Climate Dynamics (2005) 25: 189–204
DOI 10.1007/s00382-005-0027-2

**Chris Jones · Jonathan Gregory · Robert Thorpe
Peter Cox · James Murphy · David Sexton · Paul Valdes**

Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3

Received: 13 December 2004 / Accepted: 18 April 2005

Abstract FAMOUS is an unfluxadjusted coupled atmosphere-ocean general circulation model (AOGCM) based on the Met Office Hadley Centre AOGCM HadCM3. Its parametrisations of physical and dynamical processes are almost identical to those of HadCM3, but by virtue of reduced horizontal and vertical resolution and increased timestep it runs about ten times faster. The speed of FAMOUS means that parameter sensitivities can be investigated more thoroughly than in slower higher-resolution models, with the result that it can be tuned closer to its target climatology. We demonstrate a simple method for systematic tuning of parameters, resulting in a configuration of FAMOUS whose climatology is significantly more realistic than would be expected for a model of its resolution and speed. FAMOUS has been tuned to reproduce the behaviour of HadCM3 as nearly as possible, in order that experiments with each model are of maximum relevance to the physical interpretation of the other. Analysis of the control climate and climate change simulation of FAMOUS show that it possesses sufficient skill for its intended purposes in Earth system science as a tool for long-timescale integrations and for large ensembles of integrations, when HadCM3 cannot be afforded. Thus, it can help to bridge the gap between models of intermediate complexity and the higher-resolution AOGCMs used for policy-relevant climate prediction.

C. Jones (✉) · J. Gregory · R. Thorpe · J. Murphy · D. Sexton
Hadley Centre, Met Office,
FitzRoy Road, Exeter, Devon, EX1 3PB UK
E-mail: chris.d.jones@metoffice.gov.uk
Tel.: +44-1392-884514
Fax: +44-1392-885681

J. Gregory
CGAM, Department of Meteorology,
University of Reading, Reading, RG6 6BB UK

P. Cox
Centre for Ecology and Hydrology,
Winfrith, Dorset, DT2 8ZD UK

P. Valdes
School of Geographical Sciences,
University of Bristol,
Bristol, BS8 1TH UK

1 Introduction

State-of-the-art coupled atmosphere ocean general circulation models (AOGCMs), such as HadCM3 (Gordon et al. 2000), are the best tool for making predictions of future climate change over the coming century because of the detail with which they are able to represent the processes involved. However, they are very computationally expensive to run. Earth system models of intermediate complexity (EMICs, Claussen et al. 2002) are a faster type of model which have been used very successfully for simulations such as studying the stability of the THC in glacial and interglacial periods (Ganopolski and Rahmstorf 2001), or the impact of interactive vegetation in simulations of past climates (Kubatzki et al. 2000). Such models are fast because of low spatial resolution (CLIMBER, Petoukhov et al. 2000), reduced dimensionality (Bern 2.5D, Stocker et al. 1992; Marchal et al. 1998) or very simplified representation of some aspects of physics (UVIC, Weaver et al. 2001) or dynamics (ECBILT, Weber et al. 2004). Their speed makes them practical for purposes for which state-of-the-art AOGCMs are not, but the necessary simplifications may lead them to exclude processes which could be important to the aspects of climate change. For example, they may not represent gyre circulations in the ocean or 3-dimensional atmospheric dynamics, which have been suggested to be important in determining THC stability (Schiller et al. 1997; Vellinga et al. 2002; Thorpe et al. 2001).

To fill the gap between these two classes of models, what is needed is a model based on a state-of-the-art AOGCM, but significantly faster. FAMOUS (FAst Met Office/UK Universities Simulator) is such a model: it is a new, fast version of the HadCM3 AOGCM (Gordon et al. 2000; Pope et al. 2000), developed at the Met Office in collaboration with NERC-funded scientists at the universities of Reading and Bristol. It represents the same physical and dynamical processes as HadCM3 but with reduced horizontal and vertical resolution and increased timestep. As a result it runs about ten times quicker than HadCM3.

The use of FAMOUS in conjunction with HadCM3 will allow more areas of climate and Earth System science to be explored. Basing FAMOUS on HadCM3 means its results are directly traceable to the state-of-the-art model used for policy-relevant climate projections. Likewise, processes of interest in HadCM3 and uncertainty in its projections due to uncertainty in its parameter settings can be studied further in long simulations or parameter ensembles carried out with FAMOUS. For example, Thorpe et al. (2005) use FAMOUS to investigate the multi-century response of the climate and ocean circulation to a reversal of the direction of Earth's spin—an experiment not presently feasible with HadCM3 using the currently available computing power.

FAMOUS is of similar spatial resolution to various other coupled GCMs successfully used in recent climate studies, such as the Goddard Institute for Space Studies GCM (Russell et al. 1995, 2000) or the Geophysical Fluid Dynamics Laboratory (GFDL) R15 version of their climate model (Dixon and Lanzante 1999; Dixon et al. 2003). Similarly, the National Center for Atmospheric Research Community Climate System Model (CCSM, Blackmon et al. 2001) has a low resolution "Paleo-CCSM" version which has been used successfully to simulate past climates such as the Last Glacial Maximum (Shin et al. 2003; Liu et al. 2002).

The speed of FAMOUS makes it possible to use a systematic tuning procedure to optimise its climate simulation and we will demonstrate that the resulting configuration is suitably skillful to be used for scientific studies for which higher resolution GCMs such as HadCM3 would be too computationally expensive. An overview of the configuration of the FAMOUS GCM is given in Sect. 2, and the advantages of an optimal tuning algorithm are discussed in Sect. 3 along with an outline of the one developed here. The results of applying it are described in Sect. 4 and an assessment of the resulting climate of FAMOUS is given in Sect. 5. A discussion of the climate sensitivity is given in Sect. 6, followed by conclusions in Sect. 7.

2 FAMOUS formulation

FAMOUS is a GCM. It is a version of HadCM3 (Gordon et al. 2000; Pope et al. 2000) which has been widely used to simulate present day and future climate. The atmospheric component is a hydrostatic, primitive-equation, grid point model with a hybrid vertical coordinate system. It uses an Eulerian advection scheme and parametrisations of key physical processes such as clouds, precipitation, convection, radiation (for a fuller description see Stratton 1999; Pope et al. 2000). The ocean component is a 20 level version of the Cox (1984) model with parametrised horizontal eddy mixing of tracers and near surface vertical mixing. Sea-ice is represented by a simple thermodynamic scheme which contains parametrisations of ice-drift and leads (Cattle and Crossley 1995) and sea-ice dynamics (Bryan 1969).

The ocean GCM component of FAMOUS is based on the ocean component of HadCM3 (HadOM3, described in more detail in Gordon et al. 2000), but with the horizontal resolution reduced from 1.25°×1.25° to 2.5°×3.75° [the same resolution as HadCM3L, Cox et al. (2001) and HadCM2, Johns et al. (1997)]. The 20 levels in the vertical are unchanged. The use of a "distorted physics" scheme (Bryan 1969, 1984) allows a 12-h timestep to be used without degrading the simulation (Jones 2003). Other differences between the FAMOUS ocean and HadOM3 are listed in the Appendix of Jones et al. (2004); the development and testing of the ocean component is described in detail in Jones (2003).

The FAMOUS atmosphere component is based on the atmosphere component of HadCM3 (HadAM3, Pope et al. 2000). Although GCMs are physically based, many processes at unresolved spatial scales are represented by schemes whose parameters are only loosely constrained by observations. There are no changes to any of the physical parametrisation schemes, other than the tuning of some of the parameters as described in this paper. The horizontal atmospheric resolution of FAMOUS is 5°×7.5°, half that of 2.5°×3.75° used in HadAM3. The vertical resolution is reduced to 11 levels from the 19 of HadAM3. A timestep of 1 h is used compared with 30 min in HadAM3. A detailed description of the differences between FAMOUS and HadCM3 atmospheric components is given in the Appendix of Jones et al. (2004) along with some details of the coupling between the atmospheric and oceanic components.

3 Systematic tuning

Many parameters within the schemes used to represent processes at unresolved spatial scales are only loosely constrained by observations. The process of deriving the values of parameters within those constraints in order to maximise the model's skill is known as "tuning". The high computational resources required to run a GCM can mean that many GCMs are not optimally tuned because not enough parameter sensitivity tests can be run. However, the speed of FAMOUS makes it suited to the application of an optimal tuning method.

It was decided to tune FAMOUS towards its parent model, HadCM3, rather than towards observed clima-

tologies, in order to maximise "traceability" of FAMOUS to HadCM3. By "traceability" here we mean that the models simulate the same behaviour by the same physical mechanisms, although differences in resolution and parameter settings will mean they are not quantitatively identical. The aim is for FAMOUS to emulate HadCM3 sufficiently accurately that inferences can be made about HadCM3 based on simulations with FAMOUS and vice-versa.

Conventionally, tuning a model as complex as a GCM has involved a large amount of subjectivity owing to a lack of information about the effects of possible parameter perturbations. Often several parameters contribute to a given error, and a single parameter is implicated in many errors. An objective method of tuning measures the performance of the model in terms of all errors relative to climatologies and systematically tunes all the parameters to obtain the optimal model configuration. Because such a system clearly relies on hundreds or even thousands of model years to be simulated, it has hitherto only been attempted for simpler climate models than AOGCMs (Annan et al. 2005a; Price et al. 2004)

3.1 Skill score

The tuning method relies on being able to quantify the skill of a particular model configuration with a single number. The choice of which diagnostics to consider and how to assess the skill of the model with respect to each of them will have an impact on the final "optimal" parameter values determined by the tuning algorithm. These subjective decisions will depend on the intended use of the model.

The measure of skill chosen here is the "Arcsin Mielke" (Watterson 1996) score which takes into account several aspects of the fields to be compared and is calculated as follows. For a latitude–longitude field of observations, x, and a model field of the same variable, y, the area weighted mean, $\overline{x}$, and spatial standard deviation, s_x, are given by

$$\overline{x} = \frac{\mathrm{mean}(wx)}{\mathrm{mean}(w)} \tag{1}$$

$$s_x = \sqrt{\left(\frac{wx'x'}{\overline{w}}\right)} \tag{2}$$

and similarly for $\overline{y}$ and s_y, where w is a field of weights given by the relative area of each model gridbox, and $\overline{w}$ is the average of w. Bias, b, is defined as the area average difference between the fields divided by the geometric mean of their spatial standard deviations:

$$b = \frac{\overline{x} - \overline{y}}{\sqrt{s_x s_y}} \tag{3}$$

The pattern correlation coefficient between the given and modelled fields, ρ, is given by

$$\rho = \frac{\overline{x'y'}}{\sqrt{s_x s_y}} \tag{4}$$

where the overline denotes an area weighted mean and $x' = x - \overline{x}$, and $y' = y - \overline{y}$.

Finally, the "Arcsin Mielke" skill score, A, is given by

$$A = \frac{2}{\pi}\arcsin\left(\frac{2\rho}{\sigma + 1/\sigma + b^2}\right) \tag{5}$$

where σ is the ratio of the spatial standard deviations. Its possible values lie between +1 for complete agreement ($b=0$, $\sigma=1$, $\rho=1$) and −1 for complete anticorrelation ($b=0$, $\sigma=1$, $\rho=-1$). It penalises lack of resemblance between the fields ($|\rho|<1$). Fields with no corresponding features will have $\rho=0$ and a score of zero. Hence, a uniform field would score zero, even though its RMS error (a possible alternative measure of skill) would be smaller than one with pronounced but incorrect geographical variation. Since large b leads to small absolute score, it also penalises overall bias, which would not affect ρ. Finally, it penalises differences between the magnitude of spatial variation in the field; either σ or $1/\sigma$ will be large, giving small absolute score. This degrades the score for a field which has the right pattern (large ρ) but the wrong variance.

Subjective tuning of a GCM carries the risk of improving one area of the model at the expense of another. A multivariate assessment of the skill of the model reduces the chances of this risk. Thus, to determine an overall score for FAMOUS, eight model diagnostics were chosen as important variables whose skill should be maximised by the model. They are

- surface air temperature, at 1.5 m height above the ground.
- Precipitation rate (rainfall plus snowfall).
- Outgoing short-wave radiative flux at the top of the atmosphere.
- Outgoing long-wave radiative flux at the top of the atmosphere.
- Soil moisture content.
- Sea-ice concentration (i.e., the fraction of each model gridbox area occupied by sea-ice).
- NPP: net primary productivity of the terrestrial biosphere (i.e., the net uptake of atmospheric carbon by the vegetation).
- Atmospheric stream-function (at 500 mbar), from which the horizontal velocity field at that level can be calculated.

The Arcsin Mielke skill score for multi-annual means of each of these quantities from FAMOUS was calculated against corresponding means from HadCM3. The total score is simply the sum of the scores from the separate diagnostics, giving them all equal weight. The sensitivity of the results to the choice of these diagnostics, and the decision to tune the model to match HadCM3 rather than observed climatologies, will be discussed later.

3.2 The tuning parameters

Given limited computing resources, a choice must be made of a small subset of model parameters to tune. The project on Quantifying Uncertainty in Model Predictions (QUMP) at the Met Office has done much work to identify the role of many parameters in determining HadCM3's performance (Murphy et al. 2004). Based on that work a subset of parameters was chosen as being those which have most impact on the model's skill and which therefore should deliver the greatest benefit when optimally tuned. They are the following:

- RH_{crit}: the threshold of relative humidity for cloud formation (in model level 3 and above) (Smith 1990).
- Vf1: precipitating ice fall-out speed in the Heymsfield formula (Heymsfield 1977).
- C_t: the conversion rate of cloud liquid water droplets to precipitation (Smith 1990).
- C_w: the threshold values of cloud liquid water for formation of precipitation (this parameter takes different values over land and sea which are varied together by the same fractional amounts) (Smith 1990).
- Z_0 (SEA): the convective roughness length over the sea.
- Gravity wave and trapped lee wave constants within the model's parametrisation of gravity wave drag (Milton and Wilson 1996; Gregory et al. 1998). These two constants are varied together by the same fractional amounts.

Standard values of these parameters (i.e. the values used in HadCM3) are listed in Table 1 along with the "high" and "low" limits of uncertainty chosen by experts for the QUMP project, based on current knowledge of the "real" parameter value.

All these parameters are in the atmospheric component of FAMOUS. RH_{crit}, Vf1, C_t and C_w affect clouds which are among the greatest sources of systematic uncertainty in modelling of climate change (Cubasch et al. 2001). Ocean model parameters were not tuned because longer integrations, of centuries rather than decades, might be required to achieve a steady state for a new parameter value. However, ocean tuning could also influence model performance and should be considered when computing resources allow.

3.3 Estimating optimal parameter values

The tuning method developed here assesses the impact of varying model parameters on the performance of the model and iteratively moves towards the model configuration which maximises the skill score with respect to the target climatology. For a given initial configuration of the model, a control run is performed along with a series of runs in which one parameter at a time is perturbed from its control value, within preset bounds. By calculating how the skill of the model varies as a function of each parameter, a value of each parameter can be estimated which optimises the model's skill given the values of the other parameters. These can be combined to produce a new control state which should have improved skill.

This approach firstly assumes that the skill of the model for some untried parameter value can be estimated from the skill at other, nearby, parameter values. It secondly assumes that changes in skill due to single parameter changes can be added linearly when changing several parameters at once. Neither assumption will hold exactly (e.g. Stainforth et al. 2005) and so an iterative procedure is required: perturbation runs are performed on the new control state to determine further improvements.

Thirdly, it assumes that the skill in parameter space varies smoothly enough that the maximum can be found. There is a risk that the method may become "stuck" at a local maximum, when the optimal solution is far away in parameter space. This could be investigated by repeating the tuning from several different initial model configurations. The results should all converge on the same optimal solution. The tuning algorithm we have used is a rather simple one, which we adopted in order to test the concept of systematic tuning of an AOGCM. In future work it would be worthwhile investigating the use of more powerful algorithms such as the ensemble Kalman filter of Annan et al. (2005b) and the genetic algorithm of Price et al. (2004).

At each stage of the tuning we perform two experiments for each parameter with its value perturbed by $\pm 10\%$ of its expert-estimated range from its control value. We then interpolate model skill quadratically within the range of the perturbation to determine the

Table 1 Values of FAMOUS parameters at each phase of tuning

	RH_{crit}	Vf1	C_t	C_w (land, sea)	Gravity wave constants	Z_0 (sea)	Score
"High"	0.9	2.0	4.0e-4	2.0e-3, 5.0e-4	2.0e4, 3.0e5	5.0e-3	
HadCM3	0.7	1.0	1.0e-4	2.0e-4, 5.0e-5	2.0e4, 3.0e5	1.3e-3	
"Low"	0.6	0.5	0.5e-4	1.0e-4, 2.0e-5	0, 0	2.0e-4	
Phase 1	0.7	1.0	1.0e-4	2.0e-4, 5.0e-5	2.0e4, 3.0e5	1.3e-3	5.493
Phase 2	0.692	1.2	0.951e-4	1.97e-4, 4.9e-5	2.0e4, 3.0e5	1.3e-3	5.562
Phase 3	0.693	1.347	0.951–4	1.808e-4, 4.424e-5	2.0e4, 3.0e5	1.111e-3	5.565
Phase 4	0.714	1.547	0.9844e-4	1.808e-4, 4.424e-5	2.0e4, 3.0e5	1.111e-3	5.589
Phase 5	0.687	1.747	0.9414e-4	1.608e-4, 3.824e-5	2.0e4, 3.0e5	1.111e-3	5.567

Final column shows Arcsin Mielke skill score

optimal parameter value for the next phase. Although this algorithm requires two runs for each parameter, making it twice as computationally expensive as a linear descent algorithm, we favour it for two reasons. Firstly, this method tells us not only the *direction* of the optimal perturbation but also its *magnitude*—a linear descent algorithm requires a size of step in parameter space chosen somehow a priori. Secondly, the range over which the quadratic (second order) interpolation of skill holds will probably be greater than the range over which a linear (first order) interpolation would hold.

Assume that model skill, S, is locally a quadratic function of the parameter perturbation:

$$S = a + b\Delta P + c\Delta P^2 \tag{6}$$

where ΔP is the fractional change of the parameter within its allowed range. The optimal perturbation (assuming we are near a maximum rather than a minimum in skill) is therefore the one which solves $\mathrm{d}S/\mathrm{d}\Delta P = 0$:

$$\Delta P_{\mathrm{opt}} = \frac{-b}{2c} \tag{7}$$

For a control run ($\Delta P_0 = 0$) and runs perturbed by ΔP_1 and ΔP_2, the skill is denoted by S_0, S_1 and S_2, respectively. Hence

$$a = S_0 \tag{8}$$

$$a + b\Delta P_1 + c\Delta P_1^2 = S_1 \tag{9}$$

$$a + b\Delta P_2 + c\Delta P_2^2 = S_2 \tag{10}$$

whose solution leads to

$$\Delta P_{\mathrm{opt}} = \frac{1}{2}\frac{S_2\Delta P_1^2 - S_1\Delta P_2^2 - S_0(\Delta P_1^2 - {}_2^2)}{S_2\Delta P_1 - S_1\Delta P_2 - S_0({}_1 - \Delta P_2)} \tag{11}$$

We adopt perturbations of $\pm 10\%$, so $\Delta P_1 = -0.1$ and $\Delta P_2 = 0.1$ and

$$\Delta P_{\mathrm{opt}} = \frac{1}{20}\frac{S_2 - S_1}{2S_0 - S_1 - S_2} \tag{12}$$

4 Results of the optimal tuning of FAMOUS

Each phase of the optimisation comprised a 100-year control run and 12 perturbed parameter runs (two for each tuning parameter), each also for 100 year. Five iterations of the tuning algorithm were performed. The skill measure of each configuration is the sum of the Arcsin Mielke skill score for the eight diagnostics meaned over the final 50 year, compared against a long-term mean from the existing control run of HadCM3. Note that we use the score of multi-annual mean fields rather than averaging the score of individual years because of interannual variability within the run. A perfect match of FAMOUS to HadCM3 would score 8.0 since each diagnostic can score up to a maximum of 1.0, although this is not attainable in practice with runs of finite length because of interannual variability. Comparison of five periods of 50 years from the control run of HadCM3 gives a score of 7.77 ± 0.04.

Phase 1 of the tuning process started from a control run of FAMOUS using the same parameter settings as HadCM3, except that the $\mathrm{RH_{crit}}$ values were reduced to take account of the fact that reduced vertical resolution led to increased thickness of model levels 1 and 2. For model levels 3 and above, $\mathrm{RH_{crit}}$ was one of the parameters tuned subsequently. The parameter settings used are listed in Table 1. Each parameter was perturbed by $\pm 10\%$ of its allowed range from its control state except for the Gravity Wave parameters that were already at the maximum value. These were perturbed by -10% and -20%.

Although the methodology used here has been presented as objective, a degree of subjectivity was occasionally required in practice when processing the results. If the results lead to a diagnosed skill maximum outside of the parameter range tested (i.e. an extrapolation rather than interpolation), then the parameter perturbation was capped at the limit of the range tested rather than extending beyond this range to the diagnosed maximum. Occasionally, the results were not sufficient to produce a clear quadratic interpolation to a skill-maximum. In these cases a subjective decision was taken as to which parameter choice had performed the better. This occurred only once in the first two phases of the tuning, but became more common in phases 3 and 4 as the skill maximum is approached.

The skill scores of 50-year running mean periods from each run are shown in Fig. 1. The control runs (black lines) and perturbed runs (coloured lines) can be seen to improve steadily in skill from phase 1 to 4 (although occasional configurations exhibit relatively low skill). The final 50-year score for each phase is also listed in Table 1 along with the parameter values for each control run.

The phase 5 control run however appears to have slightly lower skill than phase 4. Hence it was decided to stop the tuning at this point. Thus, the phase 4 control configuration may be taken as our optimally tuned version of FAMOUS in terms of fitting this particular set of diagnostics to HadCM3. There is no guarantee however that this configuration is the optimal one by other measures (see, for example, the discussion of climate sensitivity in Sect. 6). Phase 4 does not attain the highest score for each diagnostics individually (see Table 2). For example, phases 2 and 3 have the best score for temperature and phase 5 has the best score for NPP and 500 mbar streamfunction. The quantities which show the greatest improvement during the tuning are the TOA radiative fluxes and NPP. This comes at the cost of degrading the simulation of precipitation and sea-ice fraction. Evidently a different choice or relative weighting of quantities would lead to a different tuned configuration. This is analogous to the subjective choice

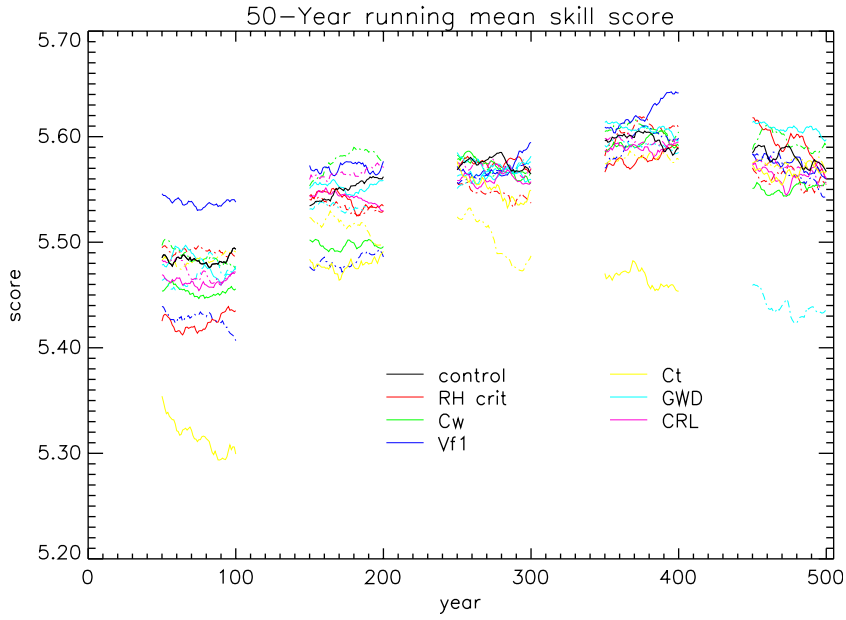

Fig. 1 The 50-year running mean Arcsin Mielke skill score for the control and perturbed runs over the five phases

of metric used to assess model performance by Murphy et al. (2004).

5 Assessment of FAMOUS

5.1 FAMOUS control climate

This section presents a comparison of the FAMOUS control climate (from phase 4) with HadCM3. The phase 4 control run was extended up to 770 years. Timeseries of global mean surface air temperature, top of the atmosphere radiative budget and the strength of the ocean's thermohaline circulation (THC: the maximum of the north Atlantic meridional overturning circulation) show that there is very little long term drift of the control climate state of FAMOUS (Fig. 2). The THC is not a target in the tuning, so it is encouraging that its strength is very similar to that in HadCM3. Thorpe et al. (2005) investigated the response of both the HadCM3 and FAMOUS THC to fresh water forcing in the North Atlantic and found very similar behaviour in both models.

Figures 3, 4, 5 and 6 show simulated annual mean fields from FAMOUS and HadCM3 compared with each other and with observed climatologies. It should be remembered that FAMOUS has been tuned towards HadCM3; thus, although it is instructive to compare with the climatologies, the primary concern here is about the performance of FAMOUS relative to HadCM3.

Figure 3 shows annual mean surface air temperature (at 1.5 m) compared with the Legates and Willmott climatology (Legates and Willmott 1990). A warm bias of FAMOUS in the tropics (especially over land) and a cold bias in the northern latitudes are apparent compared with both observations and HadCM3. In particular there are large cold errors in the north Atlantic and north Pacific, which are associated with the presence of too much sea-ice. However, some of the warm bias over high ground is an artefact of FAMOUS's representation of orography. The coarse horizontal resolution of FAMOUS necessitated some smoothing of the orography to ensure dynamical stability. Thus, the orographic maxima in FAMOUS are lowered slightly introducing a warm bias in surface temperature relative to HadCM3, particularly noticeable in the Himalayas, Andes and African highlands. Figure 3f shows improved agreement when this is taken into account.

Figure 4 shows annual mean precipitation compared with CMAP climatology (Xie and Arkin 1997). Although there are sizable differences between FAMOUS and HadCM3, especially over the ocean, it is not possible to say that one is "better" or "worse" than the other compared with the climatology. The main systematic errors appear common to both models: namely, the dry biases over north-eastern Amazonia and northern India, the wet biases over tropical Africa and China, and the pattern of differences over the oceans.

Figure 5 shows annual mean zonal mean short-wave and long-wave radiation at the top of the atmosphere compared with the ERBE climatology (Harrison et al. 1990). Both GCMs capture the main zonal mean features of the climatology. FAMOUS is closer to the SW climatology north of 70°N and south of 60°S. HadCM3 is closer in the mid- and tropical-latitudes and better captures the small peak just north of the equator. In the global mean, FAMOUS simulates too high a value (107.4 W m^{-2}) compared with an observed value of 101.9 W m^{-2}. HadCM3 is closer at 101.1 W m^{-2}. For long-wave radiation, FAMOUS is closer in the global mean (233.9 W m^{-2}) compared with an observed value of 235.2 W m^{-2} than that of HadCM3 (240.5 W m^{-2}). Here the opposite pattern from the short-wave results is

Table 2 Mean and standard deviation of error with respect to HadCM3 and Arcsin Mielke skill score compared with HadCM3 for each individual diagnostic for each control run

	HadCM3	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Temperature (K)						
Mean	286.7	288.8	287.7	287.5	288.0	287.2
b		0.136	0.071	0.058	0.087	0.039
σ		1.07	1.08	1.08	1.07	1.07
ρ		0.987	0.986	0.985	0.986	0.983
Arcsin Mielke		0.859	0.874	0.874	0.871	0.872
Precip (mm day^{-1})						
Mean	2.91	2.97	2.94	2.93	2.97	2.92
b		0.031	0.017	0.015	0.030	0.009
σ		0.847	0.824	0.822	0.841	0.820
ρ		0.836	0.839	0.831	0.832	0.819
Arcsin Mielke		0.616	0.616	0.607	0.611	0.594
NPP (gC m^{-2} day^{-1})						
Mean	0.94	0.52	0.57	0.58	0.56	0.60
b		−0.26	−0.20	−0.18	−0.21	−0.16
σ		0.73	0.76	0.77	0.75	0.79
ρ		0.79	0.82	0.82	0.81	0.82
Arcsin Mielke		0.517	0.568	0.569	0.554	0.582
SW TOA (W m^{-2})						
Mean	101.1	108.9	109.2	108.8	107.4	108.5
b		0.463	0.478	0.457	0.372	0.441
σ		0.908	0.901	0.899	0.912	0.898
ρ		0.667	0.693	0.699	0.692	0.704
Arcsin Mielke		0.410	0.424	0.433	0.446	0.440
LW TOA (W m^{-2})						
Mean	240.5	232.2	232.5	232.8	233.9	233.2
b		−0.24	−0.23	−0.22	−0.19	−0.21
σ		1.11	1.09	1.09	1.10	1.08
ρ		0.947	0.958	0.959	0.959	0.960
Arcsin Mielke		0.737	0.758	0.763	0.774	0.770
Soil moisture (mm)						
Mean	106.0	123.3	120.9	121.2	121.4	120.7
b		−0.0016	−0.0218	−0.0201	−0.0175	−0.0270
σ		0.961	0.943	0.947	0.949	0.941
ρ		0.906	0.906	0.901	0.903	0.899
Arcsin Mielke		0.720	0.720	0.712	0.715	0.708
Sea-ice fraction (%)						
Mean	4.07	5.46	5.86	5.84	5.57	5.75
b		0.081	0.102	0.101	0.087	0.096
σ		1.18	1.23	1.23	1.20	1.22
ρ		0.939	0.922	0.920	0.922	0.905
Arcsin Mielke		0.748	0.709	0.707	0.718	0.689
500 mbar streamfunction (m^2 s^{-1})						
Mean	−857	−763	−697	−671	−691	−623
b		2.7e-6	4.6e-6	5.4e-6	4.8e-6	6.9e-6
σ		1.16	1.15	1.13	1.13	1.10
ρ		0.9959	0.9957	0.9958	0.9958	0.9957
Arcsin Mielke		0.888	0.896	0.902	0.902	0.914

true: FAMOUS is closer to the climatology in the tropics, with HadCM3 being better in the high latitudes, especially in the north. The greater equator–pole contrast in outgoing long-wave radiation in FAMOUS is related to the greater temperature contrast and reduced meridional heat transport.

Sea-ice fraction (not shown) is generally a little greater in FAMOUS than HadCM3 in the northern hemisphere. FAMOUS and HadCM3 have broadly the same annual mean ice amounts around Antarctica. Soil moisture and NPP (not shown) exhibit similar patterns of differences between HadCM3 and FAMOUS. This is not surprising—it would be expected that where one model had greater (lesser) moisture available for the vegetation it would simulate greater (lesser) vegetation productivity, although temperature differences are clearly important too. Throughout Eurasia FAMOUS tends to simulate too much soil moisture and too great NPP. In Southern Africa, South America and eastern North America FAMOUS simulates too dry soil conditions and too little NPP. In general, the simulation of NPP by FAMOUS is in good agreement with HadCM3 except that it underpredicts the high level of productivity in the tropics by about 25%. Globally it underpredicts NPP by 12%.

Figure 6 shows annual mean zonal mean temperature compared with data from the ECMWF reanalyses (Gibson et al. 1997; Kallberg 1997). From the surface up to about 300 hPa, FAMOUS has a pattern of differences

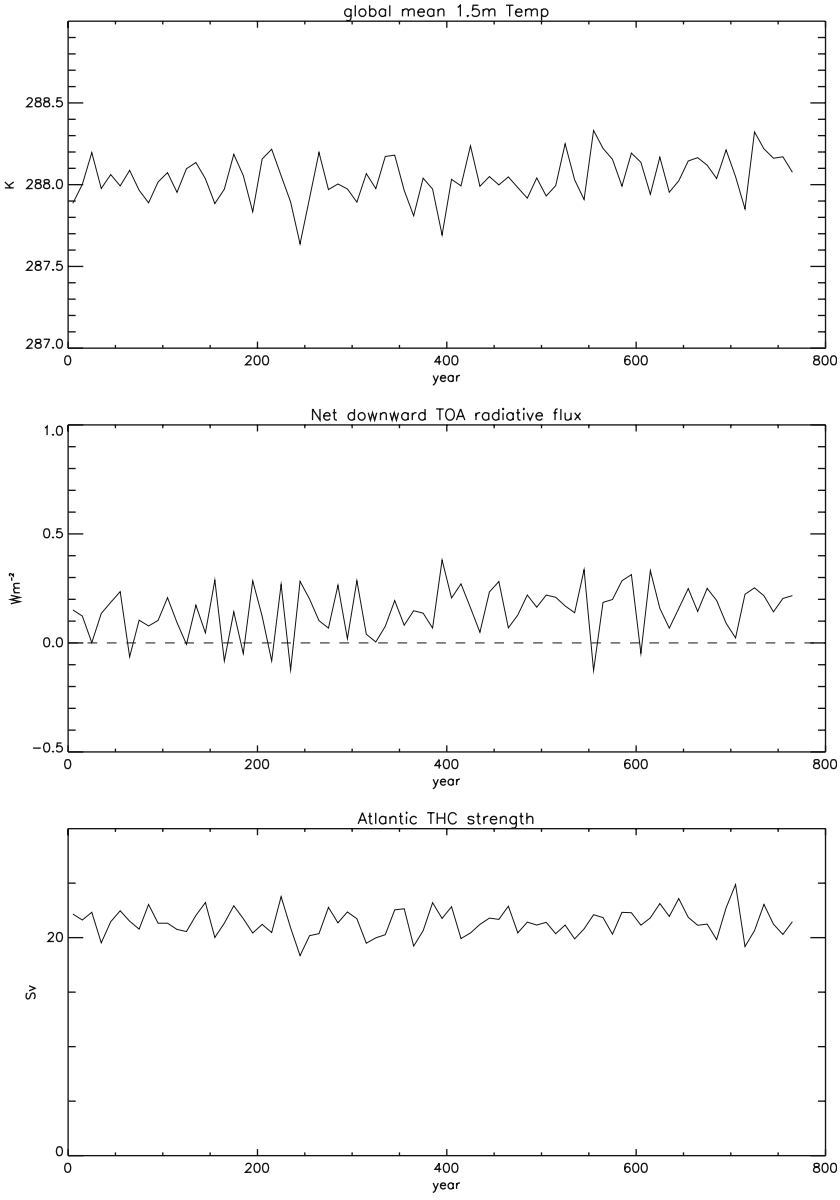

Fig. 2 770-year timeseries of decadal-mean global-mean **a** surface air temperature (K), **b** Net downward top of the atmosphere (TOA) radiative flux (W m^{-2}) and **c** THC strength (Sv) of the extended phase 4 control run of FAMOUS

from the reanalysis data similar to HadCM3, but with generally larger magnitude. As in the surface air temperature, FAMOUS is too warm in the tropics and too cold at high latitudes. This error is present throughout the depth of the troposphere up to about 300 hPa. Above this level, FAMOUS is significantly different from both HadCM3 and the reanalysis data. At these heights it is much too cold. It is likely that this poor simulation of the upper atmosphere is due to the lower vertical resolution of FAMOUS (11 levels compared with 19 in HadCM3).

5.2 Assessing FAMOUS against other models

The aim of the project was to create a GCM faster than HdCM3 without loss of too much skill. In order to determine what constitutes "loss of too much skill", it is useful to consider the skill of several models over a range of speeds. This section discusses how the skill of FAMOUS and its speed compare with other Hadley Centre GCMs (Table 3).

HadCM3L is a version of HadCM3, with the same horizontal resolution in the ocean component as in FAMOUS. It is the climate model used for the coupled climate-carbon cycle experiments (Cox et al. 2000, 2001). HadCM2 (Johns et al. 1997) has the same resolution and timestep as HadCM3L, but many of its physics schemes are older. HadCM3L and HadCM2 have been run both with and without flux adjustments: both are considered here. FAMOUS and HadCM3 are run only without flux adjustments. We also consider a combination of the FAMOUS ocean component coupled to HadAM3 (Jones 2003), to which we refer here as "HadFAM".

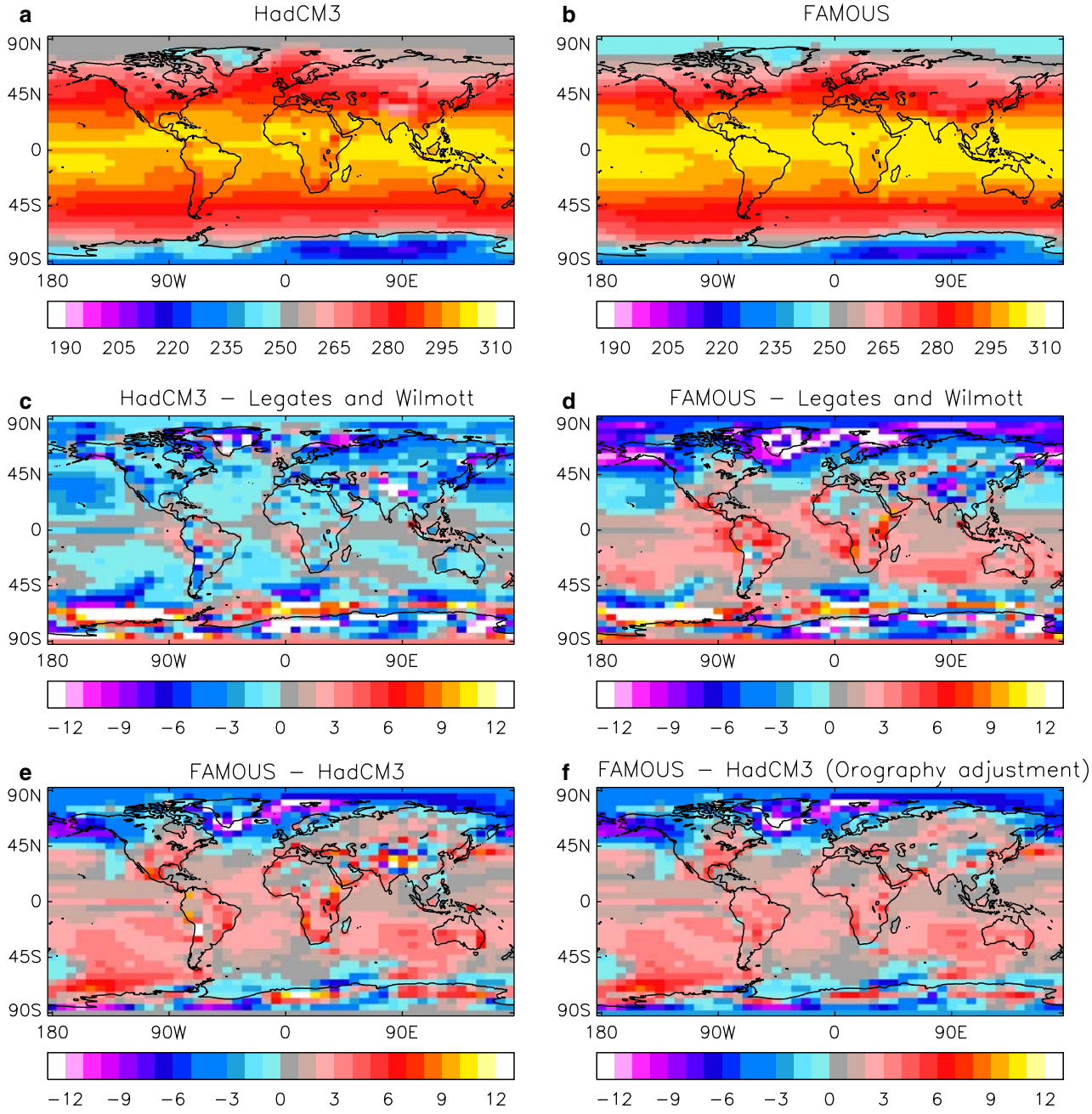

Fig. 3 Maps of annual mean 1.5 m surface air temperature compared with the Legates and Willmott climatology (1990). **a** HadCM3, **b** FAMOUS, **c** HadCM3-LW, **d** FAMOUS-LW, **e** FAMOUS-HadCM3, **f** FAMOUS-HadCM3, but with an adjustment for the different representations of orgraphy in the two models, assuming an atmospheric lapse rate of -6 K km^{-1}

Figure 7 shows the skill of each of the models as a function of model speed. The skill has been computed against observed climatologies rather than against HadCM3 (unlike in the tuning), but uses some of the same diagnostics used to tune FAMOUS (annual mean 1.5 m temperature, precipitation, top of the atmosphere shortwave and longwave, and sea-ice concentration). It might be expected a priori that there would be a steady increase in skill as a GCM increases in either sophistication or resolution (and hence becomes slower). Considering the existing unfluxadjusted models, it can be seen that there is a linear trend for increased skill as the model becomes more costly—a line of best fit is shown on the figure. FAMOUS sits well above this line for the composite of five diagnostics, and for each of the individual diagnostics too, indicating that it is much more skillful than its speed alone would imply.

Overall, with respect to these particular diagnostics, it can be said that although FAMOUS is not as good as HadCM3, it is better than HadCM2 without flux adjustments, and is of comparable quality to HadCM3L

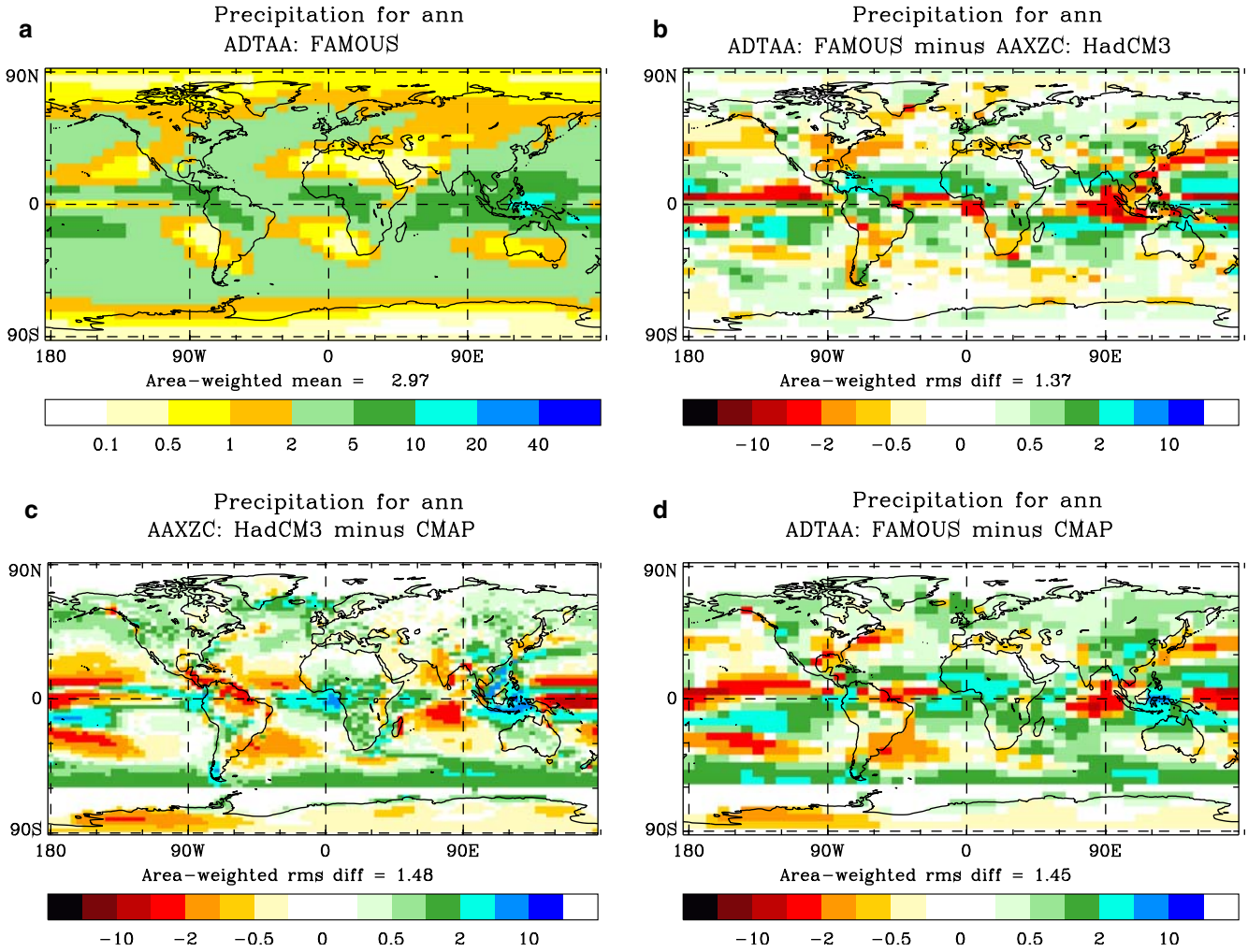

Fig. 4 Maps of annual mean precipitation compared with the CMAP climatology (Xie and Arkin 1997). **a** FAMOUS, **b** FAMOUS-HadCM3, **c** HadCM3-CMAP, **d** FAMOUS-CMAP

without flux adjustments. It is also much faster than these other models.

6 Climate sensitivity

For each control configuration of FAMOUS from the phases of the tuning, a "4×CO_2" experiment was performed in which the atmospheric CO_2 concentration is instantaneously changed to 1,160 ppm (four times the control's pre-industrial value of 290 ppm) and the model is run for a further 50 years. 4×CO_2 is chosen rather than the more conventional 2×CO_2 because it gives a bigger signal and so a more precise estimate of climate sensitivity. The climate sensitivity and patterns of climate change response of the different configurations of FAMOUS are compared with HadCM3. GFDL's R15 model, with an ocean resolution slightly coarser than FAMOUS (4.5× by 3.75× and 12 vertical levels) has been used to study climate change (Dixon and Lanzante 1999) and has been shown to have a similar transient behaviour to the R30 version (Dixon et al. 2003). So we wanted to assess whether an improved fit of the control climate to HadCM3 necessarily led to an improved agreement in the climate-change response.

The equilibrium climate sensitivity of a model (defined as the global mean change in surface temperature due to a doubling of atmospheric CO_2) can be estimated from its transient response to an instantaneous change to CO_2 without having to be run to equilibrium (Gregory et al. 2004). At any given point in the experiment, the net downward radiative flux at the top of the atmosphere, N, is related to the forcing due to the radiative forcing, Q, the climate response parameter, α, and the global mean temperature change, ΔT, by

$$N = Q - \alpha \Delta T \tag{13}$$

Plotting N against ΔT allows us to estimate α and hence the equilibrium ΔT for any given initial forcing. Figure 8 shows such a "Gregory plot" for HadCM3 and the five configurations of FAMOUS. Values of α for each phase are listed in Table 4 along with the resulting estimate of $\Delta T_{2x} = Q_{2x}/\alpha$ assuming a double-CO_2 forcing of $Q_{2x} = 3.74$ W m^{-2} as is standard for HadCM3.

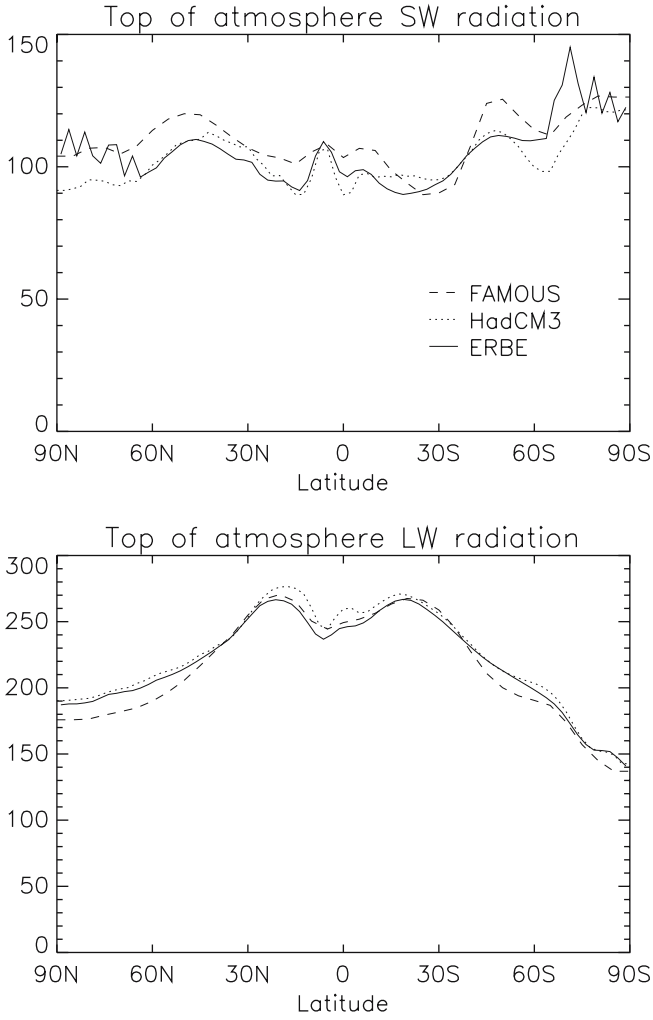

Fig. 5 Annual mean zonal mean radiative fluxes (W m^{-2}) at the top of the atmosphere compared with the ERBE climatology (Harrison et al. 1990). **a** Net downward shortwave and **b** outgoing longwave. FAMOUS in *dashed lines*, HadCM3 in *dotted*, ERBE climatology in *solid lines*

It can be seen that the early configurations of FAMOUS had much higher climate sensitivities than HadCM3. Although remaining above HadCM3, the climate sensitivity of FAMOUS reduced during the tuning, i.e., the process of the tuning makes FAMOUS more similar to HadCM3 in this respect. Despite FAMOUS phase 4 being "better" than FAMOUS phase 5 in terms of its control state, the phase 5 configuration more closely matched the climate sensitivity of HadCM3.

The climate response of FAMOUS to $4\times CO_2$ was compared quantitatively against HadCM3 by calculating the Arcsin Mielke skill score for the pattern of temperature change from years 30–50 of the experiment. The scores for each of the configurations of FAMOUS are shown in Table 5. In each case the pattern correlation is high (close to 0.8) but the skill score increases steadily from 0.29 to 0.53 as a result of the tuning. The increase in skill, therefore, is due largely to the reduction in climate sensitivity making the response generally smaller and hence closer to HadCM3.

Figure 9 shows the pattern of average surface air temperature warming and precipitation changes in HadCM3 and FAMOUS phase 5 from years 30–50 of the $4\times CO_2$ experiment. The extension of strong warming to lower latitudes in FAMOUS is associated with the greater sea-ice extent of the control simulation (Gregory and Mitchell 1997). The precipitation change is also similar and shows many common features such as the Amazon drying. Although this shows that FAMOUS has a similar pattern of climate change to HadCM3, it is likely that the fit could be made even better if the climate-change pattern of HadCM3 were to be used as one of the specific targets of the tuning process. This is an example of how different tuning targets might be used for different intended uses of FAMOUS.

7 Conclusions

A fast AOGCM has been developed. FAMOUS is a version of HadCM3 with lowered spatial resolution and increased timestep and runs about ten times faster than HadCM3. FAMOUS is intended for investigations requiring long integrations or large ensembles, for which HadCM3 is too computationally expensive, but which require the detailed simulation of physical processes permitted by a GCM. It is desirable that FAMOUS should emulate HadCM3, simulating the same behaviour by the same physical mechanisms.

An "optimal tuning algorithm" has been developed, which allows certain parameter values of a GCM to be tuned to maximise the model's fit to a target climatology. In this case, the FAMOUS GCM has been tuned to fit its parent model, HadCM3. The skill of FAMOUS, measured by the Arcsin Mielke objective skill score, was improved as a result of this method. The parameter values required for this optimal fit are not very far removed from the values used in HadCM3. As the skill score improved, the climate sensitivity of FAMOUS to CO_2 forcing became more similar to that of HadCM3, despite not being a target for training.

Although this method could be used to optimise any GCM, it is computationally intensive. Sixty-five model experiments of 100 model years each were performed here. Hence the method is best suited to "fast" GCMs such as FAMOUS. As a result, FAMOUS has a level of skill much higher than would be expected from its resolution and speed of running. It is more skillful than the older HadCM2 model, and very similar to HadCM3L that is significantly slower.

Analysis of the climate simulation of FAMOUS shows it to have a good level of skill relative to HadCM3. Surface air temperature shows a warm tropical bias and a cold northern bias possibly related to insufficient atmospheric heat transport due to the reduced horizontal resolution (Lander and Hoskins 1997).

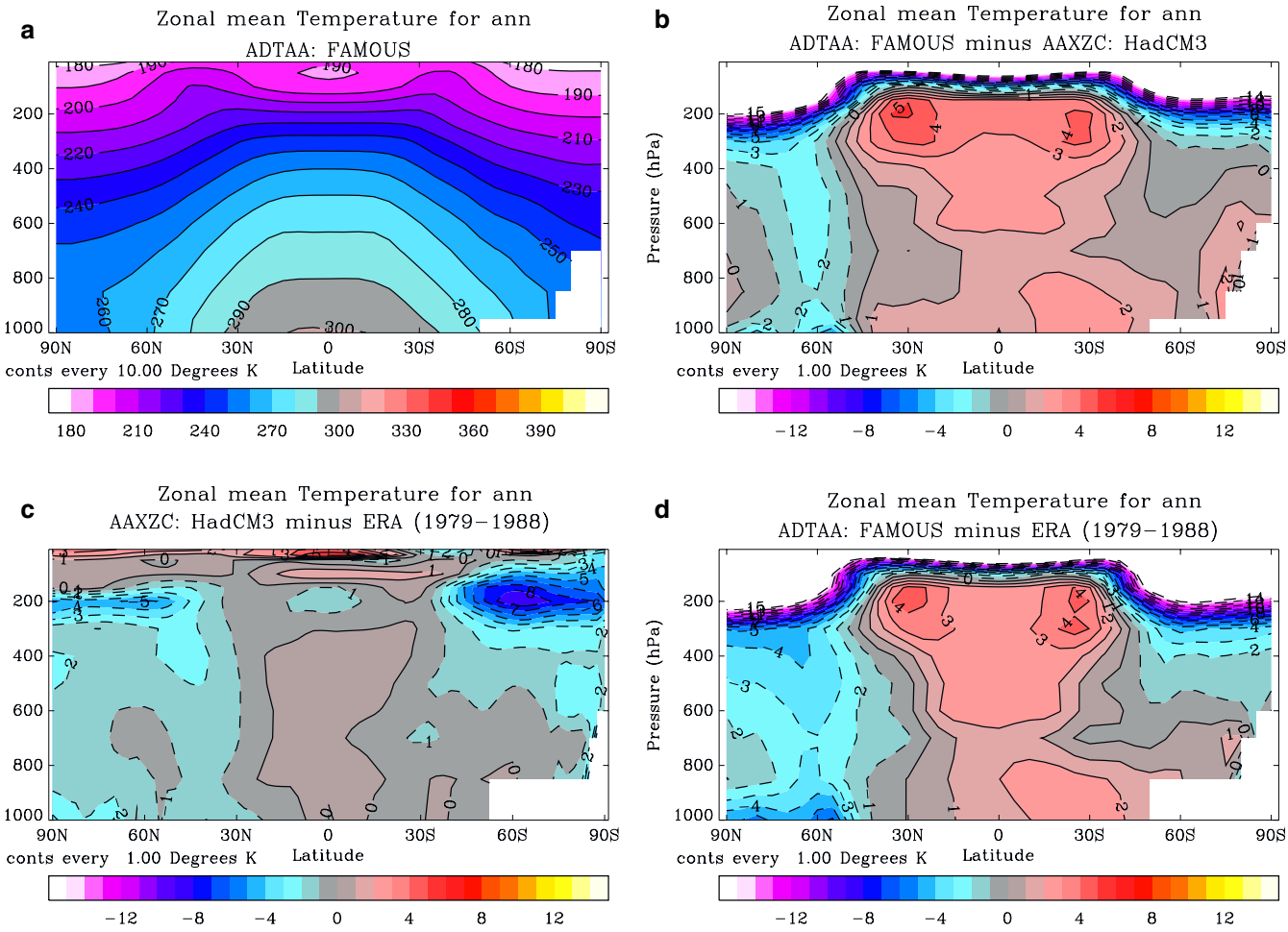

Fig. 6 Annual mean zonal mean temperature compared with ECMWF reanalysis data (Gibson et al. 1997; Kallberg 1997). **a** FAMOUS, **b** FAMOUS-HadCM3, **c** HadCM3-ERA, **d** FAMOUS-ERA

Despite its relatively low resolution and without the use of flux adjustments, FAMOUS has a stable long-term global mean temperature and meridional overturning circulation in the ocean. The main weakness of FAMOUS lies in its much reduced ability to simulate the upper atmosphere, due to the reduction in vertical resolution.

The optimal tuning approach has proved successful in steadily increasing the skill of FAMOUS. It can, however, only correct model errors that are sensitive to the values of the parameters being considered—it cannot correct errors due to other processes or mechanisms. There are FAMOUS errors remaining which are not directly related to the parameters considered here, and so have not been solved by this tuning method. Reasons for, and solutions to, remaining biases are currently still being sought. Such errors are related to the formulation of the parametrisations and the limitation on degrees of freedom from the low resolution. Systematic error arising from these structural considerations is probably an important cause of the differences in climate simulation by different AOGCMs (cf. Murphy et al. 2004).

Table 3 Resolution and timestep values of various Hadley Centre GCMs

Model	Atmosphere component			Ocean component		
	Levels	Horizontal (degrees)	Timestep (h)	Levels	Horizontal (degrees)	Timestep (h)
HadCM3	19	2.5×3.75	1/2	20	1.25×1.25	1
HadCM3L	19	2.5×3.75	1/2	20	2.5×3.75	1
HadCM2	19	2.5×3.75	1/2	20	2.5×3.75	1
HadFAM	19	2.5×3.75	1/2	20	2.5×3.75	24
FAMOUS	11	5×7.5	1	20	2.5×3.75	12

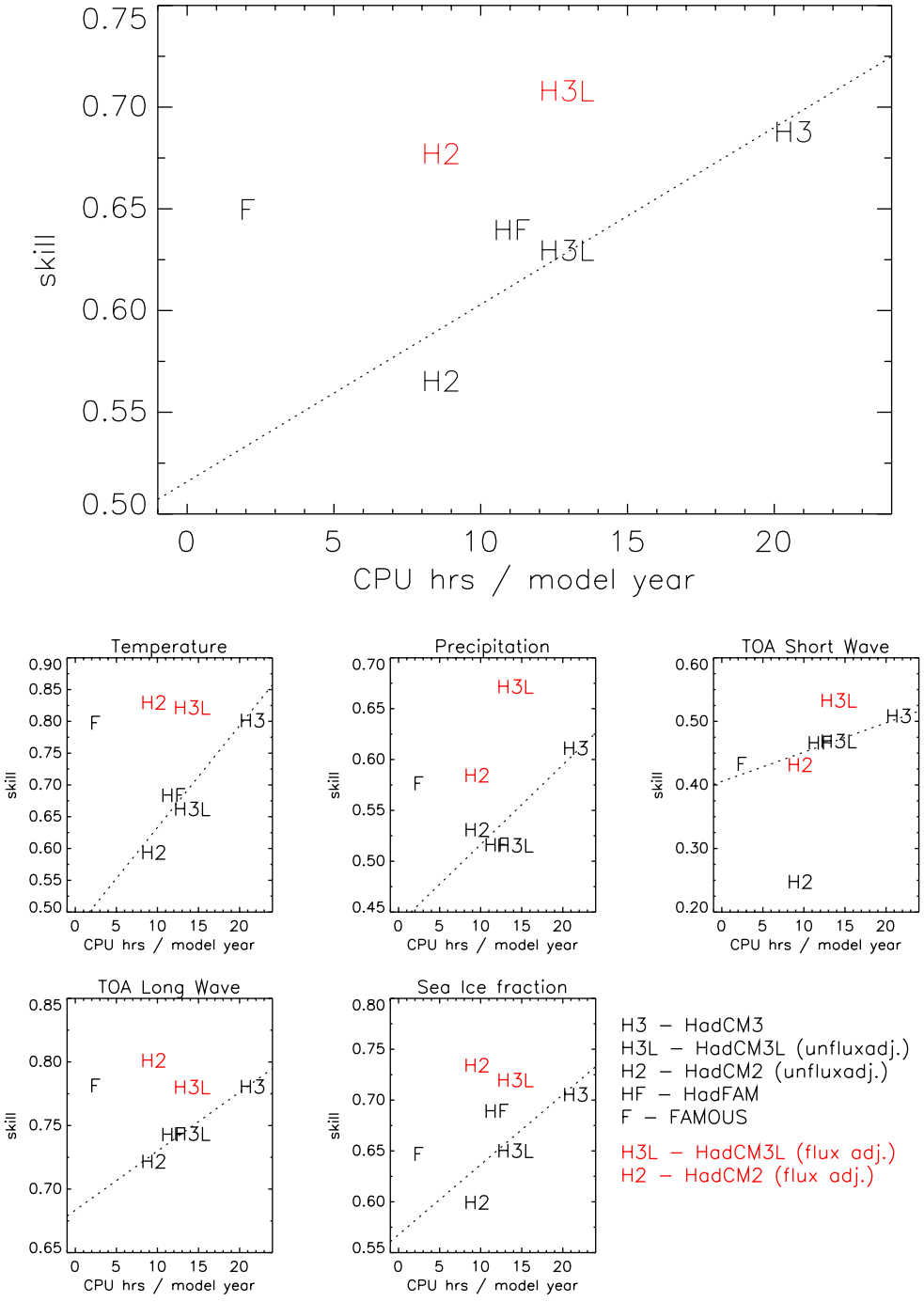

Fig. 7 Arcsin Mielke skill score for a range of Hadley Centre GCMs assessed against observed climatologies and plotted against the time taken for each model to simulate 1 year on the Met Office's T3E supercomputer. **a** Mean of the five diagnostics shown in panels **b**–**e**, **b** surface air temperature compared with Legates and Willmott climatology (Legates and Willmott 1990), **c** precipitation compared with CMAP climatology (Xie and Arkin 1997), **d** short-wave radiation at the top of the atmosphere compared with the ERBE climatology (Harrison et al. 1990), **e** long-wave radiation at the top of the atmosphere also compared with the ERBE climatology, **f** sea-ice fraction compared with HadISST (Rayner et al. 2003). The *dotted line* is the best fit through H3, unfluxadjusted H3L, unfluxadjusted H2 and HF, except in panel **d** where it excludes H2. *Red points* denote GCMs which utilised flux adjustments

The whole tuning process could be repeated starting from different initial parameter values. If the state derived above is indeed the optimal one, then other tuning experiments from different starting points would give the same result. It would also be interesting to investigate the impact on the final parameters of using a different measure of model skill. During the tuning process we made several important decisions for assessing FAMOUS against HadCM3. Introducing other diagnostics such as ocean variables or fluxes across ocean/sea-ice/atmosphere interfaces would affect the final parameter settings given by the optimisation process. Similarly, consideration of further aspects of the diagnostics than long-term means would influence the final results—here we neglect any analysis of seasonal or interannual cycles. Equally, one might choose to give more weight to some diagnostics in the final score than others. These issues have not yet been addressed for FAMOUS, but may form the basis for future work.

Several subjective decisions were made during the process of this optimisation of FAMOUS, and it would be possible in principle to keep the method more

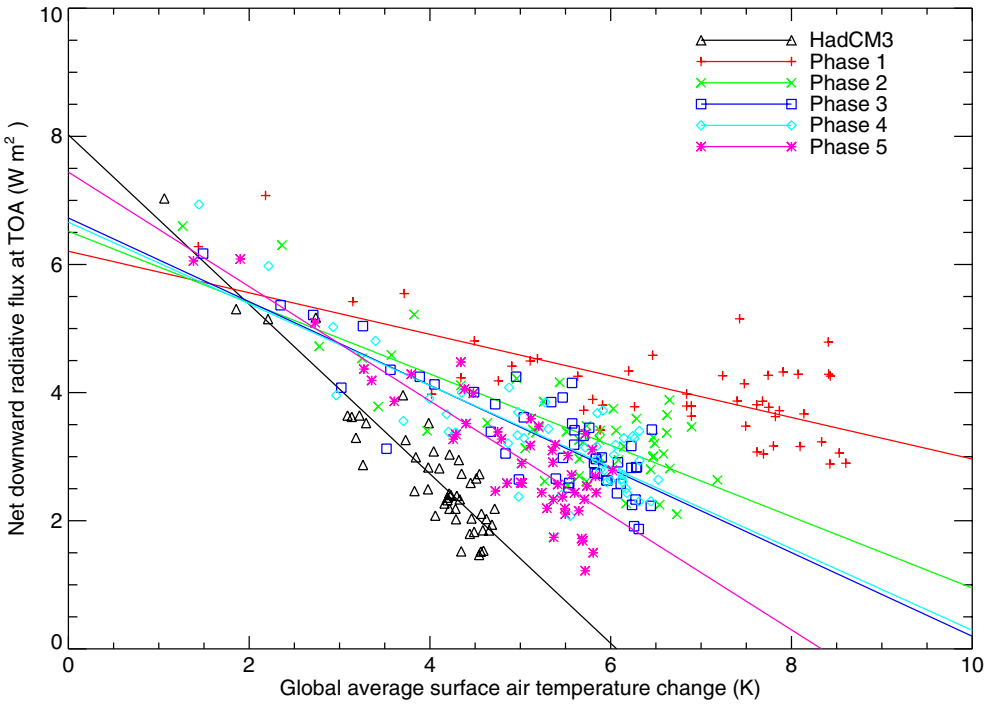

Fig. 8 "Gregory plot" showing the global mean surface air temperature response as a function of TOA radiative forcing for HadCM3 (*black line*) and the five phases of the FAMOUS tuning (*coloured lines*)

Table 4 Estimated equilibrium temperature change for double CO_2 for HadCM3 and the five FAMOUS tuning phases, as estimated from the "Gregory plot" in Fig. 8

Model	Correlation	Q (W m^{-2})	α (W m^{-2} K^{-1})	ΔT_{2x} (K)
HadCM3	0.925	8.0 ± 0.3	1.32 ± 0.08	2.83 ± 0.17
Phase 1	0.703	6.2 ± 0.3	0.32 ± 0.05	11.7 ± 1.9
Phase 2	0.797	6.5 ± 0.3	0.56 ± 0.06	6.7 ± 0.7
Phase 3	0.848	6.7 ± 0.3	0.65 ± 0.06	5.7 ± 0.5
Phase 4	0.831	6.7 ± 0.3	0.64 ± 0.06	5.9 ± 0.6
Phase 5	0.874	7.4 ± 0.4	0.89 ± 0.07	4.2 ± 0.3

Table 5 Arcsin Mielke skill score, and its components, for the five configurations of FAMOUS compared with HadCM3 for average patterns of warming from years 30–50 of the 4×CO_2 experiment

Phase	Bias	σ	ρ	Skill
1	1.217	1.548	0.818	0.294
2	0.718	1.394	0.808	0.422
3	0.571	1.403	0.803	0.457
4	0.593	1.379	0.795	0.448
5	0.402	1.275	0.824	0.533

objective. Hence this raises the possibility of a "self-tuning" GCM, which could automatically and objectively optimise itself before running a particular experiment. This would be analogous in some ways to the two-stage process of numerical weather prediction where a data assimilation is first performed to achieve initial conditions from which the model forecast itself follows. Depending on the intended use of the model, the use of different diagnostics and different model parameters to perturb, or the choice of tuning to match another model or an observed climatology, may lead to different configurations of the same model optimised differently for different uses.

It is hoped that FAMOUS will make a valuable contribution to Earth System science research by bridging the gap between higher resolution GCMs and EMICs. Its speed has allowed it to be optimally tuned and its subsequent skill will allow it to explore many areas of the Earth System previously unaccessible to a GCM.

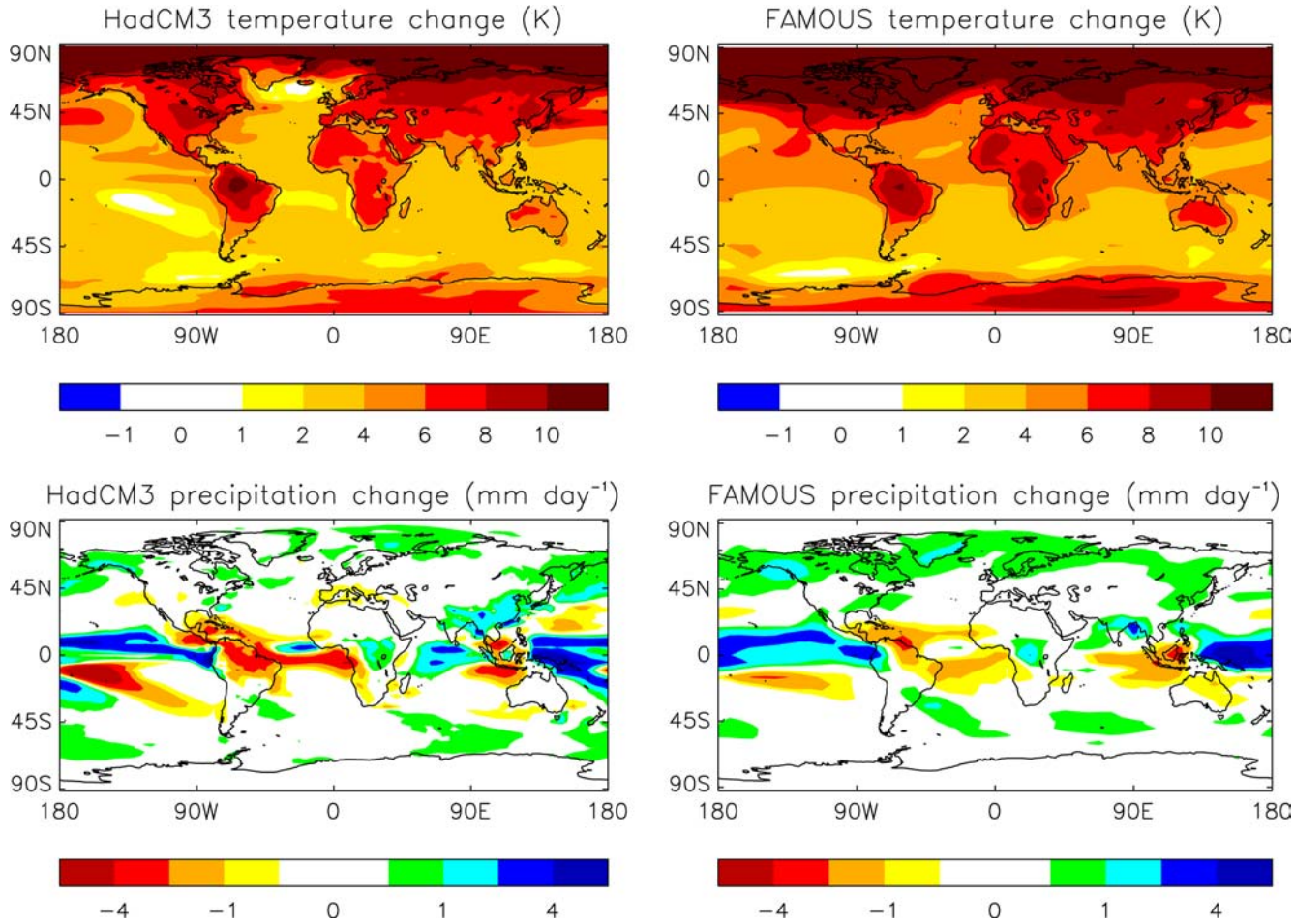

Fig. 9 Climate-change pattern from the "4×CO_2" experiments with HadCM3 (LHS) and FAMOUS (RHS). Surface air temperature (*top row*) and precipitation (*bottom row*)

Acknowledgements This work was supported by the UK Department for the Environment, Food and Regional Affairs under contract PECD 7/12/37.

References

Annan JD, Hargreaves JC, Edwards NR, Marsh R (2005a) Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. Ocean Model 8:135–154

Annan JD, Lunt DJ, Hargreaves JC, Valdes PJ (2005b) Parameter estimation in an atmospheric gcm using the ensemble Kalman filter. Nonlin Process Geophys (submitted)

Blackmon M, Boville B, Bryan F, Dickinson R, Gent P, Kiehl J, Moritz R, Randall D, Shukla J, Solomon S, Bonan G, Doney S, Fung I, Hack J, Hunke E, Hurrell J, Kutzbach J, Meehl J, Otto-Bliesner B, Saravanan R, Schneider EK, Sloan L, Spall M, Taylor K, Tribbia J, Washington W (2001) The community climate system model. Bull Am Meteorol Soc 82:2357–2376

Bryan K, Lewis L (1979) A water mass model of the world ocean. J Geophys Res 84(C5):2503–2517

Bryan K (1969) A numerical method for the study of the circulation of the world ocean. J Comput Phys 4:347–376

Bryan K (1984) Accelerating the convergence to equilibrium of ocean-climate models. J Phys Oceanogr 14:666–673

Cattle H, Crossley J (1995) Modeling Arctic climate-change. Philos Trans Roy Soc A 352:201–213

Claussen M, Mysak L, Weaver A, Crucifix M, Fichefet T, Loutre MF, Weber S, Alcamo J, Alexeev V, Berger A, Calov R, Ganopolski A, Goosse H, Lohmann G, Lunkeit F, Mokhov I, Petoukhov V, Stone P, Wang Z (2002) Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models. Clim Dyn 18(7):579–586

Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. Nature 408:184–187

Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ (2001) Modelling vegetation and the carbon cycle as interactive elements of the climate system. In: Pearce R (ed) Meteorology at the Millennium. Academic, New York, pp 259–279

Cox MD (1984) A primitive equation, three dimensional model of the ocean. Ocean Group Technical Report 1.GFDL, Princeton, 141 pp

Cubasch U, Meehl GA, Boer GJ, Stouffer RJ, Dix M, Noda A, Senior CA, Raper SCB, Yap KS (2001) Projections of future climate change. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden P, Dai X, Maskell K, Johnson CI (eds) Climate change 2001: the scientific basis. Contribution of Working Group I to the 3rd Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp 525–582

Dixon KW, Lanzante JR (1999) Global mean surface air temperature and North Atlantic overturning in a suite of coupled GCM climate change experiments. Geophys Res Lett 26:1885–1888

Dixon KW, Delworth TL, Knutson TR, Spelman MJ, Stouffer RJ (2003) A comparison of climate change simulations produced by two GFDL coupled climate models. Global Planet Change 37:81–102

Ganopolski A, Rahmstorf S (2001) Rapid changes of glacial climate simulated in a coupled climate model. Nature 409:153–158

Gibson JK, Kallberg P, Uppala S, Hernandez A, Nomura A, Serrano E (1997) ERA description. ECMWF Re-anlysis Project Report Series 1, European Centre for Medium Range Weather Forecasts

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147–168

Gregory D, Shutts GJ, Mitchell JR (1998) A new gravity-wave-drag scheme incorporating anistropic orography and low level wave-breaking: impact upon the climate of the UK Meteorological Office unified model. Q J R Meteorol Soc 124:463–494

Gregory JM, Mitchell JFB (1997) The climate response to CO_2 of the Hadley Centre coupled AOGCM with and without flux adjustment. Geophys Res Lett 24:1943–1946

Gregory JM, Ingram WJ, Palmer MA, Jones GS, Stott PA, Thorpe RB, Lowe JA, Johns TC, Williams KD (2004) A new method for diagnosing radiative forcing and climate sensitivity. Geophys Res Lett 31:L03205. DOI 10.1029/2003gl018747

Harrison EP, Minnis P, Barkstrom BR, Ramanathan V, Cess RD, Gibson GG (1990) Seasonal variation of cloud radiative forcing derived from the earth radiation budget experiment. J Geophys Res 95:18687–18703

Heymsfield AJ (1977) Precipitation development in stratiform ice clouds: a microphysical and dynamical study. J Atmos Sci 34:367–381

Johns TC, Carnell RE, Crossley JF, Gregory JM, Mitchell JFB, Senior CA, Tett SFB, Wood RA (1997) The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation. Clim Dyn 13:103–134

Jones CD, Gregory JM, Thorpe RB, Cox PM, Murphy JM, Sexton DMH, Valdes P (2004) Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3. Hadley Centre Technical Note 60, Hadley Centre, Met Office, Met Office, Exeter, EX1 3PB, UK

Jones CD (2003) A fast ocean GCM without flux adjustments. J Atmos Ocean Tech 20(12):1857–1868

Kallberg P (1997) Aspects of the re-analysed climate. ECMWF Re-Analysis Project Series, 2

Kubatzki C, Montoyo M, Rahmstorf S, Ganopolski A, Claussen M (2000) Comparison of the last interglacial climate simulated by a coupled global model of intermediate complexity and an AOGCM. Clim Dyn 16(10–11):799–814

Lander J, Hoskins BJ (1997) Believable scales and parameterizations in a spectral transform model. Mon Wea Rev 125:292–303

Legates DR, Willmott CJ (1990) Mean seasonal and spatial variability in global surface air temperature. Theor Appl Climatol 41:11–21

Liu Z, Shin SI, Otto-Bliesner B, Kutzbach JE, Brady EC, Lee DE (2002) Tropical cooling at the last glacial maximum and extratropical ocean ventilation. Geophys Res Lett 29. DOI 10.1029/2001GL013938

Marchal O, Stocker TF, Joos F (1998) A latitude-depth, circulation-biogeochemical ocean model for paleoclimate studies. development and sensitivities. Tellus B 50B(3):290–316

Milton S, Wilson C (1996) The impact of parametrized sub-grid scale orographic forcing on systematic errors in a global NWP model. Mon Wea Rev 124:2023–2045

Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430:768–772

Petoukhov V, Ganopolski A, Brovkin V, Claussen M, Eliseev A, Kubatzki C (2000) CLIMBER-2: a climate model of intermediate complexity. Part I: Model description and performance for present climate. Clim Dyn 16:1–17

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parametrizations in the Hadley Centre climate model—HadAM3. Clim Dyn 16:123–146

Price AR, Xue G, Yool A, Lunt DJ, Lenton TM, Wason JL, Pound GE, Cox SJ, The GENIE Team (2004) Tuning GENIE earth system model components using a grid enabled data management system. In: Proceedings of the UK e-Science All Hands Meeting, Nottingham, pp 593–600

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of SST, sea ice and night marine air temperature since the late nineteenth century. J Geophys Res 108(D14). DOI 10.1029/2002JD002670

Russell GL, Miller JR, Rind D (1995) A coupled atmosphere-ocean model for transient climate change studies. Atmos Ocean 33:683–730

Russell GL, Miller JR, Rind D, Ruedy RA, Schmidt GA, Sheth S (2000) Comparison of model and observed regional temperature changes during the past 40 years. J Geophys Res 105:14891–14898

Schiller A, Mikolajewicz U, Voss R (1997) The stability of the thermohaline circulation in a coupled ocean-atmosphere general circulation model. Clim Dyn 13:325–347

Shin SI, Liu Z, Otto-Bliesner B, Brady EC, Kutzbach JE, Harrison SP (2003) A simulation of the last glacial maximum climate using the NCAR-CCSM. Clim Dyn 20:127–151

Smith RNB (1990) A scheme for predicting layer clouds and their water content in a general circulation model. Q J R Meteorol Soc 116:435–460

Stainforth DA, Aina T, Christensen C, Collins M, Frame DJ, Kettleborough JA, Knight S, Martin A, Murphy J, Piani C, Sexton D, Smith LA, Spicer RA, Thorpe AJ, Allen MR (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. Nature 433:403–406

Stocker TF, Wright DG, Mysak LA (1992) A zonally averaged, coupled ocean-atmosphere model for paleoclimate studies. J Climate 5(8):773–797

Stratton RA (1999) A high resolution AMIP integration using the Hadley Centre model HadAM2b. Clim Dyn 15:9–28

Thorpe RB, Gregory JM, Johns TC, Wood RA, Mitchell JFB (2001) Mechanisms determining the Atlantic thermohaline circulation response to greenhouse gas forcing in a non-flux-adjusted coupled climate model. J Climate 14:3102–3116

Thorpe RB, Gregory JM, Johns TC, Jones CD, Rodrigues J (2005) The response of the thermohaline circulation in an idealised reverse world experiment. (manuscript in preparation)

Vellinga M, Wood RA, Gregory JM (2002) Processes governing the recovery of a perturbed thermohaline circulation in HadCM3. J Climate 15(7):764–780

Watterson IG (1996) Non-dimensional measures of climate model performance. Int J Climatol 16:379–391

Weaver AJ, Eby M, Wiebe EC, Bitz CM, Duffy PB, Ewen TL, Fanning AF, Holland MM, MacFadyen A, Matthews HD, Meissner KJ, Saenko O, Schmittner A, Wang H, Yoshimori M (2001) The UVic earth system climate model: model description, climatology, and applications to past, present and future climates. Atmos Ocean 39:361–428

Weber SL, Crowley TJ, van der Schrier G (2004) Solar irradiance forcing of centennial climate variability during the holocene. Clim Dyn 22(5):539–553

Xie P, Arkin PA (1997) Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. Bull Am Meteorol Soc 78(11):2539–2558