Utrecht University

Planbureau voor de Leefomgeving

Correcting temperature patterns in Integrated Assessment Models

Implementing local air quality and land cover temperature effects in IMAGE

Author:
L.Q. Verboom

Supervisors:
Prof. dr. M.C. Krol*
Drs. J.C. Doelman**
Drs. M. van den Berg**

A thesis presented for the degree of
Master of Science

*Institute for Marine and Atmospheric research Utrecht
Utrecht University

**Netherlands Environmental Assessment Agency

The Netherlands

December 2020

Abstract

The Dutch Environmental Assessment Agency (PBL) uses an integrated assessment model (IAM), IMAGE, to address a set of global environmental issues and sustainability challenges. IMAGE is unique compared to other IAMs because of its high spatial resolution for land-based processes. Among others, IMAGE is used to develop a set of unique scenarios, the Shared Socio-economic pathways (SSPs), to facilitate the integrated analysis of future climate impacts, vulnerabilities, adaptation, and mitigation.

In order to model temperature change at high resolution, a two-step approach is used. First, global temperature change is modelled with a reduced complexity climate model, which is scaled down using Earth System Model (ESM) results in a second step. However, SSPs studied with IMAGE can significantly deviate from the current downscaling scenarios, particularly in terms of air quality (AQ) and land cover (LC). This implies that the down-scaled temperature patterns only account for global effects of aerosol and LU-change effects, while strong local effects of AQ and LC changes are expected.

This study focuses on improving the temperature patterns by implementing temperature corrections for local AQ and LC effects. For AQ, a source receptor atmospheric chemistry model is used to model the concentration fields for aerosol components. Future concentrations are subsequently scaled with a set of controlled experiments and its temperature changes. For LC, aggregated remote sensing data on local biophysical LC change effects is used.

When the correction models are compared to ESMs in validation experiments, the AQ corrections produce results which are more or less in line with ESM results. Comparisons with fixed SST ESM runs show better agreement, with correlation coefficients up to 0.88. For LC, correlation between the corrections is small at best for some ESMs and absent in most cases.

Applying these methods on the newly developed SSPs, local AQ corrections of 0.5 K and up to 1 K are calculated in South-West, South and East Asia. For LC, the correction magnitudes are smaller and locally varying and primarily found in South America Africa. Global average temperature is only marginally affected by the corrections. Applying the temperature corrections in the IMAGE land model leads to changes in the order of 1% for crop yields and living biomass in China and India.

Acknowledgements

My interest in this topic was sparked by a guest lecture by David Gernaat, when I was taking the 'Complex Systems: Economy and the Planet' course. At this time, I was seriously considering quitting the Masters program, especially because I had not a single clue what I wanted to do after graduating. This internship (and the prospect of it at the time) gave me a better idea of this, and has been a driving force since. In other words, if it hadn't been for David, I wouldn't have ended up doing this project. He also helped with setting up the internship, for which I'm very grateful.

I would also like to thank Gregory Duveiller for providing the data for the land cover temperature corrections and meeting up with us to discuss the results and application of the land cover corrections. Also, I would like to thank Elke Stehfest, for reviewing my methods and joining the discussion on the implementation of the correction models. Moreover, I am thankful for the involvement and support of Detlef van Vuuren throughout the whole process. Not only did he help with setting up the study, he also provided feedback on the research proposal and reports and gave some very good recommendations on technical details.

A big thank you goes out my PBL supervisors Maarten van den Berg and Jonathan Doelman. After I spent a mere 3 weeks at the office, Covid-19 struck the Netherlands and working from home became the new standard. Maarten and Jonathan immediately stepped in and made sure we kept in touch by meeting up twice a week. They were also really helpful when I was writing the research proposal, a part which turned out to be quite a challenge for me. Overall I can say that I've learned a great deal from them, and without their dedication and involvement, this project wouldn't have come so far as it did. A special thank you also goes out to my Utrecht University supervisor Maarten Krol. Despite his busy schedule, Maarten was always available to help with some aerosol related issues and questions. Moreover, I've learned a great deal from his extensive feedback on my writing and presentation.

Last but not least, I would like to thank Lotte. She has been really supportive and made sure that I did not 'loose myself' throughout the whole process.

I have been thrilled to work on this subject in an environment with experienced and dedicated people. I learned a lot on all kind of topics and really enjoyed my time at PBL. Again, I thank everyone who has been involved in this project.

List of abbreviations

AerchemMIP	Aerosol Chemistry Model Intercomparison Project
AQ	Air Quality
BC	Black Carbon
BCH	Biochemical
BPH	Biophysical
CMIP6	Coupled Model Intercomparison Project (Phase 6)
ESM	Earth System Model
ISIMIP	Inter Sectoral Impact Model Intercomparison Project
IMAGE	Integrated Model to Assess the Global Environment
LPJmL	Lund-Potsdam-Jena managed Land
LC	Land cover
LU	Land use
LUMIP	Land Use Model Intercomparison Project
MAGICC	Model for the Assessment of Greenhouse Gas Induced Climate Change
MMM	Multi Model Mean
NTCFs	Near-Term Climate Forcers
OC	Organic Carbon
POM	Particulate Organic Matter
RCP	Representative Concentration Pathway
scenarioMIP	Scenario Model Intercomparison Project
SSP	Shared Socio-economic Pathway
WMGHGs	Well-Mixed Greenhouse Gases

Contents

1 Introduction

Since the start of the industrial revolution, increasing greenhouse gas emissions from human activities have led to increased longwave radiation trapping, resulting in unprecedented rapid changes in the climate on earth. Without drastic action, the impacts and associated damage of climate change will increase over time, especially in vulnerable countries (Dinar et al., 2006). In addition, mitigation attempts might fail if tipping points in the climate system are triggered (Hoegh-Guldberg et al., 2018). As time is running out, exploring and assessing climate change mitigation scenarios is becoming increasingly important, especially in the context of international treaties such as the Paris Agreement (Rogelj et al., 2018).

Currently, integrated assessment models, capable of analysing the effects of climate change and policy, are frequently used to study these problems and relied upon by the Intergovernmental Panel on Climate Change (IPCC). Although climate change modelling is one of their most important features, current models lack correct representation of local climate effects from air quality and land cover changes. As these processes have a large impact on local climate, a correct representation in integrated assessment models is important and will be the main focus of this study.

1.1 Air quality

Air quality (AQ) is a metric for the contamination of air by tropospheric ozone, particulate matter and a range of other chemical substances. For this study, particulate matter, i.e. aerosol, is the primary focus as it strongly affects air quality as well as climate.

In terms of climate, aerosols scatter and absorb radiation and have a direct overall global cooling effect of -0.45 Wm^{-2} (-0.95 to 0.05 Wm^{-2}, 5% to 95% confidence range). Additionally, aerosol particles act as cloud condensation nuclei as they facilitate the condensation of water vapour. The effect of this interaction is twofold: The increased cloud droplet number results in a higher cloud albedo as more light is scattered, and clouds have longer lifetimes as it takes longer to rain out (Twomey, 1974). The total indirect aerosol effect is estimated to have a global radiative forcing of -0.45 Wm^{-2} (-1.2 to 0.0 Wm^{-2}, 5% to 95% confidence range) (Myhre et al., 2013).

In contrast to the decadal or centennial timescales of well-mixed greenhouse gasses (WMGHGs), aerosols have lifetimes in the order of two weeks. Their environmental consequences are thus often focused in the vicinity of the emission source. For example, the Indo-Asian region is notorious for its poor air quality from high aerosol loading, caused by vast industrial and domestic emissions from fossil fuel and biomass burning. A field study in this region revealed that the total radiative aerosol forcing at the top of the atmosphere in this region (in the dry season) was approximately -5 Wm^{-2} versus the global annual forcing by aerosols of roughly -1.5 Wm^{-2} (Ramanathan et al., 2001).

Regarding air quality, the minuscule particles which form the aerosol can be inhaled and cause or worsen respiratory and cardiovascular diseases. In central China for example, the high SO_2 emissions lead to strong sulphate aerosol loading. The resulting poor air quality accounts for estimated health costs equivalent of 3% of the GDP (Xu and Yang, 2020), underlying the importance for air quality regulations.

Locally, aerosols have a net warming or cooling effect depending on the composition. Therefore it is important to take the temperature effect into account for the development of environmental policies as air quality improvements might reduce or enhance local warming (Sherwood et al., 2015).

1.2 Land cover and land-use change

Land cover and land-use change is driven by a combination of human activities (e.g. agriculture) and climate change, and has a major impact on global and local climate. The biochemical effect, describing the forcing by land-use (LU) change emissions (and effects on the carbon cycle), has been well documented and taken into account in climate treaties such as the Paris Agreement. However, the biophysical effect, describing the change of physical surface properties from land cover (LC) change, is often not included in assessment modelling, while it plays an important role as well: Between 2003 and 2012, the global mean warming on land through biophysical processes corresponds to 18% of the global biochemical temperature change from land-use changes emissions (Alkama and Cescatti, 2016).

The biochemical effects of LCLUC on temperature vary diurnally, seasonally and spatially and are highly dependent on the type of vegetation change. For instance, converting forests to grasslands will change the albedo and the amount of evapo-transpiration changes, which is balanced by sensible heat flux increase, ultimately resulting in temperature increase (West et al., 2011). To illustrate this: a local 50% loss of tree cover in a tropical forest is associated with at least 1 K temperature increase. On the contrary, deforestation in boreal zones will uncover the underlying snow, resulting in a higher surface albedo and a decrease in temperature (Prevedello et al. (2019), Duveiller et al. (2018c), Duveiller et al. (2020)). Policies involving land use and vegetation change, e.g. carbon sequestration through a- and reforestation projects, will result in extra local climate changes depending on the region. In the tropics, this effect will be beneficial for climate change mitigation or adaptation as local climate is cooled. However, in the boreal regions the biophysical effect will be counter effective and enhance local climate change.

1.3 Shared Socio-economic Pathways and Integrated Assessment Models

In order to provide a common basis for the exploration of climate change, associated impacts and related policies, the climate change research community recently adopted a new set of future scenarios, the Shared Socio-economic Pathways (SSPs) (Riahi et al., 2017). The framework consists of five unique future development pathways with specific energy, land use, and emissions implications, driven by varying global socio-economic development (O'Neill et al., 2014). For example, in SSP1 the world focuses on sustainability and reducing inequality, while the emphasis on economic growth shifts towards human well-being, resulting in low mitigation and adaptation challenges. In contrast, SSP3 is characterized by regional rivalry where countries focus on domestic issues at the expense of broader-based development, resulting in material-intensive consumption, persisting or growing inequalities and overall high challenges for mitigation and adaptation (Riahi et al., 2017).

Without interference from climate change mitigation policies, an SSP typically results in moderate to high radiative forcing (and thus climate change) at the end of the century, which is considered as the baseline case. However, policies such as pricing of greenhouse gas emissions can be implemented within SSPs to mitigate from climate change and meet a specific radiative forcing target at the end of the century. Here, a 2.6 Wm^{-2} radiative forcing target is in line with the Paris agreement, as it corresponds to a 66% probability to keep global warming under 2 K compared to pre-industrial time. Within a sustainable and cooperative SSP such as SSP1, this will be a relatively easy achievement, where for a non-sustainable and regionally conflicted scenario such as SSP3, mitigation will be relatively difficult (Riahi et al., 2017).

In order to model the socio-economic and environmental aspects of SSPs, Integrated Assessment Models (IAMs) are used. Most IAMs consist of a framework with driver components representing socio-economic factors and policies, energy and agricultural models, an earth system component and a model for the impacts of environmental change on human development. IAMs apply recursive techniques to generate a pathway with mitigation policies if an SSP is constrained with a radiative forcing target.

1.4 Climate change in Integrated Assessment Models

IAMs need information about climate change as a response from anthropogenic emissions, a process typically modelled with Earth System Models (ESMs) or their predecessors: Atmosphere Ocean coupled Global Circulation Models (AOGCMs). AOGCMs are climate models consisting of coupled atmospheric and oceanic models, and used to model the climate with high spatial and temporal resolution. Lately, AOGCM have been transformed into ESM by upgrading them with coupled representations for interactive chemistry, ocean ecology and biochemistry and plant ecology and land use. In contrast to AOGCMs, ESMs capture the physical effects of land cover change and the feedbacks of climate change on atmospheric composition.

Opposed to IAMs, ESMs can not be used to study interactions between human activities and the environment, as they are solely driven by prescribed emissions and land use. However, ESMs are used to study the climate response of changing emissions and land use, a fundamental process for IAMs. Running an ESM within IAMs is not an option regarding the complexity and computational costs of ESMs. To work around this problem, IAMs typically apply ESM emulators: reduced complexity climate models capable of calculating global average climate metrics based on emission inputs (Meinshausen et al. (2011),

Hartin et al. (2015)). In contrast to the global or regional approach of most IAMs, IMAGE (van Vuuren and Stehfest, 2017), the IAM developed and used by the Netherlands Environmental Assessment Agency, requires temperature change input at a $0.5° \times 0.5°$ resolution. Therefore, global average temperature change is downscaled with ESM generated temperature patterns in a process called geographic downscaling.

As ESMs need to be forced with scenario specific emissions, their temperature patterns are only available for the latest generation of scenarios. The continuous development of new scenarios with downscaling IAMs can therefore only rely on temperature patterns from older generations of scenarios. While the geographic downscaling approach solves the spatial resolution problem, the inconsistencies between new and old scenarios present a new one: As mentioned before, changing AQ and LC pathways will affect local temperatures. However, currently these processes are only captured within the ESM emulator, so that changes will only be accounted for on a global instead of a local scale. Especially if AQ and LC in the downscaling pattern scenario strongly deviates from the scenario studied with the IAM, local effects are likely not correctly accounted for.

1.5 Local air quality and land cover temperature effects in scenario studies

To quantify the effects of deviating air quality and land cover pathways on local climate change, special ESM experiments have recently been designed in projects such as the Coupled Model Intercomparison Project phase 6 (CMIP6) (Eyring et al., 2016). In addition to the regular SSPs and their mitigation pathways which are covered by the scenario Model Intercomparison Project (scenarioMIP) (O'Neill et al., 2016), the CMIP6 experiment portfolio also includes some SSPs with deviating AQ and LU trends. These kind of experiments generate more insight in the climate change impacts from specific AQ and LC effects within scenarios.

In order to quantify the climate impacts of Near-term Climate Forcers (NTCF; aerosol and ozone (precursors) + CH_4), the Aerosol Chemistry Model Intercomparison Project (AerchemMIP) (Collins et al., 2017) includes a scenario which is nearly identical to SSP3-baseline (or SSP370, SSP3 with 7.0 Wm^{-2} radiative forcing in 2100). However, in SSP3-lowNTCF, NTCF emissions decline after 2020 whereas in SSP3-baseline they continue to increase. By comparing the results of SSP3-baseline and SSP3-lowNTCF ESM runs, the local effects of deviating air quality regulating policies can be examined in a scenario context: Compared to SSP3-baseline, SSP3-lowNTCF shows significant global and regional air quality improvement as particulate matter and ozone levels drop. In turn, the smaller loading of cooling aerosol makes the global average temperature rise with an additional 0.23 K, while Asia warms up with an additional 0.5 K (Allen et al., 2020).

The Land Use Model Intercomparison Project (LUMIP) (Lawrence et al., 2016) includes scenarios which are forced by deviating land use compared to regular scenario. SSP126-SSP370LU, an SSP1-2.6 pathway with SSP3-7.0 land use patterns and SSP370-SSP126LU, an SSP370 Wm^{-2} pathway with SSP1-2.6 land use patterns, can be compared to their 'normal' scenario to study the effects of land use policy on local climate. By lack of literature on the comparison of results of these runs, another interesting LUMIP experiment is 'deforest globe'. Here, a large part of the world in 1850 (pre-industrial) is linearly deforested over a period of 20 years, followed by a 30 year stabilization period while emissions are kept at the 1850 level throughout the run. By comparing the results from deforest globe runs with those from a control experiment, the spatially varying biochemical effects become apparent. ESMs show a wide range of latitudinal averaged temperature differences when the deforestation runs is compared to the control, varying from -2.5 K to 0 K in the arctic, and 0 K to 1.5 K in the tropics (Boysen et al., 2020).

Including local or regional effects of AQ and LC changes can lead to considerable effects in IAMs. For example, in a sustainable scenario the regional temperature changes from WMGHGs range from 1.82K to 2.20K (assuming a climate sensitivity of 2.5K for doubling CO_2), while sulphate reduction results in a warming between 0.10K to 0.53K. These spatially varying temperature effects result in varying economic damage, especially in Asia. Here, WMGHGs account for a 3.6 billion (1990) USD welfare increase, while the strong sulphate aerosol abatement accounts for a 10.3 billion USD welfare decrease (Mendelsohn et al., 2001). The stark contrast only holds for Asia, while the global damages for WMGHGs and aerosols are 116.9 and 3.9 USD respectively, stressing the importance for regional or local climate modelling instead of a global approach. For land cover, it was demonstrated that including the physical aspects

of land cover change by accounting for the albedo effects in an IAM, significantly increases the need for emission reduction for climate change mitigation (Jones et al., 2015). This is primarily caused by radiative forcing from albedo decrease through a- or reforestation in the tropics. However, as the change in evapotranspiration is not taken into account, a- or reforestation is turned into a negative effect by the model, opposed to other studies which clearly present a positive effect on climate change mitigation.

Although the current body of literature underlines the importance of air quality and land cover on regional climate, state of the art IAMs do not cover the full effect from aerosol concentration and land cover changes, and are not in line with the latest insights from ESM studies. Correcting for AQ and LC effects would in turn impact several processes in IMAGE such as agricultural yield, energy demand for air condition and heating and the risks of diseases such as malaria. Moreover, a change in the associated climate impacts could affect future income equality (Rao et al., 2017). Hence, we expect that including these effects will lead to substantial changes in temperature related impacts and processes affecting human development. Therefore, including the AQ and LC effects might be relevant for policy assessment through IMAGE.

1.6 Relevance

In summary, ESMs, remote sensing and field studies show that air quality and land cover change effects can result in significant regional and local temperature changes. While the total or global average effect of air quality and land use is small compared to WMGHGs, effects of aerosol and land cover are far less homogeneous, stressing the relevance of proper high resolution representation of these processes in integrated assessment models.

Although current climate change models already indirectly captures these effects, local temperature change could still be represented incorrectly in IMAGE. This is especially the case if scenarios are studied where air quality and land cover pathways are highly inconsistent with those from the applied downscaling pattern. For example, if renewable energy shares grow vastly within a scenario compared to the downscaling scenario, both WMGHGs and aerosol precursor emissions will decrease and air quality will improve. Though effects on the global climate will be accounted for by the ESM emulator, local warming through aerosol reduction will not occur as the downscaling pattern does not change. Concerning land cover, the current setup only accounts for the effects on the global scale carbon cycle. If a scenario includes large scale afforestation or reforestation as a mitigation tool, this will affect global average temperatures, but local temperature changes might be far greater in reality.

To account for these dynamics, we seek to develop correction models which translate spatial and temporal changes in aerosol concentration and land cover between two scenarios into local temperature corrections.

1.7 Hypothesis

It is expected that correcting temperature patterns by including the effects of AQ and LC might be possible by developing and adapting AQ and LC models. With these models, the temperature pattern derived with the current downscaling technique can be adjusted to account for the changes in AQ and LC.

To evaluate the quality of the correction models, we aim to validate our approach against the results of ESMs, an element which is expected to be troublesome. Typically the kind of effects for which corrections are made are not studied on small scales. Moreover, the uncertainty of AQ and LC change on radiative forcing is high (compared to WMGHGs), even on a global scale (Myhre et al., 2013). On top of that, the new approach will not capture physical effects in ESMs such as feedbacks, changing circulation and teleconnections. Moreover, temperature effects are expected to be relatively small compared to background climate variability, which complicates matters even more.

Employing the correcting models for IMAGE, we expect a range of outputs. Relatively sustainable (mitigation) pathways will be subject to overall positive temperature corrections (above 0 K) from aerosol reduction and negative temperature corrections (below 0 K) from afforestation and reforestation. The opposite is expected for unsustainable pathways, where aerosol increase will yield negative temperature corrections and deforestation results in positive temperature corrections. Throughout all pathways, air

quality related corrections are expected to affect industrialized and high populated regions, while land cover corrections will be mainly limited to tropical and boreal forested areas.

Lastly, we expect that applying the temperature corrections in IMAGE will considerably affect land based processes and the carbon cycle compared to a non-corrected pathway. Moreover, larger corrections are expected to result in larger impacts in crop yield and plant carbon storage.

1.8 Aims

The overall aim of this project is to study the local effects of air quality and land cover on temperature change and how they can be fully represented in IMAGE. This includes updating the downscaling process by applying correction models for air quality and land cover related temperature effects. This study serves as a 'proof of concept', exploring the possibilities and significance of an updated downscaling process and the relevance for integrated assessment. The corresponding research questions are as follows:

- Is it possible to model the local temperature effects of AQ en LC through simple models?
- How do the AQ and LC temperature models compare to likewise driven ESMs?
- What are the temperature corrections according to these models for a range of scenarios?
- What is the effect on land-based processes for temperature corrected scenarios compared to non-corrected scenarios?

Chapter 2 discusses the applied methodology where the correction models for air quality and land cover effects and their validation experiment set up will be discussed. In Chapter 3, the results of the validation experiments and the results for corrections in IMAGE scenarios is shown. In chapter 4, the results of the validation and the corrections will be discussed. These are followed by a conclusion in chapter 5, and is closed with a number of recommendations in chapter 6.

2 Methods

This chapter starts by shortly discussing IMAGE and its current downscaling procedure. The following two sections focus on answering the first research question by explaining the new correction models for air quality and land cover effects. In both sections, the set up of the validation experiments which are used to answer the second research question are discussed as well. In the fourth section, the employment of the updated downscaling procedure and the applied input data is discussed to answer the third and fourth research question.

2.1 Geographic downscaling in IMAGE

In order to address the research questions, the Integrated Model to Assess the Global Environment (IMAGE) of the Netherlands Environmental Assessment Agency (Planbureau voor de Leefomgeving, PBL) is used. IMAGE consists of a framework of coupled submodels for socio-economic and physical processes, such as economy, agriculture, energy and earth system. An overview of the IMAGE model framework model is shown in appendix A. To model the effects of socio-economic assumptions and policies on climate with a high spatial resolution, the Atmospheric Composition and Climate component uses a two step approach, involving an ESM emulator combined with ESM temperature patterns.

First, MAGICC ('Model for the Assessment of Greenhouse gas Induced Climate Change') is used to calculate global mean temperature change. MAGICC is a reduced complexity climate model that emulates ESMs based on key physical and biological processes (Meinshausen et al., 2011). Providing emissions for well-mixed greenhouse gases, halogenated gases and ozone and aerosol precursors, MAGICC is able to quickly calculate global climate metrics.

Second, the output of MAGICC is post-processed with a geographic pattern scaling technique, i.e. downscaling. Currently, ESM patterns from RCP scenarios (van Vuuren et al., 2011) are applied, which are subsequently normalized according to output from MAGICC. The geographic downscaling process is shown in figure 1 and works as follows: For example, the global temperature change according to MAGICC is 1 K, the global average temperature change according to the ESM is 2 K and a local temperature change in a grid cell according to the ESM is 3 K. Temperature change in all cells is scaled so that the global average change in temperature matches with MAGICC. Therefore a normalization factor of $\frac{1}{2}$ will yield a local change in the grid cell of $\frac{1}{2} \cdot 3 = 1.5$ K.

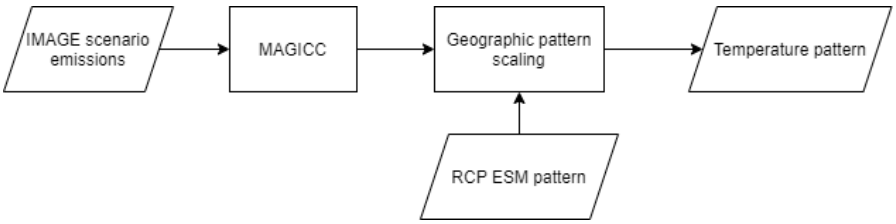

Figure 1: Flowchart of the IMAGE geographic downscaling process for modelling temperature change.

2.2 Temperature corrections for air quality change effects

2.2.1 Method

The newly developed downscaling procedure includes a correction for forcing by aerosol component concentrations. Aerosol component concentration fields are modelled with TM5-FASST, a reduced form source-receptor model of the TM5 full chemical transport model (Krol et al., 2005). FASST allows for fast modelling of future local concentrations, by modelling without major loss of accuracy compared to the full TM5 model (Van Dingenen et al., 2018). FASST applies 56 regional emission-concentration relations (derived from TM5 perturbation runs) and subsequent downscaling, to model gridded concentration fields based on regional emission inputs. This enables fast modelling of future concentrations of aerosol components for a range of future emission scenarios.

In order to model temperature corrections for aerosol precursor emission differences between two scenarios, a set of pre-industrial climate simulations from the CMIP6 portfolio is used as input data.

These 'PiClim' runs span a 30 year period where 1850 concentrations of WMGHGs and emissions of NTCFs are applied. Monthly varying sea surface temperatures (SST) and sea ice extent from a control experiment are prescribed as well, see Collins et al. (2017) for more detail. Amongst others, the PiClim experiments include runs where separate aerosol component emissions are perturbed to present day (2014) levels. The PiClim runs with individual emission perturbations in black carbon (BC), organic carbon (OC) and sulphate dioxide (SO_2) are selected to derive linear relations between their temperature and concentration response in BC, particulate organic matter (POM) and sulphate (SO_4). BC, POM and SO_4 account for the highest radiative aerosol forcing in present-day climate (Myhre et al., 2013). They are also expected to do so in the future, and are thus considered most relevant for making corrections in future scenarios.

The experiments allow to directly analyse and estimate their individual contributions to temperature change, whilst completely capturing direct and indirect aerosol effects. This is somewhat alike the method used in Mendelsohn et al. (2001). However, Mendelsohn et al. (2001) scales based on individual regional SO_2 emissions perturbation runs from a single AOGCM. However, in the new approach, temperature is scaled with gridded concentration from global emission perturbation experiments of 6 ESMs for 3 mayor radiatively relevant aerosol components. In contrast to Mendelsohn et al. (2001), the new method shown here only relies on local effects as will be discussed in more detail.

Figure 2 shows the flowchart of the new integrated air quality temperature correction model and works as follows: A perturbation in emissions of one precursor gas in single PiClim run results in a loading change in aerosol component x (calculated with FASST) $\Delta C_{x,piClim}(i,j)$, and a temperature change (from ESMs) $\Delta T_{x,piClim}(i,j)$ in grid cell i, j with respect to the control run. To calculate a correction temperature pattern for time step t, emissions from the applied RCP pattern and emissions from the scenario that is studied in the IMAGE run are taken, aggregated to FASST regions and run with FASST to model the concentration fields. Next, we determine the difference in concentration for aerosol component x between two scenarios for time step t, $\Delta C_{x,scenarios}(t,i,j)$ and scale it with the PiClim temperature response divided by the PiClim concentration difference, to calculate the expected temperature difference on grid level:

$$\Delta T_{x,scenarios}(t,i,j) = \Delta T_{x,piClim}(i,j)\frac{\Delta C_{x,scenarios}(t,i,j)}{\Delta C_{x,piClim}(i,j)} \tag{1}$$

For the scaling we assume a linear relation between component concentration and temperature response, and that corrections for individual components are additive. As the individual emission perturbations are small, the radiative response is expected to be linear (Hansen, 2005) though this might not hold for temperature responses. In the end, all temperature corrections for the individual components are summed to obtain the total correction. As long as the global average of the total correction is small, the signal can simply be added to the temperature change mapping, as overestimation of the global effect on temperature from aerosols through MAGICC and the correction is then prevented.

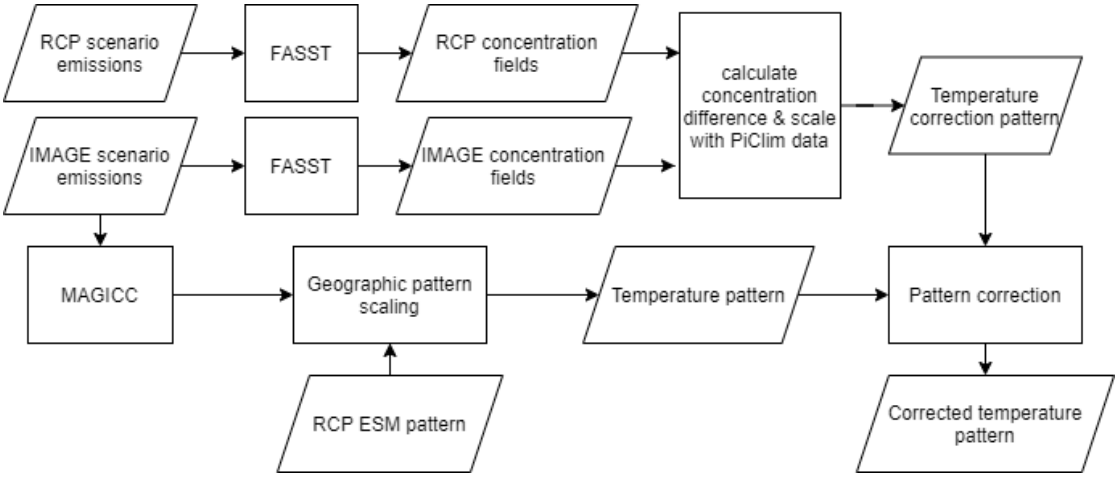

Figure 2: Air quality correction process flowchart

2.2.2 Correction model input data

To derive the input data from the PiClim experiments, monthly temperature at surface (2 meter above ground level) output is taken from seven ESMs: CNRM-ESM2-1 (Séférian et al., 2019), GISS-E2-1-G (Schmidt et al., 2014), IPSL-CM6A-LR-INCA (Boucher et al. (2020) & Hauglustaine et al. (2014), MIROC6 (Tatebe et al., 2019), MRI-ESM2-0 (Yukimoto et al., 2019), NorESM2-LM (Seland et al., 2020) and UKESM1-0-LL (Senior et al., 2019). All ESMs provide one realization (i.e. ensemble member) for the BC, OC and SO_2 perturbation runs and the corresponding control realization is selected. In order to provide a temperature set for scaling individual aerosol component effects, we pursue to derive a set of multi model means (MMM) from the ESMs mentioned above. This will prevent the temperature set to be biased towards one specific model, while enabling the derivation of a causal relationship between temperature and concentration change in the absence of large ensembles for these experiments. The temperature data is averaged over the 30 year period and interpolated to a $1° \times 1°$ grid. Next, results from the control run are subtracted from the perturbed runs to extract the temperature responses from the emission perturbations. The results of the individual ESMs for the BC, OC and SO_2 perturbations are shown in appendix B.1, along with a run in which emissions of all aerosol precursors are perturbed simultaneously (aer). When reviewing the responses for each ESM, it is observed that MIROC6 produces highly inconsistent temperature responses compared to all other ESMs (most likely caused by a faulty control run), and is left out for the calculation of the MMM temperature response.

Next to persistent patterns in some areas, the remaining six ESMs also show a small range of responses in magnitude and sign, typically in areas where the perturbation in emissions and resulting change in concentrations is small. These features can be explained by internal variability, teleconnections and differences in model parameterizations. In order to ensure a robust MMM temperature response, the gridded temperature change of the six ESMs is taken and the number of ESMs that agree on the sign is determined, the result is shown in appendix B.2. The MMMs of the temperature responses are shown in figure 3, hatched areas indicate a model agreement lower than five out of six. Here, it is observed that BC perturbations result in consistent warming over parts of North America, East Europe and the Tibetan Plateau, but cooling over central India. Both OC and SO_2 perturbations lead to overall cooling, where OC effects are only consistent over central South America and India, while SO_2 effects are consistent over Asia and the majority of North America as well.

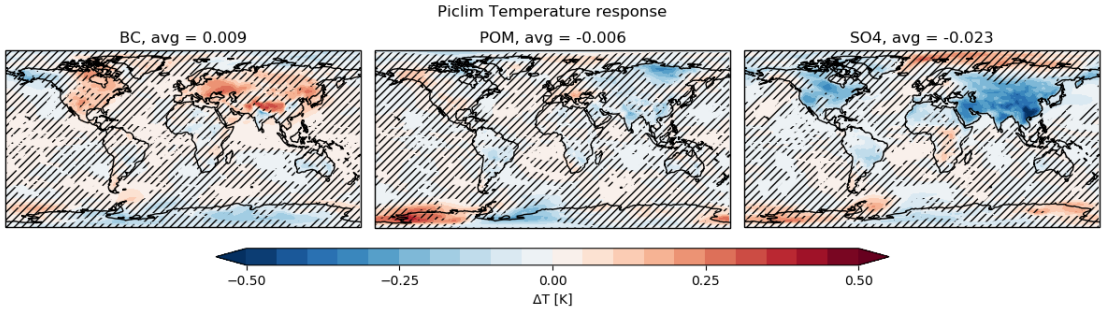

Figure 3: MMM Temperature response for the PiClim BC, OC and SO_2 emission perturbations with respect to the PiClim control simulation. Hatched areas indicate model ESM sign agreement lower than five out six.

Next, the spatial emission data from input4MIPs (https://esgf-node.llnl.gov/search/input4mips/) is used to construct the regional emission input for FASST. The control run input data is constructed by summing the emission sectors, averaging over the 12 months and aggregation from gridded to regional data for the 1850 emissions of BC, CO, NH_3, NMVOC, NO_x, OC and SO_2. FASST also requires the CH_4, CO_2 and N_2O emissions, which are not provided by input4MIPS, and are therefore passed with fill values. The emission perturbation input data is similar to the control run, except that 2014 emission levels are used for the individual BC, OC and SO_2 perturbation runs.

In order to observe the effect of the emission perturbation, the concentration data for BC, POM and SO_4 in the respective BC, OC and SO_2 perturbation runs is compared to the control run. Next, areas suitable for deriving the linear temperature-concentration response are selected, based on two criteria:

- The MMM temperature response (figure 3) has a sign agreement of at least five out of six.
- The concentration response (perturbation-control) should be significant, for which an arbitrary level of 1 $\mu g m^{-3}$ is set.

Handling these criteria ensures that corrections are based on robust MMM temperature responses, and are not hypersensitive in areas with very low concentration responses. This approach delivers corrections which only include local effects, while temperature responses in areas with negligible concentration changes (caused by teleconnections, circulation changes, etc.) are effectively filtered out.

Figure 4 shows the concentration response modelled with FASST, the areas suitable for making corrections (based on the aforementioned criteria) are marked with green contours. As can be observed, air quality-based temperature corrections will be primarily limited to South-West, South and East Asia. The patterns and magnitudes of the concentration responses correspond to those found in CNRM-ESM2-1 and MRI-ESM2-0 (the only two ESMs which provide concentration field output for these runs), confirming a correct application of FASST.

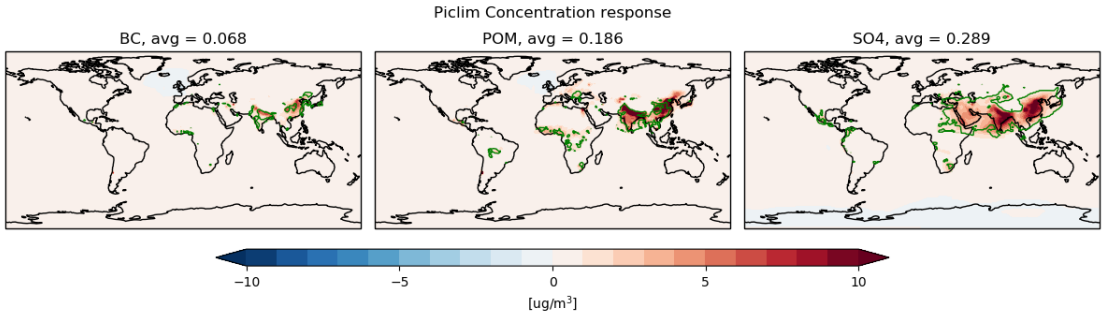

Figure 4: Concentration responses from respective emission perturbations in BC, OC and SO_2, modelled with FASST in $\mu g m^{-3}$. Green contours indicate areas with robust MMM temperature response and significant concentration response.

2.2.3 Validation

In order to validate the air quality temperature correction method, two near identical scenarios are compared, where one run features deviating emissions of aerosol (precursor) emissions. For this purpose, the SSP3-baseline SSP3-lowNTCF, SSP3SST-baseline (here SST indicates fixed sea surface temperatures and sea ice extent) and SSP3SST-lowNTCF are taken. It is expected that air quality focused mitigation such as a lowNTCF pathway will lead to additional warming due to reduced aerosol loading reduction in certain regions, compared to a scenario without air quality measures (Allen et al., 2020). To demonstrate this, temperature output from SSP3 baseline and lowNTCF ESM runs is compared, the number of realizations (i.e. ensemble members) for every scenario is shown in table 1.

Next, temperature correction is modelled with the process described in the section above, as if the baseline scenario is used for the downscaling map and lowNTCF as the actual IMAGE scenario. The correction model emission input timelines (from input4MIPS) are shown in figure 5 for regions with the highest differences between SSP3 baseline and lowNTCF. These are also the regions that are thus expected to show the largest temperature differences between the SSP3 baseline and lowNTCF scenario.

Table 1: Selected ESM data for the air quality correction method validation for SSP3 modelling exercises. Numbers indicate amount of realizations with monthly average temperature at surface output.

	SSP3		SSP3SST	
	baseline	lowNTCF	baseline	lowNTCF
BCC-ESM1(Wu et al., 2020)	3	3	1	1
CESM2-WACCM (Danabasoglu et al., 2020)	3	3	1	1
CNRM-ESM2-1 (Séférian et al., 2019)	5	3	1	1
GFDL-ESM4 (Dunne et al., 2019)	1	1	1	1
MRI-ESM2-0 (Yukimoto et al., 2019)	5	1	1	1

To demonstrate the NTCF mitigation effects in ESMs for both normal and SST runs, averages are taken for multiple realizations and interpolated to a $1° \times 1°$ grid. The data is averaged over 10 year periods, a time scale that is short enough to show a fast reduced aerosol loading effect that is expected, while sufficiently long to filter the effects from climate variability. The SSP3-lowNTCF runs span the 2015 to 2055 period, which allows to analyse four separate 10 year averaged data sets. As the correction will be compared to averaged ESM results, the corrections are also calculated based on 10 year averaged emissions in the respective time frame. Results will be presented in section 3.1.1.

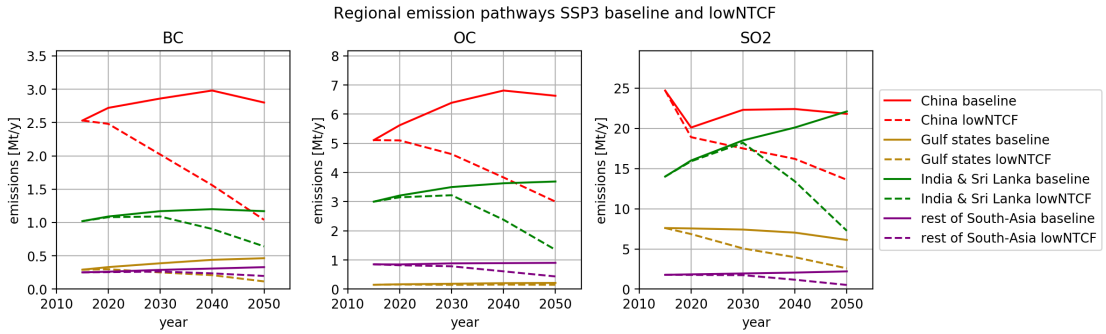

Figure 5: SSP3 baseline and lowNTCF emission timelines for the four FASST regions with the highest differences between SSP3 lowNTCF and baseline

2.3 Temperature corrections for land cover change effects

2.3.1 Method and input data

To model the effects of land cover change, a correction procedure is developed that emulates temperature effects for various land cover changes as result of land use change. The model uses data from Duveiller et al. (2020), which describes the temperature at surface (2m above ground level) change for three aggregated Tier 1 transitions; forest to grassland, forest to cropland and grassland to cropland. The data is derived with remote sensing techniques and relies on comparing temperatures over neighbouring areas with similar environmental conditions but different LC, see Duveiller et al. (2018b) for more details.

The applied data is shown in figure 6, which shows overall warming in the tropics and cooling in the boreal zone from deforestation (figures a and b), which agrees with the theoretical framework. As the grass to crop transition transition is relatively more affected by land use aspects such as irrigation, the pattern is more spatially (and especially longitudinally) varying.

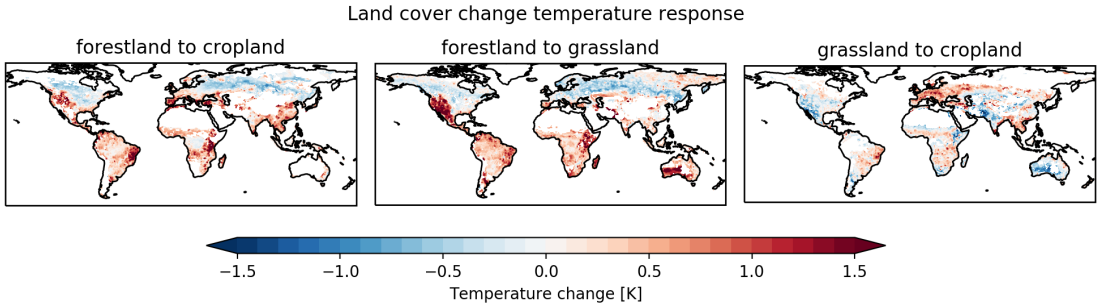

Figure 6: Temperature response for various vegetation transitions, data source: Duveiller et al. (2020).

The correction procedure is shown schematically in figure 7 and starts by taking the LC from the RCP scenario and the IMAGE scenario. The first step involves calculating the changes in area per land cover type relative to the year 2015 for both scenarios. Next, the difference in these changes (IMAGE - RCP) is taken to determine the difference in vegetation fraction changes. These differences are used to determine the fraction size of the required vegetation transitions, which are needed if the land cover from one scenario has to be converted to the other. Next, these transitions are multiplied with the respective

data from Duveiller et al. (2020) to calculate the correction temperatures, i.e. the associated temperature changes from converting the land cover from an RCP to an IMAGE scenario. Finally, the corrections for the individual transitions are summed to calculate the total temperature correction pattern, which can then be added to the temperature pattern in IMAGE.

A practical example: In a single cell, for a given year, the loss in forest fraction in the RCP and IMAGE scenario are 0.3, and 0.5, gain in grass fraction are 0.1 and 0.2 and gain in crop fraction are 0.2 and 0.3. The transition fractions in RCP are thus 0.1 forest to grass and 0.2 forest to crop, for IMAGE the transition fractions are 0.2 forest to grass and 0.3 forest to crop. The difference in transition fractions for IMAGE with respect to RCP are thus 0.1 forest to grass and 0.2 forest to crop. The temperature changes for the respective transitions are 0.3K and 0.7K, so the total local temperature correction will be $0.1 \cdot 0.3 + 0.2 \cdot 0.7 = 0.17$K.

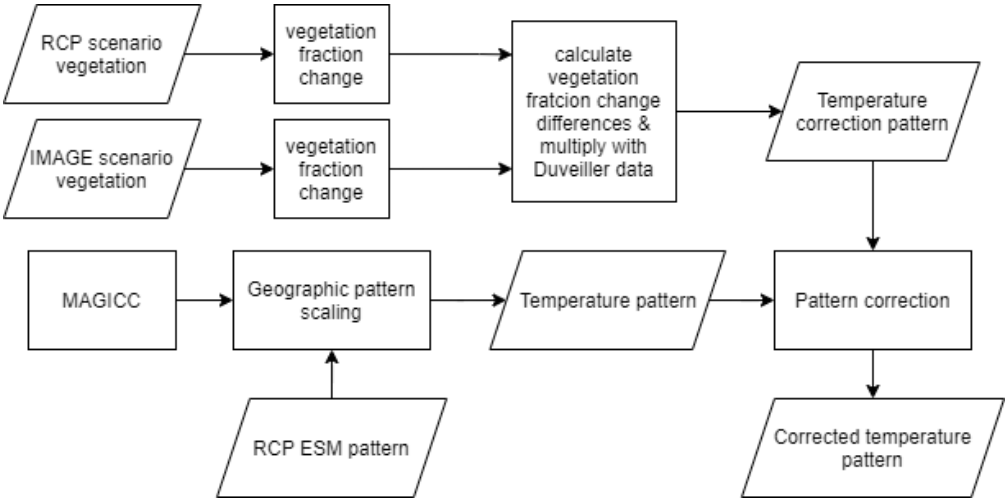

Figure 7: Flowchart of the LU correction procedure.

Next to the land cover types mentioned above, some ESMs also provide shrub land and pasture fractions. These are subsequently aggregated to the grass fractions, following the Tier 2 to Tier 1 aggregation from table 2 in Duveiller et al. (2020). Although bare soil cover undergoes significant changes throughout scenarios, (Duveiller et al., 2020) does not provide temperature response for transitions from or to this land cover type. Ignoring these changes is not an option, as transition fractions to the other cover types might be over or underestimated by the current setup. Consulting Poulter et al. (2015), grass land holds the highest bare soil share, and subsequently bare soil fractions are added to the grass land fraction.

2.3.2 Validation

To validate the land cover corrections, a set of SSPs is used including two normal and two with deviating land use. The runs that will be compared are 1) SSP126-SSP370LU (SSP1-2.6 scenario with SSP3-7.0 Land Use) with respect to SSP1-2.6 and 2) SSP370-SSP126LU (SSP3-7.0 (baseline) scenario with SSP1-2.6 Land Use) with respect to SSP3-7.0. Both land use deviating SSPs are concentration driven (concentrations from their respective scenarios), so temperature differences between a regular and LU deviating scenario are purely induced by land cover change effects. The number of available ESM realization for these runs is shown in table 2.

ESM temperature difference patterns will be constructed by taking 30 year averages and interpolation to a $1° \times 1°$ grid. The runs span the 2015 to 2100 time frame, so six 30-year averages (2015-2045, 2025-2055 etc.) are calculated for comparison. Ideally, the mean of multiple realization would be taken to effectively filter out the land cover change temperature effect from climate variability. Although numerous realizations are available for this process for the regular SSPs, climate variability will be present at a higher level in the deviating LU runs as most ESM only provide a single ensemble member. Therefore, no model means are taken for multiple realizations.

Table 2: Overview of available model data for LU correction method validation (based on ESGF) for LUMIP modelling exercises, numbers denote amount of realizations with monthly average temperature at surface output.

	SSP3		SSP1	
Source ID	7.0	7.0-SSP126LU	2.6	2.6-SSP370LU
CNRM-ESM2-1 (Séférian et al., 2019)	5	1	5	1
CanESM5 (Swart et al., 2019)	50	1	50	1
GFDL-ESM4 (Dunne et al., 2019)	1	1	1	1
IPSL-CM6A-LR (Boucher et al., 2020)	11	1	6	1
MPI_ESM1-2-LR (Mauritsen et al., 2019)	10	1	10	1
UKESM1-0_LL (Senior et al., 2019)	5	4	5	4

For modelling corrections, these scenarios will be substituted as if the regular scenario is used as the RCP downscaling map and the deviating land use scenario is the scenario that is studied with IMAGE. The modelled corrections will be compared against the difference in ESM output for the two scenarios. When comparing the ESM outputs, we expect that deforestation leads to overall warming in the tropics and cooling in the boreal zones.

Land cover type fractions from individual ESMs for SSP1-2.6 and SSP3-7.0 are taken as input data for making corrections. The global land cover change timelines for SSP126 and SSP370 are shown in figure 8. The features in figure 8 can be expected from the scenario storylines: SSP1-2.6 has lower world population, improved technology, plant-based diet shifts and afforestation efforts compared SSP3-baseline, which lead to higher crop land extent in SSP370 than in SSP126. Whereas the crop cover change differences are roughly consistent between scenarios in all ESMs, all other land cover changes are highly inconsistent throughout the ESMs. Therefore, the validation procedure is performed for each individual ESM, without taking MMMs. Results of the validation will be shown in section 3.1.2.

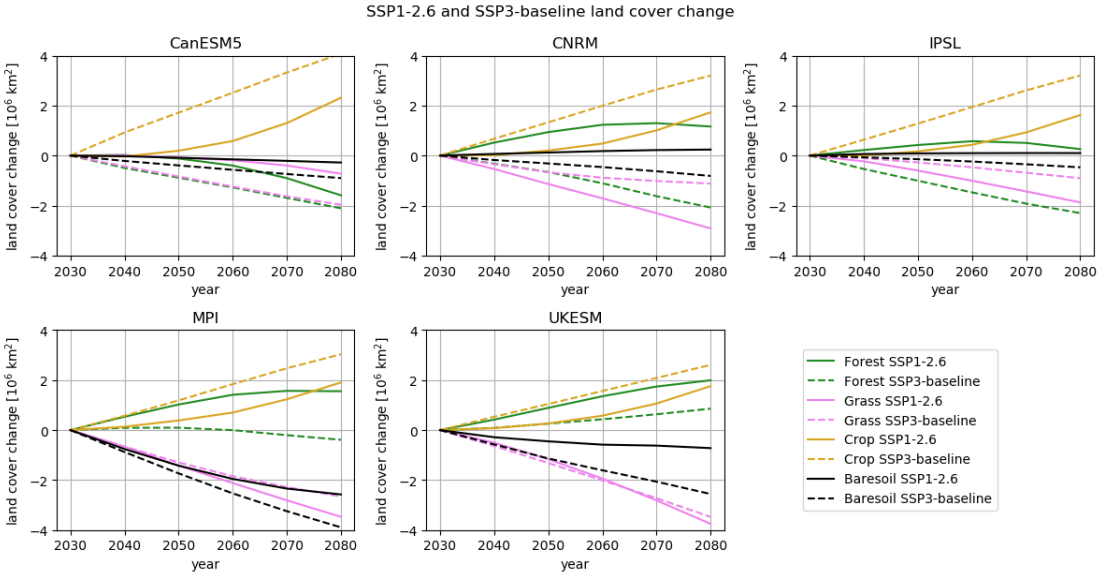

Figure 8: Global vegetation cover changes in SSP1-2.6 and SSP3-baseline

2.4 Implementing temperature corrections in IMAGE

To study the effects of employing the correction models for air quality and land use, the latest IMAGE model version is used. As the new SSP set from this model version is still under development, the range of scenarios for which corrections can be calculated is limited. To examine the corrections and their effects for integrated assessment on a large range of AQ en LC pathways, the SSP2-2.6, SSP2 baseline and SSP3 baseline scenarios are used.

Ideally, the new AQ and LC correction models would be integrated in IMAGE to capture the effect of feedbacks (e.g. additional need for mitigation in case of major AQ improvements, or afforestation as a more efficient mitigation tool). However, due to lack of time, the correction models are not integrated but forced with emissions and land cover patters from IMAGE runs, to calculate the corrections based on the applied downscaling maps and the range of studied scenarios. The corrections are calculated with a 10 year and $1° \times 1°$ resolution, and post processed by linearly interpolating to 5 year and $0.5° \times 0.5°$ resolution to comply with IMAGE. Subsequently, IMAGE runs are performed where the correction patterns are added to the downscaled patterns in the IMAGE land model. This solution is expected to capture the first order effects of the corrections for land based processes. To assess these effects, the relative difference in yields of wheat, maize and rice and biomass is calculated to indicate magnitude of agricultural pressure and effects on the carbon cycle. To put the results into perspective, effects of runs with varying RCP6.0 ESM downscaling patterns within SSP2 are also calculated.

Within IMAGE, temperature maps for RCP2.6 and RCP6.0 from GFDL-ESM2M (Dunne et al., 2012), IPSL-CM5A-LR (Dufresne et al., 2013), hadGEM2-ES (Jones et al., 2011) and MIROC5 (Watanabe et al., 2010) are available for downscaling. However, the latest downscaling configuration uses IPSL-CM5A-LR RCP6.0 maps as default, regardless of the SSP scenario. The corrections for the SSPs will thus be made relatively to RCP6.0 emissions and IPSL-CM5A-LR RCP6.0 land cover data. Although the SSPs scenarios start in 2015, corrections are made from 2020 onwards, as 2015 has no emission or land cover differences due to harmonization.

The global timelines of the emissions and land cover changes are shown in figures 9 and 10. Here the RCP2.6 and RCP6.0 emission data is taken from input4MIPS and the land cover data from the IPSL-CM5A-LR model. The SSP data is derived from IMAGE output. The original SSP emissions of BC and OC and SO_2 were modelled incorrectly through a defect in the IMAGE energy model. By lack of time, the SSP emission data is harmonized to the 2015 RCP6.0 levels to solve this problem. For more details and the original data see appendix E.

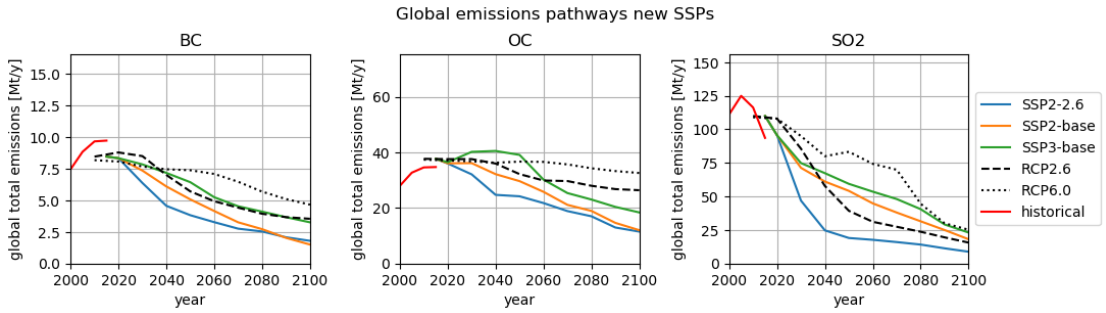

Figure 9: Global harmonized BC, OC and SO_2 emission timelines for selected SSPs, RCP2.6 & RCP6.0 and historical (van Marle et al., 2017 & Hoesly et al., 2018) in Mt/y.

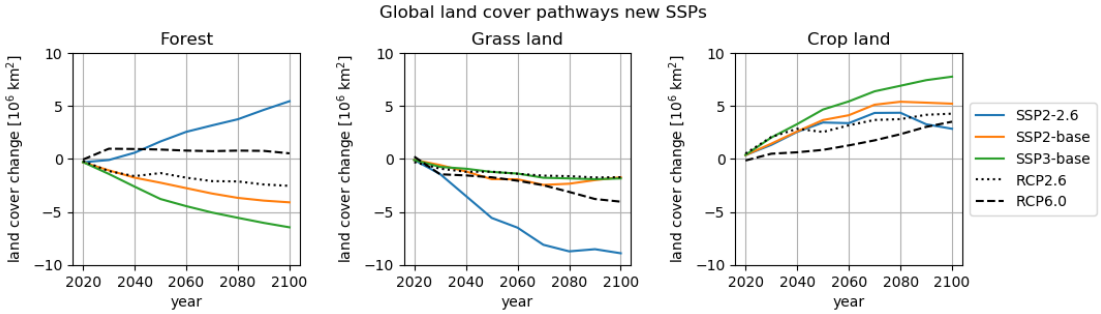

Figure 10: Global land cover change timelines for selected SSPs and IPSL-CM5A-LR RCP2.6 & RCP6.0, relative to 2015 in million km^2.

Even though SSP2-2.6 and RCP2.6 mitigate towards the same global radiative forcing, overall air quality improvements are higher in SSP2-2.6. Furthermore, all SSPs show lower emission levels than RCP6.0. This can be expected for SSP2, though surprisingly SSP3-base is also marked by overall lower aerosol precursor emissions, while the radiative forcing in 2100 is actually higher. The SSP emissions diverge from RCP6.0, and differences are most pronounced in the midst of the 21th century (especially SO_2), after which they converge again. Overall, SSP2-2.6 shows the largest differences throughout time, and is thus expected to render the highest corrections, where SSP2-baseline SSP3-baseline show smaller differences and thus likely smaller corrections. Overall, the temperature corrections are expected to be positive (above 0 K) because of the stronger aerosol precursor emission reductions. From the land cover timelines it becomes clear that SSP2-2.6 is marked by stronger afforestation and larger crop land shares at the expense of grass land, compared to RCP6.0. For SSP2-baseline and SSP3-baseline, grass land fractions barely deviate from RCP6.0, though larger crop land shares are generated through stronger deforestation. Overall, it can be expected that SSP2-2.6 will be marked by negative temperature corrections (below 0 K), while the corrections for SSP2-baseline and SSP3-baseline are expected to be positive (above 1 K).

3 Results

This section starts by showing the results of corrections for AQ and LC for the validation experiment, and how they relate to the ESM results. In the second section, the results of the corrections for new SSPs in IMAGE are shown. The third section covers the results on land-based processes from IMAGE runs with corrected patterns.

3.1 Validation

3.1.1 Air quality

Figure 11 shows the difference in concentration between the SSP3 baseline and lowNTCF (modelled with FASST), as well as the associated corrections for the year 2050 (2045-2055 average). Results for other timesteps can be found in appendix C.1. From figure 11 we observe that the lowNTCF pathway results in lower BC, POM and SO_4 concentrations, especially in South-West, South and East Asia, with respect to the baseline scenario. These result in overall warming, where BC and POM reductions show a warming up to 0.25 K in the majority of India and small parts of south-east China. The corrections for SO_4 extent over a larger area, with overall warming of roughly 0.25 K and stronger corrections for India up to 0.5 K. Overall, SO_4 differences account for the majority of the total correction, which is also the case for other time steps.

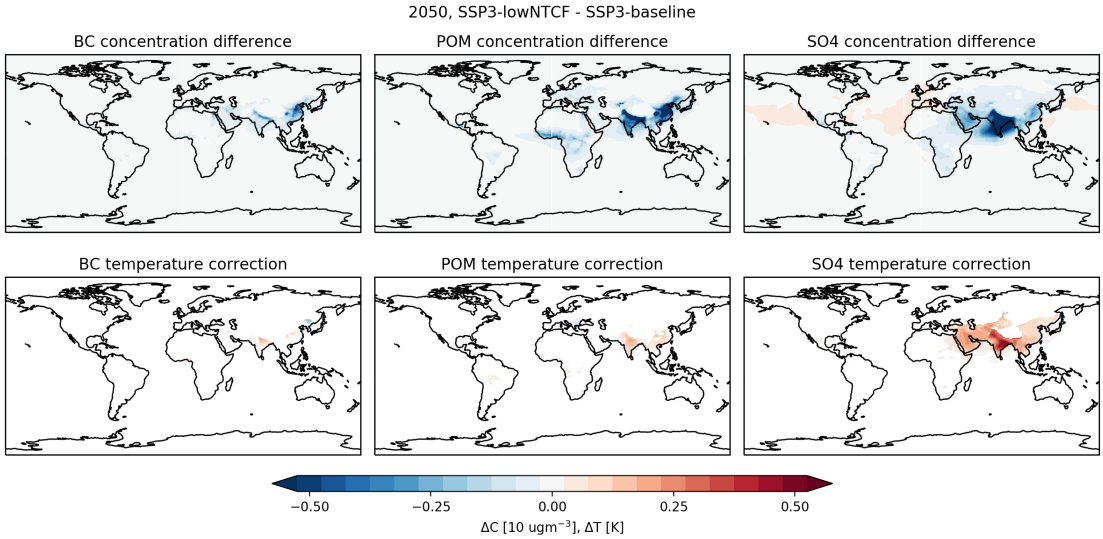

Figure 11: SSP3 baseline and lowNTCF concentration differences modelled with FASST and according temperature corrections derived with PiClim scaling for the 2045-2055 average.

In figure 12 the MMM ESM temperature difference between baseline and lowNTCF and the total temperature correction are compared for the 2045-2055 average. Results for other timesteps can be found in appendix C.2. Here, the MMM temperature difference shows warming over South-West, South and East Asia, which corresponds to the expected effect of air quality focused mitigation. However, the MMM difference also shows a strong arctic warming, which could be caused by the high arctic sensitivity to aerosol reductions (Westervelt et al., 2020). Moreover, the ESMs also show robust MMM warming over central Africa, while POM concentration changes are relatively small and corrections can not be made as a lack of ESM temperature robustness in these areas.

Patterns are similar for the other timesteps, albeit lower in magnitude. Overall, temperature corrections are only noticeable within the -5 μgm^{-3} cumulative concentration interval above land, where they roughly resemble the MMM. To accommodate a more objective comparison, scatter plots of the robust MMM temperature difference and temperature correction are shown in figure 13. Here, spatial data has been selected based on the cumulative concentration intervals shown figure 12A. Scatter plots for individual ESMs are shown in appendix C.2, figure C.7. According to the best linear fits in figure 13, the corrections

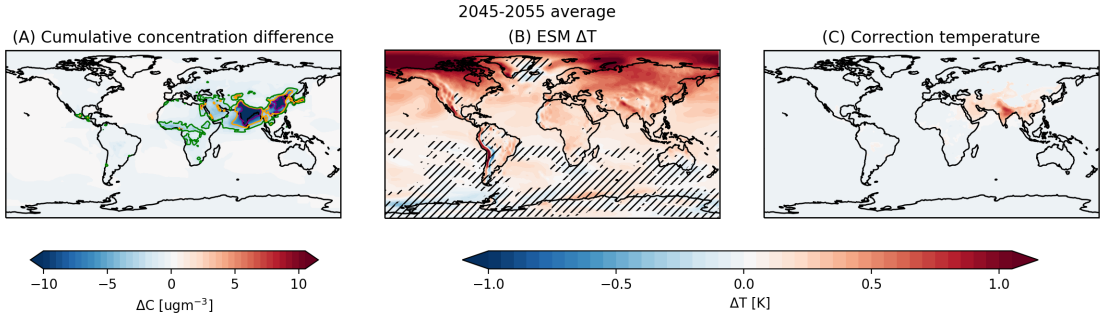

Figure 12: (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 µgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate ESM temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

tend to be to small throughout all timesteps, this is also true for all of the individual ESMs (appendix C.2). The slope of the best fit tends more towards a 1:1 relation for higher cumulative concentration intervals. In other words, regions with larger aerosol loading i.e. AQ differences are prone to corrections which show more resemblance with ESM results. However, correlations coefficients are small for individual ESMs as well as the MMMs. In order to exclude interference from arctic amplification and other indirect climate effects, the corrections are also compared to the results of the 2045-2055 SSP3 SST ESM runs in figure 14. Comparisons for other timesteps are shown in appendix C.3.

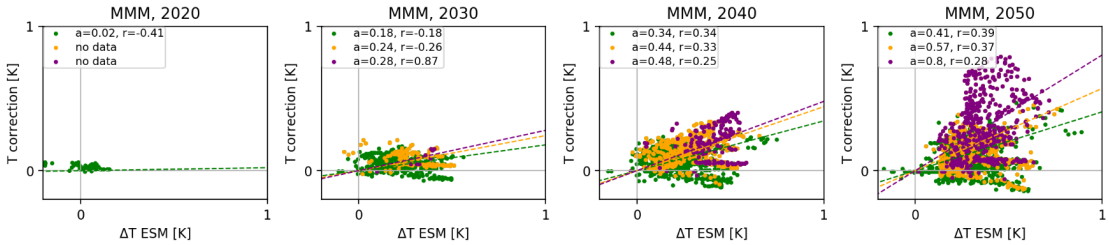

Figure 13: Scatter plots for MMM SSP3 lowNTCF-baseline temperature difference vs temperature correction on the -2, -5 and -10 µgm^{-3} (green, orange and purple) cumulative concentration change intervals. Dashed lines resemble best fit with intercept at 0 through the three cumulative concentration interval selected data sets, where a indicates the slope and r the correlation coefficient.

Again, the MMM temperature difference between SSP3-baseline and SSP3-lowNTCF with fixed SST shows warming over South-West and South Asia, while warming in East Asia is substantially smaller. However, the MMM difference still shows robust warming in northern hemisphere above land, though arctic warming is suppressed through SST and sea ice extent fixation. The concentration differences in these regions are generally negligible, so again the warming might be caused by teleconnections, or the BC on snow effect (Bond et al., 2013) might play a role.

Similar to the regular SSP3, scatter plots of the MMM SST temperature difference and temperature correction are shown in 15. Scatter plots for the individual ESMs are shown in appendix C.3 figure C.11. For the SST data, the MMM scatter plots show a much stronger correlation and better slopes than the regular SSP3 comparison. Again, correlations and slopes are higher for higher cumulative concentration intervals, which is also the case for all individual ESMs.

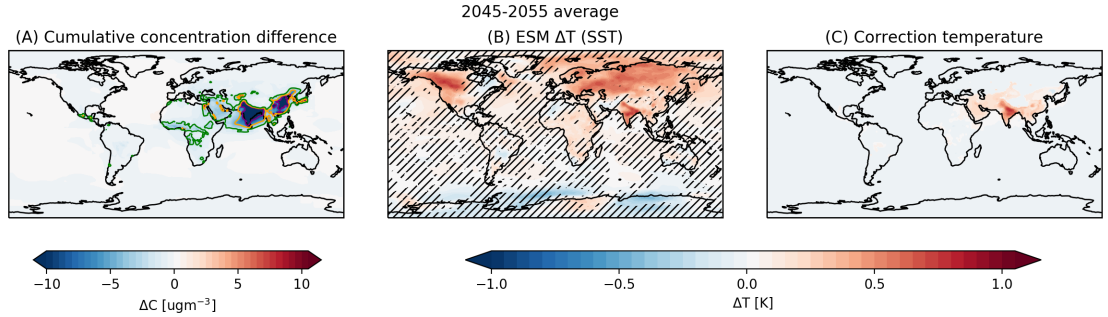

Figure 14: 2045-2055 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 µgm^{-3} intervals; (B) MMM temperature difference between SST baseline and SST lowNTCF, hatched areas indicate ESM temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

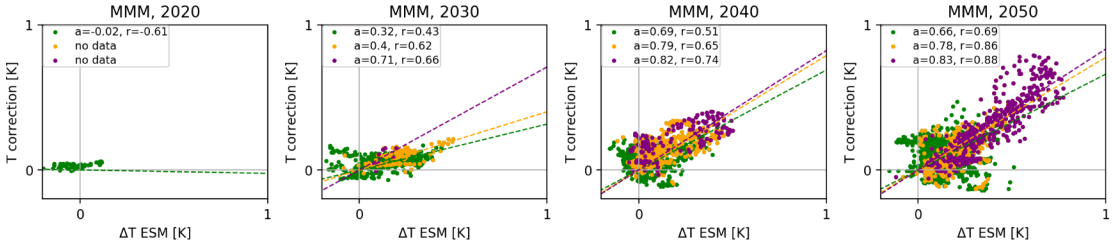

Figure 15: Scatter plots for MMM SSP3 SST lowNTCF-baseline temperature difference vs temperature correction on the -2, -5 and -10 µgm^{-3} (green, orange and purple) cumulative concentration change intervals. Dashed lines resemble best fit with intercept at 0 through the three interval selected data sets, where a indicates the slope and r the correlation coefficient.

3.1.2 Land cover

For illustrative purposes, only results for SSP126-SSP370LU with respect to SSP1-2.6 for the 2065-2095 averaged period will be shown. This time frame features the largest differences in vegetation extent, and thus the largest effects on temperature are expected. Results for SSP370-SSP126LU with respect to SSP3-7.0 are shown in appendix D.3 and roughly similar, though naturally opposite in sign. Following the method described in section 2.2.3, the differences in vegetation cover fractions in SSP2-2.6 and SSP3-7.0 are calculated and shown in figure 16 for UKESM1-0_LL. The associated transitions between SSP1-2.6 and SSP3-7.0 are shown in figure 17. Maps for other ESMs are roughly similar and shown in appendix D.1.

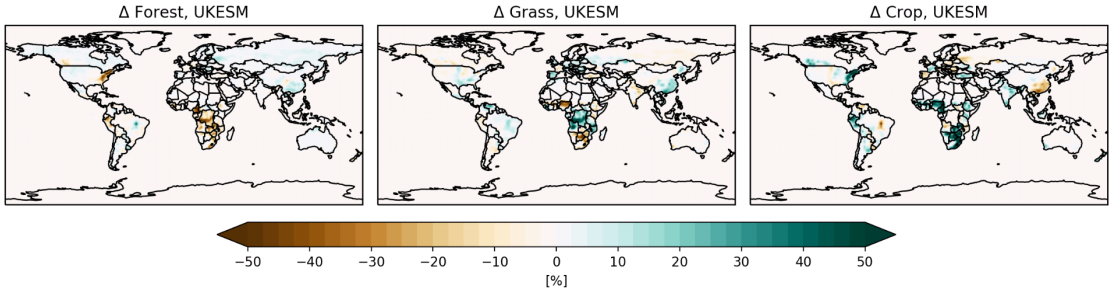

Figure 16: 2065-2095 average SSP3-7.0 - SSP1-2.6 vegetation cover fraction differences in UKESM1-0_LL.

Overall, SSP3-7.0 is marked by a stronger deforestation, where forests are primarily replaced by crops. Forests are also replaced by grass, especially in and around the Democratic Republic of the Congo. Lastly, grass lands are replaced by crops as well, especially in the African savanna. In contrast to this overall trend, south-east China is marked by stronger afforestation and grass land increase at the cost of crops in SSP3-7.0 compared to SSP1-2.6.

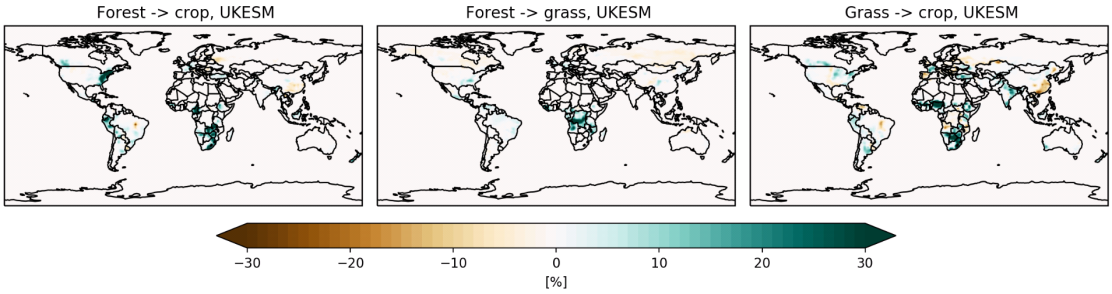

Figure 17: 2065-2095 average SSP3-7.0 - SSP1-2.6 vegetation fraction transitions in UKESM1-0_LL.

The temperature difference from ESMs is compared to the LC corrections for the difference between SSP1-2.6 and SSP3-7.0 LU in figure 18. Here, the biophysical temperature effect is filtered from background climate variability, according to the method described in appendix D.2.

A number of observations can be made from figure 18. First of all, some ESM show arctic warming, while others show arctic cooling. This is most likely the effect of increasing or decreasing atmospheric water vapour, caused by changes in the water cycle (Miller et al., 2007). Second, according to the data from Duveiller et al. (2018b) the largest temperature changes are expected in areas with high forest to grass and forest to crop transitions. In this case, one would expect a positive temperature change in central Africa, which is the case for all ESMs except UKESM1-0_LL. However, the BPH temperature change in this region for the other four ESMs is in the order of 0.5K, while the modelled correction temperature is much lower and even barely noticeable on this scale.

For objective observations, the BPH temperature changes versus the modelled correction temperature for the 2065-2095 time step are shown in figure 19. The scatter plots for all other timesteps and for SSP370LU-SSP126LU relative to SSP3-7.0 are shown and discussed in appendix D.3. The scatter plots reveal high variance in ESM local temperature responses, which is generally not represented in the temperature corrections. The changes in LC are expected to results in overall cooling in the boreal zone, and warming in the tropical zone, which is only true for CNRM-ESM2-1 and IPSL-CM6A-LR. Again, a strong correlation with a 1:1 slope is the optimal outcome, though the last two time frames from IPSL-CM6A-LR are the only data sets which remotely exhibit this behaviour.

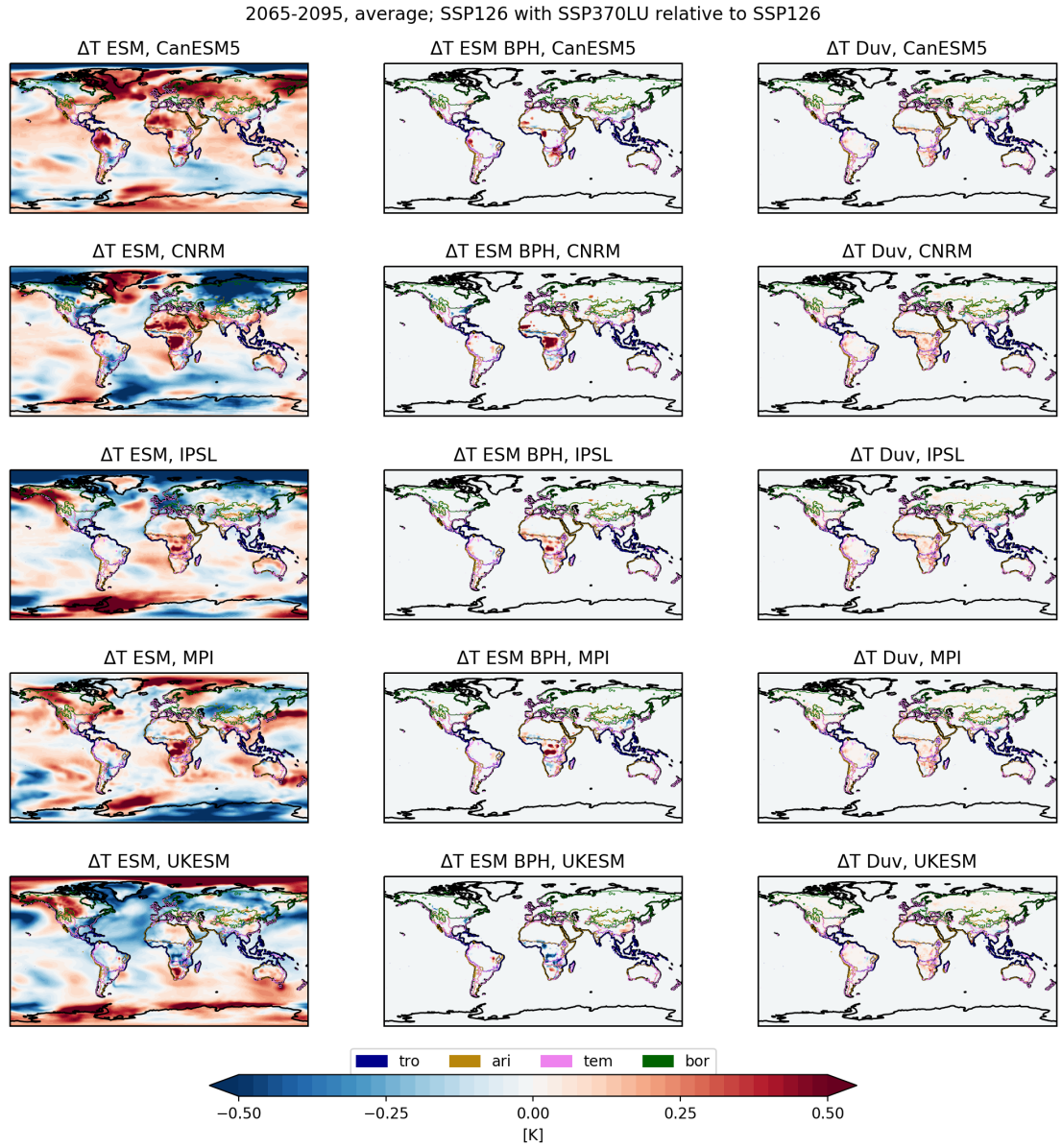

Figure 18: (left column) 2065-2095 average SSP370-SSP126LU - SSP1-2.6 ESM temperature difference; (middle column) SSP370-SSP126LU - SSP1-2.6 ESM BPH temperature difference; (right column) modelled correction temperatures (ΔT Duv). Contours indicate the boundaries of the tropical, arid, temperate and boreal zone, according to the Köppen-Geiger classification.

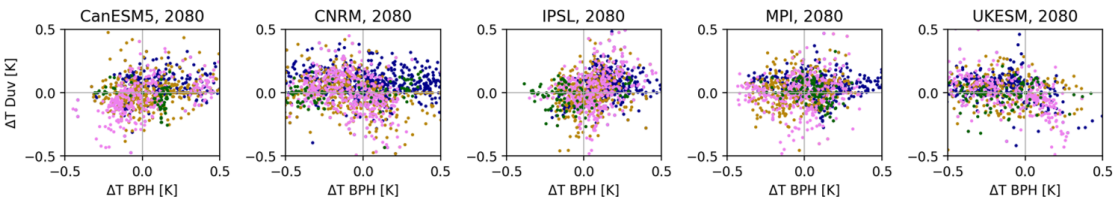

Figure 19: 2065-2095 average scatter plots of the modelled correction temperatures (ΔT Duv), vs the BPH filtered ESM temperatures. Blue, yellow, pink and green indicate data selection for the respective zones shown in figure 18.

3.2 Implementing temperature corrections in IMAGE

As described in the methodology, the temperature corrections are implemented in IMAGE-land for the selected SSPs from 2020 to 2100 with 5 year intervals. Individual aerosol concentration differences and land cover transitions are shown in appendix F. Here, only the results will be shown for the years 2060 and 2100 for SSP2-2.6 and SSP3-baseline, as these indicate the maximum range of corrections. Results for SSP2-baseline are shown in appendix G.

3.2.1 Air quality

In figures 20 to 23 the AQ corrections and their relevance, i.e. relative error, with respect to temperature change IMAGE is shown for the years 2060 and 2100 for SSP2-2.6 and SSP3-baseline. As could be expected from the emission timelines, the corrections for SSP2-2.6 are positive as aerosol loading is substantially reduced with respect to RCP6.0. SO_4 contributes the most to the total temperature correction, while BC corrections are almost negligible, as can be seen in appendix F.1. SO_4 corrections are particularly high in Pakistan and south-east China, which are marked by strong concentration differences. In this case, the corrections are highly relevant with respect to temperature change in IMAGE, as the relative error (T correction / IMAGE ΔT) in the majority of the corrected areas is close to 0.3 and even reaches up to 1. The pattern for SSP3-baseline in 2060 is roughly similar, though corrections are markedly smaller. As global temperatures change faster under SSP3-baseline, the relative error is much smaller than for SSP2-2.6. The same holds for SSP2, although corrections and relative errors are slightly higher.

In 2100, overall corrections are much smaller with respect to 2060, as concentration differences decreased through the continuous drop in RCP6.0 and stabilizing SSP aerosol precursor emissions. However, SSP2-2.6 and SSP2 now also show slight corrections in central South-America and central Africa as result of POM differences, as shown in figure 22 and figure F.5. 2100 global temperatures have slightly dropped in SSP2-2.6 as result of successful climate change mitigation, whereas global warming continuous in SSP2-baseline and SSP3-baseline. Relative errors are therefore still more pronounced in SSP2-2.6 than SSP2-baseline and SSP3-baseline.

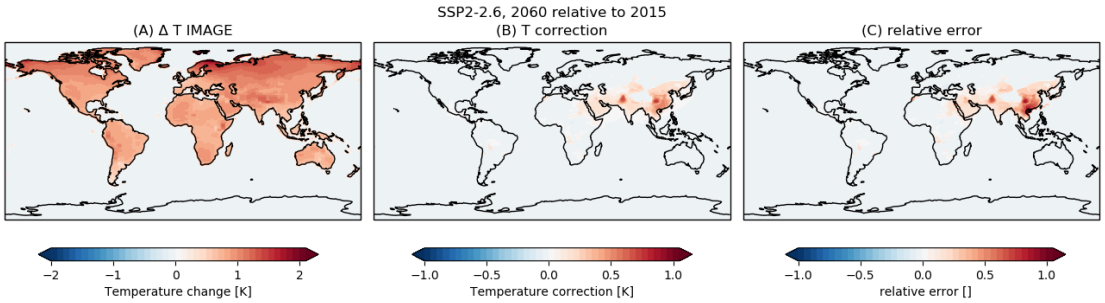

Figure 20: 2060, SSP2-2.6 (A) Temperature change in IMAGE relative to 2015; (B) AQ temperature corrections; (C) relative error: T correction / IMAGE ΔT

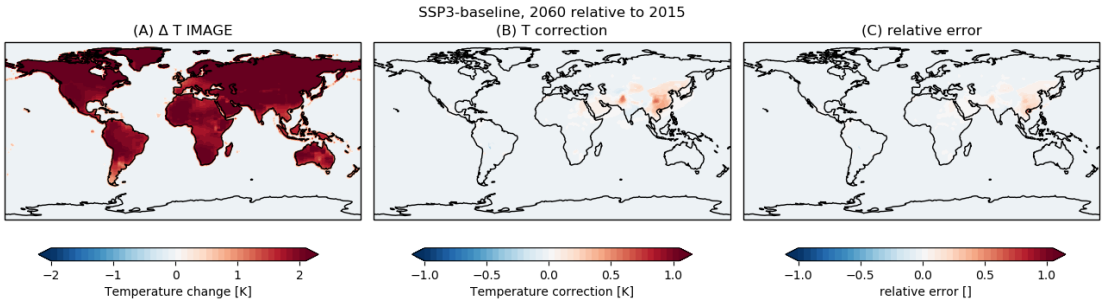

Figure 21: 2060, SSP3-baseline (A) Temperature change in IMAGE relative to 2015; (B) AQ temperature corrections; (C) relative error: T correction / IMAGE ΔT

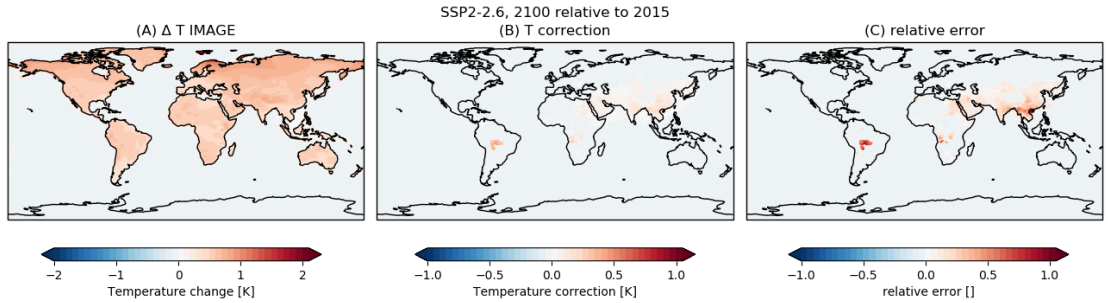

Figure 22: 2100, SSP2-2.6 (A) Temperature change in IMAGE relative to 2015; (B) AQ temperature corrections; (C) relative error : T correction / IMAGE ΔT

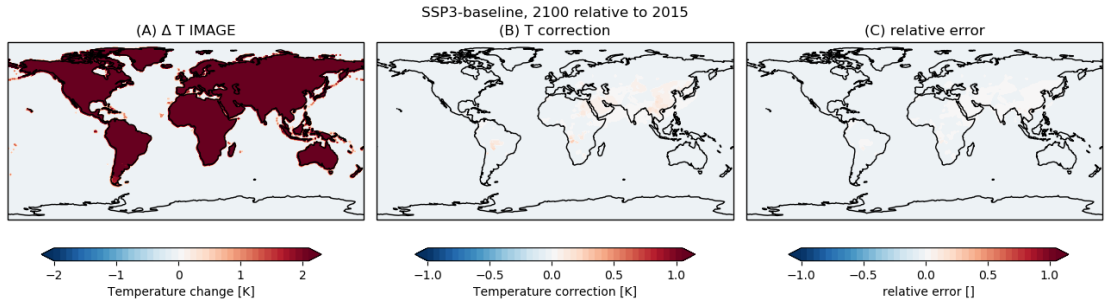

Figure 23: 2100, SSP3-baseline (A) Temperature change in IMAGE relative to 2015; (B) AQ temperature corrections; (C) relative error: T correction / IMAGE ΔT

3.2.2 Land Cover

In figures 24 to 27 the LC corrections and their relevance, i.e. relative error, with respect to temperature change IMAGE is shown for the years 2060 and 2100 for SSP2-2.6 and SSP3-baseline. Overall, the corrections patterns have smaller magnitudes and signs vary locally, in contrast to the stronger national and regional patterns for AQ corrections.

In 2060 the corrections for SSP2-2.6 are small and vary strongly in sign, as the grass to crop transition (for which the input data also strongly varies; figure 6c) is prominent, resulting in smaller relative errors compared to AQ corrections for this SSP. For SSP3-baseline, deforestation for crops and grass is present at a higher level, especially in South-America and Africa. However, relative errors are still very small, because of the high temperature rise in SSP3-baseline. The patterns for SSP2-baseline are roughly consistent with SSP3-baseline, where smaller global warming and smaller corrections yield likewise relative errors.

In 2100, the patterns become more distinct as the overall differences between RCP6.0 have grown larger. For SSP2-2.6, the corrections in South-America and Africa are now clearly negative, while some boreal areas show positive corrections and relative errors now reach up to 0.5. For SSP3-baseline, stronger positive corrections are now noticeable in South-America and Africa, though relative errors are small through higher global warming. Again, SSP2-baseline patterns are roughly consistent with SSP3-baseline, where somewhat smaller global warming and corrections result in similar relative errors.

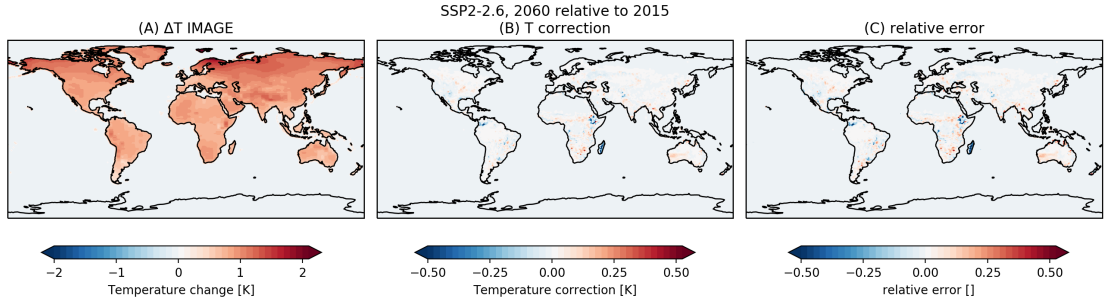

Figure 24: 2060, SSP2-2.6 (A) Temperature change in IMAGE relative to 2015; (B) LU temperature corrections; (C) relative error: T correction / IMAGE ΔT

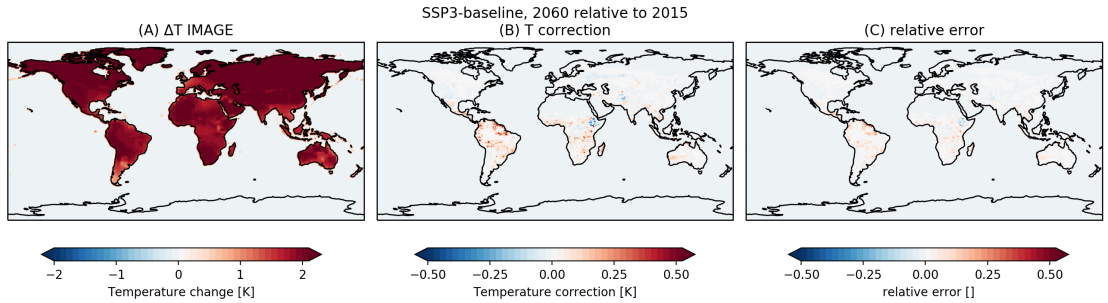

Figure 25: 2060, SSP3-baseline (A) Temperature change in IMAGE relative to 2015; (B) LU temperature corrections; (C) relative error: T correction / IMAGE ΔT

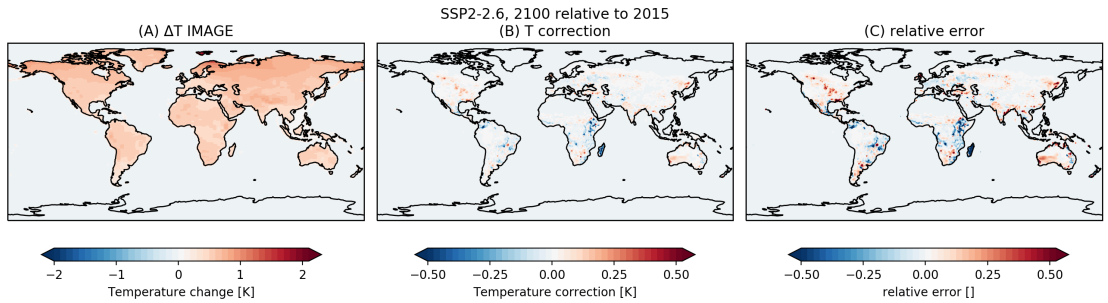

Figure 26: 2100, SSP2-2.6 (A) Temperature change in IMAGE relative to 2015; (B) LU temperature corrections; (C) relative error : T correction / IMAGE ΔT

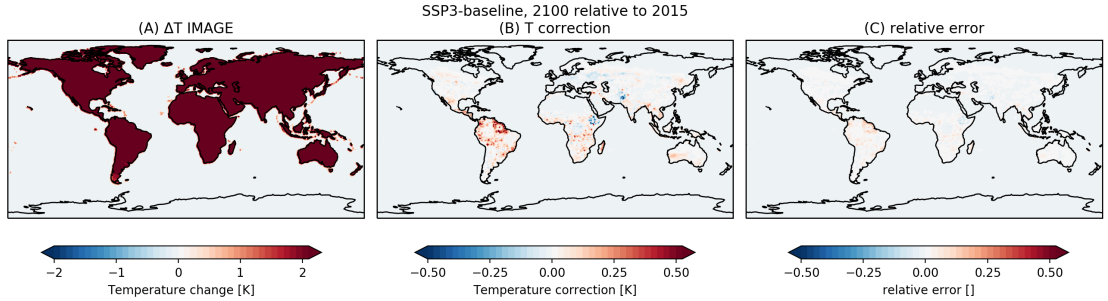

Figure 27: 2100, SSP3-baseline (A) Temperature change in IMAGE relative to 2015; (B) LU temperature corrections; (C) relative error: T correction / IMAGE ΔT

3.3 Effects on land-based processes

To asses the effects on land-based processes, the difference in crop yields and living biomass for a scenario with temperature corrections versus the same scenario without correction is calculated and shown in figure 28. Here, averages are shown for the 2050-2080 period for India, China, South-America and Africa. This period is expected to show the maximum effects, as temperature corrections for AQ peak and LC corrections have already grown substantial. India and China primarily indicate effects of AQ corrections while South-America and Africa are indicative for LC corrections.

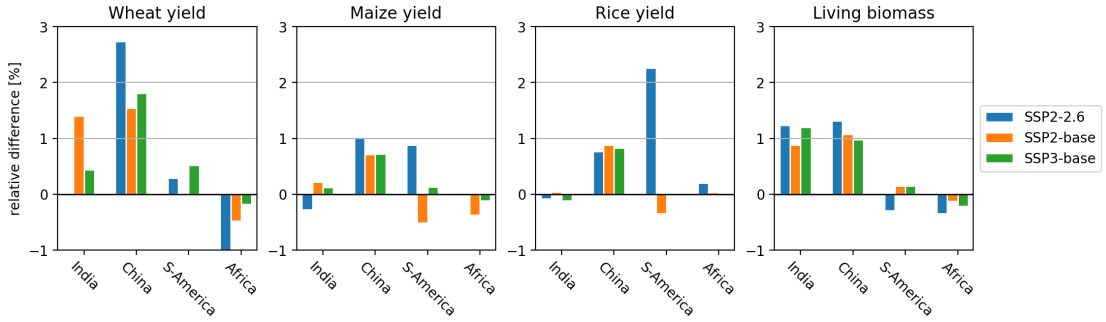

Figure 28: 2050-2080 average relative difference in wheat, maize & rice yield and living biomass for corrected vs non-correct runs in SSP2-2.6, SSP2-baseline and SSP3-baseline for India, China, South-America and Africa.

Figure 28 shows that temperature corrections for SSP2-2.6 has the highest overall agricultural and carbon cycle impacts. The effects on crop yield are most pronounced in China, though India shows the largest effects on biomass. The impacts on South-America and Africa are evidently smaller, which is most likely the consequence of their much lower average temperature correction.

4 Discussion

4.1 Air quality temperature corrections

Although this study was a first attempt to implement the local effect of aerosol loading changes with a simplified model, some interesting and promising results and insights have been produced. However, trying to capture the complex nature of aerosol effects on local climate within a simple model has proven to be difficult. Regarding the first validation experiment, we observe that the correction model is not able to reproduce the strong arctic warming which is clearly observed in ESMs. This is not necessarily a flaw of the correction model, rather a logical consequence of the simplified approach which was needed to find a solution for the original problem (limited abilities of ESM application within IAMs). The ESMs produce arctic warming as they account for feedbacks on the hydrological cycle and circulation, processes which can not be reproduced in simplified models. As expected and demonstrated in the validation experiments, higher aerosol loading differences result in higher temperature differences. However, all scatter plots also show higher slopes for higher cumulative concentration intervals. This indicates some form of non-linear concentration-temperature relationship in the PiClim data. In this case, the assumed linear response on radiation (Hansen, 2005) or the temperature response of radiation might be compromised.

As shown in the results, corrections are only in line with ESMs when comparing data in regions with high cumulative concentrations. Thus, AQ corrections are only meaningful in India and south-west China, and should be applied with consideration in other regions. The comparison with SST runs is more in line with the corrections, and fits and correlations are much better. For these runs are physically unrealistic they provide no guarantee that the model works properly for temperature correction between other scenarios. However, the sign of the model agrees with the ESMs in both experiments, which indicates that the corrections correspond to the known effects of aerosol loading reduction or increase. This promotes the application of the model in IMAGE; the effects would otherwise not be taken into account, while the correction results indicate that the effects are significant with respect to background climate change. Nonetheless, when this model for temperature corrections is applied, some other flaws should be considered as well:

1) Overall AQ corrections are limited to regions with considerable changes in concentrations in the PiClim experiments, while indirect effects (circulation changes, teleconnections, etc.) are not captured by the correction model. Moreover, the data selection also excludes correction for locations with future changes in aerosol loading. For example, if a newly emerging economy would depend on heavy industry and fossil fuels, with consequential higher aerosol loading, no cooling effect would be produced as there is no concentration-temperature relation in the current data set.

2) In addition, by not including the effects of nitrate aerosol, considerable corrections for some regions might not be included. In Europe and parts of North-America fertilizer and manure from intensive agriculture might account for relevant local corrections if nitrate aerosol effects would be included. For example, in Europe and the Po Valley, ammonium and nitrate aerosols account for a maximum -0.2 K and -0.4 K cooling respectively (Drugé et al., 2019).

3) Although using empirical data for correction modelling has considerable advantages, it also comes with some drawbacks, which might or might not account for the inconsistencies shown in validation experiment results. As the PiClim data is used as input, the scaling is based on the background climate of 1850. Temperature effects might thus not be modelled correctly, as the scaling also includes the black carbon on snow effect. Here, the assumed snow cover in some seasons and regions might not be there because of climate change. In the currently limited area for BC scaling, this might affect corrections in the southern Tibetan plateau.

4) For other aerosol components, a similar problem is also present, in the form of changes in circulation and atmospheric mixing. The concentration fields are modelled with FASST, for which the source-receptor relations are derived for the fixed meteorological year 2001. Although Van Dingenen et al. (2018) discuss the implications and possible biases for inter-annual variability, no estimates are given for the effects of climate change on meteorology. However, Xu and Lamarque (2018) demonstrate that meteorological changes result in a wet removal flux decrease, which account for a 5% to 10% higher aerosol loading at the end of the 21st century for RCP8.5. Although meteorological bias is thus a relevant factor on its own, it is still relatively small compared to the uncertainties in indirect aerosol effects (Myhre et al., 2013).

5) The PiClim runs are performed with fixed sea surface temperatures, where climate variability and climate change are effectively damped, especially in coastal areas. The effect of aerosol loading changes might thus be also underestimated in coastal areas and is completely suppressed over the ocean. The latter does not matter for IMAGE, because we are only interested in the effects above land, though one should take note of the effect in coastal areas.

4.2 Land cover temperature corrections

The simple model for LC corrections based on remote sensing data was able to reproduce the features from LC changes on local temperatures consistent with literature. However, the validation experiment shows that the data from Duveiller et al. (2018b) is not adequate to explain the differences in temperature between two scenarios with deviating LU found with ESMs. Nonetheless, this does not necessarily imply that Duveiller et al. (2018b) is not suitable for making corrections, while the inconsistencies can be caused by one or joint of the following effects:

1) The data from Duveiller et al. (2018b) is based on remote sensing data for the years 2008-2012, and therefore dependent on the background climate in this period. However, the background climate in the SSPs changes, which might result in corrections based on false BPH features. For example, the expected cooling due to smoother snow surfaces from deforestation in boreal zones might not occur in ESMs, due to future snow loss from temperature rise.

2) The ESM and Duveiller et al. (2018b) vegetation data that is used for making corrections is based on land cover, not land use. For example, irrigated land will generally be cooler than non-irrigated land due to increased latent heat release from stronger evapotranspiration (Yang et al., 2020). Future changes in land use might therefore affect local temperatures, even if the actual land cover does not change.

3) The ESMs vary in LC change patterns, which is a consequence of different criteria among ESMs for vegetation types. This is discussed in detail by Boysen et al. (2020), who show the variety in forest extent in nine ESMs, including the five used in this study in a controlled deforestation experiment. Moreover, Boysen et al. (2020) illustrate that energy balance changes from deforestation also strongly vary among ESMs. Other studies also find strong inconsistencies between land models and observations (Forzieri et al. (2018) and Duveiller et al. (2018a)). Although considerable advances have been made in land models (Clark et al., 2015), the effects on local circulation, cloud cover and other atmospheric effects remains highly uncertain. These kind of experiments show that the impacts are currently poorly understood and stress the importance of further development.

4) The data from Duveiller et al. (2018b) is aggregated by taking the average from high resolution observations. This implies that it can be best applied to estimate temperature changes from small scale and local LU changes, while it might underestimate for large scale changes, such as forests diebacks (Hoegh-Guldberg et al., 2018).

The inconsistencies between ESM temperature differences and modelled temperature corrections can thus not be attributed to the correction method. Even though validation has proven to be difficult, BPH LC effects were still applied to study the effects in IMAGE.

4.3 Implementing temperature corrections in IMAGE

As shown in the previous chapter, application of the temperature correction models for the new SSPs leads to considerable temperature corrections for some of the scenarios. Although the flaws of the models have already been discussed in the previous section, special attention is also devoted to the results of the temperature corrections for IMAGE.

Instead of integrating the models within the IMAGE framework (as suggested in the original set up) the models are forced with output from IMAGE runs as of time restrictions. However, through runs where the temperature patterns are added to the downscaled patterns, we assume to capture the first order effects of AQ and LC temperature corrections. Secondary effects, such as emission feedbacks from energy consumption or land use changes are thus not taken into account, but could play an important role as well.

The time restrictions for this study also resulted in a quick fix for the emission problems in IMAGE in the form of harmonization. The validity of harmonizing concentrations instead of emissions is checked with a small side experiment. Here, +50% and -50% emission perturbations in a RCP6.0 scenario showed

a +50% and -50% concentration field response for BC, POM and SO_2, confirming that in this case harmonization can be achieved through concentrations. However, the discrepancy in emission is caused by a defect in the IMAGE energy model, where the emissions from coal are overestimated. Here, the harmonization does not discriminate emission sectors as it is based on the total. As emissions from coal are expected to decline faster than other sectors (especially within mitigation scenarios), this might lead to an underestimation of emission in the future, as harmonization is based on the year 2015. Subsequently, the concentrations are estimated too low and the resulting temperature corrections might be too high.

4.4 Effects on land-based processes

Even though the corrections show some effect in land-based processes, the results are relatively small. For perspective, results of runs with varying RCP6.0 ESM downscaling patterns within SSP2-baseline are shown in figure 29. Here, the effects of applying the patterns from GFDL-ESM2M, hadGEM2-ES, and MIROC5 instead of the default IPSL-CM5A-LR are shown. The difference in 2050-2080 average temperature for the other ESMs with with respect to IPSL are shown in figure 30.

The same order of magnitude is also found in a similar experiment, where multiple land-vegetation models are combined with a variety of GCMs and run for the RCP scenarios (Rosenzweig et al., 2014). However, Rosenzweig et al. 2014 also demonstrates that the uncertainties are roughly in the same order of magnitude. The small effects from AQ and LC corrections can thus be considered negligible against the uncertainties of these kind of models.

Another important aspect is that the current temperature corrections are calculated on a yearly averaged basis. However, for crops and other natural processes, changes in the magnitude and frequency of extreme weather events are a far more important factor. As the correction procedures do not account for this, the effects of AQ and LC change might be underestimated.

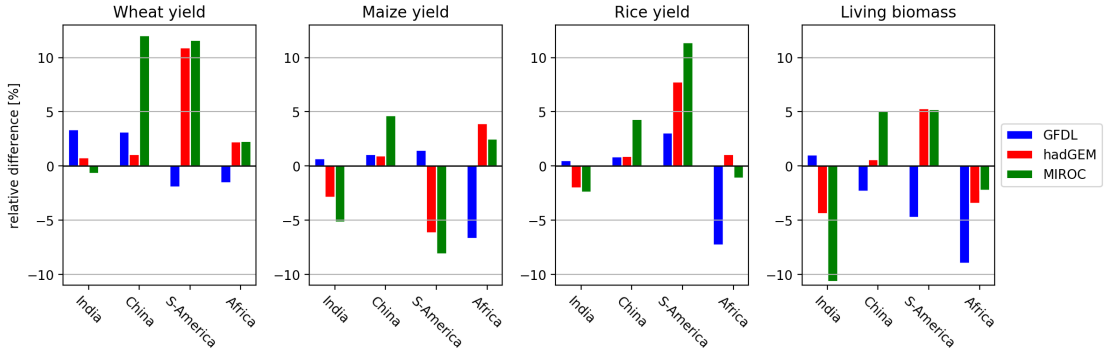

Figure 29: 2050-2080 average relative difference in wheat, maize & rice yield and living biomass for varying RCP6.0 ESM patterns with respect to IPSL-CM5A-LR (default) in SSP2-baseline for India, China, South-America and Africa.

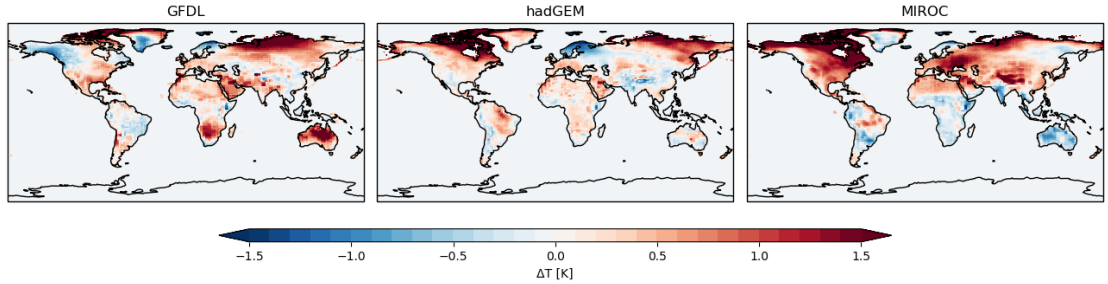

Figure 30: 2050-2080 RCP6.0 average temperature difference for GFDL-ESM2M, hadGEM2-ES and MIROC5 with respect to IPSL-CM5A-LR.

Lastly, the fact that temperature differences between downscaling patterns is far greater the modelled temperature corrections forces us to reconsider the necessity for temperature corrections in the first place:

Of course, the new method allows to incorporate the effects of AQ and LC changes, which were originally not explicitly taken into account. But as it turns out, selecting an ESM pattern is a far more important factor for local temperature change than the actual effect of changes in air quality or land use.

5 Conclusion

This research aims to identify the role of air quality and land cover change on local temperature and how these can be implemented in IAMs through simple models. Although the global radiative effects of these processes have been widely studied and documented, the actual local effect on temperature is not. Reflecting on the research questions, we conclude that:

- It is possible to model the local temperature effects of AQ and LC through simple models. However, as this research is a first attempt, corrections for AQ are still limited to certain regions due to data selection.
- For AQ, the correction model shows results which are more or less in line with MMMs from 5 ESMs. An unexpected achievement, considering the complex nature of local aerosol effects and the associated uncertainties. It is demonstrated that for a comparison between ESM and the correction model, the results are quite scattered and correlations and slopes are poor. However, for similar scenarios with fixed SST, results agree quite well. Especially for areas and time steps with strong changes in AQ, fit slopes up to 0.83 and correlation coefficients up to 0.88 are found. For LC, the correlation between the correction model and ESMs is absent in most cases and very poor at best for IPSL-CM6A-LR.
- By applying the correction models on the widest range of SSPs in IMAGE, it is demonstrated that incorporating the AQ effects lead to overall warming in the corrected regions, mainly because of SO_4 loading differences. As expected, the mitigation scenario (SSP2-2.6) yields the strongest corrections up to roughly 1 K, while the SSP3-baseline scenario shows maximum corrections up to roughly 0.5 K. For LC, applying corrections leads to overall warming in the SSP3-baseline scenario, especially central Africa and South-America. For SSP2-2.6 correction lead to overall cooling in central Africa and South America and warming in the boreal zone. For both AQ and LC the relative error (from not taking these effects into account) is highest in SSP2-2.6, as background climate change is higher in SSP3-baseline
- When looking at the relative effect on land-based processes from applying temperature corrections, we find regional average differences in the order of 1% for crop yields and living biomass. However, taking another ESM downscaling pattern within a single scenario yields differences in the order of 10%.

6 Recommendations

Recommendations are given in three categories: 1) improvements for the current IMAGE downscaling process, 2) improvements for the proposed correction methods and 3) suggestions for the development of a new generation correction method.

The current downscaling process (without corrections) could be improved by taking care of:

- Use scenarios for downscaling which are most consistent with the studied scenario at all times, i.e. RCP2.6 for SSPx-2.6 etc.. As RCP6.0 is currently used as a default, effects such as arctic amplification might lead to an overestimation of temperature increase on the northern hemisphere for mitigation scenarios.
- As ISIMIP (Inter Sectoral Impact Model Intercomparison Project) is entering the third phase, bias corrected temperature data from CMIP6 scenario runs will soon be available. This will allow the use of SSP data sets for downscaling, opposed to the current usage of RCP patterns. It is strongly recommended that SSP data sets should be used for downscaling for three reasons:
  1) The emissions and land use which are used to force ESMs are far more consistent between old and new SSP generations than among RCP and SSP scenarios with similar radiative forcing targets. This will in turn decrease the need for corrections in the first place.
  2) ESM have considerably improved since the CMIP5 generation. Not only in temporal and spatial resolution, but individual model components have been upgraded as well, and some models have been extend with new modules. Overall, as more physical and chemical processes are incorporated, overall model performance for historical experiments is better in CMIP6 than CMIP5 (e.g. Thorarinsdottir et al. (2020), Fan et al. (2020)) and ESMs are thus expected to yield more accurate future projections.
  3) For the SSPs, the amount of realizations ESM has been increased with respect to RCPs. In total, every SSPs scenario has been run between 212 and 310 times, while the amount of RCPs realizations is varying from 36 to 153. This can be attributed to increasing amount of participating ESMs, though the amount of ensemble members per ESM has increased as well: the IPSL ESM for example, provided 4, 5, 1 and 5 realizations for the four RCP scenarios, while in CMIP6 6, 11, 11, 7 and 6 runs were performed for the five SSP marker scenarios. Larger ensemble means can in turn be used to make more accurate estimates for future patterns.
- Apply 30 year moving average temperature patterns instead of linearly interpolating data between 2015 and 2100 (2070-2100 average). Currently, non-linear temporal effects are not account for, within mitigation scenarios (where temperatures peak before 2100) local warming might thus be underestimated.

The results indicate that the adapted downscaling models are able to capture the local effects of AQ and LC change on temperature consistent with literature. However, the downscaling models as proposed and applied in this paper are not flawless and it is recommended that:

- The proposed AQ and LC correction models should be implemented in IMAGE, to capture second order effects and possible feedbacks.
- AQ corrections should be based on individual ESMs. Currently, the AQ corrections are limited by the PiClim robustness condition, defined as the number of ESMs with an identical sign change in the 30 year average temperature response. By applying standard statistical methods, such as a t-test on the monthly data (as demonstrated in Zanis et al. (2020)), the significant temperature change areas could be defined for individual ESMs. Subsequently, temperature corrections could then be made for individual ESMs.
- LC temperature corrections should account for changes in background climate. A good strategy to incorporate this would be to aggregate data from Duveiller et al. (2018b) to climate zones, for

example those according to Köppen-Geiger. As these zones are defined by (seasonal) precipitation and temperature, the extend of the zones can be defined by temperature and precipitation output from IMAGE, and temperature corrections can be made accordingly.

Even after taking notice and incorporating the aforementioned recommendations, the downscaling method is still limited. Other possible steps towards the development of a next generation downscaling correction process are:

- Developing and implementing other correction models for phenomena such the Urban Heat Island effect (Akbari, 2001) and the drying of lakes such as the Aral sea (Small, 2001).
- Study the effects and possibilities of nitrate aerosol related temperature corrections.
- As ESMs have currently passed the first phases of CMIP6, more emphasis is now laid on experiments for fundamental research and more data is soon to be expected for AerChemMIP and LUMIP runs. This is expected to improve signal to noise ratios, and allow a better understanding of global (and indirect) effects of AQ and LC changes. For AQ, special attention should be given to the role of BC and OC, as their present day radiative forcing is in the same order of SO_4, but uncertainties in their effects on climate (especially locally as shown in this work) are much higher.
- In the same line of thought, experiments such as those carried out by Mendelsohn et al. (2001), could also be very helpful to improve AQ corrections. By perturbing emissions for IMAGE regions among different background climate states, more insight could be provided in the changes in circulation, teleconnections and similar transportation related effects. By isolating regional and inter-regional effects, it would be possible to also make corrections for inter-regional effects, instead of the current limited method. Moreover, by perturbing emissions within different climate states, the aerosol lifetimes feedback, BC on snow, and water vapour cycle effects would be accounted for as well. However, the amount of required runs would be enormous which is not favourable.
- Although Duveiller et al. (2018b) provides a helpful and realistic data set to model local effects of land cover change on a local scale, it misses the effects on a global scale which arise from large scale land use change. As large scale land cover change is likely to continue in the future, a good representation of these processes and their global effects is important. As considerable improvements in land models have been made lately, deriving a relation between local temperature and LC change from ESMs would be an interesting alternative for LC corrections. However, the uncertainty of these effects is not likely to improve swiftly, and this option should thus be considered with care.

References

Akbari, H. (2001). Energy Saving Potentials and Air Quality Benefits of Urban Heat Island Mitigation. *Solar Energy*, 70(3):295–310.

Alkama, R. and Cescatti, A. (2016). Biophysical climate impacts of recent changes in global forest cover. *Science*, 351(6273):600–604.

Allen, R. J., Turnock, S., Nabat, P., Neubauer, D., Lohmann, U., Olivie, D., Oshima, N., Michou, M., Wu, T., Zhang, J., Schulz, M., Tsigaridis, K., Bauer, S. E., Emmons, L., Horowitz, L., Naik, V., van Noije, T., Bergman, T., Lamarque, J.-F., Zanis, P., Tegen, I., Westervelt, D. M., Le Sager, P., Good, P., Shim, S., O'Connor, F., Akritidis, D., Georgoulias, A. K., Deushi, M., Sentman, L. T., Fujimori, S., and Collins, W. J. (2020). Climate and air quality impacts due to mitigation of non-methane near-term climate forcers. preprint, Aerosols/Atmospheric Modelling/Troposphere/Chemistry (chemical composition and reactions).

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., Lavergne, C. d., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, E., L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, 12(7):e2019MS002010. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002010.

Boysen, L., Brovkin, V., Pongratz, J., Lawrence, D., Lawrence, P., Vuichard, N., Peylin, P., Liddicoat, S., Hajima, T., Zhang, Y., Rocher, M., Delire, C., Séférian, R., Arora, V. K., Nieradzik, L., Anthoni, P., Thiery, W., Laguë, M., Lawrence, D., and Lo, M.-H. (2020). Global climate response to idealized deforestation in CMIP6 models. preprint, Earth System Science/Response to Global Change: Climate Change.

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C., Swenson, S. C., and Zeng, X. (2015). Improving the representation of hydrologic processes in Earth System Models. *Water Resources Research*, 51(8):5929–5956. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015WR017096.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J. (2017). AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6. *Geoscientific Model Development*, 10(2):585–607.

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, 12(2).

Dinar, A., Mendelsohn, R., and Williams, L. (2006). Mendelsohn R, Dinar A, Williams L. The dis-

tributional impact of climate change on rich and poor countries. Environment and Development Economics. *Environment and Development Economics*, 11:159–178.

Drugé, T., Nabat, P., Mallet, M., and Somot, S. (2019). Model simulation of ammonium and nitrate aerosols distribution in the Euro-Mediterranean region and their radiative and climatic effects over 1979–2016. *Atmospheric Chemistry and Physics*, 19(6):3707–3731.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, 40(9-10):2123–2165.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M. (2019). The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics.

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J., Cooke, W., Dunne, K. A., Harrison, M. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Phillipps, P. J., Sentman, L. T., Samuels, B. L., Spelman, M. J., Winton, M., Wittenberg, A. T., and Zadeh, N. (2012). GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics. *Journal of Climate*, 25(19):6646–6665.

Duveiller, G., Caporaso, L., Abad-Viñas, R., Perugini, L., Grassi, G., Arneth, A., and Cescatti, A. (2020). Local biophysical effects of land use and land cover change: towards an assessment tool for policy makers. *Land Use Policy*, 91:104382.

Duveiller, G., Forzieri, G., Robertson, E., Li, W., Georgievski, G., Lawrence, P., Wiltshire, A., Ciais, P., Pongratz, J., Sitch, S., Arneth, A., and Cescatti, A. (2018a). Biophysics and vegetation cover change: a process-based evaluation framework for confronting land surface models with satellite observations. *Earth System Science Data*, 10:1265–1279.

Duveiller, G., Hooker, J., and Cescatti, A. (2018b). A dataset mapping the potential biophysical effects of vegetation cover change. *Scientific Data*, 5(1):180014.

Duveiller, G., Hooker, J., and Cescatti, A. (2018c). The mark of vegetation change on Earth's surface energy balance. *Nature Communications*, 9(1):679.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, page 22.

Fan, X., Miao, C., Duan, Q., Shen, C., and Wu, Y. (2020). The Performance of CMIP6 Versus CMIP5 in Simulating Temperature Extremes Over the Global Land Surface. *Journal of Geophysical Research: Atmospheres*, 125(18):e2020JD033031. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD033031.

Forzieri, G., Duveiller, G., Georgievski, G., Li, W., Robertson, E., Kautz, M., Lawrence, P., Garcia San Martin, L., Anthoni, P., Ciais, P., Pongratz, J., Sitch, S., Wiltshire, A., Arneth, A., and Cescatti, A. (2018). Evaluating the Interplay Between Biophysical Processes and Leaf Area Changes in Land Surface Models. *Journal of Advances in Modeling Earth Systems*, 10(5):1102–1126.

Hansen, J. (2005). Efficacy of climate forcings. *Journal of Geophysical Research*, 110(D18):D18104.

Hartin, C. A., Patel, P., Schwarber, A., Link, R. P., and Bond-Lamberty, B. P. (2015). A simple object-oriented and open-source model for scientific and policy analyses of the global climate system – Hector v1.0. *Geoscientific Model Development*, 8(4):939–955.

Hauglustaine, D. A., Balkanski, Y., and Schulz, M. (2014). A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate. *Atmospheric Chemistry and Physics*, 14(20):11031–11063.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., and Zhou, G. (2018). Impacts of 1.5°C of Global Warming on Natural and Human Systems. *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*, page 138.

Jones, A. D., Calvin, K. V., Collins, W. D., and Edmonds, J. (2015). Accounting for radiative forcing from albedo change in future global land-use scenarios. *Climatic Change*, 131(4):691–703.

Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O'Connor, F. M., Andres, R. J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S., Palin, E. J., Parsons Chini, L., Raddatz, T., Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M. (2011). The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geoscientific Model Development*, 4(3):543–570.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmospheric Chemistry and Physics*, 5(2):417–432.

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., and Shevliakova, E. (2016). The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6:rationale and experimental design. *Geoscientific Model Development*, 9(9):2973–2998.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., Storch, J.-S. v., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. *Journal of Advances in Modeling Earth Systems*, 11(4):998–1038. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018MS001400.

Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L. (2011). Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration. *Atmospheric Chemistry and Physics*, 11(4):1417–1456.

Mendelsohn, R., Schlesinger, M., and Williams, L. (2001). The climate impacts of sulfate aerosols. *Integrated Assessment*, 2(3):111–122.

Miller, J. R., Chen, Y., Russell, G. L., and Francis, J. A. (2007). Future regime shift in feedbacks during Arctic winter. *Geophysical Research Letters*, 34.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H. (2013). IPCC WG1AR5 Anthropogenic and Natural Radiative Forcing. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

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

Model Development, 9(9):3461–3482.

O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R., and van Vuuren, D. P. (2014). A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Climatic Change*, 122(3):387–400.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P. (2015). Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative. *Geoscientific Model Development*, 8(7):2315–2328.

Prevedello, J. A., Winck, G. R., Weber, M. M., Nichols, E., and Sinervo, B. (2019). Impacts of forestation and deforestation on local temperature across the globe. *PLOS ONE*, 14(3):e0213368.

Ramanathan, V., Crutzen, P. J., Lelieveld, J., Mitra, A. P., Althausen, D., Anderson, J., Andreae, M. O., Cantrell, W., Cass, G. R., Chung, C. E., Clarke, A. D., Coakley, J. A., Collins, W. D., Conant, W. C., Dulac, F., Heintzenberg, J., Heymsfield, A. J., Holben, B., Howell, S., Hudson, J., Jayaraman, A., Kiehl, J. T., Krishnamurti, T. N., Lubin, D., McFarquhar, G., Novakov, T., Ogren, J. A., Podgorny, I. A., Prather, K., Priestley, K., Prospero, J. M., Quinn, P. K., Rajeev, K., Rasch, P., Rupert, S., Sadourny, R., Satheesh, S. K., Shaw, G. E., Sheridan, P., and Valero, F. P. J. (2001). Indian Ocean Experiment: An integrated analysis of the climate forcing and effects of the great Indo-Asian haze. *Journal of Geophysical Research: Atmospheres*, 106(D22):28371–28398. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2001JD900133.

Rao, N. D., van Ruijven, B. J., Riahi, K., and Bosetti, V. (2017). Improving poverty and inequality modelling in climate research. *Nature Climate Change*, 7(12):857–862.

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

Rogelj, J., Popp, A., Calvin, K. V., Luderer, G., Emmerling, J., Gernaat, D., Fujimori, S., Strefler, J., Hasegawa, T., Marangoni, G., Krey, V., Kriegler, E., Riahi, K., van Vuuren, D. P., Doelman, J., Drouet, L., Edmonds, J., Fricko, O., Harmsen, M., Havlík, P., Humpenöder, F., Stehfest, E., and Tavoni, M. (2018). Scenarios towards limiting global mean temperature increase below 1.5 °C. *Nature Climate Change*, 8(4):325–332.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T. A. M., Schmid, E., Stehfest, E., Yang, H., and Jones, J. W. (2014). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences*, 111(9):3268–3273.

Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S. E., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y.-H., Cheng, Y., Clune, T. L., Del Genio, A., de Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Lo, K. K., Matthews, E. E., Menon, S., Miller, R. L., Oinas, V., Oloso, A. O., Perlwitz, J. P., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R. A., Tausnev, N., Tsigaridis, K., Unger, N., Voulgarakis, A., Yao, M.-S., and Zhang, J. (2014). Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive: GISS MODEL-E2 CMIP5 SIMULATIONS. *Journal of Advances in Modeling Earth Systems*, 6(1):141–184.

Seland, , Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schancke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Hafsahl Karset, I. H., Landgren, O., Liakka, J., Onsum Moseid, K., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iverson, T., and Schulz, M. (2020). The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6

DECK and historical simulations. preprint, Climate and Earth System Modeling.

Senior, C. A., Jones, C. G., Wood, R. A., Sellar, A., Belcher, S., Klein-Tank, A., Sutton, R., Walton, J., Lawrence, B., Andrews, T., and Mulcahy, J. P. (2019). UK Community Earth System Modelling for CMIP6. *Journal of Advances in Modeling Earth Systems*, n/a(n/a):e2019MS002004. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002004.

Small, E. E. (2001). Changes in Surface Air Temperature Caused by Desiccation of the Aral Sea. *Journal of Climate*, 14:284–299.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, 12(11):4823–4873.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y-Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G. (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*, 11(12):4182–4227. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001791.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, 12(7):2727–2765.

Thorarinsdottir, T. L., Sillmann, J., Haugen, M., Gissibl, N., and Sandstad, M. (2020). Evaluation of CMIP5 and CMIP6 simulations of historical surface air temperature extremes using proper evaluation methods. *Environmental Research Letters*, 15(12):124041.

Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment (1967)*, 8(12):1251 – 1256.

Van Dingenen, R., Dentener, F., Crippa, M., and Leitao, J. (2018). TM5-FASST: a global atmospheric source–receptor model for rapid impact analysis of emission changes on air quality and short-lived climate pollutants. *Atmospheric Chemistry and Physics*, 18(21):16173–16211.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, 109(1-2):5–31.

van Vuuren, D. P. and Stehfest, E. (2017). Energy, land-use and greenhouse gas emissions trajectories under a green growth paradigm. *Global Environmental Change*, 42:237–250.

Watanabe, M., Suzuki, T., O'ishi, R., Komuro, Y., Watanabe, S., Emori, S., Takemura, T., Chikira, M., Ogura, T., Sekiguchi, M., Takata, K., Yamazaki, D., Yokohata, T., Nozawa, T., Hasumi, H., Tatebe, H., and Kimoto, M. (2010). Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity. *Journal of Climate*, 23(23):6312–6335.

West, P. C., Narisma, G. T., Barford, C. C., Kucharik, C. J., and Foley, J. A. (2011). An alternative approach for quantifying climate regulation by ecosystems. *Frontiers in Ecology and the Environment*, 9(2):126–133.

Westervelt, D. M., Mascioli, N. R., Fiore, A. M., Conley, A. J., Lamarque, J.-F., Shindell, D. T., Faluvegi, G., Previdi, M., Correa, G., and Horowitz, L. W. (2020). Local and remote mean and extreme temperature response to regional aerosol emissions reductions. *Atmospheric Chemistry and Physics*, 20(5):3009–3027.

Wu, T., Zhang, F., Zhang, J., Jie, W., Zhang, Y., Wu, F., Li, L., Yan, J., Liu, X., Lu, X., Tan, H., Zhang, L., Wang, J., and Hu, A. (2020). Beijing Climate Center Earth System Model version 1 (BCC-ESM1): model description and evaluation of aerosol simulations. *Geoscientific Model Development*, 13(3):977–1005.

Xu, J. and Yang, Y. (2020). Impact of SO2 Emission on the Gross Domestic Product Growth of China. *Aerosol and Air Quality Research*, 20(4):787–799.

Xu, Y. and Lamarque, J. (2018). Isolating the Meteorological Impact of 21st Century GHG Warming on the Removal and Atmospheric Loading of Anthropogenic Fine Particulate Matter Pollution at Global Scale. *Earth's Future*, 6(3):428–440.

Yang, Q., Huang, X., and Tang, Q. (2020). Global assessment of the impact of irrigation on land surface temperature. *Science Bulletin*, 65.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M. (2019). The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *Journal of the Meteorological Society of Japan. Ser. II*, 97(5):931–965.

Zanis, P., Akritidis, D., Georgoulias, A. K., Allen, R. J., Bauer, S. E., Boucher, O., Cole, J., Johnson, B., Deushi, M., Michou, M., Mulcahy, J., Nabat, P., Olivié, D., Oshima, N., Sima, A., Schulz, M., Takemura, T., and Tsigaridis, K. (2020). Fast responses on pre-industrial climate from present-day aerosols in a CMIP6 multi-model study. *Atmospheric Chemistry and Physics*, 20(14):8381–8404.

Appendices

A IMAGE model framework

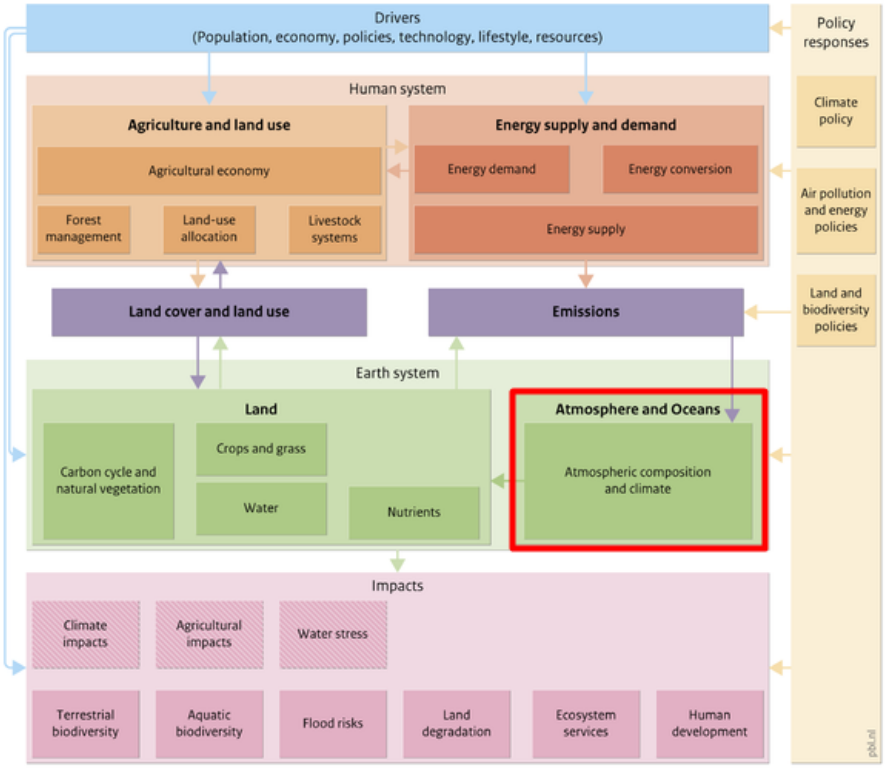

Figure A.1: The IMAGE 3.0 framework (source: PBL), the climate model component outlined in red.

B Supplementary PiClim data

B.1 ESM temperature responses

The figures in this appendix show the temperature response for the PiClim perturbation experiments with respect to the control run, for the individual ESMs as well as the MMM. A to C show the temperature response for individual perturbations in BC, OC and SO2 respectively. D shows the sum of temperature for BC, OC and SO2, e.g. total response. E shows the response of the aerosol run and F shows the difference between the total response and aerosol run. 'avg' denotes the global average temperature change.

The ESMs exhibit various temperature pattern responses to the three emission perturbations. However, some models also show arctic and/or antarctic warming or cooling, while this is not expected as sea surface temperatures and sea ice extent are fixed. This does not seem to directly affected the regions of interest i.e. South-West, South and East Asia. However, the results from MIROC6 (figure B.7) show distinct patterns over land and sea in most areas, as well as strong arctic warming in the BC, OC and SO2 runs. These patterns do not appear in the aerosol run and also do not resemble those from other ESMs. Therefore the results from MIROC6 are dismissed to calculate the MMM responses.

In order to calculate corrections, it is assumed the temperature impact of aerosol changes can be reasonably modelled by the total effects of changes in emissions of BC, OC, and SO2, as precursor emissions for smaller current-day radiative are not taken into account. Furthermore it is assumed that the temperature responses of BC, OC and SO2 are additive.

If both assumptions are correct, the total temperature response of the BC, OC, and SO2 runs should resemble that of the aerosol run, especially in the regions of interest for making corrections. Comparing figures D and E for the individual ESMs, we observe that these resemble each other more or less in most ESMs. However, the figures D and E for the MMM (figure B.8, which is ultimately used for making corrections), show patterns in South-East, South and East Asia which are highly consistent, confirming the two assumptions mentioned above.

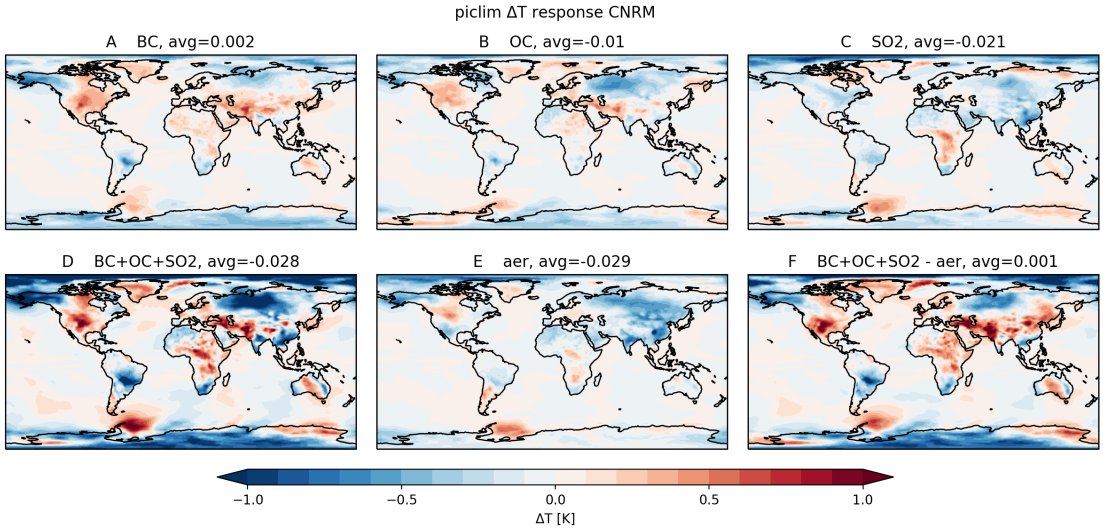

Figure B.1: PiClim CNRM-ESM2-1 temperature responses

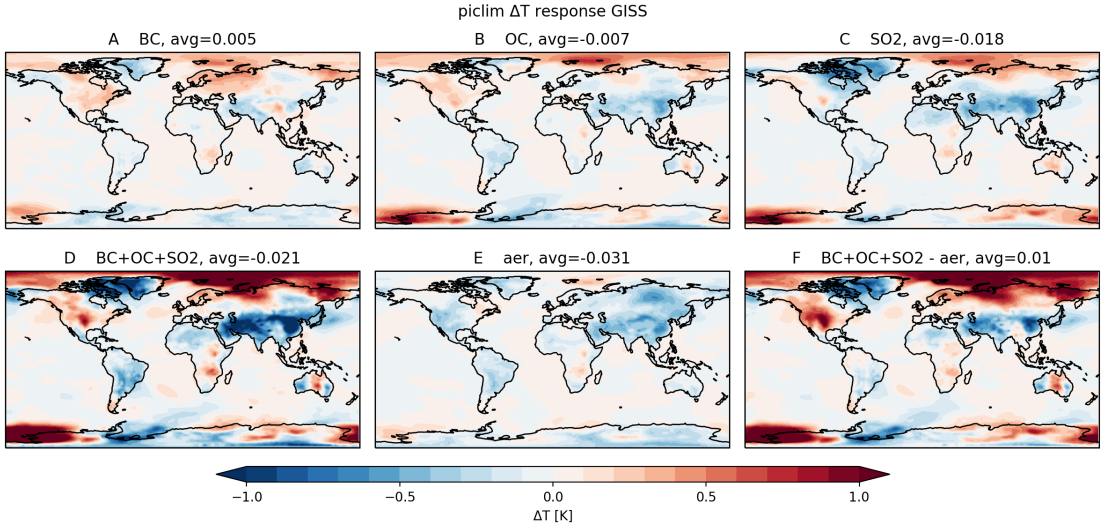

Figure B.2: PiClim GISS-E2-1-G temperature responses

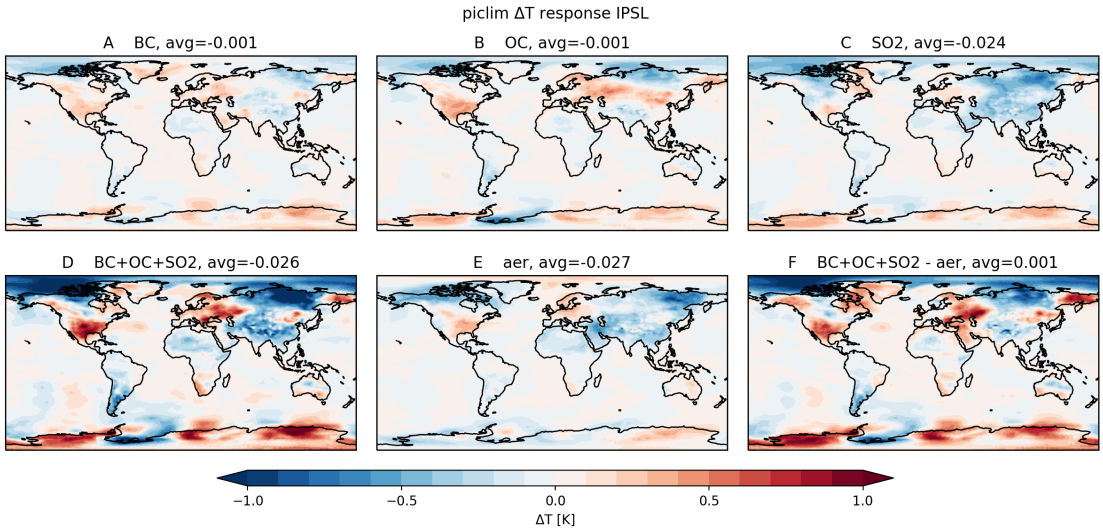

Figure B.3: PiClim IPSL-CM6A-LR-INCA temperature responses

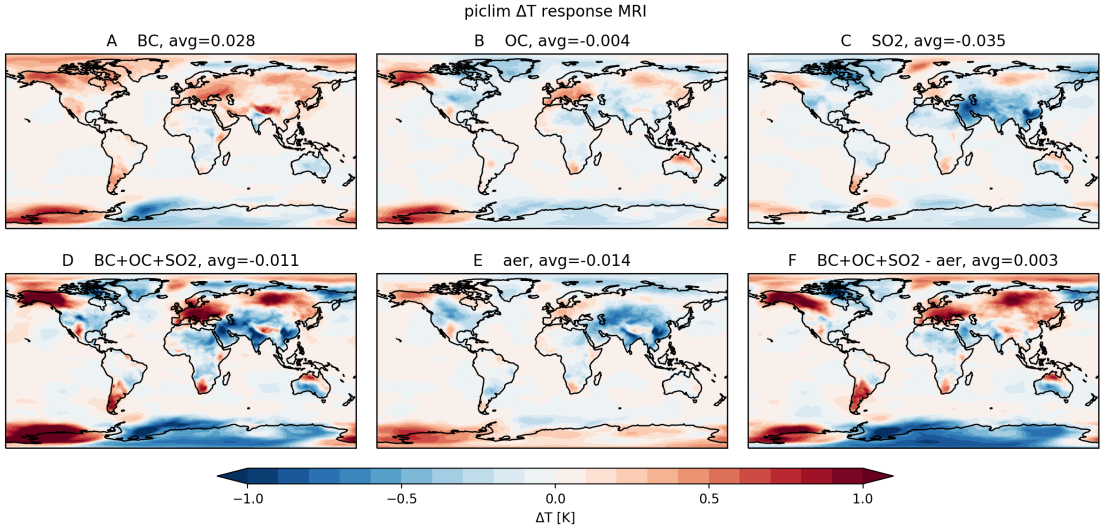

Figure B.4: PiClim MRI-ESM2-0 temperature responses

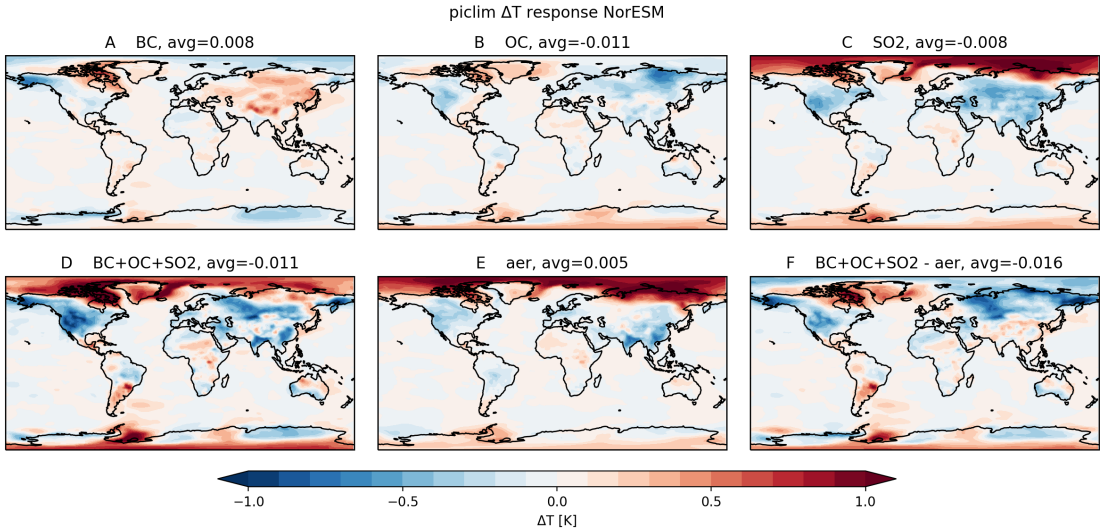

Figure B.5: PiClim NorESM2-LM temperature responses

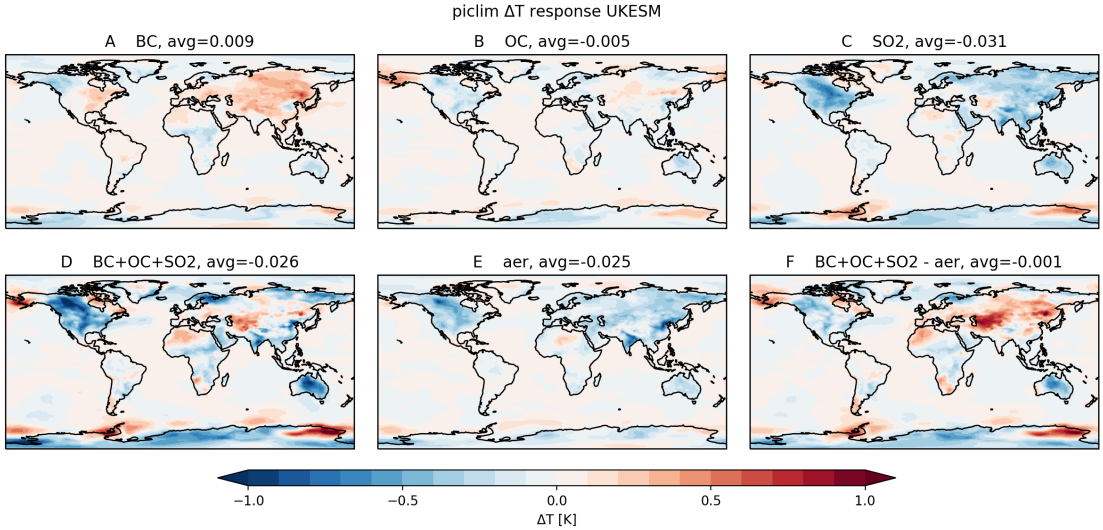

Figure B.6: PiClim UKESM1-0-LL temperature responses

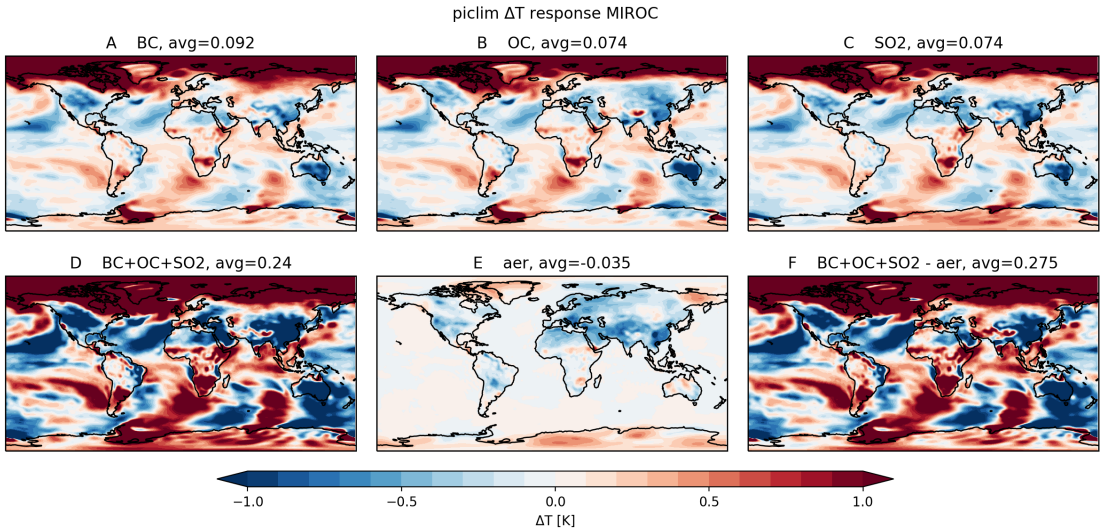

Figure B.7: PiClim MIROC6 temperature responses

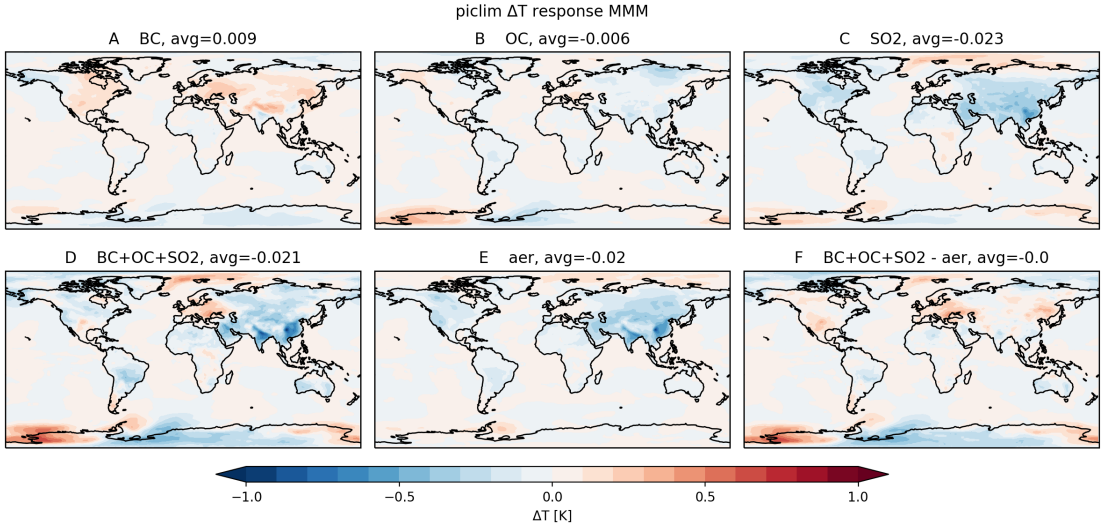

Figure B.8: PiClim MMM temperature responses (without MIROC6)

B.2 Model agreement

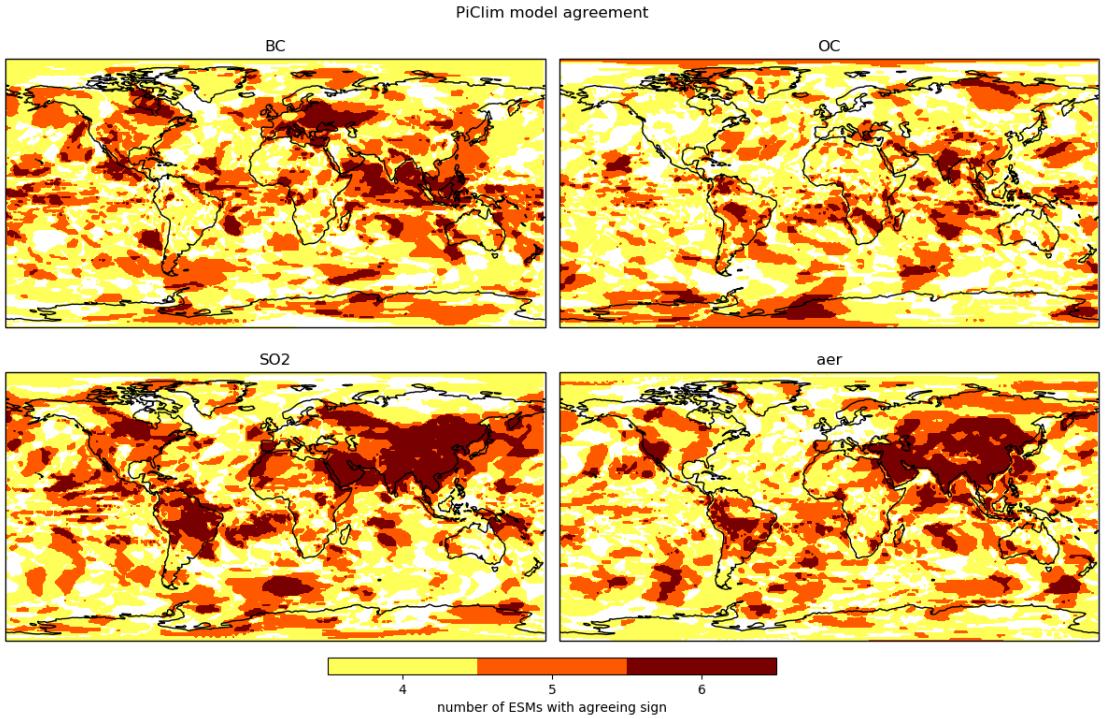

Figure B.9: Mapping of the number of models with similar temperature response sign for BC, OC, SO_2 and all aerosol precursor 'aer' emission perturbations. White indicates a three out of six model agreement.

C Air quality validation results

C.1 2020, 2030 and 2040 concentration differences and temperature corrections

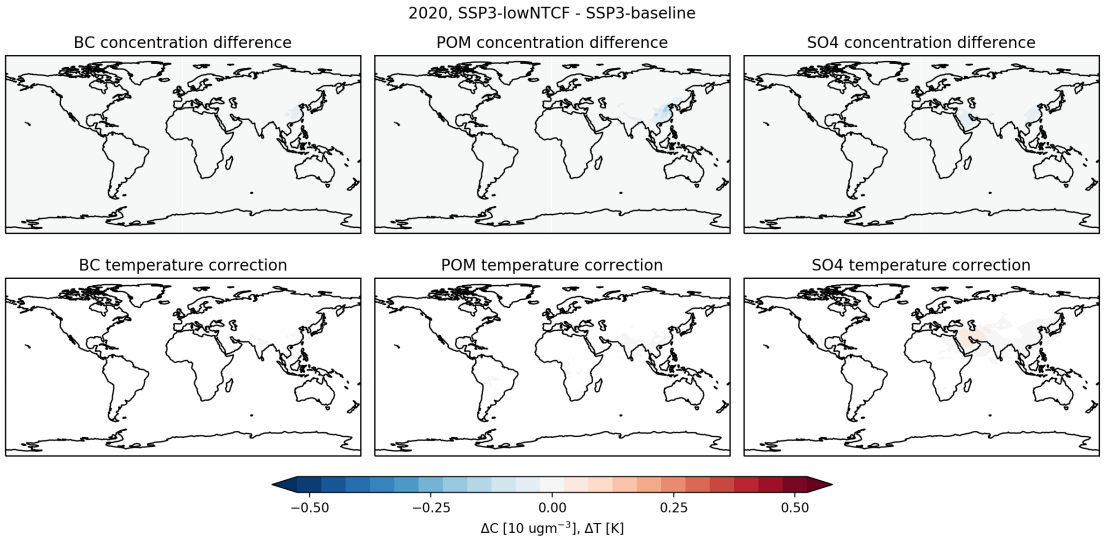

Figure C.1: SSP3 lowNTCF - baseline concentration differences modelled with FASST and respective temperature corrections derived with PiClim scaling for the 2015-2025 average.

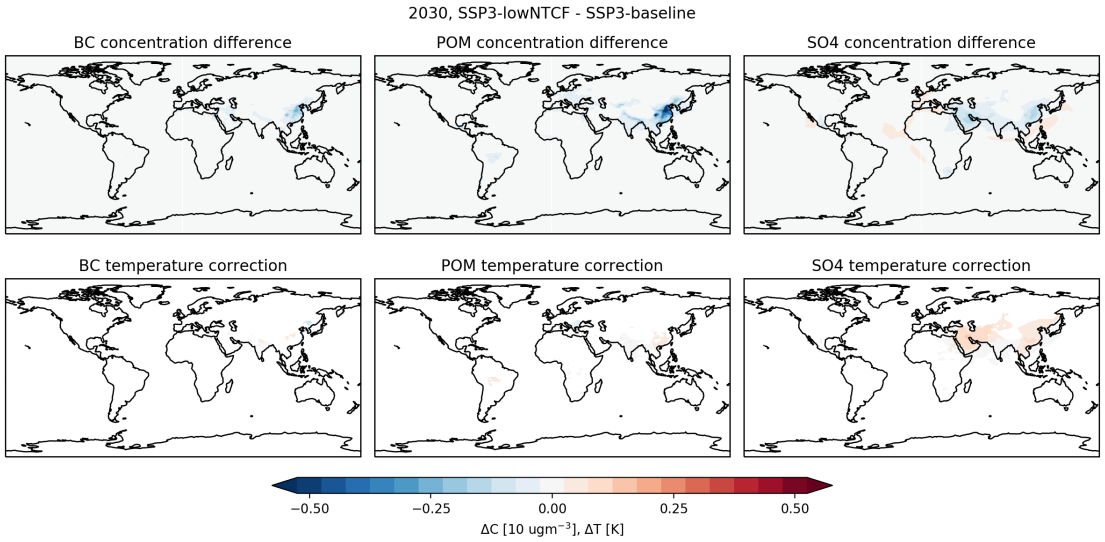

Figure C.2: SSP3 lowNTCF - baseline concentration differences modelled with FASST and respective temperature corrections derived with PiClim scaling for the 2025-2035 average.

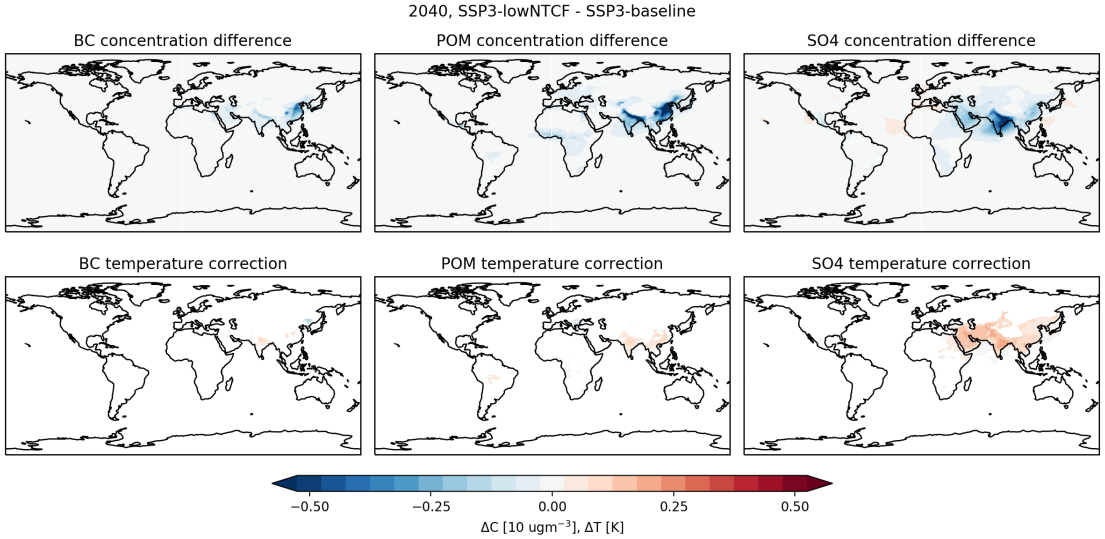

Figure C.3: SSP3 lowNTCF - baseline concentration differences modelled with FASST and respective temperature corrections derived with PiClim scaling for the 2035-2045 average.

C.2 SSP3 results

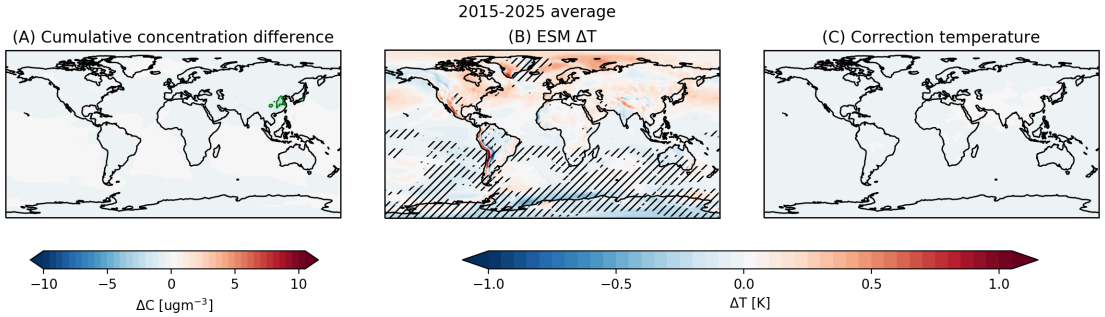

Figure C.4: 2015 - 2025 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 μgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

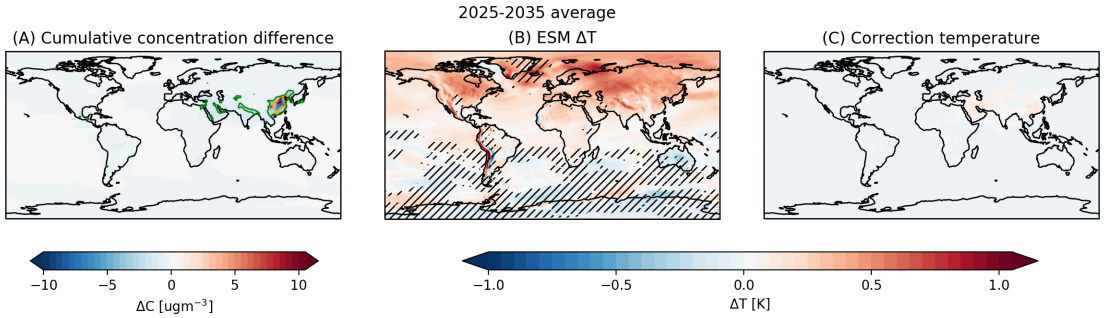

Figure C.5: 2025 -2035 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 μgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

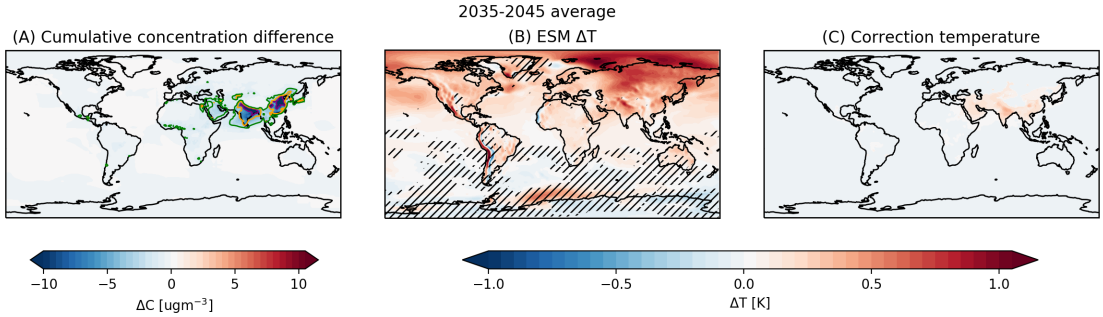

Figure C.6: 2035 -2045 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 μgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

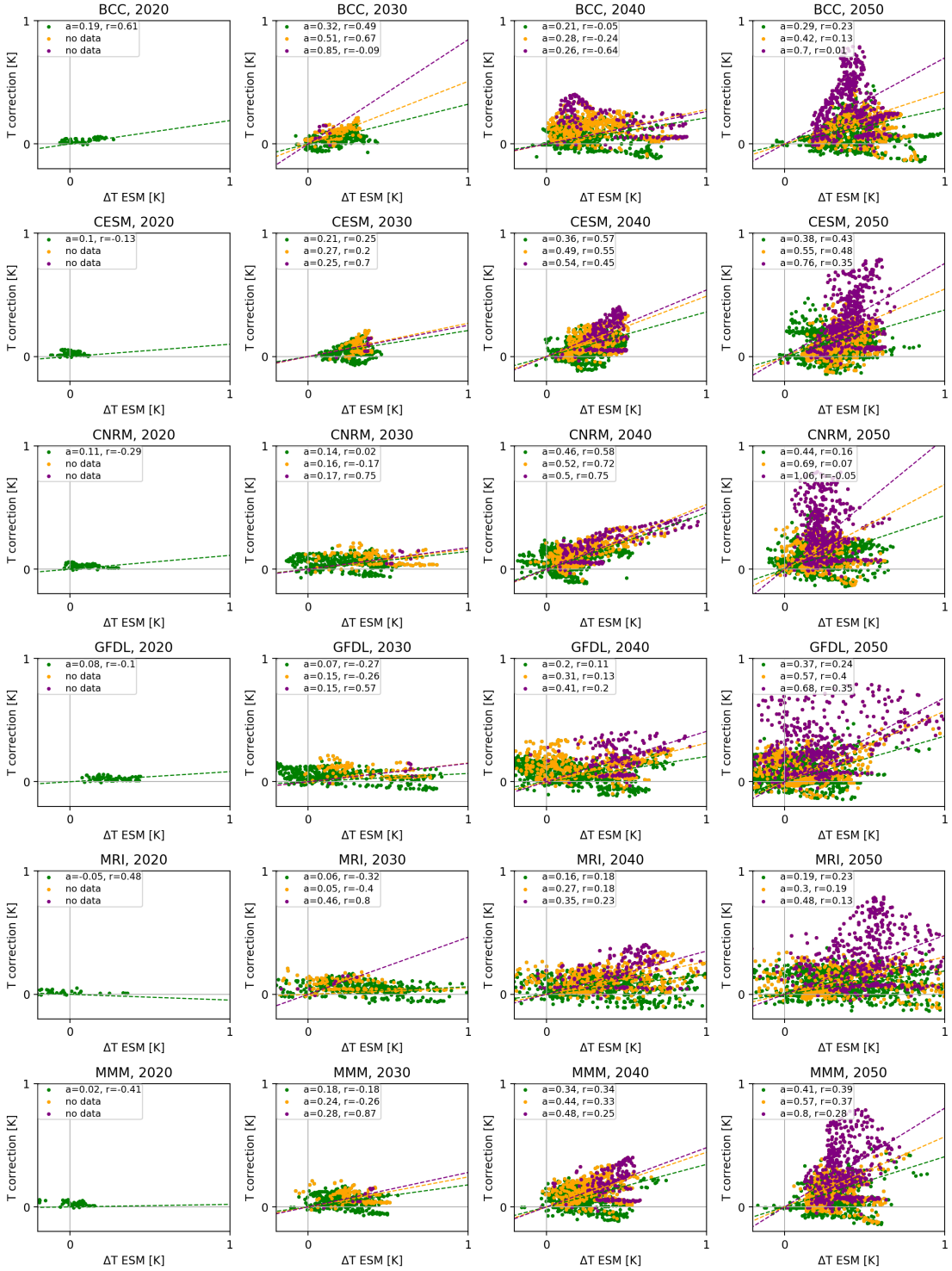

Figure C.7: Scatter plots for individual ESMs and MMM SSP3 lowNTCF-baseline temperature difference vs temperature correction on the -2, -5 and -10 μgm^{-3} (green, orange and purple) cumulative concentration change intervals. Dashed lines resemble best fit with intercept at 0 through the three cumulative concentration interval selected data sets, where a indicates the slope and r the correlation coefficient.

C.3 SSP3 SST results

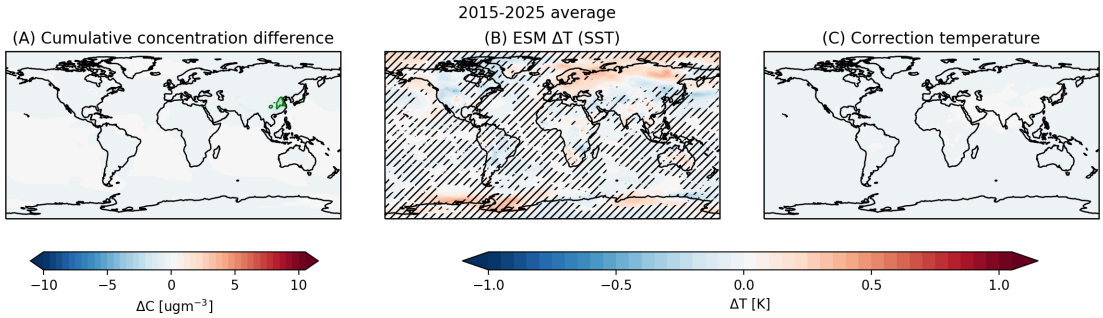

Figure C.8: 2035 -2045 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 μgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

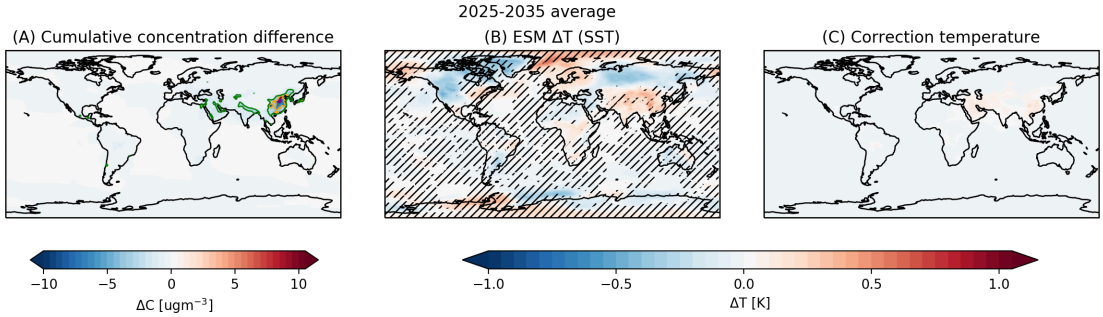

Figure C.9: 2035 -2045 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 μgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

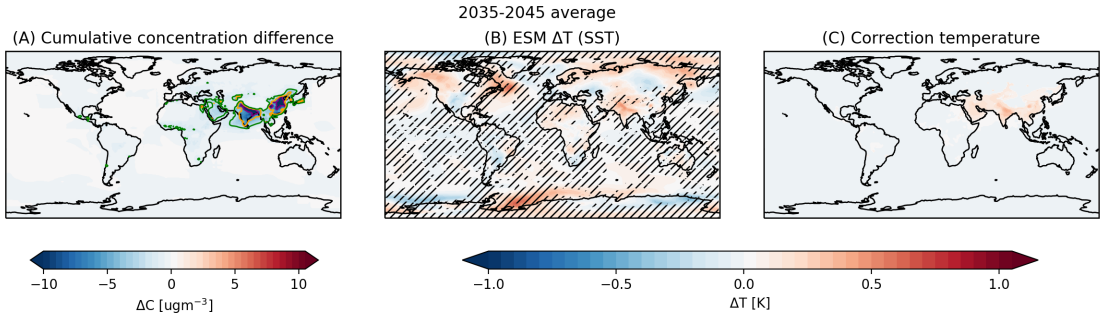

Figure C.10: 2035 -2045 average (A) Cumulative concentration difference, green, orange and purple contours indicate the -2, -5 and -10 μgm^{-3} (green, orange and purple) intervals; (B) MMM temperature difference between SSP3 lowNTCF and baseline, hatched areas indicate temperature sign agreement of less than four out of five models; (C) Temperature corrections based on the PiClim scaling method.

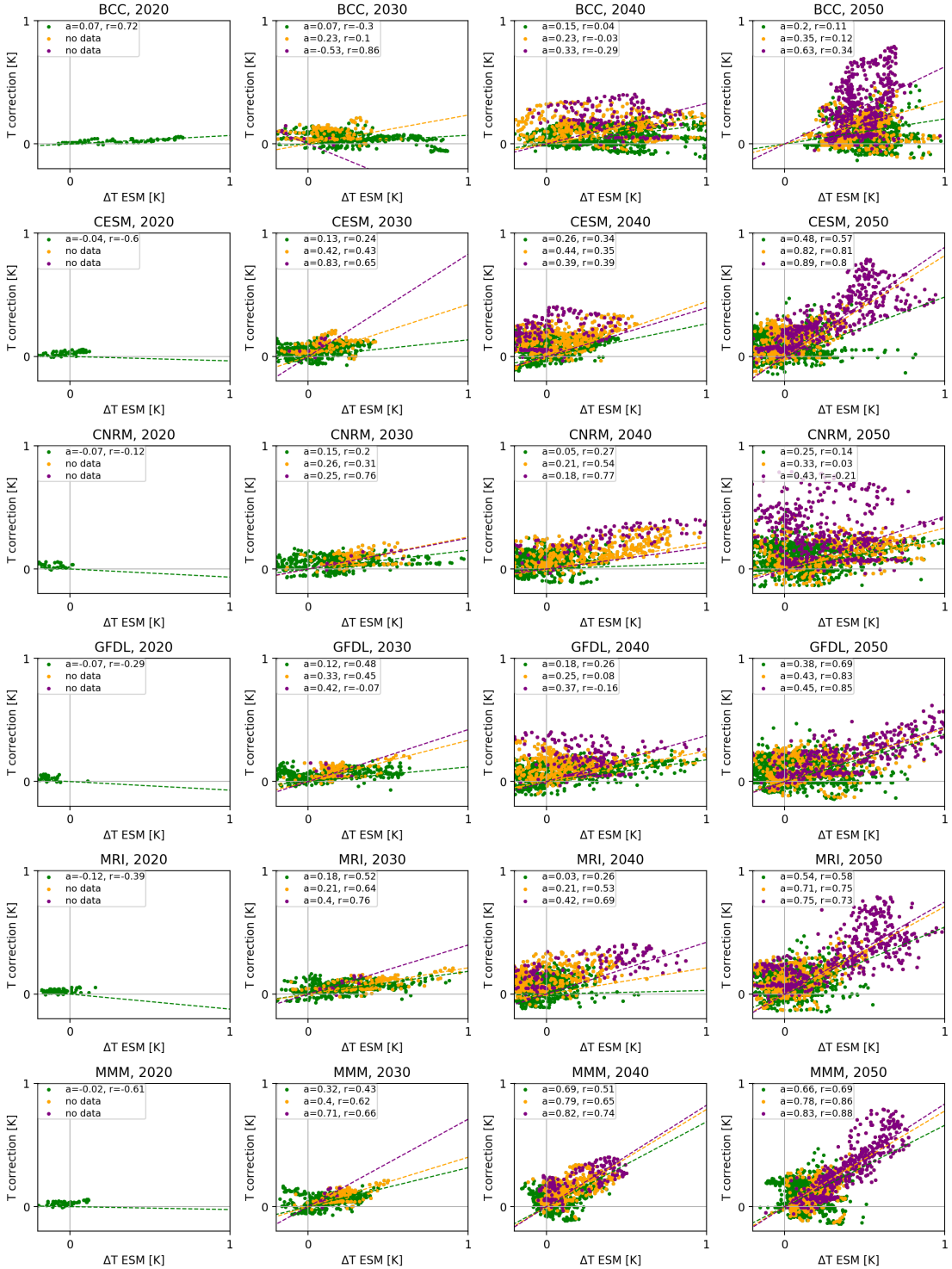

Figure C.11: Scatter plots for individual ESMs and MMM SSP3 SST lowNTCF-baseline temperature difference vs temperature correction on the -2, -5 and -10 µgm^{-3} (green, orange and purple) cumulative concentration change intervals. Dashed lines resemble best fit with intercept at 0 through the three interval selected data sets, where a indicates the slope and r the correlation coefficient.

D Land cover change validation results

D.1 Vegetation characteristics

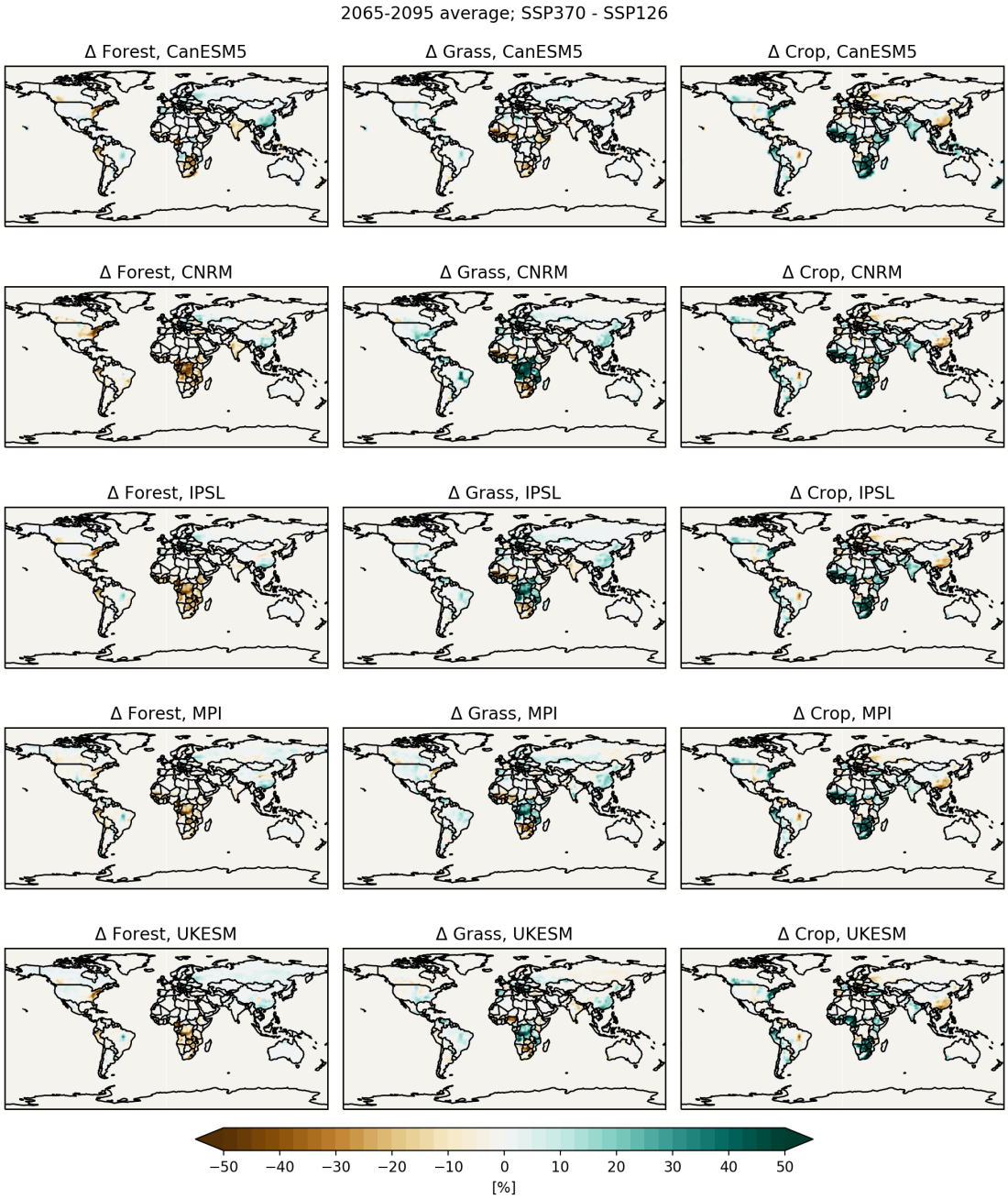

Figure D.1: 2065-2095 average SSP370-SSP126 vegetation cover fraction differences for all ESMs.

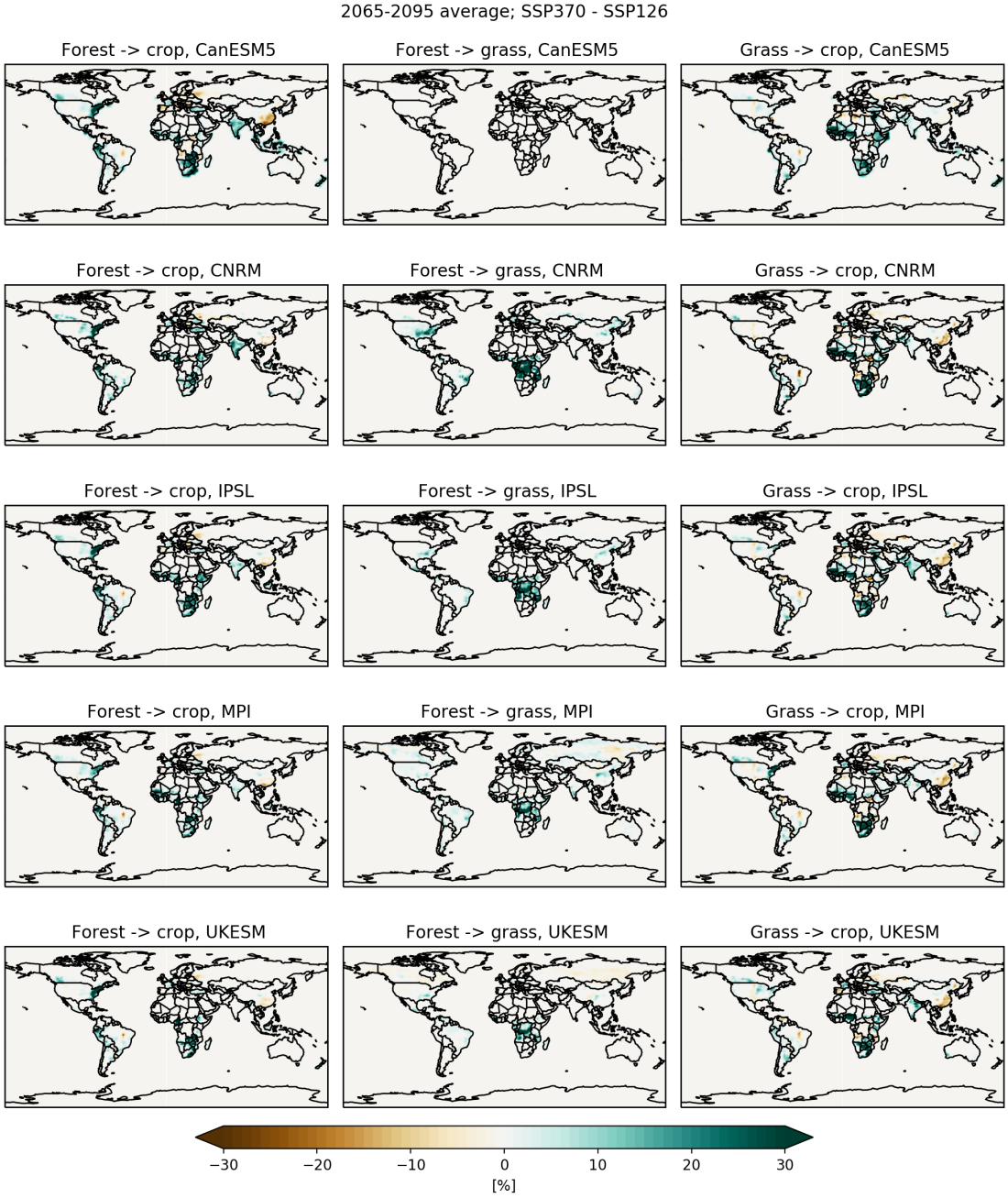

Figure D.2: 2065-2095 average SSP370-SSP126 vegetation transition fractions for all ESMs.

D.2 Filtering out background climate change

ESMs are subject to climate variability and feedback mechanisms, which will result in temperature changes in areas which are not marked by land use change. Moreover, adjacent areas will also influence each others temperature through various processes such as changes in cloud cover from changes in evapotranspiration. In order to overcome this problem, a moving window algorithm is applied to the ESM data to distinguish climate variability and other effects ΔT_{res} and the direct BPH effect ΔT_{BPH} from the total temperature change ΔT_{ESM}, similar to the method described in Alkama and Cescatti 2016:

$$\Delta T_{ESM} = \Delta T_{BPH} + \Delta T_{res} \longrightarrow \Delta T_{BPH} = \Delta T_{ESM} - \Delta T_{res} \tag{2}$$

A flowchart of the algorithm is shown in figure D.3 works as follows: For all cells, the absolute

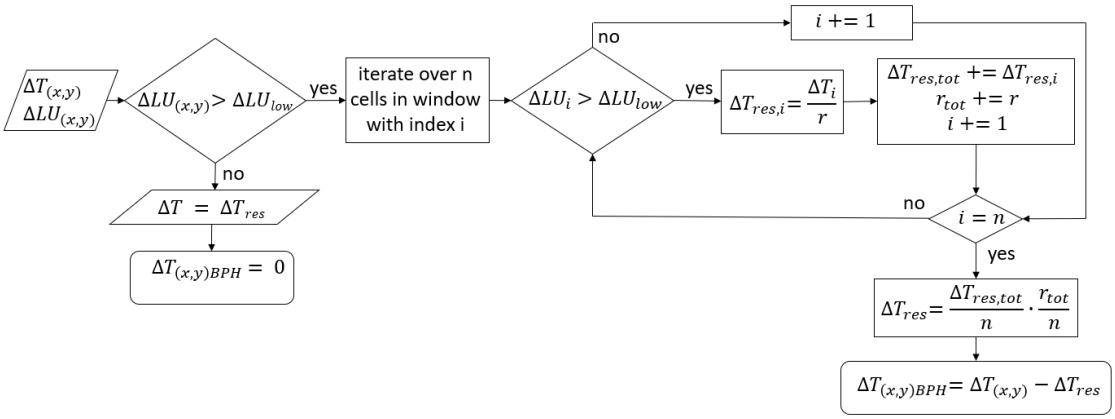

Figure D.3: Flowchart of the BPH effect filter

fractions of vegetation transitions are summed, to calculate the cumulative transition $\Delta LU_{(x,y)}$. Next an iteration is made over all cells with coordinates x,y. There are two possibilities:

1. If $\Delta LU_{(x,y)}$ in a cell is smaller than the set upper limit ΔLU_{high}, the temperature change is assumed not to be induced by LCLUC, and $\Delta T = \Delta T_{res} \longrightarrow \Delta T_{BPH} = 0$.

2. If $\Delta LU_{(x,y)}$ in a cell is larger than the set upper limit ΔLU_{high}, some part of the temperature change is induced by vegetation transition(s). Next, for a given radius R all neighbouring cells within a square window with sizes R*2+1 with the studied cell in the center are selected. Then, for every cell i in the window, ΔLU_i is compared to the set lower limit ΔLU_{low}. If ΔLU_i is smaller than ΔLU_{low}, the temperature change in cell i is assumed to be not induced by LU, and $\Delta T_{res,i}$ is weighted with the inverse of distance r to cell x,y. Cells where ΔLU_i is larger than ΔLU_{high} are omitted.

   Afterwards, all weighted $\Delta T_{res,i}$ and distances r are averaged, and the two averages are multiplied to calculate the residual temperature change. This value is subtracted from the temperature change in cell x,y, to calculate the local temperature change from BPH effects.

An example for a cell x,y is shown in figure D.4. For illustrative purposes, the chosen radius R is set to 2 and cumulative land use ΔL and temperature change ΔT is only shown for three cells within the window. The blue outlined cells indicate how the distance r is calculated, and the black outlined cell is the one for which ΔT_{BPH} is calculated. In this example, only the two red cells have a ΔLU that is smaller than ΔLU_{low}. For this example ΔT_{BPH} is calculated as follows:

$$\Delta T_{res,(x,y)} = \frac{\left(\frac{0.5}{\sqrt{1^2+1^2}} + \frac{1.2}{\sqrt{2^2+0^2}}\right)}{2} * \frac{\left(\sqrt{1^2+1^2} + \sqrt{2^2+0^2}\right)}{2} \approx 0.8 \longrightarrow \Delta T_{BPH} = 1.5 - 0.8 = 0.7K$$

In this example, of the 1.5K temperature change in cell x,y only 0.7K can be attributed to local BPH effects, while 0.8K is caused by larger scale climate variability or other effects.

Although numerous settings have been tested and analyzed, for all the results in this work a radius R of 10 cells and a threshold of 20% is applied.

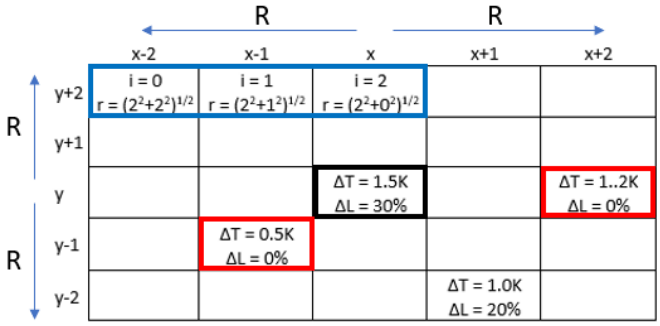

Figure D.4: ΔT_{BPH} calculation for an exemplary window

D.3 Full scatter plots for SSP126-SSP370LU and SSP370-SSP126LU

Figures D.5 and D.6 show the scatter plots for the ESM temperature difference and the modelled temperature difference based on Duveiller et al. 2018c, for both SSP126-SSP370LU with respect to SSP1-2.6 and SSP370-SSP126LU with respect to SSP370 for all 30 year averaged time steps. For both comparisons, the scatter plots show more data as time progresses. This is the result of the 20% threshold for filtering the BPH effect, which is passed more often as LC differences diverge.

As discussed before, CNRM and IPSL show a slight warming in the tropics and cooling in the boreal zone in figure D.5, while the opposite is true for figure D.6. Although the results of ESMs are inconsistent with each other and change throughout time, individual plots are thus (conversely) consistent between the two analysis. We can thus conclude that ESMs might not be consistent with the correction models or with each other, but the representation of physical vegetation effects are internally consistent. To prove this claim, the BPH filtered temperature effects for SSP126-SSP370LU with respect to SSP1-2.6 and SSP370-SSP126LU with respect to SSP370, and their sum is shown in figure D.7. Between the ESMs, patterns differ as they are forced with land use, and the resulting land cover can differ because of different model assumptions. In general, the patterns within ESMs are similar (but naturally opposite in sign) between the two comparisons. If the residuals (i.e. sum of the former two) is taken, then it can be observed for most ESMs these are relatively small. In other words, within one ESM the representation of BPH effects is consistent.

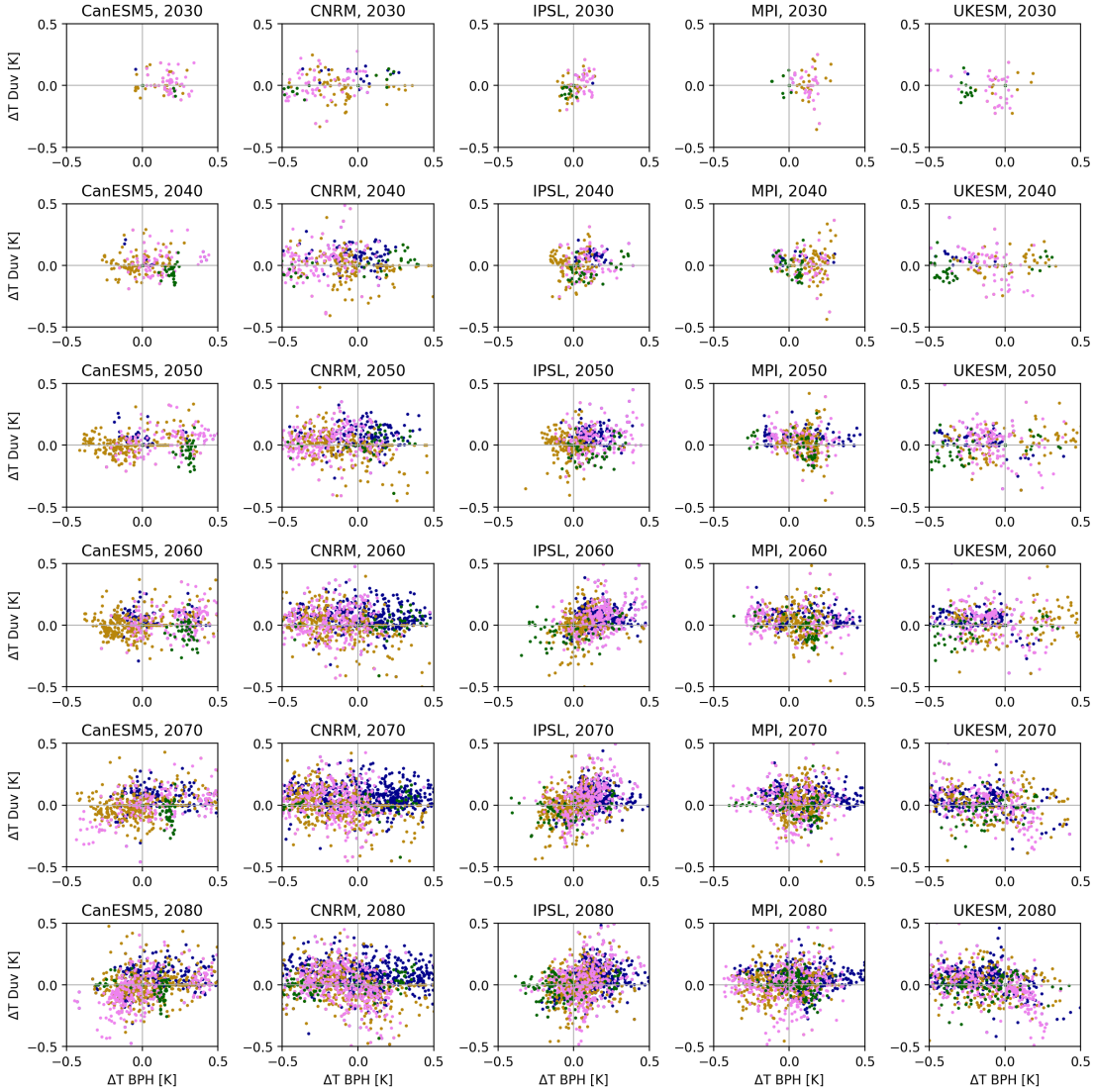

Figure D.5: Scatter plots of the modelled correction temperatures (ΔT Duv), vs the BPH filtered ESM temperatures, time stamps indicate the central year of the 30 year averages for SSP126-SSP370LU relative to SSP1-2.6. Blue, yellow, pink and green indicate data selection for the respective zones shown in figure 18

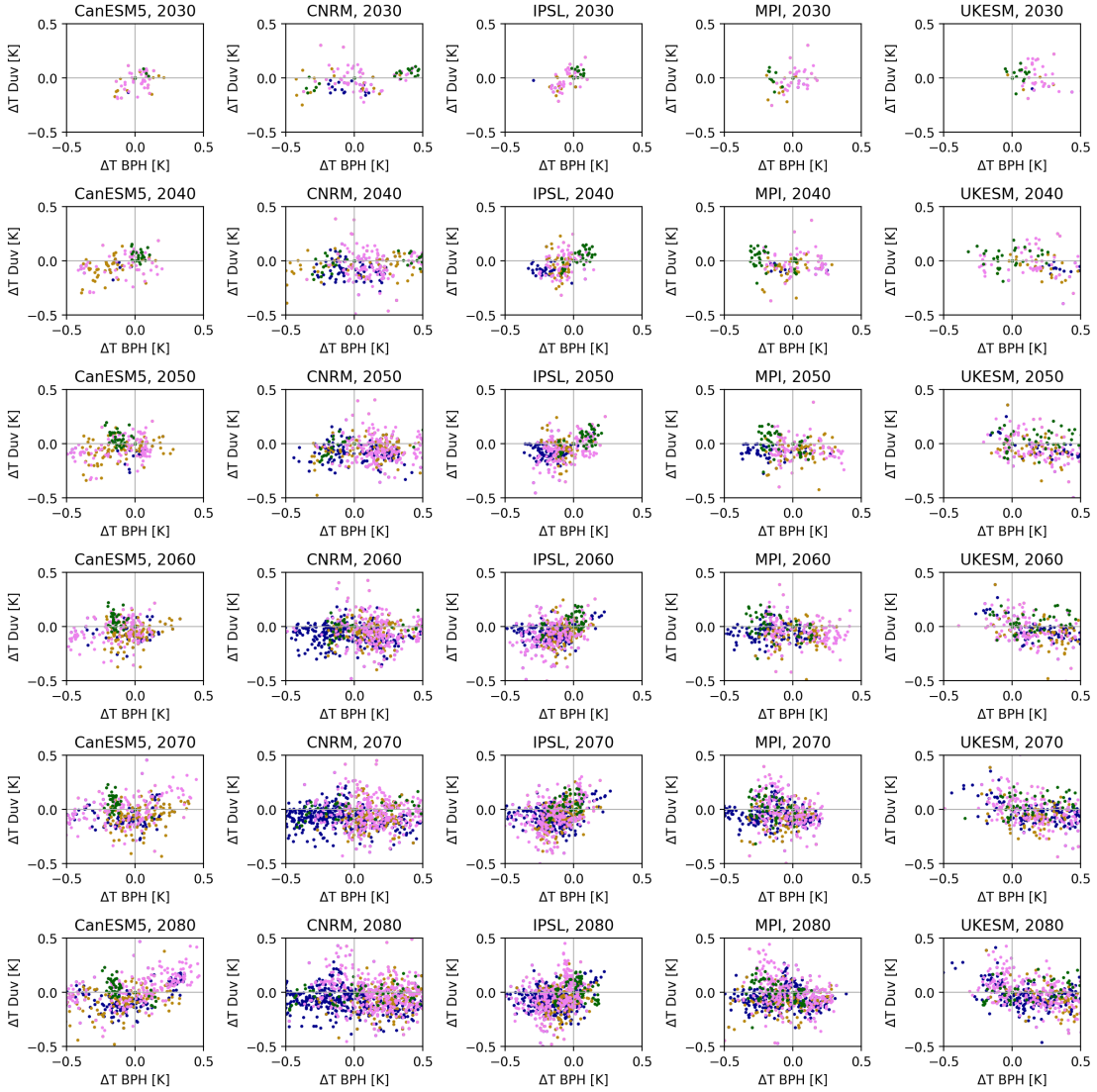

Figure D.6: Scatter plots of the modelled correction temperatures (ΔT Duv), vs the BPH filtered ESM temperatures, time stamps indicate the central year of the 30 year averages for SSP370-SSP126LU relative to SSP370. Blue, yellow, pink and green indicate data selection for the respective zones shown in figure 18

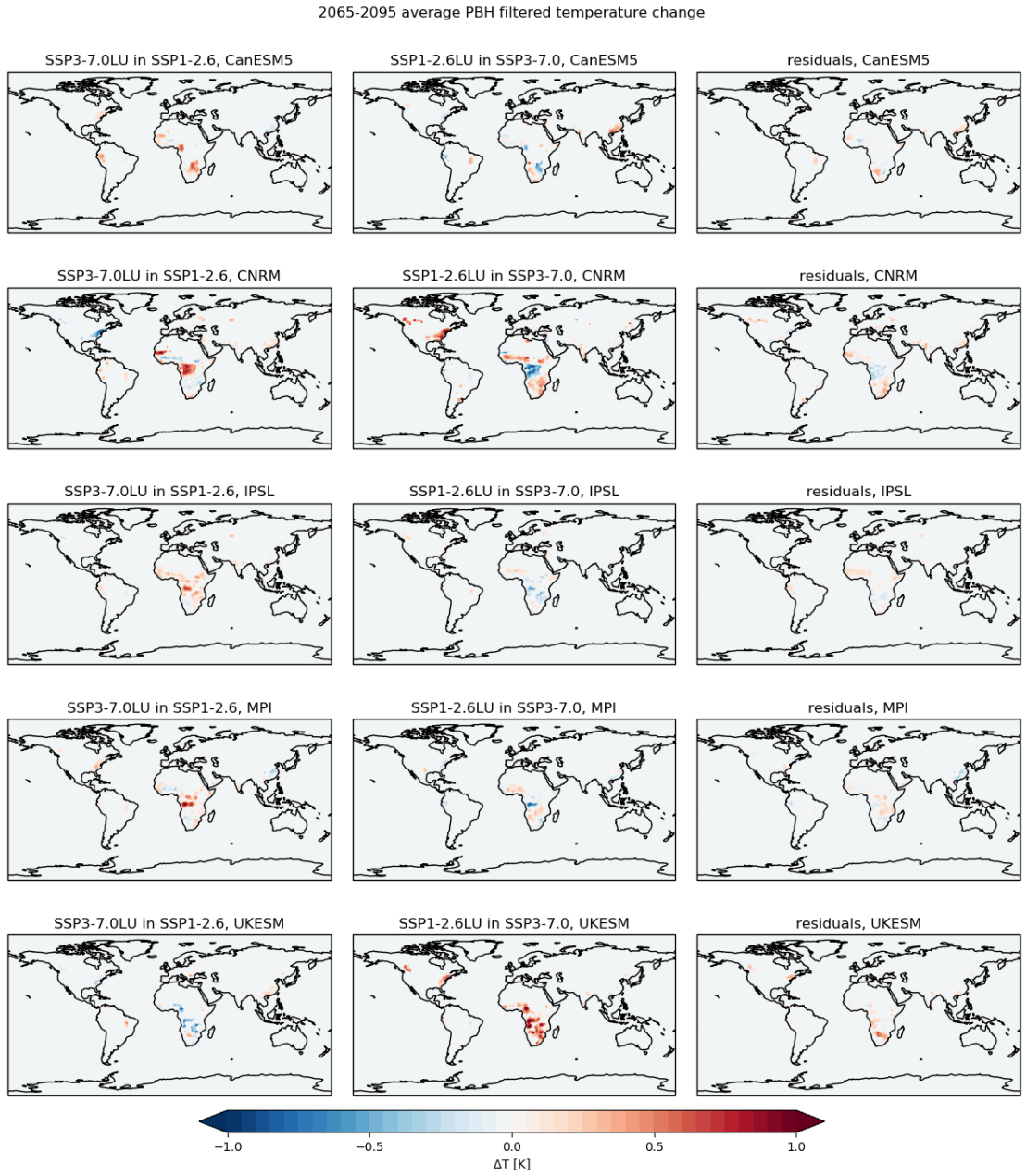

Figure D.7: BPH filtered temperature effect for the 2065-2095 average for individual ESMs: (left column) SSP126-SSP370LU relative to SSP1-2.6; (middle column) SSP370-SSP126LU relative to SSP370; (right column) residuals: sum of left and middle mappings.

E IMAGE emission harmonization for corrections

The original set of IMAGE and RCP emissions timelines is shown in figure E.1. Here can be observed that especially BC and OC emissions in IMAGE are roughly 1.5 times higher than historical and RCP at the start of the model run. These inconsistencies are caused by a fault in the IMAGE energy model, and not caused by the actual scenario characteristics. If we were to proceed with this data, corrections would be unrealistically high in the start of the century. To overcome this problem on a short time scale, the quickest solution is to harmonize the concentration fields.

Normally, harmonization would be performed through emissions, rather than concentrations. However, as the relations between concentration and emissions are linear in FASST, and harmonizing concentrations takes less time than harmonizing emissions (in this case), harmonization is achieved through simple concentration field scaling. In this process, SSP concentration fields are scaled so that the 2015 values match with those from RCP6.0 in 2015. Figure E.2 shows the resulting scaling factors, i.e., concentrations of RCP6.0 divided by the concentration of SSPs in 2015. For most of the areas which are used for corrections, the harmonization factor is roughly 0.7.

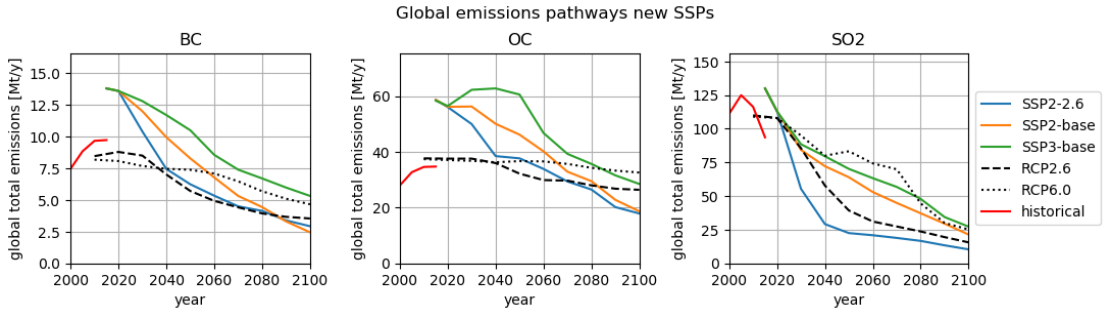

Figure E.1: Global harmonized BC, OC and SO_2 emission timelines for selected SSPs, RCP2.6 & RCP6.0 and historical (van Marle et al., 2017 & Hoesly et al., 2018) in Mt/y.

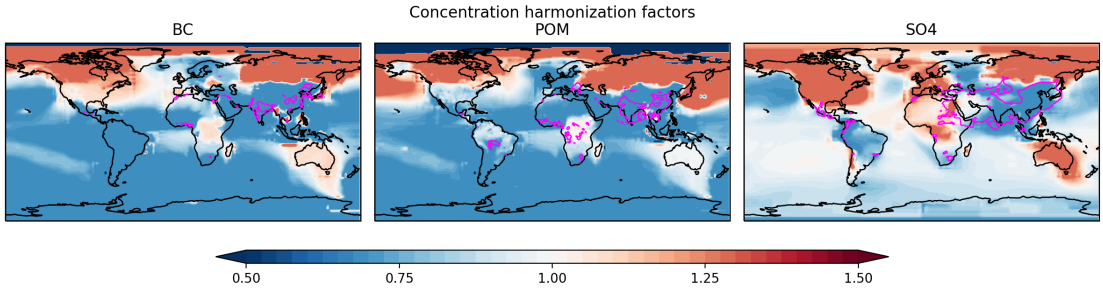

Figure E.2: Concentration harmonization factors BC, POM and SO_4 in Mt/y. Purple contour indicates the areas that are used for corrections.

F IMAGE temperature correction supplement

F.1 AQ

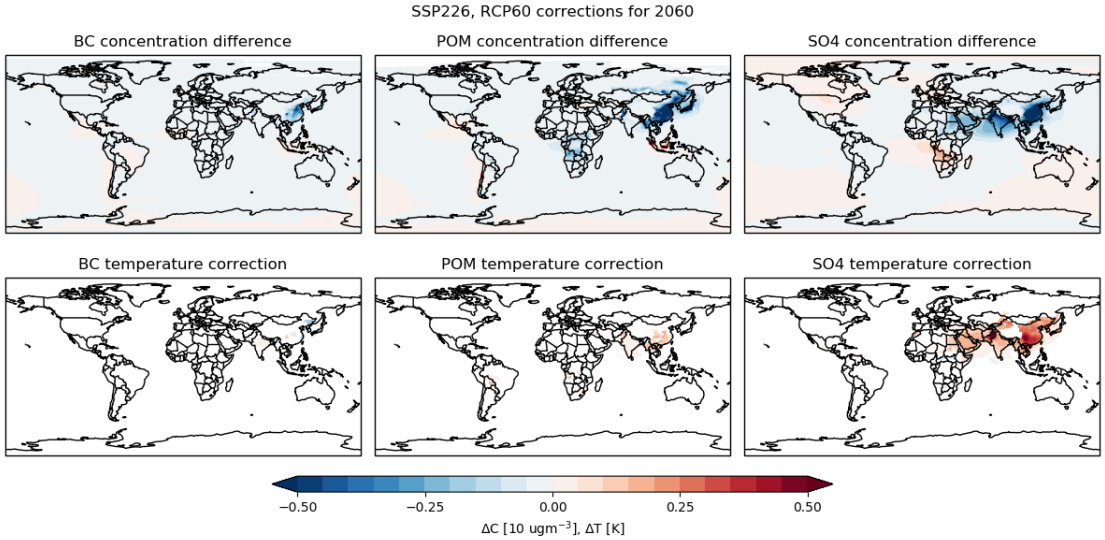

Figure F.1: Individual aerosol component differences and associated temperature corrections for SSP2-2.6 in 2060.

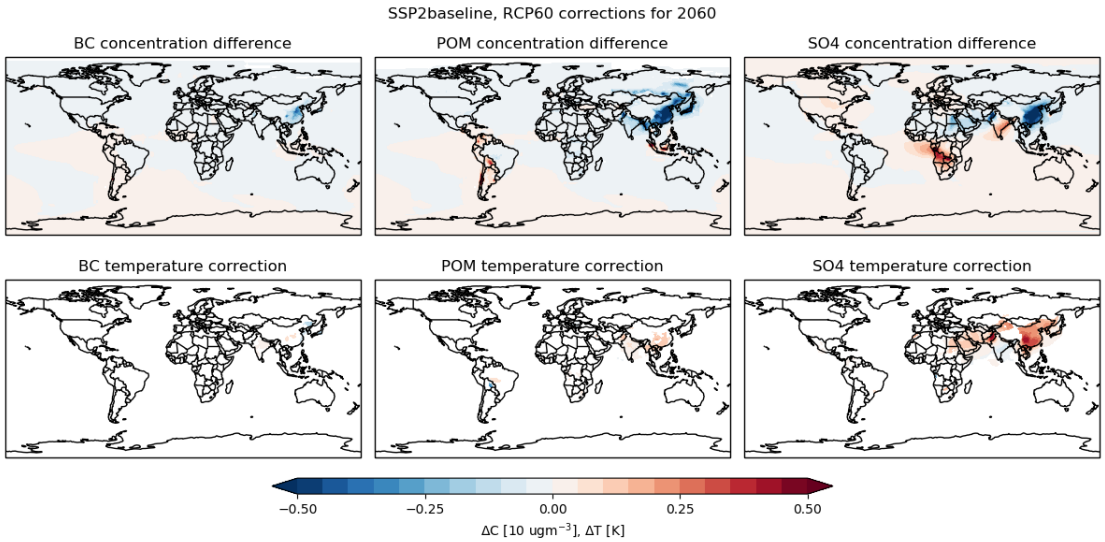

Figure F.2: Individual aerosol component differences and associated temperature corrections for SSP2-baseline in 2060.

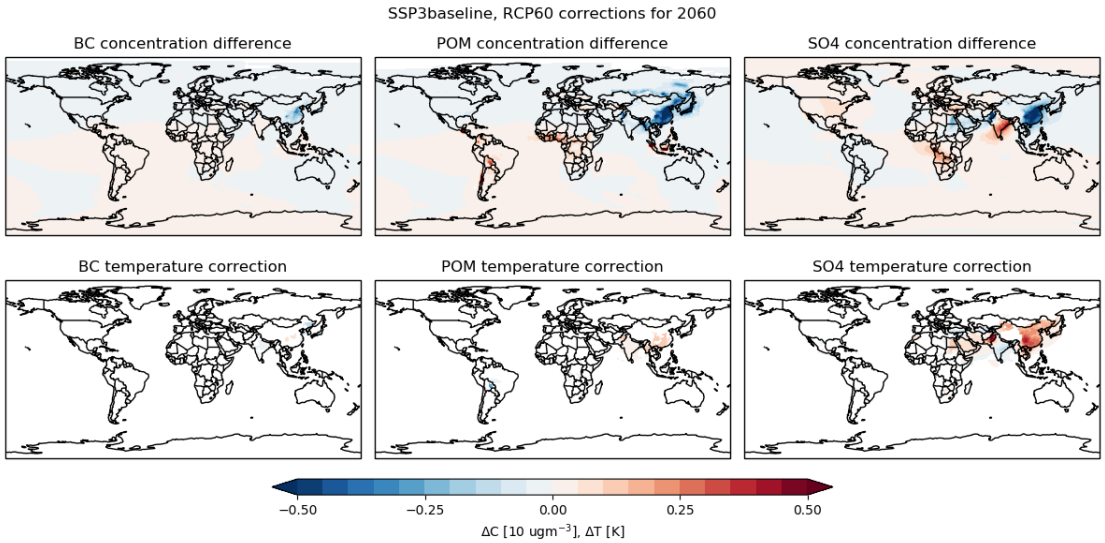

Figure F.3: Individual aerosol component differences and associated temperature corrections for SSP3-baseline in 2060.

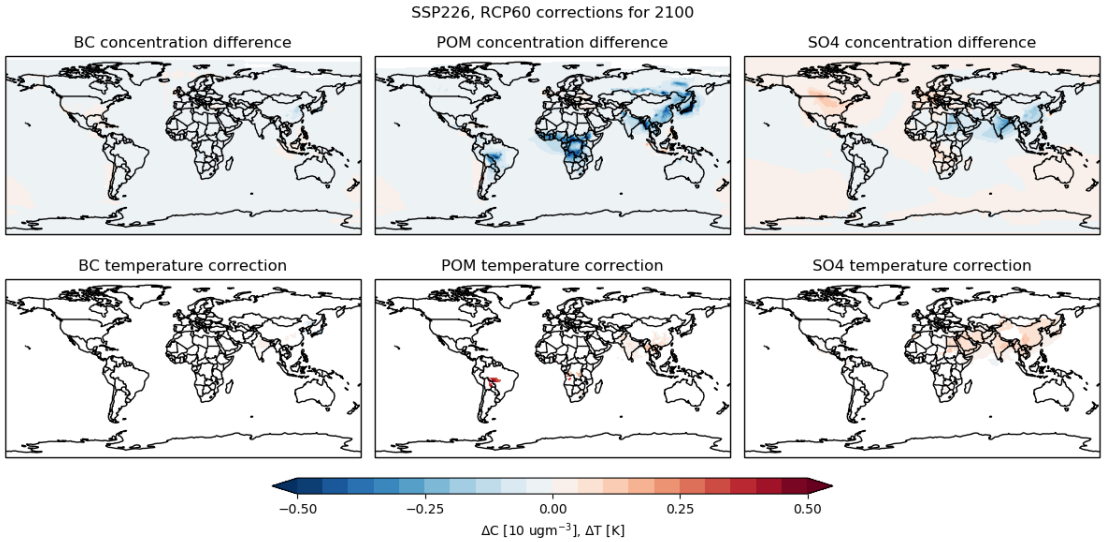

Figure F.4: Individual aerosol component differences and associated temperature corrections for SSP2-2.6 in 2100.

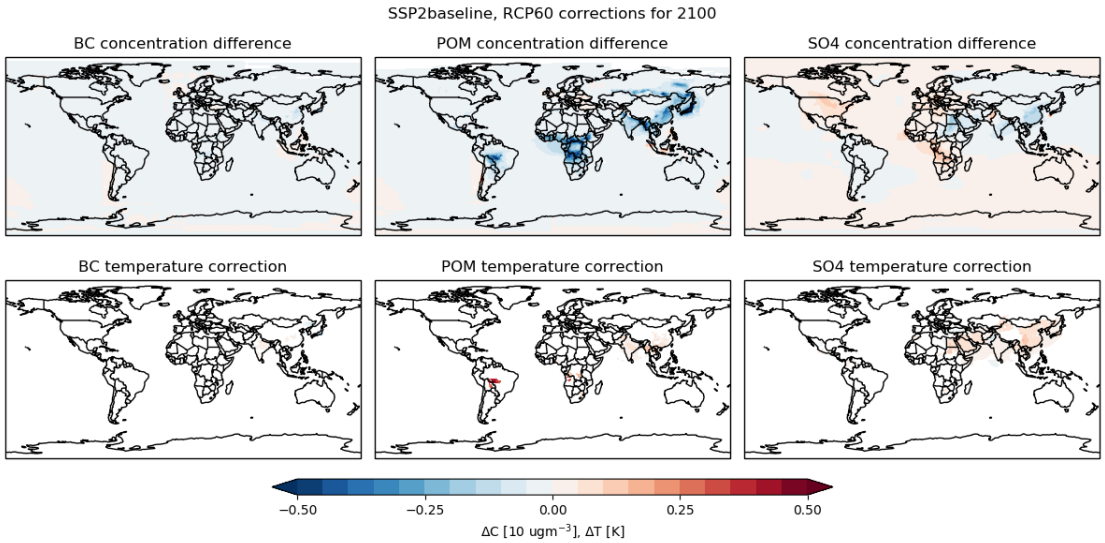

Figure F.5: Individual aerosol component differences and associated temperature corrections for SSP2-baseline in 2100.

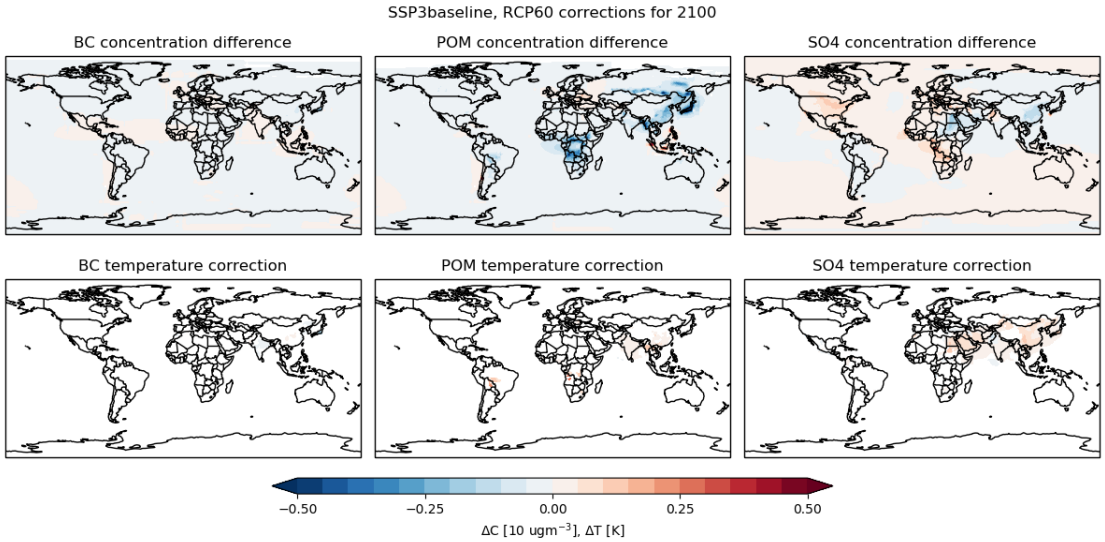

Figure F.6: Individual aerosol component differences and associated temperature corrections for SSP3-baseline in 2100.

F.2 LC

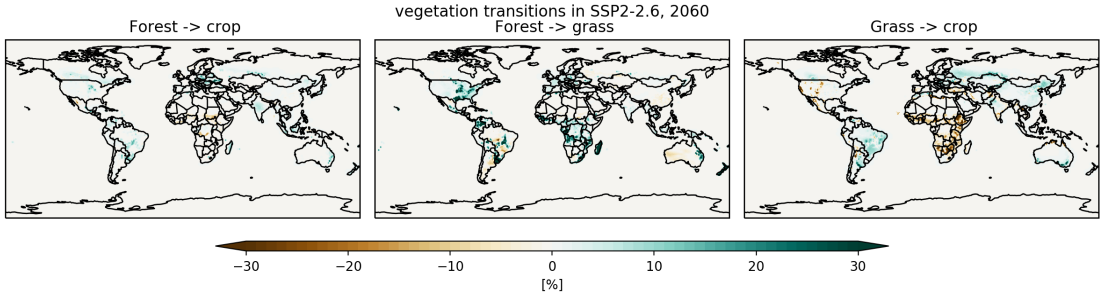

Figure F.7: Vegetation transitions for SSP2-2.6 in 2060.

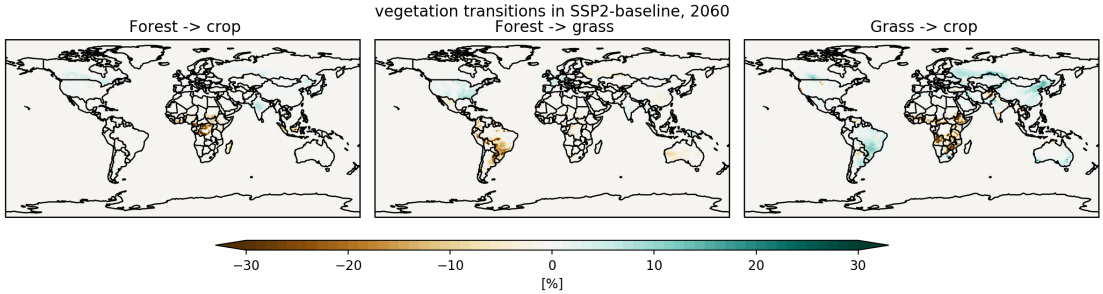

Figure F.8: Vegetation transitions for SSP2-baseline in 2060.

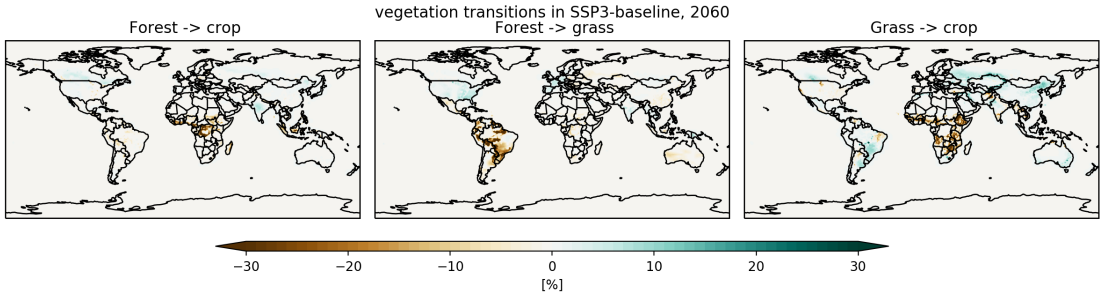

Figure F.9: Vegetation transitions for SSP3-baseline in 2060.

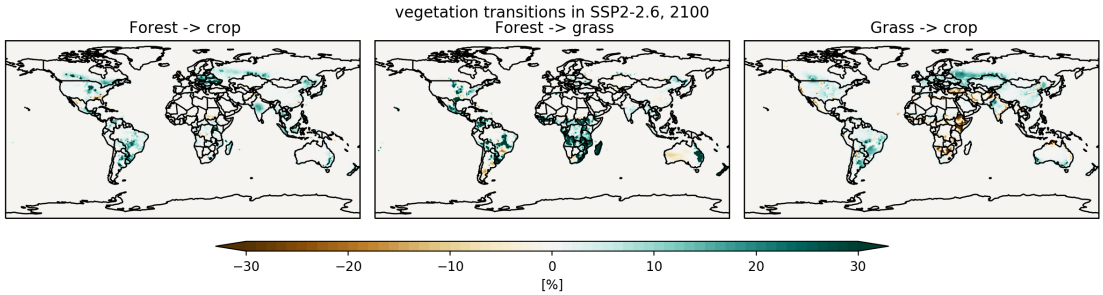

Figure F.10: Vegetation transitions for SSP2-2.6 in 2100.

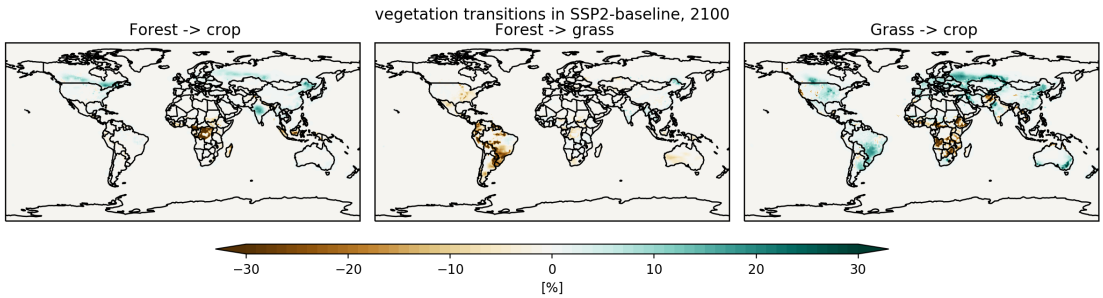

Figure F.11: Vegetation transitions for SSP2-baseline in 2100.

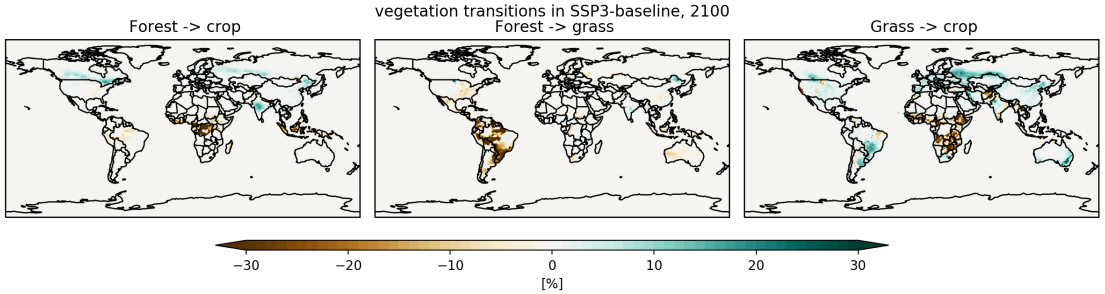

Figure F.12: Vegetation transitions for SSP3-baseline in 2100.

G IMAGE temperature correction for SSP2 baseline

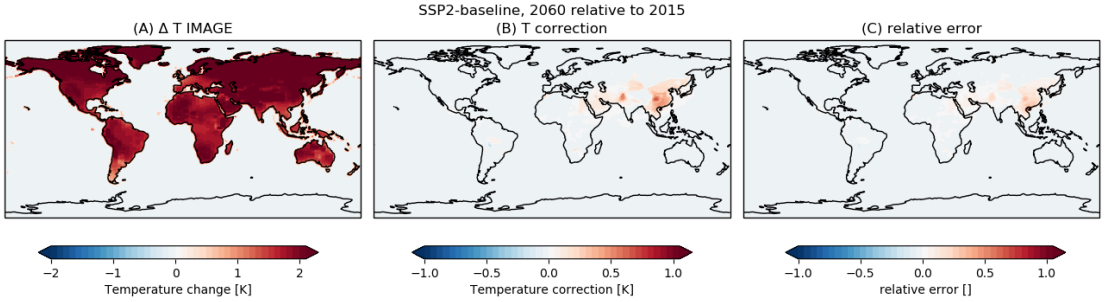

Figure G.1: 2060, SSP2-baseline (A) Temperature change in IMAGE relative to 2015; (B) AQ temperature corrections; (C) relative error: T correction / IMAGE ΔT

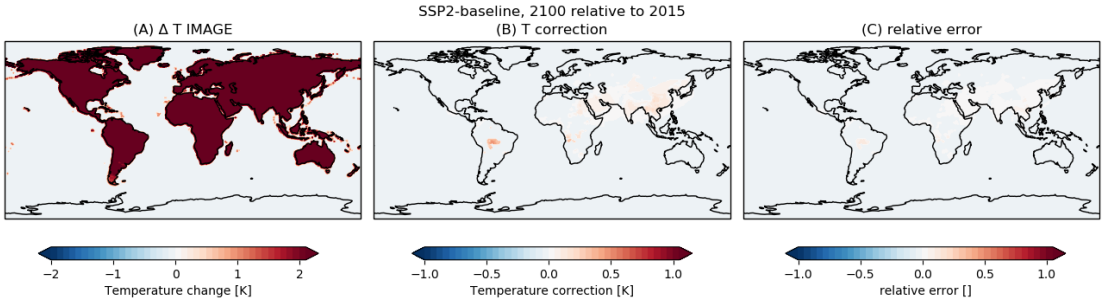

Figure G.2: 2100, SSP2-baseline (A) Temperature change in IMAGE relative to 2015; (B) AQ temperature corrections; (C) relative error: T correction / IMAGE ΔT

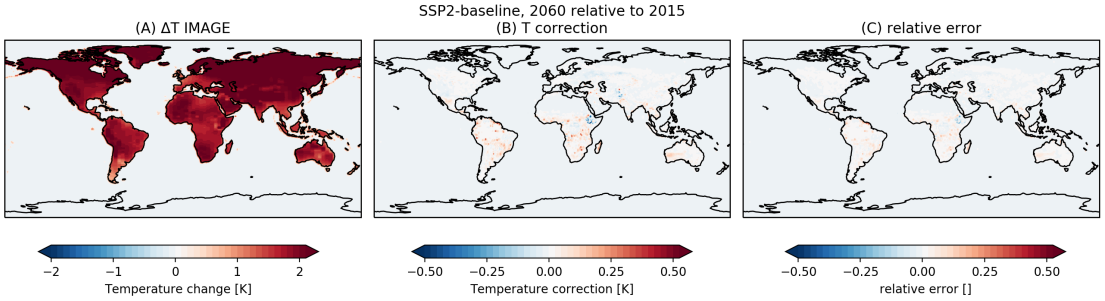

Figure G.3: 2060, SSP2-baseline (A) Temperature change in IMAGE relative to 2015; (B) LC temperature corrections; (C) relative error: T correction / IMAGE ΔT

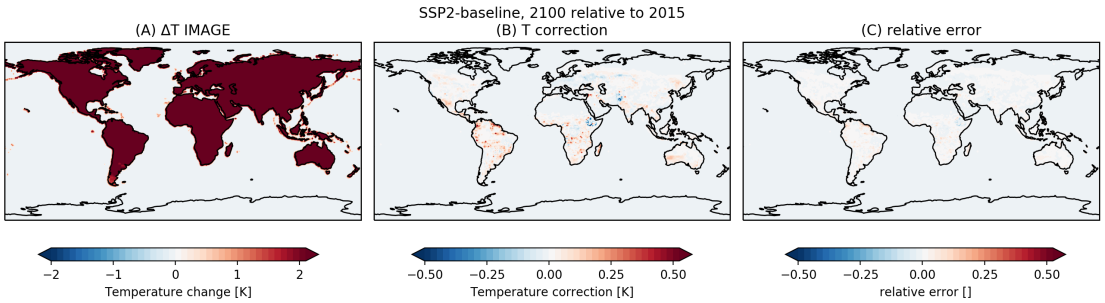

Figure G.4: 2100, SSP2-baseline (A) Temperature change in IMAGE relative to 2015; (B) LC temperature corrections; (C) relative error: T correction / IMAGE ΔT