Atmos. Chem. Phys., 11, 1417–1456, 2011
www.atmos-chem-phys.net/11/1417/2011/
doi:10.5194/acp-11-1417-2011

Atmospheric Chemistry and Physics

Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration

M. Meinshausen[1], **S. C. B. Raper**[2], **and T. M. L. Wigley**[3]

[1]Earth System Analysis, Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany
[2]Manchester Metropolitan University (MMU), Manchester, UK
[3]National Center for Atmospheric Research (NCAR), Boulder, CO, USA

Received: 20 December 2007 – Published in Atmos. Chem. Phys. Discuss.: 27 March 2008
Revised: 20 January 2011 – Accepted: 2 February 2011 – Published: 16 February 2011

Abstract. Current scientific knowledge on the future response of the climate system to human-induced perturbations is comprehensively captured by various model intercomparison efforts. In the preparation of the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC), intercomparisons were organized for atmosphere-ocean general circulation models (AOGCMs) and carbon cycle models, named "CMIP3" and "C^4MIP", respectively. Despite their tremendous value for the scientific community and policy makers alike, there are some difficulties in interpreting the results. For example, radiative forcings were not standardized across the various AOGCM integrations and carbon cycle runs, and, in some models, key forcings were omitted. Furthermore, the AOGCM analysis of plausible emissions pathways was restricted to only three SRES scenarios. This study attempts to address these issues. We present an updated version of MAGICC, the simple carbon cycle-climate model used in past IPCC Assessment Reports with enhanced representation of time-varying climate sensitivities, carbon cycle feedbacks, aerosol forcings and ocean heat uptake characteristics. This new version, MAGICC6, is successfully calibrated against the higher complexity AOGCMs and carbon cycle models. Parameterizations of MAGICC6 are provided. The mean of the emulations presented here using MAGICC6 deviates from the mean AOGCM responses by only 2.2% on average for the SRES scenarios. This enhanced emulation skill in comparison to previous calibrations is primarily due to: making a "like-with-like comparison" using AOGCM-specific subsets of forcings; employing a new calibration procedure; as well as the fact that the updated simple climate model can now successfully emulate some of the climate-state dependent effective climate sensitivities of AOGCMs. The diagnosed effective climate sensitivity at the time of CO_2 doubling for the AOGCMs is on average 2.88 °C, about 0.33 °C cooler than the mean of the reported slab ocean climate sensitivities. In the companion paper (Part 2) of this study, we examine the combined climate system and carbon cycle emulations for the complete range of IPCC SRES emissions scenarios and the new RCP pathways.

Correspondence to: M. Meinshausen
(malte.meinshausen@pik-potsdam.de)

1 Introduction

This study presents the most comprehensive AOGCM and carbon cycle model emulation exercise to date. We use an updated version of the MAGICC model, which was originally developed by Wigley and Raper (1987, 1992) and which has been updated continuously since then (see e.g. Raper et al., 1996; Wigley and Raper, 2001; Wigley et al., 2009). Several amendments to MAGICC have been spurred by new results presented in the IPCC AR4 as well as by the increased availability of comprehensive AOGCM and carbon cycle model datasets. For example, land/ocean temperature evolutions for both hemispheres were calculated for each AOGCM allowing for a more in-depth analysis of optimal heat exchange parameterizations in MAGICC. Emulations with a simple model like MAGICC6 can by no means replace research into more sophisticated carbon cycle and general

circulation models. Rather, what MAGICC6 offers primarily is a method to extend the knowledge created with AOGCMs and carbon cycle model runs in order to provide estimates of their joint responses and to extrapolate their key characteristics to a range of other scenarios.

The paper is structured as follows: First, the value and advantages of simple climate models are discussed in Sect. 2. Sect. 3 provides a brief overview of the main amendments in the climate model MAGICC6 as used here – compared to the version used in IPCC AR4. The emulation of AOGCMs is described in Sect. 4, while the emulation of the C^4MIP carbon cycle models is described in Sect. 5. Section 6 summarizes limitations of the present approach, while conclusions are given in Sect. 7. A complete description of the MAGICC6 model can be found in the Appendix A.

2 The value of simple climate models

Since the introduction of three-dimensional coupled atmosphere-ocean general circulation models (AOGCMs) (e.g. Manabe and Bryan, 1969; Manabe et al., 1975; Bryan et al., 1975; Schlesinger et al., 1985), one goal of Earth system science is to facilitate the understanding of the past and the projected future climate by building highly resolved, comprehensive models of the physical atmosphere and ocean systems, including the Earth's cryosphere, and the terrestrial and marine biosphere. Intermediate complexity or simpler models are complementary research tools that can provide focus on individual processes, span the range of parameter uncertainties with computational efficiency and extend results for multiple scenarios. After the introduction of the one-dimensional upwelling-diffusive ocean model by Hoffert et al. (1980), early applications of simple models were able to give new insights into the transient behavior of the climate system through investigation of individual feedback processes, multi-thousand year simulations, and parameter sensitivity studies (Harvey and Schneider, 1985a,b; Senior and Mitchell, 2000; Hoffert et al., 1980). Recently, this role has also been filled by intermediate complexity models (Earth System Models of Intermediate Complexity or EMICs). Shifting from their role as models in their own right, simple models started to serve four distinct purposes as exemplified in this study:

I. Emulations. Simple models may be used to emulate AOGCMs and reproduce the global or large-scale averaged results of such models (see e.g. Schlesinger and Jiang, 1990, 1991). In most cases, AOGCMs are still computationally too expensive to be able to run large ensembles (as required e.g. for probabilistic studies), simulations for large sets of emissions scenarios, and/or multiple perturbed physics experiments except in special circumstances (Allen, 1999; Stainforth et al., 2005). In the emulation of AOGCMs with simple models, a necessary condition for model credibility is that the emulation of the variables of interest is suitably accurate over a wide range of emissions or concentration scenarios actually performed with AOGCMs. Various authors (e.g. Kattenberg et al., 1996; Raper and Cubasch, 1996; Raper et al., 2001; Cubasch et al., 2001; Osborn et al., 2006) have shown, for example, that the upwelling-diffusion model MAGICC, the primary simple climate model used in past IPCC Assessment Reports, can closely match key large-scale AOGCM results over a wide range of scenarios.

II. Parametrization of structural uncertainties. One advantage of simple models is that they can be used to span structural uncertainties across more complex models. Structural uncertainties in AOGCMs arise from the way certain processes or components (such as clouds) are "parameterized" or expressed in relatively simple terms – these parameterizations are structural components of the model. Within these parameterizations there may be a number of parameters, and parametric uncertainties arise from the uncertain values of these parameters. Thus, two models can differ in their aggregated response characteristics because they have different structures (including aspects commonly referred as "parameterizations"), or because, within common structures, they use different parameter values. This distinguishes between structural and parametric sources of uncertainty. In fact, we take advantage of this in the present study by "parameterizing" the structural uncertainty range of more complex models (cf. O'Neill and Melnikov, 2008) by estimating the parametric values within the more flexible MAGICC structure that fits the AOGCM results. This approach is distinct from perturbed physics studies with intermediate complexity models or AOGCMs (Murphy et al., 2004), which often concentrate on assessing parametric uncertainties within a fixed and comparatively more rigid model structure.

III. Factor separation analysis. Simple models can assist in factor separation analysis, i.e., in separating the effects of climate or carbon cycle uncertainties from forcing uncertainties, or in investigating the effects due to different initialization choices. Thus, simple models can assist in harmonizing the results from AOGCMs and other higher complexity models by estimating their responses for unified forcing assumptions, thus making the results from different models more directly comparable. For example, a major difficulty in interpreting the multi-model AOGCM projections presented in IPCC AR4 arises from the different radiative forcings considered by the various modeling groups (see Table 10.1 in Meehl et al., 2007; Knutti et al., 2008). A major difference is in the treatment of aerosol forcing, where, for example, some models included indirect aerosol forcing

while others did not. Also, for a single forcing agent the magnitude and time-evolution of climate change differs from model to model for the same scenario, because some models applied only very weak volcanic forcing in the 20th century runs while others ignored volcanic forcing completely. Some models varied tropospheric ozone while others keep the forcing by tropospheric ozone constant for the 21st century.

Further complications arise because, for most AOGCMs, the forcing time-series are not diagnosed or documented for the model runs – exceptions are Takemura et al. (2006) and Hansen et al. (2005). Different reporting standards for radiative forcing, like reporting adjusted forcing after thermal stratospheric adjustment at the model's tropopause or at the 200 hPa level, further hinder comparability, even when some diagnostic data are provided: specifically the CO_2 forcing at the time of doubled CO_2 concentration (see e.g. Table 2 in Forster and Taylor, 2006, hereafter called F&T). In addition, studies comparing the diagnosed results from the radiative transfer schemes in AOGCMs with those from the line-by-line code found surprisingly large differences, even for well known forcing agents like CO_2 (Collins et al., 2006). In summary, imperfect knowledge with regard to the forcings in CMIP3 AOGCM experiments leads to ambiguities as to how much of the differences in their temperature projections are due to different climate responses (feedbacks, inertia, etc.) or simply an expression of different (sometimes limited or erroneous) radiative forcing implementations.

IV. Joint response and feedback analysis. Simple, but sufficiently comprehensive, models allow one to estimate the joint responses of multiple models of higher complexity. For example, for comparison purposes in the IPCC AR4, the CMIP3 AOGCMs were driven with externally calculated CO_2 concentrations and in most cases the same CO_2 concentrations were prescribed irrespective of the AOGCM climate sensitivity. However, because of climate feedbacks on the carbon cycle, a higher sensitivity AOGCM is likely to see higher concentrations under the same emissions scenario, leading to an elevated temperature response. In its coupled mode, MAGICC is internally consistent in its CO_2 concentrations because the climate feedbacks on the carbon cycle are driven by the climate model response. We can calibrate to uncoupled model component results separately and anticipate the joint response of all combinations of state-of-the-art high complexity carbon cycle models and AOGCMs in a consistent framework.

3 Model description

MAGICC has a hemispherically averaged upwelling-diffusion ocean coupled to an atmosphere layer and a globally averaged carbon cycle model. As with most other simple models, MAGICC evolved from a simple global average energy-balance equation. The energy balance equation for the perturbed climate system can be written as:

$$\Delta Q_G = \lambda_G \Delta T_G + \frac{dH}{dt} \quad (1)$$

where ΔQ_G is the global-mean radiative forcing at the top of the troposphere. This extra energy influx is partitioned into increased outgoing energy flux and heat content changes in the ocean $\frac{dH}{dt}$. The outgoing energy flux is dependent on the global-mean feedback factor, λ_G, and the surface temperature perturbation ΔT_G.

While MAGICC is designed to provide maximum flexibility in order to match different types of responses seen in more sophisticated models, the approach in MAGICC's model development has always been to derive the simple equations as much as possible from key physical and biological processes. In other words, MAGICC is as simple as possible, but as mechanistic as necessary. This process-based approach has a strong conceptual advantage in comparison to simple statistical fits that are more likely to quickly degrade in their skill when emulating scenarios outside the original calibration space of sophisticated models.

The main improvements in MAGICC6 compared to the version used in the IPCC AR4 are briefly highlighted in this section (Note that there is an intermediate version, MAGICC 5.3, described in Wigley et al., 2009). The options introduced to account for variable climate sensitivities are described in Sect. 3.1. With the exception of the updated carbon cycle routines (Sect. 3.2), the MAGICC 4.2 and 5.3 parameterizations are covered as special cases of the 6.0 version, i.e., the IPCC AR4 version, for example, can be recovered by appropriate parameter settings.

3.1 Introduction of variable climate sensitivities

Climate sensitivity ($\Delta T_{2\times}$) is a useful metric to compare models and is usually defined as the equilibrium global-mean warming after a doubling of CO_2 concentrations. In the case of MAGICC, the equilibrium climate sensitivity is a primary model parameter that may be identified with the eventual global-mean warming that would occur if the CO_2 concentrations were doubled from pre-industrial levels. Climate sensitivity is inversely related to the feedback factor λ:

$$\Delta T_{2\times} = \frac{\Delta Q_{2\times}}{\lambda} \quad (2)$$

where $\Delta T_{2\times}$ is the climate sensitivity, and $\Delta Q_{2\times}$ the radiative forcing after a doubling of CO_2 concentrations (see energy balance Eq. A45).

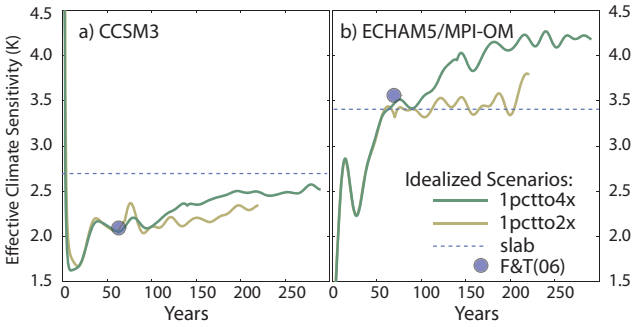

Fig. 1. The effective climate sensitivity diagnosed from low-pass filtered CCSM3 **(a)** and ECHAM5/MPI-OM **(b)** output for two idealized scenarios assuming an annual 1% increase in CO_2 concentrations until twice pre-industrial values in year 70 (1pctto2×) or quadrupled concentration in year 140 (1pctto4×), with constant concentrations thereafter. Additionally, the reported slab ocean model equilibrium climate sensitivity ("slab") and the sensitivity estimates by Forster and Taylor (2006) are shown ("F&T(06)").

The (time- or state-dependent) effective climate sensitivity (S^t) (Murphy and Mitchell, 1995) is defined using the transient energy balance Eq. (1) and can be diagnosed from model output for any part of a model run where radiative forcing and ocean heat uptake are both known and their sum is different from zero, so that:

$$S^t = \frac{\Delta Q_{2\times}}{\lambda^t} = \Delta Q_{2\times} \frac{\Delta T^t_{\mathrm{G}}}{\Delta Q^t - \frac{dH}{dt}|^t} \tag{3}$$

where $\Delta Q_{2\times}$ is the model-specific forcing for doubled CO_2 concentration, λ_t is the time-variable feedback factor, ΔQ^t the radiative forcing, ΔT^t_{GL} the global-mean temperature perturbation and $\frac{dH}{dt}|^t$ the climate system's heat uptake at time t. By definition, the traditional (equilibrium) climate sensitivity ($\Delta T_{2\times}$) is equal to the effective climate sensitivity S^t at equilibrium ($\frac{dH}{dt}|^t$=0) after doubled (pre-industrial) CO_2 concentration.

If there were only one globally homogenous, fast and constant feedback process, the diagnosed effective climate sensitivity would always equal the equilibrium climate sensitivity $\Delta T_{2\times}$. However, many CMIP3 AOGCMs exhibit variable effective climate sensitivities, often increasing over time (e.g. models CCSM3, CNRM-CM3, GFDL-CM2.0, GFDL-CM2.1, GISS-EH – see Figs. B1, B2, and B3). This is consistent with earlier results of increasing effective sensitivities found by Senior and Mitchell (2000); Raper et al. (2001) for the HadCM2 model. In addition, some models present significantly higher sensitivities for higher forcing scenarios (1pctto4×) than for lower forcing scenarios (1pctto2×) (e.g. ECHAM5/MPI-OM and GISS-ER, see Fig. 1).

In order to better emulate these time-variable effective climate sensitivities, this version of MAGICC incorporates two modifications: Firstly, an amended land-ocean heat exchange formulation allows effective climate sensitivities to increase on the path to equilibrium warming. In this formulation, changes in effective climate sensitivity arise from a geometrical effect: spatially non-homogenous feedbacks can lead to a time-variable effective global-mean climate sensitivity, if the spatial warming distributions change over time. Hence, by modifying land-ocean heat exchange in MAGICC, the spatial evolution of warming is altered, leading to changes in effective climate sensitivities (Raper, 2004) given that MAGICC has different equilibrium sensitivities over land and ocean. Secondly, the climate sensitivities, and hence the feedback parameters, can be made explicitly dependent on the current forcing at time t. Both amendments are detailed in the Appendix A (see Sects. A4.2 and A4.3). Although these two amendments both modify the same diagnostic, i.e., the time-variable effective sensitivities in MAGICC, they are distinct: the land-ocean heat exchange modification changes the shape of the effective climate sensitivity's time evolution to equilibrium, but keeps the equilibrium sensitivity unaffected. In contrast, making the sensitivity explicitly dependent on the forcing primarily affects the equilibrium sensitivity value.

Note that time-varying effective sensitivities are not only empirically observed in AOGCMs, but they are necessary here in order for MAGICC to accurately emulate AOGCM results. Alternative parameterizations to emulate time-variable climate sensitivities are possible, e.g. assuming a dependence on temperatures instead of forcing, or by implementing indirect radiative forcing effects that are most often regarded as feedbacks (see Sect. 6.2). However, this study chose to limit the degrees of freedom with respect to time-variable climate sensitivities given that a clear separation into three (or more) different parameterizations seemed unjustified based on the AOGCM data analyzed here.

3.2 Updated carbon cycle

MAGICC's terrestrial carbon cycle model is a globally integrated box model, similar to that in Harvey (1989) and Wigley (1993). The MAGICC6 carbon cycle can emulate temperature-feedback effects on the heterotrophic respiration carbon fluxes. One improvement in MAGICC6 allows increased flexibility when accounting for CO_2 fertilization. This increase in flexibility allows a better fit to some of the more complex carbon cycle models reviewed in C^4MIP (Friedlingstein et al., 2006) (see Sect. 5.1).

Another update in MAGICC6 relates to the relaxation in carbon pools after a deforestation event. The gross CO_2 emissions related to deforestation and other land use activities are subtracted from the plant, detritus and soil carbon pools (see Fig. A2). While in previous versions only the regrowth in the plant carbon pool was taken into account to calculate the net deforestation, MAGICC6 now includes an effective relaxation/regrowth term for all three terrestrial carbon pools (see Appendix A1.1).

The original ocean carbon cycle model used a convolution representation (Wigley, 1991b) to quantify the ocean-atmosphere CO_2 flux. A similar representation is used here, but modified to account for nonlinearities. Specifically, the impulse response representation of the Princeton 3D GFDL model (Sarmiento et al., 1992) is used to approximate the inorganic carbon perturbation in the mixed layer (for the impulse response representation, see Joos et al., 1996). The temperature sensitivity of the sea surface partial pressure is implemented based on Takahashi et al. (1993) as given in Joos et al. (2001). For details on the updated carbon cycle routines, see the Appendix A1.

3.3 Other additional capabilities compared to MAGICC4.2

Five additional amendments to the climate model have been implemented in MAGICC6 compared to the MAGICC4.2 version that has been used in IPCC AR4 or MAGICC5.3.

3.3.1 Aerosol indirect effects

It is now possible to account directly for contributions from black carbon, organic carbon and nitrate aerosols to indirect (i.e., cloud albedo) effects (Twomey, 1977). The first indirect effect, affecting cloud droplet size and the second indirect effect, affecting cloud cover and lifetime, can also be modeled separately. Following the convention in IPCC AR4 (Forster et al., 2007), the second indirect effect is modeled as a prescribed change in efficacy of the first indirect effect. See Sect. A3.6 in the Appendix for details.

3.3.2 Depth-variable ocean with entrainment

Building on the work by Raper et al. (2001), MAGICC6 includes the option of a depth-dependent ocean area profile with entrainment at each of the ocean levels (default, 50 levels) from the polar sinking water column. The default ocean area profile decreases from unity at the surface to, for example, 30%, 13% and 0% at depths of 4000, 4500 and 5000 m. Although comprehensive data on depth-dependent heat uptake profiles of the CMIP3 AOGCMs were not available for this study, this entrainment update provides more flexibility and allows for a better simulation of the characteristic depth-dependent heat uptake as observed in one analyzed AOGCM, namely HadCM2 (Raper et al., 2001).

3.3.3 Vertical mixing depending on warming gradient

Simple models, including earlier versions of MAGICC, sometimes overestimated the ocean heat uptake for higher warming scenarios when applying parameter sets chosen to match heat uptake for lower warming scenarios, see e.g. Fig. 17b in Harvey et al. (1997). A strengthened thermal stratification and hence reduced vertical mixing might contribute to the lower heat uptake for higher warming cases. To model this effect, a warming-dependent vertical gradient of the thermal diffusivity is implemented here (see Appendix A4.7).

3.3.4 Forcing efficacies

Since the IPCC TAR, a number of studies have focussed on forcing efficacies, i.e., on the differences in surface temperature response due to a unit forcing by different radiative forcing agents with different geographical and vertical distributions (see e.g. Joshi et al., 2003; Hansen et al., 2005). This version of MAGICC includes the option to apply different efficacy terms for the different forcings agents (see Appendix A4.4 for details and Supplement for default values).

3.3.5 Radiative forcing patterns

Earlier versions of MAGICC used time-independent (but user-specifiable) ratios to distribute the global-mean forcing of tropospheric ozone and aerosols to the four atmospheric boxes, i.e., land and ocean in both hemispheres. This model structure and the simple 4-box forcing patterns are retained as it is able to capture a large fraction of the forcing agent characteristics of interest here. However, we now use patterns for each forcing individually, and allow for these patterns to vary over time. For example, the historical forcing pattern evolutions for tropospheric aerosols are based on results from Hansen et al. (2005), which are interpolated to annual values and extrapolated into the future using hemispheric emissions. Additionally, MAGICC6 now incorporates forcing patterns for the long-lived greenhouse gases as well, although these patterns are assumed to be constant in time and scaled with global-mean radiative forcing (see Supplement for details on the default forcing patterns and time series).

4 Calibrating MAGICC to AOGCMs

In the preparation of the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4), 14 modeling groups submitted data for 23 AOGCMs, building the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset. The following subsection (4.1) describes the method used to calibrate MAGICC for 19 of these AOGCMs, i.e., those for which sufficient data were available. In subsection 4.2, the results of the calibration procedures are provided.

4.1 Climate model calibrating procedure

Three distinct calibration exercises are undertaken, optimizing a smaller (I) or larger (II, III) set of MAGICC parameters, using idealized scenarios only (I, II), or optimizing against multi-forcing scenarios as well (III).

Table 1. Overview of calibration exercises. The hemispheric land and ocean surface air temperatures and ocean heat uptake were used for each experiment.

Calibration Method	Experiments Used[a]	Calibrated Parameters[b]
I. "Basic/AR4-like"	Idealized Scenarios (1pctto2×, 1pctto4×)	$\Delta T_{2\times}$, RLO, Kz
II. "Medium"	Idealized Scenarios (1pctto2×, 1pctto4×)	$\Delta T_{2\times}$, ξ, RLO, Kz, $\frac{dKz_{\text{top}}}{dT}$, k_{NS}, k_{LO}, μ
III. "Full"	Idealized Scenarios (1pctto2×, 1pctto4×) and Multi-Forcing Runs (20c3m, COMMIT, SRESB1, SRESA1B)	$\Delta T_{2\times}$, ξ, RLO, Kz, $\frac{dKz_{\text{top}}}{dT}$, k_{NS}, k_{LO}, μ

[a] The scenarios are: 1pctto2× = 1% annual CO_2 concentration increase until CO_2 doubling, then stabilization; 1pctto4× = 1% annual CO_2 concentration increase until CO_2 quadrupling, then stabilization; 20c3m = historical 20th century run; COMMIT = year 2000 concentration stabilization; sresb1 = IPCC SRES B1 scenario; sresa1b = IPCC SRES A1B scenario.

[b] The calibrated parameters are as follows: $\Delta T_{2\times}$ = climate sensitivity (KW^{-1}m^2), i.e., warming after a doubling of CO_2 concentrations; RLO=Land-Ocean warming ratio at equilibrium; Kz = vertical diffusivity in ocean ($\text{cm}^2\ \text{s}^{-1}$); ξ = sensitivity of feedback factors λ to radiative forcing change ΔQ away from doubled pre-industrial CO_2 forcing level $\Delta Q_{2\times}$, see Eq. (A51); $\frac{dKz_{\text{top}}}{dT}$ =sensitivity of vertical diffusivity at mixed layer boundary to global-mean surface temperatures (i.e., thermal stratification). A linear diffusivity profile change is assumed for layers between the mixed and bottom layers; k_{LO} = Land-Ocean heat exchange coefficient ($\text{Wm}^{-2}\ \text{K}^{-1}$); μ = an amplification factor for the ocean to land heat exchange (see Eq. A50).

The calibration I approach mimics the procedure employed for IPCC AR4. Three key parameters (see Sect. 4.1.1 below) were calibrated to optimally reproduce the hemispheric land and ocean temperatures and ocean heat flux responses to idealized 1%/yr increasing CO_2-only scenarios (see Table 1). Secondly, an additional five parameters were optimized (calibration II) to match the idealized CO_2-only scenarios better. Thirdly, the most comprehensive calibration exercise (calibration III) employs, in addition, the AOGCM results for multi-gas scenarios, viz. the year-2000 constant concentration (COMMIT) experiments, and the SRES B1 and A1B scenarios, if available. The SRES A2 scenario is not used for calibrating MAGICC parameters, but was instead used for verification. See Table 1 for an overview of the three calibration exercises. Going beyond the match of global-mean temperatures and heat uptake that were fitted in earlier MAGICC versions, all calibration exercises also took into account hemispheric land and ocean temperatures, diagnosed from one of the ensemble members of each CMIP3 AOGCM (run 1) provided at the PCMDI database (http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php). To take account of model drift, the corresponding low pass-filtered ($1/20\ \text{yr}^{-1}$ cutoff frequency) control-run segments were subtracted from each perturbation run (see Appendix B).

4.1.1 Calibrated parameters

In the first calibration exercise (calibration 1), only three key parameters were optimized, namely the climate sensitivity $\Delta T_{2\times}$, the equilibrium land-ocean warming ratio RLO and the vertical thermal diffusivity *Kz* in the ocean. *Kz* has a large influence on the ocean heat uptake efficiency. In the second and third calibration exercises, five additional parameters in MAGICC were optimized to match the AOGCM temperature and heat-uptake results. As in any calibration exercise with multiple parameters, there is the danger of overfitting. Therefore, only a limited set with clearly distinct effects representing different physical mechanisms was chosen out of the large number of MAGICC parameters. Two of the additional five parameters are required to emulate time-varying effective climate sensitivities: namely μ, the ocean to land heat-exchange amplification, which allows the emulation of increasing effective sensitivities under global warming (Appendix A4.2); and ξ, the forcing-dependency of the feedback (see Appendix A4.3). Another parameter, the ocean stratification coefficient $\frac{dK_z}{dT}$, modulates the heat uptake efficiency under higher warming scenarios, by making the vertical diffusivity dependent on the ocean warming (see Appendix A4.7). Furthermore, the two heat-exchange parameters between land and ocean (k_{LO}) and between the hemispheres (k_{NS}) are calibrated, with the latter having no influence on the global-mean warming, but on the hemispheric warming pattern.

Several parameters were kept fixed at default values in our calibration exercises. For example, we held the sea-ice related adjustment factor α, which determines the ratio of hemispheric changes in air versus ocean mixed layer temperatures at its default value of 1.2 - based on experience with earlier versions of MAGICC (Raper and Cubasch, 1996). It is possible that this should also be a calibrated, model-specific parameter. In addition, ocean heat uptake depends on how the upwelling rate w changes over time, which varies from model to model (see Sect. A4.5). In previous versions of MAGICC this has also been a calibrated parameter (Raper et al., 2001). Here we capture the general AOGCM behavior by assuming that $w(t)$ depends linearly on the global-mean temperature, declining from an initial value of 4 m/yr to 2.8 m/yr at a warming of 8 °C (relative to the pre-industrial temperature) and remaining constant thereafter (cf. Raper et al., 2001). This simplified parameterization corresponds

approximately to a central estimate of the overturning circulation's response for the majority of CMIP3 AOGCMs in the 21st century simulations (see Fig. 10.15 in Meehl et al., 2007). We do not attempt to emulate the meridional overturning specifically for each AOGCM (Schleussner et al., 2010), thereby limiting the overall number of calibrated parameters. Using an AOGCM-specific vertical diffusivity allows us to closely emulate the AOGCM's surface temperature and ocean heat uptake responses, which are of primary interest in this study. As with the sea-ice related factor α, better fits to the AOGCMs may be obtained when emulating thermal expansion and vertical ocean temperature profiles if $w(t)$ were a calibrated, model-specific characteristic (Raper and Cubasch, 1996; Harvey, 1994).

For calibrating to each specific AOGCM, the parameter space in MAGICC is first sampled randomly with 2000 parameter sets. For each parameter set, up to five parallel runs were done, one for each of the calibration scenarios. Subsequently, the best (in a least-squared sense) parameter set is used to initialize an optimization routine with approximately 1000 iterations to find the parameter combination that minimizes the squared differences between low-pass filtered AOGCM and MAGICC time series of heat uptake, global, northern land, northern ocean, southern land and southern ocean surface air (2 m) temperatures. See Appendix B for details.

4.1.2 Calibration against idealized CO_2 scenarios

In order to successfully emulate the climate response of an AOGCM, its driving forces should be known. This is why idealized experiments, where the forcing is known, are preferred for calibration. For example, MAGICC calibrations for the IPCC TAR, as well as feedback paramater calculations by F&T, used the first 70 years of the idealized 1% runs. MAGICC 4.2 calibrations for IPCC AR4 used the full-length 1% runs (1pctto2× and 1pctto4×, cf. Fig. 1). All 19 CMIP3 AOGCMs considered here provided at least some output for such idealized forcing experiments, assuming annual 1% increases of CO_2 up to doubled and quadrupled concentrations and constant concentrations thereafter (1pctto2× and 1pctto4×, respectively) (see rows 2, 4 and 6 in Figs. B1, B2 and B3). Most AOGCMs started these experiments from pre-industrial control runs (picntrl), although four (CCSM3, MRI-CGCM2.3.2, ECHO-G, NCAR PCM) used present-day control runs (*pdcntrl*). Control-run drift was removed using the respective low pass-filtered (1/20 yr^{-1} cutoff frequency) control run segments. Assuming that the CO_2 concentration to forcing relationship is logarithmic (Shine et al., 1990; Myhre et al., 1998), the forcing is a linear ramp-function over 70 (140) years up to its forcing level $\Delta Q_{2\times}$ at doubled (or quadrupled) CO_2 concentrations and constant thereafter. $\Delta Q_{2\times}$ is estimated to be 3.71 Wm^{-2} (Myhre et al., 1998), although AOGCMs show a relatively large variation (see Table 10.2 in Meehl et al., 2007). Where available, model-specific $\Delta Q_{2\times}$ values were used during the calibration exercise (see Tables B1, B2 and B3).

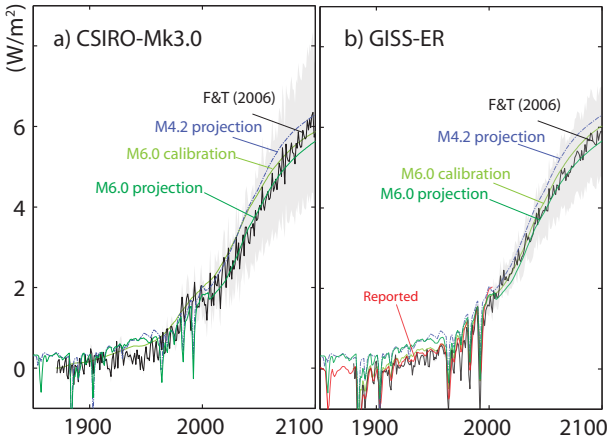

Fig. 2. Effective radiative forcing for the SRES A1B scenario from 1850 to 2100 for two CMIP3 AOGCMs. Shown are the net effective radiative forcing time-series used for calibrating MAGICC6 to CSIRO-Mk3.0 **(a)** and GISS-ER **(b)** ("M6.0 calibration"). Due to various unification adjustments and complementation of the sparse AOGCM-specific forcing sets, the effective forcings prescribed for the projections differ. Shown here is the mean for each AOGCM when combined with the ten C^4MIP carbon cycle model calibrations ("M6.0 projection"). For comparison, the forcings used in IPCC AR4 for the medium carbon cycle feedback case ("M4.2 projection") and the effective forcings (including uncertainties) as diagnosed by Forster and Taylor (2006) ('F&T, 2006') are also shown. In addition, in the case of the GISS-ER model, radiative forcing time series were made available by the modeling group ("Reported") (J. Hansen, personal communication, 2005, as reported in Forster and Taylor, 2006).

4.1.3 The difficulty posed by unknown radiative forcing

The inherent difficulty with calibrating MAGICC parameters to the multi-forcing AOGCM results, and the reason why this approach has not been used previously, is that there are large uncertainties in the actual forcings. There are two reasons of why forcings differ across AOGCMs. First, not all models used the same set of forcings. As the particular forcings used are known (see Table 2), our calibration exercises were able to take this into account. Second, even for the forcings in common, quantitative AOGCM-specific information is very limited, mostly restricted to CO_2 forcings at doubled CO_2 concentrations. The first study addressing this shortcoming in a comprehensive manner is by F&T, who diagnosed the effective forcings. However, neither forcings nor efficacies can be diagnosed from the currently available AOGCM data without making additional assumptions; for example, with regard to the models' effective climate sensitivities (see F&T).

Table 2. The subsets of forcing agents considered during the calibration III exercise to match the setup of CMIP3 AOGCM multi-forcing runs (cf. Table 10.1, Meehl et al., 2007). The forcing agents included for the multi-forcing simulations are indicated by the following symbols: C = carbon dioxide; M = methane; N = nitrous oxide; S = stratospheric ozone; T = tropospheric ozone; F = halogenated carbons (CFCs); D = direct SO_x aerosol; B = direct black carbon; O = direct organic carbon; R=nitrate aerosols; I^1 = first indirect aerosol effect; I^2 = second indirect aerosol effect; U = mineral dust; L = land use; V = volcanic; A = solar; Capital letters denote that a time-variable forcing agent is included for simulations prior to and after 2000; small letters denote that forcing agent varies with time in 20th century (20c3m) simulations and is set constant thereafter; italic letters denote that forcing was included as CO_2 equivalent, hence efficacies were set to 1 during calibration. Bracketed subscripts indicate the applied volcanic forcing scaling factor derived from Fig. 4 in F&T. Note that for some models, the forcing analysis by F&T detects no volcanic forcing ($v_{(0.0)}$), although Table 10.1 in Meehl et al. (2007) indicate the inclusion of volcanic aerosols. Future solar forcing in GISS-EH and GISS-ER was assumed to be cyclic in the original simulations, but is here assumed constant, denoted by a′. Mineral dust (U) forcings and land use albedo (L) effects have been assumed constant after 2000 in all emulations. For ECHAM5/MPI-OM and UKMO-HadCM3, denoted with a "*", the first indirect effect was assumed in the MAGICC emulations to be equal to the default joint forcing of 1st and 2^{nd} indirect effect given the substantial negative shortwave forcing analyzed by F&T for these models.

AOGCM	Forcing agents
(BCCR-BCM2.0)	CMNstFDula
(BCC-CM1)	CMNStFDlva
CCSM3	CMNSTFDBOuv$_{(0.86)}$a
CGCM3.1(T47)	CMNstFDulv$_{(0.0)}$a
(CGCM3.1(T63))	CMNstFDulv$_{(0.0)}$a
CNRM-CM3	CMNSTFDu
CSIRO-Mk3.0	C*MNSTFD*
ECHAM5/MPI-OM	CMNStFDI^{1}*
ECHO-G	CMNsTFDI^{1}v$_{(0.70)}$a
FGOALS-g1.0	CMNstFDa
GFDL-CM2.0	CMNSTFDBOulv$_{(0.70)}$a
GFDL-CM2.1	CMNSTFDBOulv$_{(0.70)}$a
(GISS-AOM)	CMNstFD
GISS-EH	CMNSTFDBORI^{2}uLv$_{(0.70)}$a′
GISS-ER	CMNSTFDBORI^{2}uLv$_{(0.70)}$a′
INM-CM3.0	CMNstDv$_{(0.2)}$a
IPSL-CM4	CMNFDI^{1}
MIROC3.2(H)	CMNSTFDBOI^{1,2}Ulv$_{(0.52)}$a
MIROC3.2(M)	CMNSTFDBOI^{1,2}Ulv$_{(0.35)}$a
MRI-CGCM2.3.2	CMNstFDv$_{(0.40)}$a
PCM	CMNSTFDv$_{(0.7)}$a
UKMO-HadCM3	CMNSTFDI^{1}*v$_{(0.0)}$a
UKMO-HadGEM1	CMNSTFDBOI^{1,2}Lv$_{(0.0)}$a
"Full Forcing" Emulation IIIc/d	CMNSTFDBOI^{1,2}ulv$_{(0.7)}$a

In the present study, given these limitations, we use informed estimates for the individual model forcings. Only the matching set of radiative forcing agents (see Table 2) together with default efficacies (see Supplement) were applied in MAGICC when calibrating each AOGCM. These reconstructed forcing time-series are not identical to the diagnosed forcings given by F&T. In the case of the GISS models, the modeling group provided an independent estimate of the radiative forcing (J. Hansen, 2005, personal communication as reported in F&T), which agrees well with the net effective forcing series used for calibration here (see Fig. 2b). A more detailed discussion of both the MAGICC4.2 and MAGICC6 forcing assumptions and emulations can be found in Part 2 of this study.

4.1.4 Special cases for multi-forcing calibration

For the individual forcing agents used in the calibration, MAGICC applies the same forcing timeseries with histories whose magnitude from 1765 to 2005 is consistent with the central estimate provided by IPCC AR4 for each individual forcing agent (see Fig. 1 in Part 2 or Table 2.12 in Forster et al., 2007). The four exceptions are:

Firstly, for volcanic forcing, the amplitude was adjusted for each AOGCM that included volcanic forcing, so that the (negative) amplitude in net effective (shortwave and longwave) volcanic forcing was approximately matched to the value calculated by F&T. In fitting to historical time series (using squared differences as the goodness-of-fit statistic), a too strong negative amplitude would result in too high a sensitivity $\Delta T_{2\times}$, and hence a future MAGICC response that is too warm. To minimize the effect of mismatching volcanic forcing series a low pass filter was applied to the temperature series before the optimization. The scaling factor for volcanic forcing was determined to be lower than unity for all models (ranging from 0.2 for INM-CM3.0 and MRI-CGCM2.3.2 to around 0.7 for most models). See Table 2.

Secondly, CO_2 related forcing is modeled slightly differently compared to other forcing agents. For the idealized scenarios, we used the actual CO_2 concentrations. To convert concentrations to forcing we set $\Delta Q_{2\times}$ to its AOGCM-specific value during the calibration exercise (see Eq. A35 and A36). For the SRES scenarios (B1 and A1B) we also drove MAGICC with concentrations rather than emissions. We assumed that CMIP3 AOGCMs prescribed CO_2 concentrations according to the Bern reference provided in the IPCC TAR. Prescribing CO_2 concentrations instead of emissions has the additional benefit of keeping the calibration of the carbon cycle (see following Sect. 5.1) strictly separate from the calibration of the climate response.

Thirdly, a special case is the second indirect aerosol effect, characterized by default in IPCC AR4 (Forster et al., 2007) as an efficacy enhancement to the first indirect aerosol effect. For AOGCMs that only included the first indirect effect (ECHAM5/MPI-OM, ECHO-G, IPSL-CM4,

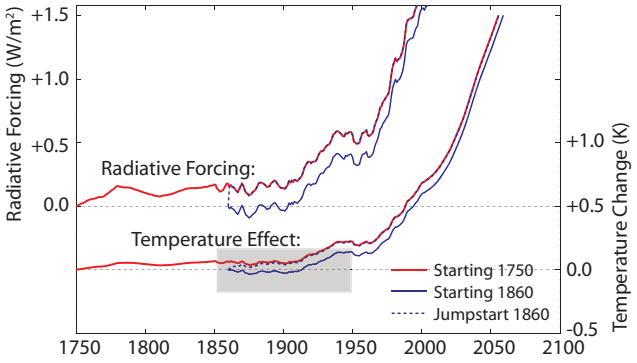

Fig. 3. Radiative forcing and temperature evolutions illustrating the "cold start problem" (Hasselmann et al., 1993). A climate model run taking into account the forcing history since 1750 (red line) provides a different future projection compared to a run taking into account deviations from a later startyear only, e.g. 1860 (blue solid line). A later common reference period, e.g. 1980–1999, or a "jump start" with radiative forcing being applied relative to 1750 (blue dashed line), minimizes this initialization problem. The temperature response for the "jump start" run asymptotically approaches the results for the run starting in 1750 (grey shaded area).

UKMO-HadCM3), the second effect was ignored during the calibration exercise. For the GISS-EH and GISS-ER models, which only included the second indirect effect (see Table 10.1 in Meehl et al., 2007), a forcing was assumed of the same magnitude as IPCC AR4's best estimate of the first indirect aerosol effect ($-0.7\,\mathrm{Wm}^{-2}$ with efficacy 0.9). For the three models MIROC3.2 (hires), MIROC3.2 (medres) and HadGEM1 that are reported to have included both indirect aerosol effects the second indirect effect is assumed to enhance the first indirect effect by two-thirds, by increasing the efficacy from 0.9 to 1.5. These (rather uncertain) default values have been chosen from the uncertainty ranges provided in IPCC AR4 for the first indirect effect's efficacy (stated to be similar to the direct aerosol effect's efficacy of 0.7 to 1.1) and the efficacy that includes both the first and second indirect effect (1.0 to 2.0), respectively (see Sect. 2.8.5.5. in Forster et al., 2007).

Fourthly, the last issue relates to the "cold start problem" (Hasselmann et al., 1993). Rather than starting in 1750, the reference year for radiative forcings, modeling groups chose years in between 1850 and 1900 as a starting point for the 20th century integrations (20c3m runs). Unfortunately, it is not documented how (or if) the AOGCM modeling groups handled any forcing differences between 1750 and the respective starting year. For example, in the default forcing series applied here (excluding volcanic forcing), a slight forcing increase of roughly $+0.2\,\mathrm{Wm}^{-2}$ occurred between 1750 and 1860. To account for this, modeling groups could have applied a "jump start", so that the model is subject to a step forcing increase in the starting year (see Fig. 3). Alternatively, models could be driven by radiative forcing changes from their starting year only, neglecting any forcing changes between 1750 and their starting year. Although the choice of initialization method does affect the fitted parameter values, the effect of these different possible initializations is small. We assumed here (based on the CMIP3 AOGCM temperature results, which show no evidence of a "jump start") that AOGCM runs were begun with zero forcing in their 20c3m starting year. However, the HadCM3LC C^4MIP coupled carbon cycle-climate model's temperature evolution suggests that it has been subject to a "jump start" in forcing and so we do likewise. Such "jump start" initializations have been used earlier as well – as documented in Johns et al. (1997) (see Fig. 30a therein).

4.2 AOGCM calibration results

This section gives the results of the three calibration exercises employed here to replicate the climate response characteristics of the AOGCMs (Sects. 4.2.1, 4.2.2, and 4.2.3). Subsequently, Sect. 4.2.4 compares climate sensitivities diagnosed for the CMIP3 AOGCMs.

4.2.1 Calibration method I – as in the AR4

This simple calibration approach I (see Table 1 for the found best-fit parameters) is able to emulate the evolution of global-mean temperatures for the idealized scenarios relatively well for most AOGCMs (see Table B1). The root mean square errors (RMSE) between emulations and the AOGCMs are well below 0.2 °C for the 1pctto2× and 1pctto4× scenarios for all but four models (UKMO HadGEM1, CCCma CGCM3.1(T47), GFDL CM2.0 and MPI ECHAM5), as shown in Fig. 5a. As can be expected, the SRES and "COMMIT" multi-forcing scenarios are less well emulated for almost all models, as their information was not used to derive the optimal parameter settings for $\Delta T_{2\times}$, *RLO* and *Kz*. This discrepancy between emulations and AOGCM multi-forcing runs is substantial for three out of the 19 emulations showing RMSE values higher than 0.35 °C. On average across all models and scenarios, the RMSE is 0.21 °C (see Fig. 5a).

In order to put this RMSE value of 0.21 °C in perspective, it is here compared to the equivalent goodness of fit statistic that would be obtained if a single AOGCM's projections were simply approximated by the global-mean temperature time-series of another randomly drawn AOGCM for the same scenario. This comparison is motivated by the common practice in many studies to make inferences from single AOGCMs, often implying that a single AOGCM is representative for a wider range of other AOGCMs. Essentially, this compares the uncertainty in fitting MAGICC to a particular model to the inter-model uncertainty. Thus, for this comparative measure of inter-model uncertainty, we computed the average RMSE between global-mean temperature series for all permutations of CMIP3 AOGCMs applying the same lowpass filter as used for the calibrations ($1/20\,\mathrm{yr}^{-1}$ cutoff

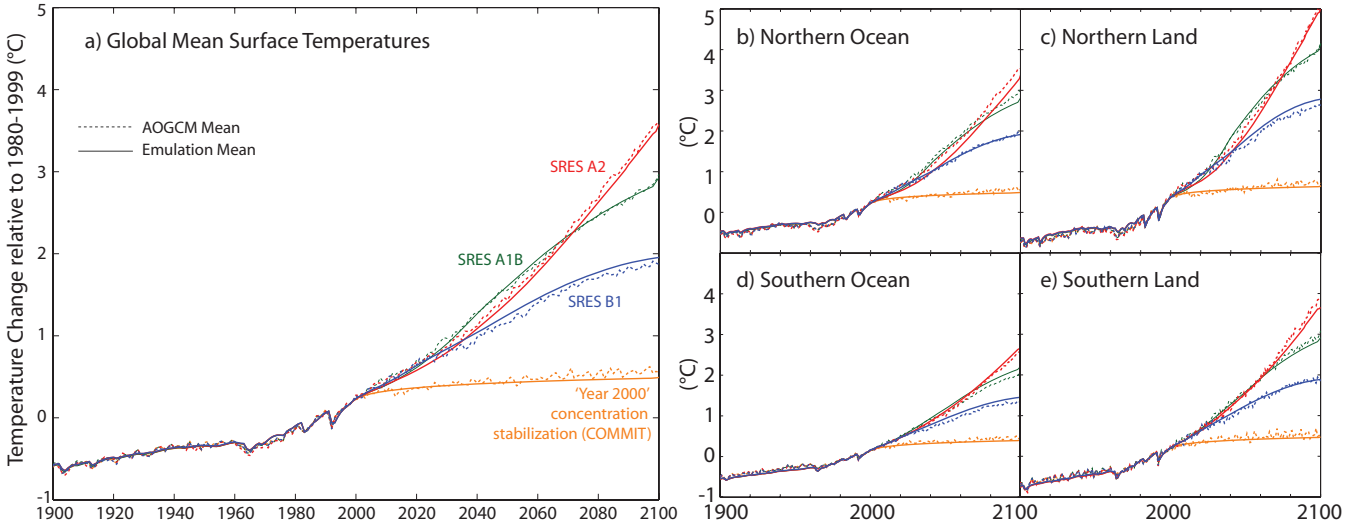

Fig. 4. Comparison of mean surface temperatures as diagnosed from CMIP3 AOGCMs (dashed) and the emulations with MAGICC6 using "like-with-like" forcings and the calibration III method (solid lines, see Sect. 4.2.3). The scenarios shown are SRES A1B (green), B1 (blue) and A2 (red) in addition to the "year 2000 concentration stabilization" (COMMIT) experiment (orange). For the different scenarios, the number of available AOGCM datasets differs, which is taken into account, so that only the mean across the corresponding set of emulations is shown. The land and ocean regions in each hemisphere were determined from the individual AOGCMs' land-ocean masks.

frequency), taking into account the full overlapping time-periods between any pair of AOGCMs. The resulting RMSE is 0.46 °C across the multi-forcing and idealized scenarios, more than twice as high compared to the RMSE of emulations following the calibration I procedure.

It is noticeable that some AOGCMs show features in their idealized scenario runs (1pctto2× and 1pctto4×) that cannot possibly be emulated satisfactorily by optimizing only three parameters $\Delta T_{2\times}$, *RLO* and *Kz*. For example, a larger best-fit effective climate sensitivity for the higher forcing 1pctto4× run than for the 1pctto2× run is apparent in the MPI ECHAM5 simulation, after these runs diverge in year 70 of the model integration (see Fig. 1, and discussion in Sect. 4.2.4). A constant climate sensitivity $\Delta T_{2\times}$ can never, therefore, match both scenarios satisfactorily. The best-fit constant climate sensitivity will be in-between the effective sensitivities for the 1pctto2× and 1pctto4× runs. Indeed, the calibration I procedure gives a climate sensitivity of 3.95 °C (see Table B1), which is in between the effective sensitivities of 3.5 and 4.2 °C towards the end of the 1pctto2× and 1pctto4× scenarios, respectively (see Fig. 1).

4.2.2 Calibration method II – using additional parameters

For some AOGCMs, the use of additional parameters in the fitting exercise did not improve the goodness of fit (MIROC3.2(hires), GISS-EH and FGOALS-g1.0). For others, the fit was improved markedly. For example, the RMSE is halved for NCAR CCSM3 and GISS-ER (see 1pctto2× and 1pctto4× scenarios in Fig. 5a and c). The enhanced ability to match the idealized scenarios of the MPI ECHAM5 model is most noticeable: under calibration I (fitting only three parameters), the RMSE values were 0.30 °C and 0.43 °C for the 1pctto2× and 1pctto4× scenarios. Under the calibration II method the idealized scenarios are now emulated with an RMSE of 0.15 °C and 0.11 °C – primarily due to the ability of MAGICC to simulate time-varying effective sensitivities (see Fig. 1). The multi-forcing scenarios are also more accurately emulated, so that the goodness of fit ranking for MPI ECHAM5 improved (see Fig. 5).

In summary, the match to the idealized scenarios improved for all those 14 models that provided 1pctto2× and 1pctto4× data, but not for those five (MIROC3.2(hires), GISS-EH and FGOALS-g1.0, UKMO-HadCM3 and CSIRO-Mk3.0) that provided only 1pctto2× data (see Figures B1, B2 and B3). The emulation skill for the multi-forcing scenarios, which were not used for calibration II, was only slightly enhanced in most cases. The average RMSE across all scenarios and models is 0.19 °C (see Fig. 5a and c), slightly improved from the 0.21 °C that resulted from the calibration I procedure.

4.2.3 Calibration method III – from CO_2-only to multi-forcing

While the inclusion of additional parameters under the calibration II procedure markedly improved the fit to the idealized experiments, the performance of the emulations for the multi-forcing runs is only slightly improved. Obviously, the emulation quality for SRES scenarios will be improved, if an appropriate goodness of fit criteria related to the SRES scenarios is included in the optimization routine. The close

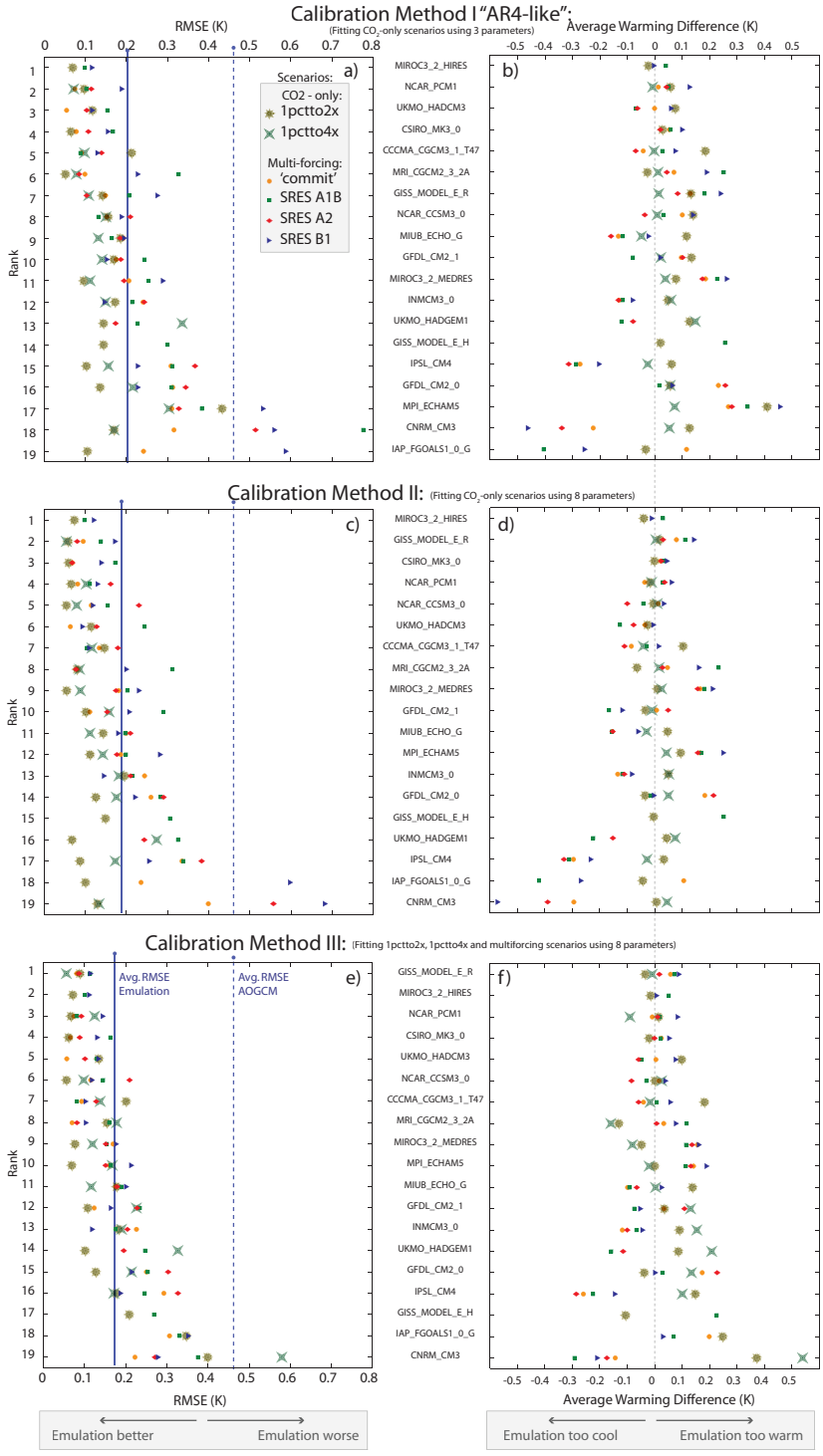

Fig. 5. The root mean square error (RMSE) and average warming differences between global-mean temperatures for individual AOGCMs and their emulations after calibrating MAGICC parameters with three different calibration procedures. Temperatures and ocean heat uptake for the 1pctto2× and 1pctto4× scenarios were fitted by calibrating three (calibration I; panel a, b) and eight (calibration II; panel c, d) MAGICC parameters, respectively (see Table 1). Calibration Method "III" (panel e,f) used in addition the multi-forcing runs SRES A1B, B1 and COMMIT when optimizing eight parameters (see Table 1). The emulations are ranked according to mean deviations (RMSE) between emulations and AOGCM data over the full length of all available scenarios. The AOGCM and MAGICC data were lowpass-filtered when calculating the RMSE values. For all emulations, "like-with-like" forcings were applied, i.e., the emulations were not subject to forcing adjustments. The mean RMSE for all emulations is given ("Avg. RMSE Emulations") and compared to the average inter-model RMSE ("Avg. RMSE AOGCM"). See text.

fit between the mean of the emulations and the mean of the AOGCM runs under the calibration III strategy is shown in Fig. 4 (see also Table 3).

Assessing our calibrations at the level of individual AOGCMs, the deviations over the full scenario durations are small, mostly $<0.2\,^{\circ}$C (see Fig. 5f). The largest deviation in global means of up to $0.5\,^{\circ}$C occurs for CNRM CM3. The emulations of CNRM CM3 show most clearly what is apparent as well for eight other AOGCMs (GISS-ER, MIROC3.2 (medres), NCAR PCM1, MPI ECHAM5, MRI CGCM2.3.2A, IPSL-CM4, INM-CM3.0 and HadGEM1), namely that the idealized scenarios are emulated too warm and the multi-forcing runs too cold or vice versa (see Fig. 5f). In the case of CNRM CM3, this may be caused by an underestimation of the net forcing in the multi-forcing runs and/or an overestimation of the CO_2 forcing in the idealized scenarios. For calibration III results, the average RMSE across all scenarios and models is further decreased to 0.17 K (c.f. 0.21 K and 0.19 K for "calibration I and II"). This is substantially lower than the AOGCM inter-model uncertainty RMSE of $0.46\,^{\circ}$C. Another useful comparison metric is the skill with which the emulations compare with the AOGCMs when averaged over all AOGCMs. The mean AOGCM versus mean emulation RMSE, over all multi-forcing runs, for 2000 to 2100, is $0.053\,^{\circ}$C. This shows that the emulations of the multi-model ensemble mean is substantially more robust than emulating a single AOGCM and is associated with only very minor biases (see Fig. 4).

As noted above, the SRES A2 scenario has not been used for calibration, but left as an independent test case for the skill of the emulations. The performance of the emulations for the high SRES A2 scenario is similar to the other two SRES scenarios, B1 and A1B, that were used in the calibration (average RMSE A2: $0.175\,^{\circ}$C; A1B: $0.190\,^{\circ}$C, B1: $0.168\,^{\circ}$C; see Fig. 5e). This is encouraging as it supports the assumption that emulations for other emissions scenarios approximately reflect what AOGCMs would project. On average across model emulations, the bias is again small, as can be seen in Figure 4 with average warming under SRES A2 being slightly lower in the emulations.

It is valuable to put these emulation errors in perspective. For the SRES scenarios, the inter-model uncertainties between AOGCMs with regard to global-mean temperatures towards the end of the 21st century (2090–2099), when expressed as two standard deviations divided by the multi-model ensemble mean, range from 49% for SRES B1 (21 models) through 41% for A1B (21 models) to 26% for A2 (17 models) (cf. Knutti et al., 2008). In comparison, the mean relative errors introduced by the emulations are substantially smaller, i.e., less than 2.2% for the ensemble means (B1:2.2%, A1B:−1.0%, A2:−0.8%) and, on average, 7% for individual AOGCM emulations over 2090 to 2099 relative to 1980 to 1999 (B1:9%, A1B:6%, A2:6%). Comparing the 2090 to 2099 warming relative to AOGCM starting years reduces differences between emulations and AOGCMs further. This is primarily because the earlier start date for the comparison removes uncertainties introduced by the strong Pinatubo volcanic forcing in the 1980 to 1999 base period. Individual AOGCMs in the last decade of the 21st century are now matched on average with a mean relative error of only 6% (B1:5%, A1B:5%, A2:7%). Half of the emulation and AOGCM pairs show deviations of only 3% on average (B1:3%, A1B:2%, A2:5%). As noted above for the example of the CNRM CM3 model, calibrations are necessarily imperfect as we do not know the precise forcings effective in the AOGCMs. This problem is likely enhanced in the calibrations towards the multi-forcing AOGCM results compared to those for the idealized CO_2 runs.

4.2.4 Comparison of climate sensitivities

Equilibrium climate sensitivity is a useful aggregate model indicator and climate system characteristic (Knutti and Hegerl , 2008). Traditionally, climate sensitivity is defined as the warming resulting from any doubling of CO_2 concentrations, irrespective of the starting concentration level. With the introduction of climate-state dependent sensitivities, we report here the climate sensitivities for a doubling of pre-industrial concentrations and compare these to other published estimates for the set of CMIP3 AOGCMs (see Table 4). Many modelling groups reported equilibrium warming results with their slab ocean model versions, stated in the first column of Table 4 taken from Randall et al. (2007). The average climate sensitivity across all 19 slab-ocean models is $3.21\,^{\circ}$C. The coupled versions of these models can exhibit different sensitivities from the slab-ocean versions, not least because the presence of a coupled ocean can alter atmospheric feedbacks (Gregory et al., 2004). Time-evolving effective climate sensitivities S^t can be diagnosed from any transient run for which the forcing and ocean heat uptake is known, as given in Eq. (3) (see Murphy and Mitchell, 1995; Raper et al., 2001; Senior and Mitchell, 2000). Gregory et al. (2004) have developed a regression technique to estimate the effective climate sensitivity even if the absolute forcing is unknown. F&T calculated climate sensitivities for the CMIP3 AOGCMs from the first 70 years of the idealized 1pctto2× scenarios (cf. Fig. 1). The average climate sensitivity following this procedure (viz. $2.76\,^{\circ}$C) is nearly half a degree cooler than that estimated for the slab-ocean models (cf. first and second column in Table 4).

MAGICC4.2 climate sensitivity results presented in IPCC AR4 (see Supplementary Table S8.1 in Randall et al. (2007) and Fig. 10.26 in Meehl et al. (2007)) and those for MAGICC6 using the calibration I method are very similar to each other (less than $0.2\,^{\circ}$C difference), except for HadGEM1, for which additional AOGCM data were available in the MAGICC6 case. For 13 out of 19 AOGCMs, these sensitivities are very similar to those in F&T, with differences less than $0.2\,^{\circ}$C. For the remaining six models analyzed by both studies, MAGICC calibrations give higher climate sensitivities,

Table 3. Comparison of global-mean temperatures from AOGCMs and emulations for three periods. The means across all available CMIP3 AOGCMs for each scenario (number of available AOGCM datasets given in column "*n*") are compared to the mean across the matching number of emulations using AOGCM-specific "like-with-like" forcings, denoted by "IIIa". The emulations with parameter settings from calibration III (see text) and applying "full" forcing emulations, averaged across all 19 emulations, are shown for comparison (column IIId). On the notation: The three methods for calibrating carbon cycle and climate parameters (see Table 1) are denoted with roman numbers I, II and III, while the application of AOGCM-specific forcing settings is denoted by a small Latin character "a", the application of standardized "full" forcings is denoted by "d" (with interim stages "b" and "c" being described in the companion paper Meinshausen et al., 2011, see Fig. 3 therein).

		Period 1: 1980–1999 vs. startyear			Period 2: 2090–2099 vs. 1980–1999			Period 3: 2100 vs. 2090–2099		
	n	AOGCM	Emulation		AOGCM	Emulation		AOGCM	Emulation	
Scenario			IIIa	IIId		IIIa	IIId		IIIa	IIId
COMMIT	16	0.56	0.59	0.50	0.58	0.48	0.57	–	0.00	0.01
SRESA1B	19	0.54	0.61	0.53	2.82	2.79	2.77	–	0.10	0.10
SRESA2	16	0.58	0.59	0.50	3.32	3.30	3.20	–	0.27	0.25
SRESB1	17	0.56	0.59	0.50	1.85	1.89	1.90	–	0.03	0.06

most notably for the MIROC3.2 (hires) model (see Table 4). The MAGICC result of around 6.0 °C is 2.0 °C higher than estimated by F&T. While the relatively short period (70 yrs) of available data for the 1pctto2× run limits the ability to make accurate estimates of the effective climate sensitivity of MIROC3.2(hires) from this 1pctto2× data set alone, the exceptionally high temperature projections for the SRES A1B and B1 scenarios for this model support our findings of a climate sensitivity around 6.0 °C (the calibration III result) rather than 3.9° results derived by F&T from the first 70 years of the idealized scenarios. An alternative explanation is that the SRES A1B and B1 forcing used by MIROC3.2(hires) could be exceptionally high compared to other AOGCMs as hypothesized by F&T. The $\Delta Q_{2\times}$ forcing for this model is, however, reported as rather low (see Table B1).

Four other climate sensitivities are estimated by the AR4 and Calibration I method to be higher than stated by F&T, namely those for CCSM3, MPI ECHAM5, GISS-EH, and GISS-ER. These models exhibit increasing effective climate sensitivities over time, so the method by F&T of deriving a fixed sensitivity over only the first 70 years of a 1pctto2× run will lead to an underestimate for the effective climate sensitivity on longer timescales and will hence result in higher forcing estimates. Lastly, the ECHO-G model is estimated to have a higher climate sensitivity than suggested by F&T possibly due to the ECHO-G heat uptake data used in the present study, which we suspect are erroneous. While the 1pctto2× scenario suggests a vertical ocean thermal diffusivity $Kz \geq 2\,\mathrm{cm^2\,s^{-1}}$, the best estimate for the vertical diffusivity under the SRES runs was more than five times smaller ($Kz = 0.43\,\mathrm{cm^2\,s^{-1}}$ – cf. Tables B3 and B1, B2). For the calibration III procedure, therefore, we excluded the ECHO-G 1pctto2× heat uptake data due to this inconsistency. When this was done, the climate sensitivity suggested by F&T is approximately confirmed (2.6 °C).

Using the calibration II procedure, the estimated climate sensitivity, $\Delta T_{2\times}$, is slightly lower for eight AOGCMs compared to calibration I results. This is largely explained by the increasing sensitivity over time in these models, a factor not accounted for in the calibration I method. The differences to the sensitivities estimated by F&T are largely reconciled by calibration II results. This is because F&T used the relatively low-forcing scenario segments up to doubled CO_2 concentrations to estimate the climate sensitivity.

The increases in effective climate sensitivities found in the present analyses confirm earlier results that the effective climate sensitivity seems often to be dependent on the climate state (see e.g. Murphy and Mitchell, 1995; Raper et al., 2001; Senior and Mitchell, 2000; Stouffer, 2004). For five AOGCMs the climate sensitivity estimate increased slightly when comparing the calibration I and calibration II results. For these AOGCMs, the data suggests no forcing dependent feedback factors (ξ=0). However, for some of these models, the calibration suggests an increasing climate sensitivity over time, parameterized by a heat exchange enhancement factor (μ>1). In this case, the transient effective sensitivity of the emulations up to doubled CO_2 concentrations is smaller than the equilibrium sensitivity at doubled pre-industrial CO_2 levels, so that this best-fit equilibrium sensitivity is estimated to be higher. Some of these calibrations to AOGCMs suggest (as well) a decreased heat uptake efficiency for higher warmings ($\frac{dK_z}{dT} \leq 0$). Thus, the warming can now be allowed to increase further compared to calibration I procedure for those AOGCMs, where an overestimation of heat uptake previously suggested a cooler warming response being optimal.

The climate sensitivity estimates under the calibration III procedure show only very minor systematic differences compared with the calibration II estimates, a slight decrease in the average sensitivity. This could be explained if the effective forcings or efficacies under the multi-gas scenarios

Table 4. Comparison of retrieved climate sensitivities for CMIP3 AOGCMs. The first column shows climate sensitivities estimated for the slab-ocean versions of the AOGCMs as given in Table 8.2 of Randall et al. (2007). The second column provides the climate sensitivities derived from the net climate feedbacks given by Forster and Taylor (2006), who use the method by Gregory et al. (2004) to retrieve feedbacks for idealized 1% CO_2 scenarios out to $2\times CO_2$. These climate feedbacks λ were converted to climate sensitivities $\Delta T_{2\times}$ using $\Delta T_{2\times}=\frac{\Delta Q_{2\times}}{\lambda}$, with the forcing $\Delta Q_{2\times}$ at doubled CO_2 concentrations taken from Table 2 in Forster and Taylor (2006), where available, and using 3.7 Wm^{-2} as default. The third column presents results for the MAGICC 4.2 calibration as done for IPCC AR4 , used as well in MAGICC5.3, and presented in Table S8.1 in Randall et al. (2007), which was methodological equivalent to the calibration method I presented here for MAGICC6. The fourth to sixth columns present this study's results using MAGICC6 under calibration exercises I, II and III (see Table 1). The last row provides the average climate sensitivities for each column.

AOGCM	IPCC AR4 Slab Ocean	Forster& Taylor	IPCC AR4 MAGICC 4.2/5.3	This Study (MAGICC6) Calibration:(I)	(II)	(III)
BCCR-BCM2.0	n/a	n/a	n/a	n/a	n/a	n/a
BCC-CM1	n/a	n/a	n/a	n/a	n/a	n/a
CCSM3	2.7	2.12	2.37	2.35	2.16[c]	2.14[c]
CGCM3.1 (T47)	3.4	2.97	3.02	3.13	3.34	2.97
CGCM3.1 (T63)	3.4[a]	3.63[b]	n/a	n/a	n/a	n/a
CNRM-CM3	n/a	2.45[b]	2.45	2.46	2.23[c]	2.98
CSIRO-Mk3.0	3.1	2.34[b]	2.21	2.18	2.17	2.24
ECHAM5/MPI-OM	3.4	3.58	3.86	3.95	3.46[c]	3.23[c]
ECHO-G	3.2	2.50[b]	3.01	3.10	3.10	2.63
FGOALS-g1.0	2.3[a]	1.98[b]	1.97	2.06	2.11	2.42[c]
GFDL-CM2.0	2.9	2.27[b]	2.35	2.41	2.32[c]	2.31[c]
GFDL-CM2.1	3.4	2.10[b]	2.28	2.34	2.19[c]	2.28[c]
GISS-AOM	n/a	n/a	n/a	n/a	n/a	n/a
GISS-EH	2.7	2.46	3.04	2.84	2.89	2.54
GISS-ER	2.7	2.22[b]	2.57	2.66	2.52[c]	2.26[c]
INM-CM3.0	2.1	2.34[b]	2.28	2.26	2.28	2.35
IPSL-CM4	4.4	3.80	3.83	3.93	4.03	4.15
MIROC3.2(H)	4.3	3.95	5.87	6.03	6.29	5.73
MIROC3.2(M)	4.0	3.73	3.93	4.12	4.15	4.00
MRI-CGCM2.3.2	3.2	3.01[b]	2.97	2.77	2.87	2.48
PCM	2.1	2.04[b]	1.88	1.94	1.83[c]	1.90
UKMO-HadCM3	3.3	3.02	3.06	3.10	3.13	3.21
UKMO-HadGEM1	4.4	2.74	2.63	3.09	2.90[c]	3.00[c]
AVERAGE	3.21	2.76	2.93	2.99	2.95	2.88

[a] These climate sensitivities were estimated from the coupled model versions available in the PCMDI database, while other values in this column denote reported equilibrium climate sensitivities of the slab-ocean model versions (Table 8.2, Randall et al., 2007).
[b] Derived feedback constant using a default 3.7 Wm^{-2} value for forcing at doubled CO_2 concentrations, given that no $\Delta Q_{2\times}$ value was available (see Table 2 of Forster and Taylor, 2006).
[c] Note that these calibrations II and III include a non-zero sensitivity parameter ξ, introducing a dependency of the sensitivity onto forcing. The effective climate sensitivity S therefore increases for forcings higher than twice pre-industrial CO_2 concentrations ($\Delta Q_{2\times}$) and decreases for lower forcings.

(SRES and COMMIT) were on average slightly overestimated, and/or, if forcings in the idealized CO_2 scenarios are underestimated. However, the potential over- or underestimations of forcings vary from AOGCM to AOGCM: in five out of the nineteen AOGCMs, multi-forcing runs are emulated warmer than the idealized scenarios, in contrast to seven AOGCMs, where idealized CO_2-only emulations are warmer (see Fig. 5f).

5 Calibrating MAGICC to carbon cycle models

The following section (Sect. 5.1) details the procedures for calibrating the MAGICC carbon cycle to ten of the eleven carbon cycle models that took part in the C^4MIP intercomparison project (Friedlingstein et al., 2006). Subsection 5.2 provides the respective calibration results.

5.1 Carbon cycle calibrating procedure

MAGICC's carbon cycle model (see Fig. A2) was calibrated in two steps. First, the climate sensitivity ($\Delta T_{2\times}$) for each

of the C^{4}MIP models was derived by prescribing each models' CO_2 concentrations (for runs that included temperature feedbacks on the carbon cycle) and calibrating MAGICC's climate sensitivity (using default MAGICC settings for other parameters) to obtain optimal (least squares) agreement with the C^{4}MIP temperature projection (see Table B4). The calibration period was the full period over which data were available, i.e., from model-specific starting years between 1765 and 1901 until 2100. Subsequently, MAGICC's main carbon cycle parameters were adjusted in order to optimally match the C^{4}MIP model-specific carbon fluxes and pool sizes for both the feedback and non-feedback cases (a total of 14 timeseries).

The initial MAGICC carbon fluxes were obtained from the available C^{4}MIP datasets, specifically the net primary productivity (NPP_{ini}) and total heterotrophic respiration ($\sum R_{ini}$ comprising R, Q_a and U). A partitioning (5:95) is assumed across all models for the initial carbon pool sizes of the detritus (D_{ini}), and soil box (S_{ini}), as only the aggregated dead carbon pool is provided for the C^{4}MIP models. C^{4}MIP's initial living carbon pool is equated to MAGICC's plant (P_{ini}) carbon pool. The start year for fertilization and temperature effects has been assumed to be the first year of the available C^{4}MIP dataseries (first model years ranging from 1765 to 1901; see Table B4.)

Using these initial conditions for carbon fluxes and pools, 13 MAGICC carbon cycle parameters were calibrated. The semi-automatic procedure involves 2000 randomly drawn parameter sets, each run once for the coupled (i.e., including temperature feedbacks) and once for the uncoupled (excluding temperature feedbacks) scenarios. The 'best match' parameter set was then chosen as initialization to an automated optimization procedure that fulfils a pre-selected error tolerance criterion after approximately another 1000 iterations. By adjusting the 13 MAGICC parameters, the procedure minimizes the weighted least-squares differences between MAGICC and 14 available time series; namely, the air-to-land, air-to-ocean, Net Primary Production (NPP), and heterotrophic respiration (R, Q_a and U) fluxes, as well as the living and dead carbon pools and CO_2 concentrations for both the with-feedback and no-feedback runs. See Appendix B for details.

The three ocean carbon cycle parameters involved in the calibration are: a) the CO_2 gas exchange rate k (yr^{-1}) between the atmosphere and the upper mixed ocean layer (Eq. A22); b) the temperature sensitivity α_T of the sea surface partial pressure (see Eq. A27); c) a scaling factor γ to scale the impulse response function r'_t for the inorganic carbon perturbation in the mixed layer (so that $r_t = \gamma r'_t/(\gamma r'_t + (1-r'_t))$ for times lower than one year and a constant scaling factor $\gamma' = (r_{t=1}/r'_{t=1})$ for longer response times, i.e., $r_t = \gamma' r'_t$ for t>1. The transition year for the scaling factor is chosen to match the transition time between the initial polynomial and subsequent exponential expression in the impulse response function representing the 3D-GFDL model. This particular two-part scaling of the impulse response function has been chosen to allow a linear scaling over medium and long timescales (cf. Fig. 7b in Joos et al., 1996), while ensuring a continuous impulse response function from year zero onwards.

The calibrated terrestrial carbon cycle model parameters determine the flux partitioning inside MAGICC; namely, the fraction of the plant box flux L going to the detritus box (ϕ_H), and the fraction of the detritus box outbound flux Q going to the soil box (Q_S). Comparison with the no-feedback runs allowed estimation of the fertilization parameters β_m and β_s, where β_m refers to whether a standard logarithmic formulation for fertilization is used (β_m=1.0), or the rectangular hyperbolic formulation (β_m=2.0), or any linear combination of these two formulations ($1.0<\beta_m<2.0$), cf. Wigley (2000). β_s denotes the fertilization factor itself (see Sect. A1.1, Eq. A15 and Eq. A20). The temperature feedback parameters σ_i of the carbon fluxes NPP, Q and U (cf. Fig. A2) were estimated by matching the difference between the with-feedback and no-feedback runs.

5.2 Carbon cycle calibration results

MAGICC has been successfully calibrated against ten of the C^{4}MIP carbon cycle models, as shown for atmospheric concentrations under the SRES A2 scenario (Fig. 6), and for all 14 available carbon pool and flux timeseries (Fig. B4). See also Table B4 for calibrated parameters. C^{4}MIP used CO_2 emissions in line with the illustrative SRES A2 scenario and treated net land-use emissions as lumped together with fossil and industrial sources, i.e., without taking into account changes in biospheric carbon pools due to deforestation. Given that not all C^{4}MIP models used exactly the same emissions, we used the model-specific emission timeseries for the calibration. The overall range across C^{4}MIP models of 2100 CO_2 concentrations (732 to 1025 ppm) is well matched by the emulations (732 ppm to 1012 ppm). For these with-feedback cases, differences in 2100 range between −23 and +2 ppm (RMSE = 10 ppm) for individual models. The match with the IPSL CM2C model in the with-feedback case is the least optimal (see Table B4). Over 2000 to 2100, the RMSE, averaged across all models, is very small, 3.5 ppm. For the no-feedback case, i.e., the runs in which the carbon cycle did not see changes in the climate, differences between emulations and the C^{4}MIP results range between −15 and +15 ppm for concentrations in 2100 (RMSE = 9 ppm, not shown in Table B4).

The additional uncertainty introduced by the emulations is more than an order of magnitude smaller than the C^{4}MIP inter-model spread. The average error (RMSE) introduced if one model's CO_2 concentration (with-feedback case) were simply approximated by another carbon cycle model's projection is 38.4 ppm over 2000 to 2100 (cf. 3.5 ppm for the emulations) and 128 ppm for 2100 concentrations (cf. 10 ppm for the emulations). While the optimization

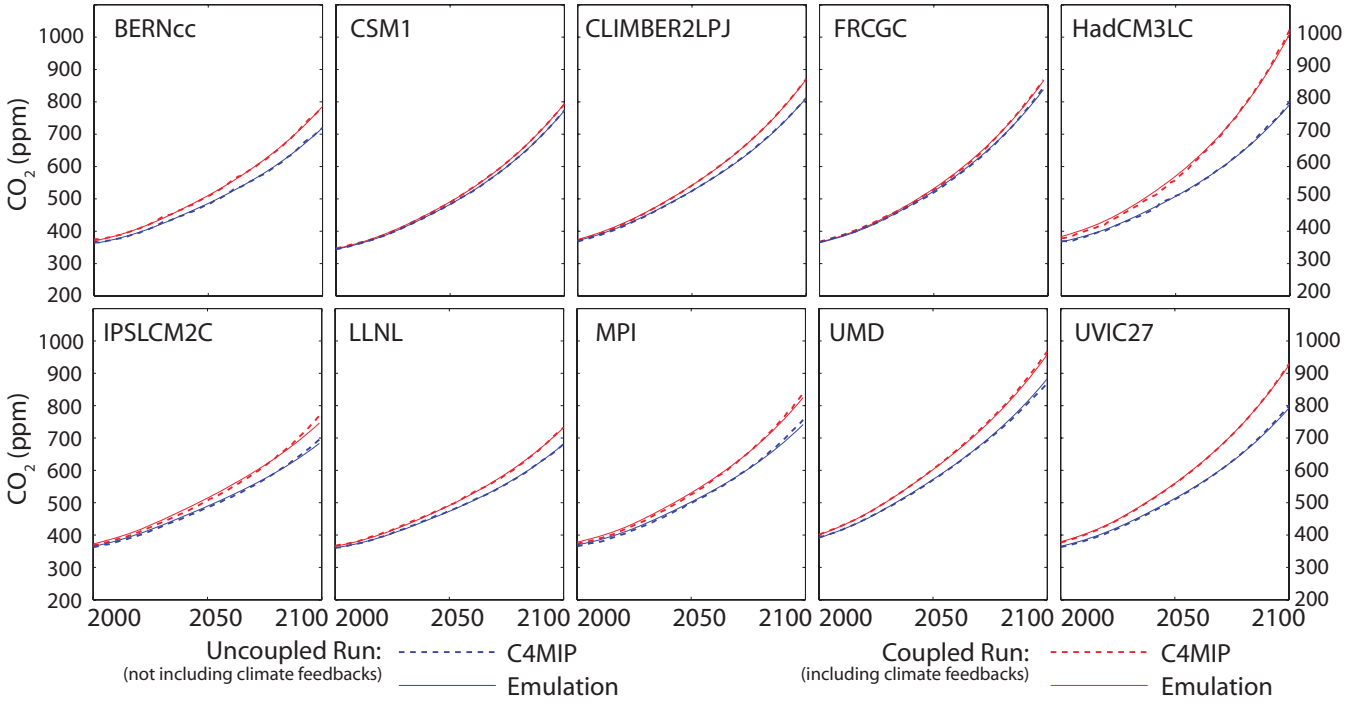

Fig. 6. Atmospheric CO_2 concentrations from 2000 to 2100 comparing C^4MIP carbon cycle model results (dashed) with the calibrated MAGICC6 (solid) model. Shown are the coupled (including climate feedbacks, red lines) and uncoupled (excluding climate feedbacks, blue lines) runs for the anthropogenic CO_2 emissions based on the IPCC SRES A2 scenario. See Fig. B4 in Appendix B for comparisons between emulations and C^4MIP models of other carbon fluxes and pools.

procedure placed the largest weights on fitting atmospheric CO_2 concentrations, the six other available C^4MIP time series, namely the terrestrial C-uptake, oceanic C-uptake, Net Primary Production (NPP), terrestrial living C-pool, terrestrial dead C-pool and the total respiration were also well matched for each model (see Fig. 6 and Fig. B4).

6 Discussion of MAGICC emulation limitations and justification for time-varying climate sensitivities

This section briefly summarizes some limitations that should be kept in mind when using the emulation results (Sect. 6.1). A possible alternative to emulating apparent time-varying climate sensitivities is briefly discussed (Sect. 6.2).

6.1 Limitations

Firstly, limitations arise in the original AOGCM and C^4MIP models themselves. Even if an emulation technique were able to perfectly match the mean responses over a wide range of scenarios, emulations can not mimic the 'real world' any better than the original models. Clearly, there are still significant developments to be expected in the realism of some aspects of both climate and carbon cycle models. The current carbon cycle models face substantial uncertainties, related to, for example, nitrogen fertilization, modeling of fire regimes, ocean circulation and chemistry, etc.

A second limitation arises from the incomplete quantitative knowledge of the forcings, including the forcing patterns, that each AOGCM was subject to, which limits our ability to correctly extract the characteristic AOGCM responses to those forcings. A consequence of this is that calibrations, even if perfect, may over- or under-estimate the climate response of an AOGCM under a given forcing depending on whether the estimated forcing is more or less than the actual AOGCM forcing. Suppose, for example, that an AOGCM includes the first indirect aerosol effect resulting in an effective radiative forcing of $-0.4\,\mathrm{Wm}^{-2}$ by 2005 relative to 1750, and that MAGICC attempts to emulate this AOGCM using the IPCC AR4 best-estimate effective forcing of $-0.7\,\mathrm{Wm}^{-2}$. MAGICC will then underestimate the temperature response of that AOGCM over the historical period, if the climate sensitivity were not adjusted upwards. The calibrated MAGICC sensitivity would then be too high. In the absence of detailed model-specific forcing information, there is no solution to this problem. Use of the independently derived forcings from F&T does not solve this problem, because these authors had to assume a climate sensitivity for each model in order to back out the forcings from temperature and heat-uptake time series. Their forcing results are thus naturally dependent on the assumed climate

sensitivities. This would lead to a somewhat circular analysis, if MAGICC attempted to back out the climate sensitivity using these forcings.

Thirdly, there are uncertainties as to how AOGCM and carbon cycle models would behave for scenarios outside the tested range. Although the SRES A2 AOGCM response was successfully emulated without having been used for calibration, the extrapolation of the calibration results to other emissions scenarios faces inherent uncertainties. This is even more important for the C^4MIP intercomparison, which was constrained to assessment of a single emissions scenario, SRES A2, and was limited to the period up to 2100 only. There are still considerable uncertainties in how the carbon cycle might react to, for example, peaking scenarios with increasing, then decreasing radiative temperatures and/or concentrations, and in long-term responses beyond 2100. Nevertheless, the choice of a relatively medium-high emissions scenario SRES A2 was useful, as it somewhat constrains the upper bound on the likely strength of the carbon cycle feedbacks during at least the 21st century. Future intercomparison projects would benefit from using a wider range of low and high emissions scenarios. While calibrations to C^4MIP have this limitation, we note that earlier (but similar) versions of MAGICC have successfully emulated other carbon cycle models over a wide range of scenarios (Wigley et al., 2007).

Fourthly, MAGICC6, by virtue of its model structure, must be limited to a subspace of the possible climate and carbon cycle responses. However, the model-by-model comparisons of key variables between the emulations and the original AOGCM and C^4MIP data did not reveal any major structural biases or limitations in MAGICC6 (see Figs. B1, B2, B3 and B4). This gives some confidence in applying MAGICC over a wide range of scenarios. Nevertheless, structural limitations might become apparent when attempting to emulate new scenarios outside the calibrated range.

Fifthly, MAGICC is limited to emulating temperature changes (and closely related variables such as oceanic heat uptake or thermal expansion). Precipitation changes, for example, are not modeled in MAGICC, even though we recognize that these are an important driver of climate change impacts. It is possible to extend MAGICC results using, for example, a pattern scaling approach (Santer et al., 1990; Mitchell, 2003) to obtain projections of the spatial patterns of temperature, sea-level pressure and precipitation, as in the MAGICC/SCENGEN software (Wigley, 2008, available here: www.cgd.ucar.edu/cas/wigley/magicc/). Future developments might refine and extend the capability to emulate variables of interest for global change analysis (see, e.g., Frieler et al., 2011).

Sixthly, the calibration procedure itself is subject to limitations. For example, due to the complexity of the AOGCM data, there may be errors in the data used for calibration (see the example noted above for the ECHO-G data in Sect. 4.2.4). Furthermore, for ocean heat uptake data we used both the net integrated ocean heat uptake as well as the total radiative flux at the top of the atmosphere, if the former data were not available, introducing small errors due to the effect of the land and cryosphere heat uptake.

Finally, there is the limitation that MAGICC is a simple model with a high level of parameterization. For example, in the C^4MIP carbon cycle calibration procedure the global-mean temperature is taken as the proxy for changes in the patterns of temperature, precipitation, cloudiness etc., which are the actual driving forces in more sophisticated carbon cycle models as well as in the real world. The skill of the emulations suggests that this is a reasonable approximation, at least for the assessed scenario.

6.2 Forcing adjustments as an alternative approach to time-variable climate sensitivity

This subsection discusses a potential alternative approach to explain and emulate phenomena currently represented by time-varying climate sensitivities. A number of recent studies suggest that there are relatively fast forcing adjustments following an increase in CO_2 forcing (Andrews and Forster, 2008; Gregory and Webb, 2008; Williams et al., 2008; Doutriaux-Boucher et al., 2009). Time-varying sensitivities might therefore be considered 'an artefact of using conventional forcings' (Williams et al., 2008).

Part of the debate may be a terminology issue, i.e., defining what is a forcing and what is regarded as a feedback. For example, cloud effects may follow tropospheric temperature and lapse rate adjustments, before noticeable changes are apparent in surface temperatures, and the question is: are these to be regarded as an indirect forcing effect or a feedback? Assuming that forcings and feedbacks could be freely redefined, then estimating a forcing value by regressing surface temperature changes against the top of the atmosphere radiative imbalance (Gregory et al., 2004) will, by construction, lead to a less time-variant diagnosed feedback parameter. However, a constant feedback that works well for medium to longer-time scales may come at the cost of not being able to emulate sufficiently well the first decades of climate response. In this respect, Williams et al. (2008) propose a time-varying forcing adjustment function, G, to emulate the initial response more closely, if the feedback parameter is assumed constant. Thus, although having gained the advantage of a simplified representation for longer-term idealized stabilization scenarios, emulating the response to more realistic scenarios with changing forcings might be equally cumbersome. Given that these forcing adjustments seem to be highly model-dependent (see e.g. Table 2 in Williams et al., 2008), the theoretical beauty of distinguishing between model-independent forcing and AOGCM-dependent feedbacks and inertia parameters is lost.

Of practical importance is whether alternative parameterizations will lead to improved emulation skill. Parameterizations based on short-term forcing adjustments could for example have substantial advantages, if they strongly differ

among forcing agents. Rather than taking into account efficacies of the conventional forcings (as this study does), one could take any fast semi-direct and indirect forcing adjustments explicitly into account. However, if time-varying climate sensitivities can be adequately calibrated in simpler models across a range of scenarios, then the emulation skill for the climate response would be no worse in comparison to an approach where fast forcing adjustments are taken into account, but with assumed constant climate sensitivity. In reality, both fast forcing adjustments and time-varying or climate-state dependent feedbacks will be at play and more research is needed to gain a better understanding of these phenomena.

Recently, Doutriaux-Boucher et al. (2009) pointed to a possible indirect forcing mechanism that is specific to CO_2: namely, the physiological response to increased CO_2 concentration by plants via stomatal conductance changes leading to a CO_2 forcing enhancement of roughly 10%. This is because the resulting reduced evaporation over land areas (in their analysis) induces a reduction in low cloud cover, which then has the forcing effect. If this is found to be a realistic effect, future versions of MAGICC will attempt to include it explicitly.

We anticipate that further studies into the fast and longer-term forcing adjustments will help to refine the optimal parameterizations required to emulate AOGCMs in the future.

7 Conclusions

In the preparation of the IPCC AR4, various resource constraints meant that only limited inter-model comparisons and syntheses were possible, both for AOGCMs and carbon cycle models. The question arises, therefore, as to how to make best use of a limited number of climate and carbon cycle model data sets, particularly with regard to their application to a wider range of emissions scenarios. A carefully calibrated model of lower complexity, which accounts realistically for key earth system components, and which is sufficiently flexible to emulate the large-scale results of more sophisticated models, is likely the most appropriate way. Thus, a simple coupled gas-cycle/climate model can function as an elaborate interpolation and (to a limited extent) extrapolation tool.

We have presented here an updated version of the simple gas-cycle/climate model, MAGICC, with enhanced representations of time-varying climate sensitivities, carbon cycle feedbacks, aerosol forcings and ocean heat uptake characteristics. MAGICC6 has been calibrated to 19 CMIP3 AOGCMs, and has been shown to closely reproduce the global-mean and hemispheric land/ocean temperature changes for both idealized and SRES multi-gas emissions scenarios. In our companion paper (Meinshausen et al., 2011), we show that for any given SRES emissions scenario, inter-model uncertainties in global-mean temperatures over the 21st century are roughly −30% to +40% – in line with the asymmetric shape of the −40% to +60% expert judgement based on multiple lines of evidence (cf. Knutti et al., 2008; Meehl et al., 2007). In comparison, the errors introduced by the emulations are substantially smaller, i.e. below 2.2% for the multi-model mean and, on average, 7% for the individual AOGCM models.

Similarly, emulations for the C^4MIP carbon cycle models were able to closely reproduce carbon pools, fluxes and atmospheric CO_2 concentrations. When climate feedbacks on the carbon cycle are included, MAGICC6 emulates 2100 CO_2 concentrations for individual C^4MIP models with a 10 ppm deviation (RMSE), which is more than an order of magnitude smaller than the inter-model range of variation (128 ppm RMSE). Thus, MAGICC6 is well suited for emulating both AOGCM and carbon cycle model responses for a variety of research purposes.

In addition, a simple model can help us to understand the behavior of and differences between AOGCMs. For example, MAGICC6 has shown here, confirming earlier studies, that the effective climate sensitivity varies over time in many AOGCMs when conventional forcing definitions are used. Possible alternative interpretations of AOGCM responses by relatively fast forcing adjustments are briefly discussed (see Sect. 6.2). As a specific example, we have shown that sensitivity estimates based on only the first 70 years of idealized 1% scenarios may be unrepresentative of longer time periods. We have also demonstrated that equilibrium sensitivities based on slab ocean versions of specific AOGCMs can differ noticeably from effective sensitivities derived from transient experiments (see Table 4).

In summary, simple coupled gas-cycle/climate models like MAGICC6, provided they are properly calibrated over a wide range of emissions scenarios against more complex climate and gas cycle models, serve as useful tools for the analysis, extension and synthesis of the results from large model intercomparison exercises. Furthermore, simple coupled models allow us to greatly expand the range of emissions scenarios that can be assessed by gas-cycle/AOGCMs, primarily because of the high computational demands of the complex models. Scientists, policy analysts and decision makers involved in the study and assessment of climate impacts, and adaptation and mitigation strategies, rely heavily on physical climate system projections that go beyond single-model, single-scenario studies. Emulation tools like MAGICC provide an important facility of benefit to both the research and stakeholder communities.

Appendix A

MAGICC model description

This appendix provides a detailed description of MAGICC6 and its different modules (see Fig. A1). Specifically,

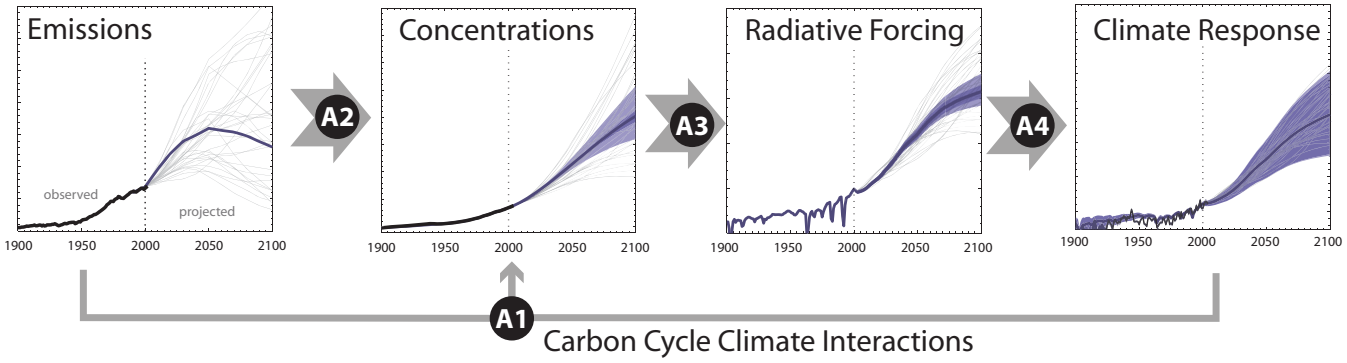

Fig. A1. Schematic overview of MAGICC calculations showing the key steps from emissions to global and hemispheric climate responses. Black circled numbers denote the sections in the Appendix describing the respective algorithms used.

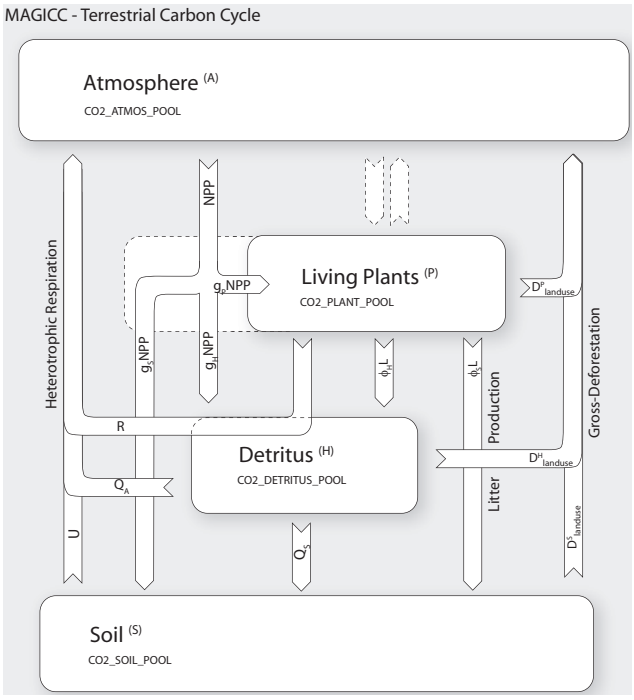

Fig. A2. The terrestrial carbon cycle component in MAGICC with its carbon pools and carbon fluxes. For description of the pools and fluxes, including the treatment of temperature feedbacks and CO_2 fertilization, see Sect. A1.1.

subsections describe MAGICC's carbon cycle (Sect. A1), the atmospheric-chemistry parameterizations and derivation of non-CO_2 concentrations (Sect. A2), radiative forcing routines (Sect. A3), and the climate module to get from radiative forcing to hemispheric (land and ocean, separately) and global-mean temperatures (Sect. A4), as well as oceanic heat uptake. Finally, details are provided on the implementation scheme for the upwelling-diffusion-entrainment ocean climate module (Sect. A5). A technical upgrade is that MAGICC6 has been re-coded in Fortran95, updated from previous Fortran77 versions. It should be noted that nearly all of the MAGICC6 code is directly based on the earlier MAGICC versions programmed by Wigley and Raper (1987; 1992; 2001).

A1 The Carbon cycle

A change in atmospheric CO_2 concentration, C, is determined by CO_2 emissions from fossil and industrial sources (E_{foss}), other directly human-induced CO_2 emissions from or removals to the terrestrial biosphere (E_{lu}), the contribution from oxidized methane of fossil fuel origin (E_{fCH_4}), the flux due to ocean carbon uptake (F_{ocn}) and the net carbon uptake or release by the terrestrial biosphere (F_{terr}) due to CO_2 fertilization and climate feedbacks. As in the C^{4}MIP generation of carbon cycle models, no nitrogen or sulphur deposition effects on biospheric carbon uptake are included here (Thornton et al., 2009). Hence, the budget Eq. (A1) for a change in atmospheric CO_2 concentrations is:

$$\frac{\Delta C}{\Delta t} = E_{\text{foss}} + E_{\text{lu}} + E_{\text{fCH}_4} - F_{\text{ocn}} - F_{\text{terr}} \qquad \text{(A1)}$$

A1.1 Terrestrial carbon cycle

The terrestrial carbon cycle follows that in Wigley (1993), in turn is based on Harvey (1989). It is modeled with three boxes, one living plant box P (see Fig. A2) and two dead biomass boxes, of which one is for detritus H and one for organic matter in soils S. The plant box comprises woody material, leaves/needles, grass, and roots, but does not include the rapid turnover part of living biomass, which can be assumed to have a zero lifetime on the timescales of interest here (dashed extension of plant box P in Fig. A2). Thus, a fraction of gross primary product (GPP) cycles through the plant box directly back to the atmosphere due to autotrophic respiration and can be ignored (dashed arrows). Only the remaining part of GPP, namely the net primary production

(NPP) is simulated. The NPP flux is channeled through the "rapid turnover" part of the plant box and partitioned into carbon fluxes to the remainder plant box (default g_P=35%), detritus (g_H=60%) and soil box (g_S=1-g_P-g_H=5%).

The plant box has two decay terms, litter production L and a part of gross deforestation D^P_{gross}. Litter production is partitioned to both the detritus (ϕ_H=98%) and soil box (ϕ_S=1-ϕ_H=2%). Thus, the mass balance for the plant box is:

$$\Delta P/\Delta t = g_P \text{NPP} - R - L - D^P_{\text{gross}} \quad \text{(A2)}$$

The detritus box has sources from litter production ($\phi_H L$) and sinks to the atmosphere due to land use (D^H_{lu}), non-land use related oxidation (Q_A), and a sink to the soil box (Q_S). The mass balance for the detritus box is thus

$$\Delta H/\Delta t = g_H \text{NPP} + \phi_H L - Q_A - Q_S - D^H_{\text{lu}} \quad \text{(A3)}$$

The soil box has sources from litter production (ϕ_SL), the detritus box (Q_S) and fluxes to the atmosphere due to land use (D^S_{gross}), and non-land use related oxidation (U). The mass balance for the soil box is thus

$$\Delta S/\Delta t = g_S \text{NPP} + \phi_S L + Q_S - U - D^S_{\text{lu}} \quad \text{(A4)}$$

The decay rates (L, Q and U) of each pool are assumed to be proportional to pool's box masses P, H and S, respectively. The turnover times τ_P, τ_H and τ_S are determined by the initial steady-state conditions for box sizes and fluxes.

$$L_0 = P_0/\tau^P_0 \quad \text{(A5)}$$
$$Q_0 = H_0/\tau^H_0 \quad \text{(A6)}$$
$$U_0 = S_0/\tau^S_0 \quad \text{(A7)}$$

Constant relaxation times τ ensure that the box masses will relax back to their initial sizes if perturbed by a one-off land use change-related carbon release or uptake – assuming no changes in fertilization and temperature feedback terms. This relaxation acts as an effective regrowth term so that deforestation ΣD_{gross}=$D^P_{\text{gross}} + D^H_{\text{gross}} + D^S_{\text{gross}}$ represents the gross land use emissions, related to net land use emissions E_{lu} by regrowth ΣG=$G^P + G^H + G^S$

$$\Sigma D_{\text{gross}} - \Sigma \text{G} = E_{\text{lu}} \quad \text{(A8)}$$
$$D^P_{\text{gross}} - \text{G}^P = d_P E_{\text{lu}} \quad \text{(A9)}$$
$$D^H_{\text{gross}} - \text{G}^H = d_H E_{\text{lu}} \quad \text{(A10)}$$
$$D^S_{\text{gross}} - \text{G}^S = d_S E_{\text{lu}} \quad \text{(A11)}$$

Gross land-use related emissions might be smaller (compared to a case where relaxation times are assumed constant) as some human land use activities, e.g. deforestation, can lead to persistent changes of the ecosystems over the time scales of interest, thereby preventing full regrowth to the initial state P_0, H_0 or S_0. A factor ψ is used to denote the fraction of gross deforestation that does not regrow ($0 \le \psi \le 1$). Thus, the relaxation times τ are made time-dependent according to the following equation:

$$\tau^P(t) = \left(P_0 - \psi \int_0^t d_P E_{\text{lu}}(t')dt'\right)/L_0 \quad \text{(A12)}$$
$$\tau^H(t) = \left(H_0 - \psi \int_0^t d_H E_{\text{lu}}(t')dt'\right)/Q_0 \quad \text{(A13)}$$
$$\tau^S(t) = \left(S_0 - \psi \int_0^t d_S E_{\text{lu}}(t')dt'\right)/U_0 \quad \text{(A14)}$$

Formulation for CO_2 fertilization

CO_2 fertilization indicates the enhancement in net primary production (NPP) due to elevated atmospheric CO_2 concentration. As described in Wigley (2000), there are two common forms used in simple models to simulate the CO_2 fertilization effect: (a) the logarithmic form (fertilization parameter β_m=1) and (b) the rectangular hyperbolic or sigmoidal growth function (β_m=2) (see e.g. Gates, 1985). The rectangular hyperbolic formulation provides more realistic results for both low and high concentrations so that NPP does not rise without limit as CO_2 concentrations increase. Previous MAGICC versions include both formulations, but used the second as default. The code now allows use of a linear combination of both formulations ($1 \le \beta_m \le 2$).

The classic logarithmic fertilization formulation calculates the enhancement of NPP as being proportional to the logarithm of the change in CO_2 concentrations C above the preindustrial level C_0:

$$\beta_{\log} = 1 + \beta_s \ln(\text{C}/\text{C}_0) \quad \text{(A15)}$$

The rectangular hyperbolic parameterization for fertilization is given by

$$N = \frac{C - C_b}{1 + b(C - C_b)} = \frac{N_0(1 + b(C_0 - C_b))(C - C_b)}{(C_0 - C_b)(1 + b(C - C_b))} \quad \text{(A16)}$$

where N_0 is the net primary production and C_0 the CO_2 concentrations at pre-industrial conditions, C_b the concentration value at which NPP is zero (default setting: C_b=31 ppm, see Gifford, 1993).

For better comparability with models using the logarithmic formulation, following Wigley (2000), the CO_2 fertilization factor β_s expresses the NPP enhancement due to a CO_2 increase from 340 ppm to 680 ppm, valid under both formulations. Thus, MAGICC first determines the NPP ratio r for a given β_s fertilization factor according to:

$$r = \frac{N(680)}{N(340)} = \frac{N_0(1 + \beta_s \ln(680/C_0))}{N_0(1 + \beta_s \ln(340/C_0))} \quad \text{(A17)}$$

Following from here, b in Eq. (A16) is determined by

$$b = \frac{(680 - C_b) - r(340 - C_b)}{(r-1)(680 - C_b)(340 - C_b)} \quad \text{(A18)}$$

which can in turn be used in Eq. (A16) to calculate the effective CO_2 fertilization factor β_{sig} at time t as

$$\beta_{sig}(t) = \frac{1/(C_0 - C_b) + b}{1/(C(t) - C_b) + b} \quad \text{(A19)}$$

MAGICC6 allows for an increased flexibility, as any linear combination between the two fertilization parameterizations can be chosen ($1 \leq \beta_m \leq 2$), so that the effective fertilization factor β_{eff} is given by:

$$\beta_{eff}(t) = (2 - \beta_m)\beta_{log} + (\beta_m - 1)\beta_{sig} \quad \text{(A20)}$$

The CO_2 fertilization effect affects NPP so that β_{eff}=NPP/NPP$_0$. MAGICC's terrestrial carbon cycle furthermore applies the fertilization factor to one of the heterotrophic respiration fluxes R that cycles through the detritus box, which makes up 18.5% of the total heterotrophic respiration ($\sum$R=$R + U_a + Q$) at the initial steady-state.

Temperature effect on respiration and decomposition

Global-mean temperature increase is taken as a proxy for climate-related impacts on the carbon cycle fluxes induced by regional temperature, cloudiness or precipitation regime changes. Those impacts are commonly referred to as "climate feedbacks on the carbon cycle", or simply, "carbon cycle feedbacks". Here, the terrestrial carbon fluxes NPP, and the heterotrophic respiration/decomposition fluxes R, Q and U are scaled assuming an exponential relationship,

$$F_i(t) = F_i'(t) \cdot \exp(\sigma_i \Delta T(t)) \quad \text{(A21)}$$

where $\Delta T(t)$ is the temperature above a reference year level, e.g. for 1990 or 1900, and F_i' (F_i) stands for the (feedback-adjusted) fluxes NPP, R, Q and U. The parameters σ_i (K^{-1}) are their respective sensitivities to temperature changes. In order to model the actual change in Q and U, the relaxation times τ for the detritus and soil pool are adjusted, respectively. Land use CO_2 emissions in many emissions scenarios (e.g. SRES, Nakicenovic and Swart, 2000) reflect the net directly human-induced emissions. At each time-step, the gross land use emissions are subtracted from the plant, detritus and soil carbon pools. The difference between net and gross land use emissions is the CO_2 uptake due to regrowth. Thus, a separation between directly human-induced (deforestation-related) emissions and indirectly human-induced effects (regrowth) on the carbon cycle is required. As both regrowth and the temperature sensitivity are modeled by adjusting the turnover times, a no-feedback case is computed separately, retrieving the regrowth, then calculating the feedback-case including the formerly calculated regrowth.

A1.2 Ocean carbon cycle

For modeling the perturbation of ocean surface dissolved inorganic carbon, an efficient impulse response substitute for the 3D-GFDL model Sarmiento et al. (1992) is incorporated into MAGICC. The applied analytical representation of the pulse response function is provided in Appendix A.2.2 of Joos et al. (1996).

The sea-to-air flux F_{ocn} is determined by the partial pressure differential for CO_2 between the atmosphere C and surface layer of the ocean ρCO_2

$$F_{ocn} = k(C - \rho CO_2) \quad \text{(A22)}$$

where k is the global average gas exchange coefficient (see Joos et al., 2001). This exchange coefficient is here calibrated to the individual C^{4}MIP carbon cycle models (default value $(7.66\,\text{yr})^{-1}$). The perturbation in dissolved inorganic carbon in the surface ocean $\Delta\Sigma CO_2(t)$ at any point t in time is obtained from the convolution integral of the mixed layer impulse response function r_s and the net air-to-sea flux F_{ocn}:

$$\Delta\Sigma CO_2(t) = \frac{c}{hA}\{\int_{t_0}^{t} F_{ocn}(t') r_s(t - t') dt')\} \quad \text{(A23)}$$

The impulse response function r_s is given for the time immediately after the impulse injection (<1 yr) by (see Appendix A.2.4 of Joos et al., 1996):

$$r_s(t) = 1.0 - 2.2617t + 14.002t^2 - 48.770t^3 + 82.986t^4 - 67.527t^5 + 21.037t^6 \quad \text{(A24)}$$

and for $t \geq 1$ year is given by:

$$r_s(t) = \sum_{i=1}^{6} \gamma_i e^{-\tau_i t} \quad \text{(A25)}$$

with the partitioning γ and relaxation τ coefficients:

$$\gamma = \begin{bmatrix} 0.01481 \\ 0.70367 \\ 0.24966 \\ 0.066485 \\ 0.038344 \\ 0.019439 \end{bmatrix}, \tau = \begin{bmatrix} 0 \\ 1/0.70177 \\ 1/2.3488 \\ 1/15.281 \\ 1/65.359 \\ 1/347.55 \end{bmatrix} \quad \text{(A26)}$$

The relationship between the perturbation to dissolved inorganic carbon $\Delta\Sigma CO_2(t)$ and ocean surface partial pressures $\Delta\rho CO_2(T_0)$ (expressed in ppm or μatm) at the preindustrial temperature level T_0 is given by Eq. (A23) in Joos et al. (2001). Furthermore, the temperature-sensitivity effect on CO_2 solubility and hence oceanic carbon uptake is parameterized with a simple exponential expression. The modeled partial pressure $\rho CO_2(t)$ increases with sea surface temperatures according to:

$$\rho CO_2(t) = [\rho CO_2(t_0) + \Delta\rho CO_2(T_0)]\exp(\alpha_T \Delta T) \quad \text{(A27)}$$

where α_T (default α_T=0.0423 K^{-1}) is the sensitivity of the sea surface partial pressure to changes in temperature (ΔT) away from the preindustrial level (see Eq. A24 in Joos et al., 2001, based on Takahashi et al., 1993).

A2 Non-CO_2 concentrations

This section provides the formulas used to convert emissions to concentrations, while Sect. A3 provides details on the derivation of radiative forcings.

A2.1 Methane

Natural emissions of methane are inferred by balancing the budget for a user-defined historical period, e.g. from 1980–1990, so that

$$E^n_{\text{ø}} = \theta(\Delta C_{\text{ø}} - C_{\text{ø}'}/\tau_{\text{tot}}) - E^f_{\text{ø}} - E^b_{\text{ø}} \quad \text{(A28)}$$

where $E^n_{\text{ø}}$, $E^f_{\text{ø}}$ and $E^b_{\text{ø}}$ are the average natural, fossil and land use related emissions, respectively; θ is the conversion factor between atmospheric concentrations and mass loadings. $C_{\text{ø}'}$ (and $\Delta C_{\text{ø}}$) are the average (annual changes in) concentrations. The net atmospheric lifetime τ_{tot} in the case of methane consists of the atmospheric chemical lifetime and lifetimes that characterize the soil and other (e.g. stratospheric) sink components according to

$$\frac{1}{\tau_{\text{tot}}} = \frac{1}{\tau_{\text{tropos}}} + \frac{1}{\tau_{\text{soil}}} + \frac{1}{\tau_{\text{other}}} \quad \text{(A29)}$$

The feedback of methane on tropospheric OH and its own lifetime follows the results of the OxComp work (tropospheric oxidant model comparison) (see Ehhalt et al., 2001, in particular Table 4.11), which provides simple parameterizations for simulating complex three-dimensional atmospheric chemistry models. As default, tropospheric OH abundances are assumed to decrease by 0.32% for every 1% increase in CH_4. The change in tropospheric OH abundances is thus modeled as:

$$\Delta\ln(\text{tropOH}) = S^{\text{OH}}_{\text{CH}_4}\Delta\ln(\text{CH}_4) + S^{\text{OH}}_{\text{NO}_x}E_{\text{NO}_x} + S^{\text{OH}}_{\text{CO}}E_{\text{CO}} + \text{S}^{\text{OH}}_{\text{VOC}}\text{E}_{\text{VOC}} \quad \text{(A30)}$$

where S^{OH}_x is the sensitivity of tropospheric OH towards CH_4, NOx, CO and VOC, with default values of −0.32, +0.0042, −1.05e-4 and −3.15e-4, respectively. Increases in tropospheric OH abundances decrease the tropospheric lifetime τ' of methane (default 9.6 yrs^{-1}), which is approximated as a simple exponential relationship

$$\tau'_{\text{CH}_4,\text{tropos}} = \tau^0_{\text{CH}_4,\text{tropos}} \exp^{\Delta\ln(\text{tropOH})} \quad \text{(A31)}$$

Approximating the temperature sensitivity of the net effect of tropospheric chemical reaction rates, the tropospheric lifetime of CH_4 is adjusted:

$$\tau_{\text{CH}_4,\text{tropos}} = \frac{\tau^0_{\text{CH}_4,\text{tropos}}}{\frac{\tau^0_{\text{CH}_4,\text{tropos}}}{\tau'_{\text{CH}_4,\text{tropos}}} + S_{\tau_{\text{CH}_4}}\Delta T} \quad \text{(A32)}$$

where $S_{\tau_{\text{CH}_4}}$ is the temperature sensitivity coefficient (default $S_{\tau_{\text{CH}_4}}$=3.16e-2°C^{-1}) and ΔT is the temperature change above a user-definable year, e.g. 1990.

A2.2 Nitrous oxide

As for methane, natural nitrous oxide emissions are estimated by a budget Eq. A28. For nitrous oxide however, the average concentrations $C_{\text{ø}'}$=$C_{\text{ø}-3}$ are taken for a period shifted by 3 years to account for a three year delay of transport of tropospheric N_2O to the main stratospheric sink. The feedback of the atmospheric burden C_{N_2O} of nitrous oxide on its own lifetime is approximated by:

$$\tau_{\text{N}_2\text{O}} = \tau^0_{\text{N}_2\text{O}}\left(\frac{\text{C}_{\text{N}_2\text{O}}}{\text{C}^0_{\text{N}_2\text{O}}}\right)^{S_{\tau_{\text{N}_2\text{O}}}} \quad \text{(A33)}$$

where $S_{\tau_{N_2O}}$ is the sensitivity coefficient (default $S_{\tau_{N_2O}}$=−5e-2) and the superscript "0" indicates a preindustrial reference state.

A2.3 Tropospheric aerosols

Due to their short atmospheric residence time, changes in hemispheric abundances of aerosols are approximated by changes in their hemispheric emissions. Historical emissions of tropospheric aerosols are extended into the future either by emissions scenarios (SO_x, NO_x, CO) or, if scenario data are not available, with proxy emissions, e.g. using CO as a proxy emission for OC and BC. As with many other emissions scenarios, the harmonized IPCC SRES scenarios do not provide black (BC) and organic carbon (OC) emissions. Hence, various *ad-hoc* scaling approaches have been applied, often scaling BC and OC synchronously (Takemura et al., 2006), sometimes linearly with CO_2 emissions. The MESSAGE emissions scenario modeling group is one of the few explicitly including BC and OC emissions in their multi-gas emissions scenarios (Rao et al., 2005; Rao and Riahi, 2006). By analyzing MESSAGE scenarios, a scaling factor was derived for this study in relation to carbon monoxide emissions (CO), varying linearly in time to 0.4 by 2100 relative to current BC/CO or OC/CO emission ratios.

A2.4 Halogenated gases

The derivation of concentrations of halogenated gases controlled under either the Kyoto or Montreal Protocol assumes time-variable lifetimes. The net atmospheric lifetime τ_i of

each halogenated gas is calculated by summing the inverse lifetimes related to stratospheric, OH-related and other sinks. Stratospheric lifetimes are assumed to decrease 15% per degree of global mean surface temperature warming, due to an increased Brewer-Dobson circulation (Butchart and Scaife, 2001). Tropospheric OH-related losses are scaled by parameterized changes in OH-abundances, matching the respective changes in the lifetime of methane. The concentration $C_{t,i}$ for the beginning of each year t is updated, using a central differencing formulation, according to:

$$C_{t+1,i}=\tau_i E_{t,i}\frac{\rho_{\text{atm}}}{m_{\text{atm}}\mu_i}(1-e^{\frac{-1}{\tau_i}})+C_{t,i}(1-e^{\frac{-1}{\tau_i}}) \quad \text{(A34)}$$

where $E_{t,i}$ is the average emissions of gas i through year t, $C_{t,i}$ the atmospheric concentration of gas i in year t, ρ_{atm} the average density of air, m_{atm} the total mass of the atmosphere (Trenberth and Guillemot, 1994), and μ_i is the mass per mol of gas i. For hydrogenated halocarbons, the tropospheric OH-related lifetimes are assumed to vary in proportion to the changes in methane lifetime.

A3 Radiative forcing

The following section highlights the key parameterizations used for estimating the radiative forcing due to human-induced changes in greenhouse gas concentrations, tropospheric ozone and aerosols. The radiative forcing applied in MAGICC is in general the forcing at tropopause level after stratospheric temperature adjustment. Efficacies of the forcings, as discussed by Hansen et al. (2005) and Meehl et al. (2007) can be applied.

A3.1 Carbon dioxide

Taking into account the "saturation" effect of CO_2 forcing, i.e., the decreasing forcing efficiency for a unit increases of CO_2 concentrations with higher background concentrations, the first IPCC Assessment (Shine et al., 1990) presented the simplified expression of the form:

$$\Delta Q_{\text{CO}_2}=\alpha_{\text{CO}_2}\ln(\text{C}/\text{C}_0) \quad \text{(A35)}$$

where ΔQ_{CO_2} is the adjusted radiative forcing by CO_2 (Wm^{-2}) for a CO_2 concentration C (ppm) above the pre-industrial concentration C_0 (278 ppm). This expression proved to be a good approximation, although the scaling parameter α_{CO_2} has since been updated to a best-estimate of 5.35 Wm^{-2} ($=\frac{3.71}{\ln(2)}\,\text{Wm}^{-2}$) (Myhre et al., 1998), used as default in MAGICC. When applying AOGCM-specific CO_2 forcing, α_{CO_2} is set to:

$$\alpha_{\text{CO}_2}=\frac{\Delta Q_{2\times}}{\ln(2)} \quad \text{(A36)}$$

A3.2 Methane and nitrous oxide

Methane and nitrous oxide have overlapping absorption bands so that higher concentrations of one gas will reduce the effective absorption by the other and vice versa. This is reflected in the standard simplified expression for methane and nitrous oxide forcing, ΔQ_{CH4} and ΔQ_{N2O}, respectively (see Ramaswamy et al., 2001; Myhre et al., 1998):

$$\Delta Q_{\text{CH}_4}=\alpha_{\text{CH}_4}(\sqrt{\text{C}_{\text{CH}_4}}-\sqrt{\text{C}^0_{\text{CH}_4}}) - f(\text{C}_{\text{CH}_4},\text{C}^0_{\text{N}_2\text{O}})-\text{f}(\text{C}^0_{\text{CH}_4},\text{C}^0_{\text{N}_2\text{O}}) \quad \text{(A37)}$$

$$\Delta Q_{\text{N}_2\text{O}}=\alpha_{\text{N}_2\text{O}}(\sqrt{\text{C}_{\text{N}_2\text{O}}}-\sqrt{\text{C}^0_{\text{N}_2\text{O}}}) - f(\text{C}^0_{\text{CH}_4},\text{C}_{\text{N}_2\text{O}})-f(\text{C}^0_{\text{CH}_4},\text{C}^0_{\text{N}_2\text{O}}) \quad \text{(A38)}$$

where the overlap is captured by the function

$$f(\text{M},\text{N})=0.47\ln(1+0.6356(\frac{\text{MN}}{10^6})^{0.75} +0.007\frac{\text{M}}{10^3}(\frac{\text{MN}}{10^6})^{1.52}) \quad \text{(A39)}$$

with M and N being CH_4 and N_2O concentrations in ppb. For methane, an additional forcing factor due to methane-induced enhancement of stratospheric water vapor content is included. This enhancement is assumed to be proportional to (default β=15%) the "pure" methane radiative forcing, i.e., without subtraction of N_2O absorption band overlaps:

$$\Delta Q^{\text{stratoH2O}}_{\text{CH}_4}=\beta\alpha_{\text{CH}_4}(\sqrt{\text{C}_{\text{CH}_4}}-\sqrt{\text{C}^0_{\text{CH}_4}}). \quad \text{(A40)}$$

A3.3 Tropospheric ozone

From the tropospheric ozone precursor emissions and following the updated parameterizations of OxComp as given in footnote a of Table 4.11 in Ehhalt et al. (2001), the change in hemispheric tropospheric ozone concentrations (in DU) is parameterized as:

$$\Delta(\text{trop}\text{O}_3)=\text{S}^{\text{O}_3}_{\text{CH}_4}\Delta\ln(\text{CH}_4) +\text{S}^{\text{O}_3}_{\text{NO}x}E_{\text{NO}x}+\text{S}^{\text{O}_3}_{\text{CO}}E_{\text{CO}}+\text{S}^{\text{O}_3}_{\text{VOC}}E_{\text{VOC}} \quad \text{(A41)}$$

where $S^{\text{O}_3}_{\text{x}}$ are the respective sensitivity coefficients of tropospheric ozone to methane concentrations and precursor emissions. The radiative forcing is then approximated by a linear abundance to forcing relationship so that $\Delta Q_{\text{trop}\text{O}_3}=\alpha_{\text{trop}\text{O}3}\Delta(\text{trop}\text{O}_3)$ with $\alpha_{\text{trop}\text{O}3}$ being the radiative efficiency factor (default 0.042).

A3.4 Halogenated gases

The global-mean radiative forcing $\Delta Q_{t,i}$ of halogenated gases is simply derived from their atmospheric concentrations C (Sect. A2.4) and radiative efficiencies ϱ_i (following Ehhalt et al., 2001, Table 4.11).

$$\Delta Q_{t,i}=\varrho_i(C_{t,i}-C_{0,i}) \quad \text{(A42)}$$

The land-ocean forcing contrast in each hemisphere for halogenated gases is assumed to follow the one Hansen et al. (2005) estimated for CFC-11. The hemispheric forcing contrast is dependent on the lifetime of the gas. For short-lived gases (<1 yr) the hemispheric forcing contrast is assumed to equal the time-variable hemispheric emission ratio. For longer lived gases (default >8 yrs), the hemispheric forcing contrast is assumed to equal the one from CFC-11 with linear scaling in between these two approaches for gases with a medium lifetime.

A3.5 Stratospheric ozone

Depletion of the stratospheric ozone layer causes a negative global-mean radiative forcing ΔQ_t. The depletion and hence radiative forcing is assumed to be dependent on the equivalent effective stratospheric chlorine (EESC) concentrations as follows:

$$\Delta Q_t = \eta_1(\eta_2 \times \Delta \mathrm{EESC}_t)^{\eta_3} \quad \text{(A43)}$$

where η_1 is a sensitivity scaling factor (default -4.49e-4 Wm^{-2}), $\Delta\mathrm{EESC}_t$ the EESC concentrations above 1980 levels (in ppb), the factor η_2 equals $\frac{1}{100}$ (ppb^{-1}) and η_3 is the sensitivity exponent (default 1.7).

EESC concentrations are derived from the modeled concentrations of 16 ozone depleting substances controlled under the Montreal Protocol, their respective chlorine and bromine atoms, fractional release factors and a bromine versus chlorine ozone depletion efficiency (default 45) (Daniel et al., 1999).

A3.6 Tropospheric aerosols

The direct effect of aerosols is approximated by simple linear forcing-abundance relationships for sulfate, nitrate, black carbon and organic carbon. Time-variable hemispheric abundances of these short-lived aerosols are in turn approximated by their hemispheric emissions, justifiable because of their very short lifetimes. The ratio of direct forcing over land and ocean areas in each hemisphere is taken from Hansen et al. (2005) (available at http://data.giss.nasa.gov/efficacy/). Specifying the direct radiative forcing patterns for one particular year, and knowing the hemispheric emissions in that year, allows us to define the future forcing as a function of future emissions.

The indirect radiative forcing, formerly modeled as dependent on SO_x abundances only (Wigley, 1991a), is now estimated by taking into account time-series of sulfate, nitrate, black carbon and organic carbon optical thickness:

$$\Delta Q_{\mathrm{Alb,i}} = r \times \mathrm{P}_{\mathrm{Alb,i}} \times \log\left(\frac{\sum_g w_g \mathrm{N}_{g,i}}{\sum_g w_g \mathrm{N}^0_{g,i}}\right) \quad \text{(A44)}$$

where $\Delta Q_{\mathrm{Alb,i}}$ is the first indirect aerosol forcing in the four atmospheric boxes i, representing land and ocean areas in each hemisphere; $\mathrm{P}_{\mathrm{Alb}}$ is the four-element pattern of aerosol indirect effects related to albedo (Twomey, 1977) in a reference year. The second indirect effect on cloud cover changes (Albrecht, 1989) is modeled equivalently – using a reference year pattern $\mathrm{P}_{\mathrm{Cvr,i}}$. The respective default patterns are derived from data displayed in Fig. 13 of Hansen et al. (2005). The scaling factor r allows one to specify a global-mean first or second indirect forcing for a specific reference year. The time-variable number concentrations of soluble aerosols $\mathrm{N}_{g,i}$ relative to their pre-industrial level in each hemisphere $\mathrm{N}^0_{g,i}$ are normed to unity in that reference year. This is done separately for sulfates, nitrates, black carbon and organic carbon. For the latter, the differential solubility from industrial (fossil fuel) and biomass burning sources is taken into account (default solubility ratio 0.6/0.8) (Hansen et al., 2005). The default contribution shares w_g of the individual aerosol types g to the indirect aerosol effect were assigned to reflect the preliminary results by Hansen et al. (2005), namely 36% for sulfates, 36% for organic carbon, 23% for nitrates and 5% for black carbon. Note, however, that these estimates of the importance of non-SOx aerosol contributions are very uncertain, not least because the solubility, e.g. for organic carbon and nitrates have large uncertainties. The number concentrations $N_{g,i}$ are here approximated by historical optical thickness estimates (as provided on http://data.giss.nasa.gov/efficacy/ see as well Supplement) and extrapolated into the future by scaling with hemispheric emissions. The general logarithmic relation between number concentrations and forcing is based on the findings by Wigley and Raper (1992); Wigley (1991a); Gultepe and Isaac (1999) and as well used in Hansen et al. (2005).

A4 From forcing to temperatures: the upwelling-diffusion climate model

In the early stages, MAGICC's climate module evolved from the simple climate model introduced by Hoffert et al. (1980). MAGICC's atmosphere has four boxes with zero heat capacity, one over land and one over ocean for each hemisphere. The atmospheric boxes over the ocean are coupled to the mixed layer of the ocean hemispheres, with a set of n-1 vertical layers below (see Fig. A1). The heat exchange between the oceanic layers is driven by vertical diffusion and advection. In the previous model versions, the ocean area profile is uniform with depth and the corresponding downwelling is modeled as a stream of polar sinking water from the top mixed layer to the bottom layer. In this study, an updated upwelling-diffusion-entrainment (UDE) ocean model is implemented with a depth-dependent ocean area (from HadCM2). For simplicity, the following equations govern the uniform area upwelling-diffusion version of the model. Section A5 provides details on the UDE algorithms.

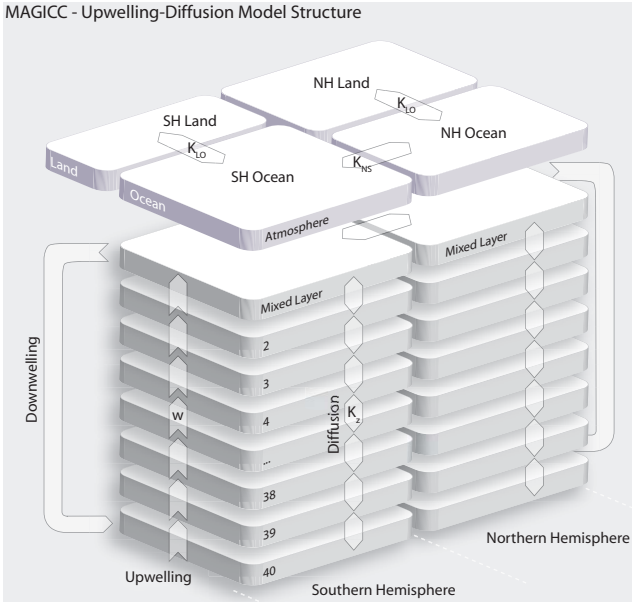

Fig. A1. The schematic structure of MAGICC's upwelling-diffusion energy balance module with land and ocean boxes in each hemisphere. The processes for heat transport in the ocean are deep-water formation, upwelling, diffusion, and heat exchange between the hemispheres. Not shown is the entrainment and the vertically depth-dependent area of the ocean layers (see Fig. A2 and text).

A4.1 Partitioning of feedbacks

In order to improve the comparability between MAGICC and AOGCMs, and following earlier versions of MAGICC, we use different feedback parameters over land and ocean. This requires an adjustable land to ocean warming ratio in equilibrium based on AOGCM results. Given that in equilibrium the oceanic heat uptake is zero, the global energy balance equation can be written as:

$$\Delta Q_G = \lambda_G \Delta T_G = f_L \lambda_L \Delta T_L + f_O \lambda_O \Delta T_O \quad \text{(A45)}$$

where ΔQ_G, λ_G and ΔT_G are the global-mean forcing, feedback, and temperature change, respectively. The right hand side uses the area fractions f, feedbacks λ, and mean temperature changes, ΔT for ocean (O) and land (L). As in earlier versions of MAGICC, the non-linear set of equations that determines λ_O and λ_L for a given set of equilibrium land-ocean warming ratio RLO (=$\Delta T_L/\Delta T_O$), global-mean feedback λ_G, heat exchange and enhancement factors (k, μ), is solved by an iterative procedure involving the set of linear Eqs. (A46–A49), seeking the solution for λ_L closest to λ_G. The procedure in version 6 has been modified slightly to take into account the time-constant radiative forcing pattern by CO_2 for the four boxes with hemispheric land/ocean regions, if prescribed.

Following Wigley and Schlesinger (1985), it is assumed that the atmosphere is in equilibrium with the underlying ocean mixed layer, so that the energy balance equation for the Northern Hemispheric ocean (NO) is:

$$\begin{aligned} f_{\mathrm{NO}}\lambda_O \Delta T_{\mathrm{NO}} = &\quad \text{:infrared outgoing flux} \\ f_{\mathrm{NO}} \Delta Q_{\mathrm{NO}} &\quad \text{:forcing} \\ +k_{\mathrm{LO}}(\Delta T_{\mathrm{NL}} - \mu \Delta T_{\mathrm{NO}}) &\quad \text{:land-ocean heat exchange} \\ +k_{\mathrm{NS}}\alpha(\Delta T_{\mathrm{SO}} - \Delta T_{\mathrm{NO}}) &\quad \text{:hemispheric heat exch.} \end{aligned} \quad \text{(A46)}$$

where ΔT_{NO} is the surface temperature change over the Northern Hemisphere ocean, ΔQ_{NO} the radiative forcing over that region, f_{NO} the northern ocean's area fraction of the earth surface, k_{LO} the land-ocean heat exchange coefficient [W m^{-2}°C^{-1}], a heat transport enhancement factor μ allowing for asymmetric heat exchange between land and ocean ($1 \leq \mu$ – see Sect. A4.2 below), k_{NS} is the hemispheric heat exchange coefficient in the mixed layer. Following Raper and Cubasch (1996) α is a sea-ice related adjustment factor to relate upper ocean temperature change to surface air temperature change (see Sect. A4.5). Correspondingly, the equilibrium energy balance equations for the Northern Hemisphere land (NL), Southern Hemisphere ocean (SO) and Southern Hemisphere land (SL) are:

$$\begin{aligned} f_{\mathrm{NL}}\lambda_L \Delta T_{\mathrm{NL}} = &\; f_{\mathrm{NL}} \Delta Q_{\mathrm{NL}} \\ &+k_{\mathrm{LO}}(\mu \Delta T_{\mathrm{NO}} - \Delta T_{\mathrm{NL}}) \end{aligned} \quad \text{(A47)}$$

$$\begin{aligned} f_{\mathrm{SO}}\lambda_O \Delta T_{\mathrm{SO}} = &\; f_{\mathrm{SO}} \Delta Q_{\mathrm{SO}} \\ &+k_{\mathrm{LO}}(\Delta T_{\mathrm{SL}} - \mu \Delta T_{\mathrm{SO}}) \\ &+k_{\mathrm{NS}}\alpha(\Delta T_{\mathrm{NO}} - \Delta T_{\mathrm{SO}}) \end{aligned} \quad \text{(A48)}$$

$$\begin{aligned} f_{\mathrm{SL}}\lambda_L \Delta T_{\mathrm{SL}} = &\; f_{SL} \Delta Q_{SL} \\ &+k_{\mathrm{LO}}(\mu \Delta T_{\mathrm{SO}} - \Delta T_{\mathrm{SL}}) \end{aligned} \quad \text{(A49)}$$

As detailed below (Sect. A4.3), if the sensitivity factor ξ is set different from zero (see Eq. A51), it is possible to make the feedback factors λ in the energy balance equation dependent on the total radiative forcing. This forcing dependence of the feedback factors and the heat exchange enhancement factors are newly introduced in this version of MAGICC. The following two sections. (A4.2 and A4.2) are intended to provide both the motivation and details of these new parameterizations.

A4.2 Revised land-ocean heat exchange formulation

This section highlights a "geometric" effect that can cause effective climate sensitivities to change over time. The global-mean sensitivity may increase simply due to decreasing land-ocean warming ratios, given that climate feedbacks over land and ocean areas are different. To control the relative temperature changes over ocean and land, a heat transport enhancement factor μ is introduced. Enhancing the ocean-to-land heat transport ($\mu \geq 1$) has the benefit that the simple climate

model can better simulate some characteristic AOGCM responses. In the idealized forcing runs, AOGCMs often show a transient land-ocean warming ratio that slightly decreases over time, but stays above unity, combined with an increasing effective climate sensitivity in some models (see bottom rows in Fig. B1, B2, and B3). The higher land than ocean warming (RLO>1) could be achieved by a smaller feedback (greater climate sensitivity) over land compared to the ocean boxes. However, as the land-ocean warming ratio decreases over time (due to less and less ocean heat uptake towards equilibrium), so would the effective global-mean climate sensitivity in previous model versions. The method used here, to allow both a RLO above unity and a non-decreasing effective climate sensitivity, assumes that ocean temperature perturbations influence the heat exchange more than land temperature changes. This asymmetric heat exchange formulation is then given by:

$$\mathrm{HX_{LO}} = k_{\mathrm{LO}}(\Delta T_L - \mu \Delta T_O) \tag{A50}$$

where $\mathrm{HX_{LO}}$ is the land-ocean heat exchange (positive in direction land to ocean), μ is the ocean-to-land enhancement factor and ΔT_L and ΔT_O are the temperature perturbations for the land and ocean region, respectively (cf. Eq. A46 ff.). Typical values for μ range between 1 and 1.4 as estimated from calibrating the CMIP3 ensemble (see Table B3).

A4.3 Accounting for climate-state dependent feedbacks

Some AOGCM runs indicate higher effective climate sensitivities for higher forcings and/or temperatures. For example, the ECHAM5/MPI-OM model shows an effective climate sensitivity of approximately 3.5°C after stabilization at twice pre-industrial CO_2 concentrations and 4°C for stabilization at quadrupled pre-industrial CO_2 concentrations (see Fig. 1b – see as well Raper et al., 2001; Hansen et al., 2005). Given that the transient land-ocean warming ratio is the same for the 1pctto2× and 1pctto4× runs (see Fig. B1 last row), the 'geometric' effect discussed in the Sect. A4.2 would not explain this increase in climate sensitivity. An alternative explanation could be that climate feedbacks are climate-state dependent. The assumption in the standard energy balance Eq. (1) with a constant global feedback (λ), with its attendant requirement that the outgoing energy flux scales proportionally with temperature change, may be an oversimplification. For example, the slow feedback due to retreating ice-sheets can lead to changes in the diagnosed effective sensitivities in AOGCMs (see e.g. Raper et al., 2001) over long time-scales. Hansen et al. (2005) show that the 100-year climate response in the GISS model is more sensitive to higher forcings than to lower or negative forcings. Hansen et al. (2005) express this effect by increasing efficacies for increasing radiative forcing. Table 1 in Hansen et al. (2005) suggests a gradient of roughly 1% increase in efficacy for each additional $\mathrm{Wm^{-2}}$ (OLS-regression of E_a versus F_a across the full range of CO_2 experiments), although some intervals (e.g. from 1.25 to 1.5×CO_2) show a slightly higher sensitivity of efficacy to forcing, i.e., 3% per $\mathrm{Wm^{-2}}$.

Rather than making the efficacies dependent on forcing, an alternative is to make the climate sensitivity dependent on the forcing level. This distinction, on whether to modify forcing or sensitivity, is not important when the climate system is at or close to equilibrium. However, if the efficacies of the forcing, instead of the feedback parameters are allowed to vary with forcing, the transient climate response after a change in forcing will be slightly faster. In this MAGICC version, if a forcing dependency of the sensitivity is assumed, the land and ocean feedback parameters λ_L and λ_O are scaled as

$$\lambda = \frac{\Delta Q_{2\times}}{\frac{\Delta Q_{2\times}}{\lambda_{2\times}} + \xi(\Delta Q - \Delta Q_{2\times})} \tag{A51}$$

where $\lambda_{2\times}$ is the feedback parameter ($=\frac{\Delta Q_{2\times}}{\Delta T_{2\times}}$) at the forcing level for twice pre-industrial CO_2 concentrations. The sensitivity factor ξ ($\mathrm{KW^{-1}\,m^2}$) scales the climate sensitivity in proportion to the difference of forcing away from the model-specific "twice pre-industrial CO_2 forcing level" ($\Delta Q - \Delta Q_{2\times}$). The 1% increase in efficacy for each additional unit forcing in Hansen's findings translates into a feedback sensitivity factor ξ of $0.03\,\mathrm{KW^{-1}\,m^2}$ (assuming a climate sensitivity $\Delta T_{2\times}$ of 3 °C). Note that this scaling convention (Eq. A51) ensures that climate sensitivities are comparable for the equilibrium warming that corresponds to twice preindustrial CO_2 concentration levels (see Table 4).

A4.4 Efficacies

Efficacy is defined as the ratio of global-mean temperature response for a particular radiative forcing divided by the global-mean temperature response for the same amount of global-mean radiative forcing induced by CO_2 (see Sect. 2.8.5 in Forster et al., 2007). In most cases, the efficacies are different for different forcing agents because of the geographical and vertical distributions of the forcing (Boer and Yu, 2003; Joshi et al., 2003; Hansen et al., 2005). The effective radiative forcing (ΔQ_e) is the product of the standard climate forcing (ΔQ_a), calculated after thermal adjustment of the stratosphere, and the efficacy (E_a). It is the effective forcings that are used in the energy balance equation (Eq. 1), although both effective and standard forcings are carried through in the MAGICC code. Note that this parameterization yields slightly faster transient climate responses compared to an approach where different climate sensitivities are applied for each individual forcing agent (cf. Sect. A4.3 above).

In MAGICC, forcings for some components differ by hemisphere and over land and ocean. Just as for the global sensitivity, this, in combination with different land/ocean feedback factors, results in MAGICC6 exhibiting efficacies different from unity for non-CO_2 forcing agents. In other

words, efficacies different from unity are in part a consequence of the geometric effect described above. MAGICC calculates these internal efficacies using reference year (default 2005) forcing patterns. After normalizing these forcing patterns to a global-mean of $\Delta Q_{2\times}$ (default 3.71 Wm^{-2}), the internal efficacy can be determined as

$$E_{\mathrm{int}} = \frac{\Delta T_{\mathrm{eff}2\times}}{\Delta T_{2\times}}, \tag{A52}$$

where $T_{\mathrm{eff}2\times}$ is the actual global-mean equilibrium temperature change resulting from a normalized forcing pattern and $\Delta T_{2\times}$ is the corresponding warming for $2\times$ CO_2 forcing, i.e., the climate sensitivity. For most forcing agents, these internal efficacies are very close to one, except for forcings with a strong land/ocean forcing contrast, such as aerosol forcings. For example, for direct aerosol forcing in the HadCM3 emulation (calibration III – see Table B3) the efficacy is 1.14. By default, these internal efficacies are taken into account when applying prescribed efficacies, so that:

$$\Delta Q_e = \frac{E_a}{E_{\mathrm{int}}} \Delta Q_a \tag{A53}$$

A4.5 The upwelling-diffusion equations

The transient temperature change evolution is largely influenced by the climate system's inertia, which in turn depends on the nature of the heat uptake by the climate system. The transient energy balance equations can be written as:

$$f_{\mathrm{NO}}(\zeta_o \frac{d\Delta T_{\mathrm{NO},1}}{dt} - \Delta Q_{\mathrm{NO}} + \lambda_o \alpha \Delta T_{\mathrm{NO},1} + F_N) = k_{\mathrm{LO}}(\Delta T_{\mathrm{NL}} - \mu\alpha \Delta T_{\mathrm{NO},1}) + k_{NS}\alpha(\Delta T_{\mathrm{SO},1} - \Delta T_{\mathrm{NO},1}) \tag{A54}$$

$$f_{\mathrm{NL}}(\zeta_L \frac{d\Delta T_{\mathrm{NL}}}{dt} - \Delta Q_{\mathrm{NL}} + \lambda_L \Delta T_{\mathrm{NL}}) = k_{\mathrm{LO}}(\mu\alpha \Delta T_{\mathrm{NO},1} - \Delta T_{\mathrm{NL}}) \tag{A55}$$

$$f_{SO}(\zeta_o \frac{d\Delta T_{\mathrm{SO},1}}{dt} - \Delta Q_{\mathrm{SO}} + \lambda_o \alpha \Delta T_{\mathrm{SO},1} + F_S) = k_{\mathrm{LO}}(\Delta T_{\mathrm{SL}} - \mu\alpha \Delta T_{\mathrm{SO},1}) + k_{\mathrm{NS}}\alpha(\Delta T_{\mathrm{NO},1} - \Delta T_{\mathrm{SO},1}) \tag{A56}$$

$$f_{\mathrm{SL}}(\zeta_L \frac{d\Delta T_{\mathrm{SL}}}{dt} - \Delta Q_{\mathrm{SL}} + \lambda_L \Delta T_{\mathrm{SL}}) = k_{\mathrm{LO}}(\mu\alpha \Delta T_{\mathrm{SO},1} - \Delta T_{\mathrm{SL}}) \tag{A57}$$

where the adjustment factor α (default 1.2) determines – over ocean areas – the ratio of hemispheric changes in air (ΔT_{xO}) versus ocean mixed layer temperatures ($\Delta T_{\mathrm{xO},1}$). Based on ECHAM1/LSG analysis (Raper and Cubasch, 1996), this sea-ice factor was first introduced by Raper et al. (2001) to account for the fact that the air temperature will exhibit additional warming, because the atmosphere feels warmer ocean surface temperatures where sea ice retreats. The bulk heat capacity of the mixed layer in each hemisphere x is $f_x\zeta_o = f_x \rho c h_m$, where ρ denotes the density of seawater (1.026×10^6 g m^{-3}), c is the specific heat capacity (0.9333 cal $\mathrm{g}^{-1}{}^{\circ}\mathrm{C}^{-1}$=4.1856×0.9333 Joule $\mathrm{g}^{-1}{}^{\circ}\mathrm{C}^{-1}$) and h_m is the mixed layer's thickness [m]. The bulk heat capacity of the land areas is $f_x\zeta_L$, here assumed to be zero. The net heat flux into the ocean below the mixed layer is denoted by F_x.

Equation (A55) can then be written as:

$$\Delta T_{\mathrm{NL}} = \frac{f_{\mathrm{NL}}\Delta Q_{\mathrm{NL}} + k_{\mathrm{LO}}\mu\alpha\Delta T_{\mathrm{NO},1}}{f_{\mathrm{NL}}\lambda_L + k_{\mathrm{LO}}} \tag{A58}$$

Substituting ΔT_{NL} in Eq. (A54) yields:

$$f_{\mathrm{NO}}(\zeta_o \frac{d\Delta T_{\mathrm{NO},1}}{dt} - \Delta Q_{\mathrm{NO}} + \lambda_o\alpha\Delta T_{\mathrm{NO},1} + F_N) = \frac{k_{\mathrm{LO}}}{\frac{k_{\mathrm{LO}}}{f_{\mathrm{NL}}} + \lambda_L}(\Delta Q_{\mathrm{NL}} - \lambda_L\mu\alpha\Delta T_{\mathrm{NO},1}) + k_{\mathrm{NS}}\alpha(\Delta T_{\mathrm{SO},1} - \Delta T_{\mathrm{NO},1}) \tag{A59}$$

Provided we know the heat flux F_N into the ocean below the mixed layer, we could now derive $d\Delta T_{\mathrm{NO},1}/dt$. The net heat flux F_N at the bottom of the mixed layer is determined by vertical heat diffusivity (diffusion coefficient K_z [$\mathrm{cm}^2\,\mathrm{s}^{-1}$=3155.76^{-1} $\mathrm{m}^2\,\mathrm{yr}^{-1}$]), and upwelling and downwelling (upwelling velocity w [m yr^{-1}]), both acting on the perturbations ΔT from the initial temperature profile $T^0_{\mathrm{NO,z}}$. If the upwelling rate w varies over time, the change in upwelling velocity $\Delta w^t = (w^t - w^0)$ compared to its initial state w^0 is assumed to act on the initial temperature profile, so that:

$$F_N = \frac{K_z}{0.5h_d}\rho c(\Delta T_{\mathrm{NO},1} - \Delta T_{\mathrm{NO},2}) - w\rho c(\Delta T_{\mathrm{NO},2} - \beta\Delta T_{\mathrm{NO},1}) - \Delta w\rho c(T^0_{\mathrm{NO},2} - T^0_{\mathrm{NO,sink}}) \tag{A60}$$

where $T^0_{\mathrm{NO,z}}$ is the initial temperature for water in layer z or in the downwelling pipe (z = "sink").

Given that the top layer is assumed to be mixed, the gradient of the temperature perturbations is calculated by the difference of the perturbations divided by half the thickness h_d of the second layer (see Fig. A2). Substituting F_N in Eq. (A59) with Eq. (A60) and transforming the equation to discrete time steps, yields:

$$\begin{aligned}\frac{d\Delta T_{\mathrm{NO},1}}{dt} \approx \frac{\Delta T^{t+1}_{\mathrm{NO},1} - \Delta T^t_{\mathrm{NO},1}}{\Delta t} = & \\ \frac{1}{\zeta_o}\Delta Q^t_{\mathrm{NO}} & \text{:forcing} \\ -\frac{\lambda_o\alpha}{\zeta_o}\Delta T^{t+1}_{\mathrm{NO},1} & \text{:feedback} \\ -\frac{K_z}{0.5h_dh_m}(\Delta T^{t+1}_{\mathrm{NO},1} - \Delta T^{t+1}_{\mathrm{NO},2}) & \text{:diffusion} \\ +\frac{w^t}{h_m}(\Delta T^{t+1}_{\mathrm{NO},2} - \beta\Delta T^{t+1}_{\mathrm{NO},1}) & \text{:upwelling} \\ +\frac{\Delta w^t}{h_m}(T^0_{\mathrm{NO},2} - T^0_{\mathrm{NO,sink}}) & \text{:variable upwelling}\end{aligned} \tag{A61}$$

$$+\frac{k_{\mathrm{LO}}(\Delta Q^t_{\mathrm{NL}}-\lambda_L\mu\alpha\Delta T^{t+1}_{\mathrm{NO},1})}{\zeta_o f_{\mathrm{NO}}(\frac{k_{\mathrm{LO}}}{f_{\mathrm{NL}}}+\lambda_L)} \quad \text{:land forcing}$$

$$+\frac{k_{\mathrm{NS}}\alpha}{\zeta_o f_{\mathrm{NO}}}(\Delta T^t_{\mathrm{SO},1}-\Delta T^t_{\mathrm{NO},1}) \text{ :inter-hemispheric ex.}$$

For the layers below the mixed layer (2≤z≤n–1), the temperature updating is governed by diffusion (first two terms in Eq. A62) and upwelling (last two terms), so that:

$$\begin{aligned}\frac{\Delta T^{t+1}_{\mathrm{NO},z}-\Delta T^t_{\mathrm{NO},z}}{\Delta t} =& \\ \frac{K_z}{0.5(h_d+h'_d)h_d}(\Delta T^{t+1}_{\mathrm{NO},z-1}-\Delta T^{t+1}_{\mathrm{NO},z}) & \\ -\frac{K_z}{h_d^2}(\Delta T^{t+1}_{\mathrm{NO,z}}-\Delta T^{t+1}_{\mathrm{NO},z+1}) & \\ +\frac{w^t}{h_d}(\Delta T^{t+1}_{\mathrm{NO},z+1}-\Delta T^{t+1}_{\mathrm{NO},z}) & \\ +\frac{\Delta w^t}{h_d}(T^0_{\mathrm{NO},z+1}-T^0_{\mathrm{NO},z}) & \quad \text{(A62)}\end{aligned}$$

where h'_d is zero for the layer below the mixed layer (z=2) and h_d otherwise, Δw^t is the change from the initial upwelling rate.

For the bottom layer ($z=n$), the downwelling term has to be taken into account, so that:

$$\begin{aligned}\frac{\Delta T^{t+1}_{\mathrm{NO},n}-\Delta T^t_{\mathrm{NO},n}}{\Delta t} =& \frac{K_z}{h_d^2}(\Delta T^{t+1}_{\mathrm{NO,n-1}}-\Delta T^{t+1}_{\mathrm{NO},n}) \\ &+\frac{w^t}{h_d}(\beta\Delta T^t_{\mathrm{NO},1}-\Delta T^{t+1}_{\mathrm{NO},n}) \\ &+\frac{\Delta w^t}{h_d}(T^0_{\mathrm{NO,sink}}-T^0_{\mathrm{NO},n}) \quad \text{(A63)}\end{aligned}$$

Corresponding to the temperature calculations shown here for the Northern Hemisphere ocean (NO), the equivalent steps apply for the Southern Hemisphere ocean (SO). For simplicity, the equations described above are for the constant-depth area profile case, which MAGICC defaults to when the depth-dependency factor ϑ is set to zero. The detailed code for the general case with $0\leq\vartheta\leq1$ is given in Sect. A5.

A4.6 Calculating heat uptake

Heat uptake by the climate system can be calculated in different ways. One method is to use the global energy balance (Eq. 1). Using the effective sensitivity as in Eq. (A45) the heat uptake F^t is estimated as:

$$\frac{dH^t}{dt}=F^t=\Delta Q^t-(f_L\lambda_L\Delta T^t_L+f_O\lambda_O\Delta T^t_O) \quad \text{(A64)}$$

For verification purposes MAGICC6 calculates heat uptake in two ways, both directly (as above) and by integrating heat content changes in each layer in the ocean (yielding identical results), given the assumed zero heat capacity of the atmosphere and land areas:

$$\Delta H^t=\sum_{i=1}^{n}\frac{1}{\rho c h_i}\frac{(f_{\mathrm{NO}}\Delta T^t_{\mathrm{NO,i}}+f_{\mathrm{SO}}\Delta T^t_{\mathrm{SO,i}})}{f_O}+\epsilon \quad \text{(A65)}$$

where h_i is the thickness of the layer, i.e., h_m for the mixed layer and h_d for the others and ϵ is a small term to account for the heat content of the polar sinking water.

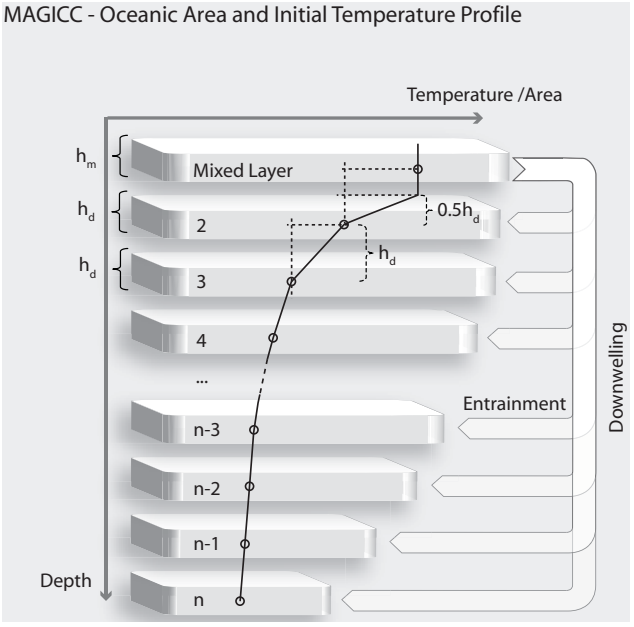

Fig. A2. The schematic oceanic area and initial temperature profiles in MAGICC's ocean hemispheres. Diffusion driven heat transport is modeled proportional to the vertical gradient of temperature, which is especially high below the mixed layer.

A4.7 Depth-dependent ocean with entrainment

Harvey and Schneider (1985b,a) introduced the upwelling-diffusion model with entrainment from the polar sinking water by varying the upwelling velocity w with depth. Building on the work by Raper et al. (2001), MAGICC6 also includes the option of a depth-dependent ocean area profile. If the depth-dependency parameter ϑ is set to 1 (default), a standard depth-dependent ocean area profile is assumed as in HadCM2 and used in Raper et al. (2001). A constant upwelling velocity is assumed and mass conservation is maintained by "entrainment" from the downwelling pipe. With ocean area decreasing with depth and constant upwelling velocity, the upwelling mass flux would also have to decrease with depth. To offset this, the amount of entrainment into layer z is assumed to be proportional to the decrease in area from the top to the bottom of each layer (cf. Fig. A2). We differ from the model structures tested by Raper et al. (2001), by equating changes in the temperature of the entraining water

to those in the downwelling pipe, namely a fraction β (default 0.2) of the mixed layer temperature $\Delta T_{x,1}^{t-1}$ of the previous timestep in Hemisphere x. For a detailed description of the code, see the following Sect. A5. Simple upwelling-diffusion models can overestimate the ocean heat uptake for higher warming scenarios when applying parameter values calibrated to match heat uptake for lower warming scenarios (see e.g. Fig. 17b in Harvey et al., 1997). To address this, MAGICC6 includes a warming-dependent vertical diffusivity gradient. The physical reasoning is that a strengthened thermal stratification and, hence, reduced vertical mixing leads to decreased heat uptake for higher warming. Thus, the effective vertical diffusivity at $K_{z,i}$ between ocean layer i and i+1 is given by:

$$K_{z,i} = \max(K_{z,\min}, (1-d_i)\frac{dK_z}{dT}(\Delta T_{H,1}^{t-1} - \Delta T_{H,n}^{t-1}) + K_z) \quad \text{(A66)}$$

where $K_{z,\min}$ 0.1 cm^2 s^{-1}); d_i is the relative depth of the layer boundary with zero at the bottom of the mixed layer and one for the top of the bottom layer; $\frac{dK_z}{dT}$ is a newly introduced ocean stratification coefficient specifying how the vertical diffusivity K_z between the mixed layer 1 and layer 2 changes with a change in the temperature difference between the top/mixed and bottom ocean layer of the respective hemisphere at the previous timestep $t-1$ ($\Delta T_{H,1}^{t-1} - \Delta T_{H,n}^{t-1}$).

A5 Implementation of upwelling-diffusion-entrainment equations

This section details how the equations governing the upwelling-diffusion-entrainment (UDE) ocean (Eqs. A62, A62, A63) are implemented and modified by entrainment terms and depth-dependent ocean area (see Fig. A2). These equations represent the core of the UDE model and build on the initial work by Hoffert et al. (1980); Harvey and Schneider (1985b,a).

The entrainment is here modeled so that the upwelling velocity in the main column is the same in each layer. Thus, the three area correction factors, θ_z^{top}, θ_z^b and θ_z^{dif}, applied below are:

$$\begin{aligned} \theta_z^{\text{top}} &= \frac{A_z}{(A_{z+1}+A_z)/2} \\ \theta_z^b &= \frac{A_{z+1}}{(A_{z+1}+A_z)/2} \\ \theta_z^{\text{dif}} &= \frac{A_{z+1}-A_z}{(A_{z+1}+A_z)/2} \end{aligned} \quad \text{(A67)}$$

where A_z is the area at the top of layer z or bottom of layer $z-1$ and the denominator is thus an approximation for the mean area of each ocean layer.

For the mixed layer, all terms in Eq. (A62) involving $\Delta T_{\text{NO},1}^{t+1}$ are collected on the left hand side in variable $A(1)$. All terms involving $\Delta T_{\text{NO},2}^{t+1}$ are collected in variable $B(1)$ on the left hand side. All other terms are held in variable $D(1)$ on the right hand side, so that the equation reads:

$$\Delta T_{\text{NO},1}^{t+1} = -\frac{B(1)}{A(1)}\Delta T_{\text{NO},2}^{t+1} + \frac{D(1)}{A(1)} \quad \text{(A68)}$$

with

$$A(1) = 1.0 + \theta_1^{\text{top}}\Delta t\frac{\lambda_O \alpha}{\zeta_o} \quad \text{:feedback over ocean} \quad \text{(A69)}$$

$$+\theta_1^b \Delta t \frac{K_z}{0.5 h_m h_d} \quad \text{:diffusion to layer 2}$$

$$+\theta_1^b \Delta t \frac{w^t \beta}{h_m} \quad \text{:downwelling}$$

$$+\theta_1^{\text{top}} \Delta t \frac{k_{\text{LO}} \lambda_L \mu \alpha}{\zeta_o f_{\text{NO}} (\frac{k_{\text{LO}}}{f_{\text{NL}}} + \lambda_L)} \quad \text{:feedback over land}$$

$$B(1) = -\theta_1^b \Delta t \frac{K_z}{0.5 h_m h_d} \quad \text{:diffusion from layer 2} \quad \text{(A70)}$$

$$-\theta_1^b \Delta t \frac{w^t}{h_m} \quad \text{:upwelling from layer 2}$$

$$D(1) = \Delta T_{\text{NO},1}^t \quad \text{:previous temp} \quad \text{(A71)}$$

$$+\theta_1^{\text{top}} \Delta t \frac{1}{\zeta_o} \Delta Q_{NO} \quad \text{:forcing ocean}$$

$$+\theta_1^{\text{top}} \Delta t \frac{\alpha k_{NS}}{\zeta_o f_{NO}} (\Delta T_{\text{SO},1}^t - \Delta T_{NO,1}^t) \quad \text{:inter-hemis. exch.}$$

$$+\theta_1^{\text{top}} \Delta t \frac{k_{LO} \Delta Q_{NL}}{\zeta_o f_{NO} (\frac{k_{LO}}{f_{NL}} + \lambda_L)} \quad \text{:land forcing}$$

$$+\theta_1^b \Delta t \frac{\Delta w^t}{h_m} (T_{\text{NO},2}^0 - T_{NO,sink}^0) \quad \text{:variable upwelling}$$

For the interior layers ($2 \le z \le n$), i.e., all layers except the top mixed layer and the bottom layer, the terms are reordered, so that A(z) comprises the terms for $\Delta T_{\text{NO,z}-1}^{t+1}$, $B(z)$ the terms for $\Delta T_{\text{NO,z}}^{t+1}$, C(z) the terms for $\Delta T_{\text{NO,z}+1}^{t+1}$ and $D(z)$ the remaining terms, according to:

$$\Delta T_{\text{NO,z}-1}^{t+1} = -\frac{B(z)}{A(z)}\Delta T_{\text{NO,z}}^{t+1} - \frac{C(z)}{A(z)}\Delta T_{\text{NO,z}+1}^{t+1} + \frac{D(z)}{A(z)} \quad \text{(A72)}$$

with

$$A(z) = -\theta_z^{top} \Delta t \frac{K_z}{0.5(h_d + h_d')h_d} \quad \text{:diffusion from layer above} \quad \text{(A73)}$$

$$B(z) = 1.0 + \theta_z^b \Delta t \frac{K_z}{h_d^2} \quad \text{:diffusion to layer below}$$

$$+\theta_z^{top}\Delta t\frac{K_z}{0.5(h_d+h'_d)h_d} \quad \text{:diffusion to layer above}$$

$$+\theta_z^{top}\Delta t\frac{w^t}{h_d} \quad \text{:upwelling to layer above} \tag{A74}$$

$$C(z)=-\theta_z^b\Delta t\frac{K_z}{h_d^2} \quad \text{:diffusion from layer below}$$

$$-\theta_z^b\Delta t\frac{w^t}{h_d} \quad \text{:upwelling from layer below} \tag{A75}$$

$$D(z)=\Delta T^t_{\mathrm{NO,z}} \quad \text{:previous temp}$$

$$+\Delta t\frac{\Delta w^t}{h_d}(\theta_z^b T^0_{\mathrm{NO,z+1}}-\theta_z^{top}T^0_{\mathrm{NO,z}}) \quad \text{:variable upwelling}$$

$$+\theta_z^{\mathrm{dif}}\Delta t\frac{w^t}{h_d}\beta\Delta T^{t-1}_{\mathrm{NO,1}} \quad \text{:entrainment}$$

$$+\theta_z^{\mathrm{dif}}\Delta t\frac{\Delta w^t}{h_d}T^0_{\mathrm{NO,sink}} \quad \text{:variable entrainment} \tag{A76}$$

where h'_d is zero for the layer below the mixed layer and h_d otherwise. For the bottom layer, the respective sum factor $A(n)$ for $\Delta T^{t+1}_{\mathrm{NO,n-1}}$, $B(n)$ for $\Delta T^{t+1}_{\mathrm{NO,n}}$ and $D(n)$ for the remaining terms is:

$$\Delta T^{t+1}_{\mathrm{NO},n-1}=-\frac{B(n)}{A(n)}\Delta T^{t+1}_{\mathrm{NO},n}+\frac{D(n)}{A(n)} \tag{A77}$$

with

$$A(n)=-\theta_n^{\mathrm{top}}\Delta t\frac{K_z}{h_d^2} \quad \text{:diffusion from layer n-1} \tag{A78}$$

$$B(n)=1.0+\theta_n^{\mathrm{top}}\Delta t\frac{K_z}{h_d^2} \quad \text{:diffusion to layer n-1} \tag{A79}$$

$$+\theta_n^{\mathrm{top}}\Delta t\frac{w^t}{h_d} \quad \text{:upwelling to layer n-1}$$

$$D(n)=\Delta T^t_{\mathrm{NO},n} \quad \text{:previous temp} \tag{A80}$$

$$+\theta_n^{\mathrm{top}}\Delta t\frac{w^t}{h_d}\beta\Delta T^{t-1}_{\mathrm{NO,1}} \quad \text{:downwelling from top layer}$$

$$-\theta_n^{\mathrm{top}}\Delta t\frac{\Delta w^t}{h_d}T^0_{\mathrm{NO},n} \quad \text{:variable upwelling}$$

$$+\theta_n^{\mathrm{top}}\Delta t\frac{\Delta w^t}{h_d}T^0_{\mathrm{NO,sink}} \quad \text{:variable downwelling}$$

With these Eqs. (A68–A80), the ocean temperatures can be solved consecutively from the bottom to the top layer at each time step.

Appendix B

Calibration result details

This appendix provides additional details on the calibration procedures and results. The results provided are the individual parameter settings for each CMIP3 AOGCM for the three calibration procedures (see Table 1, and Tables B1, B2 and B3) as well as graphical comparisons between the original CMIP3 AOGCM data and their calibration IIIa emulations (see Figs. B1, B2 and B3). In addition, detailed results are provided for the calibrations to the C^{4}MIP carbon cycle models, the optimized MAGICC parameters, and goodness-of-fit statistics (see Table B4 and Fig. B4).

By calibrating a simple model to more than a single data series, some arbitrariness arises in relation to how the overall goodness of fit is composed. In particular, fitting dataseries with different units, like temperature (K) and ocean heat uptake (W/m^2) requires some sort of normalization to avoid the situation where some data series are dominating the calibration result simply because they are measured with larger numerical values. The normalization could be done by weighting the data series by the inverse of their covariance matrix, either using observational, control run or de-drifted model output segments. For simplicity, a more pragmatic method was chosen. Weights for the root mean square errors for the available time series are chosen after a series of calibration iterations so that the contribution of each time-series to the overall goodness of fit is of similar magnitude, thereby avoiding the possibility that a single time series might dominate the calibration result. Although this approach is somewhat arbitrary, we found that the calibration results were insensitive to the chosen weights for different variables.

For the AOGCM calibrations, the chosen weights were 10 (heat-uptake series, W/m^2) and 1 (temperature dataseries, K). For calibrating to the C^{4}MIP carbon cycle models, the chosen weights are as follows: 1 (global-mean surface temperature, K): 25 (net air-to-land flux, GtC/yr): 100 (net air-to-ocean flux, GtC/yr): 50 (atm. CO_2 concentrations, ppm): 25 (NPP and heterotrophic respiration fluxes, GtC/yr): 1 (plant carbon pool, GtC): 0.5 (dead, detritus and soil carbon pools, GtC). Note that all fitted AOGCMs and carbon cycle time series were low-pass filtered in order to reduce the noise introduced by natural variability (or the modelled part thereof), as only the mean signal, not the variability, is simulated by MAGICC. The low-pass filtering method followed Mann (2004) and employed a pass band boundary of 1/20 cycle/yr and roughness constraint.

Supplement related to this article is available online at: http://www.atmos-chem-phys.net/11/1417/2011/acp-11-1417-2011-supplement.pdf.

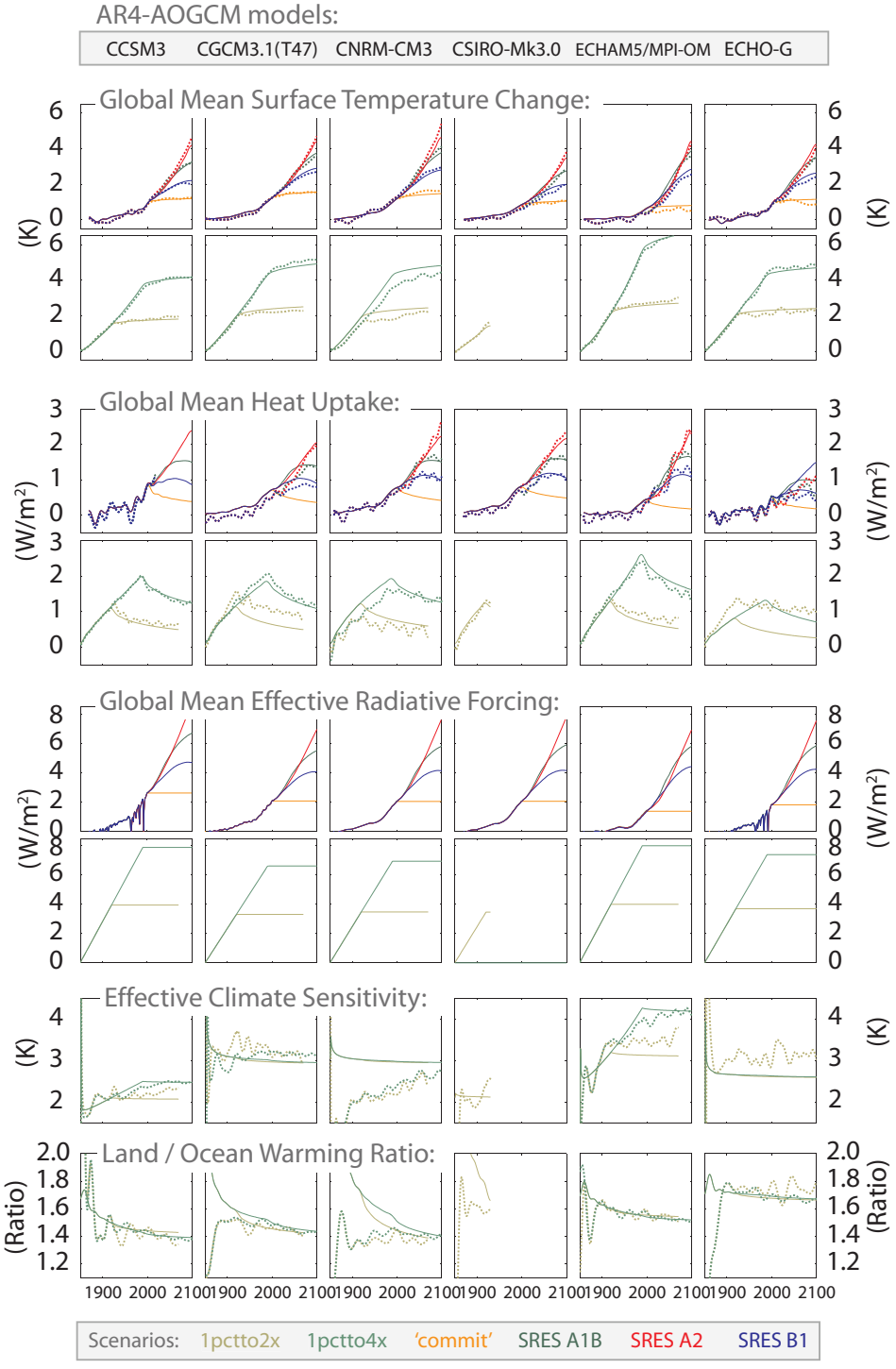

Fig. B1. Comparison of global-mean surface temperature (rows 1 and 2), heat uptake (rows 3, 4), effective radiative forcing (rows 5, 6), the effective climate sensitivity (row 7) and the land-ocean warming ratio (8), between CMIP3 AOGCM models (dotted) and the calibrated MAGICC6 (solid) model (calibration III with "like-with-like" AOGCM specific forcing) from 1850 to 2100. Shown are the comparisons for the idealized CO_2-only scenarios (1pctto2× and 1pctto4×) set to start in 1850 and the multi-forcing runs for the 20th century (20c3m), three SRES scenarios, and the commitment run. For the multi-gas scenarios, MAGICC is driven here by the AOGCM-specific subsets of forcing agents (see Table 2). AOGCM drift was removed by substracting the respective lowpass-filtered control run segments. Both the AOGCM and the MAGICC temperature outputs were lowpass-filtered using a low pass boundary of 0.05 cycle/yr and roughness constraint (Mann, 2004). See following figures for the other CMIP3 AOGCMs emulations.

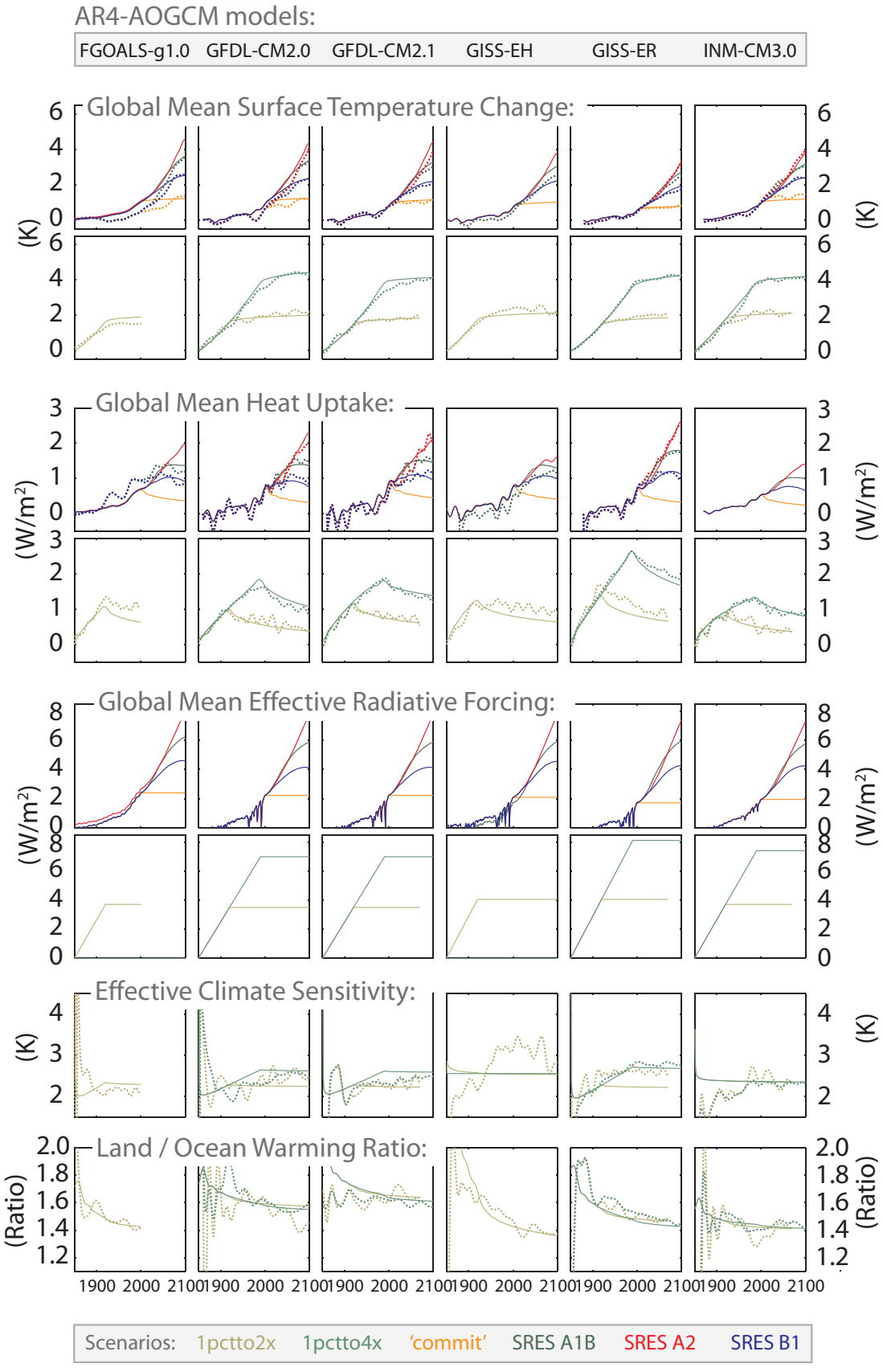

Fig. B2. As Fig. B1, but for another six of the 19 emulated CMIP3 AOGCMs.

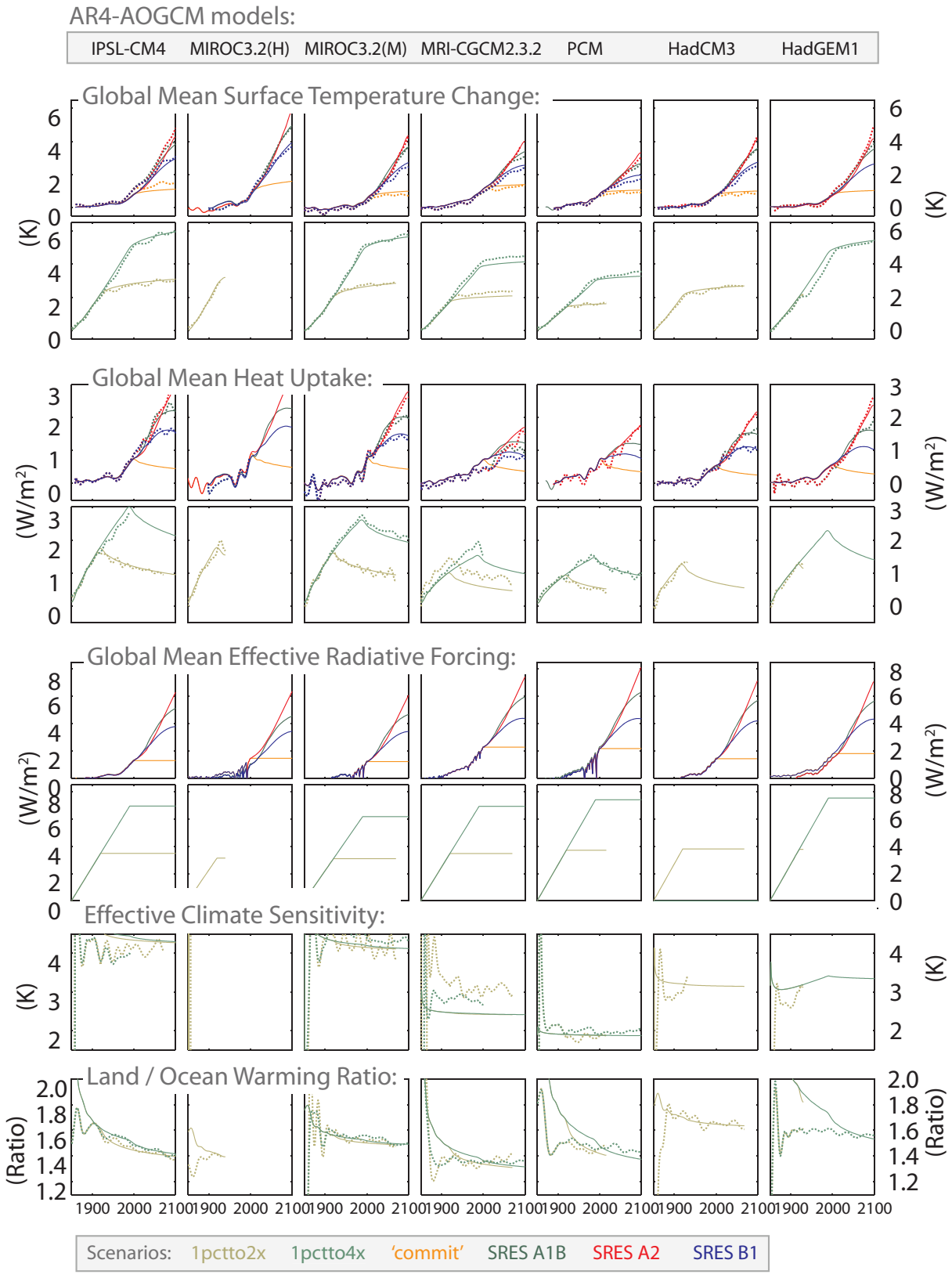

Fig. B3. As Fig. B1, but for another seven of the 19 emulated CMIP3 AOGCMs.

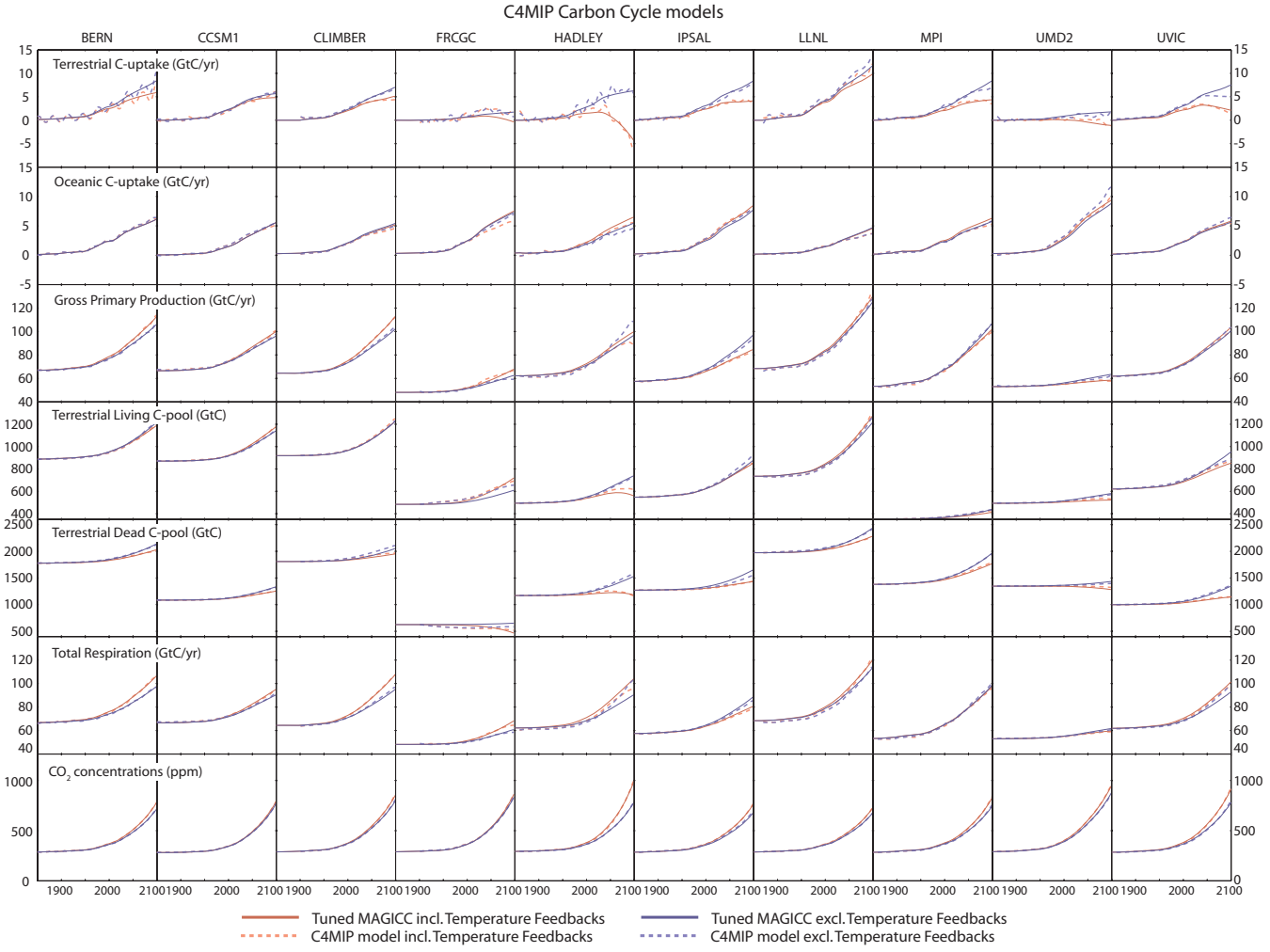

Fig. B4. Comparison of carbon cycle fluxes, pools and atmospheric CO_2 concentrations between C^4MIP carbon cycle models (dashed) and the calibrated MAGICC6 (solid) model. Shown are the coupled (including temperature feedbacks, red lines) and uncoupled (excluding temperature feedbacks, blue lines) runs for the anthropogenic CO_2 emissions based on the IPCC SRES A2 scenario. The carbon fluxes of the C^4MIP models were lowpass-filtered.

Acknowledgements. We would like to acknowledge the numerous collaborations that arose during the preparation of IPCC AR4 and beyond, specifically with J. Gregory, K. Taylor, P. Gleckler, B. Santer, J. Meehl, J. Arblaster, R. Lieberman, F. Joos and R. Knutti. F. Joos is especially thanked for providing ocean carbon cycle parameterisations as described in Joos et al. (1996, 2001). R. Knutti is as well deeply thanked for discussions and comments on an earlier manuscript. J. Lowe, T. Schneider von Deimling, R. Schofield, V. Brovkin, B. Hare and E. Kriegler are also thanked for comments on earlier drafts of this manuscript. For providing various emissions and concentration data sets for halocarbons, we thank J. Daniel and G. Velders. Many people helped in various ways in the development of MAGICC over the past 20 years, namely M. Salmon, M. Schlesinger, M. Hulme, T. Osborn, S. McGinnis and many more. Remaining errors are of course the sole responsibility of the authors.

We acknowledge the modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model output of the third coupled model intercomparison (CMIP3), and the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) for organizing the model data analysis activity. The multi-model data archive is supported by the Office of Science, US Department of Energy.

Edited by: W. E. Asher

Table B1. AOGCM calibration I results: MAGICC6 parameters required to emulate CMIP3 AOGCM models using idealized scenarios and three calibrated parameters only. See Table 1.

		Fixed [a]				Calibrated parameters		
	$\Delta Q^{b}_{2\times}$	ξ	$\frac{dKz_{top}}{dT}$	k_{NS}	k_{LO}	μ	$\Delta T_{2\times}$	RLO	Kz
AOGCM	$(\frac{W}{m^2})$	$(\frac{Km^2}{W})\times 1000$	$(\frac{cm^2}{sK})$	$(\frac{W}{m^2 K})$	$(\frac{W}{m^2 K})$		(K)		$(\frac{cm^2}{s})$
BCC-CM1	insufficient data								
BCCR-BCM2.0	insufficient data								
CCSM3	3.95	0	0	0.50	1.00	1.00	2.35	1.25	1.13
CGCM3.1(T47)	3.32	0	0	0.50	1.00	1.00	3.13	1.25	0.84
CGCM3.1(T63)	insufficient data								
CNRM-CM3	3.48	0	0	0.50	1.00	1.00	2.46	1.27	0.72
CSIRO-Mk3.0	3.47	0	0	0.50	1.00	1.00	2.18	1.24	1.35
ECHAM5/MPI-OM	4.01	0	0	0.50	1.00	1.00	3.95	1.31	0.50
ECHO-G	3.71	0	0	0.50	1.00	1.00	3.10	1.52	1.98
FGOALS-g1.0	3.71	0	0	0.50	1.00	1.00	2.06	1.23	3.74
GFDL-CM2.0	3.50	0	0	0.50	1.00	1.00	2.41	1.40	0.79
GFDL-CM2.1	3.50	0	0	0.50	1.00	1.00	2.34	1.42	1.39
GISS-AOM	insufficient data								
GISS-EH	4.06	0	0	0.50	1.00	1.00	2.84	1.20	1.94
GISS-ER	4.06	0	0	0.50	1.00	1.00	2.66	1.23	3.14
INM-CM3.0	3.71	0	0	0.50	1.00	1.00	2.26	1.30	0.59
IPSL-CM4	3.48	0	0	0.50	1.00	1.00	3.93	1.21	1.71
MIROC3.2(H)	3.14	0	0	0.50	1.00	1.00	6.03	1.17	0.73
MIROC3.2(M)	3.09	0	0	0.50	1.00	1.00	4.12	1.36	1.21
MRI-CGCM2.3.2	3.47	0	0	0.50	1.00	1.00	2.77	1.22	1.33
PCM	3.71	0	0	0.50	1.00	1.00	1.94	1.32	1.16
UKMO-HadCM3	3.81	0	0	0.50	1.00	1.00	3.10	1.35	0.65
UKMO-HadGEM1	3.78	0	0	0.50	1.00	1.00	3.09	1.35	0.64
AVERAGE	3.63	0.00	0.00	0.50	1.00	1.00	2.99	1.29	1.34

[a] The land/ocean area fractions are assumed identical to those provided in Table B3.
[b] If all-sky radiative forcing data for doubled carbon dioxide concentrations was not available (cf. Table 10.2 in Meehl et al., 2007), a default net (longwave+shortwave) forcing of 3.71 Wm^{-2} following Myhre et al. (1998) has been assumed (denoted by italics).

References

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, 1989.

Allen, M.: Do-it-yourself climate prediction, Nature, 401, 642–642, 1999.

Andrews, T. and Forster, P. M.: CO_2 forcing induces semi-direct effects with consequences for climate feedback interpretations, Geophys. Res. Lett., 35, L04802, doi:10.1029/2007GL032273, 2008.

Boer, G. J. and Yu, B.: Climate sensitivity and response, Clim. Dynam., 20, 415–429, 2003.

Bryan, K., Manabe, S., and Pacanowski, R. C.: Global Ocean-Atmosphere Climate Model. 2. Oceanic Circulation, J. Phys. Oceanogr., 5, 30–46, 1975.

Butchart, N. and Scaife, A. A.: Removal of chlorofluorocarbons by increased mass exchange between the stratosphere and troposphere in a changing climate, Nature, 410, 799–802, 2001.

Collins, W. D., Ramaswamy, V., Schwarzkopf, M. D., Sun, Y., Portmann, R. W., Fu, Q., Casanova, S. E. B., Dufresne, J. L., Fillmore, D. W., Forster, P. M. D., Galin, V. Y., Gohar, L. K., Ingram, W. J., Kratz, D. P., Lefebvre, M. P., Li, J., Marquet, P., Oinas, V., Tsushima, Y., Uchiyama, T., and Zhong, W. Y.: Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4), J. Geophys. Res., 111, D14317, doi:10.1029/2005JD006713, 2006.

Cubasch, U., Meehl, G. A., Boer, G. J., Stouffer, R. J., Dix, M., Noda, A., Senior, C. A., Raper, S., and Yap, K.: Projections of Future Climate Change, in: Climate Change 2001: The Scientific Basis, edited by: Houghton, J., Ding, Y., Griggs, D., Noguer, M., van der Linden, P., Dai, X., Maskell, K., and Johnson, C., Cambridge University Press, Cambridge, UK, p. 892, 2001.

Daniel, J. S., Solomon, S., Portmann, R. W., and Garcia, R. R.: Stratospheric ozone destruction: The importance of bromine relative to chlorine, J. Geophys. Res., 104, 23871–23880, 1999.

Doutriaux-Boucher, M., Webb, M. J., Gregory, J. M., and Boucher, O.: Carbon dioxide induced stomatal closure increases radiative forcing via a rapid reduction in low cloud, Geophys. Res. Lett., 36, L02703, doi:10.1029/2008GL036273, 2009.

Ehhalt, D., Prather, M. J., Dentener, F., Derwent, R., Dlugokencky, E., Holland, E., Isaksen, I., Katima, J., Kirchhoff, V., Matson,

Table B2. AOGCM calibration II results: MAGICC6 parameters required to emulate CMIP3 AOGCM models using idealized scenarios and eight calibrated parameters. See Table 1.

	Fixed[a]	Calibrated parameters							
AOGCM	$\Delta Q^{b}_{2\times}$ ($\frac{W}{m^2}$)	$\Delta T_{2\times}$ (K)	ξ ($\frac{Km^2}{W}$)×1000	RLO	Kz ($\frac{cm^2}{s}$)	$\frac{dKz_{top}}{dT}$ ($\frac{cm^2}{sK}$)	k_{NS} ($\frac{W}{m^2K}$)	k_{LO} ($\frac{W}{m^2K}$)	μ
BCC-CM1	insufficient data								
BCCR-BCM2.0	insufficient data								
CCSM3	3.95	2.16	71	1.35	1.36	−0.08	0.20	2.00	1.23
CGCM3.1(T47)	3.32	3.34	0	1.37	1.79	−0.50	0.20	2.00	1.32
CGCM3.1(T63)	insufficient data								
CNRM-CM3	3.48	2.23	75	1.19	0.65	−0.00	0.20	0.20	1.00
CSIRO-Mk3.0	3.47	2.17	0	1.05	2.15	−1.00	0.20	0.20	1.00
ECHAM5/MPI-OM	4.01	3.46	75	1.51	0.62	−0.00	1.00	2.00	1.26
ECHO-G	3.71	3.10	0	1.57	2.70	−1.00	0.20	2.00	1.29
FGOALS-g1.0	3.71	2.11	0	1.00	3.99	−0.00	0.20	0.20	1.00
GFDL-CM2.0	3.50	2.32	34	1.51	0.88	−0.00	0.20	2.00	1.29
GFDL-CM2.1	3.50	2.19	53	1.55	2.67	−1.00	0.20	2.00	1.32
GISS-AOM	insufficient data								
GISS-EH	4.06	2.89	0	1.13	2.67	−1.00	0.20	0.20	1.34
GISS-ER	4.06	2.52	57	1.39	4.69	−1.00	1.66	1.36	1.26
INM-CM3.0	3.71	2.28	0	1.19	0.89	−1.00	0.20	0.20	1.00
IPSL-CM4	3.48	4.03	0	1.11	2.20	−0.24	0.20	0.32	1.00
MIROC3.2(H)	3.14	6.29	0	1.28	0.76	−0.00	0.20	2.00	1.23
MIROC3.2(M)	3.09	4.15	0	1.42	1.52	−0.32	0.20	2.00	1.24
MRI-CGCM2.3.2	3.47	2.87	0	1.20	2.43	−0.68	0.46	0.85	1.00
PCM	3.71	1.83	34	1.20	1.67	−1.00	0.20	0.20	1.00
UKMO-HadCM3	3.81	3.13	0	1.65	1.50	−0.85	0.20	1.68	1.54
UKMO-HadGEM1	3.78	2.90	75	1.28	0.84	−0.00	0.20	0.20	1.00
AVERAGE	3.63	2.95	25	1.31	1.89	-0.51	0.33	1.14	1.17

[a] See note a below Table B1.
[b] See note b below Table B1.

P., Midgley, P., and Wang, M.: Atmospheric Chemistry and Greenhouse Gases, in: Climate Change 2001: The Scientific Basis, edited by: Houghton, J., Ding, Y., Griggs, D., Noguer, M., van der Linden, P., Dai, X., Maskell, K., and Johnson, C., p. 892, Cambridge University Press, Cambridge, UK, 2001.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D., Haywood, J., Lean, J., Lowe, D., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., and Van Dorland, R.: Chapter 2: Changes in Atmospheric Constituents and in Radiative Forcing, in: IPCC Fourth Assessment Report WG 1, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, 129–234, 2007.

Forster, P. M. D. and Taylor, K. E.: Climate forcings and climate sensitivities diagnosed from coupled climate model integrations, J. Climate, 19, 6181–6194, 2006.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, K., Yoshikawa, C., and Zeng, N.: Climate-Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, J. Climate, 19, 3337–3353, 2006.

Frieler, K., Meinshausen, M., Schneider von Deimling, T., Andrews, T., and Forster, P.: Changes in Global-mean Precipitation in Response to Warming, Greenhouse Gas Forcing and Black Carbon, Geophys. Res. Lett., doi:10.1029/2010GL045953, in press, 2011.

Gates, D. M.: Global biospheric response to increasing atmospheric cabron dioxide concentration, in: Direct effects of increasing carbon dioxide on vegetation, edited by: Strain, B. and Cure, J., DOE/ER0238, US Dept. of Energy, Carbon Dioxide Research Division, Washington, DC, USA, 171–184, 1985.

Gifford, R. M.: Implications of CO_2 effects on vegetation for the global carbon budget, in: The global carbon cycle, edited by: Heimann, M., Springer-Verlag, Berlin, 115, 159–200, 1993.

Gregory, J. M. and Webb, M.: Tropospheric Adjustment induces a cloud component of CO2 forcing, J. Climate, 21, 58–71, 2008.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophys. Res. Lett., 31, L03205, doi:10.1029/2003GL018747, 2004.

Table B3. AOGCM calibration III results: MAGICC6 parameters required to emulate CMIP3 AOGCM models using both idealized and multi-forcing runs with the same set of eight calibrated parameters as used in calibration II. The difference is that, here, fitting uses a wider range of climate scenario results. For a description of the calibrated parameters, see Table 1. The fixed parameters, provided here but used as fixed parameters in all three calibration methods, are $\Delta Q_{2\times}$, the AOGCM's forcing at doubled CO_2 concentration levels, and the land area fractions on the northern (F_{NL}) and Southern Hemisphere (F_{SL}).

	Fixed			Calibrated parameters							
AOGCM	$\Delta Q^{b}_{2\times}$ ($\frac{W}{m^2}$)	F_{NL} %	F_{SL} %	$\Delta T_{2\times}$ (K)	ξ ($\frac{Km^2}{W}$)×1000	RLO	Kz ($\frac{cm^2}{s}$)	$\frac{dKz_{top}}{dT}$ ($\frac{cm^2}{sK}$)	k_{NS} ($\frac{W}{m^2K}$)	k_{LO} ($\frac{W}{m^2K}$)	μ
BCC-CM1	insufficient data										
BCCR-BCM2.0	insufficient data										
CCSM3	3.95	49	25	2.14	64	1.37	1.27	−0.21	0.20	2.00	1.22
CGCM3.1(T47)	3.32	50	25	2.97	0	1.34	0.68	−0.00	1.03	0.68	1.00
CGCM3.1(T63)	insufficient data										
CNRM-CM3	3.48	46	22	2.98	0	1.20	0.89	−0.22	0.20	0.22	1.00
CSIRO-Mk3.0	3.47	45	22	2.24	0	1.39	2.57	−0.98	0.20	2.00	1.35
ECHAM5/MPI-OM	4.01	44	22	3.23	75	1.52	0.61	−0.06	1.22	2.00	1.25
ECHO-G	3.71	47	23	2.63	0	1.63	0.43	−1.00	0.29	2.00	1.20
FGOALS-g1.0	3.71	56	29	2.42	75	1.35	1.48	−0.52	0.20	2.00	1.08
GFDL-CM2.0	3.50	50	25	2.31	55	1.52	0.85	−0.02	0.20	2.00	1.30
GFDL-CM2.1	3.50	51	26	2.28	50	1.56	2.31	−1.00	0.20	2.00	1.33
GISS-AOM	insufficient data										
GISS-EH	4.06	53	27	2.54	0	1.10	2.14	−1.00	0.58	0.20	1.00
GISS-ER	4.06	53	27	2.26	75	1.39	2.61	−0.00	2.00	2.00	1.23
INM-CM3.0	3.71	53	27	2.35	0	1.38	0.65	−0.17	0.29	2.00	1.23
IPSL-CM4	3.48	56	30	4.15	0	1.27	2.00	−0.02	2.00	0.72	1.01
MIROC3.2(H)	3.14	48	24	5.73	0	1.29	0.73	−0.00	1.93	1.99	1.00
MIROC3.2(M)	3.09	47	23	4.00	0	1.42	1.65	−0.35	0.32	2.00	1.18
MRI-CGCM2.3.2	3.47	57	29	2.48	0	1.28	1.05	−0.24	0.22	2.00	1.18
PCM	3.71	55	29	1.90	0	1.17	1.36	−0.18	0.20	0.20	1.00
UKMO-HadCM3	3.81	47	23	3.21	0	1.59	0.70	−0.00	0.56	2.00	1.39
UKMO-HadGEM1	3.78	46	23	3.00	75	1.32	0.77	−0.00	0.20	0.20	1.00
AVERAGE	3.63	50	25	2.88	25	1.37	1.30	-0.31	0.63	1.49	1.16

[a] See Table note b in Table B1.
[b] If available, the land area fractions were retrieved from the land area fraction for the pre-industrial control runs as given in the CMIP3 database. If not available, a standard land-sea mask has been used, as available for the land-sea mask function within NCL, the NCAR Command Language.

Gultepe, I. and Isaac, G. A.: Scale Effects on Averaging of Cloud Droplet and Aerosol Number Concentrations: Observations and Models, J. Climate, 12, 1268, doi: 10.1175/1520-0442(1999)012¡1268:SEOAOC¿2.0.CO;2, 1999.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, J. Geophys. Res., 110, D18104, doi:10.1029/2005JD005776, 2005.

Harvey, L. D. D.: Managing atmospheric CO_2, Clim. Change, 15, 343–381, 1989.

Harvey, L. D. D.: Transient Temperature and Sea-Level Response of a 2-Dimensional Ocean-Climate Model to Greenhouse-Gas Increases, J. Geophys. Res.-Ocean., 99, 18447–18466, 1994.

Harvey, L. D. D. and Schneider, S. H.: Transient Climate Response to External Forcing on 10(0)-10(4) Year Time Scales. 1. Experiments with Globally Averaged, Coupled, Atmosphere and Ocean Energy-Balance Models, J. Geophys. Res., 90, 2191–2205, 1985a.

Harvey, L. D. D. and Schneider, S. H.: hklhklhkfdsldfs, J. Geophys. Res., 90, 2207–2222, 1985b.

Harvey, L. D. D., Gregory, J. M., Hoffert, M., Jain, A., Lal, M., Leemans, R., Raper, S. C. B., Wigley, T. M. L., and de Wolde, J.: An Introduction to Simple Climate Models used in the IPCC Second Assessment Report: IPCC Technical Paper 2, Tech. rep., Intergovernmental Panel on Climate Change, 1997.

Hasselmann, K., Sausen, R., Maierreimer, E., and Voss, R.: On the Cold Start Problem in Transient Simulations with Coupled Atmosphere-Ocean Models, Clim. Dynam., 9, 53–61, 1993.

Hoffert, M. I., Callegari, A. J., and Hsieh, C. T.: The Role of Deep-Sea Heat-Storage in the Secular Response to Climatic Forcing,

Table B4. C^{4}MIP calibration: Calibrated and fixed MAGICC6 parameters required to emulate ten of eleven C^{4}MIP carbon cycle models. The upper part of the Table lists calibrated parameters for MAGICC's ocean and terrestrial carbon cycle, with the latter including four parameters for the carbon flux partitions, four temperature sensitivity parameters and two parameters determining fertilization behavior. The lower part of the Table provides the applied fixed parameters during the calibration procedure, such as reference years for each model, from which their scenarios started (Yr_{ref}), initial carbon fluxes for net primary production (NPP_{ini}), the total heterotrophic respiration ($\sum R_{ini}$), initial pool sizes for the plant box (P_{ini}), the detritus box (D_{ini}) and the soil box (S_{ini}). For both the coupled and uncoupled model runs, some goodness of fit statistics are provided: Atmospheric CO_2 concentrations are given for year 2100 for the original C^{4}MIP data (C^{4}MIP) and for the emulation by MAGICC (Emul.), as well as the root mean square error (RMSE) of concentrations over the whole 21st century (cf. Figs. 6 and B4) .

	Calibrated parameters[b]													
	Ocean				Flux Partition				Temp. Sensitivity				Fertilization	
Model	$\Delta T_{2\times}$ (K)	k (yr^{-1})	r	α (K^{-1})	ϕ_H	q_S $\times 100$	g_P	g_H	σ_{NPP} (K^{-1})	σ_R (K^{-1})	σ_Q (K^{-1})	σ_U (K^{-1})	β_m	β_s
BERNcc	2.41	0.24	0.95	0.04	1.00	0.10	0.45	0.40	0.01	0.07	−0.14	0.15	1.10	0.65
CSM1	1.66	0.09	1.68	0.04	1.00	0.10	0.48	0.35	0.01	−0.02	−0.28	0.15	1.81	0.44
CLIMBER2LPJ	3.57	0.26	2.87	0.00	0.98	0.10	0.45	0.41	0.04	0.10	−0.26	0.23	1.95	0.52
FRCGC[c]	4.37	0.17	1.12	0.00	1.00	6.18	0.66	0.23	0.02	−0.20	0.40	0.22	2.00	0.26
HadCM3LC	5.75	0.17	2.71	0.00	0.38	0.10	1.00	0.00	0.00	0.19	−0.03	0.04	2.00	0.57
IPSLCM2C	5.38	0.01	0.01	0.00	1.00	0.10	0.87	0.00	−0.06	−0.20	−0.10	0.09	2.00	0.75
LLNL	5.96	0.04	1.37	0.02	1.00	0.10	0.57	0.31	0.00	−0.01	−0.22	0.22	1.00	1.03
MPI	6.69	0.26	1.23	0.00	1.00	0.10	0.22	0.57	−0.03	−0.02	−0.06	0.07	1.00	1.05
UMD	2.76	0.03	0.57	0.00	1.00	0.10	0.89	0.00	−0.03	0.02	−0.30	0.20	1.44	0.17
UVIC27	5.39	0.19	1.74	0.01	1.00	0.10	0.69	0.00	−0.02	0.01	−0.18	0.11	1.95	0.57

	Fixed parameters						Goodness of fit CO_2 (Coupled)			(Uncoupled)		
Model	Yr_{ref} (yr)	NPP_{ini} ($\frac{GtC}{yr}$)	$R_{P,ini}$ ($\frac{GtC}{yr}$)	P_{ini} (GtC)	D_{ini} (GtC)	S_{ini} (GtC)	C4MIP (ppm)	Emul. (ppm)	RMSE (ppm)	C4MIP (ppm)	Emul. (ppm)	RMSE (ppm)
BERNcc	1765	66	12.3	885	93	1682	784	786	2.0	719	720	2.0
CSM1	1820	67	12.3	870	57	1028	792	794	0.8	773	773	0.4
CLIMBER2LPJ	1901	64	11.9	919	95	1714	871	870	1.6	812	812	1.6
FRCGC[c]	1901	48	8.9	484	33	592	868	865	2.5	845	838	2.7
HadCM3LC	1860	61	11.3	495	61	1109	1025	1012	7.3	801	792	3.1
IPSLCM2C	1860	57	10.6	548	66	1205	769	746	8.0	695	683	4.7
LLNL	1870	67	12.4	735	103	1870	732	732	1.8	681	681	1.3
MPI	1860	53	9.9	351	72	1308	839	825	5.9	756	741	5.6
UMD	1860	53	9.8	493	70	1277	967	958	3.8	869	884	4.0
UVIC27	1860	62	11.5	621	52	947	930	926	1.6	801	793	3.4

[a] For a detailed description of the C^{4}MIP carbon cycle models, see Friedlingstein et al. (2006) and references therein. For some models data was not available up to 2100 (MPI, IPSAL: up to 2099; FRCGC: up to 2098). In such cases, the latest available model years are provided in the Goodness of fit section. The RMSE is calculated over the last 101 years available.

[b] In the automated calibration procedure, the climate module parameters vert. ocean diffusivity ($K_z = 2.3\,cm^2\,s^{-1}$), and ocean-to-land heat exchange enhancement ($\mu = 1.4$), and land-ocean warming ratio (RLO = 1.3) were kept at default values. Climate sensitivity ($\Delta T_{2\times}$)was calibrated in a first step to match temperatures of the coupled run for the model-specific CO_2 concentrations.

[c] The terrestrial carbon pools of the FRCGC model were not used in the calibration routines as they show significant drift at the start of their coupled and uncoupled runs.

Journal of Geophys. Res., 85, 6667–6679, 1980.

Johns, T. C., Carnell, R. E., Crossley, J. F., Gregory, J. M., Mitchell, J. F. B., Senior, C. A., Tett, S. F. B., and Wood, R. A.: The second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation, Clim. Dynam., 13, 103–134, 1997.

Joos, F., Bruno, M., Stocker, T., Siegenthaler, U., Le Quere, C., and Sarmiento, J.: An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake, Tellus, 48B, 397–417, 1996.

Joos, F., Prentice, I. C., Sitch, S., Meyer, R., Hooss, G., Plattner, G.-K., Gerber, S., and Hasselmann, K.: Global Warming feedbacks on terrestrial carbon uptake under the Intergovernmental Panel on Climate Change (IPCC) emission scenarios, Global Biogeochem. Cy., 15, 891–907, 2001.

Joshi, M., Shine, K., Ponater, M., Stuber, N., Sausen, R., and Li, L.: A comparison of climate response to different radiative forcings in three general circulation models: towards an improved metric of climate change, Clim. Dynam., 20, 843–854, 2003.

Kattenberg, A., Giorgi, F., Grassl, H., Meehl, G., Mitchell, J. F. B., Stouffer, R. J., Tokika, T., Weaver, A. J., and Wigley, T. M. L.: Climate models - Projection of future climate. pp. 285-357 (Chapter 6), in: Climate Change 1995: The Science of Climate Change, Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change, edited by Houghton, J., Meira Filho, L., Callander, B., Harris, N., Kattenberg, A., and Maskell, K., Cambridge University Press, Cambridge, p. 572, 1996.

Knutti, R., Allen, M. R., Friedlingstein, P., Gregory, J. M., Hegerl, G. C., Meehl, G. A., Meinshausen, M., Murphy, J. M., Plattner, G.-K., Raper, S. C. B., Stocker, T. F., Stott, P. A., Teng, H., and Wigley, T. M. L.: A review of uncertainties in global temperature projections over the twenty-first century, J. Climate, 21, 2651–2663, 2008.

Knutti, R., Hegerl, G. C.: The equilibrium sensitivity of the Earth's temperature to radiation changes, Nature Geoscience, 1, 735-743, 2008.

Manabe, S. and Bryan, K.: Climate Calculations with a Combined Ocean-Atmosphere Model, J. Atmos. Sci., 26, 786–789, 1969.

Manabe, S., Bryan, K., and Spelman, M. J.: Global Ocean-Atmosphere Climate Model. 1. Atmospheric Circulation, J. Phys. Oceanogr., 5, 3–29, 1975.

Mann, M. E.: On smoothing potentially non-stationary climate time series, Geophys. Res. Lett., 31, L07214, doi:10.1029/2004GL019569, 2004.

Meehl, G. A., Stocker, T. F., Collins, W., Friedlingstein, P., Gaye, A., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J., Noda, A., Raper, S. C. B., Watterson, I., Weaver, A., and Zhao, Z.-C.: Chapter 10: Global Climate Projections, in: IPCC Fourth Assessment Report, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, 747–845, 2007.

Meinshausen, M., Wigley, T. M. L., and Raper, S. C. B.: Emulating atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 2: Applications, Atmos. Chem. Phys., 11, 1457–1471, doi:10.5194/acp-11-1457-2011, 2011.

Mitchell, T. D.: Pattern scaling – An examination of the accuracy of the technique for describing future climates, Clim. Change, 60, 217–242, 2003.

Murphy, J. M. and Mitchell, J. F. B.: Transient-Response of the Hadley-Center Coupled Ocean-Atmosphere Model to Increasing Carbon-Dioxide. 2. Spatial and Temporal Structure of Response, J. Climate, 8, 57–80, 1995.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Myhre, G., Highwood, E. J., Shine, K. P., and Stordal, F.: New estimates of radiative forcing due to well mixed greenhouse gases, Geophys. Res. Lett., 25, 2715–2718, 1998.

Nakicenovic, N. and Swart, R., eds.: IPCC Special Report on Emissions Scenarios, Cambridge University Press, Cambridge, United Kingdom, 2000.

O'Neill, B. C. and Melnikov, N. B.: Learning about parameter and structural uncertainty in carbon cycle models, Clim. Change, 89, 23–44, 2008.

Osborn, T. J., Raper, S. C. B., and Briffa, K. R.: Simulated climate change during the last 1000 years: comparing the ECHO-G general circulation model with the MAGICC simple climate model., Clim. Dynam., 27, 185–197, 2006.

Ramaswamy, V., Boucher, O., Haigh, J., Hauglustaine, D., Haywood, J., Myhre, G., Nakajiama, T., Shi, G., and Solomon, S.: Radiative Forcing of Climate Change, in: Climate Change 2001: The Scientific Basis, edited by Houghton, J., Ding, Y., Griggs, D., Noguer, M., van der Linden, P., Dai, X., Maskell, K., and Johnson, C., Cambridge University Press, Cambridge, UK, p. 892, 2001.

Randall, D., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R., Sumi, A., and Taylor, K. E.: Chapter 8: Climate Models and Their Evaluation, in: IPCC Fourth Assessment Report, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., Cambridge University Press, Cambridge, UK, 589–662, 2007.

Rao, S. and Riahi, K.: The role of Non-CO2 greenhouse gases in climate change mitigation: Long-term scenarios for the 21st century, Energy J., 177–200, 2006.

Rao, S., Riahi, K., Kupiainen, K., and Klimont, Z.: Long-term scenarios for black and organic carbon emissions, Environ. Sci., 2, 205–216, 2005.

Raper, S. C. B.: Interpretation of Model Results that Show Changes in the Effective Climate Sensitivity with Time, in: IPCC Workshop on Climate Sensitivity, 131–133, IPCC, available online at: http://www.ipcc.ch/pdf/supporting-material/ipcc-workshop-paris-july-2004.pdf, Paris, France, 2004.

Raper, S. C. B. and Cubasch, U.: Emulation of the results from a coupled general circulation model using a simple climate model, Geophys. Res. Lett., 23, 1107–1110, 1996.

Raper, S. C. B., Wigley, T. M. L., and Warrick, R. A.: Global Sea-level Rise: Past and Future, in: Sea-Level Rise and Coastal Subsidence: Causes, Consequences and Strategies, edited by: Milliman, J. and Haq, B., Kluwer, Dordrecht, The Netherlands, 11–45, 1996.

Raper, S. C. B., Gregory, J. M., and Osborn, T. J.: Use of an upwelling-diffusion energy balance climate model to simulate and diagnose A/OGCM results, Clim. Dynam., 17, 601–613, 2001.

Santer, B., Wigley, T., Schlesinger, M., and Mitchell, J. F. B.: Developing Climate Scenarios from Equilibrium GCM Results, Report No. 47, Tech. Rep. 47, Max-Planck-Institut für Meteorologie, 1990.

Sarmiento, J. L., Orr, J. C., and Siegenthaler, U.: A perturbation simulation of CO_2 uptake in an ocean general circulation model, J. Geophys. Res., 97, 3621–3645, 1992.

Schlesinger, M. E. and Jiang, X. J.: Simple-Model Representation of Atmosphere-Ocean Gcms and Estimation of the Time Scale of CO_2-Induced Climate Change, J. Climate, 3, 1297–1315, 1990.

Schlesinger, M. E. and Jiang, X. J.: Revised Projection of Future Greenhouse Warming, Nature, 350, 219–221, 1991.

Schlesinger, M. E., Gates, W. L., and Han, Y.-J.: The role of the ocean in CO2-induced climatic warming: Preliminary results from the OSU coupled atmosphere-ocean GCM, in: Coupled Ocean-Atmosphere Models, edited by: Nihoul, J., 447–478, Elsevier, Amsterdam, The Netherlands, 1985.

Schleussner, C. F., Frieler, K., Meinshausen, M., Yin, J., and Levermann, A.: Emulating Atlantic overturning strength for low emission scenarios: consequences for sea-level rise along the North American east coast, Earth Syst. Dynam. Discuss., 1, 357–384, doi:10.5194/esdd-1-357-2010, 2010.

Senior, C. A. and Mitchell, J. F. B.: The time-dependence of climate sensitivity, Geophys. Res. Lett., 27, 2685–2688, 2000.

Shine, K., Derwent, R., Wuebbles, D., and Morcrette, J.-J.: Radiative forcing of climate, in: Climate Change: The IPCC Scientific Assessment, edited by Houghton, J., Jenkins, G., and Ephraums, J., Cambridge University Press, New York, USA, 41–68, 1990.

Stainforth, D. A., Aina, T., Christensen, C., Collins, M., Faull, N.,

Frame, D. J., Kettleborough, J. A., Knight, S., Martin, A., Murphy, J. M., Piani, C., Sexton, D., Smith, L. A., Spicer, R. A., Thorpe, A. J., and Allen, M. R.: Uncertainty in predictions of the climate response to rising levels of greenhouse gases, Nature, 433, 403–406, 2005.

Stouffer, R. J.: Time scales of climate response, J. Climate, 17, 209–217, 2004.

Takahashi, T., Olafson, J., Goddard, J. G., Chipman, D. W., and Sutherland, S. C.: Seasonal variation of CO2 and nutrients in the high-latitude surface oceans: A comparative study, Global Biogeochem. Cy., 7, 843–878, 1993.

Takemura, T., Tsushima, Y., Yokohata, T., Nozawa, T., Nagashima, T., and Nakajima, T.: Time evolutions of various radiative forcings for the past 150 years estimated by a general circulation model, Geophys. Res. Lett., 33, L19705, doi:10.1029/2006GL026666, 2006.

Thornton, P. E., Doney, S. C., Lindsay, K., Moore, J. K., Mahowald, N., Randerson, J. T., Fung, I., Lamarque, J.-F., Feddema, J. J., and Lee, Y.-H.: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences, 6, 2099–2120, doi:10.5194/bg-6-2099-2009, 2009.

Trenberth, K. E. and Guillemot, C. J.: The Total Mass of the Atmosphere, J. Geophys. Res., 99, 23079–23088, 1994.

Twomey, S.: The influence of pollution on the shortwave albedo of clouds, J. Atmos. Sci., 34, 1149–1152, 1977.

Wigley, T. M. L.: Could reducing fossil-fuel emissions cause global warming?, Nature, 349, 503–506, 1991a.

Wigley, T. M. L.: A Simple Inverse Carbon Cycle Model, Global Biogeochem. Cy., 5, 373–382, 1991b.

Wigley, T. M. L.: Balancing the Carbon Budget – Implications for Projections of Future Carbon-Dioxide Concentration Changes, Tellus, 45B, 409–425, 1993.

Wigley, T. M. L.: Stabilization of CO_2 concentration levels, in: The Carbon Cycle, edited by: Wigley, T. M. L. and Schimel, D., Cambridge University Press, Cambridge, UK, 258–276, 2000.

Wigley, T. M. L.: MAGICC/SCENGEN 5.3: USER MANUAL (version 2), www.cgd.ucar.edu/cas/wigley/magicc/, 2008.

Wigley, T. M. L. and Raper, S. C. B.: Thermal-Expansion of Sea-Water Associated with Global Warming, Nature, 330, 127–131, 1987.

Wigley, T. M. L. and Raper, S. C. B.: Implications for climate and sea level of revised IPCC emissions scenarios, Nature, 357, 293–300, 1992.

Wigley, T. M. L. and Raper, S. C. B.: Interpretation of high projections for global-mean warming, Science, 293, 451–454, 2001.

Wigley, T. M. L. and Schlesinger, M.: Analytical Solution for the effect of increasing CO_2 on global mean temperature, Nature, 315, 649–652, 1985.

Wigley, T. M. L., Richels, R., and Edmonds, J.: Overshoot pathways to CO_2 stabilization in a multi-gas context, in: Human Induced Climate Change: An Interdisciplinary Assessment, edited by: Schlesinger, M., Kheshgi, H., Smith, J., de la Chesnaye, F., Reilly, J., Wilson, T., and Kostad, C., Cambridge University Press, Cambridge, UK, 84–92, 2007.

Wigley, T. M. L., Clarke, L. E., Edmonds, J. A., Jacoby, H. D., Paltsev, S., Pitcher, H., Reilly, J. M., Richels, R., Sarofim, M. C., and Smith, S. J.: Uncertainties in climate stabilization, Clim. Change, 97, 85–121, 2009.

Williams, K. D., Ingram, W. J., and Gregory, J. M.: Time variation of effective climate sensitivity in GCMs, J. Climate, 21, 5076–5090, 2008.