Reference	Reference	Reference	1
*ssp370SST-**lowNTCF***	AGCM_AER	Reference	**Clean**	**Clean**	1
*ssp370SST-**lowAer***	AGCM_AER	Reference	**Clean**	Reference	2
*ssp370SST-**lowBC***	AGCM_AER	Reference	Reference (non BC) **Clean** (BC)	Reference	2
*ssp370SST-**lowO3***	AGCM_CHEM^{T}	Reference	Reference	**Clean**	2
*ssp370SST-**lowCH4***	AGCM_CHEM^{T}	**Low**	Reference	Reference	2
*ssp370SST-ssp126**Lu***	AGCM_AER	Reference	Reference (anthropogenic)		2

Table 5: List of future prescribed SST simulations. Experiments cover the period 2015 to 2055. The "AER" suffix means models should calculate tropospheric aerosols driven by emission fluxes The "CHEM^{T} suffix means that interactive tropospheric chemistry is required. The species columns refer to the specifications for concentrations (CH_4) or emissions (Aerosol and Ozone precursors). "Reference" means the concentrations or emissions should evolve as for the SSP3-7.0, "Clean" means the concentrations or emissions should evolve following a version of SSP3-7.0 with cleaner air quality policies. CFC/HCFC concentrations should follow the SSP3-7.0 scenario in all cases. For *ssp370SST-lowLu* the land use mask from the SSP1-2.6 scenario should be used for the interactive natural emission schemes (and for the climate); anthropogenic emissions should follow the SSP3-7.0 scenario.

Experiment ID	Minimum model configuration	CH_4	N_2O	Aerosol Precursors	Ozone Precursors	CFC/ HCFC	Tier
piClim[†]	AGCM-AER	1850	1850	1850	1850	1850	1
*piClim-**NTCF***[†]	AGCM-^{T}AER	1850	1850	**2014**	**2014**	1850	1
*piClim-a**er***	AGCM-AER	1850	1850	**2014**	1850	1850	2
*piClim-**BC***	AGCM-AER	1850	1850	1850 (non BC) **2014** (BC)	1850	1850	2
*piClim-**O3***	AGCM-CHEM^{T}	1850	1850	1850	**2014**	1850	2
*piClim-**CH4***	AGCM-CHEM$^{T/S}$	**2014**	1850	1850	1850	1850	1
*piClim-**N2O***	AGCM-CHEM^{S}	1850	**2014**	1850	1850	1850	2
*piClim-**HC***	AGCM-CHEM^{S}	1850	1850	1850	1850	**2014**	1
*piClim-**NOX***	AGCM-CHEM^{T}	1850	1850	1850	1850 (non NO_X) **2014** (NO_X)	1850	3
*piClim-**VOC***	AGCM-CHEM^{T}	1850	1850	1850	1850 (non CO/VOC) **2014** (CO/VOC)	1850	3
*piClim-**SO2***	AGCM-AER	1850	1850	1850 (non SO2) **2014** (SO2)	1850	1850	3
*piClim-**OC***	AGCM-AER	1850	1850	1850 (non OC) **2014** (OC)	1850	1850	3
*piClim-**NH3***	AGCM-AER	1850	1850	1850 (non NH3) **2014** (NH3)	1850	1850	3

Table 6: List of fixed SST ERF simulations. These are timeslice experiments of 30 years total, using pre-industrial climatological average SST and sea-ice distributions. The "AER" suffix means that models should calculate tropospheric aerosols driven by emission fluxes. The "CHEM^{T}" or "CHEM^{S}" suffix means that tropospheric or stratospheric chemistry are required. The species columns refer to the specifications for concentrations (CH_4, N_2O and CFC/HCFC) or emissions (Aerosol and Ozone precursors). A year means the concentrations or emissions should be fixed to that year. [†] *piClim* is identical to that in RFMIP, *piClim-NTCF* is identical to piClim-aerO3 (in RFMIP) for models with interactive tropospheric and stratospheric chemistry.

Experiment ID	Minimum model configuration	Flux to be doubled	Tier
piClim-***2xdust***	AGCM-AER	Dust	2
piClim-***2xss***	AGCM-AER	Sea salt	2
piClim-***2xDMS***	AGCM-AER	Oceanic DMS	3
piClim-***2xfire***	AGCM-AER	Fire (NOx, BC, OC, CO, VOCs…)	3
piClim-***2xNOX***	AGCM-CHEM^{T}	Lightning NO$_X$	3
piClim-***2xVOC***	AGCM-CHEM^{T}	Biogenic VOCs	3

Table 7: List of fixed SST simulations for ERFs of natural emitted species. These are timeslice experiments of 30 years total, using climatological average SST and sea-ice distributions. The "AER" suffix means that at least interactive aerosols are required, interactive chemistry should be active if available. The "CHEM^{T}" suffix means that both interactive aerosols and (at least tropospheric) chemistry are required. The specified natural emission flux should be doubled compared to the pre-industrial control, either scaling the parameterisations in an interactive scheme or the data files for specified emissions. All other forcing agents should be as in pre-industrial control.

Anonymous Referee #1

This manuscript describes the Aerosol Chemistry Model Intercomparison Project (AerChemMIP) - a CMIP6 endorsed MIP. The motivation for this MIP is to quantify the climate forcing and response due to near term climate forcers including tropospheric ozone, aerosols and methane, and N2O in a coordinated manner within the Coupled Model Intercomparison Project (CMIP). Although the contribution of tropospheric aerosols and ozone to climate forcing was recognized since the early 1990s, attribution of climate forcing due to these species within the CMIP framework was not performed until CMIP5. And even within the CMIP5, the radiative forcing due to tropospheric aerosols and ozone was calculated by a combination of methods (offline, double calls) partly because not all models included chemistry-aerosols and partly because all the required diagnostics were not available from CMIP. AerChemMIP represents the first coordinated effort towards quantifying climate forcing and response, and air quality impacts from tropospheric aerosols and ozone within the mainstream physical climate model intercomparison project, the CMIP, a major step forward. The manuscript is generally well-written, provides motivation for why modeling centers should participate in this computationally expensive (over 3000 simulation years) MIP, and provides examples of scientific studies that can be conducted beyond the goals of the MIP. However, I note some issues that should be addressed to improve the manuscript.

Given that the multi-model output will be made publicly available for the wider community and users, and the complexity of chemistry-climate models, I suggest indicating that a model documentation/evaluation paper is required for participation in this MIP. This will address two issues - 1) often errors in the submitted data are not revealed until the data is compared against observations; model evaluation and documentation will ensure that the data (or at least the key atmospheric chemistry and aerosol fields) has been looked at before being submitted to the CMIP6 database, 2) each model is unique and data users are not in a position to know the details of all the parameters/processes/chemical schemes that are included in a model; a model documentation paper would be a source of this information.

We agree that model documentation and evaluation is extremely important. CMIP does not request individual model evaluation papers in advance of submission of the model output to ESGF since this would cause a substantial delay in making the model output available to the community. We therefore will not require this for AerChemMIP either. However, model documentation will occur through the ES-DOC component of CMIP6, see https://www.earthsystemcog.org/projects/es-doc-models/ and the models will be centrally evaluated with the ESMValTool and PCMDI metrics package as soon as the output is submitted to the ESGF (see Eyring et al. 2016). This will include the evaluation of chemistry and aerosols with diagnostics implemented into the ESMValTool. We have added this to the text in section 4.6 (page 16).

Some thought needs to be given to the third scientific question that AerChemMIP aims to answer - Can the uncertainties associated with anthropogenic emissions be quantified? Based on section 2.3, it is not clear to me how perturbing emissions by 10% will inform us about the dependence of radiative forcing on anthropogenic emissions. Also, it is not clear why a 10% perturbation is chosen.

We realise that the 10% perturbations would have imposed extra work on the modelling groups. We have instead added extra single species simulations to SO2, NH3 and OC in section 3.3 in order to cover the main NTCFs. These simulations are also necessary to characterise the individual ERFs fully. This will generate sufficient data to answer the question whilst limiting the computational requests. This section (section 2.3, page 6) now reads: "The primary focus of this question is to understand the sensitivity of present-day ERF to uncertainties in estimates of historical NTCF emissions. Indeed, while all proposed simulations rely on the usage of a central estimate, it is clear that there is a range of emission estimates (as discussed in Granier et al., 2011; Smith et al., 2011; Bond et al., 2013) that needs to be considered. While this uncertainty will clearly be region, sector and species dependent, it would be unrealistic to explore the full spectrum of variations. For that purpose, we will make use of perturbation (pre-industrial to present-day) simulations. This is likely to provide an upper bound on the impact of emission uncertainties. Results from the simulations can be directly compared to the simulations in section 3.1 and analysed for differences in radiative forcing as well as

air quality and overall atmospheric composition. Inter-model differences will document their varying sensitivities to emissions."

Specific and Minor Comments:

P1, L3-4: According to Myhre et al (2013), NTCFs are defined as species "whose impact on climate occurs primarily within the first decade after their emissions" and includes methane in addition to ozone, aerosols and their precursors. Suggest revising this sentence to avoid confusion.

We have made this revision (page 1, line 23)

P2, L25: Insert "tropospheric" in "...increases in ozone.." and insert "stratospheric in "...decreases in ozone. . ."

We did not make this change and ozone precursors can affect the stratosphere, and ODSs can affect the troposphere.

P2, L30: Ozone and O3 are used interchangeably in the text. Recommend sticking with one and specifically "ozone" throughout the manuscript. Same holds for methane.

We now use "ozone", "methane" and "nitrous oxide" throughout instead of the chemical formulae. (multiple locations)

P3, L3: Please reference a chapter in the IPCC (2014).

We now reference Clarke et al. 2014 (Chapter 6 of WGIII). (page 3, line 13)

P3, L5-7: I am not sure what the authors are trying to convey in this sentence. Is it that the "regional/local" air pollution policies should informed by the projections of air pollutants from CMIP6 models? If so, I am not convinced that this would be possible given the large model diversity in projected atmospheric composition (e.g, Young et al., 2013).

This has been rephrased (page 3, line 16): "CMIP6 will provide comprehensive information on the future large-scale evolution of atmospheric composition thus updating the knowledge base used to manage air pollution."

P3. L7-9: This sentence appears to represent an incomplete thought. Please elaborate or cut.

This sentence has been moved up where it fits better (page 3, lines 14-17).

P3, L30: Define the acronym ACCENT here rather than on Page 4, since it is used here for the first time in the text.

This change has been made. (page 3, line 40)

P3, L31: This sentence assumes that the relationship between IPCC and CMIP is commonly known. Please clarify how CMIP3 is related to the IPCC Fourth Assessment report.

This has been clarified (page 4 ,line 1)

P4, L11: A few of the newer generation CCMs available now with coupled ocean-atmosphere-chemistry have been documented and can be cited here. For eg., John et al. (2012); Shindell et al. (2013).

These references have been added. (page 4, line 23)

P4, L11-13: This sentence is giving short-shrift to the tremendous amount of work done under ACCMIP to assess the performance of CCMs with coupled strat-trop chemistry. The model shortcomings highlighted by ACCMIP studies is being addressed by more detailed analysis within CCMI as noted on http://blogs.reading.ac.uk/ccmi/.

Suggest revising this sentence to recognize earlier work regarding the assessment of the performance of CCMs with coupled strat-trop chemistry.

This paragraph was focussing on stratospheric ozone. We are not aware of ACCMIP studies that addressed this.

P4, L15-16: I am not sure if the inclusion of time-varying ozone in the climate models has led to an improvement in the climate forcing by stratospheric ozone. What has really improved since AR4 is the representation of climate forcing by stratospheric ozone in the CMIP5 models. Please revise sentence.

We have revised the sentence as suggested. (page 4, line 27)

P4, L28-32: Sentence is confusing. Please rephrase.

This sentence has been broken down into shorter ones. (page 4, line 40 – page 5, line 1)

P4, section 1.2: A key point missing in this section is that the radiative forcings due to short-lived species provided in Myhre et al. (2013) were decoupled from the CMIP5 climate model simulations that provided the basis for IPCC-AR5 chapters on climate change (historical - Bindoff et al., 2013; future - Kirtman et al., 2013; Collins et al. 2013). This made it difficult to relate the temperature responses to radiative forcing due to NTCFs and also to constrain the climate sensitivity. AerchemMIP is designed to fill in this information gap to inform the IPCC-AR6. Suggest adding a sentence to note this point.

We have added the suggested sentence to the text. (page 5, lines 6-9)

P5, L10: CO is the largest sink of OH followed closely by methane . Suggest rephrasing to "...is a dominant sink of the hydroxyl radical (OH), the primary tropospheric oxidizing agent. . ."

We have rephrased this as suggested. (page 5, line 26)

P5, L11: References for methane's influence on aerosol oxidation and natural aerosol precursors would be helpful.

We have added a reference to Shindell et al. (2009). (page 5, line 29)

P5, L12-13: Methane also influences stratospheric ozone directly by converting reactive chlorine into the reservoir hydrochloric acid HCl (e.g., Pawson et al., 2014).

We have added the suggested sentence. (page 5, line 30)

P5, section 2.1: Diverging multi-model results lead to more questions than answers. An example of model diversity from the multimodel analysis in ACCMIP was the opposing changes in trends in hydroxyl radicals simulated by the models (Naik et al., 2013; Voulgarakis et al., 2013) which did little to alleviate the uncertainty in our understanding of OH trends. Similarly, models that resolve stratospheric chemistry could produce diverging results on the impact of stratospheric ozone depletion on regional climate, reinforcing the uncertainty in our understanding. The issue of diversity in model results should be recognized in the text.

We have added the text: "As a multi-model exercise AerChemMIP will identify areas of consensus and disagreement in the answers."

P5, L40: Commas needed after NTCF emissions and CH4.

These have been added(page 5, line 13)

P6, L5: Need references for SSP3-7.0 scenario. Does this scenario include land-use changes, and if so, do they influence air pollutant emissions?

Most of the SSP3-7.0 description is in section 3.2 where we have added a reference to Fujimori et al. (2016) (page 10, line 29)

P6, L15: This in my opinion is the least framed question in the AerChemMIP proposal. It is not clear to me what perturbing emissions by 10% will tell us about the dependence of radiative forcing on uncertainties in anthropogenic emissions. Besides the uncertainty in radiative properties of NTCFs, the uncertainty in the NTCF radiative forcing comes from their spatial and temporal distributions which in turn are a function of their emission estimates. A wide variety of chemistry and physics parameterizations are implemented in CCMs which then lead to diversity in simulated atmospheric composition and therefore radiative forcing, despite implementing the same anthropogenic emissions (e.g., different tropospheric ozone, aerosols, methane lifetimes in ACCMIP). It would help to elaborate what inter-model differences in radiative forcings from 10% increase in anthropogenic emissions will tell us about the uncertainty in radiative forcing induced by emissions. Further, it is not clear from the tables 1-6 which experiments could be used to perform this analysis.

We realise that the 10% perturbations would have imposed extra work on the modelling groups. We have instead added extra single species simulations in section 3.3 in order to cover the main NTCFs. These simulations are also necessary to characterise the individual ERFs fully. This will generate sufficient data to answer the question whilst limiting the computational requests. This section (section 2.3, page 6) will now read: "The primary focus of this question is to understand the sensitivity of present-day ERF to uncertainties in estimates of the historical NTCF emissions. Indeed, while all proposed simulations rely on the usage of a central estimate, it is clear that there is a range of emission estimates (as discussed in Granier et al., 2011; Smith et al., 2011; Bond et al., 2013) that needs to be considered. While this uncertainty will clearly be region, sector and species dependent, it would be unrealistic to explore the full spectrum of variations. For that purpose, we will make use of the perturbations (pre-industrial to present-day) simulations. This is likely to provide an upper bound on the impact of uncertainties. Results from the simulations can be directly compared to the simulations in section 3.1 and analysed for differences in radiative forcing as well as air quality and overall atmospheric composition. Inter-model differences will document their varying sensitivities to emissions."

P6, L24-25: Please complete this sentence.

This should have finished "… biogeochemical feedbacks." (page 7, line 7)

P6, L29: Insert space before 5).

We have added a space.

P7, L9-10: Does the CMIP6-specified stratospheric ozone dataset use the CMIP specified forcings? If not, how will this impact results from tropospheric-only chemistry models. Will the stratospheric ozone dataset for future scenarios considered here also use future CMIP6 scenarios.

No, the new CMIP6 ozone forcing dataset will not follow CMIP6 scenarios, since these were not ready in time for implementation in the models that had to be run to generate the ozone fields. Instead CMIP5 scenarios were used with different RCPs into the future. The changes in historical emissions are deemed to have only a small effect on atmospheric chemistry, while for the future differences will exist. The full effects will need to be quantified when suitable model simulations become available. However, the main advantage of the CMIP6 over the CMIP5 ozone forcing database is that it was produced using stratosphere-troposphere resolving chemistry-climate models (CCMs) instead of patching together fields from one stratospheric and one tropospheric chemistry climate model, so that the influence of the tropospheric chemistry on the stratospheric ozone (and vice-versa) is now included appropriately.

P8, section 3.1.1: Given that the natural precursor emissions are tied to model meteorology, the perturbed NTCF simulations (hist-piNTCF, hist-piAer, hist-1950HC) will likely simulate different natural emissions compared to that in the historical simulation. Some information on how this natural component will be separated from the anthropogenic component in the simulated climate response will be helpful here.

Perturbing NTCFs will change the climate, and a changed climate will affect the NTCFs through natural emissions, chemistry, transport etc. These feedbacks are included in the response and a sentence clarifying this has been added to the text. (page 9, line 25)

P8, L33-34: Some justification for why observed SSTs and sea ice from 1870 to 2014 (available as a forcing from CMIP) could not be used for these forcing calculations, similar to the work of Andrews (2014), should be provided here.

The DECK AMIP experiment (and Andrews et al. 2014) only starts in 1979 and we wish to categorise forcings from the start of the historical (1850).

P8,L37-39: Without a reference to RFMIP simulation, this sentence is confusing. Either cut this discussion (as it is being discussed in section 5.1) or provide a reference to RFMIP here.

We have clarified this: (page 9, line 36) "Use of historical SSTs rather than pre-industrial will eliminate any effects of using an inconsistent background climate state (such as different cloud cover and natural emissions) which could affect concentrations of aerosols and reactive species, and the ERFs. The impact of background state on the diagnosis of ERFs is likely to be small (Forster et al., 2016)."

P9, section 3.2.1: Some discussion of how natural emissions will be impacted in these simulations is needed here.

A sentence has been added to explain this: "Where natural emissions are modelled interactively, these will vary with the evolving climate and will differ between ssp370 and ssp370-lowNTCF as the climate diverges." (page 11, line 9)

P10, L10: I probably misunderstand this experiment (ssp370SST-lowLu). The chemistry impacts of land-use changes can only be diagnosed if the precursor emissions (biogenic VOCs, fire emissions, dust) in the model are tied to model land-surface characteristics. If a model does not include land-surface and emission coupling then it is not clear to me how the chemical impacts of land-use changes can be simulated. It would be helpful to clarify this point.

A sentence has been added to clarify this: " … for models which include interactive schemes for emission and deposition. Not all models will model all these processes interactively." (page 11, line 26)

P10, L20-21: How different would the pre-industrial to present-day ERFs be if they were calculated using results (10 year means for PI and PD) from the transient simulations described in section 3.1.2? If the results are similar then this would avoid having to run additional time-slice simulations and considerably pare down the number of requested AerChemMIP simulations.

A sentence has been added to clarify this: "These simulations differ from those in 3.1.2 in that they use pre-industrial SSTs and maintain the same emissions (or concentrations) for 30 years. They therefore give a more accurate representation of pre-industrial conditions to present ERF than would be obtained from portions of the transient historical ERF simulations." (page 11, line 41)

P12, section 4.2: This section appears to be primarily focused on diagnostics to document aerosol forcing. There is no mention of diagnostics for gaseous NTCF forcings (e.g., ozone). Suggest adding a short paragraph on diagnostics needed to better quantify ozone, methane forcings.

We now explicitly mention gases in the discussion of ERF (page 14, line 1). We have also added the sentence: "For gaseous pollutants ozone molar mixing ratios and methane lifetime are requested in order to diagnose forcing offline." (page 14, line 10)

P13, L31: Suggest revising this sentence to: A thorough documentation of natural emissions and 3D fields of reactive gases and aerosols is needed.

The sentence has been revised as suggested(page 15, line 11)

P13, L35-36: Need a reference here.

A reference to Collins et al. (2011) has been added. (page 15, line 17)

P14, section 4.6: Which simulations will be used to evaluate the skill of models in simulating chemical composition and aerosol radiative parameters? Since modelers are encouraged to increase model output during the AMIP period, would the DECK AMIP simulations be the cornerstone for model evaluation? If so, then it would be helpful to recommend (or even make it a requirement) that each model group evaluate and document chemistry and aerosols in their AMIP simulations, to ensure data quality.

We agree that model documentation and evaluation is extremely important. CMIP does not request individual model evaluation papers in advance of submission of the model output to ESGF since this would cause a substantial delay in making the model output available to the community. We therefore will not require this for AerChemMIP either. However, model documentation will occur through the ES-DOC component of CMIP6, see https://www.earthsystemcog.org/projects/es-doc-models/ and the models will be centrally evaluated with the ESMValTool and PCMDI metrics package as soon as the output is submitted to the ESGF (see Eyring et al. 2016b). This will include the evaluation of chemistry and aerosols with diagnostics implemented into the ESMValTool. We have added this to the text. (page 16, last paragraph)

P14, L28-29: The ability of models to simulate climate is measured in terms of how well they simulate global temperature and precipitation. Are there equivalent metrics for the simulation of chemistry and aerosols in the face of large uncertainties and data gaps in observations of short-lived chemicals? Suggest adding examples of robust metrics that could be used to measure the skill of CCMs.

We have added a reference to the SPARC CCMVal Report where process-oriented diagnostics and performance metrics have been defined and successfully applied to the CCMVal ensemble. (page 16, line 17)

P14, L31-32: A reference to Lee et al. (2013) would be useful here.

This paragraph has been revised and no longer refers to BC observations explicitly. Therefore the Lee et al. reference is not needed. (section 4.6, 1st paragraph)

P16, Data availability: DOIs should be assigned to each model’s output for AerchemMIP so that proper credit can be given to each modeling group providing the data. Are there any plans for assigning DOIs to contributions to AerchemMIP?

AerChemMIP is a CMIP6-Endorsed MIP. All CMIP6 simulations will be assigned with DOIs as the sentence in the ‘Data Availability’ section already says: “The climate model output from AerChemMIP experiments described in this paper will be distributed through the Earth System Grid Federation (ESGF) with DOIs assigned.”

P16, L27-28: https://pcmdi.llnl.gov/projects/input4mips/ can be referenced here.

This has been referenced(page 18, line 34)

Anonymous Referee #2

This manuscript will serve as a key reference for those participating in AerChemMIP, a formal CMIP6 activity, which is expected to feed into the next IPCC report. The science goals of this MIP are to address the impact of aerosols and chemically-reactive gases on climate resulting from (1) historical anthropogenic emissions, (2) future policies on climate, air quality and land use, (3) climate feedbacks on natural emissions. AerChemMIP will also address uncertainties associated with anthropogenic emissions. The manuscript lays out the specific simulations needed to address these scientific goals, with a prioritization of which simulations are most critical to ensuring the success of AerChemMIP in achieving these goals.

We wish to thank reviewer 2 for taking the time to make valuable comments on the AerChemMIP description.

General comments. Overall the paper is well-written, though the final two sections need some editing to help a less-initiated reader parse the jargon as it seems to assume the reader is immersed in all the details of many MIPs past and present. The tables usefully organize the key information that a modeling center needs to decide which experiments they wish to perform.

We have followed the reviewer's suggestion to reduce the jargon in the final two sections.

Based on the tiered rankings in the Tables, it seems that some of the science questions listed in the abstract are higher priority than others and it would be useful to indicate that there are goals that AerChemMIP will certainly achieve, and then others that will be possible to achieve if the modeling community responds with a sufficient set of Tier 2 and 3 experiments. It's important then to provide a strong motivation for answering these questions, and the current phrasing of both questions 3 and 4 could be improved unless the authors feel these are best posed as requiring a yes/no answer? Maybe something like, 'How important are climate feedbacks occurring through changes in natural emissions relative to anthropogenic perturbations of the climate system' ? For Question 3, what aspects of uncertainty are addressed (historical?); isn't the scenario approach of CMIP/IPCC designed to span a range of uncertainty in anthropogenic emissions?

We have rephrased question 4 as suggested by the reviewer. Question 3 will become "How do uncertainties in historical NTCF emissions affect radiative forcing estimates?" (Abstract) Regarding the tiers, we have added: "We have arranged the experiments into 3 tiers to reflect their priority. Tier 1 experiments are those necessary to answer science questions 1 and 2 in terms of overall impacts of NTCFs and reactive well-mixed gases. Tier 2 experiments will answer question 4 and provide further detail on questions 1 and 2 by separating the effects of aerosol and ozone precursors. Tier 3 experiments contribute to question 3 and provide additional detail and speciation." (page 8, lines 23-26)

We realise that the 10% perturbations would have imposed extra work on the modelling groups. We have instead added extra single species simulations in section 3.3 in order to cover the main NTCFs. These simulations are also necessary to characterise the individual ERFs fully. This will generate sufficient data to answer the question whilst limiting the computational requests. This section (section 2.3, page 6) will now read: "The primary focus of this question is to understand the sensitivity of present-day ERF to uncertainties in estimates of historical NTCF emissions. Indeed, while all proposed simulations rely on the usage of a central estimate, it is clear that there is a range of emission estimates (as discussed in Granier et al., 2011; Smith et al., 2011; Bond et al., 2013) that needs to be considered. While this uncertainty will clearly be region, sector and species dependent, it would be unrealistic to explore the full spectrum of variations. For that purpose, we will make use of perturbation (pre-industrial to present-day) simulations. This is likely to provide an upper bound on the impact of emission uncertainties."

While Section 5 is devoted to discussing overlap with other MIPs, it might make it easier for readers and for modeling groups to prioritize their overall contributions to the many CMIP6 MIPs if some of this discussion could be incorporated into the Tables. For instance, it could be noted in a different color or in footnotes which simulations are identical to those requested by other MIPs. Alternatively, the authors may wish to create a new Table based on the information in Section 5 that allows readers to quickly identify simulations from this MIP that overlap with other MIPs. If a group can only perform a limited number of Tier 2 or 3 experiments, they may wish to prioritize simulations that address multiple MIPs and this would ease their task of identifying those simulations, at least for AerChemMIP.

Notes identifying simulations that overlap with other MIPs have been added to the table captions.

Similarly, it'd be useful to have a table explaining which DECK experiments or simulations from other MIPs are required for participation in AerChemMIP.

The DECK requirements are already described in section 3. The required simulations from other MIPS are already listed in the experiment tables. We have highlighted these in the table captions.

Somewhere it would help to articulate the rationale for the Tier categorizations. For instance, why is N2O lower priority than CH4 in Table 2?

N2O is lower priority since its chemical effects are less important than for methane. We have revised the Tier categorisations as follows:

"We have arranged the experiments into 3 tiers to reflect their priority. Tier 1 experiments are those necessary to answer science questions 1 and 2 in terms of overall impacts of NTCFs and reactive well-mixed gases. Tier 2 experiments will answer question 4 and provide further detail on questions 1 and 2 by separating the effects of aerosol and ozone precursors. Tier 3 experiments contribute to question 3 and provide additional detail and speciation." (page 8, lines 23-26)

Specific comments.

P2 L5-7. The definition of NTCF given here doesn't seem to align with that from IPCC AR5 WG1 Chapter 8 Box 8.2, which explicitly includes methane as a major motivation for using NTCF versus short-lived climate forcers. It's important to note that methane falls in both WMGG and NTCF categories, at least as defined in the last IPCC report. If the authors are revising this definition, it's important to explain this, particularly as several of the authors were lead authors on this recent IPCC chapter. But on P5 L40, NTCF emissions include methane.

We have modified the text to include methane. (page 2, line 10)

P3 L5-7, "The knowledge base used to manage air pollution to date must be updated. . .". This seems reversed to me, and in any case the phrasing could be improved. Isn't it rather that air pollution policies are driving major changes in NTCFs and we need to be sure we understand the global atmospheric composition and climate impacts from implementing these policies? With the exception of methane, it's hard to imagine that climate policies are going to have a bigger effect on NTCFs than healthmotivated air pollution policies. Or maybe the authors are simply trying to make point here that the CMIP6 scenarios will be more relevant for air quality planning than the CMIP5 RCPs?

We have revised this text: "CMIP6 will provide comprehensive information on the future large-scale evolution of atmospheric composition thus updating the knowledge base used to manage air pollution." (page 3, line 16)

P3 17-20. "Undiscovered feedback processes. . .' It's not clear what this means. Will AerChemMIP discover these processes?

We have deleted the word "Undiscovered" (page 3, line 26)

P4 L15-16. How do we know that the climate forcing from stratospheric ozone is improved?

This has been rephrased: "This has led to substantial improvements in the representation of climate forcing by stratospheric ozone in climate models since the AR4." (page 4, line 27)

P5 Section 2.1. Are the historical emissions going to be the same as what was used in CMIP5?

No. CMIP6 will use an updated historical emission dataset, see https://pcmdi.llnl.gov/projects/input4mips/ and http://www.globalchange.umd.edu/ceds/ceds-cmip6-data/.

P5 L 28-30 vs L32-33 seem like contradictory statements?

We have improved the text to remove the contradiction: (page 6, 2nd paragraph)

"A detectable regional response to inhomogeneous climate forcing concerns the Southern hemisphere summertime surface circulation changes and is induced by the Antarctic ozone hole as an indirect response to stratospheric ozone depletion from increasing halocarbons. These changes have been argued to lead to changes in the subtropical jet position, rainfall patterns, ocean circulation, and possibly sea-ice cover (Arblaster and Meehl, 2006; McLandress et al., 2011; Polvani et al, 2011). The relative role of these ozone-induced changes for observed Southern hemisphere summertime climate compared to other anthropogenic forcings and natural variability is not fully resolved by the scientific community with some contradictory studies in particular for the Antarctic sea-ice response (WMO, 2014). Hence there is a need for a multi-model ensemble of simulations that resolve stratospheric chemistry to isolate the role of stratospheric ozone depletion."

P6 L5 How do these SSP emission scenarios for air pollution compare with those used from IIASA in the ECLIPSE project (e.g., Stohl et al., Evaluating the climate and air quality impacts of short-lived pollutants, ACP, 2015)?

This discussion has been moved to section 3.2(page 10, lines 24-28): "The medium strength of pollution control corresponds to following current legislation (CLE) until 2030 and progressing three-quarters of the way towards maximum technically feasible reduction (MTFR) thereafter. Strong pollution control exceeds CLE and progresses ultimately towards MTFR. The rate of progress is different for high, medium and low income countries. Weak pollution controls assume delays to the implementation of CLE and make less progress towards MTFR than the medium scenario. For more details see Rao et al. (2016)."

P6 L12. Please elaborate on what is meant by 'mixed results'.

We have removed the comment on "mixed results" (page 6, line 29)

P6 L14. What are the statistically significant differences between? (2015 versus 2055 or between SSP3-7.0 and SSP3-7.0 with maximum feasible reductions applied?)

This referred to increasing the number of ensemble members, but we have removed this sentence as it is not relevant to section 2.2. (page 18, line 34)

Section 2.3. How will artificially increasing present-day emissions help with quantifying uncertainty? Isn't there more uncertainty associated with the time-evolution of emission changes (i.e., when BC emissions versus SO2 emissions peak)? This question and approach to answering it should be elaborated on.

The reviewer raises relevant issues, some of which are outside of the scope of AerChemMIP, which focuses on uncertainties related to atmospheric chemical processes. AerChemMIP results can then be used in future work to access the influence of time trends (e.g. Smith and Bond 2014, ACP, 14 537–549.). As noted elsewhere in this response, the 10% perturbation experiments have been replaced in order to reduce the modelling burden.

Section 2.4. It would help to provide more rationale for the selection of the 6 sensitivity simulations proposed. While marine biogenic aerosols are noted, only DMS (not organics) is considered. A large climate feedback is likely to be through methane from wetlands and yet that is not mentioned. What about feedbacks via N2O or halogens on stratospheric (and tropospheric) ozone? Maybe this is limited to what processes are typically included in current climate models, or some of these processes are addressed in other MIPs, but this should be stated.

We agree that methane and N_2O feedbacks are important, but we do not need extra experiments to characterise these. As noted elsewhere, we will be able to diagnose any changes in, for example, natural methane emissions in the historical experiments due to feedbacks. We have added: "To do this it will be necessary to quantify the climate response to the heterogeneous forcing patterns from naturally emitted short-lived species (the climate responses to WMGHGs are already covered in section 2.1)." (page 7, line 10)

P7 L12-15 Include this point in the tables so it's very clear that the more complex con- figuration is always encouraged for AerChemMIP as opposed to preferring minimum configuration for inter-model consistency.

We have added this point in each table: "The "AER" suffix means that at least interactive aerosols are required, interactive chemistry should be active if available."

P7 L35. How realistic are these requests? Is it possible to further prioritize some of the Tier 1 into lower Tiers? How many years are needed for DECK + other MIP simulations needed for entry to AerChemMIP?

We have added the text "This includes 30 years for pre-industrial fixed SST control in common with RFMIP. In addition, models should have been run for the DECK experiments (501 years excluding control) and 3 ensembles of SSP3-7.0 for ScenarioMIP (41 years each)." (page 8, line 27)

P8 L13-14. But does the net impact of NOx emissions depend on how it affects nitrate?

Yes it does. We have clarified this: "even though it will generate both ozone and nitrate aerosol in models." (page 9, line 8)

P8 L20. Where is 'historical' defined?

We have clarified this: "which is a simulation from 1850 to 2014 with all forcings applied (Eyring et al. 2016)" (page 9, line 14)

P9 L8-11. Confusing. Is this assuming everyone is also participating in DAMIP? Explaining this in the tables, or with a new table, would help here.

No, we are just explaining that AerChemMIP does not address responses to homogeneous forcing. Modelling groups are free to choose whether or not they wish to participate in DAMIP hist-GHG. We have clarified this (page 10, line 8): "The ERF pattern from these simulations is expected to be relatively homogeneous, although their chemical effects on ozone and secondary aerosols may be less so therefore AerChemMIP does not include any experiments to derive the climate responses to methane or nitrous oxide forcing. The climate response to homogeneous forcing is quantified in Detection and Attribution Model Intercomparison Project (DAMIP) from the hist-GHG simulation (as historical, but only the WMGHG forcings evolve)."

We have also clarified this in section 5 (page 17, line 2): "AerChemMIP is self-contained in so far as the questions posed can be answered by running only the experiments listed here, the DECK and historical. For a full analysis of the past and future climate-composition interactions (including unreactive greenhouse gases) in the CMIP6 chemistry climate models we recommend that as many as possible of the tier 1 simulations of RFMIP (Pincus et al., 2016), DAMIP (Gillett et al. 2016), ScenarioMIP (O'Neill et al., 2016), C4MIP (Jones et al. 2016) and LUMIP (Lawrence et al., 2016) are run with the AER CHEM model configuration and with AerChemMIP diagnostics."

P9 L20 explain the 7.0 next to SSP3. Will this paper be updated to reflect the ongoing discussions with ScenarioMIP? Otherwise should be sure to encourage readers to check the AerChemMIP website.

We have clarified the reference scenario (page 10, line 28): "we choose the reference scenario to be SSP3-7.0 "Regional Rivalry" without climate policy (7.0 Wm-2 at 2100), (Fujimori et al. 2016)" and have explained that the AIM group are generating the perturbation scenario (page 10, line 32). "The perturbation experiment to this within AerChemMIP will be generated by the AIM group using the same socio-economic scenario as in Fujimori et al. (2016), but with "Strong" levels of air quality control measures SSP3-7.0-lowNTCF"

P10 L5-6. What exactly is being compared from the previous coupled model simulations with the magnitudes and patterns of ERF?

We have clarified this: (page 11, line 17) "Comparison between the magnitudes and patterns of ERF with surface temperature and precipitation from the previous coupled model simulations will allow quantification of the efficacy of the NTCFs to affect climate."

P10 L7. What are the different groups of NTCFs?

We have clarified that these are "emissions (aerosols, ozone precursors, black carbon)" (page 11, line 22)

P10 L8-9. Methane isn't Tier 1 – why?

The ERF of the methane change will approximately scale with concentration, and so could be derived from *histSST-piCH4* or *piClim-CH4*. The *ssp370SST-lowCH4* simulation will however provide extra information on non-linearities and the air quality benefits from methane mitigation under the SSP3-7.0 scenario.

P10 L13. How is land use affecting NTCFs? Do these models have their NTCF emissions tied to specific land-use categories imposed in the model? How would this work for groups using MEGAN driven by present-day base emission capacity maps?

We have clarified this: "… for models which include interactive schemes for emission and deposition. Not all models will model all these processes interactively." (page 11, line 26)

General question on Section 3.2.2 versus 3.3. What is the difference between ERF simulations and prescribed SSTs? Can these terms be used interchangeably?

We now refer to the quantities in section 3.2.1 (and 3.1.2) as "transient ERFs" to distinguish them from the ERFs in section 3.3. We have also added text to section 3.3: "These simulations differ from the transient ERF simulations in 3.1.2 in that they use pre-industrial SSTs and maintain the same emissions (or concentrations) for 30 years. They therefore give a more accurate representation of the pre-industrial to present ERF than would be obtained from portions of the transient historical ERF simulations." (page 11, line 41)

P11 L9-11. How good is the assumption of the same climate response to ERF from any species?

The feedback parameters calculated here (in W/m2/K) do not make any assumptions about the climate response to ERF. How these parameters are then subsequently used would depend on assumptions made, but that is moving out of the scope of this paper.

Section 4 is clearly written assuming the reader has some knowledge to parse all the jargon and may benefit from re-organization. It might help to include a table translating 'Mon3d', 'Mon2d', etc. What is 'MonDay2d' – typo? It gets even worse in Section 4.1

We have rewritten this section avoiding jargon and using the recently agreed CMIP6 nomenclature

P12 L5-6. What is COSP simulator data? Similarly, the long section 4.2 might be possible to shorten by moving information there to an appropriate table.

We have added text to explain what the COSP simulator is: "To facilitate the exploitation of A Train satellite data in numerical models, a system has been developed that allows simulation of the signal that CloudSat/CALIPSO would see in a model-generated world. It is a flexible tool to simulate active instruments in models (climate, forecast, cloud-resolving)." (page 13, line 25)

P12 L35-38 seems to jump from talking about quantifying natural emissions to anthropogenic. Aren't separate diagnostics needed for natural versus anthropogenic? Please clarify.

This was trying to explain that we would ask for total and anthropogenic emissions (from which we could calculate natural). The anthropogenic emissions are specified by CMIP6, so we have modified these sentences to say that we ask for total emissions and assume anthropogenic emissions are as CMIP6. (page 14, line 15)

P14 L9. What are these new transport tracers? Define them in a table or a separate section to call attention to modelers that they may need to add some new development to participate in some of the science for this MIP?

We have added a link to the definition of these tracers: 'In addition, two artificial transport tracers will help to track changes in tropospheric transport between hemispheres (aoa_nh and nh_50; see definitions in the CCMI-1 data request at http://blogs.reading.ac.uk/ccmi/data-requests-and-formats/).'(page 15, line 32)

P14 L9 and L15-16 repeat requests for loss rates of methane, CO, N2O.

This repetition has been removed.

P15 L26-30. Why can't the DAMIP and AerChemMIP requests be the same to reduce the number of requests to modeling groups?

The only overlap is between hist-piAer (AerChemMIP) and hist-aer (DAMIP). The DAMIP protocols will not work for interactive ozone chemistry. For historically evolving NTCF and ODS runs the chemistry will need to see historically evolving methane and N_2O concentrations in order to reproduce the historical ozone evolution, whereas in DAMIP these are fixed to PI levels. The stratospheric temperature differences between PI and present would also mean the DAMIP setup would fail to reproduce the present day stratospheric ozone when running their hist-stratO3 experiment with interactive chemistry.

P16 L20 what is 1pctco2 / abrupt4co2?

This is now explained: "…the DECK 1% yr^{-1} CO_2 (1pctco2) or 4×CO_2 (abrupt4co2) simulations…"(page 18, line 23)

Figure 1. Not sure the figure adds much beyond what is in the Table, and it's hard to read. Consider converting to a regular table.

We have converted this to a regular table.

Figure 2. This implies that the net impact of air quality controls are to lower surface temperature, but wouldn't reductions in aerosols actually warm in the near term? Is warming from aerosol reductions being offset by substantial methane controls here?

This schematic originally assumed methane controls as well. It has now been redrawn to show a warming from aerosol mitigation.

Alex Archibald:

Collins et al. present a nice summary of the rational and design of the Aerosol Chemistry Model Intercomparison Project (AerChemMIP), a CMIP6 endorsed project that is designed to quantify the climate and air quality impacts of aerosols and chemically reactive gases in the atmosphere. AerChemMIP is a huge project and reading through the number of simulations in the tables it's clear that a huge amount of thought has gone into this and by and large I feel that the requests are reasonable for answering the overarching science questions stated. However, one of the primary reasons for undertaking these experiments (and previous experiments like ACCMIP, CCMI etc) is to understand the functioning of coupled chemistry-climate models in a multi model framework. The climate community are in a great position with the DECK experiments in that they can use these to understand how as a function of time the evolution of models leads to improvements/changes in our understanding of the physical climate system. But what parallel is there for the aerosol and chemistry community? What is our DECK experiment(s)? I think a short discussion on this would make a useful addition to this paper.

The science goal of AerChemMIP is to quantify the contributions of reactive gases and aerosols to the climate in the CMIP6 simulations. Since the climate modelling centres will be contributing to a wide range of CMIP6 MIPs we deliberately limited the number of experiments and focussed these on the above science goals. Process-oriented evaluation of chemistry-climate models on the other hand is covered by the more specialised aerosol and chemistry projects Aerocom and CCMI, which will feed information into AerChemMIP with respect to model performance and how it may affect results. We agree that the aerosol and chemistry modelling communities would benefit from the design of at least some additional idealised experiments to characterise model performance from one generation to another, and suggest that Aerocom and CCMI remain the appropriate projects to pursue this. As a minimum step in this direction, we note that the AerChemMIP timeslice ERF experiments (piClim-…) will allow us to explore the model responses to step changes in emissions of individual species or groups of species. We have added text to clarify this additional value of the ERF experiments.

Also, I noted the following minor errors in the manuscript:

Page 3 line 10: "contribute" should be "contributes". Page 3 line 38: after "chemistry-climate models" could you add "(CCMs)"?

Page 3 line 38: similarly, after "chemistry transport models" could you add "(CTMs)"?

Page 4 line 3: Why is "Chemistry-Transport Models" (a) capitalised? (b) hyphenated?

Page 4 line 7: adoptions above mean you could remove "(CCMs)".

Page 4 line 21: The reference list has the "(IPCC 1995)" citation as being published in 1996.

Page 6 line 2: consider replacing "...)" with "etc)"

Page 6 line 25: remove "in" at the end of the line.

We have made all these changes.

M. Michou martine.michou@meteo.fr

1 More important remarks/questions

• p.14 para 5.1 The authors indicate that there are ’considerable synergies between AerChemMIP and RFMIP’. However the experiments to diagnose transient ERFs are diffently designed, even though authors note that “the impact of different approaches ...have been estimated to be small”. Given the amount of work involved in managing these CMIP6 simulations, could not the protocol for these specific simulations under ’prescribed SST experiments’ be the same? A similar comment applies to paragraph 5.2 : what is the justification for asking for two different protocols in AerChemMIP and DAMIP in some coupled model experiments with regards to species fixed at PI levels or historically evolving?

For DAMIP the **overlap is between hist-piAer (AerChemMIP) and hist-aer (DAMIP). For RFMIP the overlap is between histSST-piNTCF (AerChemMIP) and piClim-histAerO3 (RFMIP). The RFMIP and DAMIP protocols will not work for interactive ozone chemistry. For historically evolving NTCF and ODS runs the chemistry will need to see historically evolving methane and N_2O concentrations in order to reproduce the historical ozone evolution, whereas in DAMIP and RFMIP these are fixed to PI levels. The stratospheric temperature differences between PI and present would also mean the DAMIP setup would fail to reproduce the present day stratospheric ozone when running their hist-stratO3 experiment with interactive chemistry.**

• p.12 l.8 : the DynVarMIP project is mentioned here: simulations that are of interest to several MIPs are of special interest to the modelling community at large. The AerChemMIP paper should detail both the simulations and the diagnostics that are behind these simulations. I understand this is an additional burden to the authors of the article, but this would benefit to the entire community. If not done in the paper, then each individual modelling team will have this burden, which in the end will result in a much larger community burden. The same comment applies to other parts of the article when other MIPs are mentioned.

The remark in our paper was intended to encourage modellers to also contribute to other MIPs. Diagnostics should be chosen such that all requested output from linked MIPs are included.

• Another issue in the data request is the vertical coordinate: in the excel files of the aerocom wiki page, it is mentioned that 3D data should be provided on model levels. In our case, our model has 91 model levels with about half of the model levels in the troposphere. What is the scientific justification to provide tropospheric aerosol information on stratospheric levels?

Tropospheric parameters are strictly not needed for stratospheric levels, and vice versa. Output for all model levels had been chosen to limit output complications, confusion and inconsistent data. We have added a comment to those variables in AerChemMIP for which we need output of 3d data just in the troposphere. To include all tropospheric data under all conditions of mixing we define troposphere to levels up to ca 20 km altitude.

• in the article a distinction is made between models without and with interactive gas-phase chemistry. It would be clearer to distinguish between four types of models (1) without interactive chemistry (2) with interactive tropospheric chemistry only (3) with interactive stratospheric chemistry only (4) with both tropospheric and stratospheric chemistry. Our model includes interactive aerosols and stratospheric chemistry with the chemistry calculated down to the mid-troposphere (560hPa). So for us, so-called NTCF simulations and Aer simulations are the same. But we will rather name our simulations xxxNTCF that are Tier1 simulations. Thank you for any comment you may have on this choice.

We have clarified whether tropospheric or stratospheric chemistry is required: “The suffix “CHEM^T” or “CHEM^S” indicates interactive tropospheric or stratospheric chemistry is the minimum needed for these experiments.”. We have also added a comment on the NTCF/Aer simulations: “For models without interactive tropospheric chemistry, the tier 1 …NTCF and tier 2 …Aer simulations will be identical so only need to be run once.” (page 7, line 36)

- the names of the experiments in the paper and in the official data request web page (https://www.earthsystemcog.org/projects/wip/CMIP6DataRequest) are not the same, and a few experiments appear in the paper and not in the data request web page or vice-versa. In our case at CNRM, we have chosen to use the data request to build as automatically possible experiment designs (names, list of variables, ...). In the case here, what are the official experiments IDs? We need this information in the coming couple of weeks as for a number of constraints our CMIP6 simulations will start on 1 November.

The names in the paper are the correct ones. We have submitted these changes to the data request web page.

- a comment similar to the previous one concerns the diagnostics: p.11 l.30: it is mentioned that the diagnostics are assembled in two excel files (https://wiki.met.no/aerocom/aerchemmip/diagnostics), and that the definite request will be found in the CMIP6 Data request web page. At this stage there is no obvious link between the two lists. For instance in the CMIP6 Data request web page the data are not presented in 6 sheets as specified in the article p11 l35. A second example, is that there does not seem to be any request for 2D zonal monthly mean data in the CMIP6 Data request web page while there is the excel files of wiki.met.no. We can hope that the CMIP6 Data request web page will coincide at some point with the wiki excel files but when will that be? In the mean time, I would suggest to add as an appendix to the paper the final list of variables with all their characteristics (CMOR names, units, method of calculation if required, etc...) to which the CMIP6 data request will comply. Some MIP papers do not include this list, but others such as the OMIP or the C4MIP papers do, and in the end the entire community, both the one that puts together the MIP simulations, and then the one that will analyse the simulation outputs will benefit from that.

The updated excel tables on the AerChemMIP wiki pages were unfortunately also not up-to-date earlier in the summer. They have been updated now. We will keep the excel tables updated on the wiki page and they shall be consistent with the data request on November 1st. We prefer to not add an appendix to the paper, because the data request is still under review and in exchange with BADC.

- what is the recommendation for the aerdaily data: average from 6hourly data, or instantaneous data once per day, or?

We have added a note to the data request for each variable indicating the preferred time averaging method.

- p.12 l.19: please detail here which specific CMIP6 variables will come out of this additional call to the radiation call. Please indicate the name, and the method of calculation (equation xx from Ghan 2012 for instance) if appropriate. Such details are presented in other CMIP6 MIP description papers, such as the OMIP paper, and in the end it ensures common grounds for these variables which were not part of CMIP5. Such details should appear at least for all non CMIP5 variables.

For the repeated aerosol-free call to the radiation code we have now added the necessary variables in the data request.

- For the forcing, how are such fields as the swtoaasaer (that appears in the CMIP6 Data request) generated in the course of the simulation? The same question applies to diagnostics such as the swtoaasdust.

We agree these variables are indeed difficult to output during most of the simulations. They have therefore been removed from the standard output request. A remark has been added that the ERF simulations for the different anthropogenic aerosol components, in combination with aerosol-free calls to the radiation code, will be used to derive forcing for aerosol components.

- For the CMIP5 variables, there could/should be some coordination between AerChemMIP and other MIPs such as DAMIP or RFMIP to ensure that variables are not requested twice under two different names. For instance, ozone is requested as the tro3 and as the o3 CMOR variable, and it is not clear what the justification is for providing the same variable under different names.

These variables are being rationalised with the other MIPs to avoid duplication as much as possible.

- p.9 l.31: Could you explain why AerChemMIP future simulations should end in 2055 ? I admit it would have an additional cost to continue them until 2100, but it could be considered at least for a few simulations.

We have clarified this: "SSP3-7 does run out until 2100 as part of ScenarioMIP but, to save computation expense, AerChemMIP is only requesting simulations out to 2055. This is the time period over which the divergence in air quality policies is expected to be largest." (page 10, line 37)

- p.13 l.6: 'Speciated AOT diagnostics are suggested': I could repeat here my previous comment. Thank you for listing diagnostics explicitly, describing how to obtain these diagnostics (additional calls to the radiative code, specific simulations...)

The way of computation is now detailed in the paper following recommendations from earlier AeroCom discussions on how to do this. (page 14)

- p.13 l.19: same comment with regards to "with additional radiation calls": thank you for listing diagnostics explicitly

Diagnostics for additional radiation calls are now listed in the paper and in the data request. (page 15, 1st paragraph)

- p.13 l.23: same comment as above with regards to "Fluxes for this repeated call have to be stored separately" : what are the names of fluxes?

Extra flux output variables have been added to the data request

- p.13 p.24: please list explicitly the aerosol-oriented ERF experiments

These are the experiments with the prefix "piClim". This has been added to the paper. (page 15, line 4)

- p.14 para4.6 : again here, it would be very useful to have a clear list of the diagnostics concerned.

This section now includes more details on the diagnostics and references to CCMI and aerocom. We have complemented the text by explaining better the most relevant diagnostics, however the full detail is contained in the CMIP6 data request and our excel tables. (page 16)

2 Additional remarks/questions

- abstract l.32: please indicate that a number of additional simulations, and not only specific diagnostics, are part of AerChemMIP

We already mention the extra simulations in the abstract.

- p.2 l.16: please indicate a reference for the ERF here

We have moved the Myhre et al. reference up here. (page 2, line 22)

- p.3 l.10: the contributions listed appear rather different from the ones in the abstract. Why is it so?

We have removed the numbering so it is clear this is just a description, not a list. (page 3, lines 19-24)

- p.3 l.21: "Finally, additional..." : the sentence does not appear to be logical with the rest of the paragraph. A reformulation would certainly facilitate the reading.

We have rewritten this: "Finally, uniform evaluation of the models will expose systematic biases and better constrain our overall goal of quantifying the role of aerosols and reactive gases on climate forcing." (page 3, line 30)

- p.3 l.36: "the model setups for CMIP5 and ACCMIP tended to be different". Could you give more details on these differences ?

We have clarified this: “In practice, the model setups for CMIP5 and ACCMIP tended to be different (in terms of resolution and complexity)” (page 4, line 7)

- p.6 l.18: I may be wrong, but I have not seen in this paper any simulation with an increase of 10%. Please include here simulation names for the sake of clarity.

We realise that the 10% perturbations would have imposed extra work on the modelling groups. We have instead added extra single species simulations in section 3.3 in order to cover the main NTCFs. These simulations are also necessary to characterise the individual ERFs fully. This will generate sufficient data to answer the question whilst limiting the computational requests. This section now reads: “The primary focus of this question is to understand the sensitivity of present-day ERF to uncertainties in estimates of the historical NTCF emissions. Indeed, while all proposed simulations rely on the usage of a central estimate, it is clear that there is a range of emission estimates (as discussed in Granier et al., 2011; Smith et al., 2011; Bond et al., 2013) that needs to be considered. While this uncertainty will clearly be region, sector and species dependent, it would be unrealistic to explore the full spectrum of variations. . For that purpose, we will make use of the perturbations (pre-industrial to present-day) simulations. This is likely to provide an upper bound on the impact of uncertainties. Results from the simulations can be directly compared to the simulations in section 3.1 and analysed for differences in radiative forcing as well as air quality and overall atmospheric composition. Inter-model differences will document their varying sensitivities to emissions.”

- p.6 l.24: As far as volcanic SO2 emissions are concerned, is there any work on these emissions considered in AerChemMIP ? Or at least any dataset provided ?

Volcanic emissions have no specific focus in AerChemMIP, but are one of the forcing datasets for CMIP6 (see https://pcmdi.llnl.gov/projects/input4mips/). There is also a CMIP6-Endorsed MIP that specifically looks at volcanic emissions (VolMIP).

- p.8 l.41: could you precise if there is any specific spin-up period for the simulations beginning in 1950 ?

The simulations beginning in 1950 should start from the appropriate date in the *historical* as specified in the paper.

- p.10 para 3.2.2: the experiment concerning CH4 should be mentioned in this subsection, as it appears in Table 4

This has been added. (page 11, line 24)

- p.11 l.20: the article indicates that “the data request is structured according to overarching analysis subjects”: please indicate that these categories are presented later in the paper.

We do not use these labels anymore, but instead refer forward to the relevant subsections of section 4. (page 13, line 2)

- p.11 l.36: as of today, there is no reference to the 6 sheets listed in the article in the CMIP6 Data request web page.

The data request web page is being updated to be consistent with this paper.

- p.12 l.6 : in which document can we find these tables? The tables in the CMIP6 Data request web page are different (Amon, AmonAdj, Lmon, OMon, aerannual, aerdaily, aerfixed, aerhourly, aermonthly, cfDay, cfMon, cfSites, day)

The data request web page is being updated to be consistent with this paper.

- It seems to me that almost all specific AerChemMIP variables are given in Priority 1. Is there any possibility to share out these variables between the three priorities? Or at least in terms of temporal frequency or number of vertical levels?

We have already removed tier 2 and 3 variables from the official data request for simplicity. These may be added later through specific requests from AeroCom and CCMI to participating modellers. Requesting specific output for specific experiments will make the data request for the modellers too complicated.

- in tables 5 and 6, it is said “using pre-industrial climatological average SSTs”: what is the time length recommended for the average?

We now clarify that these are 30 year averages for consistency with RFMIP.